Lecture Notes in Electrical Engineering

Volume 138

For further volumes:
http://www.springer.com/series/7818

Xudong Wang · Fuzhong Wang
Shaobo Zhong
Editors

Electrical, Information Engineering and Mechatronics 2011

Proceedings of the 2011 International Conference on Electrical, Information Engineering and Mechatronics (EIEM 2011)

Springer

Xudong Wang
School of Electrical Engineering
and Automation
Henan Polytechnic University
Jiaozuo
People's Republic of China
e-mail: xudwang2011@163.com

Shaobo Zhong
College of Elementary Education
Chongqing Normal University
Chongqing
People's Republic of China
e-mail: sbzhong2010@163.com

Fuzhong Wang
School of Electrical Engineering
and Automation
Henan Polytechnic University
Jiaozuo
People's Republic of China
e-mail: fzwang2010@163.com

ISSN 1876-1100
ISBN 978-1-4471-2466-5
e-ISSN 1876-1119
e-ISBN 978-1-4471-2467-2
DOI 10.1007/978-1-4471-2467-2
Springer London Heidelberg New York Dordrecht

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2012931421

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Welcome to the proceedings of the International Conference on Electrical, Information Engineering and Mechatronics 2011 (EIEM 2011), held on 23–25 December 2011, in Jiaozuo, Henan, China.

EIEM 2011 was a venue for leading academic and industrial researchers to exchange their views, ideas, and research results on innovative technologies and sustainable solutions on Electrical, Information Engineering and Mechatronics.

The objective of EIEM 2011 was to facilitate an exchange of information about the best practices for the latest research advances in the area of Electrical, Information Engineering and Mechatronics, which mainly includes the intelligent control, efficient management, or optimal design of access network infrastructures, home networks, terminal equipment, etc. EIEM 2011 was a forum for engineers and scientists in academia, industry and government to address the most innovative research and developments including technical challenges, social and economic issues. Attendees presented and discussed their ideas, results, works in progress and experiences on all aspects of Electrical, Information Engineering and Mechatronics.

There were a very large number of paper submissions (868), representing five countries and regions. All submissions were reviewed by at least three Program and Technical Committee members or external reviewers. It was extremely difficult to select the presentations for the conference because there were so many excellent and interesting submissions. In order to allocate as many papers as possible and keep the quality of the conference high, we finally decided to accept 255 papers for presentations, reflecting a 29.3% acceptance rate; and 85 papers were included in this volume. We believe that all of these papers and topics not only provided novel ideas, new results, work in progress and state-of-the-art techniques in this field, but also stimulated the future research activities in the area of Electrical, Information Engineering and Mechatronics.

The exciting program for this conference was the result of the hard and excellent work of many others, such as Program and Technical Committee members, external reviewers and Publication Chairs under a very tight schedule. We are also grateful to the members of the Local Organizing Committee for

supporting us in handling so many organizational tasks, and to the keynote speakers for agreeing to come to the conference with enthusiasm. Last but not least, we hope you enjoy the conference program.

With our warmest regards,

23–25 December 2011
Jiaozuo, China

EIEM 2011 General and Program Chairs

Contents

Part X Detection Measurement Technologies

Part XIII Power Engineering and Transmission

Part XIV Control System

Part XVII Optical, Electrical, Magnetic and Composite Materials

Organization

EIEM 2011 was organized by Henan Polytechnic University, co-organized by Chinese Academy of Sciences, Society of Intelligent Aerospace Systems (SIAS), Linear Motor Committees of China Electrotechnical Society, China, Zhejiang University, Southeast University, BeiHang University, Northwestern Polytechnical University, China University of Mining and Technology, ZhengZhou University and et al. It was held in cooperation with *Lecture Notes in Electrical Engineering* (LNEE) of Springer.

Executive Committee

General Chairs:	Prof. Shiying Yuan, Henan Polytechnic University, China Yuhang Yang, Shanghai Jiao Tong University, China
Program Chairs:	Yunyue Ye, Zhejiang University, China Fashan Yu, Henan Polytechnic University, China Liucheng Jiao, Henan Normal University, China Ming Fan, University of Washington, USA
Local Arrangement Chairs:	Xudong Wang, Henan Polytechnic University, China Fuzhong Wang, Henan Polytechnic University, China
Steering Committee:	MaodeMa, Nanyang Technological University, Singapore Nadia Nedjah, State University of Rio de Janeiro, Brazil Lorna Uden, Staffordshire University, UK Dechang Chen, Uniformed Services University of the Health Sciences, USA Mei-Ching Chen, Tatung University, Taiwan Rong-Chang Chen, National Taichung Institite of Technology, Taiwan Chi-Cheng Cheng, National Sun Yat-Sen University, Taiwan Naohiko HANAJIMA, Muroran Institute of Technology,

Japan
Shumin Fei, Southeast University, China
Yingmin Jia, BeiHang University, China
Weiguo Liu, Northwesten Polytechnic University, China
Yongji Wang, Huazhong University of Science and Technology, China
Xiaoping Ma, China University of Mining and Technology, China
Jie Wang, Zhengzhou University, China
Yanbo Hui, Henan University of Technology, China
Guangzhao Cui, Zhengzhou University of light Industry, China
Xuejun Li, Hunan University of Science and Technology, China
Pengjun Mao, Henan University of Science and Technology, China
DongyunWang, Zhongyuan University of Technology, China
Wenjiang Du, Chongqing Normal University, China
Qixing Xu, Henan Institute of Engineering, China
Xiaohong Fan, Henan University of Urban Construction, China
Bo Zhao, Henan Polytechnic University, China
Zongpu Jia, Henan Polytechnic University, China
Ziyi Fu, Henan Polytechnic University, China
Xuemao Guan, Henan Polytechnic University, China
Zheng Zhen, Henan Polytechnic University, China
Quanyu Kang, Henan Polytechnic University, China
Yongle Ai, Henan Polytechnic University, China
Ming Fan, Foster School of Business of University of Washington, USA
Juergen Bruess, AutoTXT, Germany
Bahram Honary, Lancaster University, UK
Michael Darnell, Warwick University, UK
Plamen Angelov, Lancaster University, UK
Farideh Honary, Lancaster University, UK
T.R. Melia, Cisco Systems, Switzerland

Program/Technical Committee

Yuan Lin — Norwegian University of Science and Technology, Norwegian

Yajun Li	Shanghai Jiao Tong University, China
Yanliang Jin	Shanghai University, China
Mingyi Gao	National Institute of AIST, Japan
Yajun Guo	Huazhong Normal University, China
Haibing Yin	Peking University, China
Jianxin Chen	University of Vigo, Spain
Miche Rossi	University of Padova, Italy
Ven Prasad	Delft University of Technology, Netherlands
Mina Gui	Texas State University, USA
Nils Asc	University of Bonn, Germany
Ragip Kur	Nokia Research, USA
On Altintas	Toyota InfoTechnology Center, Japan
Suresh Subra	George Washington University, USA
Xiyin Wang	Hebei Polytechnic University, China
Dianxuan Gong	Hebei Polytechnic University, China
Chunxiao Yu	Yanshan University, China
Yanbin Sun	Beijing University of Posts and Telecommunications, China
Guofu Gui	CMC Corporation, China
Haiyong Bao	NTT Co., Ltd., Japan
Xiwen Hu	Wuhan University of Technology, China
Mengze Liao	Cisco China R&D Center, China
Yangwen Zou	Apple China Co., Ltd., China
Liang Zhou	ENSTA-ParisTech, France
Zhanguo Wei	Beijing Forestry University, China
Hao Chen	Hu'nan University, China
Lilei Wang	Beijing University of Posts and Telecommunications, China
Xilong Qu	Hunan Institute of Engineering, China
Duolin Liu	ShenYang Ligong University, China
Xiaozhu Liu	Wuhan University, China
Yanbing Sun	Beijing University of Posts and Telecommunications, China
Yiming Chen	Yanshan University, China
Hui Wang	University of Evry in France, France
Shuang Cong	University of Science and technology of China, China
Haining Wang	College of William and Marry, USA
Zengqiang Chen	Nankai University, China
Dumisa Wellington Ngwenya	Illinois State University, USA
Hu Changhua	Xi'an Research Insti. of Hi-Tech, China
Juntao Fei	Hohai University, China

Zhao-Hui Jiang	Hiroshima Institute of Technology, Japan
Michael Watts	Lincoln University, New Zealand
Tai-hon Kim	Defense Security Command, Korea
Muhammad Khan	Southwest Jiaotong University, China
Seong Kong	The University of Tennessee USA
Worap Kreesuradej	King Mongkuts Institute of Technology Ladkrabang, Thailand
Uwe Kuger	Queen's University of Belfast, UK
Xiao Li	CINVESTAV-IPN, Mexico
Stefa Lindstaedt	Division Manager Knowledge Management, Austria
Paolo Li	Polytechnic of Bari, Italy
Tashi Kuremoto	Yamaguchi University, Japan
Chun Lee	Howon University, Korea
Zheng Liu	Nagasaki Institute of Applied Science, Japan
Michiharu Kurume	National College of Technology, Japan
Sean McLoo	National University of Ireland, Ireland
R. McMenemy	Queens University Belfast, UK
Xiang Mei	The University of Leeds, UK
Cheol Moon	Gwangju University, Korea
Veli Mumcu	Technical University of Yildiz, Turkey
Nin Pang	Auckland University of Technology, New Zealand
Jian-Xin Peng	Queens University of Belfast, UK
Lui Piroddi	Technical University of Milan, Italy
Girij Prasad	University of Ulster, UK
Cent Leung	Victoria University of Technology, Australia
Jams Li	University of Birmingham, UK
Liang Li	University of Sheffield, UK
Hai Qi	University of Tennessee, USA
Wi Richert	University of Paderborn, Germany
Meh shafiei	Dalhousie University, Canada
Sa Sharma	University of Plymouth, UK
Dong Yue	Huazhong University of Science and Technology, China
YongSheng Ding	Donghua University, China
Yuezhi Zhou	Tsinghua University, China
Yongning Tang	Illinois State University, USA
Jun Cai	University of Manitoba, Canada
Sunil Maharaj Sentech	University of Pretoria, South Africa
Mei Yu	Simula Research Laboratory, Norway
Gui-Rong Xue	Shanghai Jiao Tong University, China
Zhichun Li	Northwestern University, China

Lisong Xu	University of Nebraska-Lincoln, USA
Wang Bin	Chinese Academy of Sciences, China
Yan Zhang	Simula Research Laboratory and University of Oslo, Norway
Ruichun Tang	Ocean University of China, China
Wenbin Jiang	Huazhong University of Science and Technology, China
Xingang Zhang	Nanyang Normal University, China
Qishi Wu	University of Memphis, USA
Jalel Ben-Othman	University of Versailles, France

Part I
Advanced Control Algorithm

Chapter 1
The Track Initiation Method Based on Ant-Based Clustering Algorithm and Hough Transform Method

Huigang Xu, Benlian Xu, Fei Wang and Jihong Zhu

Abstract In terms of the phenomenon which occurs in the track initiation method based on Hough transform for multi-target tracking in clutter environment, multiple tracks are initiated by only one target. A novel track initiation method based on ant-based clustering algorithm and Hough transform method is proposed. Firstly, it utilizes the ant-based clustering to remove many clutters and ghosts, and then extracts the possible candidate points originated from true targets. Finally the Hough transform method is applied to initiate the track of the extracted points. Numerical simulations indicated the effectiveness of the proposed method.

Keywords Bearings-only · Target tracking · Track initiation · Ant-based clustering · Hough transform

1.1 Introduction

Track initiation is the first difficult issue in the field of multi-target bearings-only tracking. The sequential track initiation methods [1], such as the heuristic method and logic method, have better initial performance with low false rate; however,

H. Xu (✉) · B. Xu · F. Wang
School of Electric and Automatic Engineering,
ChangShu Institute of Technology, ChangShu 215500, Jiangsu,
People's Republic of China
e-mail: xuhuigang@cslg.edu.cn

J. Zhu
School of Automation, NanJing University of Science & Technology,
Nanjing 210094, Jiangsu, People's Republic of China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_1,

when the number of targets is large or clutter density is heavy, sequential methods have the disadvantage of large computation which goes against real-time tracking. The batch processing methods, such as Hough transform and modified Hough transform method, are applicable to track in a straight line in clutter background, but it needs scan many times to get better initial tracks, which is an obstacle for fast track initiation. In order to solve the problems mentioned above, many track initiation methods are proposed. Especially, with the development of evolutionary optimization algorithms, they are utilized to solve many practical applications and gradually emerged in target tracking field [2–5], and the [2, 3] have been used in track initiation successfully.

In general, the number of targets is unknown in target tracking system. The ant-based clustering with the ability of global optimization [6], which originated from the picking up and dropping behaviors of ants, does not rely on specific issues and require the center of clusters. So it may be very suitable for the stage of track initiation. In this paper, a hybrid track initiation method based on Hough transform and ant-based clustering algorithm is proposed. Firstly, the ant-based clustering algorithm is applied to all measurements and extracted all candidate points of interest. Secondly, Hough transform method is used to initiate the tracks.

1.2 Checking the PDF File Track Initiation Method Based on Hough Transform and Ant-Based Clustering

In this section, the main idea of proposed method is introduced, which composes three steps, i.e., attribute construction, candidate discrimination and track initiation method based on Hough transform method.

Step 1: Attribute construction. The attributes of each candidate are constructed similarl to [7].

Step 2: Candidate discrimination. Each candidate is regarded as an object, and the ith object is defined as $o_i(\theta_i^t, \omega_i^t)$, where θ^t denotes the azimuth change rate and ω^t denotes the ratio of change rate and range, the collection of all objects are composed of the data set $X\left(\bigcup_{i=1}^{n_0} o_i\right)$. The data set X and ants are placed on a two-dimentional board M randomly, where M is denoted by a matrix with $m_b \times m_b$ $\left(m_b = \sqrt{4 \times n_0}\right)$ size, and each cell can only contain an ant and a number of targets; ants are allowed to travel from one cell to another freely. There are four types of situation for ants forming heaps $\tilde{X}$ with some of objects of data set X [6]:

(1) If the heap $\tilde{X}$ contains object o_j, for any object o_i, if condition (1.1) holds, a heap of two objects is created by dropping o_i on o_j.

$$\frac{D(o_i, o_j)}{D_{\max}} < T_{\text{create}} \tag{1.1}$$

where $D(\cdot)$ denotes Euclidean distance, $D_{\max}(\tilde{X}) = \max_{X_i, X_j \in \tilde{X}} D(o_i, o_j)$ denotes the maximum distance of objects in heap $\tilde{X}$, and T_{create} is a given parameter.

(2) For a heap $\tilde{X}$, when the Eq. 1.2 holds, the object $O_{\text{dissim}}(\tilde{X})$ which is the farthest one to the center of heap $O_{\text{center}}(\tilde{X})$ will be removed, and a new heap will be formed by the left objects.

$$\frac{D(O_{\text{dissim}}(\tilde{X}), O_{\text{center}}(\tilde{X}))}{D_{\text{mean}}(\tilde{X})} > T_{\text{remove}} \tag{1.2}$$

where $O_{\text{center}}(\tilde{X}) = \frac{1}{n_{\tilde{X}}} \sum_{o_i \in \tilde{X}} o_i$, $n_{\tilde{X}}$ denotes the total number of objects in $\tilde{X}$, and $D_{\text{mean}}(\tilde{X}) = \frac{1}{n_{\tilde{X}}} \sum_{o_i \in \tilde{X}} D(o_i, O_{\text{center}}(\tilde{X}))$ represents the mean distance between the objects of $\tilde{X}$ and the center of the mass of the heap.

(3) For any o_i and $\tilde{X}$, if condition (1.3) holds, the o_i is dropping on $\tilde{X}$

$$D(o_i, O_{\text{center}}(\tilde{X})) < D(O_{\text{dissim}}(\tilde{X}), O_{\text{center}}(\tilde{X})) \tag{1.3}$$

(4) For heap $\tilde{X}$, if inequality (1.4) holds, ant drops heap $\tilde{X}'$ onto another heap $\tilde{X}$, and they form a new heap.

$$\frac{D(O_{\text{center}}(\tilde{X}), O_{\text{center}}(\tilde{X}'))}{D_{\max}} < T_{\text{createofheap}} \tag{1.4}$$

After N iterations finish, there formed n' heaps on board M, i.e. $X = \bigcup_{i=1}^{n'} \tilde{X}_i$. According to the reason introduced in [7], the smaller θ_i^t and ω_i^t are, the more possibility the corresponding candidate point comes from the true target. In addition, the performance of ant-based clustering algorithm depends on the parameters set, such as the number of iterations. However, if little iteration is set, the clustering effect will be greatly reduced, on the other hand, it cannot meet the timeliness requirements of tracking initiation. Under this condition, a compromise is applied to get the clustering result, i.e. any two heaps are similar when inequality (1.5) is satisfied while relatively little iteration is utilized, ε is a small positive parameter.

$$D(O_{\text{center}}(\tilde{X}_i), O_{\text{center}}(\tilde{X}_j)) \le \varepsilon \tag{1.5}$$

A new data set $\tilde{\Omega}$ formed by total similar heaps, and the number of $\tilde{\Omega}$ is $n_0'(n_0' \le n_0)$. Every object in $\tilde{\Omega}$ is the candidate point.

Table 1.1 The proposed track initiation method

1 Construct attributes of all candidate points
2 Initialize the board, data set, the location, and direction of ant
3 Repeat
3.1 For each ant do
3.1.1 The ants move along the direction to one of its neighbors
3.1.2 If the ants do not carry any object then search its neighbors, and picks up the object in the cell with a probability, else
3.1.3 Drop the object with a probability
4 Extract all possible candidate points
5 Using Hough transform to initiate tracks of these extracted points

Step 3: Track initiation method based on Hough transform. For all objects in $\tilde{\Omega}$, the Hough transform method proposed in work [1] is used to initiate tracks of targets. The track initiation based on ant-based clustering algorithm and Hough transform is presented in Table 1.1.

1.3 Additional Information Required by the Volume Editor

The initial information of targets and state of sensors are selected from [9], and other parameters, based on numerous simulation result, are selected as follows: $T = 6$, $t_0 = 10\,\text{s}$, $\varepsilon = 0.002$, $r_0 = 3$, $N_\theta = 180$, $N_\rho = 50$, $I_{\text{iteration}} = 500$, $P_{\text{direction}} = 0.3$, $T_{\text{create}} = 0.5$, $P_{\text{drop}} = 0.2$, $P_{\text{destroy}} = 0.3$, $P_{\text{load}} = 0.3$, and $T_{\text{remove}} = 1.5$.

In order to show the effectiveness of the proposed method, two evaluation indices are selected from work [2], i.e. the false initiation probability and the true initiation probability. Figures 1.1, 1.2, 1.3, and 1.4 show one of all results of Hough transform and the proposed method in this paper. Figure 1.1 gives the measurement results of three targets on conditions that measurement noise is set 0.05°. Figure 1.2 gives the initiation results of traditional Hough transform. From Fig. 1.2 we can see that the drawback of traditional Hough transform method is that multiple tracks are obtained from one true target even under the condition of low clutter density. Figure 1.3 shows all the possible candidate points which may come from true target after using the ant-based clustering algorithm. Obviously, a mount of clutters and ghosts are removed, while candidate points that come from true targets are guaranteed as much as possible. Figure 1.4 shows the result of the proposed method in this paper.

Table 1.2 shows the false and true initiation probability with different measurement noise and Monte Carlo runs is set 20 in each measurement noise scenario. From Table 1.2, we can say that when measurement noise is set at 0.01, 0.05 and 0.1° respectively, the minimum true initiation probability is 0.88 and the maximum false initiation probability is 0.56. When measurement noise is set to 0.2°, the minimum true initiation probability is 0.8 and false initiation probability

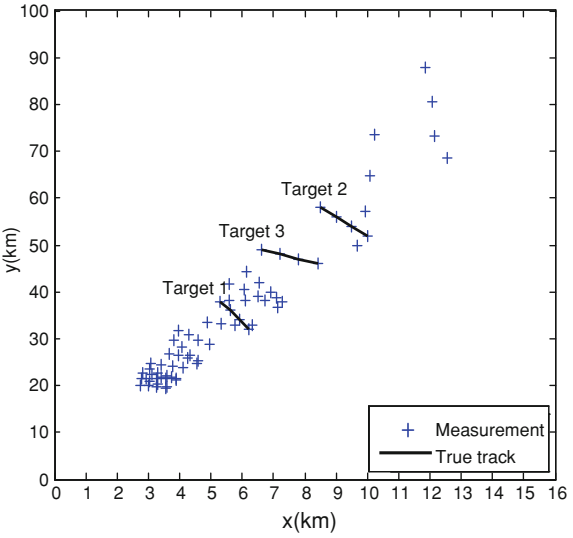


Fig. 1.1 True position of three targets

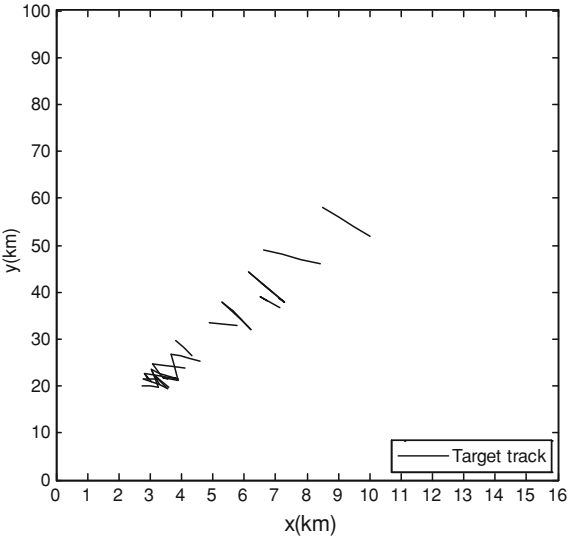


Fig. 1.2 Result of Hough transform

is 0.87. So with the increase of measurement noise, the true initiation probability is decreased and false initiation probability is increased. When the measurement noise increases to a certain level, such as 0.2°, the performance of the proposed method is greatly reduced.

Table 1.3 shows performance comparison of the proposed method, logic method, Hough transform method, and ant colony optimize method [2] with measurement noise 0.05. From Table 1.3 we can say that the minimum true initiation probability of the proposed method is better than logic method and Hough

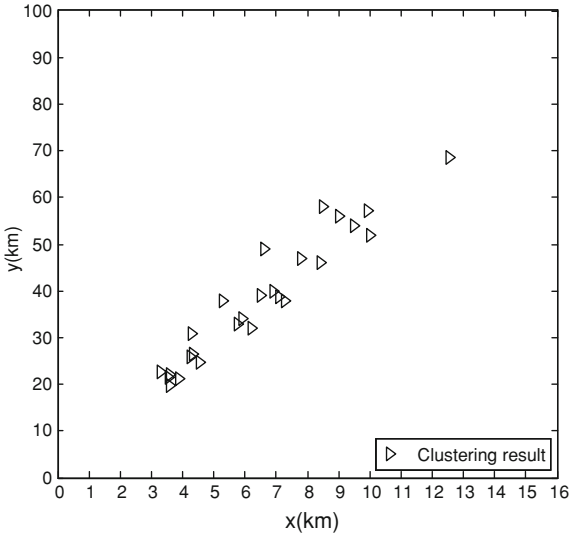


Fig. 1.3 Result of ant-based clustering

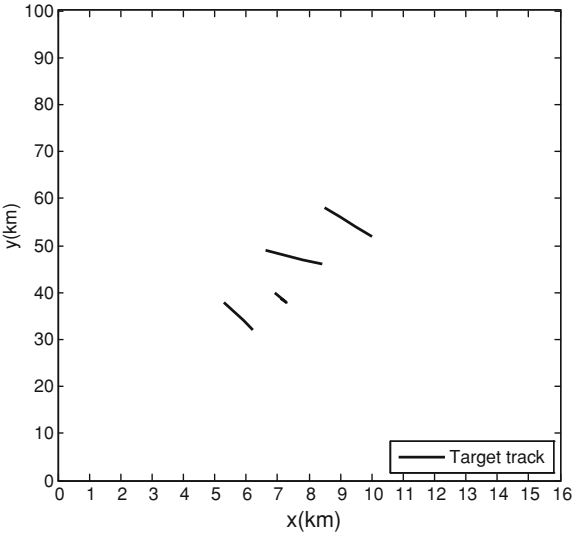


Fig. 1.4 Result of hybrid method

Table 1.2 Probability of track initiation with different measurement noise

η	0.01°	0.05°	0.1°	0.2°
F	0.30	0.35	0.56	0.87
C_1	0.98	0.96	0.88	0.45
C_2	0.96	0.92	0.90	0.58
C_3	1.0	0.99	0.94	0.80

transform. The false initiation probability is 0.35, which is similar to logic method and Hough transform and worse than that in work [2]. In addition, work [2] has the maximum run time of 7.89 s, but in the proposed method it is 5.15 s(less than

Table 1.3 Performance comparison with different track initiation method

Evaluation indices	Logic method	Hough transform	Work [4]	The proposed method
F	0.33	0.35	0.24	0.35
C_1	0.67	0.64	1.0	0.96
C_2	0.70	0.61	0.96	0.92
C_3nn	0.63	0.71	0.93	0.99

sampling interval 10 s). So, the proposed track initiation method is effective in such scenario.

1.4 Conclusions

In term of the drawback of traditional track initiation method, a hybrid track initiation method, which is based on ant-based clustering and Hough transform, is proposed. Simulation results show that the proposed method is able to extract true tracks in the cluttered environment when low measurement noise is considered. Compared to other methods of track initiation, the proposed method has better performance of initiation than that of logic method and Hough transform, so the hybrid method of track initiation based on ant-based clustering and Hough transform is effective. The future research will focus on a method of adaptive parameter, from which universality and robustness can be achieved.

Acknowledgments This work is supported by National Natural Science Foundation of China (No. 60804068), Natural Science Foundation of Jiangsu Province (No.BK2010261), and Cooperation Innovation of Industry, Education, and Academy of Jiangsu Province (No. BY2010126).

References

1. Leung H, Hu Z, Blanchette M (1996) Evaluation of multiple target track initiation techniques in real radar tracking environments [J]. IEE Proc Radar Sonar Navig 143(18):246–254
2. Xu BL, Chen QL, Wang ZQ (2008) A new track initiation method for multi-target tracking [A]. In: Proceedings of the 3rd IEEE conference on industrial electronics and applications [C], pp 1551–1556
3. Xu BL, Chen QL, Wang ZQ (2008) Ants for track initiation of bearings-only tracking [J]. Simul Model Pract Theory 16(6):626–638
4. Turkmen I, Guney K (2004) Multi-sensor single target tracking using genetic algorithm [A]. In: Proceedings of the IEEE 12th signal processing and communications applications conference [C], pp 591–594
5. Yang L, Hu WW, Yang SY, Pu SJ (2006) Application of particle swarm optimization in multi-sensor multi-target tracking [A]. The 1st international symposium on systems and control in aerospace and astronautics [C], pp 715–719

6. Monmarché N, Slimane M, Venturini G (1999) On improving clustering in numerical databases with artificial ants [A]. Advances in artificial life, 5th European conference, ECAL'99. In: Proceedings (Lecture Notes in Artificial Intelligence, vol 1674)[C], pp 626–635
7. Zhu JH, Xu BL, Wang ZQ (2010) Hybrid method of track initiation for Bearings-only target tracking [J]. Inf Control 39(5):596–600

Chapter 2
On the Design of Incremental Predictive Functional Control Algorithm

Wei Wu, Qiang Yang and Wenjun Yan

Abstract To improve the predictive functional control (PFC) algorithm, this paper presents an incremental PFC algorithm designed through the combination of incremental control and PFC, and hence it has the advantages of both approaches. The Internal Model Control (IMC) framework is presented from the control law and is used to analyse and evaluate the performance of the controller design and closed-loop system. The proposed algorithm could effectively guarantee rapid unbiased-tracking with enhanced system robustness. The suggested algorithm has been assessed through a set of numerical experiments with a focus on the impact of the number of basis functions in the control law on the system performance and the simulation result proves the effectiveness of the algorithm.

Keywords Incremental predictive functional control · IMC · Unbiased-tracking

2.1 Introduction

The Model Predictive Control (MPC), as a model-based advanced control technology, was originally proposed in [1] and has become one of the most widely used control strategies in a variety of industrial applications. The Predictive Functional Control (PFC) is considered the third generation of MPC and it is different from the conventional approaches, as the framework of control input is a critical aspect for the control law. MPC enables rapid response and accuracy in the

W. Wu · Q. Yang · W. Yan (✉)
College of Electrical Engineering, Zhejiang University, Hangzhou,
People's Republic of China
e-mail: wj.yan@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_2,

control performance. PFC is used in rapid-tracking system, e.g. robot system [2] and slow time-varying industrial control systems [3, 4]. Based on PFC, many extension proposals have been developed in the literature to further improve the algorithm in the past decades, including the combination with intelligent control [5–9], e.g. fuzzy control, neural networks and support vector machine. In general, the control input is a linear combination of a collection of basis functions in PFC to make the input more regular. However, the control performance is often unsatisfactory in the case of model mismatch [10] and can barely guarantee unbiased-tracking in the consideration of constraint handling. In this paper, we attempt to propose and evaluate a novel incremental PFC algorithm to address this technical challenge.

2.2 Incremental Predictive Functional Control

To start with the algorithmic design, we first consider the following CARIMA (controlled autoregressive and integrated moving average) model:

$$A(q^{-1})y(k) = B(q^{-1})u(k-1) + C(q^{-1})\xi(k)/\Delta \tag{2.1}$$

where $\Delta = 1 - q^{-1}$, $A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n}$, $B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m}$, $C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_t q^{-t}$ with the following notations: $u(k)$-system input, $y(k)$-system output, $\{\xi(k)\}$-white noises with zero mean.

2.2.1 Multi-Step Incremental Predictive Model

In practice, the noise after k instants, $\xi(k+i)$, is unmeasurable, so the exact model cannot be obtained. In order to approximate the true model, we adopt the estimation of model parameters, $\hat{A}(q^{-1})$ and $\hat{B}(q^{-1})$, to set up the following predictive model

$$\hat{A}(q^{-1})y_m(k) = \hat{B}(q^{-1})u(k-1) \tag{2.2}$$

where $y_m(k)$ is the model predictive value of $y(k)$. The error between the actual value and predictive value of output is denoted as $d(k) = y(k) - y_m(k)$.

The potential errors cannot be measured due to the existence of noise, therefore, the predictive output after i steps can be expressed as follows:

$$\begin{aligned}\hat{y}(k+i) = &-\hat{a}_1 y(k+i-1) - \cdots - \hat{a}_n y(k+i-n) \\ &+ b_0\hat{u}(k+i-1) + \cdots + b_m\hat{u}(k+i-1-m) + d(k)\end{aligned} \tag{2.3}$$

Considering multi-steps prediction and defining the step size hp, the predictive output can be expressed as follows:

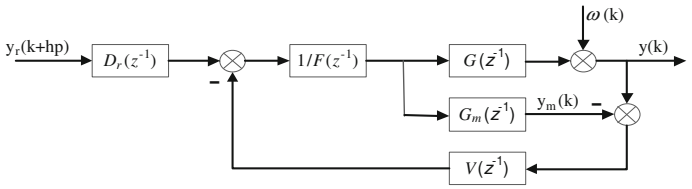


Fig. 2.1 The closed-loop system framework where $\omega(k) = \xi(k)/\Delta$

$$[A_1 \quad A_2]\begin{bmatrix} y_p \\ y_f \end{bmatrix} = [B_1 \quad B_2]\begin{bmatrix} u_p \\ u_f \end{bmatrix} + Cd_f \tag{2.4}$$

where

$$[A_1 \quad A_2] = \begin{bmatrix} \hat{a}_n & \cdots & \hat{a}_1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & \hat{a}_n & \cdots & \hat{a}_1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \hat{a}_n & \cdots & \hat{a}_1 & 1 & 0 \\ 0 & \cdots & 0 & 0 & \hat{a}_n & \cdots & \hat{a}_1 & 1 \end{bmatrix}_{hp\times(n+hp)} \tag{2.5}$$

A_1 is the left $hp \times n$ matrix, and A_2 is the right $hp \times hp$ matrix.

$$[B_1 \quad B_2] = \begin{bmatrix} \hat{b}_m & \cdots & \hat{b}_1 & \hat{b}_0 & 0 & 0 & \cdots & 0 \\ 0 & \hat{b}_m & \cdots & \hat{b}_1 & \hat{b}_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \hat{b}_m & \cdots & \hat{b}_1 & \hat{b}_0 & 0 \\ 0 & \cdots & 0 & 0 & \hat{b}_m & \cdots & \hat{b}_1 & \hat{b}_0 \end{bmatrix}_{hp\times(m+hp)} \tag{2.6}$$

B_1 is the left $hp \times n$ matrix, B_2 is the right $hp \times hp$ matrix, and C is a $hp \times hp$ identity matrix.

$y_p = [y(k+1-n), y(k+2-n), \cdots, y(k)]^T$, the past output vector;

$y_f = [\hat{y}(k+1), \hat{y}(k+2), \cdots, \hat{y}(k+hp)]^T$, the future predictive output vector;

$u_p = [u(k-m), u(k+1-m), \cdots, u(k-1)]^T$, the past input vector;

$u_f = [u(k), u(k+1), \cdots, u(k+hp-1)]^T$, the control input vector to be solved;

$d_f = [d(k), 0, \cdots, 0]^T_{1\times hp}$, the future error.

It is obvious that A_2 is unsingular, so y_f can be transformed as:

$$y_f = -A_2^{-1}A_1 y_p + A_2^{-1}B_1 u_p + A_2^{-1}B_2 u_f + A_2^{-1}Cd_f \tag{2.7}$$

As a result, the multi-step predictive model is obtained.

To obtain the incremental predictive model, y_f is separated in the following form:

$$
\begin{aligned}
y_f &= \begin{bmatrix} y(k+1) \\ y(k+2) \\ \vdots \\ y(k+hp) \end{bmatrix} = \begin{bmatrix} y(k)+\Delta y(k+1) \\ y(k)+\Delta y(k+1)+\Delta y(k+2) \\ \vdots \\ y(k)+\Delta y(k+1)+\cdots+\Delta y(k+hp) \end{bmatrix} \\
&= \begin{bmatrix} y(k) \\ y(k) \\ \vdots \\ y(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta y(k+1) \\ \Delta y(k+2) \\ \vdots \\ \Delta y(k+hp) \end{bmatrix}
\end{aligned} \tag{2.8}
$$

define $S = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ & & \ddots & & \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}_{hp\times hp}$, $Y(k) = \begin{bmatrix} y(k) \\ y(k) \\ \vdots \\ y(k) \end{bmatrix}_{hp\times 1}$, the above formula is simply denoted as:

$$
\begin{aligned}
y_f &= Y(k) + S\Delta y_f \\
&= Y(k) + \left[-SA_2^{-1}A_1\Delta y_p + SA_2^{-1}B_1\Delta u_p + SA_2^{-1}B_2\Delta u_f + SA_2^{-1}C\Delta d_f\right]
\end{aligned} \tag{2.9}
$$

So the incremental predictive model is obtained.

2.2.2 Reference Trace

The ultimate control objective is to make the system output change in a designed reference trace which is in a first-order form.

$$
y_r = \begin{bmatrix} y_{rl}(k+1) \\ \vdots \\ y_{rl}(k+hp) \end{bmatrix},\ y_{rl}(k+i) = \alpha^i y(k) + (1-\alpha^i)w(k+i),\ y_{rl}(k) = y(k) \tag{2.10}
$$

where $y_{rl}(k+i)$, $w(k+i)$and α is the value after trace i steps, the set-point after i steps, and the adjustable parameter.

The above reference trace can make the control slighter and system output tend to the set-point smoothly. The big value of α can improve the system flexibility and robustness, while it can degrade the rapid-tracking.

2.2.3 Predictive Functional Control

The control input $u(k)$ consists of a set of basis functions, which can be expressed by $u(k+i) = \sum_{j=1}^{L} \mu_j f_j(i)$ in PFC to make the control more regular and to guarantee that the output can satisfy the control requirement. Instead of this conventional practice, we choose $\Delta u(k)$ as the control variable and separate it into basis functions. The control law is as follows:

$$\Delta u(k+i) = \mu_0 + \mu_1(i+1) + \cdots + \mu_L(i+1)^L \tag{2.11}$$

$$u(k) = u(k-1) + \Delta u(k) \tag{2.12}$$

Define $\mu = [\,\mu_0 \quad \cdots \quad \mu_L\,]^T$, Δu_f can be written as the combination of $\mu_0, \cdots, \mu_L$. If we can compute μ, the control law easily can be obtained. In order to guarantee μ can be solved, $L+1 \le hp$ must be satisfied.

$$\Delta u_f = \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \Delta u(k+hp-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^L \\ \vdots & \vdots & \vdots & \vdots \\ 1 & hp & \cdots & hp^L \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_L \end{bmatrix} \tag{2.13}$$

Define $T = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^L \\ \vdots & \vdots & \vdots & \vdots \\ 1 & hp & \cdots & hp^L \end{bmatrix}_{hp\times(L+1)}$, where $T_{ij} = i^{j-1}$, the above formula can be expressed as:

$$\Delta u_f = T\mu \tag{2.14}$$

Taking the above formula to the formula (2.9), we can obtain:

$$\begin{aligned} y_f &= Y(k) + S\Delta y_f \\ &= Y(k) + \left[-SA_2^{-1}A_1\Delta y_p + SA_2^{-1}B_1\Delta u_p + SA_2^{-1}B_2T\mu + SA_2^{-1}C\Delta d_f\right] \end{aligned} \tag{2.15}$$

Considering the constraint handling, we choose the quadratic performance index with weighted control variables as the cost function:

$$J = (y_f - y_r)^T Q(y_f - y_r) + \Delta u_f^T R \Delta u_f \tag{2.16}$$

The control objective is to minimize the cost function J and the optimal control law can be solved and expressed as:

$$\begin{aligned} \mu = &\left[(SA_2^{-1}B_2T)^T Q(SA_2^{-1}B_2T) + T^T RT\right]^{-1}(SA_2^{-1}B_2T)^T Q \\ &\times (y_r - Y(k) + SA_2^{-1}A_1\Delta y_p - SA_2^{-1}B_1\Delta u_p - SA_2^{-1}C\Delta d_f) \end{aligned} \tag{2.17}$$

Define $K_{uf} = SA_2^{-1}B_2T$, $K_{yp} = SA_2^{-1}A_1$, $K_{up} = SA_2^{-1}B_1$, $K_{df} = SA_2^{-1}C$, the control law could be expressed as:

$$\mu = (K_{uf}^T Q K_{uf} + T^T R T) K_{uf}^T Q (y_r - Y(k) + K_{yp}\Delta y_p - K_{up}\Delta u_p - K_{df}\Delta d_f) \quad (2.18)$$

$$u(k) = u(k-1) + \Delta u(k) = u(k-1) + \sum_{i=0}^{L} \mu_i \quad (2.19)$$

2.2.4 The IMC Framework of the Incremental Predictive Functional Control

Define $G(z^{-1}) = B(z^{-1})/A(z^{-1}), G_m(z^{-1}) = \hat{B}(z^{-1})/\hat{A}(z^{-1})$, the system output can be expressed as:

$$y(k) = G_m(k)u(k) + d(k) \quad (2.20)$$

Based on the control law (2.18) and (2.19),

$$\Delta u(k) = \sum_{i=0}^{L} \mu_i = [\,1 \quad \cdots \quad 1\,]_{1\times(L+1)} [\,\mu_0 \quad \cdots \quad \mu_L\,]^T \quad (2.21)$$

By defining $H = [\,1 \quad \cdots \quad 1\,]_{1\times(L+1)}$, the above fomula can be written as:

$$\begin{aligned}(1 - z^{-1})u(k) &= H(K_{uf}{}^T Q K_{uf} + T^T R T) K_{uf}{}^T Q (y_r \\ &\quad - Y(k) + K_{yp}\Delta y_p - K_{up}\Delta u_p - K_{df}\Delta d_f) \\ &= H'(y_r - Y(k) + K_{yp}\Delta y_p - K_{up}\Delta u_p - K_{df}\Delta d_f)\end{aligned} \quad (2.22)$$

where $H' = H(K_{uf}{}^T Q K_{uf} + T^T R T) K_{uf}{}^T Q$,

$$\Delta y_p = \begin{bmatrix} y(k+1-n) - y(k-n) \\ y(k+2-n) - y(k+1-n) \\ \vdots \\ y(k) - y(k-1) \end{bmatrix} == \begin{bmatrix} z^{-(n-1)} - z^{-n} \\ z^{-(n-2)} - z^{-(n-1)} \\ \vdots \\ 1 - z^{-1} \end{bmatrix} (G_m(z^{-1})u(k) + d(k))$$
$$= M_{11}(z^{-1})u(k) + M_{12}(z^{-1})d(k)$$

$$\Delta u_p = \begin{bmatrix} u(k-m) - u(k-m-1) \\ u(k+1-m) - u(k-m) \\ \vdots \\ u(k-1) - u(k-2) \end{bmatrix} = \begin{bmatrix} z^{-m} - z^{-(m+1)} \\ z^{-(m-1)} - z^{-m} \\ \vdots \\ z^{-1} - z^{-2} \end{bmatrix} u(k)$$

$$\Delta d_f = \begin{bmatrix} d(k) - d(k-1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - z^{-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} d(k)$$

$$y_r = \begin{bmatrix} y_{rl}(k+1) \\ y_{rl}(k+2) \\ \\ y_{rl}(k+hp) \end{bmatrix} = \begin{bmatrix} z^{-(hp-1)} \\ z^{-(hp-2)} \\ \vdots \\ 1 \end{bmatrix} y_{rl}(k+hp)$$

$$Y(k) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad y(k) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} (G_m(z^{-1})u(k) + d(k)) = M_{51}(z^{-1})u(k) + M_{52}(z^{-1})d(k)$$

Therefore, Eq. 2.22 can be further derived as:

$$\begin{aligned} (1 - z^{-1})u(k) &= H'(y_r - Y(k) + K_{yp}\Delta y_p - K_{up}\Delta u_p - K_{df}\Delta d_f) \\ &= H'[(M_4 y_{rl}(k+hp) - (M_{51}u(k) + M_{52}d(k)) \\ &\quad + K_{yp}(M_{11}u(k) + M_{12}d(k)) - K_{up}M_2u(k) - K_{df}M_3d(k)] \\ &= H'[M_4 y_{rl}(k+hp) - (M_{51} - K_{yp}M_{11} + K_{up}M_2)u(k) \\ &\quad - (M_{52} - K_{yp}M_{12} + K_{df}M_3)d(k) \end{aligned} \tag{2.23}$$

The above formula can be further transformed as:

$$u(k) = \frac{1}{F(z^{-1})}\left[D_r(z^{-1})y_r(k+hp) - V(z^{-1})d(k)\right] \tag{2.24}$$

where $F(z^{-1}) = 1 - z^{-1} + H'(M_{51} - K_{yp}M_{11} + K_{up}M_2)$,$D_r(z^{-1}) = H'M_4(z^{-1})$, $V(z^{-1}) = H'(M_{52} - K_{yp}M_{12} + K_{df}M_3)$.

Each of $F(z^{-1})$, $D_r(z^{-1})$, $V(z^{-1})$ can be expressed by finite linear combination of polynomials of z^{-1}.

It can be seen from (2.24) that the stability of controller is determined by $F(z^{-1})$. Figure 2.1 shows the system framework, and it indicates the IMC framework, which makes the system robust.

2.2.5 The Performance of Closed-Loop System

The output function and error function of closed-loop system can be derived from Fig. 2.1.

$$\begin{aligned} y(k) = & \frac{G(z^{-1})D_r(z^{-1})}{F(z^{-1}) + V(z^{-1})(G(z^{-1}) - G_m(z^{-1}))} y_r(k+hp) \\ & + \frac{F(z^{-1}) - V(z^{-1})G_m(z^{-1})}{F(z^{-1}) + V(z^{-1})(G(z^{-1}) - G_m(z^{-1}))} \omega(k) \end{aligned} \tag{2.25}$$

$$\begin{aligned} e(k) &= y_r(k) - y(k) \\ &= (z^{-hp} - \frac{G(z^{-1})D_r(z^{-1})}{F(z^{-1}) + V(z^{-1})(G(z^{-1}) - G_m(z^{-1}))}) y_r(k+hp) \\ &- \frac{F(z^{-1}) - V(z^{-1})G_m(z^{-1})}{F(z^{-1}) + V(z^{-1})(G(z^{-1}) - G_m(z^{-1}))} \omega(k) \end{aligned} \tag{2.26}$$

The eigenvalue equation is $F(z^{-1}) + V(z^{-1})(G(z^{-1}) - G_m(z^{-1})) = 0$. In the case of system unstability, we can set a filter $G_v(z^{-1})$ in the feedback loop to solve the problem.

In the steady state, $k \to \infty$, $z^{-1} \to 1$, so, $F(z^{-1}) \to H' G_m(1)$, $D_r(z^{-1}) \to H'$, $V(z^{-1}) \to H'$, (2.26) derives $e(\infty) \to 0$. This implies that no matter whether the model mismatches or not, system steady error is always zero.

2.3 Numerical Experiments and Result

In this section, a simulation scenario is adopted to evaluate and assess the proposed algorithm with a focus on exploiting the impacts of parameter L on the system performance.

To simplify the analysis, we define $C(q^{-1}) = 1$ in system model (1), and the chosen model is

$$y(k) = 1.7y(k-1) - 0.9y(k-2) + 0.6u(k-1) + 0.4u(k-2) + 0.02\xi(k)/\Delta \tag{2.27}$$

where $\xi(k)$ is a white-noise signal with amplitude of 1, with a set of predefined values $hp = 5$, $\alpha = 0.95$, $Q = R = 1$, the set-point is 1.

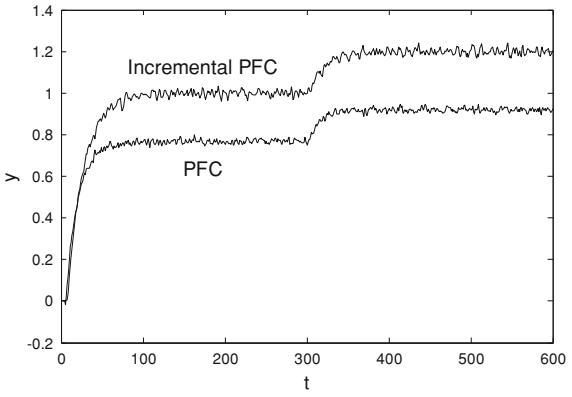


Fig. 2.2 Incremental and conventional PFC

2.3.1 The Closed-Loop Performance

In general, the basis functions in PFC is the combination of step function and ramp function, that is $u(k+i) = \mu_0 + \mu_1 i$. Figure 2.2 presents the result through comparing the proposed algorithm with the conventional PFC algorithm where the set-point changes to 1.2 from 1 when $t = 300$ and the parameter $L = 1$ in the incremental PFC. Under the condition of restricted input, the Incremental PFC eliminates the steady error perfectly and remains good unbiased-tracking subject to set-point changes.

Figure 2.3 shows the system output when $t = 300$, model parameter a_1 changes from 1.7 to 1.8 and $L = 1$, which demonstrates that the proposed algorithm can guarantee system stability and good unbiased-tracking, and hence enhanced system robustness.

2.3.2 The Effect of Parameter L on System Performance

In incremental PFC control law, the number of basis functions L is adjustable, which could affect the overall system performance. We carry out three simulation experiments with $L = 0$, $L = 1$, $L = 4$ to study the impact of the number of basis functions on the system performance. The outputs are consistent with the result given in Fig. 2.2. To illustrate the difference in a more clear manner, we present the result with the range $t = 200$–400, as shown in Figs. 2.4, 2.5 and 2.6. Figures 2.7, 2.8 and 2.9 show corresponding control input. When $t = 300$, a_1 changes from 1.7 to 1.8. Figures 2.10, 2.11 and 2.12 present three simulation results with the value of $L = 0$, $L = 1$, $L = 4$. It is obviously indicated from Figs. 2.4, 2.5, 2.6, 2.10, 2.11 and 2.12 that the big L makes system output fluctuate

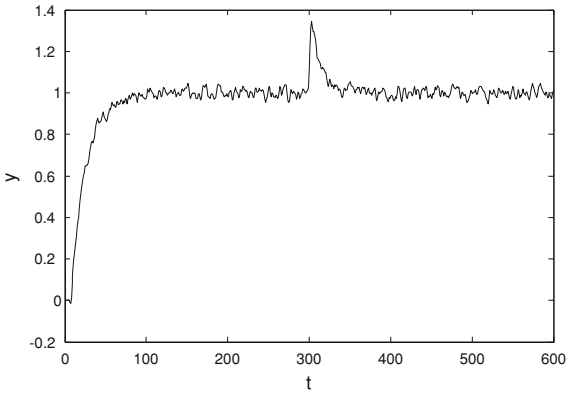


Fig. 2.3 Output in the case of parameter mutation

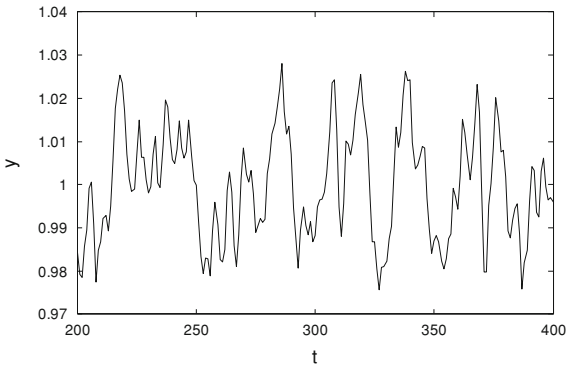


Fig. 2.4 System output ($L = 0$)

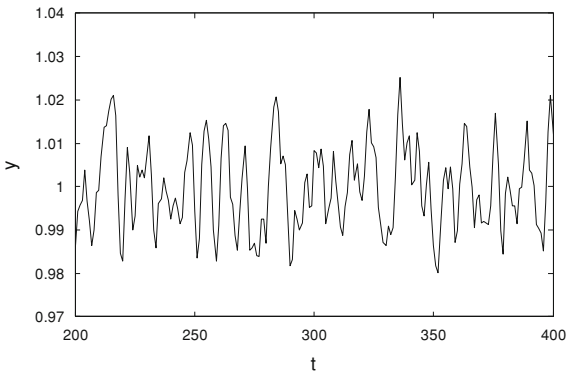


Fig. 2.5 System output ($L = 1$)

more smoothly and help improve system robustness when model mismatches. Actually, the increase of basis function contributes to the fitting of control input to the optimal value so that it reduces the fluctuation when the reference trace is

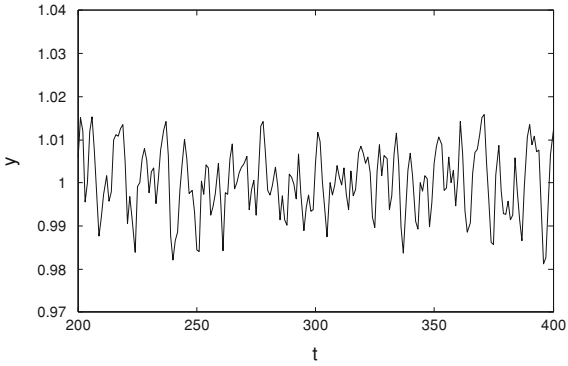


Fig. 2.6 System output ($L = 4$)

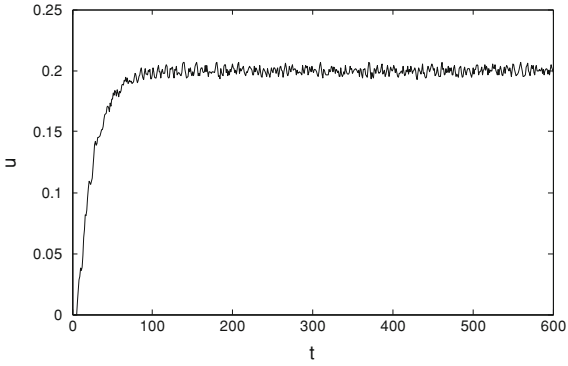


Fig. 2.7 Input when $L = 0$

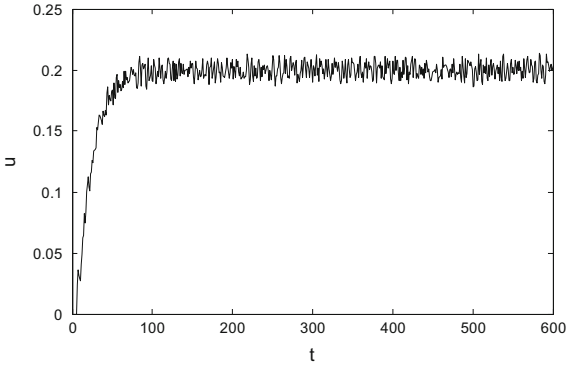


Fig. 2.8 Input when $L = 1$

smooth and strengthens system robustness. Figures 2.7, 2.8 and 2.9 illustrate that control input fluctuates more fiercely as the value of L is bigger in this simulation case, which is bad for system performance. It can also be known that the big L can cause more compute works.

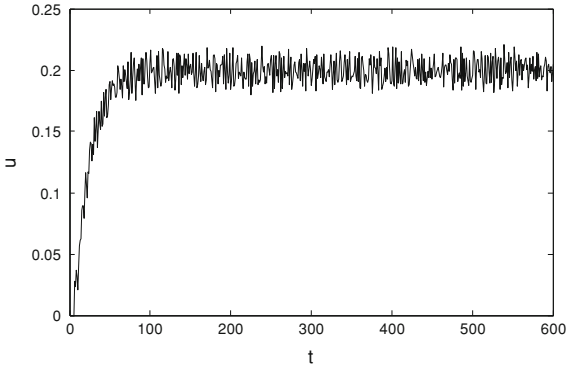


Fig. 2.9 Input when $L = 4$

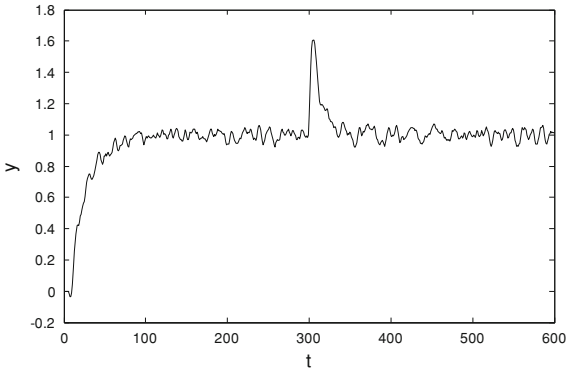


Fig. 2.10 System output in the case of parameter mutation ($L = 0$)

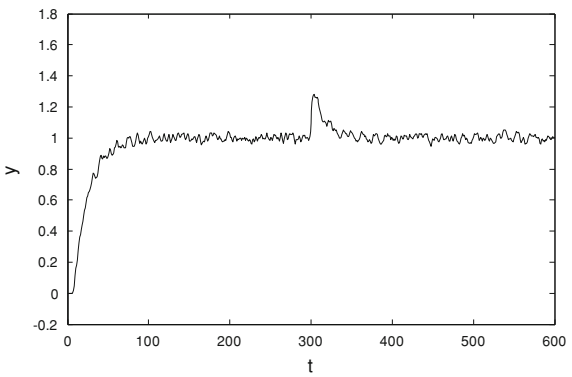


Fig. 2.11 System output in the case of parameter mutation ($L = 1$)

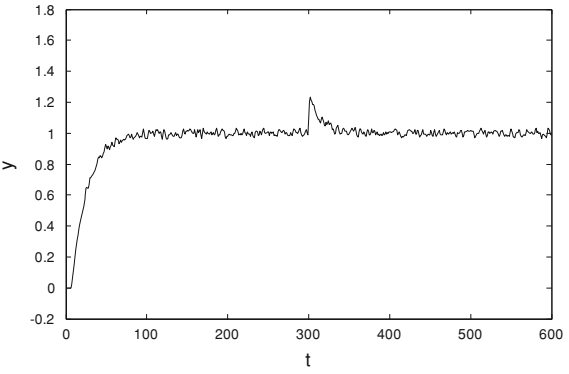


Fig. 2.12 System output in the case of parameter mutation ($L = 4$)

2.4 Conclusions and Remarks

An incremental PFC algorithm is proposed in this paper. The IMC framework is provided and the closed-loop performance is studied. The suggested algorithm design could guarantee unbiased-tracking, rapid response, accuracy and enhanced robustness even under the situation of model mismatch. The proposed algorithm has been implemented and assessed through numerical simulation experiments. The result demonstrates its effectiveness and indicates that the increase of basis function could improve smoothness of system output and system robustness while it will cause more compute works and may deteriorate the smoothness of control input so that we should choose L according to the actual need.

Acknowledgements This work is supported in part by the National High Technology Research and Development Program of China ("863" Program) with grant No. 2007AA05Z232 and National Natural Science Foundation of China (No. 51107113).

References

1. Richalet J, Rault A, Testud J, Papon J (1978) Model predictive heuristic control: application to industrial process. Automatic 14:413–428
2. Vivas A, Poignet P (2005) Predictive functional control of a parallel robot. Control Eng Pract 13:863–874
3. Zhang R et al (2010) A simplified linear iterative predictive functional control approach for chamber pressure of industrial coke furnace. J Process Control 20:464–471
4. Bouhenchir H, Cabassud M, Le Lann MV (2006) Predictive functional control for the temperature control of a chemical batch reactor. Comput Chem Eng 30:1141–1154
5. Zhao H, Cao J, Li Z, Liu Y (2008) Nonlinear predictive functional control of recursive subspace model using support vector machine. In: Chinese control and decision conference, pp 4909–4913
6. Skrjanc I, Matko D (2001) Fuzzy predictive functional control in the state space domain. J Intell Rob Syst 31:283–297

7. Skrjanc I, Blazic S (2005) Predictive functional control based on fuzzy model: design and stability study. J Intell Rob Syst 43:283–299
8. Li X, Ling H, Zhu J (2009) Predictive functional control based on artificial neural networks and it's application of coordinated control systems of fossil power plant. In: 21st Chinese control and decision conference. pp 5847–5852
9. Zhang R, Wang S (2008) Support vector machine based predictive functional control design for output temperature of coking furnace. J Process Control 18:439–448
10. Hu J, Pan H, Su H, Chu J (2001) Performance analysis for predictive functional control system. Control Theory Appl 18:774–778

Chapter 3
A Novel Fuzzy Data Association Algorithm with the Heading-Angle Aided

Pengfei Li, Tian Qin and Li Ren

Abstract In this chapter, a new data association method based on fuzzy logic is proposed for the multi-target tracking. Firstly, the heading-angle and distance are analyzed to be the effective parameters for separating the multi-target, the definition of heading-angle and the method of how to calculate the heading-angle have been introduced, the state including the heading-angle is updated using the cubature Kalman filter, at last the fuzzy logic inference method is used for data association. Simulation results show that the proposed method have better performance in data association than the Nearest Neighbor (NN) algorithm and Joint Probability Data Association (JPDA) algorithm.

Keywords Data association · Multi-target tracking · Fuzzy logic · Heading angle

P. Li (✉) · T. Qin · L. Ren
Air Defense Forces Command Academy,
Zhengzhou 450052, People's Republic of China
e-mail: roc5683@gmail.com

P. Li
ATR Key Lab, Shenzhen University,
Shenzhen 518060, People's Republic of China

T. Qin
e-mail: qintian790206@sina.com

L. Ren
e-mail: renli_xu@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_3,

3.1 Introduction

Data Association is the highlight in the multi-target tracking problem, the scholars have done many deep researches, and proposed a lot of algorithms, such as the Nearest Neighbor (NN) algorithm [1, 2], the Joint Probability Data Association (JPDA) algorithm [3] and the Multi-hypothesis Testing algorithm [4]. Recently, scholars use the fuzzy method to deal with the data association, one is the fuzzy clustering, and the other is the fuzzy logic. The fuzzy logic data association method normally using the residual of distance, azimuth and elevation as the input variables for the fuzzy inference system, but the residual of azimuth and elevation could not reflect the target trajectories directly. The heading angle has been paid more and more attention in the multi-target tracking [5]. Because the heading could reflect the target flight course, using the heading angle could differentiate the different batches of target. Based on this point, a heading-angleaided data association algorithm is proposed in this chapter. The distance and heading angle are used as the input variables of the fuzzy logic inference system for the data association.

3.2 The Definition of the Heading Angle

The course angle is the angle between the projection of target velocity in the level surface and the reference direction. In this chapter, the reference direction is the positive direction of the y axis. c_k is the heading angle in Fig. 3.1.

As Fig. 3.1 shows, the radar locates at the origin, T denotes the target trajectory, $\mathbf{z}_{k+1}$ denotes the radar measurement at the $k+1$th scan, $\mathbf{z}_{k+1} = [\, r_{k+1} \quad \alpha_{k+1} \,]$, r_{k+1} is the distance between the radar and the target, α_{k+1} is the azimuth measured by radar. $\hat{\mathbf{x}}_{k|k} = [\, x_{k|k} \quad \dot{x}_{k|k} \quad y_{k|k} \quad \dot{y}_{k|k} \,]^T$ denotes the $k+1$th estimate state, $x_{k|k}$ and $y_{k|k}$ is the target position at the x direction and y direction, $\dot{x}_{k|k}$ and $\dot{y}_{k|k}$ is the velocity. c'_{k+1} is called measured heading angle, it is defined as follows:

$$c'_{k+1} = \arctan\left(\frac{r_{k+1}\cos(\alpha_{k+1}) - x_{k|k}}{r_{k+1}\sin(\alpha_{k+1}) - y_{k|k}}\right). \tag{3.1}$$

3.3 The Data Association Based on Fuzzy Logic

In this chapter, the distance and heading angle will be used as the input variables of the fuzzy logic inference system, and the algorithm is as follows:

Step 1. Calculation of the measurements predictions and covariance of innovations

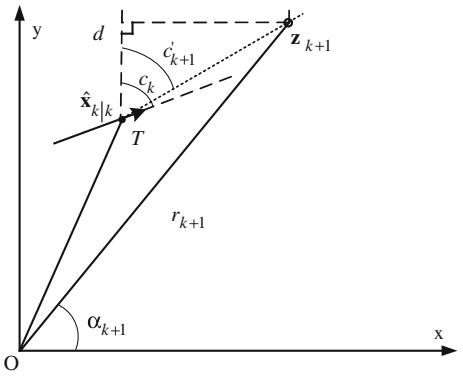


Fig. 3.1 The definition of course angle

Suppose targets have been gained by the track initiation, and measurements also gained at the *k*th moment. In order to use the tracking thresholds and fuzzy logic for data association, firstly, the nonlinear filter—cubature Kalman filter (CKF) is used to calculate the prediction and innovation covariance.

$$\mathbf{P}^j_{k-1|k-1} = \mathbf{S}^j_{k-1|k-1}\left(\mathbf{S}^j_{k-1|k-1}\right)^T \tag{3.2}$$

$$\boldsymbol{\chi}^j_{i,k-1|k-1} = \mathbf{S}^j_{k-1|k-1}\boldsymbol{\xi}_i + \hat{\mathbf{x}}^j_{k-1|k-1} \tag{3.3}$$

$$\boldsymbol{\chi}^j_{i,k|k-1} = f\left(\boldsymbol{\chi}^j_{i,k-1|k-1}, \mathbf{u}_{k-1}\right) \tag{3.4}$$

$$\hat{\mathbf{x}}^j_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\chi}^j_{i,k|k-1} \tag{3.5}$$

$$\mathbf{P}^j_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\chi}^j_{i,k|k-1}\boldsymbol{\chi}^{jT}_{i,k|k-1} - \hat{\mathbf{x}}^j_{k|k-1}\hat{\mathbf{x}}^{jT}_{k|k-1} + \mathbf{Q}_{k-1} \tag{3.6}$$

$$\mathbf{P}^j_{k|k-1} = \mathbf{S}^j_{k|k-1}\left(\mathbf{S}^j_{k|k-1}\right)^T \tag{3.7}$$

$$\boldsymbol{\chi}^j_{i,k|k-1} = \mathbf{S}^j_{k|k-1}\boldsymbol{\xi}_i + \hat{\mathbf{x}}^j_{k|k-1} \tag{3.8}$$

$$\mathbf{z}^j_{i,k|k-1} = h\left(\boldsymbol{\chi}^j_{i,k|k-1}, \mathbf{u}_k\right) \tag{3.9}$$

$$\hat{\mathbf{z}}^j_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{z}^j_{i,k|k-1} \tag{3.10}$$

$$\mathbf{P}^{j}_{zz,k|k-1} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{z}^{j}_{i,k|k-1}\mathbf{z}^{jT}_{i,k|k-1} - \hat{\mathbf{z}}^{j}_{k|k-1}\hat{\mathbf{z}}^{jT}_{k|k-1} + \mathbf{R}_k \tag{3.11}$$

where $m = 2n$, n is the dimension of the state vector; $\boldsymbol{\xi}_i$ is the ith cubature point. The cubature points will be given as [6]: $\boldsymbol{\xi}_i = \sqrt{\frac{m}{2}}[1]_i$.

Step 2. The first correlation with the tracking threshold

To limit the numbers of associations, firstly, only those measurements which lie in the validation region will be dealt with. In this chapter, the eclipse threshold is used for data association. In the kth moment, the ith measurement, $\mathbf{z}^{i}_{k}(i = 1, \ldots, m_k)$, is validated for the jth target if its Mahalanobis distance from the predicted location of $\hat{\mathbf{z}}^{j}_{k|k-1}$of a jth target is less than a threshold, i.e

$$d^2_{ij} = v'_{ij}\left(\mathbf{P}^{j}_{zz,k|k-1}\right)^{-1}\boldsymbol{v}_{ij} \le \gamma \tag{3.12}$$

where $\boldsymbol{v}_{ij}$ is the innovation $\boldsymbol{v}_{ij} = \mathbf{z}^{i}_{k} - \hat{\mathbf{z}}^{j}_{k|k-1}$.

$\mathbf{P}^{j}_{zz,k|k-1}$ is the innovation covariance of the jth target. The threshold γ is selected to produce a predetermined probability of erroneous rejection of the correct return. It is assumed for simplicity that the true return is a Gaussian random vector. If the ith measurement correctly corresponds to the jth target, then d^2_{ij} has a χ_2 distribution with n degrees of freedom, where n is the order of the state vector.

Step 3. The final correlation with the fuzzy logic

When the measurements are at a high density, the measurement may not be selected for the jth target because several measurements are probably validated for the tracking threshold of the jth target. So it is necessary to do another association in order to confirm the final measurement which matches the jth target best. In this chapter, the distance and heading angle is used for the final correlation. Suppose the measurement at kth moment is $\mathbf{z}^{i}_{k} = \begin{bmatrix} r^{i}_{k} & \alpha^{i}_{k} \end{bmatrix}^T$, and the $c^{'}_{k}$ is calculated using Eq. 3.1, the prediction state has been obtained by Step 1, $\mathbf{x}^{j}_{k|k-1} = \begin{bmatrix} x^{j}_{k|k-1} & \dot{x}^{j}_{k|k-1} & y^{j}_{k|k-1} & \dot{y}^{j}_{k|k-1} \end{bmatrix}^T$, then the difference of distance and heading angle are defined as follows:

$$\Delta r_{\mathrm{ij}} = \sqrt{\left(r^{i}_{k}\cos\left(\alpha^{i}_{k}\right) - x^{j}_{k|k-1}\right)^2 + \left(r^{i}_{k}\sin\left(\alpha^{i}_{k}\right) - y^{j}_{k|k-1}\right)^2} \tag{3.13}$$

$$\Delta h_{ij} = \left| c^{i}_{k} - \arctan\left(\frac{\dot{x}^{j}_{k|k-1}}{\dot{y}^{j}_{k|k-1}}\right)\right| \tag{3.14}$$

The magnitude of Δr_{ij} and Δh_{ij} may vary from target to target. In order to design a general fuzzy association algorithm, it is necessary to normalize the Δr_{ij}

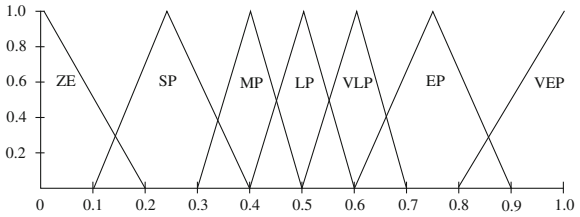


Fig. 3.2 Membership function for the difference of distance

and Δh_{ij} in form of $\Delta r_{ij}'$ and $\Delta h_{ij}'$. Suppose m_p measurements are validated for the tracking threshold of the jth target, we compute $\Delta r_{ij}'$ and $\Delta h_{ij}'$ where

$$\Delta r_{ij}' = \frac{\Delta r_{ij}}{\max\limits_{i=1}^{m_p}(\Delta r_{ij})} \tag{3.15}$$

$$\Delta h_{ij}' = \frac{\Delta h_{ij}}{\max\limits_{i=1}^{m_p}(\Delta h_{ij})} \tag{3.16}$$

After the normalization, we define the membership functions by experience for the difference of distance and the difference of heading angle, also for the output variables, which is shown in Figs. 3.2, 3.3 and 3.4. The values are mapped into some fuzzy sets, labeled in the linguistic terms of zero (ZE), small positive (SP), medium positive (MP), large positive (LP), very large positive (VLP), extremely positive (EP) and very extremely positive (VEP).

With the input and output variables defined above, the fuzzy rules can be expressed as follows:

$$\text{IF } \Delta r_{ij}' = \text{A} \text{ AND } \Delta h_{ij}' = \text{B} \text{ THEN OUTPUT} = \text{C}$$

Here A is one of the fuzzy sets in Fig. 3.2, B is one of the fuzzy sets in Fig. 3.3 and C is one of the fuzzy sets in Fig. 3.4.

Since every such rule used by the proposed fuzzy association can be considered as an implication that defines a fuzzy association, they are concisely written as a linguistic map or fuzzy associative memory here, as shown in Table 3.1.

The rules are designed based on the principle that the smaller the distance between the measurements and the target, the higher the magnitude of the output variables. The smaller the difference of heading angle is, the higher the magnitude of the output variables are.

The magnitude of the output variables are evaluated through the Max–Min compositional rule of inference technique and center of gravity (COG) defuzzification. The measurement with the most membership magnitude will be chosen as the validated measurement for the jth target.

Step 4. The updating of target state and covariance

In this step, the CKF is used to update the state and the covariance.

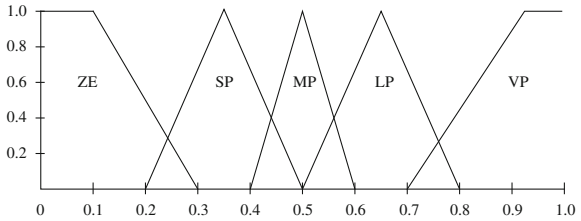


Fig. 3.3 Membership function for the difference of heading angle

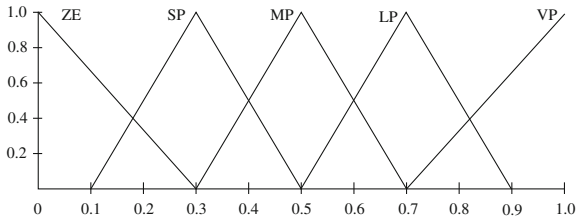


Fig. 3.4 Membership function for the output variables

Table 3.1 Fuzzy associations for output variables

Output variables		$\Delta r_{ij}^{'}$						
		ZE	SP	MP	LP	VLP	EP	VEP
$\Delta h_{ij}^{'}$	ZE	VP	VP	VP	LP	MP	MP	SP
	SP	VP	VP	LP	MP	MP	MP	SP
	MP	VP	LP	MP	MP	MP	SP	ZE
	LP	LP	MP	MP	MP	SP	ZE	ZE
	VP	LP	MP	MP	SP	ZE	ZE	ZE

$$\mathbf{P}_{xz,k|k-1}^{j} = \sum_{i=1}^{m} \omega_i \boldsymbol{\chi}_{i,k|k-1}^{j} \mathbf{z}_{i,k|k-1}^{jT} - \hat{\mathbf{x}}_{k|k-1}^{j} \hat{\mathbf{z}}_{k|k-1}^{jT} \tag{3.17}$$

$$\mathbf{W}_{k}^{j} = \mathbf{P}_{xz,k|k-1}^{j} \left(\mathbf{P}_{zz,k|k-1}^{j} \right)^{-1} \tag{3.18}$$

$$\hat{\mathbf{x}}_{k|k}^{j} = \hat{\mathbf{x}}_{k|k-1}^{j} + \mathbf{W}_{k}^{j} \boldsymbol{v}_{k+1}^{ij} \tag{3.19}$$

$$\mathbf{P}_{k|k}^{j} = \mathbf{P}_{k|k-1}^{j} - \mathbf{W}_{k}^{j} \mathbf{P}_{zz,k|k-1}^{j} \mathbf{W}_{k}^{jT} \tag{3.20}$$

Then the jth target and the measurement matching to it are deleted from the target and measurement lists. We can do the association for the rest of the targets and measurements in the same way, till all the targets find their matching measurements.

3.4 Simulation Results and Analysis

In order to evaluate the performance of the proposed method, one scenario is performed. In the scenario, the proposed method is compared to the JPDA algorithm in the clutter environment.

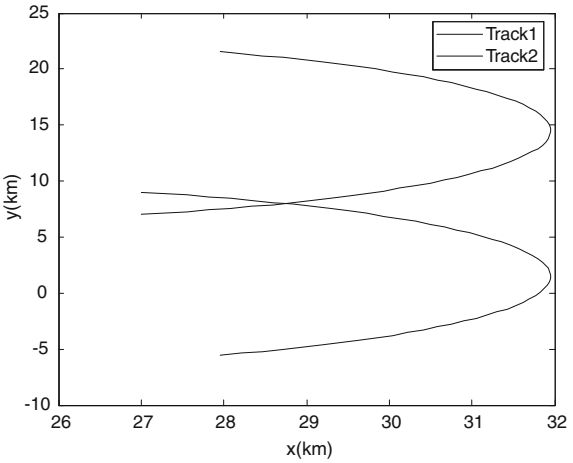


Fig. 3.5 Two original trajectories

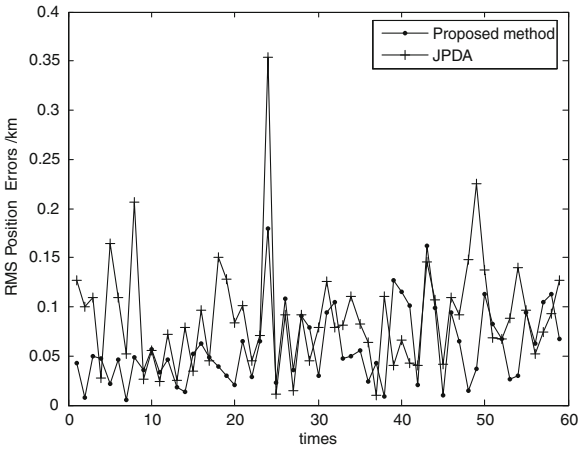


Fig. 3.6 The RMSE of Track 1 position using two methods

In this simulation, we generate two trajectories using the model as follows:

$$\mathbf{x}_k = \begin{bmatrix} 1 & \frac{\sin \Omega T}{\Omega} & 0 & -\frac{(1-\cos \Omega T)}{\Omega} \\ 0 & \cos \Omega T & 0 & -\sin \Omega T \\ 0 & \frac{1-\cos \Omega T}{\Omega} & 1 & \frac{\sin \Omega T}{\Omega} \\ 0 & \sin \Omega T & 1 & \cos \Omega T \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{v}_k \tag{3.21}$$

where the state $\mathbf{x} = [x \quad \dot{x} \quad y \quad \dot{y}]^T$, x and y denote positions, and $\dot{x}$ and $\dot{y}$ denote velocities in the x and y directions, respectively; Ω denote the turn rate, T is the time-interval between two consecutive measurements; the process noise is the $v_k \sim \mathrm{N}(0, \quad Q)$ with a nonsingular covariance $Q = \mathrm{diag}[q_1 M \quad q_2 M \quad q_2 T]$, where $M = \begin{pmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{pmatrix}$; The scalar parameters q_1 and q_2 are related to process

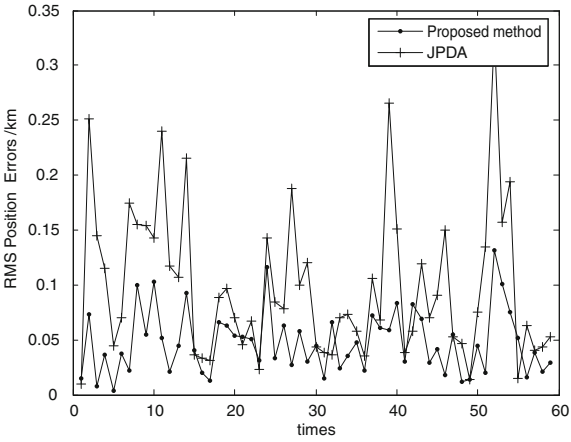


Fig. 3.7 The RMSE of Track 2 position using two methods

noise intensities. The time-interval $T = 3$ s. Initial states of the two targets are as follows: $\times 0 = [27{,}000\text{ m},\ 280\text{ m/s},\ 9{,}000\text{ m},\ -120\text{ m/s}]T$, $\times 1 = [27{,}000\text{ m},\ 280\text{ m/s}, 7{,}000\text{ m}, 120\text{ m/s}]T$, and the turn rate of the four targets are -0.11186 rad/s and 0.11186 rad/s. Clutter measurements, whose number are Poisson distributed with parameter $\lambda = 1$ [number of false measurements per unit area (km^2)], are generated uniformly within the ellipsoid region centered in the predicted measurement for each target. Figure 3.5 shows the two original trajectories. Fifty Monte Carlos runs have been performed.

Figures 3.6 and 3.7 show the root mean square errors (RMSE) of positions for the two trajectories. From the two figures, we can see the proposed method outperforms the JPDA in tracking accuracy. Besides, the mean execution time using JPDA and the proposed method, each in 50 random runs are 25.75 and 10.624 s. It proves that the proposed method is the efficient approach for real application.

3.5 Conclusions

In this chapter, a heading-angle-aided data association algorithm based on fuzzy logic is proposed for multi-target tracking. We analyze that the heading angle is a very important parameter to differentiate the different batches of target. As we know, the distance parameters are commonly used for the data association. So both the heading angle and the distance are used for the data association based on the fuzzy logic. Firstly, the CKF is used to predict the target positions for each trajectory. To the predicted locations as the center, the measurements whose Mahalanobis distances from the appointed predicted location are more than the threshold will be eliminated. Then the heading angle and the distance are used as the input parameters of the fuzzy logic inference system for final correlation. At last, the targets' state vectors will be updated using the CKF. The simulation

results show that the proposed method performs excellently in data association, and it is an efficient algorithm for the real application.

References

1. Singer RA, Sea RG (1971) A new filter for optimal tracking in dense multitarget environment. In: Proceedings of the 9th allerton conference circuit and system theory, University of Illinois, Urbana-Champaign, USA
2. Singer RA, Stein JJ (1971) An optimal tracking filter for processing sensor data of imprecisely determined origin in surveillance system. In: IEEE proceedings of the 10th conference on decision and control
3. Bar-Shalom Y, Fortmann TE (1988) Tracking and data association. Academic Press, Boston
4. Reid DB (1979) An algorithm for tracking multiple targets. IEEE Trans Autom Control 24:843–854
5. Wang J, Jin YG, Dong HC (2009) Course information auxiliary target data association improved method. J Syst Simul 21:3344–3347
6. Arasaratnam I, Haykin S (2009) Cubature Kalman filters. IEEE Trans Autom Control 54:1254–1269

Chapter 4
The Parametric Design Based on Bee Dual Population Evolution on Chaos

Cao Chun-hong, Tang Chuan, Zhao Da-zhe and Zhang Bin

Abstract Geometric constraint problem is equivalent to the problem of solving a set of nonlinear equations substantially. We propose a new optimization algorithm—Bee Dual Population Evolution on Chaos Algorithm and apply it into the geometric constraint solving. By the use of properties of ergodicity, randomicity and regularity of chaos motion, genetic algorithm of bee dual population evolution based on chaos is proposed. On the basis of genetic algorithm based on bee dual population evolution, the algorithm adopts chaos optimization to improve species' initial quality and utilizes the annealing chaotic mutation operator to replace the standard mutation operator in order to avoid the search being trapped in local optimum. The experiment shows that Bee Dual Population Evolution on Chaos has good capability in the geometric constraint solving.

Keywords Parametric design · Geometric constraint solving · Bee Dual population evolution · Chaos

C. Chun-hong (✉) · Z. Bin
Collge of Information Science and Engineering, Northeastern University, Shenyang 110819, People's Republic of China
e-mail: caochunhong@ise.neu.edu.cn

T. Chuan
State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, People's Republic of China

C. Chun-hong
State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, People's Republic of China

Z. Da-zhe
Key Laboratory of Medical Image Computing of Ministry of Education, Northeastern University, Shenyang, Lining 110819, People's Republic of China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_4,

4.1 Introduction

The parametric design is a geometric constraint-solving problem. Geometric constraint-solving approaches are made of three approaches: algebraic-based solving, approach-based solving approach and graph-based solving approach. One constraint describes a relation that should be satisfied. Once a user defines a series of relations, the system will satisfy the constraints by selecting proper state after the parameters are modified. The idea is named model-based constraints. Constraint solver is a segment for the system to solve the constraints.

Based on the conventional Genetic Algorithm (GA), inspired by the process of bee population evolution, we proposed some mechanisms which are fit for the improvement of GA and then improve on them in this paper [1]. A genetic algorithm of bee population evolution has been proposed [2, 3]. Chaos is an phenomena in the nonlinear system and it is universal but it also has an exquisite inherence. Chaos optimization uses the chaotic variable to search. The search process is based on the rules of the chaos motion which has the properties of ergodicity, randomicity and regularity, so it has a better chance to avoid the search being trapped in local optimum [4–7]. Some scholars have done some work in the combination of chaos search and GA, such as, in the literature [6] the mutation in the GA was introduced in chaos and makes use of it in the neural network weights optimization, in the literature [7] it uses the Chaos GA in the multi-objective optimization problem, etc.

4.2 Genetic Algorithm of Bee Dual Population Evolution Based on Chaos

4.2.1 Genetic Algorithm and Chaotic Search

GA is a kind of search method that evolved by the reference of biological evolution law. It mainly consists of selection, crossover and mutation. Selection is to select the individual whose fitness is higher from the current group of individuals; it is higher quality individual operations. Crossover is used to generate new individuals, the general value of crossover probability is $0.5 \sim 0.95$. Variation is mainly used to maintain the population diversity, mutation rate is generally $0.001 \sim 0.2$.

Chaos has a wealth of time and space dynamics, it can go by its own rule within a certain range through all of the state and not repeated, so to optimize the use of chaotic variables search has undoubtedly great advantages.

In the chaos the logistic mapping is frequently used, the following is its iterative:

$$x_{n+1} = \mu_{xn}(1\, x_n) \tag{4.1}$$

Of which: state of the control parameter μ plays a key role in the system. Many experiments show that when x0 $\in$(0,1), $\mu = 4$, the system has been in complete

chaos, then the system can transverse the entire search space not repeatedly, and the ergodicity can improve the global optimization to avoid the search process being trapped in local minimum. Chaotic search is introduced in the GA and it can improve the premature convergence of GA to improve the search efficiency [4].

4.2.2 Genetic Algorithm of Bee Dual Population Evolution

Reference [2] from the proposal of GA, it has solved many problems successfully, and achieved a wide range of applications. But the simple GA in the application process is prone to premature convergence, "prematurity" phenomenon. The main reason that caused the "premature" is that in the early stage of the evolution the diversity of the individual in the crowd decreased rapidly, making the algorithm premature convergence, which may lose some optimum points and search points and step into a local optimum. The fundamental goal of using GA to solve the problem is to find out the global optimal solution in the shortest time. Therefore, at the end of the algorithm, we are often most concerned about the performance of the best individual. From the probability point of view, the degree of affinity between the outstanding individuals of the groups and the global optimal solution is greater than that between the other individuals of the groups and the global optimal solution, and, the individuals who have greater affinity with the excellent individuals should also have a higher degree of fitness. Therefore, whether the outstanding feature information embodied in individuals can be taken full advantage is an important factor to determine the optimal ability of the algorithm.

For the evolution of bees, which has been well known, the queen bee plays a leading role in breeding and evolution of the colony. The drone obtained the mating opportunities with the queen bee through a competition, and the worker bees do not have a substantial role in the reproductive peak of the honey. Traditional and improved GA usually have only one species in the iterative operation, and parallel distributed processing technology has become the focus of research in recent years. Multi-population model, on one hand, can avoid that an individual with a larger fitness rapidly spreads before the global optimal solution is found, then some key information will be lost and results in the premature convergence of the search; on the other hand, it can also improve the quality and enhance the evolutionary speed of the GA. In summary, this paper introduced bee dual population evolutionary mechanisms in the GA, and aims to improve the performance of GA. The whole process of Genetic algorithm of bee dual population evolution can be described as follows:

Step 1 Randomly generated initial population A (0), the number of it is N, $t = 0$, the crossover takes rate p_c, assigned for the mutation probability P_m.

Step 2 Compute the fitness of the individual in the population A (0), the optimal individuals (i.e. the first 0 generation queen) saved to Queen1, the optimal individual of the current generation Queen = Queen1.
Step 3 If meet stop standards, the algorithm putout the results and stop running, otherwise, continue.
Step 4 Set the iteration times $t = t + 1$.
Step 5 Randomly generated population B that contains N individuals, the optimal individuals saved to *Queen*2.
Step 6 Generated r_1, r_2 respectively by the randomly generated way in the range from 0 to 1.
Step 7 Use the ways of random selection, select r_1*N/2, R_2*N/2 individuals from the population A and B, respectively.
Step 8 *Queen*1 is crossover with the r_1*N/2 individuals that has selected at the last step
Step 9 *Queen*2 is crossover with the r_2*N/2 individuals.
Step 10 for the population that choose from the step 8 and step 9, use the proportion choice to select the first N individuals, denoted as C (t).
Step 11 Perform mutation operation on C (t) and get population D (t).
Step 12 Compute the fitness of the individuals in the population D (t), denoted the individual whose fitness is the largest as *Queen-New*.
Step 13 If the fitness of *Queen-New* is larger than both Queen1 and Queen2, then *Queen*1 = *Queen-New*, $D(t)$ is the tth generation population A(t); otherwise, choice the larger one from Queen1 and Queen2 to instated the worst individual of D(t), so can get the population A(t) and the optimal individual of the current generation.
Step 14 Turn to Step 3.

4.2.3 The Combination of Chaos and Genetic Algorithm

In order to combine the chaos and GA, it mainly uses the following two measures:

Generate initial population by using the chaos optimization the issue of GA's convergence is related with the initial population. Because of a considerable part individuals of the randomly generated initial population are away from the optimal solution, limiting the efficiency of the algorithm for solving. Using the ergodicity of chaos for coarse-grained global search, we can usually get better results than the random search so as to improve the quality of the individuals of the initial population and the computation efficiency.

Assigned to x_n which appeared in Eq. (4.1) to m different initial values, and the initial values have a few difference among them and $0 \leq x_n \leq 1 (n = 1, 2, \ldots, m)$, generated m chaotic variables $\{x_n, i, i = 1, 2, \ldots, m\}$ and they all have different track. Then based on Eq. (4.2) to map m chaotic variables into the range of the optimization variables so that it becomes chaotic variables $x^*_{n,\ i}$.

$$x*_{n,i} = a_i + (b_i a_i) * x_{n,i} \tag{4.2}$$

The fixed $n, x_n^* = \left(x_{n,1}^*, x_{n,2}^*, \ldots, x_{n,m}^*\right)$ represents a feasible solution computing the fitness of every feasible solution, and choice N individuals whose fitness are the highest to compose the initial population. In order to ensure the chaotic variables can travel adequately, in the initial population, and during the process of the chaos generated, the chaotic sequence should take sufficient number of iterations (it generally 400–500).

Degenerate the chaotic mutation operator

The purpose of mutation is to continuously generate the individuals that of higher fitness, and lead the evolution of the whole population. GA specializes in generating the population of high average fitness, but lack of good way to produce the optimal individual with a greater fitness. Introducing degradation chaotic mutation operator can simulate the process of biological evolution to chaos; at the same time, due to degradation chaotic mutation operator uses chaotic variables to traverse the solution space, and the search process is based on its own chaotic motion law, it is more easy to find another better solution at the near- optimal solution of the current generation. So GA always has the forward momentum, it can effectively overcome the defect that the speed is significantly slower when search is near the global optimum. Its main process is as follows:

Mapping the population $(y_{n1}, y_{n2}, \ldots, y_{nm})$ of the nth generation whose current solution space is (a, b) to the interval of the chaotic variable [0, 1], and formatting the chaotic variable space vector Y_n^*

$$Y_n^* = (y_{n1}^* y_{n2}^* \ldots y_{nm}^*) \tag{4.3}$$

In the equation, $y_{ni}^* = (y_{ni} - a)/(b - a)\,(i = 1, 2, \ldots, m)$

Degenerated and added the ith chaotic variable to the individual mapping y_n at a probability, producing the map of chaotic mutation individuals on [0, 1].

$$Z_{ni}^* = y_{ni}^* + \rho \mathrm{x}_{n\,i} \tag{4.4}$$

Of which, ρ is the degradation operator,

$$\rho = 1 - \left|\frac{n-1}{n}\right|^k \tag{4.5}$$

In the equation, n is the number of iterations, k is an integer.

Finally, mapping the chaotic mutation individual obtained from Eq. (4.4) to the solution space (a, b) conversely, and complete a mutation operation:

$$\mathrm{Z}_{ni} = a + (ba)\ \mathrm{Z}_{ni}^* \tag{4.6}$$

The Eqs. (4.4) and (4.5) show that degeneration and variance operators simulate the evolution process of the species in the natural world. In the initial stage of evolution, the mutation rate is high and often has more evolutionary test, it leads to population diversity. When the degradation appeared at the late of the evolution,

the role of the mutation operator reduced generally, and the role of the crossover is more prominent, the population became more conservative. Crossover and selection operators can be combined and make local search at the solution space accurately.

4.2.4 Genetic Algorithm of Bee Dual Population Evolution Based on Chaos

Based on the above analysis, we tried an algorithm that combined the improved GA with chaos so as to expect better outcomes at the aspect of the accuracy of the algorithm and timing performance. Here the improved GA is based on bee dual population evolution mechanism. Here is genetic algorithm of bee dual population evolution based on chaos:

Algorithm 2

Step 1 Set the encoding and related parameters. Use Metric real-coded to represent the variables. This representation is naturally intuitive, and with the advantages of saving of time and space overhead and high efficiency. Set the population size, crossover probability, mutation probability, chaos iteration and other parameters.

Step 2 Generate the initial population A with chaos optimize.

Step 3 Compute the fitness, according to the objective function or make appropriate transformation on the objective function, and get the individual's fitness and the queen (best individual in colony).

Step 4 Randomly generate the alien species B (the same population size and A), and in accordance with the fitness to select the queen of the population.

Step 5 Selection operation. Use the roulette wheel selection method, and combine with the best retention strategies. We select drones which are involved in cross from the two populations. The number is determined by the proportion.

Step 6 Crossover operation. According to a certain crossover probability, randomly select individuals from the population which generated with the method of copy as a parent. Use a linear crossover operator to crossover with the two beeswaxes and generate population of the next generation.

Step 7 Mutation. According to some mutation probability, do chaos mutation operation according to Eqs. (4.4) and (4.5), and generate the filial generation.

Step 8 Determine the abeyant condition. If the total number of the evolution is met or after through n iterations the optimal individual is the same, then the algorithm is stop. Otherwise go to step 3 and make the next iteration.

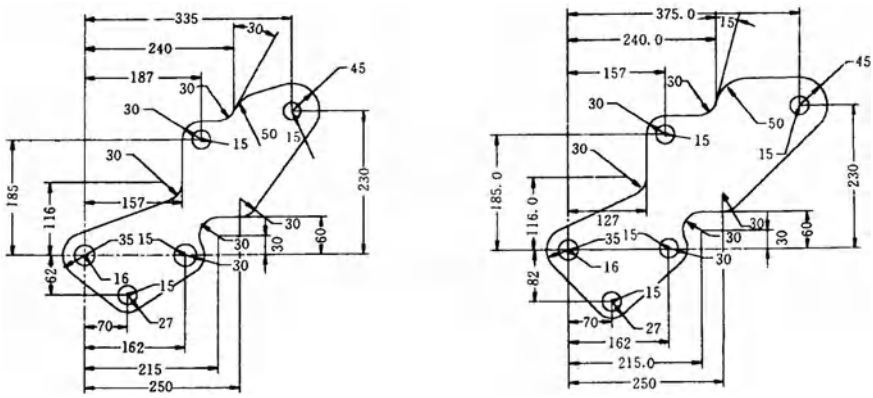


Fig. 4.1 **a** A design instance. **b** Solving result

4.3 The Accomplishment of Bee Dual Population Evolution Based on Chaos in the Geometric Constraint Solving

The constraint problem can be formalized a ← s (*E*, *C*) [8], here E = (e_1, e_2, ……, e_n), it can express geometric elements, such as point, line and circle; $C = (c_1, c_2, \ldots, c_m)$, c_i is the constraint set in these geometric elements. Usually one constraint is represented by an algebraic equation, so the constraint can be expressed as follows:

$$\begin{cases} f_1(x_0, x_1, x_2, \ldots, x_n) = 0 \\ f_m(x_0, x_1, x_2, \ldots, x_n) = 0 \end{cases} \tag{4.7}$$

$X = (x_0, x_1, \ldots, x_n)$, X_i are some parameters, for example, planar point can be expressed as(x_1, x_2). Constraint solving is to get a solution x to satisfy formula (4.7).

$$F(X_j) = \sum_{1}^{m} |f_i| \tag{4.8}$$

Apparently if X_j can satisfy $F(X_j) = 0$, then X_j can satisfy formula (4.7). So the constraint problem can be transformed to an optimization problem and we only need to solve min $\left(F(X_j)\right) < \varepsilon$. ε is a threshold.

4.4 Application Instance

The two graphs in Fig. 4.1 are drafts in engineering design. Figure. 4.1(b) is an auto-produced graph after some sizes of the Fig. 4.1(a) are modified by Bee Dual Population Evolution on Chaos Algorithm. Parameters are changed from 62 to 82,

from 187 to 157, from 335 to 375, from 30 to 15 and from 157 to 127, respectively. From Fig. 4.1 we can realize from the above figures that once a user defines a series of relations, the system will satisfy the constraints by selecting proper state after the parameters are modified by Bee Dual Population Evolution on Chaos Algorithm.

4.5 Conclusion

In this paper we propose a new method using the geometric constraint solving—Bee Dual Population Evolution on Chaos Algorithm. The comparison result shows that this algorithm has a strong capability of searching the best solution. We can get the initial population by the code in GA combined with chaos as to enhance the quality of the population of the individual. Chaos degenerate mutation operators are introduced. And it can solve the problems that GA is easily trapped into local optimum and convergence is slow. Bee Dual Population Evolution on Chaos Algorithm has advantages of traditional GA that it can search in many points and has a small memory. In addition, it can overcome the tradition GA premature and is easily in the local the best and stagnate. The introduction of two populations can expand the capability of algorithm exploring and is helpful for the practical application.

Acknowledgments This work is supported by "the Fundamental Research Funds for the Central Universities" (Project number: N100404002). This work is supported by Opening fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (Chengdu University of Technology)(Project number: SKLGP2011K004).

References

1. Xueyan L, Jianmin Z (2009) Genetic algorithm of bee dual population evolution based on chaos. Comput Appl Softw 26(11):263–271
2. Xueyan L, Yongquan Z (2008) Genetic algorithm based on double bee population evolutionary [J]. Comput Eng Appl 44(11):71–74
3. Wei M, Xuedong H, Bingrong H (2006) Bee evolutionary genetic algorithm [J]. Acta Electronica Sin 34(7):1294–1300
4. Yan W, Zhao F (2008) Adaptive invading genetic algorithm based on chaos search [J]. J Comput Appl 28(1):101–103
5. Dongping T, Hongqin C (2007) An adaptive genetic algorithm combining with chaos searching. Comput Appl Softw 24(12):189–190
6. You Y, Shengw X, Wang SN (2002) Study of chaos genetic algorithms and its application in neural networks [C]//IEEE TENCON. IEEE Press, Beijing, pp 232–235
7. QIRB, QIANF, LISJ et al (2006) Chaos-genetic algorithm formulti-objective optimization[C]// 6thWorld congress on intelligent control and automation. IEEE Press, Dalian pp 1563–1566
8. Shengli L, Min T, Jinxiang D (2003) Geometric constraint satisfaction using genetic simulated annealing algorithm. J Image Graphics 8(8):938–945

Chapter 5
Indirect Adaptive Dynamic Surface Control

Yunjian Wang, Xiaodong Liu and Guodong Wang

Abstract To study the tracking problem for the uncertain nonlinear systems, a scheme, namely an improved indirect adaptive dynamic surface control, is proposed. The program used integral Lyapunov function can avoid the controller singularity, canceled the derivative of virtual input, and induced nonlinearity item which can make better use of nonlinear information to improve the performance of loop transition completely. The method is of characteristic with little calculation for unknown parameters to identify untidily. Finally theory is proved that the closed system is stable.

Keywords Dynamic surface control · Adaptive control · Uncertainty · Nonlinear system

5.1 Introduction

Adaptive controller design to realize function which is using general function to approach unknown functions, such as neural network and fuzzy system, has become focus on the uncertain nonlinear systems [1–3]. For strict-feedback

Y. Wang (✉) · X. Liu · G. Wang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo Henan 454000, People's Republic of China
e-mail: yunjian_wang@163.com

X. Liu
e-mail: liuxiaodong_1@126.com

G. Wang
e-mail: wgd@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_5,

systems, an adaptive backstepping control is proposed to obtain global stability and asymptotic tracking in [4]. An integral adaptive control algorithm is suggested to avoid the controller singularity using a group of integral Lyapunov functions in chapter [5]. And in [6, 7], The Nussbaum gain method used to study adaptive controller design is to stabilize nonlinear systems which is of unknown gain sign. Although backstepping control has the advantages of systematization and structuralization, but also has the problem of control terms explosion.

The dynamic surface control technology can solve the "explore" of backstepping control terms. Based on the research of Swaroop, Song and so on, an adaptive neural network control technology is proposed combining the dynamic surface control technology with adaptive control in chapters [8, 9]. And based on the dynamic surface control technology, a robust controller designed with L2 gain is presented if uncertain terms is of linear upper bound in chapter [10].

Based on chapter [5], an improved adaptive dynamic surface control is suggested in this chapter, to study the tracking problem for a class of uncertain nonlinear systems. A nonlinear term in virtual inputs and control law is adopted to improve the dynamic performances of the close system. And more, theory proved that the close system is stable.

5.2 Description of Uncertain Nonlinear System

The system is expressed as

$$\begin{cases} \dot{x}_1 = \theta_{11} g_1(x_1) x_2 + \theta_{12} f_1(x_1) + d_1(t) \\ \dot{x}_2 = \theta_{21} g_2(x_1, x_2) x_3 + \theta_{22} f_2(x_1, x_2) + d_2(t) \\ \cdots \\ \dot{x}_n = \theta_{\mathrm{n}1} g_n(x) u + \theta_{\mathrm{n}2} f_n(x) + d_n(t) \\ y = x_1 \end{cases} \tag{5.1}$$

where $x = [x_1, x_2, \cdots, x_n]^T$ is state variable, n is system order, u is control value, y is output, $f_i(x_1, \cdots x_i)$ and $g_i(x_1, \cdots x_i)$ are certain nonlinear vector function, $d_i(t)$ is external disturbance. $\theta_{\mathrm{i}1}$ and $\theta_{\mathrm{i}2}$ are unknown row vector of the corresponding dimension.

Note $g_i(x_i) = \theta_{\mathrm{i}1} g_i(x_1, \cdots x_i) \quad f_i(x_i) = \theta_{\mathrm{i}2} f_i(x_1, \cdots, x_i)$.

Assume 1: $g_{\mathrm{ai}} \le g_i(x_1, \cdots x_i) \le g_{\mathrm{bi}}$ denotes that unknown function is bounded, where g_{ai} and g_{bi} are known positive, and $g_i(x_1, \cdots x_i)$ is smooth and continuous.

Assume 2: $f_i(x_1, \cdots x_i)$ is smooth and continuous.

Control target is given enough smooth and continuous scheduled track $r(t)$, designing controller u makes system(1) track scheduled track progressively, namely $\lim_{t \to \infty} |y - r(t)| = 0$.

5.3 Design of Adaptive Dynamic Surface Controller

Step 1. defining track error, $S_1 = y - r$, then $\dot{S}_1 = g_1(x_1)x_2 + f_1(x_1) + d_1 - \dot{r}$. Choosing smooth positive definite function as

$$V_1 = \int_0^{S_1} \sigma d\sigma / g_1(\sigma + r) \tag{5.2}$$

According to integral mean value theorem, Eq. 5.2 can expressed as

$$V_1 = S_1^2 / 2g_1(\varepsilon S_1 + r), \quad 0 \le \varepsilon \le 1 \tag{5.3}$$

According to Assume (1), we can know $g_1(\varepsilon S_1 + r) > 0$ and $V_1 \ge 0$. By deviating Eq. 5.2, it can be obtained that

$$\dot{V}_1 = S_1 \dot{S}_1 / g_1(x_1) + \int_0^{S_1} \sigma \frac{\partial g_1^{-1}(\sigma + r)}{\partial r} \dot{r} d\sigma \tag{5.4}$$

By the equation $\frac{\partial g_1^{-1}(\sigma+r)}{\partial r} = \frac{\partial g_1^{-1}(\sigma+r)}{\partial \sigma}$, $\int_0^{S_1} \sigma \frac{\partial g_1^{-1}(\sigma+r)}{\partial r} \dot{r} d\sigma = \frac{\dot{r} S_1}{g_1(x_1)} - \dot{r} S_1 \int_0^1 \frac{1}{g_1(\varepsilon S_1 + r)} d\varepsilon$ is obtained. And then Eq. 5.4 can be expressed as $\dot{V}_1 = S_1(\dot{S}_1 + \dot{r}) / g_1(x_1) - \dot{r} S_1 \int_0^1 d\varepsilon / g_1(\varepsilon S_1 + r) = S_1(x_2 + h_1)$, where $h_1 = f_1(x_1)/g_1(x_1) + d_1/g_1(x_1) - \dot{r} \int_0^1 d\varepsilon / g_1(\varepsilon S_1 + r) + d_1/g_1(x_1) - \dot{r} \int_0^1 d\varepsilon / g_1(\varepsilon S_1 + r)$. Choose virtual input as

$$\bar{x}_2 = -\lambda_1 S_1 - \rho_1 - \hat{h}_1 \tag{5.5}$$

where controller gain is $\lambda_1 > 0$, also $\hat{h}_1$ is estimate of h_1, and error estimated is noted $\tilde{h}_1 = h_1 - \hat{h}_1$, and ρ_1 is odd function, such as choosing $\rho_1 = \frac{S_1}{1 + e^{-x_1^2}}$. The design of the following part is same.

$x_{2\mathrm{d}}$ is the target track of x_2 produced by $\bar{x}_2$, through the low-pass filter as

$$\tau_2 \dot{x}_{2\mathrm{d}} + x_{2\mathrm{d}} = \bar{x}_2, x_{2\mathrm{d}}(0) = \bar{x}_2(0) \tag{5.6}$$

Defining filter error $\xi_2 = x_{2\mathrm{d}} - \bar{x}_2$, then $\dot{x}_{2\mathrm{d}} = -\tau_2^{-1} \xi_2$ is obtained. Finally, $\dot{V}_1$ can be expressed as

$$\dot{V}_1 = -\lambda_1 S_1^2 + S_1(x_2 - x_{2\mathrm{d}}) + S_1 \xi_2 - \rho_1 S_1 + \tilde{h}_1 S_1 \tag{5.7}$$

Step $i : 2 \le i \le n - 1$ defining error $S_i = x_i - x_{\mathrm{id}}$, then $\dot{S}_i = g_i(x_i)x_{\mathrm{i}+1} + f_i(x_i) + d_i - \dot{x}_{\mathrm{id}}$, Choosing smooth positive definite function as

$$V_i = \int_0^{S_i} \sigma d\sigma / g_i(\sigma + x_{id}) \tag{5.8}$$

According to integral mean value theorem and Assume 1, we have $V_i \geq 0$. As same as step 1, Eq. 5.9 can be obtained.

$$\begin{aligned} \dot{V}_i = S_i \Bigg[& x_{i+1} + g_i^{-1}(x_i)(f_i(x_i) + d_i) - \tau_i^{-1}\xi_i \int_0^1 g_i^{-1}(\varepsilon S_i + x_{id}) d\varepsilon \\ & + S_i \int_0^1 \left(\sum_{j=1}^{i-1} \frac{\partial g_i^{-1}(\varepsilon S_i + x_{id})}{2\partial x_j} \dot{x}_j \right) d\varepsilon \Bigg] = S_i(x_{i+1} + h_i) \end{aligned} \tag{5.9}$$

where $h_i = g_i^{-1}(x_i)(f_i(x_i) + d_i) - \tau_i^{-1}\xi_i \int_0^1 g_i^{-1}(\varepsilon S_i + x_{id}) d\varepsilon - \tau_i^{-1}\xi_i \int_0^1 g_i^{-1}(\varepsilon S_i + x_{id}) d\varepsilon$. Choose virtual input as

$$\bar{x}_{i+1} = -\lambda_i S_i - \rho_i - \hat{h}_i \tag{5.10}$$

$x_{(i+1)d}$ is the target track of x_{i+1} produced by $\bar{x}_{i+1}$, through the low-pass filter as

$$\tau_{i+1}\dot{x}_{(i+1)d} + x_{(i+1)d} = \bar{x}_{i+1}, \quad x_{(i+1)d}(0) = \bar{x}_{i+1}(0) \tag{5.11}$$

Define filter error $\xi_{i+1} = x_{(i+1)d} - \bar{x}_{i+1}$, then $\dot{x}_{(i+1)d} = -\tau_{i+1}^{-1}\xi_{i+1}$ is obtained. Finally, $\dot{V}_i$ can be expressed as

$$\dot{V}_i = S_i(x_{i+1} - x_{(i+1)d}) + S_i\xi_{i+1} - \lambda_i S_i^2 - \rho_i S_i + \tilde{h}_i S_i \tag{5.12}$$

Step n : define error $S_n = x_n - x_{nd}$, then $\dot{S}_n = g_n(x)u + f_n(x) + d_n - \dot{x}_{nd}$. Choosing smooth positive definite function as

$$V_n = \int_0^{S_i} \sigma d\sigma / g_n(\sigma + x_{nd}) \tag{5.13}$$

According to integral mean value theorem and Assume 1, we have $V_i \geq 0$. The same as above steps,

$$\begin{aligned} \dot{V}_n = S_n \Bigg(& u + g_n^{-1}(x)(f_n(x) + d_n) - \tau_n^{-1}\xi_n \int_0^1 g_n^{-1}(\varepsilon S_n + x_{nd}) d\varepsilon \\ & + S_n \int_0^1 \left(\sum_{j=1}^{n-1} \frac{\partial g_n^{-1}(\varepsilon S_n + x_{nd})}{2\partial x_j} \dot{x}_j \right) d\varepsilon \Bigg) = S_n(u + h_n) \end{aligned} \tag{5.14}$$

Choosing control law expression as

$$u = -\lambda_n S_n - \rho_n - \hat{h}_n \tag{5.15}$$

Substituting Eq. 5.15 into Eq. 5.14, then

$$\dot{V}_n = -\lambda_n S_n^2 - \rho_n S_n + \tilde{h}_n S_n \tag{5.16}$$

Making $h = [h_1, \cdots, h_n]^T$ and using a BP neural network approaches the unknown function h. Describing BP neural network as

$$h = \hat{W}^T \phi\left(\hat{V}_g^T Z\right) - \tilde{W}^T\left(\hat{\phi} - \hat{\phi}' \hat{V}_g^T Z\right) - \hat{W}^T \hat{\phi}' \tilde{V}_g^T Z \tag{5.17}$$

where $Z = [x^T, S^T, 1]^T$, $S = [S_1, \cdots, S_n]^T$. $\hat{W} = [\hat{w}_1, \cdots \hat{w}_n]$ and $\hat{V}_g = [\hat{v}_{g1}, \cdots, \hat{v}_{gl}]$ denote the value estimated of $W^*$and V_g^*, respectively. And their error estimated was defined $\tilde{W} = \hat{W} - W^*$ and $\tilde{V}_g = \hat{V}_g - V_g^*$, also l is number of hidden layer nodes, and $\phi = [\phi_1(\hat{v}_{g1}^T Z), \cdots, \phi_l(\hat{v}_{gl}^T Z), 1]^T$, Activation function is $\phi_j(\alpha) = \frac{1}{1+e^{-\beta\alpha}}$, $j = 1, \cdots, l$, constant $\beta > 0$. $\hat{\phi}' = \text{diag}\{\hat{\phi}_1', \cdots, \hat{\phi}_l', 0\}$, $\hat{\phi}_j' = \frac{d\phi_j(z)}{dz}\Big|_{z=\hat{v}_{gj}^T Z}$. W^* and V_g^* are ideal weighting matrix.

Making $\hat{h} = \hat{W}^T \phi(\hat{V}_g^T Z)$, which is the value estimated of h, then

$$\tilde{h} = -\tilde{W}^T\left(\hat{\phi} - \hat{\phi}' \hat{V}_g^T Z\right) - \hat{W}^T \hat{\phi}' \tilde{V}_g^T Z \tag{5.18}$$

Choosing Lyapunov function

$$V = \sum_{i=1}^{n} V_i + \frac{1}{2}\gamma_w^{-1}\text{trace}(\tilde{W}^T \tilde{W}) + \frac{1}{2}\gamma_v^{-1}\text{trace}\left(\tilde{V}_g^T \tilde{V}_g\right) \tag{5.19}$$

By Deviating Eq. 5.19, we have

$$\begin{aligned} \dot{V} \leq & -(\lambda_1 - 0.5)S_1^2 - \sum_{i=2}^{n-1}(\lambda_i - 1)S_i^2 - (\lambda_n - 0.5)S_n^2 + \sum_{i=2}^{n}\xi_i^2 - S^T \tilde{W}^T\left(\hat{\phi} - \hat{\phi}' \hat{V}_g^T Z\right) \\ & - S^T \hat{W}^T \hat{\phi}' \tilde{V}_g^T Z - \sum_{i=1}^{n}\rho_i S_i + \gamma_w^{-1}\text{trace}(\dot{\tilde{W}}^T \tilde{W}) + \gamma_v^{-1}\text{trace}\left(\dot{\tilde{V}}_g^T \tilde{V}_g\right) \end{aligned} \tag{5.20}$$

Noting that $\gamma_w^{-1}\text{trace}(\dot{\tilde{W}}^T \tilde{W}) = S^T \tilde{W}^T(\hat{\phi} - \hat{\phi}' \hat{V}_g^T Z), \gamma_v^{-1}\text{trace}(\dot{\tilde{V}}_g^T \tilde{V}_g) = S^T \hat{W}^T \hat{\phi}' \tilde{V}_g^T Z$, then the regulation of neural network is derived as

$$\begin{cases} \dot{\hat{w}}_i = -\gamma_w S_i\left(\hat{\phi} - \hat{\phi}' \hat{V}_g^T Z\right) \\ \dot{\hat{v}}_{gj} = -\gamma_v \varphi_j Z \end{cases} \tag{5.21}$$

where $i = 1, \cdots, n$, $j = 1, \cdots, 2n+1$, $\varphi = S^T \hat{W}^T \hat{\phi}'$, φ_j is the j element of φ, $\gamma_w > 0$ and $\gamma_v > 0$ are learning factors.

Because ξ_i is filter error of the low-pass filter, if the input of the low-pass filter is bounded, then ξ_i is bounded. And ρ_i is odd function of S_i, so if $S_i \neq 0$, then $\rho_i S_i^{-1} > 0$, or if $S_i = 0$, then $\rho_i = 0$. Substituting Eq. 5.21 into Eq. 5.20, we obtain $S_i \neq 0$, then

$$\begin{aligned} \dot{V} \leq & -\left(\lambda_1 + \rho_1 S_1^{-1} - 0.5\right) S_1^2 - \sum_{i=2}^{n-1} \left(\lambda_i + \rho_i S_i^{-1} - 1\right) S_i^2 \\ & - \left(\lambda_n + \rho_n S_n^{-1} - 0.5\right) S_n^2 + \sum_{i=2}^{n} \xi_i^2 \end{aligned} \tag{5.22}$$

if $S_i = 0$, then

$$\dot{V} \leq -(\lambda_1 - 0.5) S_1^2 - \sum_{i=2}^{n-1} (\lambda_i - 1) S_i^2 - (\lambda_n - 0.5) S_n^2 + \sum_{i=2}^{n} \xi_i^2 \tag{5.23}$$

Choosing $\lambda_i > 1$, Eqs. 5.22 and 5.23 denote that the closed-loop system is stable.

Based on the above analysis, the conclusion can be obtained as: If the system(1) satisfies the assumes, then a dynamic surface controller which consists of Eq. 5.15, filter expressed by Eq. 5.11 and adaptive regulation Eq. 5.20, exists to ensure the close system tracking the target $r(t)$ gradually.

5.4 Conclusion

The tracking problem for uncertain nonlinear systems is investigated in this chapter, and an improved indirect adaptive dynamic surface control is proposed, which can avoid the controller singularity. Then the nonlinearity item ρ_i in the virtual input is adopted to improve the performance of the closed system. Finally theory is proved that the closed system is stable.

References

1. Huang JT (2009) Hybrid-based adaptive NN backstepping control of strict-feedback systems. Automatica 45:1497–1503
2. Chen WS, Zhang ZQ (2010) Globally stable adaptive backstepping fuzzy control for output-feedback systems with unknown high-frequency gain sign. Fuzzy Sets Syst 161:821–836
3. Wu TZ, Juang YT (2008) Adaptive fuzzy sliding-mode control of uncertain nonlinear systems. ISA Trans 30:279–285
4. Kanellakipoulos I, Kokotovic PV, Morse AS (1991) Systematic design of adaptive controllers for feedback linearizable systems. IEEE Trans Autom Control 36:1241–1253

5. Zhang T, Ge SS, Hang CC (2000) Adaptive neural network control for strict-feedback nonlinear systems using backstepping design. Automatica 36:1835–1846
6. Ge SS, Yang CG, Lee TH (2008) Adaptive robust control of a class of nonlinear strict-feedback discrete-time systems with unknown control directions. Syst Control Lett 57: 888–895
7. Cheng gang (2006) Output feedback control of perturbed nonlinear systems with unknown high-frequency gains. Control Theor App 23:779–782
8. Hongchun L, Tianping Z, Yan S (2007) Indirect adaptive neural network block control using dynamic surface control. Electr Mach Control 11:275–281
9. Wang D, Huang J (2005) Neural network-based adaptive dynamic surface control for a class of uncertain nonlinear systems in strict-feedback form. IEEE Trans Neural Networks 16: 195–202
10. Song B (2007) Robust stabilization of decentralized dynamic surface control for a class of interconnected nonlinear systems. Int J Control, Autom Syst 5:138–146

Chapter 6
An EPON Dynamic Bandwidth Allocation Algorithm Based on the Multiple Traffic Prediction and Error Correction

Ziyi Fu and Juanjuan Bai

Abstract According to the upstream TDM in the system of Ethernet passive optical network (EPON), a dynamic bandwidth allocation algorithm (DBA) based on the mechanism of the multiple traffic prediction-correction is presented in this chapter. To improve the real-time performance of the bandwidth allocation, this algorithm forecasts the traffic of varying priority services, then corrects forecast error according to its Gaussian distribution characteristics, which will make traffic prediction closer to the real traffic. The simulation results show that proposed algorithm is better than the existing DBA algorithm. Not only can it meet the delay requirement of high priority services, but also control the delay abnormity of low priority services. In addition, with rectification scheme it is obvious in improving the bandwidth utilization.

Keywords Dynamic bandwidth allocation · The multiple traffic prediction · Error correction · Gaussian distribution

6.1 Introduction

With the development of network technology, the popularity of high-bandwidth real-time multimedia applications like video-on-demand, video telephony, distance learning, and the concept of triple play put forward, the end-user for access

Z. Fu · J. Bai (✉)
School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: bjjbaijuanjuan@126.com

Z. Fu
e-mail: fuzy@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_6,

network bandwidth and the requirements of service quality is daily on the increase. Ethernet passive optical network (EPON) based on the advantages of its bandwidth-rich, low cost, easy to expand and maintain, as the best carrier of IP, has become the best way to solve the access network "last mile".

Dynamic bandwidth allocation (DBA) algorithm is the key technology for EPON, and directly related to the bandwidth utilization in upstream channel and quality of information. An enhanced bandwidth allocation algorithm (DBA2) [1] has improved the traffic delay, provided preallocated bandwidth to the incoming EF traffic during the waiting time, and perfected the scheduling mode that the optical line terminal (OLT) can assign timely enough bandwidth for light-load of optical network units (ONUs), On the contrary, OLT cannot assign bandwidth for overloaded ONUs until it receives all REPORT messages from each ONU sending ones. A distributed bandwidth allocation algorithm [2] predict the next bandwidth allocation according to the total of all queues for each ONU, of which the disadvantage is that even if there is a lot of low priority traffic of certain ONU, without high priority traffic or little, they would still get excessive bandwidth allocation from OLT, which cannot reflect the real-time traffic and traffic class. To solve these problem, this chapter introduces a new dynamic bandwidth allocation algorithm based on Multiple Services-prediction correction-allocation scheme, also called MPA-DBA.

6.2 MPA-DBA Algorithm

6.2.1 Design Principles of MPA-DBA

Based on point to multipoint transmission structure of EPON, the working group IEEE802.3 ah makes the MPCP agreement located on the MAC control layer [3], relying on the GATE and REPORT Ethernet control frame to assign bandwidth for each ONU. OLT uses GATE frame assign transmission Windows to ONU, and ONU uses REPORT frame to report current condition for OLT. As shown in Fig. 6.1.

According to the GATE/REPORT mechanism, the wait timeslot of each ONU is defined as the time from sending REPORT message to receive GATE message, namely T_{wait}. As shown in Fig. 6.1, take ONUi for example, the service interval of ONUi, say m, ranges from time t1 to time t3, namely T_{cycle}. Time t1 to time t2 is to send data window for ONUi in service interval m, while time t2 to time t3 is waiting time for ONUi in service interval m, during which ONUi is idle, and more data frames from the local users are enqueued in the buffer [4]. At the moment, even if the uplink is idle, the arrived data during the time cannot be transmitted until at least one more service interval. The reason is what the REPORT message sent at time t2 is the data that is waiting for being sent in the current buffer of ONUi, which is the basis of bandwidth allocation in the next cycle. Therefore, the key problems this chapter want to solve is that when OLT assigns bandwidth for

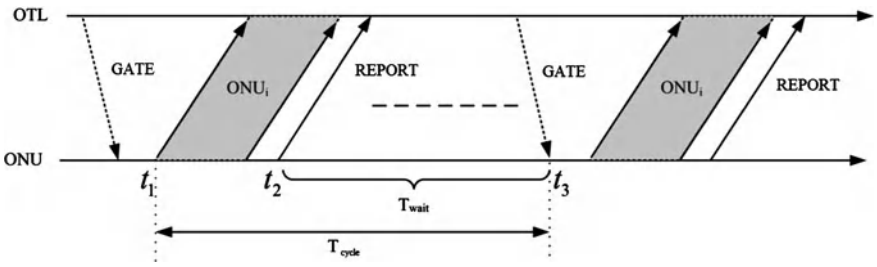


Fig. 6.1 The REPORT/GATE mechanism

different traffic classes, OLT needs to make a prediction about the arrived data in the wait timeslot first and correct the prediction error, then OLT will assign matching bandwidth for different traffic class, to improve real-time of the data sent and the entire bandwidth utilization.

6.2.2 Realization of MPA-DBA

In this section we describe the MPA-DBA scheme. It is well known that the goal of bandwidth allocation is to divide timeslot between ONU and OLT effectively and fairly, and achieve the QoS of different traffic. Residing in customer premises, an ONU must be capable of supporting broadband data, voice, and video services to the local users. Categorizing the traffic of an ONU into different classes is a practical approach for provisioning differentiated services. We borrow the traffic classes from the literature [5], in which the high priority class is the expedited forwarding (EF) which is very sensitive to delay and needs enough bandwidth to ensure highest priority, the medium priority class is assured forwarding (AF) which is not sensitive to delay but has to ensure that bandwidth, and the low priority class is best effort (BE) without delay requirement and bandwidth requirements. Different classes of data frame are enqueued into each of the corresponding priority buffer echelon. All queues share the same buffer. The multiple queue status at an ONU is reported to OLT by means of the REPORT message.

6.2.2.1 Bandwidth Prediction at an ONU

According to the GATE/REPORT mechanism, an ONU sends the bandwidth requirement to OLT, but it has no consideration of the new increased bandwidth requirements of different traffic classes during the wait timeslot from ONUi sending REPORT message to receive GATE message.

To make a certain prediction for the new increased traffic during the wait timeslot, in MPA-DBA, it predicts the frames arrived during the waiting timeslot in terms of predicted bandwidth as

$$B_{i,c}^{p}(m-1) = T_{i,c}(m) - T_{i,c}(m-1),\ C \in (\text{EF, AF, BE}) \tag{6.1}$$

where $B_{i,c}^{p}$ is the prediction of new increased traffic class C for ONUi during the wait timeslot, $T_{i,c}$ is the requested bandwidth of traffic class C by ONUi in the service interval m, C is the traffic class sets of EF, AF and BE.

6.2.2.2 Amendment Mechanism of Prediction Error

It is inevitable to have the excessive bandwidth, if we assign bandwidth based on the predicted bandwidth $B_{i,c}^{p}$. In order to make more excess bandwidth released in the time of bandwidth allocation, amendment mechanism of prediction error is introduced by MPA-DBA. In this chapter we analyze the network performance in terms of the success probability of bandwidth prediction. The prediction error plays a key factor on network performance. We call the bandwidth prediction successful if $e_{i,c} = B_{i,c}^{w} - B_{i,c}^{p} \le 0$, where $B_{i,c}^{w}(m)$ is the real arrived data frame of the C class traffic for ONUi in service interval m. Therefore the success probability of bandwidth prediction is $P_{i,c}^{s} = \{e_{i,c} \le 0\}$. The LMS-based adaptive prediction error is Gaussian [6]. Assume the prediction error has mean $m_{i,c}$, and variance $\sigma_{i,c}^{2}$, i.e.,$e_{i,c} \sim N(m_{i,c}, \sigma_{i,c}^{2})$, the success probability of bandwidth prediction is

$$\begin{aligned} P_{i,c}^{s} = \{e_{i,c} \le 0\} &= \frac{1}{\sqrt{2\pi}\sigma_{i,c}} \int_{-\infty}^{0} e^{-(x-m_{i,c})^2/2\sigma_{i,c}^2} dx = 1 - Q\left(-\frac{m_{i,c}}{\sigma_{i,c}}\right) \\ &= Q\left(\frac{m_{i,c}}{\sigma_{i,c}}\right),\ C \in (\text{EF, AF, BE}) \end{aligned} \tag{6.2}$$

The probability that the prediction failure is

$$P_{i,c}^{f} = 1 - P_{i,c}^{s},\ C \in (\text{EF, AF, BE}) \tag{6.3}$$

The distribution of traffic class C (EF, AF, BE) can be seen by the prediction error. When $e_{i,c} < 0$, it explains that the assigned bandwidth by OLT is enough for the queued frames, and the prediction is deemed a success. When $e_{i,c} = 0$, it is the ideal situation of bandwidth allocation that the actual arrived frames are equal to the prediction result without correction. Therefore if the smaller $e_{i,c}$, the more the success probability of prediction $P_{i,c}^{s} = \{e_{i,c} \le 0\}$ is close to 1. That is, predicted result is closer to actual arrived frames. However, when $e_{i,c} < 0$, it means OLT assign excess frames to ONU, namely prediction fail. So, at the moment, MPA-DBA will make $B_{i,c}^{p}$ corrected in the next timeslot and the amended value is defined as $\Delta_{i,c}(m)$, as follows:

$$\begin{aligned} B_{i,c}^{p}(m) &= B_{i,c}^{p}(m) - \Delta_{i,c}(m),\ C \in (\text{EF, AF, BE}) \\ \Delta_{i,c}(m) &= \frac{B_{i,c}^{p}(m)}{B_{i,c}^{p}(m-1)} P_{i,c}^{f},\ C \in (\text{EF, AF, BE}) \end{aligned} \tag{6.4}$$

6.2.2.3 The Process of Bandwidth Allocation for an ONU

Bandwidth allocation is defined as OLT calculate bandwidth for traffic class C (EF, AF, BE), when it receives the REPORT message by ONUi sanding. To avoid the delay data of high priority traffic with sensitive delay, OLT will assign prior bandwidth for EF traffic. In order to ensure the bandwidth requirement of EF traffic, we reserve a basal bandwidth for EF traffic which is defined as B_b. Once there is available bandwidth, we assign orderly bandwidth to EF traffic and BE traffic according to their priority classes.

After receiving the REPORT message, OLT begin bandwidth allocation immediately. Assume it is defined as $B_i^{\text{Total}}(m)$ that the OLT assign to the total bandwidth of ONUi, namely

$$B_i^{\text{Total}}(m) = \sum B_{i,EF}(m) = B_{i,\text{EF}}(m) + B_{i,\text{AF}}(m) + B_{i,\text{BE}}(m), \quad C \in (\text{EF}, \text{AF}, \text{BE}) \tag{6.5}$$

So, different traffic class will obtain the corresponding bandwidth in terms of the priority classes, Specific allocation process is as follows:

$$\begin{aligned}
B_{i,\text{EF}}(m) &= \min\left\{B_i^{\text{Total}}(m), \min\left\{T_{i,\text{EF}}(m-1) + B_{i,\text{EF}}^P(m-1) - \Delta_{i,\text{EF}}(m), B_b\right\}\right\} \\
B_{i,\text{AF}}(m) &= \min\left\{B_i^{\text{Total}}(m) - B_{i,\text{EF}}(m), T_{i,\text{AF}}(m-1) + B_{i,\text{AF}}^P(m-1) - \Delta_{i,\text{AF}}(m)\right\} \\
B_{i,\text{BE}}(m) &= \min\left\{B_i^{\text{Total}}(m) - B_{i,\text{EF}}(m) - B_{i,\text{AF}}(m), T_{i,\text{BE}}(m-1)\right. \\
&\qquad \left. + \mathrm{B}_{i,\text{BE}}^P(m-1) - \Delta_{i,\text{BE}}(m)\right\}
\end{aligned} \tag{6.6}$$

The process continues until all ONUs have been assigned, and the MPA-DBA will arrange orderly the upload time sequence and report the time of each ONU in the unstable degree list.

6.3 Simulation Results and Analysis

This chapter analyzed the performance of the DBA2 and MPA-DBA in terms of simulation results in the system of OPNET. The system model is set up with the OPNET simulator with one OLT and 16 ONUs. The downstream and upstream channels are both 1 GB/s. The maximum distance from an ONU to the OLT is assumed to be 20 km. Each ONU has a queue buffer of which the length is 5 Mbits. The maximum rate of subscribers is 100 Mb/s and the maximum period is 2 ms. For the traffic model considered here, an extensive study shows that most network traffic can be characterized by self-similarity. Then the traffic profile is as follows: 20% of the total generated traffic is considered for EF traffic, and the remaining 80% is equally distributed between AF traffic and BE traffic. The traffic load of network, defined as the ratio of the average arrival to the traffic rate, ranges from 0.1 to 0.9.

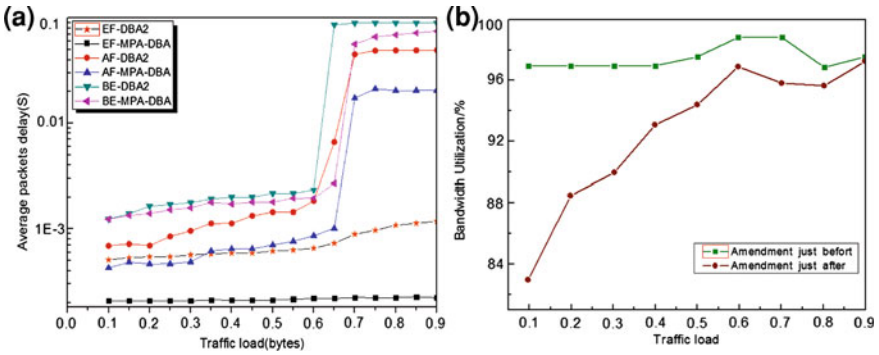


Fig. 6.2 Simulation results. **a** Before removing of strong signals; **b** after removing of strong signals

Figure 6.2a illustrates the relationship between the average packets delay and the network traffic load, which compared the average packets delay of *C* (EF, AF, BE) class traffic of DBA2 to MPA-DBA. Taking EF traffic as example, when traffic load is approximately 0.8, the average packets delay is close to 1 ms in DBA2, while the average packets delay is not more than 0.1 ms with the load ranging from 0.1 to 0.9 in MPA-DBA. Thereby, it is evident that the average packets delay of MPA-DBA is lower compared with DBA2.

Comparing the bandwidth utilization between original prediction error and amendatory prediction error in MPA-DBA is as shown in Fig. 6.2b. Obviously, introducing the amendment scheme for prediction error, it is more reasonable to assign bandwidth from OLT to ONU. The less the idle bandwidth, the more the bandwidth utilization improves.

6.4 Conclusion

To further promote the EPON system, the question we have to resolve is how to use efficient dynamic bandwidth allocation algorithm to guarantee QoS of the network. This chapter designs a new DBA algorithm on the basis of priority classification to different traffic class, predicts various types of traffic, corrects forecast error, and assigns transmission bandwidth according to their priority. This algorithm reduces the waiting time in the process of bandwidth allocation, and improves the delay of the different priority services effectively to ensure the QoS of entire EPON network and the scientific of bandwidth allocation; thereby bandwidth utilization is further improved.

Acknowledgments This chapter is supported by key scientific and technological project in Hennan Province: 112102210004. The author thanks the Editor for helpful comments that greatly improved the chapter.

References

1. Zhang G, WU X (2008) Research of in ethernet passive optical networks. Comput Eng 34(10):123–125
2. Chen G, Fan D, Yang J (2006) A kind of optimized dynamic bandwidth allocation algorithm for distributed optical network. J Lanzhou Jiaotong Univ 25(6):23–26
3. LANPMAN Standards Committee (2004) IEEE Standard 802.3 ah 2004[S]
4. Wang H, Li W, Zhang C, Huang H (2007) Dynamic bandwidth algorithm applying traffic prediction. Opt Electr Eng 34(5):117–120
5. Fu Z, Lv S (2010) A EPON dynamic bandwidth assignment algorithm share-based the multiple traffic prediction. Third international symposium on electronic commerce and security workshops, 29–31 July 2010, Guangzhou, China, pp 112–115
6. Hwang I-S (2008) A novel early DBA mechanism with prediction-based fair excessive bandwidth allocation scheme in EPON. Comput Commun 31:1814–1823

Chapter 7
Synchronization of Chaotic Systems with Double Strange Attractors via Passivity Approach

Yunzhong Song and Yanyan Li

Abstract This paper is devoted to address passivity-based approach synchronization between master and slave Newton–Leipnik chaotic systems, each of which has double co-existing strange attractors. Unlike the already existing results, Lyapunov stability and Linear Matrix Inequality (LMI) approach is firstly employed here. Synchronization was realized in virtue of control action added in the slave system, which was designed in an easy going way under auspices of the solution of the LMI, which in essence is a convex optimization. On account of which the identity of slaved system model and master system model was facilitated and guaranteed. A numerical example is given to demonstrate the effect of the proposed synchronization scheme.

Keywords Chaos synchronization · Newton–Leipnik chaotic system · Passivity approach

7.1 Introduction

Since the pioneering work of Pecora and Carrol [1], chaos synchronization has drawn much attention both in theoretical studies and in potential engineering areas, see literature [2, 3] and references there in. Recently, a new chaotic system the Newton–Leipnik system with double strange attractors has aroused interest among

Y. Song (✉) · Y. Li
Complex Networks Laboratory, School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China
e-mail: yunzhong.song@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_7,

researchers. The Newton–Leipnik system was firstly obtained by modifying Euler's rigid body equations with the addition of a linear feedback [4]. Since then, many scholars had touched upon this exciting system. For example, dynamics and bifurcation analysis were considered in [5, 6]; fine structure of periodic orbits in the Newton–Leipnik system was investigated in [7]; stabilizing a desired motion within one strange attractor as well as taking the system dynamics from one attractor to another one was executed in [8–11].

Here, synchronization of Newton–Leipnik systems is addressed. Lyapunov stability and linear matrix inequality (LMI) approach is firstly employed here. Synchronization was realized in virtue of control action added in the slave system, which was designed in an easy going way under auspices of the solution of the LMI, which in essence is a convex optimization.

The rest of this paper is arranged as follows. First, basic concept of passivity is given in Sect. 7.2; then, problem formation is available after in Sect. 7.3; controlling law and the relevant theoretical proof is discussed in Sect. 7.4 afterwards; and then numerical simulation is provided on Sect. 7.5 to show the advantages and effectiveness of the proposed strategy; finally, Sect. 7.6 serves as conclusion.

7.2 Basic Concept of Passivity

In case of brevity, the basic concept of passivity can be referred to [10] and the corresponding references there in.

7.3 Problem Formulation

We will use the passivity technique to control Newton–Leipnik equations. The controller proposed in this study is based on LMI. Using the passive method, it is very easy to prove the stability of the closed-loop system. The dynamical model of Newton–Leipnik system is given by

$$\begin{aligned} \dot{x}_1 &= -\gamma x_1 + x_2 + 10x_2x_3, \\ \dot{x}_2 &= -x_1 - 0.4x_2 + 5x_1x_3, \\ \dot{x}_3 &= \sigma x_3 - 5x_1x_2. \end{aligned} \tag{7.1}$$

where x_1, x_2 and x_3 represent angular velocities about principal axes, which are fixed in the body, γ and σ are positive parameters that are related to the eigenvalues of inertia moments of the body. The Newton–Leipnik system model shows a double strange attractor for the nominal parameters $\gamma = 0.4$ and $\sigma = 0.175$ when control action is not applied. For the initial states $x_0^{(1)} = (0.349, 0, -0.160)^{\mathrm{T}}$ and $x_0^{(2)} = (0.349, 0, -0.180)^{\mathrm{T}}$ two strange attractors are obtained, see Fig. 7.1.

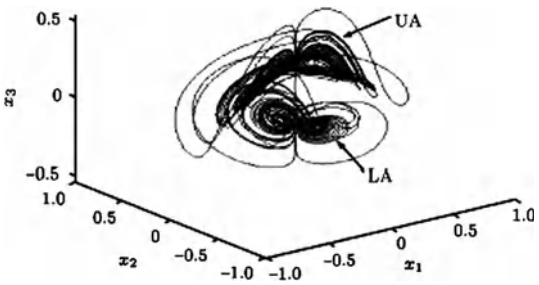


Fig. 7.1 Phase locus of Newton–Leipnik system

As discussed in the previous work on the Newton–Leipnik system [8–11], the attractor A1 in the upper space is referred to as the upper attractor (UA) and the attractor A2 in the lower space is referred to as the lower attractor (LA), respectively. The equilibrium points of the system (7.4) with the nominal parameters $\gamma = 0.4$ and $\sigma = 0.175$ are

$$\begin{aligned}
x_e^{(1)} &= (0, 0, 0)^{\mathrm{T}}, \\
x_e^{(2)} &= (0.2390, 0.0308, 0.2103)^{\mathrm{T}}, \\
x_e^{(3)} &= (-0.2390, -0.0308, 0.2103)^{\mathrm{T}}, \\
x_e^{(4)} &= (0.0315, -0.1224, -0.1103)^{\mathrm{T}}, \\
x_e^{(5)} &= (-0.0315, 0.1224, -0.1103)^{\mathrm{T}}.
\end{aligned}$$

All these equilibrium points are unstable. Note that the points $x_e^{(2)}$ and $x_e^{(3)}$ lie in the eye-centres of UA, while the points $x_e^{(4)}$ and $x_e^{(5)}$ are situated in the eye-centres of LA.

The synchronization problem of system (7.1) is considered using the drive-response configuration. The system (7.1) is considered as the drive system. According to the drive-response concept, the controlled response (or slave) system is given by

$$\begin{aligned}
\dot{\bar{x}}_1 &= -\gamma \bar{x}_1 + \bar{x}_2 + 10\bar{x}_2\bar{x}_3 + u_1(t), \\
\dot{\bar{x}}_2 &= -\bar{x}_1 - 0.4\bar{x}_2 + 5\bar{x}_1\bar{x}_3 + u_2(t), \\
\dot{\bar{x}}_3 &= \sigma \bar{x}_3 - 5\bar{x}_1\bar{x}_2 + u_3(t),
\end{aligned} \tag{7.2}$$

where $u_1(t)$, $u_2(t)$ and $u_3(t)$ are the nonlinear controllers.

Define the synchronization error as

$$\begin{aligned}
e_1(t) &= \bar{x}_1 - x_1, \\
e_2(t) &= \bar{x}_2 - x_2, \\
e_3(t) &= \bar{x}_3 - x_3,
\end{aligned} \tag{7.3}$$

Then we obtain the following synchronization error system

$$\begin{aligned}\dot{e}_1(t) &= -\gamma e_1(t) + e_2(t) + 10\bar{x}_2\bar{x}_3 - 10x_2x_3 + u_1(t),\\ \dot{e}_2(t) &= -e_1(t) - 0.4e_2(t) + 5\bar{x}_1\bar{x}_3 - 5x_1x_3 + u_2(t),\\ \dot{e}_3(t) &= \sigma e_3(t) - 5\bar{x}_1\bar{x}_2 + 5x_1x_2 + u_3(t),\end{aligned} \tag{7.4}$$

which can be recast as

$$\dot{e}(t) = Ae(t) + f(t) + u(t), \tag{7.5}$$

where $e(t), u(t)$ and $f(t)$ are defined by

$$e(t) = \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix},$$

$$u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{bmatrix},$$

$$A = \begin{bmatrix} -\gamma & 1 & 0 \\ -1 & -0.4 & 0 \\ 0 & 0 & \sigma \end{bmatrix},$$

$$f(t) = \begin{bmatrix} 10\bar{x}_2\bar{x}_3 - 10x_2x_3 \\ 5\bar{x}_1\bar{x}_3 - 5x_1x_3 \\ -5\bar{x}_1\bar{x}_2 + 5x_1x_2 \end{bmatrix}.$$

7.4 Controller Design

The LMI problem for achieving the passivity-based synchronization is presented in the following theorem.

Theorem 1 *For a given $Q = Q^T > 0$, if there exist $X = X^T > 0$ and Y such that*

$$\begin{bmatrix} AX + XA^T + Y + Y^T & X \\ X & -Q^{-1} \end{bmatrix} < 0. \tag{7.6}$$

Then the error system (7.5), under the control

$$u = YX^{-1}e(t) - f(t) + v(t). \tag{7.7}$$

where $v(t)$ is an external input, is passive from the external input $v(t)$ to the output $y(t)$, which is defined as

$$y(t) = 2Pe(t). \tag{7.8}$$

Proof The closed-loop error system with the control can be written as

$$\dot{e}(t) = (A + K)e(t) + v(t). \tag{7.9}$$

where $K = YX^{-1}$. Consider a Lyapunov function

$$V(e(t)) = e^T(t)Pe(t). \tag{7.10}$$

where $P = P^T > 0$. Its time derivative along the trajectory of (7.9) is

$$\begin{aligned}\dot{V}(e(t)) &= \dot{e}^T(t)Pe(t) + \dot{e}^T(t)Pe(t) \\ &= e^T(t)[A^TP + PA + PK + K^TP]e(t) + 2e^TPv(t) \\ &= e^T(t)[A^TP + PA + PK + K^TP]e(t) + y^T(t)v(t).\end{aligned}$$

If the following matrix inequality is satisfied

$$A^TP + PA + PK + K^TP + Q < 0. \tag{7.11}$$

We have

$$\dot{V}(e(t)) < -e^T(t)Qe(t) + y^T(t)v(t). \tag{7.12}$$

Integrating both sides of Eq. (7.15) from 0 to t gives

$$V(e(t)) - V(e(0)) < -\int_0^t e^T(\tau)Qe(\tau)d\tau + \int_0^t y^T(\tau)v(\tau)d\tau. \tag{7.13}$$

Let $\beta = V(e(0))$. Since $V(e(t)) > 0$,

$$\begin{aligned}\int_0^t y^T(\tau)v(\tau)d\tau + \beta &> \int_0^t e^T(\tau)Qe(\tau)d\tau + V(e(t)) \\ &\geq \int_0^t e^T(\tau)Qe(\tau)d\tau.\end{aligned} \tag{7.14}$$

the relation (7.14) satisfies the passivity definition according to [10]. Therefore, the error system (7.5) is rendered to be passive from the external input $v(t)$ to the output $y(t)$ under the feedback control. From Schur complement, the matrix inequality (7.11) is equivalent to

$$\begin{bmatrix} A^TP + PA + PK + K^TP & I \\ I & -Q^{-1} \end{bmatrix} < 0. \tag{7.15}$$

Pre- and post-multiplying (7.15) by diag (P^{-1}, I) and $Y = KP^{-1}$, Eq. (7.15) is equivalently transformed into the LMI (7.6). Then the gain matrix of the control $u(t)$ is given by $K = YX^{-1}$. This completes the proof.

Comment 1 Since nonlinear parts of error system is counteracted through directly, scheme supposed here may have some bit of conservativeness in case that the nonlinear part has damping effect to the error system; however, if the nonlinear part is kept, the extra work must be done to verify its damping effect in detail, that cannot be seen as an easy task.

Corollary 1 (zero-input error response) *If the external input $v(t)$ is zero, the closed-loop error system is asymptotically stable.*

Proof When $v(t) = 0$, from Eq. (7.12) we obtain

$$\dot{V}(e(t)) < -e^T(t)Qe(t) \leq 0. \tag{7.16}$$

From lyapunov theory, this guarantees

$$\lim_{t\to\infty} e(t) = 0. \tag{7.17}$$

And by this way, the proof is completed.

Corollary 2 (non-zero-input error response) *If the external input $v(t)$ is selected to be*

$$v(t) = -\lambda y(t), \lambda > 0. \tag{7.18}$$

The closed-loop error system is asymptotically stable.

Proof For $v(t) = -\lambda y(t)$, from Eq. (7.13) the time derivative of $V(e(t))$ satisfies

$$\dot{V}(e(t)) < -e^T(t)Qe(t) - \lambda y^T(t)y(t) \leq 0. \tag{7.19}$$

And from Lyapunov theory, this guarantees the asymptotical stability. This ends the proof.

Comment 2 The LMI problem given in Theorem 1 is to determine whether the solution exists or not. It is called the feasibility problem. The LMI problem can be solved efficiently by using recently developed convex optimization algorithms. In this study, in order to solve the LMI problem, we utilize the solver of feasp in MATLAB LMI Control Toolbox, which implants state-of-the-art interior-point algorithms.

7.5 Example and Simulation Result

To verify and demonstrate the effectiveness of the proposed method, we discuss the simulation result for synchronizing Newton–Leipnik system under different initial conditions. For the numerical simulation, we select $Q = I$, where I

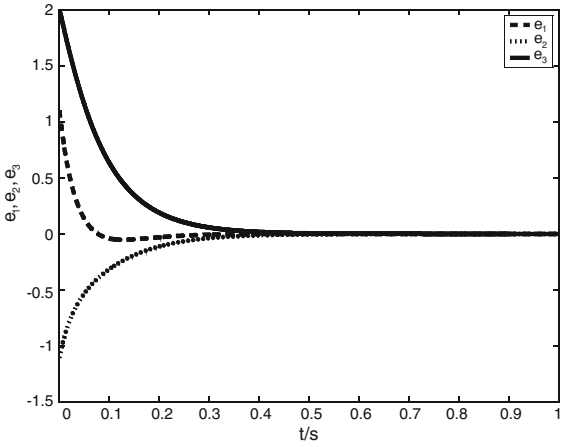


Fig. 7.2 Synchronization error trajectories

is an identity matrix. Applying Theorem 1 to the Newton–Leipnik system yields

$$\begin{aligned} X &= \begin{bmatrix} 4.3741 & 0 & 0 \\ 0 & 2.9899 & 0 \\ 0 & 0 & 3.3221 \end{bmatrix}, \\ Y &= \begin{bmatrix} -3.5208 & -0.3345 & 0 \\ 0.2286 & -3.4784 & 0 \\ 0 & 0 & -3.6649 \end{bmatrix}. \end{aligned} \tag{7.20}$$

Comment 3 In order to obtain the ideal dynamical characteristics, we can assign the expected poles of the closed system, which is determined by the eigenvalues of matrix $(A + K)$, in this paper we design the eigenvalues of matrix $(A + K)$ as $[-15.8, -10.8, -12]$, and based upon the assumption, matrix P can be solved through LMI solver in MATLAB toolbox, and then X and Y can be obtained, thereafter.

Figure 7.2 shows the state trajectories for drive and response systems when the parameter $\lambda = 0$ and the initial conditions are given by

$$\begin{aligned} \begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix} &= \begin{bmatrix} -0.1 \\ 0.1 \\ 0 \end{bmatrix}, \\ \begin{bmatrix} \bar{x}_1(0) \\ \bar{x}_2(0) \\ \bar{x}_3(0) \end{bmatrix} &= \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}. \end{aligned} \tag{7.21}$$

From Fig. 7.2, it can be seen that drive and response systems are indeed achieving chaos synchronization. It can be deduced that the parameter λ only influences the transient response and the larger λ gives the faster response.

7.6 Conclusion

In conclusion, with the passive control, synchronization of the nominal master Newton–Leipnik system and the slave Newton–Leipnik system is realized. Based on Lyapunov theory and LMI approach, the proposed method guarantees to make the closed-loop error system passive and also asymptotically stable. Furthermore, a numerical simulation is given to illustrate the effectiveness of the proposed scheme.

Acknowledgments This work was partially supported by the Program for Excellent Talents in Henan Polytechnic University (Grant No 649033); The Fund for Innovation Talents of Universities and Colleges in Henan Province, China (Grant No. 2009HASTIT021). Leonid Freidovich from Umea University in Sweden has given some constructive suggestions, which are indispensable in improving the quality of this paper. The authors express their thanks for Leonid Freidovich's help.

References

1. Pecora LM, Carrol TL (1990) Synchronization in chaotic systems. Phys Rev Lett 64(8):821–824
2. Wu X, Chen G, Cai J (2007) Chaos synchronization of the master-slave generalized Lorenz systems via linear state error feedback control. Physics D 229:52–80
3. Li CD, Chen G, Liao X, Fan Z (2006) Chaos quasi-synchronization induced by impulses with parameter mismatches. Chaos 16:1–8
4. Leipnik RB, Newton TA (1981) Double strange attractors in rigid body motion with linear feedback control. Phys Lett A 86(2):63–67
5. Lofaro T (1997) A model of the dynamics of the Newton–Leipnik system. Int J Bifur Chaos 7(12):2723–2733
6. Lofaro T (1997) The dynamics of symmetric bimodal maps. Int J Bifur Chaos 7(12):2735–2744
7. Marlin BA (2002) Periodic orbits in the Newton–Leipnik system. Int J Bifur Chaos 12(3):511–523
8. Richter H (2002) Controlling chaotic systems with multiple strange attractors. Phys Lett A 300:182–188
9. Wang XD, Tian LX (2006) Bifurcation analysis and linear control of the Newton–Leipnik system. Chaos Solit Fract 27:31–38
10. Song YZ, Zhao GZ, Qi DL (2006) Passive control of chaotic systems with multiple strange attractors. Chin Phys 15:2266–2270
11. Song YZ, Zhao GZ, Qi DL (2006) The open plus closed loop method for chaotic systems with multiple strange attractors. Chin Phys 16:1918–1922

Chapter 8
Adaptive Inverse Control Method Based on SVM-Fuzzy Rules Acquisition System for Twin-Lift Spreader System

Xixia Huang, Fanhuai Shi and Hui Zhang

Abstract This paper proposes a new method of adaptive inverse control based on Support Vector Machine-Fuzzy Rules Acquisition System (SVM-FRAS) for twin-lift spreaders. In this control mechanism, an identifier is established based on SVM-FRAS, and an inverse controller based on SVM-FRAS is designed. The proposed adaptive inverse control method can automatically extract control rules from the process data. Comprehensibility is one of the required characteristics for a complex twin-lift spreader. We use the proposed SVM-FRAS-based adaptive inverse control method to obtain the rule-based process and the control model of the twin-lift spreader. Based on the simulation experiments for twin-lift spreaders, the SVM-FRAS adaptive inverse control method is found to be effective.

Keywords Adaptive inverse control · Support vector machine · Fuzzy rule · Twin-lift spreaders

X. Huang (✉) · H. Zhang
Marine Technology and Control Engineering Key Laboratory, Shanghai Maritime University, 1550 Pudong Ave., Shanghai 200135, China
e-mail: merryhuang1975@gmail.com

H. Zhang
e-mail: 357021885@qq.com

F. Shi
College of Electronics and Information Engineering, Tongji University, 4800 Cao'an Highway, Shanghai 201804, China
e-mail: fhshi@tongji.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_8,

8.1 Introduction

Compared with single-lift spreader, twin-lift spreader has fast speed, and averagely each lift cost only about 60–80 s. Currently, twin-lift spreader can move two 40' containers or 45' containers or four 20' containers at a time. Additionally, for both spreaders, independent lifting and the situation that one is lifting and the other stays still can be realized, as well as the handling operation of single-side spreaders, namely one of both spreaders is empty. Because of different container weight and lifting speed, and the difference in motor features, advantages of twin-lift yard bridges and twin-lift spreaders cannot bring into full play.

Therefore, the synchronized control of twin-lift spreader movements is very important to increase handling efficiency. Currently, twin-lift spreaders are mostly used for handling operations, especially for empty containers. While loading on board, twin-lift spreaders are mostly used for the operation of empty 40' containers. The imbalance in the loading and unloading operations of twin-lift spreaders, and the imbalance between full and empty containers are the main issues. The core technology of lifting yard bridges is to control the synchronized motions of yard bridges, but the motion of the spreaders of lifting yard bridges is driven by motors, so to control the synchronous motion of twin-lift spreaders, the key is to control the rev of the two asynchronous motors.

When using twin-lift yard bridges and twin-lift spreaders at ports, both spreaders are connected by mechanical rods, so the purpose of mechanical control of the synchronized motions of both spreaders is realized. During the operation, the two motors are controlled artificially, so good coordinative motion would not be realized. Additionally, for the handling imbalance and the disunity of container weight needs to be adjusted, while at the container yard, empty and full containers and containers of different types are sorted artificially, and containers with quite different weight should be sorted out as much as possible after loading on board. Although to some extent, it is helpful for the coordinative motions of twin-lift spreaders, additional manpower and material should be invested, as it lacks enough flexibility and suitability, and cannot handle accidents and not fit with the intelligent and unmanned concept in the future. Based on these problems, the paper underlines the control of twin-lift yard bridges and twin-lift motors, and studies the automatic adjustment that twin-lift motors make in case of mechanical control and artificial preset. It lays the foundation for the establishment of intelligent and unmanned port.

Support vector machine (SVM), developed by Vapnik [1], is gaining popularity due to its various attractive features and promising empirical performance. Originally, SVM is developed for pattern recognition problems, but in recent years, with the introduction of ε-insensitive loss function, SVM has been extended to solve nonlinear regression estimation, time-series prediction, system nonlinear identification, and control [2].

Recently, Huang et al. [3, 4] have proposed a fuzzy modeling method based on support vector machine (SVM) for the arc welding process. In [3, 4] SVM is used to

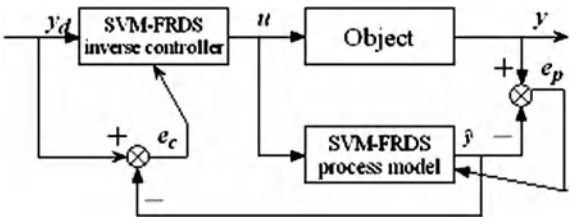


Fig. 8.1 Adaptive inverse control architecture based on SVM-FRAS

extract If–Then rules; the fuzzy basis function inference system is adopted as the fuzzy inference system. So the approach possesses good comprehensibility as well as satisfactory generalization capability. Based on the previous modeling work, we apply adaptive inverse control3 method based on SVM-FRAS to twin-lift spreaders control in this paper. In this control mechanism, an identifier is established based on SVM-FRAS, and an inverse controller based on SVM-FRAS is designed. Finally, the Hard-in-loop experiment on twin-lift spreaders is also conducted.

8.2 SVM-FRAS-Based Adaptive Inverse Control

The architecture of the proposed SVM-FRAS-based adaptive inverse control is shown in Fig. 8.1. The control algorithm can be separated into two sub-stages, namely, the identification stage and the control stage, where the (1) SVM-FRAS algorithm to train the system model was used, and (2) u was obtained by using the SVM-FRAS inverse model to control output y according to reference input y_d.

8.2.1 SVM-FRAS-Based Identification

The mathematical model of the discrete-time nonlinear single-input single-output system can be described by Eq. (8.1):

$$\begin{aligned} y(k+1) &= f(y(k), \ldots, y(k-n), \\ &u(k), \ldots, u(k-m)) \end{aligned} \tag{8.1}$$

where $y \in R^n$, $u \in R^m$, $m \leq n$, u and y represent the control input and output of the nonlinear system, respectively.

Define

$$\begin{aligned} X(i) &= (y(i), y(i-1), \ldots, y(i-n), u(i), \\ &u(i-1), \ldots, u(i-m)) \end{aligned} \tag{8.2}$$

where $i = 1, 2, \ldots, N$ refers to the number of samples. Equation (8.1) can be represented as

$$y(i+1) = f(X(i)) \tag{8.3}$$

where f is the desired function. Constructing the learning sample $(X(i), y(i+1))$ and using the supporting vector regression, the common form of fuzzy rules has been extracted:

$$\begin{aligned} R_j &: \text{IF } X \text{ is } X(j), \\ &\quad \text{THEN } y \text{ is } y(j+1) \end{aligned} \tag{8.4}$$

where $j = 1, 2, \ldots, N\_R$, N_R is the rule number, X is the input variable, and y is the output variable of the fuzzy system.

The model of $y(k+1)$ is

$$\begin{aligned} y(k+1) &= \sum_{j=1}^{N\_R} p_j(x) \cdot \bar{z}_j \\ p_j(x) &= \exp\left(-r\|X(j) - X(k)\|^2\right) \end{aligned} \tag{8.5}$$

In this model, $p_j(x)$ is the fuzzy basis function which measures the extent to which the sample approaches the j rule, and $\bar{z}_j$ is the output of the R_j Rule.

8.2.2 SVM-FRAS-Based Inverse Controller

Reference [5] is a detailed study on the invertible problem of Eq. (8.1). From such, the following conclusion is made:

Theorem 1 *If Eq. 8.1 is strictly monotone for u, the system is invertible at* $[y(k), \ldots, y(k-n), u(k), \ldots, u(k-m)]^T$.

If the control process is revertible, then it is possible to construct the inverse model of the process using an appropriate method. This inverse model is the controller designed in this paper. Suppose that the system is invertible, then the inverse model of the system is the controller, that is,

$$\begin{aligned} u(k) = f^{-1}&(y(k+1), y(k), \ldots, y(k-n), \\ &u(k-1), \ldots, u(k-m)) \end{aligned} \tag{8.6}$$

where f^{-1} *is the desired nonlinear function.*

As for the nonlinear system, it is difficult to obtain the analytical solution of the inverse model. However, since the process is invertible, there must exist an inverse model which may be estimated by the SVM-FRAS method.

Define

$$\begin{aligned} Y(k) = (&y(k+1), y(k), y(k-1), \ldots, y(k-n), \\ &u(i-1), \ldots u(k-m)) \end{aligned} \tag{8.7}$$

In constructing the learning sample $(Y(i), u(i))$, we extract the fuzzy rule model through the SVM-FRAS method:

$$\begin{aligned} R_j : &\text{IF } Y \text{ is } Y(j), \\ &\text{THEN } u \text{ is } u(j) \end{aligned} \tag{8.8}$$

where $j = 1, 2, \ldots, N\_R$, N_R is the rule number, Y is the output variable, and u is the output variable of the fuzzy system.

The model of $u(k)$ is

$$\begin{aligned} u(k) &= \sum_{j=1}^{N\_R} p_j(x) \cdot \bar{z}_j \\ p_j(x) &= \exp\left(-r\|Y(j) - Y(k)\|^2\right) \end{aligned} \tag{8.9}$$

In this case, N_R is the rules number, and $p_j(x)$ is the fuzzy basis function which measures the extent to which the sample approaches the j rule. $\bar{z}_j$ is the consequent of R Rule.

SVM-FRAS is applied to approximate the nonlinear function of f and f^{-1}. Furthermore, forward and inverse models are constructed, respectively. The accurate inverse model constructed by SVM-FRAS makes a steady-state gain between the controller and the model to be ranked 1. Thus, this guarantees no steady-state error in the tracking characteristic of the control system.

8.3 SVM-FRAS-Based Adaptive Inverse Control for Twin-Lift Spreaders

8.3.1 Experiment Platform

The experiment platform control system is the three-layer distributed control system, just like the figure. The system is based on Profibus on-site bus network, connects different equipments, realizes the coordination and the sharing of data and information, and achieves the experiment platform control and management function. DC motor of the multi-motor drive experiment platform is used to simulate different container weight; the number of motor turns is used to simulate container position; two driver motors are used to simulate the position synchronization, and simulate different motor features. See Fig. 8.2 for experiment platform system structure.

We first give the following analysis about the input/output variables for modeling of the twin-lift spreader. Currently, twin-lift spreaders are mainly used to unload the ship, especially for empty containers. Now the main problem is that different container weight, different lift position caused by different container types and different motor features result in the handling imbalance and the imbalance

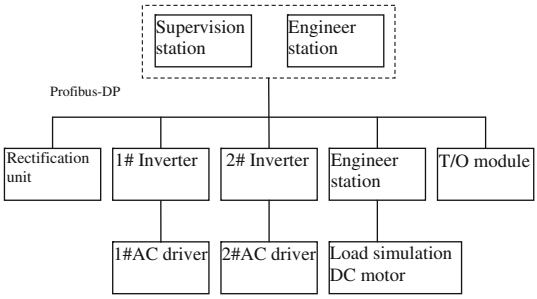


Fig. 8.2 Structural drawing of the control system of multi-motor drive experiment platform

between empty and full containers. To solve the problem, by studying control technology of twin-lift yard bridges and twin-lift spreaders, and by applying the track control system at semiclosed-loop position, just like Fig. 8.3, the purpose of synchronization of twin-spreader positions is realized. Through the timely information of track respond parameter from main motor and submotor, position incoordination is detected, and the solution for the imbalance and low efficiency of twin-lift spreaders comes out.

The input variables should be Speed (V). The output of the model should be Position (S). The history information (including Speed and Position) should also be included in the model inputs.

8.3.2 Controller Design Based on SVM-FRAS

When $m = n = 1$, the relation model between V and S is obtained from Eq. (8.10):

$$V(k) = f^{-1}(S(k+1), S(k), S(k-1), V(k-1)) \tag{8.10}$$

In this study, the training set (500 data) and the test set (260 data) are randomly split to investigate the performance of the model. The controller design is first made in the training set using the SVM-FRAS inverse model.

The rules set is obtained using margin support vectors through SVM.

The reasoning model of the GTAW process controller is

$$\begin{aligned} V(k) &= \sum_{j=1}^{N\_R} p_j(x) \cdot \overline{V_j} \\ p_j(x) &= \exp\left(-r\|Y(j) - Y(k)\|^2\right) \\ Y(k) &= (S(k+1), S(k), S(k-1), V(i-1)) \end{aligned} \tag{8.11}$$

where N_R is the number of rules. In addition, $p_j(x)$ is the fuzzy basis function which measures the extent to which the sample approaches the j rule.

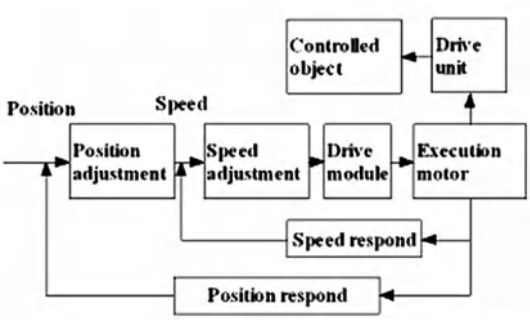


Fig. 8.3 Position control system

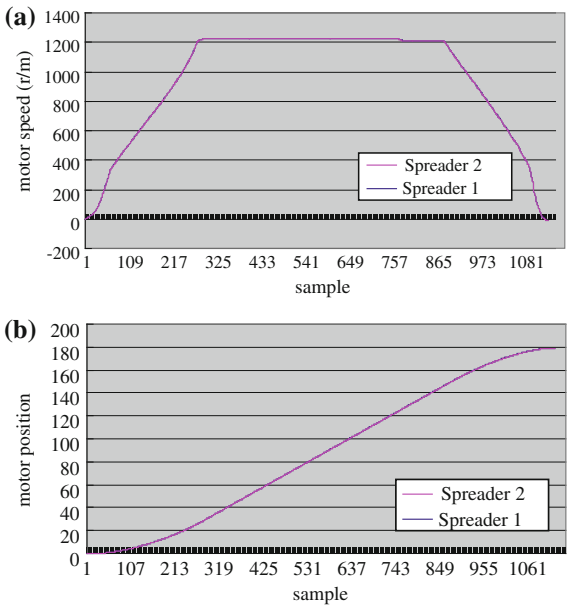


Fig. 8.4 Synchro control (a) motor speed (b) motor position

8.3.3 Hard-in-Loop Experiment of Adaptive Inverse Control for Twin-Lift Spreaders

During the operation of twin-lift yard bridges, the process of start, stable speed, speed reduction, and stop is usually experienced. See Fig. 8.4 for the simulation curve for the motor speed and position of Spreader 1 and the motor position n_2 of Spreader 2. Spreader 2 can follow the motions of Spreader 1 reliably. Synchronization error of both spreaders is little and the synchronization effect is good.

8.4 Conclusion

A new method of adaptive inverse control based on the Support Vector Machine-Fuzzy Rules Acquisition System (SVM-FRAS) for twin-lift spreaders is proposed in this study. The proposed adaptive inverse control method can automatically extract control rule from twin-lift spreaders data. Overall, the Hard-in-loop experiments for twin-lift spreaders show the efficacy of the SVM-FRAS adaptive inverse control method.

Acknowledgments This work is supported by Shanghai Leading Academic Discipline Project (S30602), National High-tech R&D Program of China (863 Program) by Ministry of Science & Technology of China (No. 2009AA043000) and National Natural Science Foundation of China (No. 60805018).

References

1. Smola AJ, Scholkopf B (2004) A tutorial on support vector regression. Stat Comput 14:199–222
2. Mittul S, Jivitej C, Puneet A, Suresh C (2011) Reduced twin support vector regression. Neurocomputing 74:1474–1477
3. Huang XX, Shi FH, Gu W, Chen SB (2009) SVM-based fuzzy rules acquisition system for pulsed GTAW process. Eng Appl Artif Intel 22:1245–1255
4. Huang XX, Gu W, Shi FH, Chen SB (2009) An adaptive inverse control method based on SVM-fuzzy rules acquisition system for pulsed GTAW process. Int J Adv Manuf Technol 44:686–694
5. Widrow B, Walach EW (1996) Adaptive inverse control. Prentice-Hall, Upper Saddle River

Chapter 9
SVM-Based Fuzzy Rules Acquisition System for Twin-Lift Spreader System

Xixia Huang, Fanhuai Shi and Hui Zhang

Abstract This paper proposes a support vector machine-based fuzzy rules acquisition system (SVM-FRAS) for twin-lift spreaders. The character of SVM in extracting support vector provides a mechanism to extract fuzzy If–Then rules from the training data set. We construct the fuzzy inference system using fuzzy basis function. The gradient technique is used to tune the fuzzy rules and the inference system. Modeling is one of the key techniques in the automatic control of twin-lift spreaders, and is still a very difficult problem. Comprehensibility is one of the required characteristics in modeling for the complex twin-lift spreaders. We use the proposed SVM-FRAS to obtain the rule-based model of twin-lift spreaders. Experimental results show the SVM-FRAS model possesses good generalization capability as well as high comprehensibility.

Keywords Modeling · Fuzzy rules · Support vector machine · Twin-lift spreader

X. Huang (✉) · H. Zhang
Marine Technology and Control Engineering Key Laboratory,
Shanghai Maritime University, 1550 Pudong Ave., Shanghai 200135, China
e-mail: merryhuang1975@gmail.com

H. Zhang
e-mail: 357021885@qq.com

F. Shi
College of Electronics and Information Engineering, Tongji University,
4800 Cao'an Highway, Shanghai 201804, China
e-mail: fhshi@tongji.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_9,

9.1 Introduction

Existing yard bridge system of container terminals at home and abroad are all used for handling operations of single containers (operations of single lifting and single spreading, only one 40′ container can be handled once), and the handling efficiency is very low. But compared with single-lift spreader, twin-lift yard bridge can lift two 40′ container or four 20′ containers. Twin-lift spreader has fast speed, and averagely each lift only cost about 60–80 s. Twin-lift spreader can not only move on-deck containers, but can also move under-deck containers. Besides moving two similar containers at the same time, twin-lift spreader can move the containers that do not have similar types, so the handling efficiency increases dramatically.

For most Rubber Tyred Gantries (RTG) and Rail Mounted Gantries (RMG) at home and abroad, the hoisting capacity is 40 t. Only one 20′ or 40′ container can be lifted at a time, but more and more twin-lift spreaders are selected as shore container cranes at wharf apron. The front and the back do not match, so handling and container yard efficiency are influenced heavily. Thames Container Terminal in Britain together with companies of Morris and Automation has finished the application test of the first twin-lift spreader in unpiloted and full automatic RMG. Its rated capacity is 50 t, and it can move two 20′ containers simultaneously. ZPMC Company developed the double-container RTG in 2002. Its lifting capacity was 60 t and it can move two 20′ containers. Proved by two years test, it had good effect. Most existing container cranes use special hydraulic spreaders. At present, all electric container spreaders begin to draw people's attention and new related products are launched one by one. For example, ZPMC, shipping department with ministry of communications, Mitsui Company and many European companies all launch electric spreaders. Because of electric container spreader, container yard cranes do not have to use hydraulic systems, so the crane deadweight decreases, and equipment investment and running cost is saved. According to statistics, for the same RTG, if electric spreader is installed, compared with hydraulic spreader, 10% of energy is saved.

On October 25, 2006, Sheko Container Terminal (SCT) ordered two shore container loading bridges (bridge crane or shore bridge in short) from Shanghai Zhenhua Port Machinery Company, and both bridges were transported to the terminal successfully. They were the most advanced and the biggest shore container loading bridges in the world. Both can install two retractable twin-container spreaders and unload two 40′ containers or four 20′ containers. Lifting height reached 43 m from the rail level; the outreach was 65 m; maximum load-bearing under spreader reached 100 t. They can handle the largest post panamax container ship.

According to reports from spreader manufacturers around the world, as container transport grows quickly, the number of container loading bridges increases quickly and the demand for spreaders also grows dramatically. Many container terminal companies have to select new container spreaders with bigger load and more powerful functions for handling operations of panamax or post

panamax container ships. Therefore, port handling capacity increases with the lowest investment. To a certain extent, the application of twin-lift spreaders increases the container-handling efficiency, but because of some objective reasons, its strengths do not bring into play completely.

Currently, container twin-lift spreader can move two $40'$ containers or $45'$ containers or four $20'$ containers at a time. Additionally, for both spreaders, independent lifting and the situation that one is lifting and the other stays still can be realized, as well as the handling operation of single-side spreaders, namely one of both spreaders is empty. Because of different container weight and lifting speed, and the difference in motor features, advantages of twin-lift yard bridges and twin-lift spreaders cannot bring into full play. Therefore, the synchro control of twin-lift spreader movements is very important to increase handling efficiency. Currently, twin-lift spreaders are mostly used for handling operations, especially for empty containers. While loading on board, twin-lift spreaders are mostly used for the operation of empty $40'$ containers. The imbalance in the loading and unloading operations of twin-lift spreaders, and the imbalance between full and empty containers are the main issues. The core technology of lifting yard bridges is to control the synchronized motions of yard bridges, but the motion of the spreaders of lifting yard bridges is driven by motors, so to control the synchronous motion of twin-lift spreaders, the key is to control the rev of the two asynchronous motors.

When using twin-lift yard bridges and twin-lift spreaders at ports, both spreaders are connected by mechanical rods, so the purpose of mechanical control of the synchronized motions of both spreaders is realized. During the operation, the two motors are controlled artificially, so good coordinative motion would not be realized. Additionally, for the handling imbalance and the disunity of container weight that needs to be adjusted, while at the container yard, empty and full containers and containers of different types are sorted artificially, and containers with quite different weight should be sorted out as much as possible after loading on board. Although to some extent it is helpful for the coordinative motions of twin-lift spreaders, additional manpower and material should be invested, as it lacks enough flexibility and suitability, and cannot handle accidents and cannot fit with the intelligent and unmanned concept in the future. Based on these problems, this paper underlines the modeling of twin-lift yard bridges and twin-lift motors, and studies the automatic adjustment that twin-lift motors are made of in case of mechanical control and artificial preset. It lays the foundation for the establishment of intelligent and unmanned port.

Support vector machine (SVM), developed by Vapnik [1], is gaining popularity due to many attractive features and promising empirical performance. Originally, SVM is developed for pattern recognition problems. Recently, with the introduction of ε-insensitive loss function, SVM has been extended to solve nonlinear regression estimation, time-series prediction, system nonlinear identification and control [2].

Recently, Huang et al. [3, 4] proposed a fuzzy modeling method based on SVM for the arc welding process. In [3, 4] SVM is used to extract If–Then rules; the fuzzy basis function inference system is adopted as the fuzzy inference system.

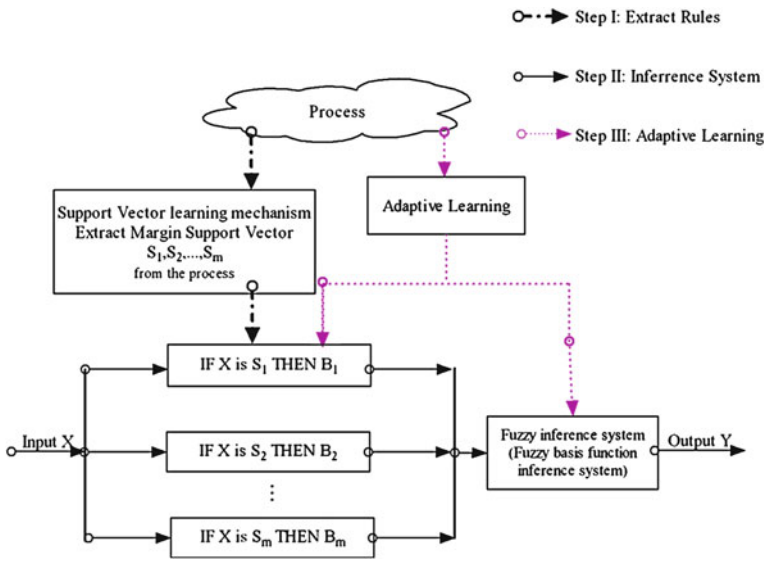


Fig. 9.1 Schematic diagram of SVM-based fuzzy rules discovery system

So the approach possesses good comprehensibility as well as satisfactory generalization capability. Based on the previous modeling work, we apply SVM-FRAS for modeling of the twin-lift spreaders in this paper.

9.2 SVM-Based Fuzzy Rules Acquisition System

In this section, we introduce a fuzzy rule-based inference system utilizing SVM. The architecture of the proposed SVM-based fuzzy inference is shown in Fig. 9.1. For this approach, the first step is to use SVM to extract If–Then rules; the next step is to adopt the fuzzy basis function inference system as the fuzzy inference system; the third step is the adaptive learning for the fuzzy rules and the parameters in the fuzzy inference system.

9.2.1 Extract Fuzzy Rules Using SVM

SVM learning mechanism provides a solid foundation to extract support vectors for further use in the If–Then rules generation [3, 4]. As described in Sect. 9.2, only a small part of training data, the support vectors, contribute to the solution.

Geometrically, they are points that lie on or outside the ε-tube. In other words, they are in "key position". Note that we choose only margin support vectors $(0 \prec |\alpha_i - \alpha_i^*| \prec C)$ because error support vectors $(|\alpha_i - \alpha_i^*| = C)$ may be far from the regression curve, as the empirical error should be equal or more than ε in this case.

Our basic idea is to extract the fuzzy rules depending on some important "key points", which condense decision information from learning process and provides a set of "basis" to express the overall fuzzy system. In fact, the essential goal is to extract the appropriate "key points" using the appropriate kernels. Various real-valued function estimation problems need various sets of approximating functions. Therefore, it is important to choose or even construct special kernels that reflect special properties of the approximating functions.

9.2.2 Fuzzy Inference System

We adopt the following fuzzy basis function inference system as the fuzzy inference system:

$$\begin{aligned} y = f(x) &= \sum_{j=1}^{N\_R} p_j(x)\bar{z}_j; \\ p_j(x) &= \exp\left(- \sum_{k=1}^{N\_D} r_k \cdot \left(x_k - \bar{x}_{jk}\right)^2 \right) \Big/ \sum_{j=1}^{N\_R} \exp(- \sum_{k=1}^{N\_D} r_k \cdot (x_k - \bar{x}_{jk})^2) \end{aligned} \tag{9.1}$$

where $f : R^{N\_D} \rightarrow R$, N_D is the number of the dimension of the input/output variables, N_R is the number of rules, $p_j(x)$ is the fuzzy basis function, which indicates how close the sample is to the jth rule, $\bar{z}_j$ is the output of the jth rule, r_k is the coefficient of the kth dimension, x_k is the value of the kth dimension of the sample, $\bar{x}_{jk}$ is the value of the kth dimension of the jth rule.

The most important advantage of using fuzzy basis functions is that an If–Then rule is directly related to a fuzzy basis function. A fuzzy basis function expansion provides a natural framework for combining both numerical and linguistic information in a uniform fashion. The inference system reflects the human habit of thinking: the closer the sample is to the rule, the closer the output of the sample is to the output of the rule.

9.3 Modeling for the Twin-Lift Spreader Using SVM-FRAS

The key technology of double container cranes is to control the double cranes move synchronized. In order to keep the two cranes synchronized the most important is to control the motors' speeds and positions.

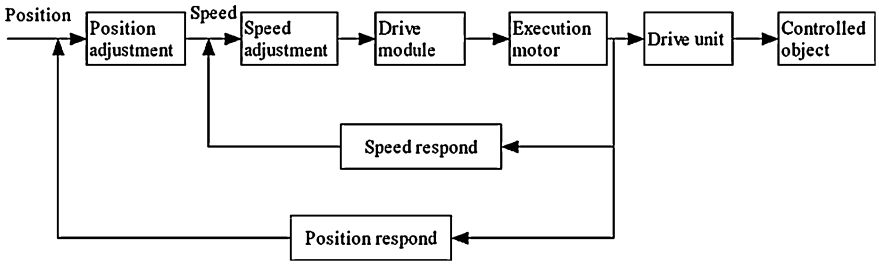


Fig. 9.2 Position control system

9.3.1 Select Input/Output Variables for Modeling of Twin-Lift Spreader

We first give the following analysis about the input/output variables for modeling of the twin-lift spreader. Currently, twin-lift spreaders are mainly used to unload the ship, especially for empty containers. Now the main problem is that different container weight, different lift position caused by different container types and different motor features result in the handling imbalance and the imbalance between empty and full containers. To solve the problem, by studying control technology of twin-lift yard bridges and twin-lift spreaders, and by applying the track control system at semi-closed loop position, just like Fig. 9.2, the purpose of synchronization of twin-spreader positions is realized. Through the timely information of track respond parameter from main motor and sub-motor, position incoordination is detected, and the solution for the imbalance and low efficiency of twin-lift spreaders comes out.

The input variables should be Speed (V). The output of the model should be Position (S).The history information (including Speed and Position) should also be included in the model inputs.

9.3.2 Acquire the Raw Data for Modeling of the Twin-Lift Spreader

Based on the analysis about the twin-lift spreader V is selected as the input signals. Random signals are considered as the optimal input signals to the twin-lift spreader.

9.3.3 Extract Fuzzy Rules Using SVM-FRAS for Modeling of Twin-Lift Spreader

We adopt the RBF kernel $K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2), \gamma > 0$ as the kernel function of SVM. The twin-lift spreader is inherently nonlinear. The RBF kernel nonlinearly maps samples into a higher dimensional space, so it can handle the case when the relation between inputs and outputs is nonlinear.

The fuzzy inference function adopted in this application is Eq. (9.1). After linearly scaling each attribute to the range $[-1, +1]$, we set $r_k(k = 1, 2, \ldots, 15) = 0.75$.

9.3.4 Experiments Results

For the purpose of validating the models, the data set is randomly split into 300 training samples and 200 test samples. For SVM-FRAS, the error is 3.3%. Considering the complexity of twin-lift spreader the proposed approaches, especially that with adaptive learning, are quite acceptable.

9.4 Conclusion

This paper proposes a SVM-based fuzzy modeling method for the twin-lift spreader. SVM is used to select significant fuzzy basis functions, linear combination of fuzzy basis functions is used to construct the final fuzzy rules acquisition system. We use the proposed SVM-FRAS to obtain the rule-based model of twin-lift spreader. Experimental results show that the SVM-FRAS model is effective.

Acknowledgments This work is supported by the National High-tech R&D Program of China (863 Program), by the Ministry of Science & Technology of China (No. 2009AA043000), Shanghai Leading Academic Discipline Project (S30602) and the National Natural Science Foundation of China (No. 60805018).

References

1. Smola AJ, Scholkopf B (2004) A tutorial on support vector regression. Stat Comput 14: 199–222
2. Mittul S, Jivitej C, Puneet A, Suresh C (2011) Reduced twin support vector regression. Neurocomputing 74:1474–1477
3. Huang XX, Shi FH, Gu W, Chen SB (2009) SVM-based fuzzy rules acquisition system for pulsed GTAW process. Eng Appl Artif Intel 22:1245–1255
4. Huang XX, Gu W, Shi FH, Chen SB (2009) An adaptive inverse control method based on SVM-fuzzy rules acquisition system for pulsed GTAW process. Int J Adv Manuf Technol 44:686–694

Chapter 10
Application of Particle Swarm Optimization in the Decision-Making of Manufacturers' Production and Delivery

Lingxiao Yang and Liangyou Shu

Abstract The aim of this paper is to study the loading decision problem in manufacturer's product distribution. Owing to the order fulfillment optimization condition of the manufacturer, the decision-making model of manufacturers' production and delivery has been founded. This paper has given out the algorithm finding the solution based on particle swarm optimization. The results indicate that the decision-making of manufacturers' production and delivery is a complicated N-P decision-making problem and finding the solution is also very difficult. The solving algorithm based on particle swarm optimization is effective to the model of this paper.

Keywords Management engineering · Particle swarm optimization · Optimization model

10.1 Introduction

With increasing users' high demands to timely delivery of products, the response speed of supply chain to user demands is significant, which has become a decisive factor to obtain inside track in competition and win the competitive advantage.

L. Yang (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, China
e-mail: yanglx@hpu.edu.cn

L. Shu
School of Economics and Management, Henan Polytechnic University,
Jiaozuo, China
e-mail: lyshu@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_10,

Long respond cycle to the supply chain results in high costs [1–3]. Reducing response time can enhance the response ability in market fluctuations [4]. The key of implementing Efficient Consumer Response (ECR) is to ensure the supply according to demand of channel partners [5, 6], which can maintain production flexibility, improve overall supply chain efficiency and reduce time and costs [7].

Taken into impacts of the back distribution risk and time competition of manufacturers, Shu studied the order fulfillment optimization model of manufactures based on time competition [8]. The solution algorithm of the order fulfillment optimization model is not discussed. Therefore, this paper discussed the method of solving of the model.

Particle swarm optimization (PSO) invented by Eberhart and Kennedy is based on the behavior research of birds preying [9]. PSO is fast and has high-quality solutions, and brevity code, etc., especially the field of multi-dimensional continuous space optimization problems, neural network training [10] and application in power system [11]. It is not easy to express particle position and velocity. This restricted its application in the combinatorial optimization field. Based on the traditional speed-displacement search model, Gao haibing etc. analyzed the PSO mechanism, and proposed a generalized PSO model, so that the PSO algorithm can be applied to the fields of discrete and combinatorial optimization [12]. So, the author attempts to use the PSO algorithm to solve the model

10.2 Problem Description and Basic Assumption

In the given time, the manufacturers who adopt the MTO management policy will begin to produce after receiving the customers' order about some kind of products. Customers give lead-time to the order of the products and have request about the delivering time. They will ask delay fees if the manufacturer provides the goods late, and the manufacturers have original material purchasing lead-time. The products which have been completed can be delivered by transporters. The products which cannot be delivered in time can temporarily be stored in the manufactory. When subjected to the demands of customers, requests of quick respond, the efficiency of manufacturers' produces and various delivery expenses and abilities, manufacturers optimize producing and delivery plans to reach the goal of total costs' minimum concluding the production cost, delivery cost, cost of storing in the manufactory and punitive fee of delaying.

For the convenience of study, following assumptions are made: (1) the costs and time can be different for different transporters in completing the task; the repeated use of transport vehicles considered as different vehicle; (2) storage costs during transport including the transportation costs; (3) manufacturers doing production only for a single product orders; (4) products of manufacturer can be shipped immediately or in the following time. Products shipped in the following time can temporarily stored in the manufacturers'; the loading time of products is out of restrictions, the time needed during loading is omitted and loading

completed at one-time; (5) only the punishment costs in distribution delay of users is considered; (6) the internal transport between members is considered as the internal management of the members, the costs incurred included in the corresponding costs of its members.

10.3 Order Fulfillment Optimize Model of Manufacturers

Set of product transporting companies from the manufacturer to the user is E. Highest producing capacity of the manufacturers is Q. Cost per unit production is C; number of vehicles used by transporter j is D_j, the fixed cost of using each vehicle is CT_j, the largest shipment of each vehicle as QF_j, per unit cost of product transportation is U_j, the time required to complete one transport from the manufacturer to the user is T_j, the largest transport capacity to the user is YQ_j; the quantity of product which the user orders is BQ; the fee paying by manufacturers for an unit equivalent back-order is ω; lead time of user ordered is BT; the longest-lead time for manufacturers ordering the raw material is MT; the time manufacturers need to produce per unit product is PT; an unit equivalent cost for inventory time of product in the manufacturers' is h; N^+ is defined as a non-negative integer set.

The shipment amount of the dth vehicle using by transporter j from the manufacturer to the users' is defined as z_{jd}, γ_j the decision-making variable of transporter. If transporter j is choose to transport products, then $\gamma_j = 1$; otherwise, $\gamma_j = 0$.

Owing to that the above-mentioned hypothesis and the idea of the order fulfillment optimization model [8], the decision-making model of manufacturers' production and delivery can be founded.

This optimal model is as follows:

$$\begin{aligned}
&\min QF \times C + \sum_{j\in E}\sum_{d\in D_{jv}} \left(z_{jd} \times U_{jv}\right) + D_j \times CT_j \times \gamma_j \\
&+ \sum_{j\in E} \frac{h \times PT}{2} \left[\sum_{d\in D_j} \left(z_{jd}\right)^2 - \sum_{d_j} z_{jd}\right] \times \gamma_j \\
&+ \sum_{j\in E}\sum_{d\in D_j} \max\left(\left(PT \times \left(\sum_{l=1}^{d} z_{jl}\right) + T_j - BT + MT\right) \times z_{jd} \times \omega, 0\right) \times \gamma_j
\end{aligned} \tag{10.1}$$

Subject to the following constraints:

$$\sum_{j\in E}\left(\sum_{d\in D} z_{jd}\right) \times \gamma_j \le Q \tag{10.2}$$

$$z_{jd} \le QF_j,\ d = 1, 2, \ldots, D_j,\ j \in E \tag{10.3}$$

$$\sum_{j\in E}\sum_{d=1}^{D_j} z_{jd} \times \gamma_j = BQ \tag{10.4}$$

$$\sum_{d=1}^{D_j} z_{jd} \le YQ_j,\ j \in E \tag{10.5}$$

$$\sum_{j\in E} \gamma_j = 1,\ \gamma_j \in \{0,1\} \tag{10.6}$$

$$z_{jd} \in N^+,\ d = 1,2,\ldots,D_j,\ j \in E. \tag{10.7}$$

The first item in the objective function is production costs; the second is product transport cost; transport costs can also be divided into fixed costs and variable costs; the third is inventory cost; the fourth is punishment fee of products back delivery. Formula (10.2) in the subjects is producing capacity of manufacturers. Formula (10.3) is shipment capability subject from the manufacturer to the user. Formula (10.4) is subject of user demands. Formula (10.5) is subject of transport capacity. Formula (10.6) is subject to the selection of transporter. Formula (10.7) is non-negative subject.

10.4 An Improved Particle Swarm Optimization

In this section, an improved PSO algorithm is proposed for solving the decision-making of manufacturers' production and delivery.

PSO starts with a population of random solutions 'particles' in a D-dimension space. The ith particle is represented by $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$. Each particle keeps track of its coordinates in hyperspace, which are associated with the fittest solution it has achieved so far. The value of the fitness for particle i is stored as $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$ that its best value is represented by (pbest). The global version of the PSO keeps track of the overall best value (gbest), and its location, obtained thus far by any particle in the population. PSO consists of, at each step, changing the velocity of each particle toward its pbest and gbest according to Eq. (10.8). The velocity of particle i is represented as $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$. Acceleration is weighted by a random term, with separate random numbers being generated for acceleration toward pbest and gbest. The position of the *i*th particle is then updated according to Eq. (10.9) [13, 14]:

$$v_{id}(t+1) = \omega v_{id}(t) + c_1 r_1 (p_{id} - x_{id}(t)) + c_2 r_2 \left(p_{gd} - x_{id}(t)\right) \tag{10.8}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{10.9}$$

where, p_{id} and p_{gd} are pbest and gbest. It is concluded that gbest version performs best in terms of median number of iterations to converge. However, pbest version

with neighborhoods of two is most resistant to local minima. The results of past experiments about PSO show that x was not considered at an early stage of PSO algorithm. However, x affects the iteration number to find an optimal solution. If the value of x is low, the convergence will be fast, but the solution will fall into the local minimum. On the other hand, if the value will increase, the iteration number will also increase and therefore the convergence will be slow. Usually, for running the PSO algorithm, value of inertia weight is adjusted in training process. It was shown that PSO algorithm is further improved via using a time decreasing inertia weight, which leads to a reduction in the number of iterations [14]. In Eq. (10.8), term of $c_1 r_1(p_{id} - x_{id}(t))$ represents the individual movement and term of $c_2 r_2\left(p_{gd} - x_{id}(t)\right)$ represents the social behavior in finding the global best solution.

The PSO algorithm is described as follows:

1. Initialize a population of N particles with random positions and random velocities with D dimensions in a given searching space.
2. While a specified stop condition (the optimal solution is found or the maximal number of iterations is reached) is not met do
3. Evaluate the fitness of each particle in the populations according to the objective function of the problem. For each particle id do
4. Update the personal best position p_{id} for each particle and the global best position p_{gd} for all particles.
5. By Eq. (10.8), update the velocity of each particle.
6. By Eq. (10.9), update the position of each particle.
7. End for
8. End while

In order to improve the solution quality, the inertial weight ω, the cognition learning factor c_1 and the social learning factor c_2 are varying with time, not using the fixed coefficients. PSO algorithm can get better solution by letting ω vary from higher to lower by the increase of time [15, 16]. Let ω be varying with time by the following linear decreasing function.

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{\text{MaxIter} - 1} \times (\text{Iter} - 1) \tag{10.10}$$

where MaxIter is the maximum iterations during the evolutionary process, Iter is the current iteration number, $\omega_{\max}$ is the initial value of ω and $\omega_{\min}$ is the final value of ω, respectively.

10.5 Experimental Results

Assuming that the manufacturer has received order that the user needs 200 products, maximum production capacity of manufacturers is 260 in the period, cost of unit production is 200 yuan, lead time products order is 500 h, the time consumed in producing one unit by manufacturer is 2 h, inventory costs per unit

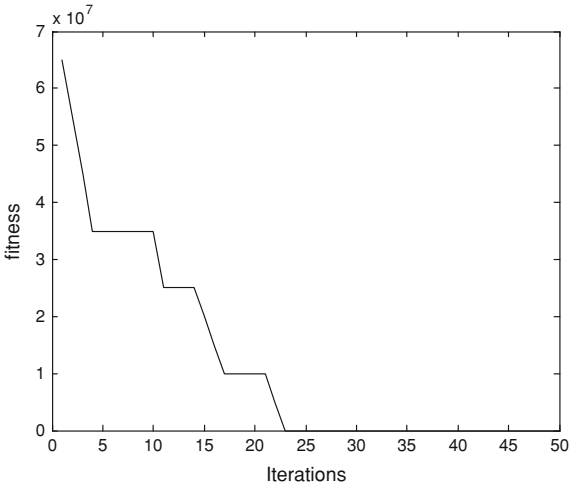


Fig. 10.1 Curve of fitness

product is 0.5 yuan, back order cost of one unit product per hour is 0.5 yuan, lead time order of raw materials used by manufacturer is 30 h; transport cost per unit product by the first transporter is 22 yuan, 7 vehicles can be used, the largest quantity of per vehicle is 30, the largest quantity is 210, the time used to complete one transportation is 6 h, fixed costs of the vehicles is 300 yuan; transport cost per unit product by the second transporter is 20 yuan, seven vehicles can be used, the largest quantity of per vehicle is 30, the largest quantity is 210, the time used to complete one transportation is 7 h, fixed costs of the vehicles is 400 yuan.

We coded the improved PSO algorithm in MATLAB language, and simulated it with a 1.73 GHz Intel Pentium PC. The parameters used during the experimental process in Eqs. (10.8) and (10.9) are defined as follows. $c_1 = c_1 = 1.0$; ω (the inertia weight) is decreased linearly from 0.9 to 0.1 for a run. And the population size of the swarm is set to 20. Every instance is executed by PSO for 10 runs. Most of the instances only need 25 iterations in each run, but a few of them need 40 iterations in each run (Fig. 10.1). We can see that during the process of manufacturer completing the order, the total cost of production and delivery is 50423 yuan $z_{11} = z_{12} = z_{13} = z_{14} = 14$, $z_{15} = z_{16} = z_{17} = 13$, $z_{2i} = 15$, $i = 1, \ldots, 7$.

10.6 Conclusions

Owing to the idea of the order fulfillment optimization model in the document 8, the article has founded the decision-making model of manufacturers' production and delivery and has given out the algorithm for finding the solution based on PSO. Bat, the optimization decision-making model applied to variety of products, is required to do further study.

References

1. Sabri EH, Beamon BM (2000) A multi-objective approach to simultaneous strategic and operational planning in supply chain design. Omega 28:581–598
2. Marsh J, Hole D (1996) Re-designing a complex multi-customer supply chain. Logist Inf Manag 9:223–249
3. Arntzen C, Brown G, Harrison TP (1995) Global supply chain management at digital equipment corporation. Interfaces 25:56–89
4. Bowersox DJ, Stank TP, Daugherty PJ (1999) Lean launch managing product introduction risk through response-based logistics. J Prod Innov Manag 16:557–568
5. Weeks D, Crawford FA (1994) Efficient consumer response: a mandate for food manufacturers? Food Process 55:312–354
6. Kurnia S, Johnston RB (2001) Adoption of efficient consumer response: the issue of mutuality. Supply Chain Manag 6:230–241
7. Lohtia R, Xie FT, Subramaniam R (2004) Efficient consumer response in Japan: industry concerns, current status, benefits, and barriers to implementation. J Bus Res 57:306–311
8. Shu L, Yang L (2009) Order fulfillment optimization model of manufactures based on time competition. In: International conference of management science and information system, vol 3, pp 393–396
9. Kennedy J, Eberhart RC (1995) Particle swarm optimization. In: Proceedings of IEEE international conference on neutral networks, Australia, vol 4, pp 1942–1948
10. Eberhart RC, Kennedy J (2001) Swarm intelligence. Morgan Kaufmann, San Diego, ISBN 1-55860-595-9
11. Yoshida H, Kawata K, Fukyyama Y et al (1999) A particle swarm optimization for reactive power and voltage control considering voltage stability. In: Torres GL, Alves da Silva AP (eds) ISAP'99, Rio de Janeiro, brazil, pp 117–121
12. Gao HB, Zhou C, Gao L (2005) Generalized model of particle swarm optimization. Chin J Comput 28:1980–1987
13. Shayeghi H, Jalili A, Shayanfar HA (2008) Multi-stage fuzzy load frequency control using PSO. Energy Convers Manag 49:2570–2580
14. Clerc M, Kennedy J (2002) The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans Evolut Comput 6:58–73
15. Kuo IH, Horng SJ, Kao TW, Lin TL, Lee CL, Terano T et al (2009) An efficient flow-shop scheduling algorithm based on a hybrid particle swarm optimization model. Expert Syst Appl 36:7027–7032
16. Xia W, Wu Z (2005) A hybrid particle swarm optimization approach for the jobshop scheduling problem. Int J Adv Manuf Technol 29:360–366

Chapter 11
MPPT Algorithm Under Partial Shading Conditions

Guihua Liu, Panbao Wang, Wei Wang and Qi Wang

Abstract This paper proposes a novel maximum power point tracking (MPPT) algorithm based on analyzing the output characteristics of PV array under uniform irradiance and partial shading conditions (PSC). The algorithm adds partial shadow judging conditions in variable step perturb and observes (P&O) method. The system runs the variable step P&O method to realize MPPT when photovoltaic array is under uniform irradiance. Under PSC, the control method can eliminate the interference of local maximum power point (MPP) to make the PV array running at global MPP. Experimental results verify that the proposed algorithm can automatically track the global MPP under uniform irradiance and PSC.

Keywords MPPT algorithm · Output characteristic · Photovoltaic array · PSC

11.1 Introduction

Photovoltaic (PV) power generation system occupies an important position in renewable energy for its flexible configuration, easy installation and no pollution. Except for establishing large-scale PV station, building integrated PV and household PV power generation are the important forms in future city using solar energy. In order to improve the photoelectric conversion efficiency, maximum power point tracking (MPPT) control technology is used in PV power generation system. Under uniform irradiance, the output characteristic curves of PV array

G. Liu (✉) · P. Wang · W. Wang · Q. Wang
School of Electrical Engineering and Automation,
Harbin Institute of Technology, No.92, West Da-Zhi Street,
Harbin 150001, China
e-mail: liuguihua@hit.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_11,

have the similar shape such as single PV cell, and power-voltage curve has a unique maximum power point (MPP). As PV power generation systems become popular applications in urban buildings and family roofs, the environmental conditions of PV panels become more complex, and the surrounding tall buildings, trees may produce shadow on PV panels in some time during the day. Under partial shading conditions (PSC), the current-voltage and power-voltage characteristic curves will appear multiple steps and peaks, and using MPPT control technology to track real MPP becomes a difficult task.

The traditional MPPT methods include constant voltage (CV), perturb and observe (P&O), incremental conductance (IncCond), and so on. When the PV array is under PSC, traditional MPPT control methods can obtain the local MPP, but may not be able to extract global MPP and cause power loss. Because partial shading conditions are quite common, exploring effective MPPT method which suits uniform irradiance and PSC becomes a pressing need to make full use of PV cell [1–6]. Some researchers have proposed some methods to solve it [7–12], and the solutions are mainly divided into two kinds: hardware method and software method. The hardware solution usually adopts AC or DC module structure and each panel has independent MPPT controller, its shortcomings are complex structure and high cost. The software solution generally adopts software control algorithm to track global MPP, and does not change the original structure of PV power generation system. Combined with variable step P&O method and output voltage current information of PV array, this paper proposes a novel MPPT method to track the global MPP under uniform irradiance and PSC.

11.2 Output Characteristic of PV Array Under PSC

Single PV panel has limited power and working voltage, so PV panels are usually connected in series–parallel to ensure the desired voltage, current, and power. The current–voltage equation of a PV panel can be written as

$$I = I_{\mathrm{ph}} - I_{\mathrm{s}}\left\{\exp\left[\frac{q}{AKT}(V + IR_{\mathrm{s}})\right] - 1\right\} - \frac{V + IR_{\mathrm{s}}}{R_{\mathrm{sh}}} \tag{11.1}$$

where I_{ph} is the photocurrent, I_{s} is the diode saturation current, q is the electronic charge (1.6×10^{-19} C), A is the diode ideality factor, K is the Boltzmann's constant (1.38×10^{-23} J/K), T is the junction temperature, R_{s} is the series resistance, and R_{sh} is the shunt resistance.

The value of shunt resistance (R_{sh}) is large, so the last term of Eq. 11.1 can be neglected. At the same time, the short-circuit current (I_{sc}) is approximately equal to photocurrent (I_{ph}). So Eq. 11.1 can be expressed as

$$I = I_{\mathrm{sc}} - I_{\mathrm{s}}\left\{\exp\left[\frac{q}{AKT}(V + IR_{\mathrm{s}})\right] - 1\right\} \tag{11.2}$$

When the PV panel is in the open-circuit condition, its output current is equal to zero, and diode saturation current can be derived as

$$I_s = \frac{I_{sc}}{\exp(\frac{qV_{oc}}{AKT}) - 1} \tag{11.3}$$

where V_{oc} is the open-circuit voltage.

Substitute Eq. 11.3 into Eq. 11.2 and ignore the effect of '1' in Eq. 11.2, then current-voltage approximate expression can be written as follows

$$I = I_{sc}\left\{1 - \exp\left[\frac{q}{AKT}(V + IR_s - V_{oc})\right]\right\} \tag{11.4}$$

Equation 11.4 is implicit function of voltage and current. To facilitate analysis, it can be rewritten as

$$V = \frac{AKT}{q}\ln\left(1 + \frac{I}{I_{sc}}\right) - IR_s + V_{oc} \tag{11.5}$$

Then the output power can be expressed as

$$P = VI = \frac{AKTI}{q}\ln\left(1 + \frac{I}{I_{sc}}\right) - I^2R_s + IV_{oc} \tag{11.6}$$

According to Eqs. 11.5 and 11.6, the output characteristic curves (I–V and P–V) can be easily obtained. When the PV panel is connected in series–parallel to constitute PV array, the total output voltage, output current, open-circuit voltage, short-circuit current, and series resistance of PV array can be expressed as

$$V_A = N_S V,\ I_A = N_P I,\ V_{ocA} = N_S V_{oc},\ I_{scA} = N_P I_{sc},\ R_{sA} = (N_S/N_P)R_s \tag{11.7}$$

where N_S is series panel number and N_P is parallel panel number.

When the PV panel works in series under uniform irradiance, the load current is equal to output current of PV array. If one or more pieces of PV panels appear shadow, the output current of shadowed panels become lower and easily occur hot-spot phenomenon. Hot-spot can accelerate aging or even damage PV panel, so most PV panel is integrated with a bypass diode. When the load current is lower than the maximum output current of the shadowed PV panel, the bypass diode does not work and the output voltage of PV array can be written as

$$V_A = \sum_{i=1}^{Ns} V_i = \sum_{i=1}^{Ns}\left[\frac{AKT}{q}\ln\left(1 + \frac{I_i}{\lambda I_{sc}}\right) - I_iR_s + V_{oc}\right] \tag{11.8}$$

where V_i, I_i, I_{sc}, V_{oc}, and R_s are the output voltage, output current, short-circuit current, open-circuit voltage, and series resistance of each PV panel in series, λ represents irradiance level and it is equal to the ratio of actual irradiance to standard irradiance.

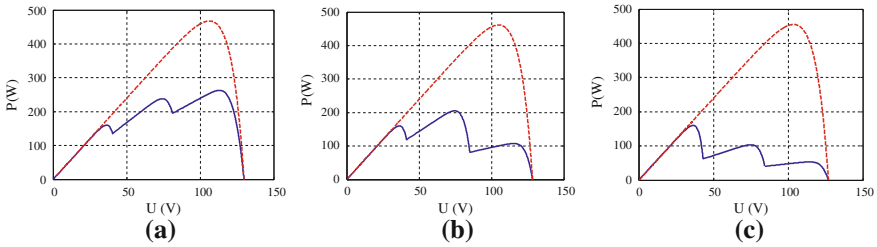


Fig. 11.1 P-V curves of PV panels under uniform irradiance and PSC. **a** $\lambda = 1, 0.7, 0.5$; **b** $\lambda = 1, 0.6, 0.2$; **c** $\lambda = 1, 0.3, 0.1$ (The *dotted line* and *solid line* are corresponding to uniform irradiance and PSC, respectively)

If load current is larger than the maximum output current of the shadowed PV panel, the parallel bypass diode will be conductive. Assuming the turn-on voltage of bypass diode is zero, the output voltage, open-circuit voltage and series equivalent resistance of PV array can be expressed as follows:

$$V_{\mathrm{A}} = (N_{\mathrm{S}} - N_{\mathrm{D}})V,\ V_{\mathrm{ocA}} = (N_{\mathrm{S}} - N_{\mathrm{D}})V_{\mathrm{oc}},\ R_{\mathrm{sA}} = (N_{\mathrm{S}} - N_{\mathrm{D}})R_{\mathrm{s}} \tag{11.9}$$

where N_{D} is the shadowed number of PV panel.

When the PV array is series–parallel structure, blocking diode is usually added in each series branch to prevent current circumfluence. Under PSC, output current of the shadowed series branch becomes lower, but the total output voltage of PV array is equal to highest voltage of the series branch.

According to the above formulas, the output characteristic curves can be expressed under uniform irradiance and PSC. The *P–V* curves of three PV panels (simulation parameters come from the PV panel BP3160) in series are shown in Fig. 11.1. It can be seen from the figures that the global MPP locates in different position when the PV panel accepts different irradiance level (λ). In order to make full use of PV panels, effective MPPT algorithm is needed to research.

11.3 MPPT Algorithm Under PSC

The main problem to realize MPPT under PSC is ruling out the interference of local MPP and automatically tracking the global MPP. The MPPT algorithm in this paper adds partial shadow judging rules in variable step P&O method. If the shadow appears in PV array, its output current and voltage will change and Eq. 11.10 shows the judgment criterion. The current threshold (I_{TH}) and voltage threshold (V_{TH}) are set according to the structure of PV array.

$$\begin{cases} \Delta I = I(n) - I(n-1) < I_{\mathrm{TH}} \\ \Delta V = V(n) - V(n-1) < V_{\mathrm{TH}} \end{cases} \tag{11.10}$$

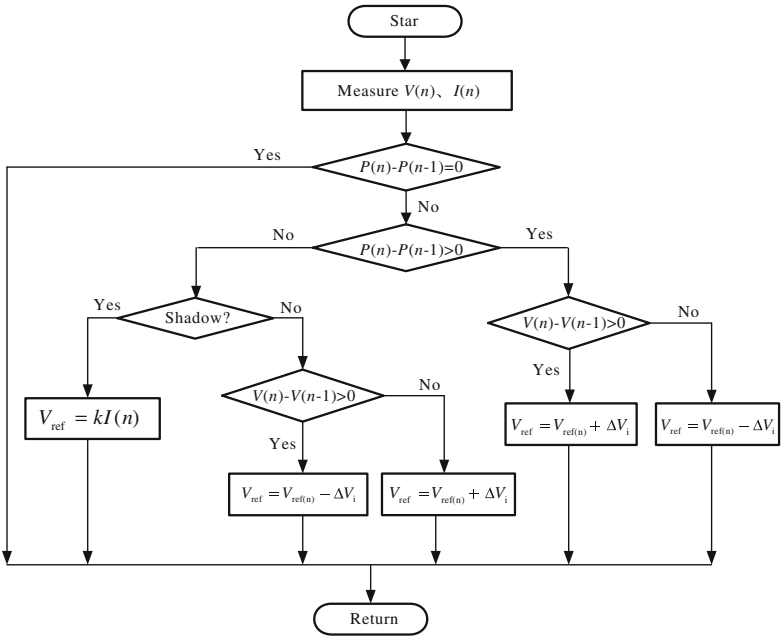


Fig. 11.2 MPPT algorithm flowchart under uniform irradiance and PSC

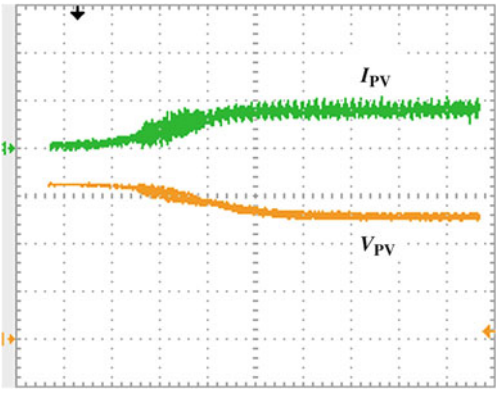


Fig. 11.3 Voltage and current waveforms during starting period (voltage: 40 V/div; current: 5 A/div; time: 500 ms/div)

Once the photovoltaic array appears shadow, the reference value of output voltage (V_{ref}) can be decided by the linear function as shown in Eq. 11.11.

$$V_{\text{ref}} = kI(n) \tag{11.11}$$

where $k = V_{\text{ocA}}/I_{\text{scA}}$.

In order to obtain compromise of speed and accuracy, the variable step P&O method is used when the PV panel is under uniform irradiance. During the starting period, large step is used to reach the nearly MPP quickly, and then small step is used to minimize the oscillation to get relatively accurate MPP. Figure 11.2 is the

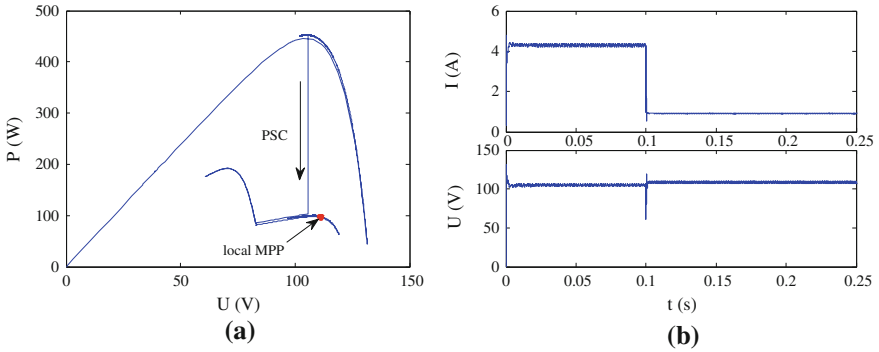


Fig. 11.4 P-V, voltage and current curves with variable step P&O method. **a** *P-V* curve under PSC; **b** current and voltage curves

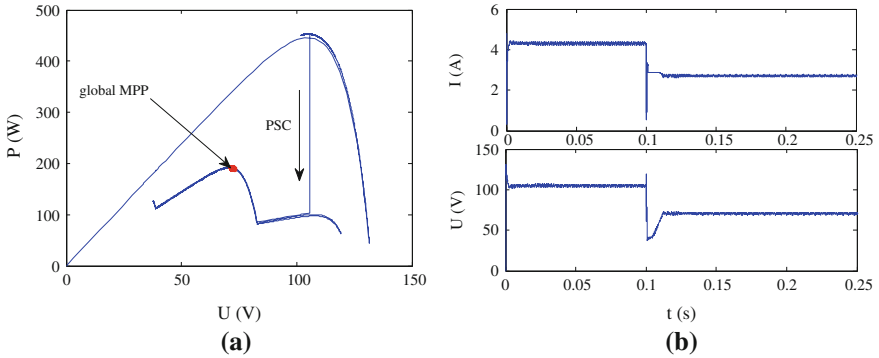


Fig. 11.5 P-V, voltage and current curves with the designed MPPT algorithm. **a** *P-V* curve under PSC; **b** current and voltage curves

MPPT algorithm flowchart in steady state. In the flowchart, the shadow judgment is according to the formula (11.10), and the step (ΔV_i) is much smaller than the step during starting period.

11.4 Simulation and Experimental Results

In order to validate the effectiveness of the proposed MPPT algorithm under PSC, three PV panels (BP3160) in series are used. Figure 11.3 shows the voltage and current waveforms during starting period. It can be seen from the figure that quick system tracing speed can be obtained with large step.

Figures 11.4 and 11.5 show the simulation results with variable step P&O method and the designed MPPT algorithm when the PV panels are under PSC (irradiance level $\lambda = 1, 0.6, 0.2$). The simulation results verified that only variable

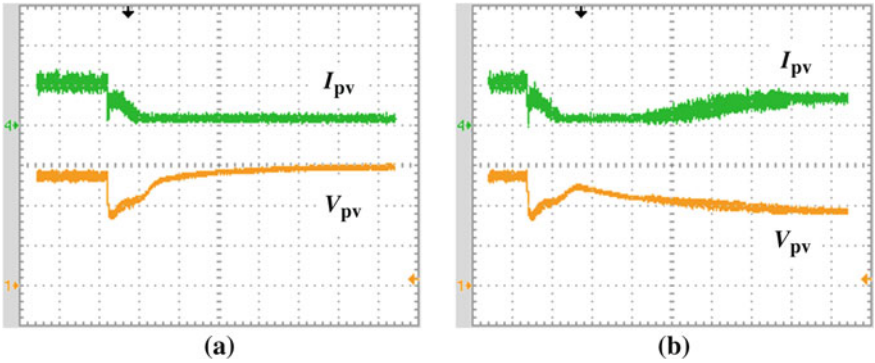


Fig. 11.6 Voltage and current waveforms with variable step P&O method and the designed MPPT algorithm. **a** With variable step P&O method; **b** with designed MPPT algorithm (voltage: 40 V/div; current: 5 A/div; time: 500 ms/div)

step P&O method reaches local MPP and the designed MPPT algorithm can track global MPP under PSC.

Figure 11.6 is the experimental results with variable step P&O method and the designed MPPT algorithm when the PV panels under PSC. The experimental results verify that the variable step P&O method fails to track the global MPP and causes power loss. The designed MPPT algorithm can eliminate the interference of local MPP to make the PV array running at real global MPP.

11.5 Conclusion

In order to improve the photoelectric conversion efficiency of PV array in complex external conditions, this paper proposes a novel MPPT algorithm based on analyzing the output characteristics of PV array under uniform irradiance and PSC. When the PV array accepts the uniform irradiance, the variable step P&O method is used to realize MPPT. If partial shadow appears in PV array, the novel algorithm can eliminate the interference of local MPP to make the PV array run at global MPP. The algorithm uses software to realize shadow judging and control, and does not need to add hardware circuits. The simulation and experimental results verify that the proposed algorithm has good control effect under PSC.

Acknowledgments This work was supported by China Postdoctoral Science Foundation funded project (20100471036), the National Natural Science Funds (51077017), and Hei Long Jiang Postdoctoral Foundation.

References

1. Miyatake M, Inada T, Hiratsuka I et al (2004) Control characteristics of a fibonacci-search-based maximum power point tracker when a photovoltaic array is partially shaded. In: The 4th international power electronics and motion control conference, vol 2. Xi'an, China, pp 816–821
2. Kobayashi K, Takano I, Sawada Y (2003) A study on a two stage maximum power point tracking control of a photovoltaic system under partially shaded insolation conditions. In: IEEE power engineering society general meeting, vol 4. Toronto, ON, Canada, pp 2612–2617
3. Al-Diab A, Sourkounis C (2010) Multi-tracking single-fed PV inverter. In: IEEE mediterranean electrotechnical conference, Valletta, Malta, pp 1117–1122
4. Patel H, Agarwal V (2008) Maximum power point tracking scheme for PV systems operating under partially shaded conditions. In: IEEE transactions on industrial electronics, vol 55. Seoul, Korea, pp 1689–1698
5. Cheng Z, Zhou H, Yang H (2010) Research on MPPT control of PV system based on PSO algorithm. In: Chinese control and decision conference. Xuzhou, China pp 887–892
6. Mishima T, Ohnishi T (2002) Power compensation system for partially shaded PV array using electric double layer capacitors. In: IEEE industrial electronics society conference, vol 4. Sevilla, Spain, 3262–3267
7. Liu B, Liang C, Duan S (2008) Design considerations and topology selection for dc-module-based building integrated photovoltaic system. In: IEEE conference on industrial electronics and applications, Singapore, pp 1066–1070
8. Karatepe E, Hiyama T, Boztepe M et al (2007) Power controller design for photovoltaic generation system under partially shaded insolation conditions. In: International conference on intelligent systems applications to power systems, pp 1–6
9. Dzung N, Lehman B (2008) A reconfigurable solar photovoltaic array under shadow conditions. In: Twenty-third annual IEEE applied power electronics conference and exposition, Austin, TX, USA, pp 980–986
10. Khajehoddin SA, Bakhshai A, Jain P (2008) A Novel topology and control strategy for maximum power point trackers and multi-string grid connected PV inverters. In: Twenty-third annual IEEE applied power electronics conference and exposition, Austin, TX, USA, pp 173–178
11. Ji Y-H, Jung D-Y, Won C-Y et al (2009) Maximum power point tracking method for PV array under partially shaded condition. In: IEEE energy conversion congress and exposition, San Jose, CA, USA, pp 307–312
12. Patel H, Agarwal V (2008) Maximum power point tracking scheme for PV systems operating under partially shaded conditions. In: IEEE transactions on industrial electronics, vol 55, Seoul, Korea, pp 1689–1698

Chapter 12
Echo State Network-Based Internal Model Control for Pneumatic Muscle System

Jun Wu, Yongji Wang, Jian Huang, Hanying Zhou and Hong Cai

Abstract Pneumatic muscle (PM) has strong time varying characteristic. The complex nonlinear dynamics of PM system poses problems in achieving accurate modeling and control. To solve these challenges, we propose an echo state network (ESN)-based internal model control (IMC) for PM system in this paper. Here, ESN is employed for identifying the plant model and constructing controller. Recursive least squares (RLS) is used for the online update of ESN. The ESN-based IMC fully embodies the virtues of ESN and IMC. It can build high accurate plant model without detailed model information, as well as attain strong robustness by online self-tuning of controller and internal model. Experiment demonstrates the effectiveness of the proposed control algorithm. The results show that ESNBIMC achieves satisfactory tracking performance.

Keywords Echo state network · Internal model control · Pneumatic muscle

12.1 Introduction

As an interesting actuator, pneumatic muscle (PM) is similar to human skeletal muscle in size, weight, and high power/weight output. Moreover, it can generate inherently compliant force and thus has excellent safety potential. These attractive features make that PM is feasible for using in rehabilitation engineering, which requires great compliance and high safety for patients [1, 2]. However,

J. Wu · Y. Wang (✉) · J. Huang · H. Zhou · H. Cai
Key Laboratory of Image Processing and Intelligent Control,
Department of Control Science and Engineering,
Huazhong University of Science and Technology, 430074 Wuhan, China
e-mail: wangyjch@mail.hust.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_12,

the PM application also faces some challenges. Comparing to electric motor, PM has slower response time, time varying parameters depending on the load position and speed. The complex nonlinear dynamics of the PM poses problems in achieving accurate modeling and control. This has restricted its widespread use in the past.

Internal model control (IMC) is proposed in Ref. [3]. Due to its simple structure, easy tuning property, strong robustness, and the ability of eliminating unpredictable disturbance, IMC could be used in linear systems successfully. However, it is difficult to obtain satisfactory control performance when IMC is directly introduced in nonlinear systems due to the inherent complexity of nonlinear systems. With the development of artificial neural networks (ANNs) for nonlinear modeling, the idea of using ANN for nonlinear IMC has been considered by Bhat and McAvoy [4]. ANNs have an inherent ability to approximate an arbitrary nonlinear function, and become an attractive way to construct complex nonlinear system. Thus, ANNs can be used for building the plant model of complex system and controller in nonlinear IMC.

Recently, the echo state network (ESN) has been introduced as a novel approach by Herbert Jaeger for time series modeling and nonlinear system identification [5, 6]. The notable character of ESN is that its internal layer is composed of a large number of neurons and these neurons are sparsely connected to each other. This layer can map inputs into high-dimensional space and reserve information of the past. It makes the ESN not only hold self-organization, self-learning, and adaptive ability, but also has short-term memory. The recursive least square (RLS) algorithm can be applied to online training of the ESN without affecting the previous weight structure, which is very suitable for real-time control problems. Nevertheless, to the best of our knowledge there are few applications of utilizing an ESN-based method in the control filed.

In this paper, we propose ESN based IMC (ESNBIMC) for the control of PM system. It fully embodies the virtues of ESN and IMC. Here, ESN is employed for identifying the plant model and constructing controller. This control algorithm can build high accurate plant model without detail model information, as well as attain strong robustness by online self-tuning of internal model.

This paper is organized as follows. Section 12.2 describes the PM control system. Section 12.3 presents the ESN. In Sect. 12.4, we take advantage of both ESN and IMC to propose ESNBIMC for the control of PM system. In Sect. 12.5, the experiment of the trajectory tracking control of PM system is undertaken by using ESNBIMC. The results show that ESNBIMC is effective and can achieve satisfactory tracking performance. Finally, the discussion and conclusion are given.

12.2 PM Control System

The experimental schematic diagram of PM platform is shown in Fig. 12.1a. As the rubber bladder expands due to an increase in pressure, the diameter of the combined sheath and bladder assembly easily changes in the radial direction and the

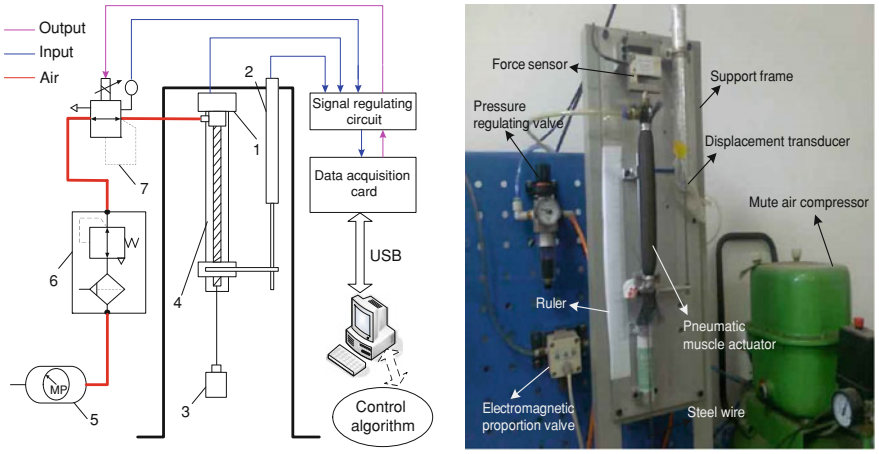


Fig. 12.1 **a** The experimental schematic diagram of PM platform (*1*-force sensor, *2*-displacement transducer, *3*-load/weights, *4*-pneumatic muscle, *5*-mute air compressor, *6*-pressure regulating valve, *7*-electromagnetic proportion valve). **b** The prototype of PM platform

muscle shortens in the axial direction. Thus, the force exerted on the environment occurs in the axial direction. The change of PM length is measured by a displacement transducer which is attached to the moving lower end of the PM. In addition, there is also a pressure sensor integrated in the electric proportional valve to provide the pressure feedback. Data are collected by a multichannel real-time data acquisition card. The prototype of the experimental platform is shown in Fig. 12.1b.

12.3 Echo State Network

As both work using the mechanism of reservoir computing, the ESN proposed by Jaeger [5] shares some similarities with the Liquid State Machine (LSM) [7]. The ESN structure is shown in Fig. 12.2. An ESN is a recurrent discrete-time neural network with K input units, N internal (reservoir) units, and L output units. The activations of input, internal, and output units at time step k are denoted by $u(k) = (u_1(k), \ldots, u_K(k))^T, x(k) = (x_1(k), \ldots, x_N(k))^T$ and $y(k) = (y_1(k), \ldots, y_L(k))^T$, respectively. The real-valued input, internal, and output weights are collected in a $N \times K$ matrix W^{in}, a $N \times N$ matrix W, a $L \times (K + N + L)$ matrix W^{out}, respectively. The output units may optionally project back to internal units with connections whose weights are collected in a $N \times L$ matrix W^{back}.

The next activation state $x(k + 1)$ of internal units is updated according to

$$x(k+1) = f(W^{\text{in}}u(k+1) + Wx(k) + W^{\text{back}}y(k)) \tag{12.1}$$

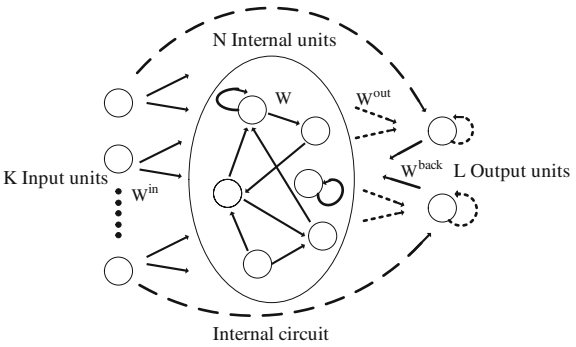


Fig. 12.2 The ESN structure

where $f = (f_1, \ldots, f_N)$ are the internal units' activation functions.

The next output $y(k+1)$ of ESN is computed according to

$$y(k+1) = f^{\text{out}}(W^{\text{out}}(u(k+1), x(k+1), y(k))) \tag{12.2}$$

where $f^{\text{out}} = (f_1^{\text{out}}, \ldots, f_L^{\text{out}})$ are the output units' activation functions. $(u(k+1), x(k+1), y(k))$ denotes the concatenation vector made from input, internal, and output activation vectors. In our study, we choose sigmoid neurons to form the hidden layer, and linear units to form the output layer, and assume the ESN without feedback connections from the output to the reservoir as used in [8, 9]. This common-used ESN structure not only keeps the nonlinear approximation capability and the echo state property, but also to some extent decreases computational complexity for real-time control.

The most important property of ESN is the echo state property. For a practical use, we simply apply the conclusion obtained in [5]. That is, it is sufficient to ensure this property by scaling the spectral radius of internal weight matrix to $|\lambda_{\max}| < 1$ for an ESN with normal sigmoid units in the internal layer.

12.4 ESNBIMC for the Control of PM System

In the ESNBIMC, we employ the framework of IMC, while the identification model of plant and internal model controller are structured by ESN.

12.4.1 Internal Model Control

IMC is a mathematical model of plant/process-based control strategy for the design of controller. The classical structure of IMC is demonstrated in Fig. 12.3. In the diagram, R is the reference input of the control system, Y is the system

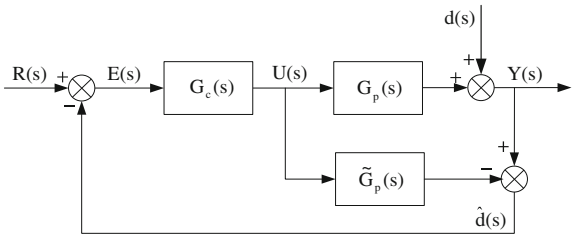


Fig. 12.3 The structure of IMC

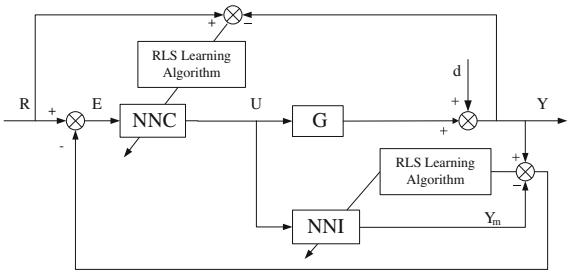


Fig. 12.4 The structure of ESNBIMC

output, U is the controller output, d is the unknown disturbance, G_c is the controller, G_p is the plant, and $\tilde{G}_p$ is plant model, which is used as the internal model.

According to the IMC architecture, the output of the closed loop IMC system is:

$$Y(s) = \frac{G_c(s)G_p(s)R(s) + \left[1 - G_c(s)\tilde{G}_p(s)\right]d(s)}{1 + \left[G_p(s) - \tilde{G}_p(s)\right]G_c(s)} \tag{12.3}$$

From (12.3), we can see that if there exists the perfect match between the plant model and plant, that is $\tilde{G}_p(s) = G_p(s)$, perfect tracking of R and Y is attained, as well as disturbance rejection is achieved by selecting $G_c(s) = \tilde{G}_p(s)^{-1}$. Note that, theoretically, even if $G_p(s) \neq \tilde{G}_p(s)$, perfect disturbance rejection can still be realized when $G_c(s) = \tilde{G}_p(s)^{-1}$.

12.4.2 ESN-Based IMC

The proposed ESNBIMC is shown in Fig. 12.4, the plant model and controller are constructed by ESN in this paper. Here, the plant model can be termed as the neural network identifier (NNI) of plant, which can be approximated according to the input and output data of plant by ESN. As the Fig. 12.4 shows, Y is the plant output, and Y_m is the NNI output of the plant. Through the error between the real output Y and the computed output Y_m, the weight of NNI is updated online. It makes the NNI can accurately represent the dynamic characteristic of the plant.

The internal model controller is the inverse of the plant, which can be termed as the neural network controller (NNC). It can also be approximated according to the input and output data of plant by ESN. Unlike the NNI, the output data of plant is considered as the NNC input, the input data of plant is considered as the NNC output for NNC training. Through the error between the desired output R and the real output Y, the weight of NNC is also adjusted online.

After the training of the NNI and NNC, the control strategy of ESNBIMC can be implemented. Due to the online adjustment of NNI and NNC, the two ESNs can optimize the weights to achieve higher accuracy. Besides, if there are some uncertainties or disturbances making the system change, the online adjustment can also make the control system have strong disturbance rejection.

12.4.3 Online Training Algorithm

The RLS algorithm has been extensively used in adaptive identification, prediction, filtering, and many other fields. It can be applied to the online training of the ESN without affecting the previous weight structure, which is very much suitable for real-time control problems. In our study, ESN is conjoined with the RLS algorithm for online self-adaption after the initial steps.

If $W^{\text{out}}(k)$ is the output weight of ESN at the kth step, while the desired output of ESN is $y_d(k)$. $e(i|k) = y_d(i) - W^{\text{out}}(k)x(i|k)$ represents the i-step-ahead prediction error based on the kth step. The goal of RLS algorithm is to minimize the following objective function at the kth step.

$$J = \sum_{i=1}^{k} \gamma^{k-i} e^2(i|k) \tag{12.4}$$

where γ is forgetting parameter, it is usually set to the value smaller or equal to 1.0.

According to the principle of least mean square error, RLS algorithm can be expressed as follows.

$$\lambda(k) = \frac{P(k-1)v(k)}{v^T(k)P(k-1)v(k) + \gamma} \tag{12.5}$$

$$P(k) = \gamma^{-1}\big(P(k-1) - \lambda(k)v^T(k)P(k-1)\big) \tag{12.6}$$

$$e(k) = y_d(k) - \hat{y}(k) = y_d(k) - W^{\text{out}}(k-1)v(k) \tag{12.7}$$

$$W^{\text{out}}(k) = W^{\text{out}}(k-1) + \lambda(k)e(k) \tag{12.8}$$

where λ stands for the innovation vector calculated in every time step, W^{out} is the vector of output weight, v is the vector of state activities of ESN, y_d is the vector of target values, and $\hat{y}$ is the output vector which is calculated by the input vector and

the ESN. P is the inverse correlation matrix initialized with large diagonal elements and updated in every time step. It is initialized as:

$$P(0) = \delta^{-1} I \quad (\delta > 0) \tag{12.9}$$

where δ is a small positive number. I is used to denote a unit matrix.

In addition, the reset of the covariance matrix is introduced to avoid instability problem caused by long-term non-static signal [10]. If the track of the matrix $P(k)$ is lower than a small positive constant, then $P(k)$ is reset.

The detail algorithm used in ESNBIMC can be described as follows.

Step 1: Initialize the weights of ESNs and the error covariance matrix$P$;

Step 2: Attain the vector v of state activities of ESN, and then calculate the innovation vector λ;

Step 3: Calculate the error between the output of the plant and the output of ESN for NNI, while calculate the error between the reference input and the output of plant for NNC;

Step 4: Update the output weights of the NNI or the NNC by using RLS;

Step 5: Update the error covariance matrix P;

Step 6: The controller produces the new control value to drive the plant;

Step 7: Collect the new sample values for the next control;

Step 8: If the control process is not completed, then go to Step 2, otherwise go to Step 9;

Step 9: Quit the control process successfully.

12.5 Experiment

The experiment is to demonstrate the validity of the proposed ESNBIMC. First, the experiment is set up according to Fig. 12.1. In our experiment, the initial length of PM is 200 mm. The initial diameter is 12.26 mm, the initial angle of mesh grid is 22°, and the thickness of rubber sleeve is 1.64 mm. The main experimental apparatus includes electromagnetic proportional valve, pressure regulating valve, displacement transducer, acquisition card, and so on. The physical parameters of experiment equipments are summarized and listed in Table 12.1. The software is developed by VC.

In order to compare the control effectiveness and attain the training data for identifying NNI and constructing NNC in ESNBIMC, both the performance of the PID and the ESNBIMC controllers are evaluated using sine and square waves as the desired trajectory for PM control. The parameters of PID controller are: $k_p = 0.012, k_i = 0.0002$ and $k_d = 0.0002$. The collected data by using PID controller is used for the training of NNI and NNC in ESNBIMC. The number of input units, internal units, and output units of NNI are 1, 20, and 1, respectively. The NNC has the same network structure with NNI.

Table 12.1 The main performance parameters of the equipments

Name	Model	Performance	Manufacturer
Mute air compressor	FB-0.017/7	Rating exhaust pressure: 0.7 Mpa	Shanghai Jiebao Co., Ltd
Electromagnetic proportion valve	ITV1030-211BS	Input: 0–5 V Output: 0.005–0.5 Mpa	SMC Co., Ltd
Pressure regulating valve	AW20-02BCG	Range: 0.05–0.85 Mpa	SMC Co., Ltd
Data acquisition card	USB-4716	8 DI, 8 DO, 16 AI, 2 AO Sampling rate: 200 kS/s	Advantech Co., Ltd
Displacement transducer	DA-75	Range: 0–150 mm Linearity: <0.1% (= 0.09%) Output: 0–5 V	Beijing King Sensor Technology Co., Ltd
Others	Joints connector, air pipe, etc.		SMC Co., Ltd

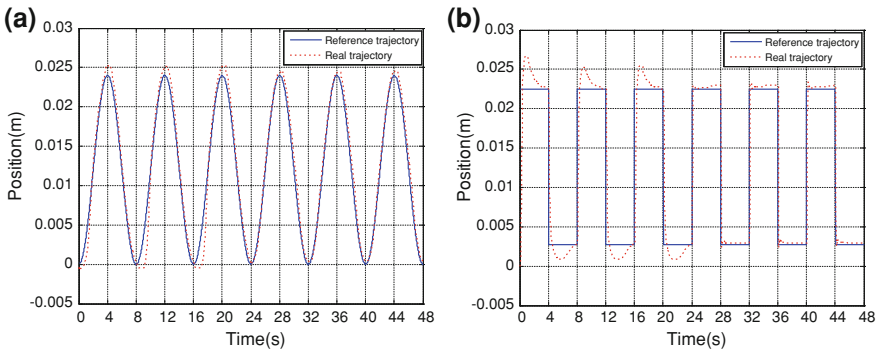


Fig. 12.5 **a** The sine tracking result of PM system by employing PID and ESNBIMC. **b** The square tracking result of PM system by employing PID and ESNBIMC

The tracking results of sine and square waves for PM system are shown in Fig. 12.5. At the beginning of the ESNBIMC control strategy, we use the PID control and make the ESN update the internal state without the update of the output weight from 0 to 24 s (the primitively three periods of tracking trajectory). Then the controller is switched to the ESNBIMC control with self-adaption. In the practical experiment, after several sampling control periods, the output of ESNBIMC can present a good performance. However, we still select the end of the third period of tracking trajectory as the switch time for the following reasons.

(1) To conveniently compare the performance of the PID control and the ESNBIMC control. In the primitive of three periods of tracking trajectory, we investigate the performance of PID control, while we investigate the performance of ESNBIMC control in the rest time;

(2) To provide enough time for updating the internal state of ESNs and to obtain enough input history for using short term memory capability of ESN.

From these results, it shows that the performance of the ESNBIMC is satisfactory.

12.6 Conclusions

In this paper, ESN-based IMC is proposed for the control of PM system. IMC has the advantages, such as simple structure, easy tuning property, strong robustness, the ability of eliminating unpredictable disturbance, and so on. The traditional framework of IMC is designed by employing the plant model. However, PM system is complex nonlinear and time-varying system. It is difficult to establish the accurate mathematic model of plant. Fortunately, the ESN has the echo state property and the strong ability of nonlinear approximation. The proposed ESN-BIMC algorithm is an optimized method which can fully embody the virtues of ESN and IMC. Based on the ability of ESN approximation, we do not need too much priori knowledge about the detailed model of the plant to identify the plant and construct the controller. Moreover, in ESNBIMC algorithm, the identified plant model and the constructed controller can be updated by the online learning of ESN for achieving accurate control.

The experimental results demonstrate that the effectiveness of the ESNBIMC algorithm is satisfactory. Besides, it has better performance than the PID controller.

Acknowledgments This work has been supported in part by Hi-tech R&D Program of China under Grant 2007AA04Z204 and Grant 2008AA04Z207, and in part by the Natural Science Foundation of China under Grant 60674105, 60975058, and 61075095. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Noritsugu T, Tanaka T (1997) Application of rubber artificial muscle manipulator as a rehabilitation robot. IEEE/ASME Trans Mech 2(4):259–267
2. Tsagarakis NG, Caldwell DG (2003) Development and control of a 'soft-actuated' exoskeleton for use in physiotherapy and training. Auton Robot 15(1):21–33
3. Garcla CE, Morarl M (1982) Internal model control: 1: a unifying review and some new results. Ind Eng Chem Pro Des Dev 21(2):308–323
4. Bhat N, McAvoy TJ (1990) Use of neural nets for dynamical modelling and control of chemical process system. Comput Chem Eng 14(4–5):573–582
5. Jaeger H (2001) The "echo state" approach to analysing and training recurrent neural networks. Technical report-GMD report, number 148

6. Jaeger H (2002) Short term memory in echo state networks. Technical report-GMD report, number 152
7. Maass W, Natschläger T, Markram H (2002) Real-time computing without stable states: a new framework for neural computing based on perturbations. Neural Comput 14(11): 2531–2560
8. Buehner M, Young P (2006) A tighter for the echo state property. IEEE Trans Neural Netw 17(3):820–824
9. Rodan A, Tiňo P (2011) Minimum complexity echo state network. IEEE Trans Neural Netw 22(1):131–144
10. Tham MT, Mansoori SN (1988) Covariance resetting in recursive least squares estimation. In: International conference on control, pp 128–133

Chapter 13
A Self-Optimizing Cumulative Inversion Based on Combination Weight Method for Particle Size Distribution

Jian Qiu and Peng Han

Abstract Using cumulative inversion to get particle size distribution based on dynamic light scattering is a very important and popular method. However, noise, dust and drifts can interfere with this information heavily. In this paper, after the introduction of fundamentals of dynamic light scattering and the method of conventional cumulative inversion, a self-optimizing cumulative inversion based on the minimum weighted mean combined with fitting standard deviation and fitting correlation coefficient is proposed. By contrasting it with the cumulative inversion of Brookhaven Instrument in simulation experiments according to four noise factors, the results show the validity and feasibility of the proposed approach.

Keywords Self-optimizing cumulative inversion · Weighted mean · Dynamic light scattering · Software simulation

13.1 Introduction

Dynamic light scattering (DLS) technology is one of the most important optical method in measuring particle size distribution. The main procedure of DLS is to detect the fluctuation of scattered light intensity from a fixed angle, which is produced by particles in solution randomly go inside or outside of the incident light bean [1]. After this fluctuating signal is processed by autocorrelation, the size distribution will be achieved by some inversion algorithms at last [2].

J. Qiu (✉) · P. Han
Department of Electronics, School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China
e-mail: jianchild@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_13,

Until now, there are several useful inversion algorithms about particle size distribution, such as cumulative inversion (CI) [3], Nonnegative Least Squares inversion (NNLS) [4], CONTIN inversion [5], and so on. As is well known, the process of such inversion consists in the Laplace transform equation, which can be severely influenced by the ill-condition arising from measurement noise dark count rates and baseline drifts. To solve this problem, there are more and more new approaches being developed. Respectively, Yu [6] improved CONTIN algorithm by Iterative procedure, Li [7] investigated the Gray-code genetic algorithm and Ren [8] applied real-coded hybrid genetic algorithm in inversing particle size distribution.

Within these inversions, CI is the most basic and popular inversion method, which is characterized by clearness in principle and simpleness in inverse process. Within the results of CI, the average size ($\bar{d}$) and the polydispersity (Poly) are applied as the primary results in scientific paper and report. Therefore, the further improvement of the precision and stability of CI is very necessary. In this paper, a self-optimizing cumulative inversion (SOCI) based on minimum weighted mean combined with fitting standard deviation and correlation coefficient is presented. Under the premise of rapid computation, it can highly improve the precision, stability and reliability of CI.

13.2 Dynamic Light Scattering Interpretation

Based on the principle of DLS, each scattered wave should arrive to the detector at different times according to the different positions of the particle respectively. Because of the interference effects of these waves, the output signal from detector is an average value with superimposed fluctuations as shown in Fig. 13.1. The decay times of this fluctuations are relative to the particles sizes. Small particles moving rapidly cause faster decaying fluctuations than large particles moving slowly. In order to obtain the decay times in the time domain, the correlator is the most efficient method. Then the scattered intensity autocorrelation function (ACF) $G^{(2)}(\tau)$ is always given by [9]

$$G^{(2)}(\tau) = B\left(1 + f^2 |e^{-\Gamma\tau}|^2\right) \tag{13.1}$$

where B is the baseline, f^2 is an instrumental constant of constraint signal noise ratio, Γ is the decay rate, τ is the delay time. Provided Γ has been obtained, the particle size can be solved by Eq. 13.2.

$$\Gamma \times d = \frac{16\pi kTn^2}{3\eta\lambda_0^2} \sin^2\left(\frac{\theta}{2}\right) \tag{13.2}$$

Here, particle size d is relative to the scattering angle θ, the wavelength of the laser light λ_0, and the index of refraction n of the suspending liquid. The other constants

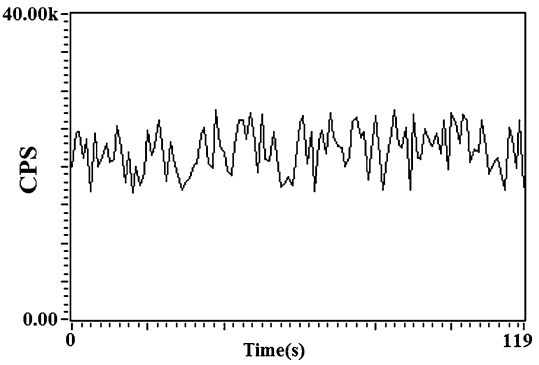


Fig. 13.1 Fluctuations about the average scattered intensity

are the Boltzmann's constant k, the temperature T in °K and the viscosity of the liquid η. This equation assumes the concentration of liquid is dilute enough so that particles are moving independently without multiple collisions.

In reality, particle size distribution cannot be monodispersed absolutely, the measured ACF should be formulated by Eq. 13.3, which means sample has a certain probability distribution [10].

$$|g^{(1)}(\tau)| = \int_0^\infty G(\Gamma)e^{-\Gamma\tau}d\Gamma \tag{13.3}$$

where $G(\Gamma)$ is the probability distribution function of decay rate. On the left-hand side of Eq. 13.3 is the normalization of the measured ACF (N-MACF), called Electric Field Autocorrelation Function (EFACF). On the right-hand side, it is the Laplace transform equation whose solution is nontrivial under the ill-conditioning combined with measurement noise, dark count rates and baseline drifts. Therefore, better flexibility and higher statistical stability of inversions are needed.

The solution of Eq. 13.3 by CI method is based on two relationships. One is the equivalent relationship between EFACF and moment generation function (MGF), another is between natural logarithmic EFACF and cumulative generating function (CGF). It is clear that the right-hand side of Eq. 13.3 is equivalent to MGF, and the natural logarithm of its left-hand side is equivalent to CGF. Through the expansion of natural logarithmic EFACF in a Taylor series, the average Γ is just the first cumulant $\kappa_1(\Gamma)$ and the variance is the second cumulant $\kappa_2(\Gamma)$. They are shown as Eq. 13.4 respectively.

$$\kappa_1(\Gamma) \equiv \overline{\Gamma}, \quad \kappa_2(\Gamma) = \overline{\Gamma^2} - \left(\overline{\Gamma}\right)^2 \tag{13.4}$$

And then, the $\bar{d}$ can be obtained by substituting $\bar{\Gamma}$ in Eq. 13.2. Due to the implication of the size distribution width in $\kappa_2(\Gamma)$. So the practical parameter of the size distribution width is defined as follows [10, 11]:

$$\text{Poly} = \kappa_2(\tau) \Big/ \overline{\Gamma}^2 \tag{13.5}$$

where Poly is a relative width of size distribution and without units. It is small (0.000–0.1) for monodisperse samples or narrow distributions, and larger for broader distributions.

13.3 Self-Optimizing Cumulative Inversion Proposal

Conventional cumulative inversion (CCI) method is based on the "Intercepted Channel" Method, which the inversed data from N-MACF would be discarded from the first data less than zero [12]. Thus the inversed data are compared with zero in turns from the short delay time to the long. Yet in practice, the results of CCI would vary so widely, particularly the ACF is interfered with serious noise or dust so that the curve becomes more dispersive. For example, as shown in Fig. 13.2, because of the large scale of baseline drifts and the discrete curve, it must result in large fluctuation of the cumulative results.

For this reason, a SOCI by obtaining the minimum weighted mean involved the fitting standard deviation (parameter σ) and the fitting correlation coefficient (parameter R) is proposed. The specific procedure of this method is divided into four steps.

First step, according to intercepted channel method, it achieves all the inversed data from the short τ to the long. Suppose the first data less than zero is the one of channel $(n + 1)$ in N-MACF, then the effective data is from channel 0 to n in N-MACF.

In the second step, it calculates each σ and R according to different starting channels. The specific procedure is as follows. The first time CI proceeds from channel 1 to n, and the first σ_1 and R_1 are calculated and recorded, which are both between cumulative fitting polynomial and the n effective data. Second, CI is from channel 2 to n, then σ_2 and R_2 are also calculated and recorded and so on. The last time, CI is from channel $(n - 10)$ to n, and σ_{n-10} and R_{n-10} are obtained at the end.

In the third step, the weighted mean (WM) array is obtained by Eq. 13.6, in which the parameters of σ and R have also become the array respectively after the second step.

$$WM_k = W_1 \times \sigma_k + W_2 \times (1 - R_k)/(W_1 + W_2) \tag{13.6}$$

where W_1 is the weight of σ and W_2 is the weight of $(1 - R)$. As well know, the stronger correlation is, the closer to "1" correlation coefficient is. Therefore, the expression of $(1 - R)$ in Eq. 13.6 means the minimum weighted value corresponds to the best fitting.

In the fourth step, the optimizing results could be achieved by finding the minimum value in the WM array. In other words, the sequence number of the minimum WM is an index to the optimizing results, and the SOCI based on the minimum weighted mean (SOCI-WM) is finished.

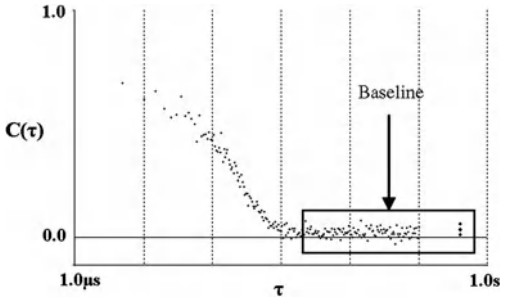


Fig. 13.2 Example of measured ACF curve with noise disturbance and baseline drift

13.4 Simulation and Analysis

For verifying the proposed method, a contrast experiment between the cumulative inversion of BI (BI-CI) and SOCI-WM in this paper was carried out by using 9KDLSW32 software of BI. This software involves a simulation source of DLS and BI-CI, and so on.

The simulation procedure consists of three parts. First, simulation data of the N-MACF is produced by the simulation source and saved as file. At the same time, the results of BI-CI are obtained. Second, the SOCI-WM realized by LabVIEW8.5 reads out simulation data from the file and outputs results. Third, these results can be analyzed in contrast. To study the accuracy and statistical stability of these inversions in different noise intensities, it can change the noise factor (NF) of simulation source whose maximum value is 1. Here, NF was set to 0, 0.2, 0.5 and 0.8 in the simulation.

The default experimental conditions are as follows: the incident wave length is 632.8 nm, solvent is water, temperature T of sample cell is 25°C and scattering angle θ is 90°. In addition to these parameters involved in CI, there are several settings on the simulation source. The simulation Γ is 1,900 rad/s, the amplitude of N-MACF is 0.7, the average photon count rates per second is 25 Kcps, the whole simulation experiment time is 5 min. To sum up these parameters, the ideal particle diameter is about 90.11 nm and its distribution is monodisperse.

The N-MACF curves of scattered intensity is partly shown in Fig. 13.3. Figure 13.3a–d respectively indicates the N-MACF curves according to different values of NF. Obviously, N-MACF curves became more and more dispersed following the increase of NF. The noise interference and baseline drifts could be seen clearly in Fig. 13.3c and d, so that the difficulty of CI can be predicted easily.

This kind of contrast analysis is carried out in three aspects. (1) It would analyze the influence on introducing R only in SOCI (SOCI-R) in contrast with BI-CI. (2) It should surely contrast the changes on just introducing σ in SOCI (SOCI-σ). (3) It is the comparative analysis of the different performance among BI-CI, SOCI-R, SOCI-σ and SOCI-WM. It is necessary to mention that the results of SOCI-R and SOCI-σ can be obtained easily by setting W_1 and W_2 to 0 respectively.

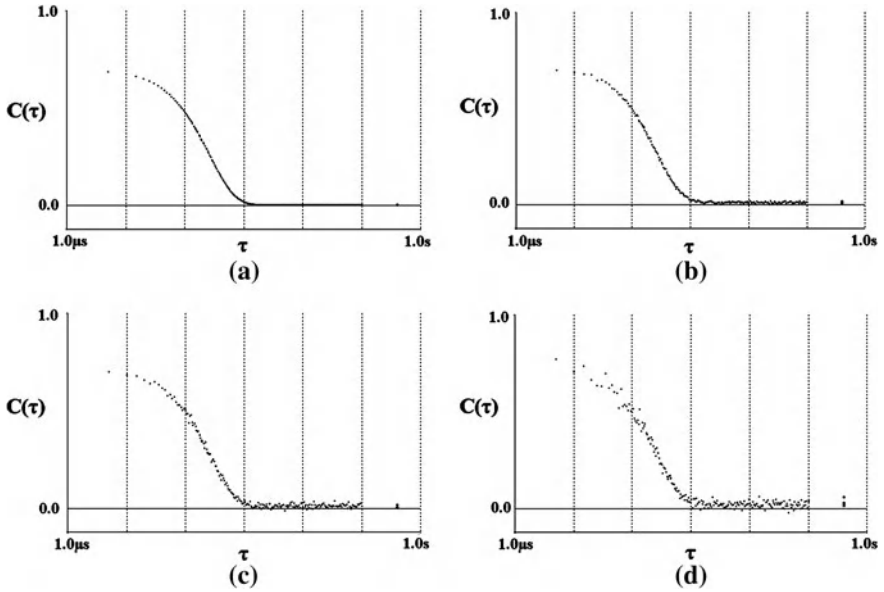


Fig. 13.3 The N-MACF curves of scattered intensity when NF is 0.0, 0.2, 0.5, 0.8 respectively. **a** NF = 0. 0; **b** NF = 0. 2; **c** NF = 0.5; **d** NF = 0.8

Table 13.1 gives the results of these experiment with different NFs which is equal to 0, 0.2, 0.5, 0.8 separately. As shown in the table, each value of NF was simulated five times continuously and the results among four methods are listed.

First, in order to confirm that SOCI-R, SOCI-σ and SOCI-WM are going to work exactly in ideal when NF is zero. From the first part (NF = 0) in Table 13.1, each result of SOCI-R, SOCI-σ and SOCI-WM is 90.1 nm and the entire RMS errors are zero. It can be seen that the results are accurate. However, the reason that the results of BI-CI is 90.2 nm, is because the different software system, between BI and LabView, would have different calculating precision or approximation.

Second, compared with the results of BI-CI, with the increase of NF from 0.2 to 0.8, the average diameter of SOCI-R is much similar to the ideal result of 90.1 nm basically. Actually, when NF = 0.8, these two results are almost the same. But it may conclude that the results of SOCI-R disperse much greatly than those of BI-CI, particularly when NF is large. Although the precision has been improved, the variation of SOCI-R is the biggest of the four.

Third, the average diameter values of SOCI-σ deviate from 90.1 nm a little more than those of SOCI-R do. But unlike the state of SOCI-R, the RMS errors of SOCI-σ are smaller than those of SOCI-R . So it is the basis of applying weighted mean combined with the fitting standard deviation and the fitting correlation coefficient.

Table 13.1 Second order cumulant results of comparative experiment with NF = 0, NF = 0.2, NF = 0.5 and NF = 0.8

Times	NF = 0				NF = 0.2			
	BI-CI	SOCI-σ	SOCI-R	SOCI-WM	BI-CI	SOCI-σ	SOCI-R	SOCI-WM
First	90.2	90.1	90.1	90.1	91.2	91.9	90.7	90.7
Second	90.2	90.1	90.1	90.1	88.8	88.3	89.6	90.6
Third	90.2	90.1	90.1	90.1	89.9	88.9	91.4	90.4
Fourth	90.2	90.1	90.1	90.1	90.3	88.5	88.9	91
Fifth	90.2	90.1	90.1	90.1	89.3	91	90	89.8
Avg(d)	90.2	90.1	90.1	90.1	89.9	89.72	90.12	90.5
RMS	0	0	0	0	0.925	1.625	0.968	0.447
Avg (Poly)	0.003	1.08E−06	1.10E−06	9.70E−07	0.029	0.017	0.074	0.068
Times	NF = 0.5				NF = 0.8			
	BI-CI	SOCI-σ	SOCI-R	SOCI-WM	BI-CI	SOCI-σ	SOCI-R	SOCI-WM
First	86.8	92.1	92.1	90.6	83.9	84.7	83.3	86.3
Second	89.3	88.2	88.3	89.1	89.3	88.1	89.8	89.8
Third	87	90.9	88.8	90.2	88.6	90.9	94.1	90.8
Fourth	85.3	90.7	88.5	89.4	93.3	93.3	93.8	92.3
Fifth	87.7	91	91	90.4	92.6	88.1	86.3	90.8
Avg(d)	87.22	90.58	89.74	89.94	89.54	89.02	89.46	90
RMS	1.455	1.438	1.707	0.654	3.75	3.248	4.701	2.253
Avg (Poly)	0.122	0.02	0.099	0.038	0.132	0.041	0.136	0.061

Fourth, adjust the weighted coefficients in Eq. 13.6 by setting W_1 to 1 and W_2 to 10. Clearly from Table 13.1, the results of SOCI-WM are the best fit to the ideal diameter no matter how many the NF is. Meanwhile, they are the most consistent of all according to the RMS errors. That is to say, in application of minimum weighted mean, the SOCI-WM method not only improves much more precision of the results, but also increases their stability and reliability.

Finally, the last line in Table 13.1 lists the average value of Poly in simulation experiment. It is found that when NF is equal to 0, the Poly of these three self-optimizing methods are very small just as BI-CI. In three other cases of NF, the results of Poly of SOCI-R, SOCI-σ and SOCI-WM are almost less than 0.1except one. According to the criterion for Poly as mentioned above, the results of SOCI-σ and SOCI-WM can accurately indicate the monodisperse sample which is the default setting. Although the results of Poly of SOCI-WM are not the minimum value, the difference of the Poly between SOCI-σ and SOCI-WM should be neglected.

13.5 Summary

In this paper the principles of DLS and the CCI were elaborate. It analyzed and discussed the reasons of unstable results of CCI in detail. Then a SOCI based on minimum weighted mean combined with fitting standard deviation and fitting

correlation coefficient (SOCI-WM) was proposed. After the introduction of the fundamentals of this SOCI, the detailed procedure was presented step by step. Through the comparative simulation experiment with the BI-CI, it proved the way of using minimum weighted mean to realize the process of finding the best fitting results automatically is valid and feasible. The greatest advantage of this approach is a good solution to the baseline drifts and noise interference, particularly in high intensity of noise. It obviously improved the precision and stability of CI and had good value of application and popularity.

References

1. Pecora R (1985) Dynamic light scattering: application of photon correlation spectroscopy. Plenum Press, New York
2. Brown W (1993) Dynamic light scattering: the method and some applications. Clarendon Press, Oxford
3. Koppel DE (1972) Analysis of macromolecular polydispersity in intensity correlation spectroscopy: the method of cumulants. J Chem Phys 57(11):4814
4. Lawson CL, Hanson RJ (1974) Solving least squares problems. Pritice-Hall, Englewood Cliffs
5. Provencher SW (1976) An eigenfunction expansion method for the analysis of exponential decay curves. J Chem Phys 64(7):2772
6. Yu L (2006) Iterative CONTIN algorithm for particle sizing in dynamic light scattering. Opto Electr Eng 33(8):64–69
7. Shaoxin Li (2008) Inversion of particle size distribution from dynamic light scattering data with Gray-code genetic algorithm. Chin J Comput Phys 25(3):323–329
8. Ren S, Han P, Yang G (2011) Real-coded hybrid genetic algorithm for particle sizing with photon correlation spectroscopy. Chin J Comput Phys 28(1):87–93
9. Bertero M, Pike ER (1991) Exponential-sampling method for Laplace and other dilationally invariant transforms: II. Examples in photon correlation spectroscopy and Fraunhofer diffraction. Inv Probl 7(1):21
10. Information on http://www.bic.com/
11. Grabowski E, Morrison I (1983) Particle size distributions from analysis of quasi-elastic light scattering data. In: Dahneke B (ed) Measurements of suspended particles by quasi-elastic light scattering. Wiley-Interscience, New York
12. Qiu J, Yang G, Han P et al (2011) Self-optimizing cumulative inversion of particle size distribution with photon correlation spectroscopy. Chin J Comput Phys 28(2):268–274

Chapter 14
LQR-Based Adaptive Control Strategy for the Planar Double Inverted Pendulum

Jiaolong Zhang and Wei Zhang

Abstract In this paper, based on previous research results and the linear quadratic optimal theory, the improved LQR adaptive controller was designed that can optimize control effect using coordinated factor. The mathematical model of planar double inverted pendulum is established by means of analytical dynamics method, then the optimal controller is presented with LQR theory, and further the output of the LQR adaptive controller is refined through coordinator, which is the function of the states of planar pendulum, and on account of that, control action exerted on the pendulum is improved. Simulation results together with pilot scale experiment in the lab verify the efficacy of the suggested scheme, show that the controller designed to be simple, real-time is good, and can also ensure that the different operating conditions with high control precision, fast response, good stability and robustness.

Keywords Inverted pendulum · LQR adaptive · Coordinator

14.1 Introduction

Inverted pendulum system has the characteristics of nonlinear, redundancy, uncertainty, strong coupling and the natural characteristics of instability. All these features make it the ideal model of advanced control theory and typical experiment

J. Zhang (✉) · W. Zhang
School of Electrical Engineering and Automatization,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zjl@hpu.edu.cn

W. Zhang
e-mail: zwei1563@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_14,

platform of test control results. With the development of control theory, fuzzy control, variable structure control, energy control, adaptive control methods are generally accepted. In recent years, compound control combining several control methods has become the central issue of inverted pendulum control research, and can take full advantage of their strengths. Scholars from various countries has done many studies on inverted pendulum control [1, 2]. These control methods have some limitations. At the same time more complex control algorithms, the worse the effect of real-time control, will also increase the difficulties of implementation.

This paper analyzes the characteristics of planar inverted pendulum based on the linear quadratic optimal theory, optimal controller is designed. In control of this controller, control effect of the inverted pendulum is not good enough, LQR self-adjusting controller with optimize factor is presented. It automatically adjusts the size of control according to actual conditions, further the output of the LQR controller is refined through optimize factor, which is the function of the states of planar pendulum, and control action exerted on the pendulum is improved. It is easy to achieve stable control of inverted pendulum.

14.2 Mathematical Model of Inverted Pendulum

This object of study is the Googol company GPIP2002 planar inverted pendulum. The mathematical model of planar double inverted pendulum is established by means of analytical dynamics method. Inverted pendulum system is mainly made up of controlled object, the direction of horizontal rails, servo motor, drive shaft and electric drive equipment. Control object is composed of the car, the downside pendulum, the upside pendulum, and two rotation shafts which connected to the car and pendulums as shown in Fig. 14.1. Based on the fulcrums of pendulum, coordinates $o_1x_1y_1z_1$ and $o_2x_2y_2z_2$ are established. In condition of ignore effects of air resistance and friction, etc. Planar inverted pendulum can seen as comprising car platform, shaft mass, the uniform downside pendulum, and the uniform upside pendulum.

The basic parameters are defined as follows: l_1, l_2 are the length of the downside pendulum and the upside pendulum respectively. m_1, m_2 are the quality of the downside pendulum and the upside pendulum respectively. m_3 is the quality of link mass linked to the pendulum. M_x, M_y are the quality of the platform motion part and swing bearing of X direction and Y direction respectively. Specific values are: $l_1 = 0.2\,\text{m},\ l_2 = 0.55\,\text{m},\ m_1 = 0.06\,\text{kg},\ m_2 = 0.13\,\text{kg},\ m_3 = 0.27\,\text{kg}, \text{g} = 9.8\,\text{m/s}^2$.

In this paper, differential equations of inverted pendulum system is established by the Lagrange equation [3]. The Lagrange operator is as follows:

$$L(q, \dot{q}) = T(q, \dot{q}) - V(q, \dot{q}) \tag{14.1}$$

where q is the generalized coordinate of system. T is kinetic energy of the system. V is potential energy of the system. From the generalized coordinates q and L, Lagrange equations can be expressed as:

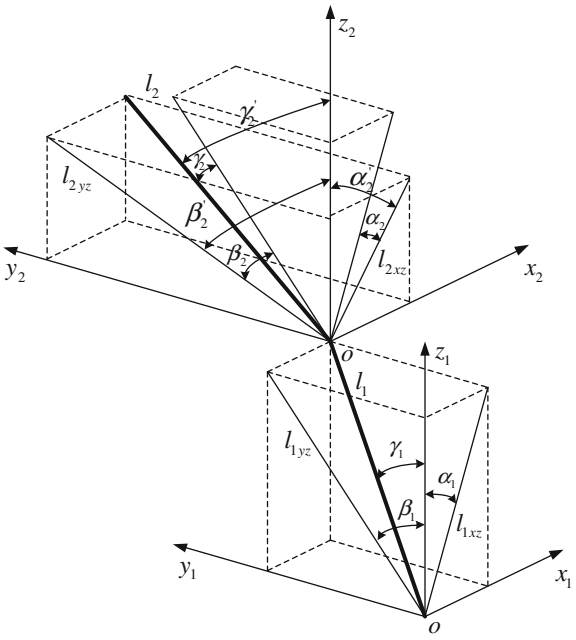


Fig. 14.1 Schematic of planar double inverted pendulum

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = \tau_i \tag{14.2}$$

where i is system variable label, $i = 1, 2, \ldots, n$, $q = \{q_1, q_2, q_3, \ldots\}$ is generalized variable, τ_i are system generalized external forces along the direction of generalized coordinate. For planar double inverted pendulum, the generalized coordinates are: $x, y, \alpha_1, \beta_1, \alpha_2, \beta_2$. The total kinetic energy of the system is:

$$T = T_M + T_{m_1} + T_{m_2} + T_{m_3}$$

The total potential energy of the system is: $V = V_M + V_{m_1} + V_{m_2} + V_{m_3}$.

Relevant parameters are used in (14.1), Since the external forces of the generalized coordinates $\alpha_1, \beta_1, \alpha_2, \beta_2$ is zero, and equations can be established by Eq. (14.2). $\ddot{\alpha}_1, \ddot{\beta}_1, \ddot{\alpha}_2, \ddot{\beta}_2$ are worked out. In the equilibrium position $(q = \dot{q} = 0)$, the Taylor series expansion of $\ddot{\alpha}_1, \ddot{\beta}_1, \ddot{\alpha}_2, \ddot{\beta}_2$ and linearization are executed, parameter values are brought into, decoupled state equations in X and Y directions of double inverted pendulums respectively are solved.

$$\begin{cases} \dot{X}_x = A_x X_x + B_x u_x \\ Y_x = C_x X_x \end{cases} \qquad \begin{cases} \dot{X}_y = A_y X_y + B_y u_y \\ Y_y = C_y X_y \end{cases}$$

where control actions of X and Y direction are $u_x = \ddot{x}, u_y = \ddot{y}$, State variables are: $X_x = [x, \alpha_1, \alpha_2, \dot{x}, \dot{\alpha}_1, \dot{\alpha}_2]^T, X_y = \left[y, \beta_1, \beta_2, \dot{y}, \dot{\beta}_1, \dot{\beta}_2\right]^T$.

$$A_x = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 50.2235 & -14.7272 & 0 & 0 & 0 \\ 0 & -50.8908 & 49.4875 & 0 & 0 & 0 \end{pmatrix}, \quad B_x = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -5.1248 \\ 5.1926 \end{pmatrix},$$

$$C_x = \mathrm{diag}(1, 1, 1, 1, 1, 1), \quad A_y = A_x, B_y = B_x, C_y = C_x.$$

From state equations, linear models in X and Y directions of double inverted pendulums have been decoupled respectively, and can be controlled separately. Each system has one input, six outputs, and the control object is greatly simplified. Controllability principle can be applied to verify the controllability of the system.

14.3 Design of LQR Adaptive Controller

After linearization of the mathematical models, linear optimal control theory is applied to design LQR controller for the state equation of X direction [4]. Y direction uses the same control algorithm. For the form as fellow, state feedback control can be found:

$$u(t) = -Kx(t)$$

So as to minimize the performance index function

$$J = \frac{1}{2}\int_0^{T_i} \left[x^T(t)Qx(t) + u^T(t)Ru(t)\right] \mathrm{d}t$$

where Q is semi-definite matrix, R is positive definite matrix, Q and R are the weighting matrixes of state variables and input variables respectively. For the smallest performance index function, at first Hamilton function is constructed, derivation of this function was obtained and made equal to zero, which can determine the optimal control rate:

$$u(t) = -Kx(t) = -R^{-1}B^T Px(t)$$

where P is the only positive definite symmetric solution which meets the Riccati equation PA + ATP – PBR − 1BTP + Q = 0.

When the system is stable, inverted pendulum control objective is that the downside pendulum and the upside pendulum are upright, the car is at the central. So according to the actual situation, control strategy should stabilize the upside pendulum as a priority, the downside pendulum second, and final control car displacement. After extensive simulation, when Q = diag [200, 400, 600, 1, 1, 1],

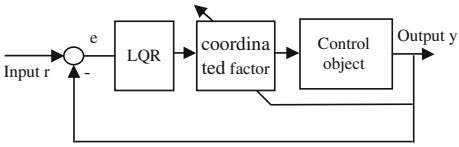


Fig. 14.2 Schematic of LQR adaptive controller

$R = 1$, by means of MATLAB simulation, LQR state feedback gain matrix is calculated, that is $K = [14.14, 66.83, 97.19, 13.92, 19.73, 16.98]$.

In general, LQR controller is designed on the basis of the weighting matrixes Q and R system performance reaches optimal in such conditions. But there are ways to optimize its control action. The output of the LQR controller is refined through coordinated factor $(a - be^{-|x|}), (a > 0, b > 0)$ in this paper, and the basic principle is shown in Fig. 14.2. The controller can adjust the size of control action adaptively according to the change of state variables, and ensure that systems have fast response and robustness.

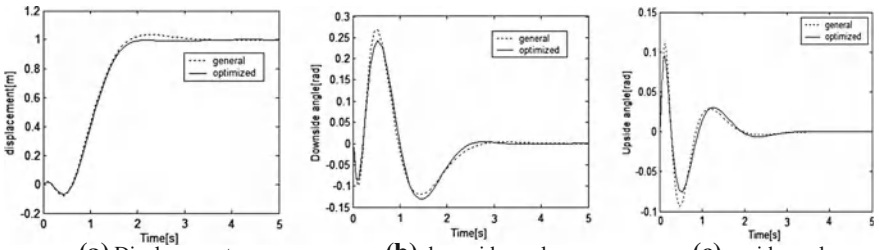
According to coordinated factor expression $(a - be^{-|x|})$, when the system is unstable, state variable such as the upside pendulum angle will increase, $e^{-|x|}$ and also decrease with it, and optimize factors will increase at this time and strengthen control action. This makes for stabilize inverted pendulum faster. When the inverted pendulum is gradually stabilized, state variable such as the upside pendulum angle will decrease $e^{-|x|}$ also will increase with it, and optimize factors will decrease at this time and weaken control action. This can avoid excessive vibration of inverted pendulum and reduce the system error. The system is more stable. But the parameters a and b in factor need to be optimized according to the actual situation.

14.4 Experimental Results

14.4.1 Simulation Results

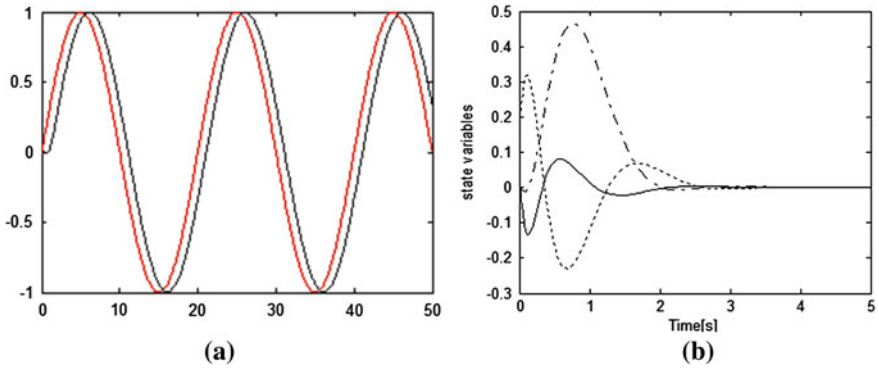
According to the state equation of inverted pendulum, simulation model can be constructed in the SIMULINK environment [5]. When state variable of optimize factor refers to the upside pendulum angle, at this time,and initial state equal to zero. After several simulations, coordinated factor $(a - be^{-|x|})$ parameter $a = 1.98$, $b = 0.45$, $K = [14.14,66.83, 97.19,13.92,19.73,16.98]$., the displacement of car can track unit step signal, downside pendulum and upside pendulum reach to stable within 3 s. SIMULINK curves of car displacement, downside angle and upside angle are shown in Fig. 14.3.

Carry out square wave signal tracking experiment in the zero state conditions, sine wave signal amplitude is 1 m, frequency is 0.05 Hz. The simulation curve of displacement is shown in Fig. 14.4a. From the figure we can see that fast response of LQR adaptive controller is better than general LQR controller and system has better stability.



(a) Displacement; **(b)** downside angle; **(c)** upside angle

Fig. 14.3 Simulation curves of state variables. **a** Displacement; **b** downside angle; **c** upside angle



(a) **(b)**

Fig. 14.4 Simulation curves of displacement and state variables. **a** Displacement tracking sine wave; **b** state variables

When the external disturbance is zero, and initial state is [0,0.2,0,0,0,0], at this time the downside angle is 0.2 rad, car displacement and the upside angle are zero, response curves of system zero input are shown in Fig. 14.4b.

14.4.2 Real-Time Results

Based on the simulation, real-time control test is implemented by using Real-Time Workshop (RTW) in SIMULINK. The control object is a GPIP2002 inverted pendulum system. Dynamic link library (DLL) developed from interface MEX generated by Visual C++, is the communicate intermediary of MATLAB and PCI data acquisition card. Simulation model and M documents are drawn up to achieve the user's own control algorithm. Control parameters are modified easily.

Planar inverted pendulum has no ability to swing automatically, needs to manually put up the pendulum to the equilibrium position. When real-time control

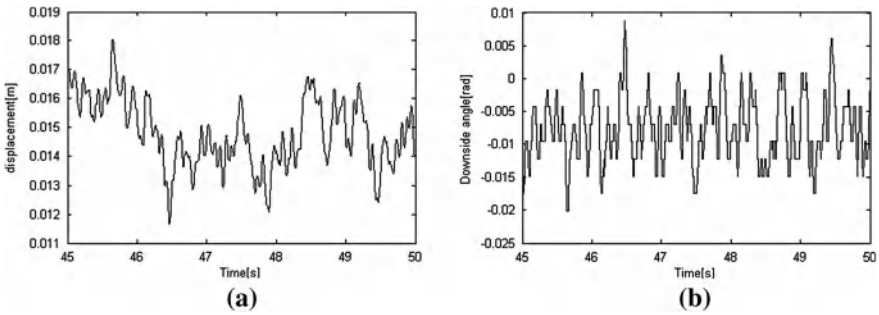


Fig. 14.5 Real-time curves state variables. **a** Displacement; **b** downside angle

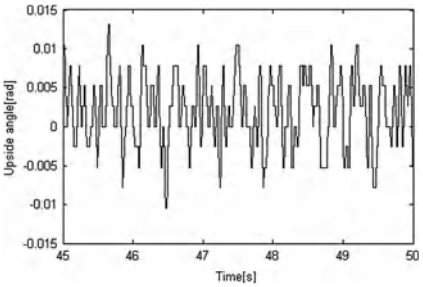


Fig. 14.6 Real-time curves upside angle

module is to be started, release the pendulum pole after exert control action to car, the pendulum will maintain balance in control of computer.

Figures 14.5 and 14.6 show respectively real-time curves of displacement, downside angle and upside angle in control of LQR adaptive controller.

14.5 Conclusion

The mathematical model of planar double inverted pendulum is established by means of Lagrange equation. LQR self-adjusting controller is proposed in this paper, and has good control effect. The simulation model is constructed in SIMULINK, which can reduce the programming effort and easily change the control parameters. This LQR self-adjusting controller has successfully achieved the control of inverted pendulum physical system by using RTW which calls DLL. The controller this paper presented simplifies the control algorithm and has good real-time. Experimental results show the effectiveness of the design.

References

1. Yang KZ, Piao G, Cheng S (2005) Control of double inverted pendulum by using state varieties synthesis fuzzy neural network. Cont Decis 22(1):123–128
2. Li H, Miao Z, Wang J (2005) Variable universe adaptive fuzzy control of fourfold inverted pendulum. Sci China Ser E 32(1):65–75
3. Luo C, Hu D, Zhu X, Dong G (2005) Quintuple inverted pendulum control based on LQR and fuzzy piecewise interpolation. Cont Decis 20(4):392–397
4. Niemann H (2005) Design and analysis of controllers for a double inverted pendulum. ISA Trans 44(1):145–163
5. Mira C (2008) LQR control for a rotary inverted pendulum. In: Proceedings of the 2008 international conference on genetic and evolutionary methods, GEM 2008, pp 212–215

Chapter 15
Simulation Research on Self-Adaptive Dynamic Optimizing Control for Step Change of Parameters of Control Plant

Guoqiang Li, Yuanfeng Zhang, Yongqin Liu and Wenjiang Liu

Abstract All effects on self-adaptive dynamic Optimizing control system from step change and drift of every parameter of control plant and control method are discussed in detail. Theoretical analysis and simulation research show that whichever suitable step period T_0 is selected and the parameter of pure time $\tau < T_0/2$ delay is satisfied, the simple, convenient and practical self-adaptive dynamic optimizing method for an extremum value control system is feasible, regardless of the parameters of control plant step and drift. It can solve a never-evading problem in the practical industrial process by using the previous method. The problem is that accurate identification of the parameters and order of linear part in the extremum value control system is too difficult. The method in this paper only needs much less a priori information concerning the controlled plant. This new method can automatically not only identify the parameters of control plant, but can also automatically adapt to the drift of the parameters. So it is very simple, convenient and practical to be realized in the real industrial process.

Keywords Step of parameters · Simulation · Self-adaptive search · Simple and practical

G. Li (✉) · Y. Zhang · Y. Liu
School of Physics and Electrical Engineering,
Weinan Normal University, Weinan 714000, China
e-mail: lgq@wntc.edu.cn

Y. Liu
e-mail: liuyongqin@163.com

W. Liu
Automatic Control Department,
Xi'an Jiaotong University, Xi'an 710049, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_15,

15.1 Introduction

As for the extremum value controlled plant, a combustion system of the boil and the heating furnace is a typical example, their parameters of linear part in these controlled plant will be a bigger change after a overhaul, in other words, the step change of the parameters will take place, as pointed out in Ref. [1]. After the step change of the parameters the control system cannot work well so it has to stop to identify the parameters of the control plant to adjust the parameters of the control system again by using the previous methods [2–4] which is a great hindrance for the extremum value control system to be used in the industrial process. Therefore, a self-adaptive dynamic optimization method is presented by us so that the problem is solved theoretically after the step change of the parameters. However, it is necessary to know accurately the order N of linear part in the extremum value control plant by this method. It is very difficult to obtain accurately the order in the real industrial process. So we presented a simple, convenient and practical self-adaptive dynamic optimizing method for an extremum value control system. This new method is based on the linear part of the control plant being approximated as a cascade of a lower order dynamic element and a pure time-delay element. Whether this simple and convenient method can adapt automatically the step and drift of the parameters of control plant or not is necessary to simulate.

15.2 Brief Review of a Simple and Convenient Method for Self-Adaptive Dynamic Optimization Not, It is Necessary to Simulate

Let a step-by-step controller be used in the extremum value control system. Assuming the step-by-step period is changeless, the extremum value controlled plant may be expressed as a cascade form of a nonlinear element and a linear element as shown in Fig. 15.1.

Assuming that the transfer function of linear part in the extremum value control system is given as:

$$G(s) = \frac{1}{(T_1 S + 1)(T_2 S + 1) \cdots (T_N S + 1)}, \quad T_1 > T_2 > T_3 \cdots\cdots > T_N > 0 \tag{15.1}$$

The corresponding unit step response function becomes as follows [5–7]:

$$\hbar(t) = A_0 + \sum_{i=1}^{N} A_i e^{-\frac{t}{T_i}},\ A_0 = G(s)|_{s=0} = 1 \text{ and } A_i = \left(s + \frac{1}{T_i}\right)\frac{1}{s} G(s)|_{s=-\frac{1}{T_i}},$$
$$i = 1, 2, \ldots, N$$

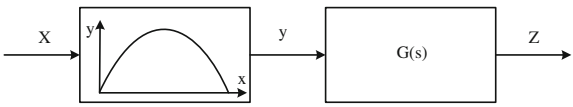


Fig. 15.1 Block diagram of extremum value control system

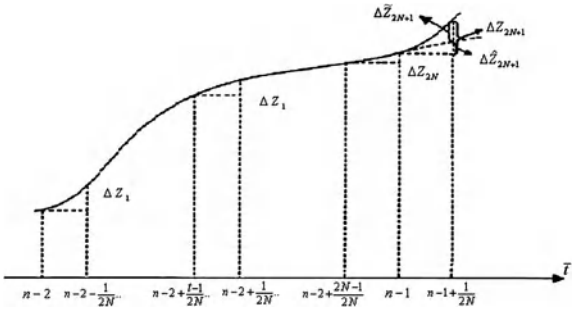


Fig. 15.2 Self-adaptive dynamic optimizing algorithm diagram

if time t is quantized with step-by-step T_0, it then becomes:

$$\hbar(\bar{t}) = 1 + \sum_{i=1}^{N} A_i e^{-\alpha_i \bar{t}}, \quad \alpha_i = \frac{T_0}{T_i};\ i = 1, 2, \ldots, N;\ \text{and}\ \bar{t} = \frac{t}{T_0} \tag{15.2}$$

Assuming each step-by-step period is divided into $2N$ equal intervals as shown in Fig. 15.2, the sample values are taken at divided point. $2N + 1$ sample value would be obtained, and the pre-estimating comparative point is the first sampling point after exploring step.

$\Delta\hat{Z}_{2N+1}$: pre-estimating difference value of the system output if the probing step is not exerted at $\bar{t} = n - 1$.

ΔZ_{2N+1}: the real difference value of the system output as the probing step is exerted at $\bar{t} = n - 1$.

$\Delta\tilde{Z}_{2N+1}$: the deviation in value between the real difference value of the system output at the comparing point.

As given in Ref. [2], comparing value becomes:

$$\Delta\tilde{Z}_{2N+1} = \begin{vmatrix} \Delta Z_1 & \Delta Z_2 & \cdots & \Delta Z_{N+1} \\ \Delta Z_2 & \Delta Z_3 & \cdots & \Delta Z_{N+2} \\ \cdots & \cdots & \cdots & \cdots \\ \Delta Z_{N+1} & \Delta Z_{N+2} & \cdots & \Delta Z_{2N+1} \end{vmatrix} \Bigg/ \begin{vmatrix} \Delta Z_1 & \Delta Z_2 & \cdots & \Delta Z_{N+1} \\ \Delta Z_2 & \Delta Z_3 & \cdots & \Delta Z_{N+2} \\ \cdots & \cdots & \cdots & \cdots \\ \Delta Z_{N+1} & \Delta Z_{N+2} & \cdots & \Delta Z_{2N+1} \end{vmatrix} \tag{15.3}$$

$$\Delta\tilde{Z}_{2N+1} = \Delta Z_{2N+1} - \Delta\hat{Z}_{2N+1} \tag{15.4}$$

Since the direction of the step-by-step increment can be determined by using Eq. (15.4) with the output sample value, the expression of judging the step-by-step is as follows::

$$\mathrm{sgn}[\Delta x_n] = \mathrm{sgn}\left[\Delta\tilde{Z}_{2N+1}\Delta x(n-1)\right] \tag{15.5}$$

where $\mathrm{sgn}[\Delta x_n]$ implies the sign of x

In Eq. (15.1), if $T_1 \gg T_2 > T_3 \cdots\cdots > T_N > 0$, it then becomes:

$$\begin{aligned} G(s) &\approx \frac{1}{(1+T_1 s)e^{T_2 s}e^{T_3 s}\cdots e^{T_N s}} = \frac{1}{(1+T_1 s)e^{\tau s}} = \frac{e^{-\tau s}}{1+T_1 s}\tau \\ &= T_2 + T_3 + \cdots + T_{\mathrm{N}} \end{aligned} \tag{15.6}$$

Therefore, a real, accurate and higher order linear part of the controlled plant could be approximated by a cascade of a lower order dynamic element and a pure time-delay element. As for the effect of a time-delay element, as pointed out in Refs. [1, 8, 9], that the moment when the step-by-step is exerted in advance of time τ compared with that when there is no pure time-delay element in the controlled plant. We could delete the effect of pure time-delay τ in using Eq. (15.5) It does not make any effect on the self-adaptive dynamic optimizing method [1] by this supplementing method, the method of judging the direction of the step-by-step increment described in [1] can be used all the same. Furthermore, after simplifying a complex and accurate model of the real controlled plant, $N = 1$, Eq. (15.5) of calculating the comparative value can simply be expressed as follows:

$$\begin{aligned} &\Delta\tilde{Z}_{2N+1} = \Delta\tilde{Z}_3 = \begin{vmatrix} \Delta Z_1 & \Delta Z_2 \\ \Delta Z_2 & \Delta Z_3 \end{vmatrix} \quad \Delta Z_1 \\ &N = 2, \quad \Delta\tilde{Z}_{2N+1} = \Delta\tilde{Z}_5 = \begin{vmatrix} \Delta Z_1 & \Delta Z_2 & \Delta Z_3 \\ \Delta Z_2 & \Delta Z_3 & \Delta Z_4 \\ \Delta Z_3 & \Delta Z_4 & \Delta Z_5 \end{vmatrix} \quad \begin{vmatrix} \Delta Z_1 & \Delta Z_2 \\ \Delta Z_2 & \Delta Z_3 \end{vmatrix} \end{aligned} \tag{15.7}$$

15.3 Simulation Block Diagram and Flowchart Extremum Value Controller

A simulation block diagram of extreme control system is shown in Fig. 15.3. Extreme nonlinear characteristics is expressed by Eq. (15.7):

$$y(x) = 0.4x(10 - x) \tag{15.7}$$

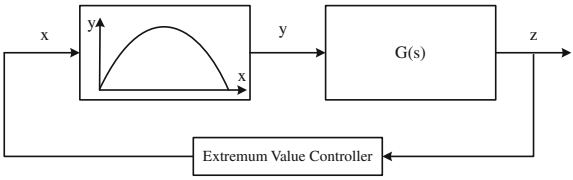


Fig. 15.3 Block diagram showing extremum value control system

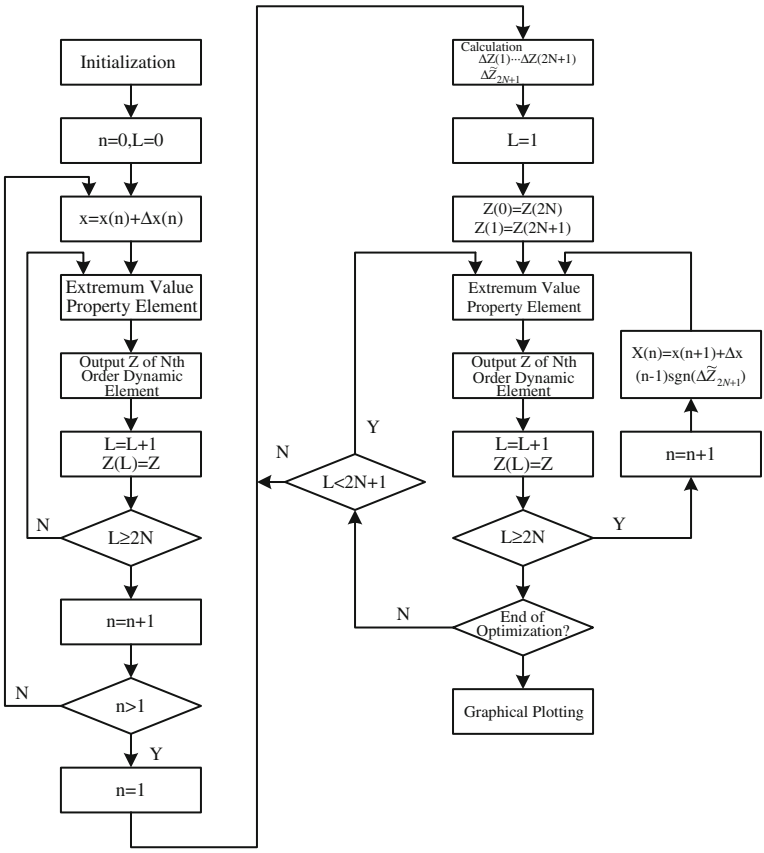


Fig. 15.4 Flowchart of simulation program

When there is no drift extreme characteristics, the initial extreme point is (5, 10), extremum value obtained by Eqs. (15.4) and (15.5). Simulation block diagram shown in Fig. 15.4.

Description of Simulation block diagram:

(1) The main task of initializing is to set the variable initial value, and input parameters.
(2) *N*-order inertial element is obtained by the method of fourth-order Runge–Kutta.

(3) According to the requirements of self-adaptive dynamic optimization algorithm, a step cycle is divided into $2N$ equal parts of each sampling cycle $\Delta t = T_0/2N$, calculating $\Delta\tilde{Z}_{2N+1}$ when $L = 2N + 1$, and determine the direction of the incremental step, meanwhile, prepared for the next delivery step. When $L = 2N$, according to $\Delta\tilde{Z}_{2N+1}$ and signs of $\Delta x(n-1)$ to send the N-step.

15.4 Simulation

In the actual industrial production process, it is extremely difficult to accurately identify the parameters of the control object or even its accuracy order. In this paper, a simple, efficient, and practical self-adaptive dynamic optimization method is presented, whose linear part based on control object can be approximated as a low-level inertial element series with a pure delay element. This simple, practical dynamic optimization method can automatically adapt to the control object parameters of the drift, but a simulation needs to be made whether it can also automatically adapt the control parameters of the jump target. At the same time, extreme characteristics will be slowly drift because of the perturbation, and the optimal operating point will drift too, so the control system's task is not only control object parameters jump, but also automatically search for the most excellent operating point continuously in a dynamic process. Therefore, the purpose of simulation is to test whether the control system can work well to complete the task.

(1) Dynamic optimization for extremum value controlling objects of a third-order inertial element, the linear part transfer function is:

$$G(s) = \frac{1}{(T_1 s + 1)(T_2 s + 1)(T_3 s + 1)}, T_1 = 100\text{ s}, T_2 = 30\text{ s}, T_3 = 10\text{ s}$$

Approximate it as:

$$G(s) \approx \frac{1}{(1 + T_1 s)e^{T_2 s}e^{T_3 s}} = \frac{1}{(1 + T_1 s)e^{\tau s}} = \frac{e^{-\tau s}}{1 + T_1 s}, \tau = T_2 + T_3 = 40\text{ s},$$

Extreme regulator parameters are: step $\Delta X = 2$, step cycle T0 = 90 s, $\tau < T_0/2$, due to $N = 1$, and sampling period $\Delta t = T_0/2N = 45$ s, while changed inertia time constant $T_1 = 100$ s, $T_2 = 40$ s to $T_1 = 150$ s, $T_2 = 20$ s, $T_3 = 15$ s in the process of the dynamic optimization, Fig. 15.5 shows all the dynamic optimization process. It can be seen from the figure that it has no effect on the reliability of the dynamic optimization by changing the time constant of control object.

(2) Control object parameters and extreme regulator parameters are the same as that of Eq. (15.1) . By using this simple, practical self-adaptive dynamic optimization method, a simulation is made for analyzing the ability of extreme regulator tracking the drift of extreme characteristics. Figure 15.6 shows the

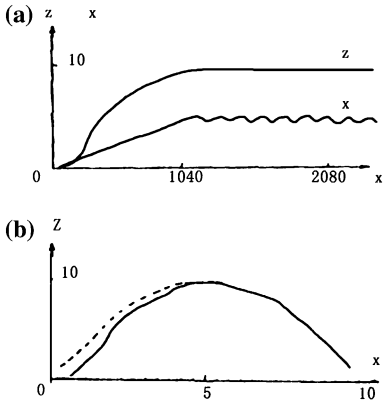


Fig. 15.5 Dynamic optimization process

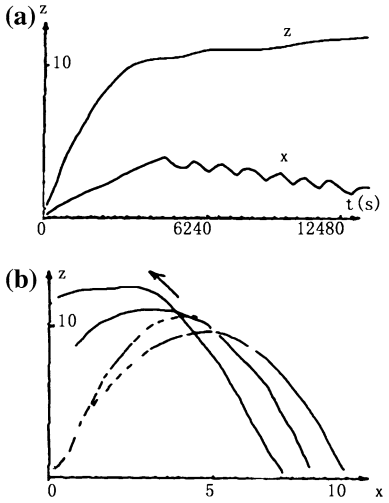


Fig. 15.6 Dynamic optimization process extreme

dynamic optimization process, it can be seen from Fig. 15.6 that the extreme regulator using this method is very good to track the drift of extreme characteristics.

15.5 Conclusion

In actual industrial processes, when boilers, furnaces and other furnace combustion system have a major overhaul, its linear part of the control object parameters will have a greater change,and both the jump and drift will appear. It is very difficult to obtain the linear part parameters and the order of a control object of the extremum value accurately. However, we can simplify the model of the linear part of control

object of the extreme regulation, and extremum value regulated by using the self-adaptive dynamic optimization method with a low-level inertia element series of a pure time delay part instead of an accurate high-level model, as long as a suitable choice of a good step cycle T_0, it is feasible to use this method, regardless of how the parameters of the control object jump and drift. The most prominent advantage of this new method is it requires very little prior knowledge of the control object model, and only needs to know that the linear part of the control object can be approximated by a first-order inertial element and a pure time delay in series or a two-order inertial element with a pure delay part in series. Therefore, the actual debugging process will be much easier, and it will be easy to implement in the industrial production process.

Acknowledgments This work was supported by the National Natural Science Foundation of China (NO. 69374030), Natural Science Foundation of Science and Technology Department of Shaanxi Province (2011JM8021), and Natural Science Foundation of Education Department of Shaanxi Province (2010JK533).

References

1. Li G, Zhu Z (2011) Self-adaptive dynamic optimizing control for step change of parameters of control plant. Mod Electr Technol 2:106–108 (in Chinese)
2. Li G, Jin Y (2004) A self-adaptive control for dynamic seeking optimum of extremum value control system. J Shanxi Normal Univ 3:50–53 (in Chinese)
3. Yuan T, Liu W (2001) A self-tuning extreme regulator with application based on Hammerstein model. Chin J Sci Instrm 22(4):413–415
4. Li G, Zhu Z (2004) Self-adaptive dynamic optimization in an extremum value control system. Bull Okayama Univ Sci Nat Sci 40:161–166
5. Li G (2003) Method of pre-estimating comparative theory and dynamic seeking optimum. Mod Electr Technol 11:52–54
6. Xia X, Shao J (2002) Design of control system for annular gap mill. J Wuhan Yejin Univ Sci Technol 1:51–54
7. Xu F, Jiang Z et al (2007) Adjustable reactor controller based on self-optimizing control. Conv Technol Electr Tract 6:62–68
8. Li G, Jin Y (2003) Simulating research for pre-estimating dynamic optimum of extremum value control system. J Shaanxi Normal Univ 4:46–50 (in Chinese)
9. Li G, Jin Y (2003) Research for simplifying model and decreasing order of extremum value control plant. Mod Electr Technol 3:40–42 (in Chinese)

Chapter 16
Sparse Multipath Channel Estimation Using Regularized Orthogonal Matching Pursuit Algorithm

Rui Wang and Jing Lu

Abstract Conventional linear channel estimation methods, such as the least squares, are unavailable in practical applications since all entries of the solution obtained by these methods were nonzeros. Using the sparsity of channel, several methods have been proposed, such as orthogonal matching pursuit (OMP) and convex program. However, OMP algorithm is unstable in the case of highly redundant dictionary and the convex program method is hard to implement due to its complexity. In this paper, a novel sparse channel estimate strategy named as regularized orthogonal matching pursuit algorithm is proposed. This algorithm combines the advantage of OMP algorithm and convex program methods. Numeral experiments demonstrate that the proposed algorithm is effective for the problem of sparse multi-path channel estimation.

Keywords Sparse multi-path channel (SMPC) · Regularized orthogonal matching pursuit (ROMP) · Sparse approximation

16.1 Introduction

Sparse multi-path channels (SMPC) are frequently encountered in wireless communication applications [1–3]. A sample of SMPC impulse response can be seen in Fig. 16.1. Channel estimation for the purposes of equalization is an active problem in communications and signal processing. Channel estimation has received

R. Wang (✉) · J. Lu
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: wangrui@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_16,

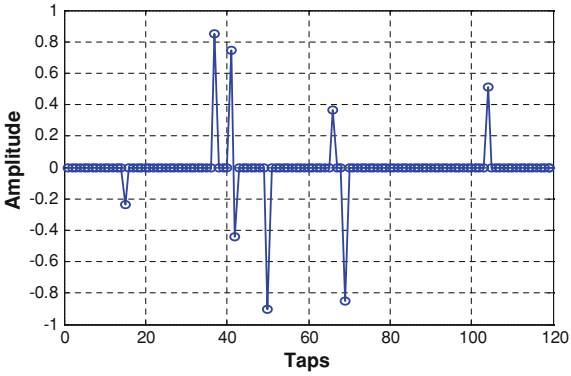


Fig. 16.1 A sample of sparse multi-path channel impulse where its length is 120 and the number of nonzero taps is 8

considerable interest in the last few years for its advantages in high bit rate transmissions over multi-path channel [4].

Exploiting the sparse property, several methods have proposed such as orthogonal matching pursuit (OMP) [5, 6], and convex program [7, 8]. They have advantages and disadvantages.

The major advantages of OMP algorithm are fast and easy to implement. However, Donoho and his collaborators [9] suggest that OMP should be less stable than the convex program. But no theoretical results have been known in either the positive or negative direction [10]. Tropp and Gilbert [6] investigated the performance of OMP algorithm for a set of $n \times p$ $(n \ll p)$ sensing matrices $\mathbf{X}$ that if satisfying four properties, then OMP can recover the signals h with high probability provided that the number of measurements is proportional to the sparsity S of the signals: $n \sim S \cdot \log p$. What a pity that their argumentation is only suitable for some sets of signal which are statistically independent from rows of sensing matrix. In other words, OMP's condition on sensing matrix given in [6] is more restrictive than the restricted isometry condition (RIC). From the Mathematical point of view, Kunis and Rauhut [11] claimed in their empirical works that the first iteration of OMP, based on the number of measurements $n \sim S \cdot \log p$, can identify a correct column from the sensing matrix. Unfortunately, because of the unavoidable correlation between columns of the sensing matrix, it is difficult to analyse subsequent iterations of the OMP algorithm. For example, sparse channel estimation using OMP method can acquires each nonzero tap with high probability, but with high probability failing to estimate all nonzero taps of a channel. This instability of the algorithm easy leads to instability of the wireless communication systems. However, the convex program method can change the embarrassing circumstance of the instability.

The convex program method is based on linear programming, which has its advantages and disadvantages. The main advantage of the convex program method is that it has strong guarantees of exact SMPC estimation. Once the sensing matrix satisfies the RIC, this method works correctly for all sparse signals. But it has disadvantages in computational cost and implementation complexity [12].

One thinks of linear programming as a black box, and any development of fast solvers will reduce the running time of the SMPC estimation method. On the other hand, it is not clear what this running time is, as there is no strongly polynomial time algorithm in linear programming yet.

In this paper, we consider a novel SMPC estimation method using regularized orthogonal matching pursuit (ROMP) algorithm introduced recently by Needell and Vershynin [10]. ROMP algorithm combines advantages of the OMP with the convex program. This new algorithm for SMPC estimation will perform correctly for all sensing matrix $\mathbf{X}$ satisfying RIC. In Sect. 16.4, we will demonstrate that a variety of simulation results confirm the effectiveness of the algorithm which was employed on SMPC estimation.

The remainder of the paper is organized as follows. In Sect. 16.2, sparse channel model is presented. Section 16.3 will restate the regularized orthogonal matching pursuit algorithm of [10]. In Sect. 16.4, we give a variety of simulation results that are performance comparisons of the ROMP with OMP and convex program (such as Lasso method). Finally, conclusions are given in Sect. 16.5.

Notation: In this paper, the superscript T stands for transposition. Bold capital letters denote a matrix where bold small letters indicate a vector. The arithmetical notation $|\cdot|$ stands for the absolute value of a scalar or each component of a vector. Finally, h% and $\mathbf{h}$ indicate an estimate and a real value of channel vector, respectively.

16.2 Sparse Multipath Channel Model

Reliable wireless communications frequently requires accurate knowledge of the underlying SMPC, which was always disturbed by noise. Sparse channel estimation is the problem of processing the input training sequence and channel output to recover channel impulse. Let us transmit the training sequence (called sensing matrix in this paper) $\mathbf{x}(t)$, $t = 0, 1, \ldots, n-1$, through a stationary SMPC. The training sequence symbols $x(t)$ for $t<0$ can be obtained from the previous estimates or for the first arriving frame they are assumed to be zero [13]. In a typical sparse multipath channel estimation scenario, we try to estimate a channel vector $\mathbf{h} \in \mathbb{R}^{\kappa}$ based on the standard underdetermined linear models

$$y_i = \sum_{t=0}^{p-1} \mathbf{x}(t-i)h_t + z_i \tag{16.1}$$

where $i = 0, 1, \ldots, n-1$, h_t is the channel impulse response of length p, z_i is additive white Gaussian noise with zero mean and variance σ^2. In the matrix form, we have

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{z} \tag{16.2}$$

where

$$\begin{aligned}\mathbf{h} &= [\,h_0 \quad h_1 \quad \ldots \quad h_{p-1}\,]^T,\\ \mathbf{y} &= [\,y_0 \quad y_1 \quad \ldots \quad y_{n-1}\,]^T,\\ \mathbf{z} &= [\,z_0 \quad z_1 \quad \cdots \quad z_{n-1}\,]^T\end{aligned}$$

and $\mathbf{X}$ is the known training matrix given by

$$\mathbf{X} = \begin{bmatrix} s(0) & s(-1) & \cdots & s(1-p) \\ s(1) & s(0) & \cdots & s(2-p) \\ \vdots & \vdots & \ddots & \vdots \\ s(n-1) & s(n-2) & \cdots & s(n-p) \end{bmatrix} = [\,\mathbf{x}_0 \quad \mathbf{x}_1 \quad \cdots \quad \mathbf{x}_{p-1}\,].$$

Denote the number of nonzero taps of $\mathbf{h}$ as S. The channel $\mathbf{h}$ is sparse if $S \ll p$ is satisfied. In the context of sparse component analysis (SCA), $\mathbf{X}$ is called sensing dictionary and the column vector $\mathbf{x}_t$ is called atom. As a result of short length training sequence, which improves throughput efficiency for the systems where transmitted packet length is short, the dictionary is highly redundant [14]. In other words, the dimension of the received base-band signal vector $\mathbf{y}$ is much smaller than the number of atoms in the dictionary, i.e., $n \ll p$.

16.3 Sparse Multipath Channel Estimation with ROMP

Regularized orthogonal matching pursuit [10] is an iterative algorithm which is fundamentally based on OMP with some modifications. This method was proved as stable as convex program in [10]. Therefore, this fast and stable method is very suitable for wireless sparse multipath channel estimation. In practical procedures, one can use ROMP [10] to approximately compute a 2S-sparsity vector that is close to the 2S-term approximation $\mathbf{h}_{2S}$ of the SMPC $\mathbf{h}$.

On sparse channel estimation, the ROPM algorithm [10] is summarized as follows:

(1) Initialize: let the channel residual vector $\mathbf{r}_t = \mathbf{y}$, the index set $I_t = \varnothing$, and start the iteration counter with$t = 1$.
(2) Identify: choose a set J of nonzero taps coordinates in magnitude of the observation vector $\mathbf{y} = \mathbf{X}^*\mathbf{r}_t$.
(3) Regularize: divide the set J into subsets J_k which satisfies

$$|u(i)| \le 2 \cdot |u(j)|, \text{ for all } i,j \in J_k,$$

and choose the subset J_0 with the maximum energy of $\left\|u_{|J_0}\right\|_{\ell_2}^2$.

(4) Update: set $I_t = I_{t-1} \cup J_0$. Calculate the new output approximation by solving the least square equation

$$\mathbf{h}\% = \arg\min_{\mathbf{h}\%} \|\mathbf{y} - \mathbf{X}\mathbf{h}\%\|_{\ell_2}.$$

Update the residual: $\mathbf{r} = \mathbf{y} - \mathbf{X}\mathbf{h}\%$.

(5) Stopping: check the stopping criterion, if not, then keeps increasing $t = t + 1$.

16.4 Simulation Results and Discussion

In this section, to gain some insight into the effect of sparseness on SMPC estimation accuracy we carried out the simulation described as follows. Nonzero taps of the SMPC are drawn from a uniform distribution on $[-1, -0.2] \cup [0.2, 1]$ and the number of nonzero taps is 8, that is to say, $\|\mathbf{h}\|_{l_0} = S = 8$. These nonzero taps are generated randomly over $\mathbf{h}$ of order 100. The number of measurements of the Toeplitz structural sensing matrix $\mathbf{X}$ is $n = 80$.

In the first place, we consider overall and nonzero taps estimation error of LS, Lasso, OMP, ROMP and ideal (know the nonzero taps position) using RMSE standard. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{m=1}^{M} \|\mathbf{h} - \hat{\mathbf{h}}_m\|_2^2},$$

where $M = 1000$ is the number of Monte Carlo runs in the simulation. This process is repeated for a large number of channel realizations and the RMSE curve is generated. For each case five average RMSE curves are computed for five different choices of the estimated impulse response for $\mathbf{h}\%_{\text{LS}}$, $\mathbf{h}\%_{\text{OMP}}$, $\mathbf{h}\%_{\text{Lasso}}$, $\mathbf{h}\%_{\text{ROMP}}$ and $\mathbf{h}\%_{\text{Ideal}}$ respectively. In the second place, we consider different sparsity level and estimation error of ROMP using RMSE standard. In the last section, we compare the RMSE of the LS, Lasso, OMP and ROMP which use cumulative density function (CDF) [15].

16.4.1 Comparison of Taps Estimation Error

Figures 16.2 and 16.3 provide the overall taps and nonzero taps estimation error using RMSE standard and cumulative density function for LS, Lasso, OMP, ROMP and ideal. The true channel and the estimation for run $m = 1000$ are denoted by $\mathbf{h}$ and $\hat{\mathbf{h}}_m$ at different SNR, respectively.

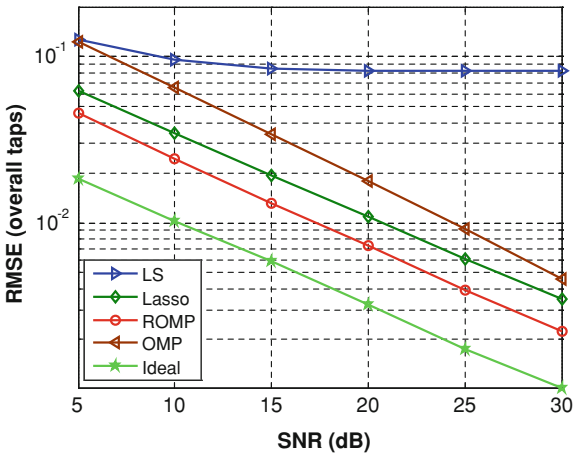


Fig. 16.2 Performance comparison via RMSE of the overall taps with different channel estimation methods at different SNR

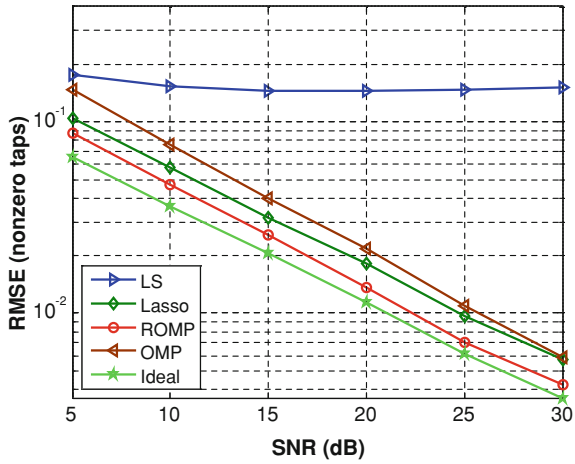


Fig. 16.3 Performance comparison via RMSE of the nonzero taps with different channel estimation methods at different SNR

From the two figures, we can conclude that the estimation performance of ROMP absolutely outperforms the LS and OMP. The estimation performances of the ROMP and Lasso are approximate in theory. However, we can also see that the performance of the ROMP outperforms the Lasso because choosing a proper error bound of the Lasso is very difficult. In these simulations, the error bound of the Lasso was choosing from [16].

16.4.2 Different Sparsity Levels

Figures 16.4 and 16.5 provide the overall and nonzero taps estimation error using RMSE standard for ROMP. From these figures, we can conclude that channel

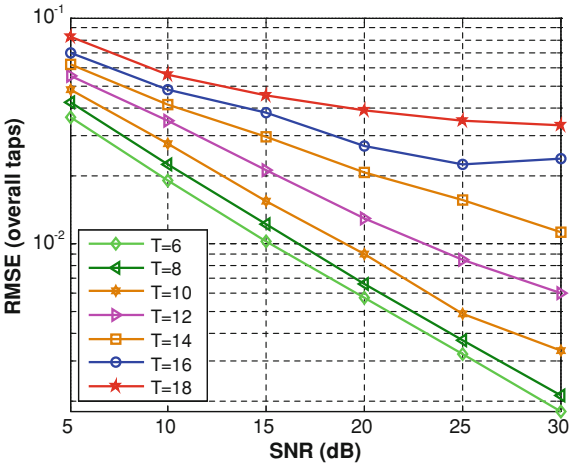


Fig. 16.4 Performance analysis of the ROMP via RMSE standard of the overall taps with different number of nonzero taps

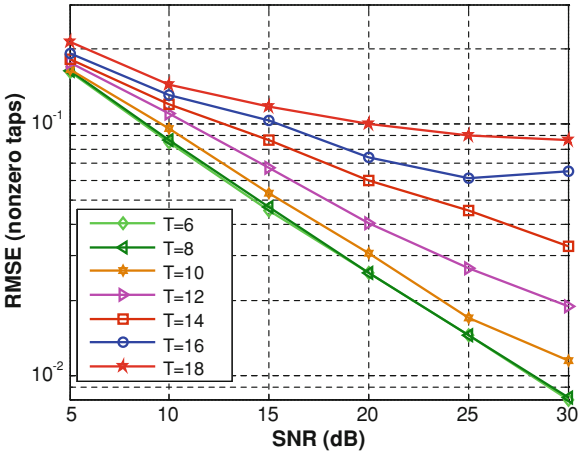


Fig. 16.5 Performance analysis of the ROMP via RMSE standard of the nonzero taps with different number of nonzero taps

estimation performance is better for less sparsity level at the same SNR [17]. As the same estimation performance, smaller nonzero taps need less number of measurements. For example, the RMSE of the line of sparsity level $T = 6$ is always smaller than other lines in Fig. 16.4 at same SNR.

16.5 Conclusions

In this paper, we have presented a novel sparse multipath channel estimation method using regularized orthogonal matching pursuit algorithm. The proposed method has combined with both the greedy algorithm and convex program

advantages. Our results demonstrate that ROMP makes it possible to improve the stability of greed algorithm and to reduce the computational complexity. The effectiveness of the proposed scheme is demonstrated by comparing its performance with existing schemes even if the number of variables is much larger than the number of observations.

References

1. Schreiber WF (1995) Advanced television systems for terrestrial broadcasting: some problems and some proposed solutions. IEEE Proc 83:958–981
2. Mallat SG, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Trans ASSP 41(12):3397–3415
3. Kocic M et al (1995) Sparse equalization for real time digital underwater acoustic communications. In: Proceedings of the OCEANS'95, San Diego, CA, pp 1417–1422
4. Carbonelli C, Vedantam S, Mitra U (2007) Sparse channel estimation with zero tap detection. IEEE Trans Wirel Commun 6(5):1743–1763
5. Cotter SF, Rao BD (2002) Sparse channel estimation via matching pursuit with application to equalization. IEEE Trans Commun 50(3):374–377
6. Tropp JA, Gilbert AC (2007) Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory 53(12):4655–4666
7. Gui G, Wan Q, Huang AM, Jaing CG (2008) Partial sparse multipath channel estimation using L1_LS algorithm (submitted for IEEE TENCON)
8. Bajwa WU, Haupt J, Raz G, Nowak R (2008) Compressed channel sensing. In: Proceedings of the 42nd annual conference information sciences and systems (CISS'08), pp 19–21
9. Donoho D, Elad M, Temlyakov V (2006) Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Trans Inf Theory 52(1):6–18
10. Tropp J (2004) Greed is good: algorithmic results for sparse approximation. IEEE Trans Inf Theory 50:2231–2242
11. Kunis S, Rauhut H (2006) Random sampling of sparse trigonometric polynomials ii—orthogonal matching pursuit versus basis pursuit. Arxiv: preprint math.CA/0604429v2
12. Nguyen NH, Tran TD (2007) The stability of regularized orthogonal matching pursuit algorithm. http://www.dsp.ece.rice.edu/cs/Stability_of_ROMP.pdf
13. Cotter SF, Rao BD (2002) Sparse channel estimation via matching pursuit with application to equalization. IEEE Trans Commun 50(3):277–374
14. Schnass K, Vandergheynst P (2008) Dictionary preconditioning for greedy algorithms. IEEE Trans Inf Theory 56(5):1994–2002
15. Candès E, Tao T (2005) Decoding by linear programming. IEEE Trans Inf Theory 51: 4203–4215
16. Candès E, Tao T (2007) The Dantzig selector: statistical estimation when p is much larger than n. Ann Stat 35:2313–2351
17. Bajwa WU, Haupt J, Raz G, Wright S, Nowak R (2007) Toeplitz structured compressed sensing matrices. In: Proceedings of the 14th IEEE/SP workshop on statistical signal processing (SSP'07), pp 294–298

Part II
Control Theory

Chapter 17
FSM Compensation Technology for High Accuracy Airborne Stabilized Sighting System

Feifei Xu, Ming Ji, Qingqing Xu, Chuangshe Zhao and Shaojun Shou

Abstract The fast-steering mirror (FSM) is a device used to control the direction of light beam accurately. As a sensitive instrument, it is very significant in astronomical telescope, laser communication, autoadaptation optical system and laser weapon, etc. Aiming at the difficulty of improving stabilization accuracy for airborne stabilization technology, an FSM driven by piezoelectric ceramic is proposed to compensate residual error of traditional integrated stabilized platform in order to improve system stabilization accuracy greatly. Through simulation and experiment validation, it shows that this method can improve stabilization accuracy by more than a magnitude.

Keywords Airborne stabilized sighting system · Fast steering mirror · High accuracy

17.1 Introduction

An electro-optical stabilized sighting system is a mechanical-electronic servo mechanism which provides LOS axis stabilization and aiming capability for electro-optical detectors (such as FLIR, Laser, TV, LLL and visible sight) through isolating posture disturbance caused by such mobile carriers as vehicles, ships and aircrafts [1, 2]. As it will affect imaging quality of detectors and aiming accuracy greatly, the system has become a key part of modern mobile detection and guidance arms used for anti-vehicles, anti-aircraft and anti-warship.

F. Xu (✉) · M. Ji · Q. Xu · C. Zhao · S. Shou
Xi'an Institute of Applied Optics, 710065 Shannxi, China
e-mail: Fajir@sohu.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_17,

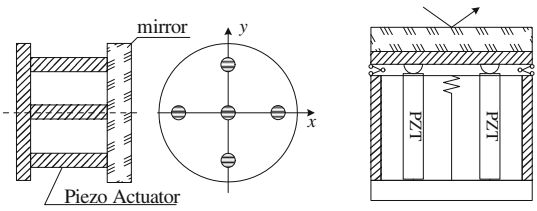


Fig. 17.1 Principle of PZT fast-steering mirror

Airborne electro-optical stabilized sighting system usually calls for high accuracy transmission, big transmission travel, high response speed and high mechanical−electronic integration [3]. As performance of electro-optical systems improves continuously, traditional stabilization method cannot meet application requirements. How to meet high performance requirements for future electro-optical stabilized sighting system and how to solve "bottle neck" of improving stabilization accuracy have become the focus of theoretical and engineering study. This chapter studies the application of FSM based on piezoelectric ceramic into high accuracy airborne stabilized sighting system to compensate residual errors of integrated stabilization platform in order to achieve higher stabilization accuracy.

17.2 Working Principle of FSM Based on Piezoelectric Ceramic

FSM consists of light mirror, piezoelectric ceramic actuator, flexible hinge, sensor and circuits [4]. Figure 17.1 is the schematic drawing of FSM based on piezoelectric ceramic. Mirror is supported by five supporting points on substrate with four supporting points around which are piezoelectric actuators and every two of which are grouped as actuating mechanisms and the middle one as a fixed point to achieve stabilization function. Taking the example of *y* axis rotating, when it works, one actuator extends while one shortens which transforms linear motion into rotation motion through flexible hinge and makes pushing and pulling motion to drive mirror rotation around *y* axis. The rotation around *x* axis is achieved by the same method as *y* axis. Mirror has two DOF in azimuth and elevation.

17.3 System Architecture

The principle prototype of high accuracy airborne stabilized sighting system is shown in Fig. 17.2, which consists of coarse stabilized platform and FSM assembly. The coarse stabilized platform has architecture of outer azimuth and internal elevation with payload of electro-optical sensors placed in coarse stabilized platform. The structure of platform enables optical path to make optical hinge. FSM is also placed in coarse stabilized platform and installed in optical path which needs to be stabilized accurately. The mirror can be adjusted in two freedoms of azimuth and elevation with a range of five mrad.

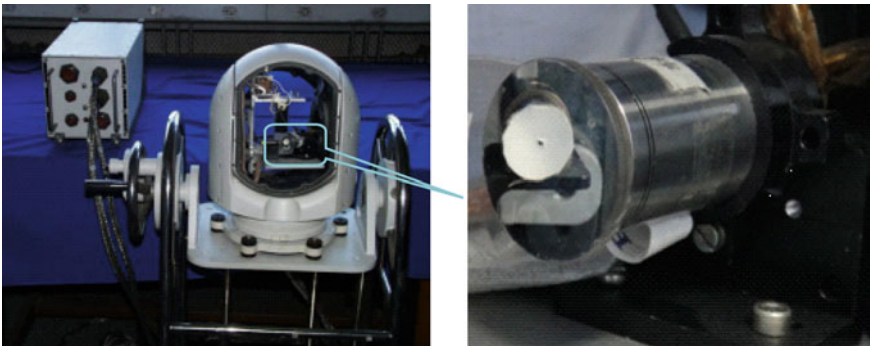


Fig. 17.2 Principle prototype of high accuracy line of sight stabilized system

When pedestal causes disturbance, disturbance moment occurs along stabilized axis which forces coarse stabilized platform drifted from original inertial space. Gyro installed in coarse stabilized platform senses the deviation and output the deviation signal which is processed to drive platform moment motors to produce control moment of equal magnitude and opposite direction with disturbance moment. In this way disturbance moment is counteracted and platform is stabilized. As coarse stabilized system is system with error, its stabilized error is far away from the requirements. Thus an accurate sensor is installed in system to sense the relative error of LOS and residual error of gyro is compensated through FSM assembly to achieve high accurate stabilization.

17.4 Design of System Control Loop

The schematic of servo control loop for high accuracy airborne stabilized sighting system is shown in Fig. 17.3 [5]. W_1, W_2 are controller for FSM assembly and coarse stabilized platform respectively and W_3 is forward feedback controller; K_1, K_2 are transfer function for coarse stabilized platform and FSM assembly respectively; G_m is transfer function for motor. G_2 is transfer function for piezoelectric ceramic and mirror; G_P, G_g are transfer function for micro displacement sensor and gyro respectively; $\frac{T_{d1}}{J_g}$, $\frac{T_{d2}}{J_m}$ are ratios of disturbance moment acted on coarse stabilized platform and mirror with their moment of inertia respectively; $\dot{\theta}_{\text{cmd}}$, θ_{los} are command input angle velocity and LOS output angle displacement. The coarse stabilized platform has big residual error due to big moment of inertia and low resonant frequency while small moment of inertia for mirror rotation enables high servo bandwidth and rapid response. Through rational design of control loop, the residual error of stabilized platform can be compensated. The system takes gyro as inertia reference and mirror assembly is used to make forward feedback for coarse platform. λ is optical coefficient and $\lambda = 2$ for planar mirror. Through design of controller of W_1, W_2, W_3, the steady error of LOS can be compensated completely.

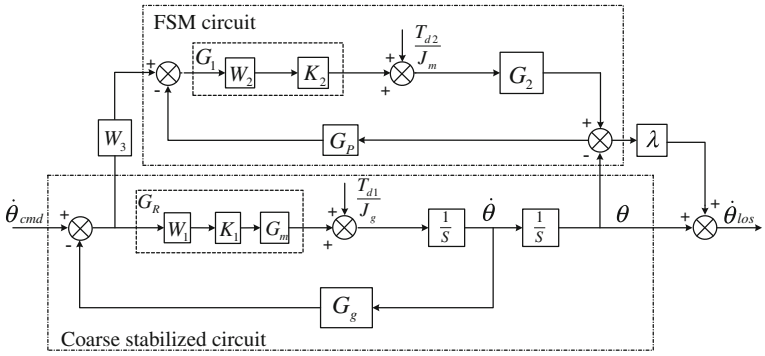


Fig. 17.3 Schematic of system servo control principle

Deducing from the above control schematic, it's not difficult to get the response for disturbance input and expected input. Table 17.1 shows the system response of ordinary coarse stabilized platform and the ones after introducing FSM assembly.

FSM assembly is a kind of forward feedback solution for ordinary coarse stabilized platform. The design of W_3 controller combining coarse stabilized platform with FSM assembly loop can make residual error of coarse stabilized platform compensated completely. According to response of system to disturbance of T_{d1}, in complete compensation:

$$-W_3 G_1 G_2 G_g \lambda s + (1 - \lambda + G_1 G_2 G_p) = 0. \tag{17.1}$$

While making $W_3 = \dfrac{k}{s}$, the condition of complete compensation is gotten as:

$$k = \frac{1 - \lambda + G_1 G_2 G_p}{G_1 G_2 G_g \lambda} \tag{17.2}$$

Thus,

$$W_3 = \frac{1}{s} \cdot \frac{1 - \lambda + G_1 G_2 G_p}{G_1 G_2 G_g \lambda} \tag{17.3}$$

The response of system to disturbance T_{d2} acted on FSM assembly can be rejected through increasing $G_1 G_2 G_p$ gain of FSM assembly loop. Compared with loop of coarse stabilized platform, the bandwidth of loop for FSM assembly can be made very high and its capability to reject disturbance can be made stronger.

Through simulation design of system, the frequency characteristics of loop for coarse stabilized platform and FSM assembly are given in Fig. 17.4 with performance parameters listed in Table 17.2. From this we can tell that due to mechanical structure, electronic noise, stabilization and other influencing factors, the control bandwidth of coarse stabilized platform is about 40 Hz and phase allowance $\gamma = 60^{\circ}$ while the control bandwidth of FSM assembly loop can be 380 Hz. The ratio of control bandwidth between coarse stabilized platform and FSM assembly

Table 17.1 System response of coarse stabilized platform and ones after introducing FSM assembly

	Coarse stabilized platform	Introducing FSM assembly
$\frac{\theta_{\text{Los}}}{T_{d1}/J_g}$	$\frac{1}{s^2 + G_R G_g s}$	$\frac{-W_3 G_1 G_2 G_g \lambda s + (1 - \lambda + G_1 G_2 G_p)}{s(1 + G_1 G_2 G_p)(s + G_R G_g)}$
$\frac{\theta_{\text{Los}}}{T_{d2}/J_m}$	–	$\frac{\lambda G_2}{s + G_1 G_2 G_p}$
$\frac{\theta_{\text{Los}}}{\dot{\theta}_{\text{cmd}}}$	$\frac{G_R}{s^2 + G_R G_g s}$	$\frac{W_3 G_1 G_2 \lambda s^2 + G_R(1 - \lambda + G_1 G_2 G_p)}{s(1 + G_1 G_2 G_p)(s + G_R G_g)}$

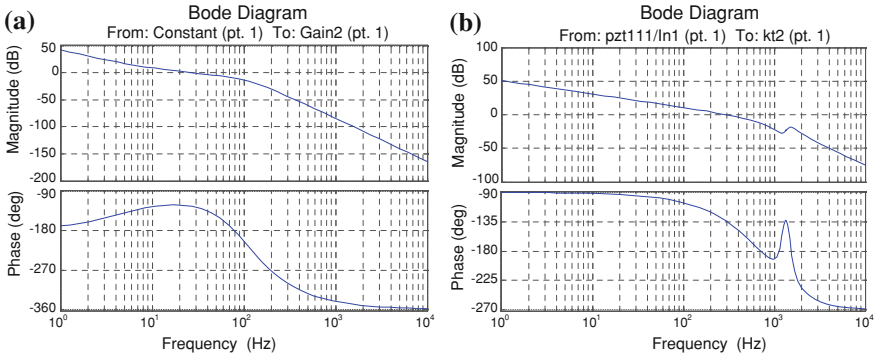


Fig. 17.4 Performance of control loop for coarse stabilized platform (**a**) and FSM assembly (**b**)

Table 17.2 Performance simulation parameter of control loop

	Coarse stabilized platform	FSM assembly loop
Phase margin (°)	60	53
Amplitude margin (dB)	7.22	12.4
Bandwidth (Hz)	40	381

is: $\omega_{\text{FSM}} \geq 9\omega$. The bigger the ratio, the stronger the error rejection capability can become and better the system stabilization performance can achieve.

17.5 System Performance Analyses

We make modeling to high accuracy airborne stabilized sighting system, simulate the above control loop and add the following disturbance moment:

(1) linear disturbance

The model is from linear vibration spectrum of a helicopter. The disturbance goes through shock absorber and mass imbalance coefficient and finally acts on coarse stabilized platform. The linear vibration spectrum is shown in Fig. 17.5a.

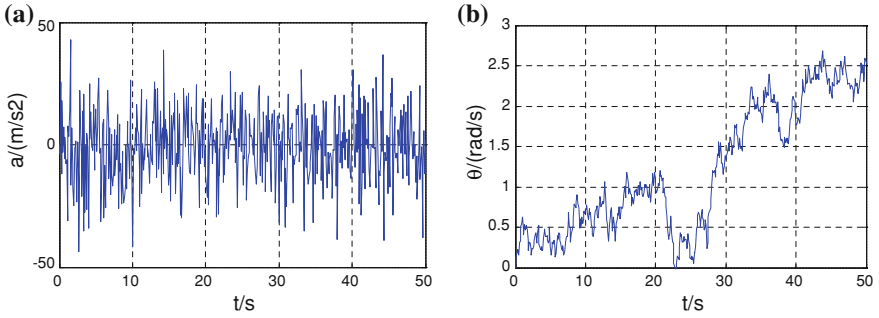


Fig. 17.5 Disturbance model **a** linear disturbance model, and **b** angle disturbance model

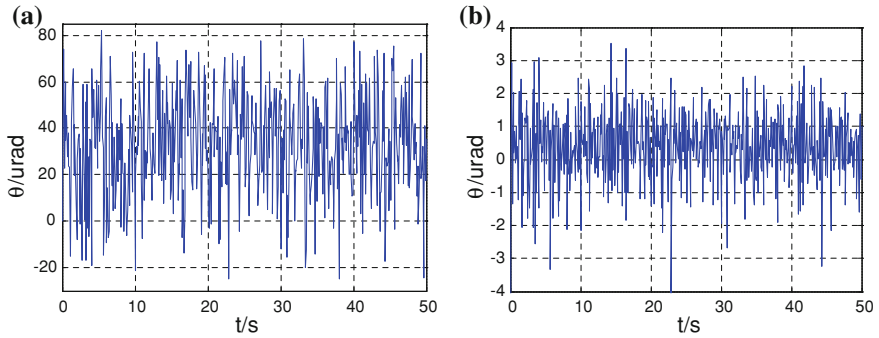


Fig. 17.6 Stabilization error of Coarse stabilized platform (**a**) and ones after introducing FSM assembly (**b**)

(2) Angle disturbance

The model is from angle vibration spectrum of a helicopter as shown in Fig. 17.5b.

(3) Friction

The model is classical friction model. The maximum static friction moment is 0.2 N m while dynamic friction moment is 0.425 times as big as static friction.

(4) Elastic moment

The elastic moment of cable is 0.1 N m/rad.

Simulation is made in SIMULINK toolbox of MATLAB software, five order *Runge–Kuta* arithmetic is used and duration is 50 s. The step can be adjusted automatically. Through simulation, the error of coarse stabilized platform is shown in Fig. 17.6a. The stabilization error after FSM compensation is shown in Fig. 17.6b. The simulation results show that stabilization error of coarse stabilization error is 110 μrad while the system stabilization error after FSM compensation is smaller than 8 μrad and the stabilization accuracy is improved by 14 times compared with coarse stabilized platform.

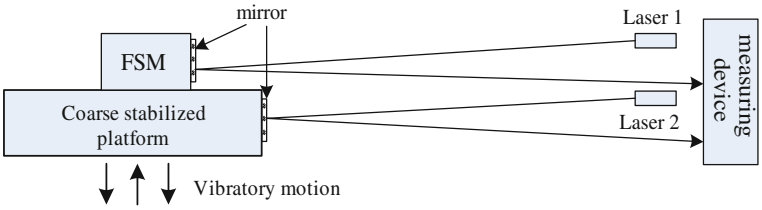


Fig. 17.7 Schematic of system measurement

After simulation analyses, the system parameters are used in prototype shown in Fig. 17.2 and the stabilization accuracy is measured. The measurement principle is shown in Fig. 17.7. The measurement of stabilization error for coarse stabilization platform is 116 μrad while the measurement of stabilization error after FSM compensation is smaller than 9 μrad. The stabilization accuracy has been improved by nearly 13 times which is almost the same as the simulation results.

17.6 Conclusion

The application of FSM assembly has changed the shortcomings of big moment of inertia and small bandwidth of traditional optical servo system and improved greatly the accuracy and bandwidth. FSM based on piezoelectric ceramic is used in airborne stabilized sighting system and multi-loop control is designed which not only improves the system stabilization accuracy by a magnitude but also enables integration and miniaturization of airborne stabilized sighting system.

References

1. Ming JI (1996) Research on high accuracy stabilization technology for integrated multi-spectrum electro-optical system. J Appl Opt 17(3):55–60
2. Zhang Z, Fan D, Fan S (2006) Servo system design for E-O stabilization and tracking devices. Opt Precis Eng 14(4):681–688
3. Marth H, Donat M (1991) Latest experience in design of piezoelectric driven fine steering mirrors. SPIE Act Adapt Opt Compon 1543:248–261
4. Sweeney M, Rynkowski G et al (2002) Design considerations for fast steering mirrors (FSMs). Proc SPIE Opt Scan 4773:63–73
5. Watkins RJ, Chen H-J, Agrawal NB et al (2004) Optical beam jitter control. Proc SPIE 5338:204–213

Chapter 18
Decoupling Control of MIMO System Using Neural Network Based on APSO

Shufang Sun, Jiahai Zhang and Jianhui Wang

Abstract To improve the decoupling and dynamic property of MIMO strong-coupling system, this chapter presents a modification of the particle swarm optimization algorithm (APSO) intended to combat the problem of premature convergence observed in many applications of PSO. Then the algorithm is used for training decoupling neural network. To prove the training outcome a typical MIMO system is used to be decoupled using the neural network trained by APSO. Finally PID controller is used to control the decoupled MIMO system. In the simulation experiment, the basic PSO and APSO is used for training decouple neural network separately; the controlling result indicated that neural network training by APSO is more stabilized and effective.

Keywords Neural network · Decoupling · APSO · MIMO

18.1 Introduction

In the field of industrial there are many multi-variable systems that have multi-input and multi-output. In those systems the input variables influence each other, at the same time one output influences the other output, that is to say

S. Sun (✉) · J. Zhang
College of Electric and Automatic Engineering, Sanjiang University,
Nanjing 210012, China
e-mail: sunshufang_2003@163.com

J. Wang
College of Information Engieering, Northeastern University, Shenyang, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_18,

they are coupled. The controlling of multi-variable systems has been a difficult problem.

Neural network has a strong ability of parallel processing, approximation and fast adaptive [1–3]. However before using neural network the connection weights and the thresholds must be settled. The traditional methods to determine the parameters are easy to fall into local optimum and sensitive to the initial condition or slow convergence.

In order to get satisfying decoupling neural network parameters a new adaptive PSO algorithm is proposed. The decoupling network is used to decouple a multi-variable coupled system [4].

18.2 BP Neural Network

BP neural network is a kind of multi-layer feed forward neural network with one input layer, one or several hidden layers and one output layer with every layer having a number of neurons. It is one of the most widely used neural networks.

In BP neural network between the left and right layers the neurons are fully connected. In the learning process when there are inputs in the input layers, the neuron is activated, and the value is transmitted from the input layer to the output layer. In the output layer there are corresponding network outputs in response to the input mode. Then the network weights and the thresholds are updated following the principle that reduces the error between the desired output and the actual output. The active function is *S*-type function in the BP neural network hidden layer, while the output neuron activation function can be *S*-type function or linear function. The typical *S*-type function is as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{18.1}$$

where x is the input weight function of the neurons.

18.3 The New Adaptive PSO Algorithm

18.3.1 Basic PSO Algorithm

The PSO algorithm is an evolutionary computation technique introduced by Kennedy and Eberhart in 1995. The underlying motivation for the development of PSO was the social behavior of animals such as bird flocking. In PSO, each individual of the population, called particle, has an adaptable velocity, according to the search space which it moves over. Each particle keeps track of its coordinate in hyperspace, associated with the solution (flatness value) it has achieved so far.

This value is called p_{best}. Another "best" value is called g_{best} that is obtained so far by any particle in the population and stored the overall best value.

In the basic version of the PSO algorithm each particle in the population manipulated according to the following assignment statements:

$$v_{id}^{t} = wv_{id}^{t-1} + c_1 r_1(\,)(p_{id} - x_{id}^{t-1}) + c_2 r_2(\,)(p_{gd} - x_{id}^{t-1}) \tag{18.2}$$

$$x_{id}^{t} = x_{id}^{t-1} + v_{id}^{t} \triangle t \tag{18.3}$$

where v_{id}^{t} and x_{id}^{t} are the velocity and position of the ith particle in the tth iteration; p_{id} is the best position the ith particle has accomplished at the $(t-1)$th iteration, and p_{gd} is the global best position achieved in the particle at the $(t-1)$th iteration. c_1 and $c_2$are two positive constants called acceleration constants. r_1 and r_2 are two different random numbers in the range of 0 to 1.

18.3.2 A New Adaptive PSO Algorithm

The basic PSO algorithm has disadvantages such as convergence to local optimum and hard to get higher convergence accuracy. To solve the problems an adaptive inertia weight and the maximum velocity is introduced into the basic PSO algorithm to improve and restrict PSO algorithm. At the same time dynamic neighborhood and linearity congruence-based adequate distribution initial strategy is used to establish an adaptive PSO.

The maximum velocity $v_{\max}$ determines the maximum change one particle can take during iteration, and determines the precision between current position and the global best position. If $v_{\max}$ is large value, the particle may fly beyond the best solution; If $v_{\max}$ is small value, particle cannot precede enough searches outside the partial good zone and sinks into the local optimized value. Usually we set the range of the particle as $v_{\max}$ and unified maximum velocity can also be set up, and can set the each dimension maximum velocity $v_{\max}$ according to dimension.

The inertia weight w keeps the movement inertial for the particle. It describes the influence of the previous velocity to the current velocity, makes it have the trend to extend the search space and have the ability to explore the new district, and there is the function to adjust the rate of velocity of particle. The inertia weight is decreased linearly from 0.9 to 0.4.

Linearity variety of the w:

$$w = w_{\max} - \left(w_{\max} - w_{\min}/G_{\max}\right)G \tag{18.4}$$

where: $w_{\max}$ is the maximum inertia weight, usually $w_{\max} = 0.9$; $w_{\min}$ is the minimum inertia weight, usually $w_{\min} = 0.4$; $G_{\max}$ is the maximum number of iteration; G is the current number of iterations.

In order to increase the searching precise when the Algebra is increasing, a variational neighborhood operator is introduced. The specific procedure is: first computing the distance between the candidate individual and all the other individuals, there dist (l) is the distance between the l one. The maximum distance is , and define a fraction namely frac $= 0.6 + 3.0G_n/G_{\max}$ which is concerned with current generation G_n If frac > 0.9 and frac $>$ dist$(l)/$ max inst, l_{best} is used in searching process, else g_{best} will be used.

In a word the basic step of adaptive algorithm is as follows:

(1) Setting parameters: m, $G_{\max}$, ε c_1 c_2 w ($w_{\max}$ and $w_{\min}$);
(2) Initialing population, the initial velocity and the position of particles with a series of random number;
(3) Computing the fitness value for each particle f_i;
(4) If f_i the current position is better than the position of the individual optimum, then p_{best} = the current position;
(5) If f_i the current position is better than the position of the swarm optimum, then g_{best} = the current position;
(6) Updating the velocity: Computing the velocity of the next generation according to Eq. (18.2), and using maximum velocity to restrict the updating velocity,

$$\begin{aligned} &\text{if } v_{id} < -v_{\max,d}, \quad \text{then } v_{id} = -v_{\max,d}, \\ &\text{else if } v_{id} > v_{\max,d}, \quad \text{then } v_{id} = v_{\max,d}; \end{aligned}$$

(7) Updating the position: Computing the next generation position according to Eq. (18.3), and in response to the position span restrict the new position;
(8) Decreasing the inertia weight according to Eq. (18.4);
(9) Computing the distance between the candidate individual and all other individuals, if frac > 0.9 and frac $>$ dist$(l)/$ max inst, searching l_{best}, else will be used g_{best};
(10) If the adaptive threshold of fitness value is satisfied, then stop; otherwise continue;
(11) If the maximum number of iterations does not reach the setting value, then go back to step (3), otherwise end.

18.4 Training Decoupling Neural Network by APSO

18.4.1 Coding the Neural Network Parameter in APSO

Training neural network in fact is a process of weight and threshold searching and optimizing using global searching capability. The self-learning and self-organizing ability of neural network is improved. In the decoupling model of neural network based on APSO, the weight and thresholds between the neurons, the error weight

of the hidden layer and the error weight of the output layer are defined as the positions of the particles in the algorithm [5].

18.4.2 Determine the Fitness Function

The purpose of neural network learning is expected to minimize the error between the network's actual output and the desired output. When training the neural network using APSO the sample mean square deviation is chosen as fitness function. Namely,

$$E = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i')^2 \tag{18.5}$$

where N is the total number of the training sample; y_i is the ideal output of the training sample i; y' is the actual output of the training sample i.

The APSO searching process is to find the minimum fitness value of the particle according to the iterative equation [6].

18.4.3 Training Decoupling Neural Network by APSO

According to the APSO algorithm which is introduced above, the searching step is as follows:

(1) Setting parameters: m, $G_{\max}$, ε c_1 c_2 w ($w_{\max}$ and $w_{\min}$);
(2) Initialing population, setting velocity vector, and then position vector which is a coded by weight of neural network;
(3) Step phrase is inputted and then computing the fitness value of each particle f_i by Eq. (18.6);
(4) If $f_i < f_{p_{\text{best}}}$, then p_{best} = current position; If $f_i < f_{g_{\text{best}}}$, then g_{best} = the current position;
(5) Updating the velocity: computing the velocity of the next generation according to Eq. (18.2), and using maximum velocity to restrict velocity,

$$\begin{aligned} &\text{if } v_{id} < -v_{\max,d}, \text{ then } v_{id} = -v_{\max,d}, \\ &\text{else if } v_{id} > v_{\max,d}, \text{ then } v_{id} = v_{\max,d}; \end{aligned}$$

(6) Updating the position: computing the position of the next generation according to Eq. (18.3), and in response to the position span restrict the new position;
(7) Decreasing the inertia weight according to the Eq. (18.4);

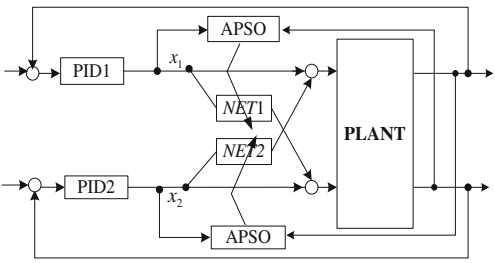


Fig.18.1 The structure of the decoupling system

(8) Computing distance between the candidate individual and all other individuals, if frac > 0.9 and frac $>$ dist$(l)/$ max inst, searching l_{best}, else will be used g_{best};
(9) If the threshold of fitness value is satisfied, then stop; otherwise continue;
(10) If the maximum number of iterations does not reach the setting value, then go back to step (3), otherwise end.

At the end of the training procedure, all the parameters should be recorded. There the couple system has been decoupled and coming into two SISO systems.

18.5 The Simulation of MIMO Decoupling Control System by Neural Network APSO based

18.5.1 The Structure of MIMO Decoupling Control System

A typical two-input two-output system is proposed to test the decoupling control method based neural network [7]. The structure of the control system is given in Fig. (18.1). In the design system the controller, the decoupler, and the control object are included. The control object mathematical model is presented in Eq. (18.6) and also the ideal decoupling output of the mathematical function is as Eqs. (18.7) and (18.8). The decoupler is made of two parts. The first is built by two neural networks, namely NET1 and NET2. The second part is APSO which training the neural network. Although the system is two-input and two-output decoupling system, after the decoupling process, it will be a SISO system, and PID controller which is common and effective is used to control the decoupled system.

$$\begin{aligned} Y(k) = &\begin{bmatrix} 0.2424\,0.0997 \\ 0.1322\,0.4265 \end{bmatrix} Y(k-1) + \begin{bmatrix} 0.3426\,0.2105 \\ 0.1070\,0.2404 \end{bmatrix} Y(k-2) \\ &+ \begin{bmatrix} 0.0576\,0.0424 \\ 0.0295\,0.0751 \end{bmatrix} U(k-1) + \begin{bmatrix} 0.0509\,0.0319 \\ 0.0372\,0.0656 \end{bmatrix} U(k-2) \end{aligned} \tag{18.6}$$

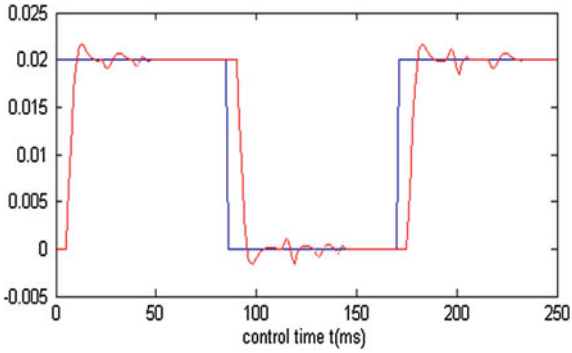


Fig.18.2 The control conclusion of basic PSO

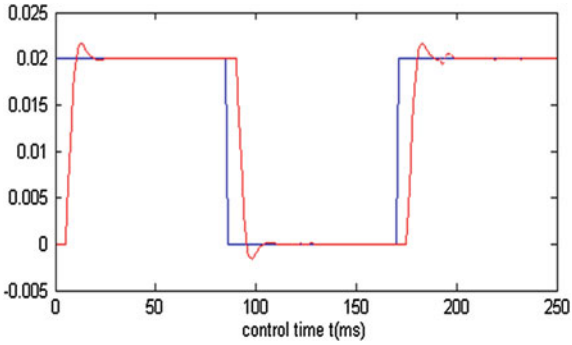


Fig.18.3 The control conclusion of APSO

$$y_1(k) = 0.6689y_1(k-1) + 0.4928y_1(k-2) - 0.1659y_1(k-3) - 0.0598y_1(k-4) + 0.0576u_1(k-1) + 0.0293u_1(k-2) - 0.0256u_1(k-3) - 0.0044u_1(k-4) \quad (18.7)$$

$$y_2(k) = 0.6689y_2(k-1) + 0.4928y_2(k-2) - 0.1659y_2(k-3) - 0.0598y_2(k-4) + 0.0751u_2(k-1) + 0.0530u_2(k-2) - 0.0329u_2(k-3) - 0.0190u_2(k-4) \quad (18.8)$$

18.5.2 Simulation of MIMO Decoupling Control System by Neural Network Based on APSO

In the experiment MATLAB7.0 is used to simulate the work system of the decoupling system. In order to decouple the generalized control object, first the controller should break with the object, so the neural network is connected with the control object. The structure of the decoupling neural network is one input layer, three hidden layers and one output layer.

The parameter of neural network is adjusted by adaptive PSO. In decoupling training process the input signal is phase step, namely when phase step signal carries in from x_1, there should be 0 in output y_2; when phase step signal carries in from x_2, there should be 0 in output y_1.

So the index function of decoupler NET1 is:

$$f_1 = \frac{1}{2}\sum_{k=0}^{N}[y_2(k)]^2 \tag{18.10}$$

Similarly, the index function of neural network *NET*2 is:

$$f_2 = \frac{1}{2}\sum_{k=0}^{N}[y_1(k)]^2 \tag{18.11}$$

The decouple steps of the two input two output system by neural network is as the former part stated. But when training the MIMO system the fitness function is Eqs. (18.10) and (18.11); namely in step (3), the computing process should use the two equations.

After the decoupling training, it connects the controller with decoupled system. That is to say the PID controller will control the SISO system which is got by decoupling. In simulation the given input is square wave. Figure 18.2 is the decoupling conclusion of basic PSO algorithm training neural network. Figure 18.3 is the decoupling conclusion of APSO algorithm training neural network. Also the output is delay to the input and the APSO is more effective and faster than the basic PSO.

18.6 Conclusion

In the experiment a new adaptive PSO algorithm is adopted to training neural network, and a decoupler is designed by using the neural network to decouple the MIMO decoupling system, and the after-decoupled system is controlled by PID controller. Through the simulation it is obvious that this control system can get a satisfying control result, and solve the intensive couple problem effectively.

References

1. Dai X, Zhang X, Liu G (2004) Decouple control of induction motor based on neural networks inverse. Proc Csee 24(1):112–117
2. Xu-hua SHI (2005) A design method for adaptive generalized predictive decoupling controller based on the a neural network. Acta Simulata Systematica Sinica 17(1):178–180
3. Dai X, Dan H, Zhang T, Zhang K (2003) ANN generalized inversion for the linearization and decoupling control of nonlinear systems. IEE Proc-Control Theor Appl 150(3):267–277

4. Dai XZ, He D, Fan LI, Li NH, Chen H (1999) Improved ANN αth-order inverse TCSC controller for enhancing power system transient stability. IEE Proc C Gener Transm Distrib 146(6):550–556
5. Aziz MFA, Hamed HNA, Shamsuddin SMH (2008) Augmentation of elman recurrent network learning with particle swarm optimization. In: Second Asia international conference on modelling & simulation. pp 625–630
6. Ge H-w, Li X-l, Liang Y-c, He X-d (2008) Immune PSO-based dynamic recurrent neural network for identifying and controlling nonlinear systems. J Jilin Univ Eng Technol Edn 38(4):858–864
7. Shu D, Feng C, Yin Y (1996) Neuron network decoupling control of multivariable system with generalized predicative controller and its application. Electric Drive 26(1):44–51

Chapter 19
Mean-Square Exponential Stability Control for Networked Control Systems with Interval Distribution Time Delays

Hejun Yao and Fushun Yuan

Abstract To study the problem of exponential stability control for a class of networked control systems with interval distribution time delays, a new approach is given to model the networked control systems with the stochastic time delays which is assumed to be satisfying a interval Bernoulli distribution. Based on linear matrix inequality approach, the mean-square exponential stability controller design method is presented, and the controller gain matrix is obtain by solving a linear matrix inequalities. Moreover, a Lyapunov functional is used, and some stack matrices, which bring much flexibility in solving LMI, are introduced during the proof. A numerical example is given to demonstrate the validity of the results.

Keywords Networked control systems · Stochastic delays · Linear matrix inequality (LMI)

19.1 Introduction

Networked control systems (NCSs) are systems where the feedback loop is closed via a communication network in which information, from various components such as sensors, controllers and actuators, is exchanged with limited bandwidth. NCSs have received increasing attentions in recent years due to their low cost, simple installation and maintenance and high reliability [1, 2].

However, the network itself is dynamic system that exhibits characteristics such as network-induced delays. The delays come from the time sharing of the

H. Yao (✉) · F. Yuan
School of Mathematics and Statistics, Anyang Normal University,
Henan 455002, China
e-mail: yaohejun@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_19,

communication medium as well as the computation time required for physical signal coding and communication processing. As is known, network-induced delays can degrade a system's performance and even cause system instability. Many researchers have studied stability analysis and controller design for NCSs [3, 4]. It is quite common in practice that the time delays occur in a random way, rather than a deterministic way. Based on a similar Bernoulli stochastic model, [5, 6] study NCSs with both sensor-to-controller and controller-to-actuator stochastic delays, and design the controller gain. One step stochastic delays or one stochastic packet dropout is considered in these papers.

The stabilization problem for a networked control system with Markov communication delays existing in both the system state and the mode signal is considered in [7]. The problem of the stabilization of NCSs with packet dropout is studied in [8]. A networked predictive control method for networked systems with stochastic delay and data dropout is proposed in [9, 10] to compensate the networked-induced delay. This paper considers the problem of mean-square exponential stability control for a class of networked control systems with interval distribution time delay. A new approach is given to model the networked control systems. Based on the LIM approach and Lyapunov stability theorem, the mean-square exponential stability condition is given.

19.2 Problem Formulation

Consider the following control system with delay

$$\begin{aligned} \dot{x}(t) &= Ax(t) + A_d x(t-d) + Bu(t) \\ x(t) &= \phi(t)\, t \in [-d,\, 0] \end{aligned} \tag{19.1}$$

where $x(t) \in R^n$ is the state vector, $u(t) \in R^m$ is the input vector, d is state delay $A, A_d \in R^{n\times n}$ are known real constant matrices, $B \in R^{n\times m}$ is input matrix, $\phi(t) \in R^n$ is the given initial state on $[-d,\, 0]$.

Throughout this note, we suppose that all the system's states are available for a state feedback control. In the presence of the control network, data transfers between the controller and the remote system, e.g., sensors and actuators in a distributed control system will induce network delay in addition to the controller proceeding delay. We introduce stochastic delay $\tau(t)$ to denote the network-induced delay. In this note we make the following assumptions:

- Assumption 1: Sensor and controller are clock-driven;
- Assumption 2: Actuator is event-driven.

We will design the state feedback controller

$$u(t) = Kx(t - \tau(t)) \tag{19.2}$$

Inserting the controller (19.2) into system (19.1), we obtain the closed system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + A_d x(t-d) + BKx(t-\tau(t)) \\ x(t) &= \psi(t) \quad t \in [-\overline{d}, 0] \end{aligned} \tag{19.3}$$

The initial condition of the state is supplemented as $x(t) = \psi(t)$, where $\psi(t)$ is a smooth function on $[-\overline{d}, 0]$, $\overline{d} = \max\{\tau, d\}$. Therefore, there exists a positive constant $\overline{\psi}$ satisfying

$$||\dot{\psi}(t)|| \leq \overline{\psi} \quad t \in [-\overline{d}, 0]$$

It is assumed that there exists a constant $\tau_1 \in [0, \tau]$ such that the probability of $\tau(t)$ taking values in $[0, \tau_1)$ and $[\tau_1, \tau]$ can be observed. In order to employ the information of the probability distribution of the delays, the following sets are proposed first

$$\Omega_1 = \{t : \tau(t) \in [0, \tau_1)\}, \quad \Omega_2 = \{t : \tau(t) \in [\tau_1, \tau]\}$$

Obviously, $\Omega_1 \cup \Omega_2 = R^+$ and $\Omega_1 \cap \Omega_2 = \Phi$.

Then we define two functions as:

$$h_1(t) = \begin{cases} \tau(t) & t \in \Omega_1 \\ 0 & t \notin \Omega_1 \end{cases}, \quad h_2(t) = \begin{cases} \tau(t) & t \in \Omega_2 \\ \tau_1 & t \notin \Omega_2 \end{cases} \tag{19.4}$$

Corresponding to $\tau(t)$ taking values in different intervals, a stochastic variable $\beta(t)$ is defined

$$\beta(t) = \begin{cases} 1 & t \in \Omega_1 \\ 0 & t \in \Omega_2 \end{cases} \tag{19.5}$$

where we suppose that $\beta(t)$ is a Bernoulli distributed sequence with $\Pr ob\{\beta(t) = 1\} = E\{\beta(t)\} = \beta$, where $\beta \in [0, 1]$ is a constant.

With (19.4, 19.5), we know that the systems (19.3) is equivalent to

$$\begin{aligned} \dot{x}(t) &= Ax(t) + A_d x(t-d) + \beta(t)BKx(t-h_1(t)) + (1-\beta(t))BKx(t-h_2(t)) \\ &= \bar{A}\xi(t) \\ x(t) &= \psi(t) \quad t \in [-\bar{d}, 0] \end{aligned} \tag{19.6}$$

where

$$\begin{aligned} \bar{A} &= [A\, A_d\, \beta(t)BK(1-\beta(t))BK] \\ \xi^T(t) &= [x^T(t), x^T(t-d), x^T(t-h_1(t)), x^T(t-h_2(t))] \end{aligned}$$

19.3 Main Results

Lemma 1 [2] *For any vectors a, b and matrices N, X, Y, Z with appropriate dimensions, if the following matrix inequality holds*

$$\begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix} \geq 0$$

then we have

$$-2a^T Nb \leq \inf_{X,Y,Z} \begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} X & Y-N \\ Y^T - N^T & Z \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Lemma 2 [4] *The LMI* $\begin{bmatrix} Y(x) & W(x) \\ * & R(x) \end{bmatrix} > 0$ *is equivalent to*

$$R(x) > 0,\ Y(x) - W(x)R^{-1}(x)W^T(x) > 0.$$

where $Y(x) = Y^T(x), R(x) = R^T(x)$ *depend on x.*

Theorem 1 *For the given constants* $\alpha > 0,\ 1 \geq \beta \geq 0$, *if there exist positive-definite matrices* $P, Q, R \in R^{n\times n}$ *and matrices* $K \in R^{m\times n}$, X, Y *with appropriate dimensions, such that the following matrix inequality holds*

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ * & \Theta_{22} \end{bmatrix} < 0 \tag{19.7}$$

where

$$\Theta_{11} = \begin{bmatrix} \mathrm{PA} + A^T P + Q + 2\alpha P & \mathrm{P\bar{A}}_d + \tau_2 X_{12} + \tau_2 A^T \mathrm{RA}_d \\ +\tau_2 X_{11} + \tau_2 A^T \mathrm{RA} & \\ * & -e^{-2\alpha d} Q + \tau_2 X_{22} + \tau_2 A_d^T \mathrm{RA}_d \end{bmatrix}$$

$$\Theta_{12} = \begin{bmatrix} P\beta\mathrm{BK} + Y_1 + \tau_2 X_{13} + \tau_2 A^T R\beta\mathrm{BK} & P(1-\beta)\mathrm{BK} + \tau_2 X_{14} - Y_1 + \tau_2 A^T R(1-\beta)\mathrm{BK} \\ Y_2 + \tau_2 X_{23} + \tau_2 A_d^T R\beta\mathrm{BK} & \tau X_{24} - Y_2 + \tau_2 A_d^T R(1-\beta)\mathrm{BK} \end{bmatrix}$$

$$\Theta_{22} = \begin{bmatrix} \tau_2 X_{33} + Y_3 + Y_3^T + \tau_2 K^T B^T R\beta\mathrm{BK} & -Y_3 + Y_4^T + \tau_2 X_{34} \\ * & \tau_2 X_{44} - Y_4 - Y_4^T + \tau_2 K^T B^T R(1-\beta)\mathrm{BK} \end{bmatrix}$$

with the controller (19.2), *the network control systems* (19.6) *is mean-square exponentially stable.*

Proof Choose a Lyapunov functional candidate for the system (19.6) as follows: system (19.6) as follows:

$$\begin{aligned} V(t) = {} & x^T(t)Px(t) + \int_{t-d}^{t} x^T(s)Qe^{2\alpha(s-t)}x(s)ds \\ & + \int_{-\tau_2}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)dsd\theta \end{aligned}$$

where P, Q, R positive-definite matrices in Theorem 1. Then, along the solution of system (19.6) we have

$$\begin{aligned}
&\dot{V}(t) + 2\alpha V(t) \\
&\quad = 2x^T(t)P\dot{x}(t) + x^T(t)Qx(t) - x^T(t-d)Qe^{-2\alpha d}x(t-d) \\
&\qquad + \tau\dot{x}^T(t)R\dot{x}(t) + 2\alpha x^T(t)Px(t) - \int_{t-\tau_2}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)ds
\end{aligned} \tag{19.8}$$

With

$$x(t-h_1(t)) - x(t-h_2(t)) - \int_{t-h_2(t)}^{t-h_1(t)} \dot{x}(s)ds = 0$$

For any $4n \times n$ matrix

$$N = \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix}$$

We know

$$0 = \xi^T(t)N\left[x(t-h_1(t)) - x(t-h_2(t)) - \int_{t-h_2(t)}^{t-h_1(t)} \dot{x}(s)ds\right] \tag{19.9}$$

With Lemma 1 and (19.9), we obtain

$$\begin{aligned}
0 &\le 2\xi^T(t)N[x(t-h_1(t)) - x(t-h_2(t))] \\
&\quad + \int_{t-h_2(t)}^{t-h_1(t)} \begin{bmatrix} \xi(t) \\ \dot{x}(s) \end{bmatrix}^T \begin{bmatrix} X & Y-N \\ Y^T - N^T & Re^{2\alpha(s-t)} \end{bmatrix} \begin{bmatrix} \xi(t) \\ \dot{x}(s) \end{bmatrix} ds \\
&\le 2\xi^T(t)Y[x(t-h_1(t)) - x(t-h_2(t))] + \tau_2\xi^T(t)X\xi(t) \\
&\quad + \int_{t-\tau_2}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)ds
\end{aligned} \tag{19.10}$$

Inserting (19.10) into (19.8), we have:

$$\begin{aligned}
\dot{V}(t) + 2\alpha V(t) &\le x^T(t)[PA + AP + Q + 2\alpha P]x(t) \\
&\quad + 2x^T(t)PA_dx(t-d) + 2x^T(t)P\beta(t)BKx(t-h_1(t)) \\
&\quad + 2x^T(t)P(1-\beta(t))BKx(t-h_2(t)) \\
&\quad - x^T(t-d)Qe^{-2\alpha d}x(t-d) + 2\xi^T(t)Y[0 \quad 0 \quad I \; -I]\xi(t) \\
&\quad + \tau_2\xi^T(t)X\xi(t) + \tau_2\dot{x}^T(t)R\dot{x}(t)
\end{aligned} \tag{19.11}$$

Where

$$\dot{x}^T(t)R\dot{x}(t) = \tau_2\xi^T(t)\begin{bmatrix} A^T\mathrm{RA} & A^T\mathrm{RA}_d & A^TR\beta(t)BK & A^TR(1-\beta(t))BK \\ * & A_d^T\mathrm{RA}_d & A_d^TR\beta(t)BK & A_d^TR(1-\beta(t))BK \\ * & * & \beta^2(t)K^TB^T\mathrm{RB}K & \beta(t)(1-\beta(t))K^TB^T\mathrm{RBK} \\ * & * & * & (1-\beta(t))^2K^TB^T\mathrm{RB} \end{bmatrix}\xi(t) \tag{19.12}$$

Obviously

$$2\xi^T(t)\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix}[0 \quad 0 \quad I \quad -I]\xi(t) = \xi^T(t)\begin{bmatrix} 0 & 0 & Y_1 & -Y_1 \\ * & 0 & Y_2 & -Y_2 \\ * & * & Y_3+Y_3^T & -Y_3+Y_4^T \\ * & * & * & -Y_4-Y_4^T \end{bmatrix}\xi(t) \tag{19.13}$$

Inserting (19.12–19.13) into (19.11), we obtain

$$E\{\dot{V}(t)+2\alpha V(t)\} \le \sum_i^r\sum_j^r \mu_i(z(t))\mu_j(z(t))\xi^T(t)\Theta\xi(t)$$

With matrix inequality (19.7), we know

$$E\{\dot{V}(t)\} < -2\alpha E\{V(t)\}$$

therefore

$$E\{V(t)\} < E\{V(0)\}\mathrm{e}^{-2\alpha t} \le [\lambda_{\max}(P)+d\lambda_{\max}(Q)+\tau\lambda_{\max}(R)\bar{\psi}^2]E\{||\psi(t)||^2\}\ \mathrm{e}^{-2\alpha t} \tag{19.14}$$

Obviously

$$E\{V(t)\} \ge \lambda_{\min}(P)E\{||x(t)||^2\} \tag{19.15}$$

From (19.14–19.15), we obtain

$$E\{||x(t)||\} < \sqrt{\frac{\lambda_{\max}(P)+d\lambda_{\max}(Q)+\tau\lambda_{\max}(R)\bar{\psi}^2}{\lambda_{\min}(P)}}E\{||\psi(t)||\}\ \mathrm{e}^{-\alpha t}$$

With the Lyapunov stability theorem and the above inequality, we know that the system (19.6) is mean-square exponentially stable.

Theorem 2 *For the given constants* $\alpha > 0, 1 \ge \beta \ge 0$, *if there exist positive-definite matrices* $\bar{P}, \bar{Q}, \bar{R} \in R^{n\times n}$ *and matrix* $\bar{K} \in R^{m\times n}$, $\bar{X}, \bar{Y}$ *with appropriate dimensions, such that the following linear matrix inequality holds*

$$\Xi = \begin{bmatrix} \Xi_{11} & \Xi_{12} \\ * & \Xi_{22} \end{bmatrix} < 0 \tag{19.16}$$

where

$$\Xi_{11} = \begin{bmatrix} A\bar{P} + \bar{P}A^T + \bar{Q} & A_d\bar{P} + \tau_2\bar{X}_{12} & \beta B\bar{K} + \tau_2\bar{X}_{13} + \bar{Y}_1 \\ +2\alpha\bar{P} + \tau_2\bar{X}_{11} & & \\ * & -e^{-2\alpha d}\bar{Q} + \tau_2\bar{X}_{22} & \tau_2\bar{X}_{23} + \bar{Y}_2 \\ * & * & \tau_2\bar{X}_{33} + \bar{Y}_3 + \bar{Y}_3^T \end{bmatrix}$$

$$\Xi_{12} = \begin{bmatrix} (1-\beta)B\bar{K} - \bar{Y}_1 + \tau_2\bar{X}_{14} & \tau_2\beta\bar{P}A^T & \tau_2(1-\beta)\bar{P}A^T \\ \tau_2\bar{X}_{24} - \bar{Y} & \tau_2\beta\bar{P}A_d^T & \tau_2(1-\beta)\bar{P}A_d^T \\ \tau_2\bar{X}_{34} + \bar{Y}_4^T - \bar{Y}_3 & \tau_2\beta\bar{K}^T B^T & 0 \end{bmatrix}$$

$$\Xi_{12} = \begin{bmatrix} \tau_2\bar{X}_{44} - \bar{Y}_4 - \bar{Y}_4^T & 0 & \tau_2(1-\beta)\bar{K}^T B^T \\ & -\tau_2\beta\bar{R} & 0 \\ & & -\tau_2(1-\beta)\bar{R} \end{bmatrix}$$

with the controller $u(t) = \bar{K}\bar{P}^{-1}x(t)$, *the* systems (19.6) is mean-square exponentially stable.

Proof

$$\Theta = \Theta_0 + \alpha_1^T \frac{1}{\tau_2\beta} R^{-1}\alpha_1 + \alpha_2^T \frac{1}{\tau_2(1-\beta)} R^{-1}\alpha_2$$

where

$$\Theta_0 = \begin{bmatrix} \mathrm{PA} + A^T P + Q + 2\alpha P & \mathrm{PA}_d + \tau_2 X_{12} & P\beta\mathrm{BK} + Y_1 + \tau_2 X_{13} & P(1-\beta)\mathrm{BK} + \tau_2 X_{14} - Y_1 \\ +\tau_2 X_{11} & & & \\ * & -e^{-2\alpha d}Q + \tau_2 X_{22} & Y_2 + \tau_2 X_{23} & \tau_2 X_{24} - Y_2 \\ * & * & \tau_2 X_{33} + Y_3 + Y_3^T & -Y_3 + Y_4^T + \tau_2 X_{34} \\ * & * & * & \tau_2 X_{44} - Y_4 - Y_4^T \end{bmatrix}$$

$$\alpha_1 = [\,\tau_2\beta\mathrm{RA} \quad \tau_2\beta RA_d \quad \tau_2\beta\mathrm{RBK} \quad 0\,]$$
$$\alpha_2 = [\,\tau_2(1-\beta)RA \quad \tau_2(1-\beta)\mathrm{RA}_d \quad 0 \quad \tau_2(1-\beta)\mathrm{RBK}\,]$$

With Lemma 2, we know that inequality $\Theta < 0$ is equivalent to

$$\Sigma = \begin{bmatrix} PA + A^T P + Q & PA_d + \tau_2 X_{12} & P\beta BK + Y_1 + \tau_2 X_{13} & P(1-\beta)BK + \tau_2 X_{14} - Y_1 & \tau_2\beta A^T R & \tau_2(1-\beta)A^T R \\ +2\alpha P + \tau_2 X_{11} & & & & & \\ * & -e^{-2\alpha d}Q + \tau_2 X_{22} & \tau_2 X_{23} + Y_2 & \tau_2 X_{24} - Y_2 & \tau_2\beta A_d^T R & \tau_2(1-\beta)A_d^T R \\ * & * & \tau_2 X_{33} + Y_{13} + Y_{13}^T & \tau_2 X_{34} + Y_4^T - Y_3 & \tau_2\beta K^T B^T R & 0 \\ * & * & * & \tau_2 X_{44} - Y_4 - Y_4^T & 0 & \tau_2(1-\beta)K^T B^T R \\ * & * & * & * & -\tau_2\beta R & 0 \\ * & * & * & * & * & -\tau_2(1-\beta)R \end{bmatrix} < 0 \tag{19.17}$$

Pre- and Post-multiplying the inequality (19.17) by $diag\{P^{-1}\ P^{-1}\ P^{-1}\ P^{-1}\ R^{-1}\ R^{-1}\}$, and giving some transformations:

$$\bar{P} = P^{-1}\ \bar{Q}{=}P^{-1}QP^{-1}\ \bar{K} = KP^{-1}\ \bar{X} = P^{-1}XP^{-1}\ \bar{Y} = P^{-1}YP^{-1}\ \bar{R} = R^{-1}$$

we know that inequality $\Sigma < 0$ is equivalent to (19.16). Therefore, the linear matrix inequality (19.16) is equivalent to (19.7). With Theorem 1, the systems (19.6) is mean-square exponentially stable.

19.4 Simulation

Consider the networked control systems in the form of (19.7), where

$$A = \begin{bmatrix} -3 & 1 \\ 0 & -4 \end{bmatrix},\ A_d = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix},\ B = \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix},\ \tau = 1,\ \tau_1 = 0.5,\ \beta = 0.5,$$

$\alpha = 0.1$, $d = 0.1$. Solving the linear matrix inequality (19.16), we can obtain the gain matrix K of the stabilizing controller $u(t)$

$$K = [-0.7645 \quad 2.5692]$$

From the Theorem 2, we know that the systems (19.6) is mean-square exponentially stable.

19.5 Conclusion

This paper considers the problem of mean-square exponential stability control for a class of networked control systems with interval distribution time delay. Based on the Lyapunov stability theorem, a sufficient condition and the controller design approach are given in term of LMI.

Acknowledgments The author would like to thank the associate editor and the anonymous reviewers for their constructive comments and suggestions to improve the quality and the presentation of the paper. This work was supported by National Nature Science Foundation under Project Code 61073065; Henan Province Nature Science Foundation under Project Code 092300410145.

References

1. Xia Y, Fu M, Liu B, Liu G (2009) Design and performance analysis of networked control systems with random delay. J Sys Eng Electron 20:807–822
2. Xia X, Zhang D, Zheng L, Zheng N (2008) Modeling and stabilization for a class of nonlinear networked control systems, a T-S fuzzy approach. Progr Nat Sci 18:1031–1037

3. Zhang H, Yang J, Su C (2007) T-S fuzzy model based robust H_∞ design for networked control systems with uncertainties. IEEE Trans Indus Inf 3:289–301
4. Jiang X, Han Q (2008) On designing fuzzy controllers for a class of nonlinear networked control systems. IEEE Trans Fuzzy Syst 16:1050–1060
5. Schenato L (2006) Optimal estimation in networked control systems subject to random delay and packet loss. In: Proceedings of the 45th IEEE conference on decision and control. San Diego, pp 5615–5620
6. Li JG, Yuan JQ, Lu JG (2010) Observer-based H_∞ control for networked nonlinear systems with random packet loss. ISA Trans 49:39–46
7. Liu M, Hoa DW (2009) Stabilization of markov jump linear system over networks with random communication delay. Automatica 45:416–421
8. Xiong JL, Lam J (2007) Stabilization of linear systems over networks with bounded packet loss. Automatica 43:80–87
9. Zhao Y, Liu G, Rees D (2008) A predictive control based approach to networked wiener systems. Int J Innov Comp Inf Control 4:2793–2802
10. Xia Y, Liu GP, Fu M, Rees D (2009) Predictive control of networked systems with random delay and data dropout. IET Control Theor Appl 3:1476–1486

Chapter 20
Designing and Realizing of Measurement Temperature System Based on Kingview

Zhang Yingqi, Wang Junfeng, Ai Yongle, Xu Shuai and Liu Xuebin

Abstract This paper designs the temperature collecting system by combination Kingview with digital temperature sensor DS18B20 based on 51 SCM. First, this paper illustrates the main hardware circuit of the temperature measuring system. Second, the engineering temperature measurement system software is developed with the KingView software. The characteristics of the system are realized temperature curve real-time display, historical curve enquiry, alarm warning of temperature more than the value and so on.

Keywords Kingview · Historical curve enquiry · Temperature sensor

20.1 Introduction

Two kinds of temperature collecting systems are widely used at home and abroad at present. One is that the temperature is measured by the temperature sensor DS18B20, then the temperature data is collected by the SCM and displayed by

Z. Yingqi (✉) · W. Junfeng · A. Yongle · X. Shuai · L. Xuebin
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: Zhang_Yingqi_zyq@hpu.edu.cn

W. Junfeng
e-mail: Wang_Junfeng_zyq@hpu.edu.cn

A. Yongle
e-mail: Ai_Yongle_zyq@hpu.edu.cn

X. Shuai
e-mail: Xu_Shuai_zyq@hpu.edu.cn

L. Xuebin
e-mail: Liu_Xuebin_zyq@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_20,

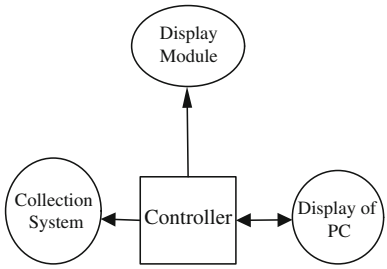


Fig. 20.1 Block diagram of system scheme

LED or LCD, whose disadvantage is inflexible [1, 2]. The other is that the temperature data measured by DS18B20 is transferred to the PC through remote transmission of PLC and processed by hand or man—machine software, whose disadvantages are the poor reliability, difficulty of exchanging data with the outside and high cost. Based on the above analysis, 51 SCM, Kingview and temperature sensor are adopted in the design of temperature collecting system, whose advantage is that the temperature data can be processed and displayed by LED and Kingview, realizing functions such as real-time display of temperature curve, inquiry of history curve, etc.

20.2 Scheme Design

Four parts (controller, collecting system, display module and display of the PC) as shown in Fig. 20.1 are included in the design. Temperature collecting system is mainly responsible for the information collection and conversion, and then wait for the call and control controller is the hardware core and responsible for receiving and processing data that is to be displayed and delivered to the PC. The display module is mainly responsible for displaying the collected data. The display of the PC is to process and display the data from the controller with the Kingview [3].

20.3 Hardware Circuit Design

The design of the main hardware is the temperature collecting circuit, LED display and level conversion circuit used in the serial communication [4].

20.3.1 Temperature Collecting Circuit

The temperature data collected by DS18B20 is transferred to 51 SCM. The hardware circuit is shown in Fig. 20.2. DS18B20 is a single-Bus digital interface

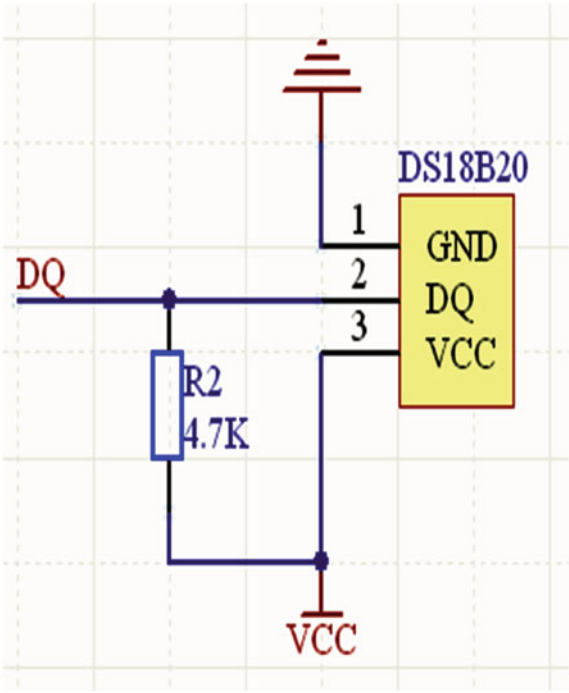


Fig. 20.2 Temperature collection circuit of DS18B20

device. One two-way I/O port of AT89S52 is used by S18B20. The communication between DS18B20 and AT89S52 is through the data line DQ.

20.3.2 LED Display Module

Four bits of temperature data are displayed, which are two integers and two decimals. The functions which are to realize real-time display of data collected, and compare correct with the data displayed by the PC are completed. In order to meet the requirements of brightness of LED, the transistor 8550 is used to drive LED with limiting resistor 2 k as shown in Fig. 20.3.

20.3.3 Serial Communication Circuit

The communication between the SCM and the PC is serial, and the chip MAX232 is used to convert level. There is a power supply voltage converter in the chip MAX232, which means that +5 V supply voltage can be transformed into +10 V or −10 V, which is required by RS232. The level conversion circuit of Serial communication between the SCM and the PC is shown in Fig. 20.4.

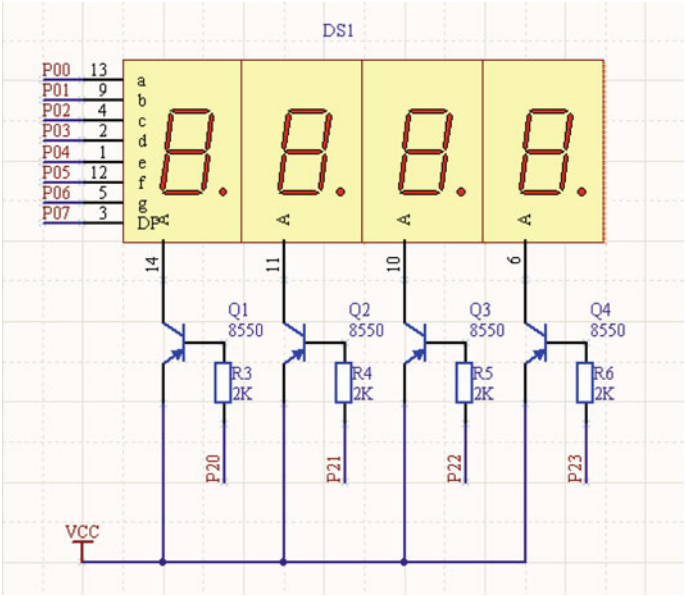


Fig. 20.3 Display circuit of LED

20.4 Software Design

The task of software design includes two parts which are PC software design and down computer software [5, 6]. The main program flowchart of the down computer is shown in Fig. 20.5.

After the initialization of the main program, temperature collecting program starts running, of which serial interrupt is turned off. The SCM dose not respond to any inquiry from Kingview. After the collecting program is completed, serial interrupt is turned on and the SCM can respond to requests sent by Kingview according to protocol and starts to send data. The temperature data collected by DS18B20 can be displayed real-time by Kingview. Timeout parameter settings of Kingview need to be carefully and repeatedly adjusted in the process of debugging to find a suitable delay time and adapt to different real-time requirements.

20.4.1 Design of Temperature Collecting Program

DS18B20 is used to complete the Single-Bus data transmission. Single-Bus protocol is the key to achieve temperature collecting program. The temperature collecting program flowchart is shown in Fig. 20.6.

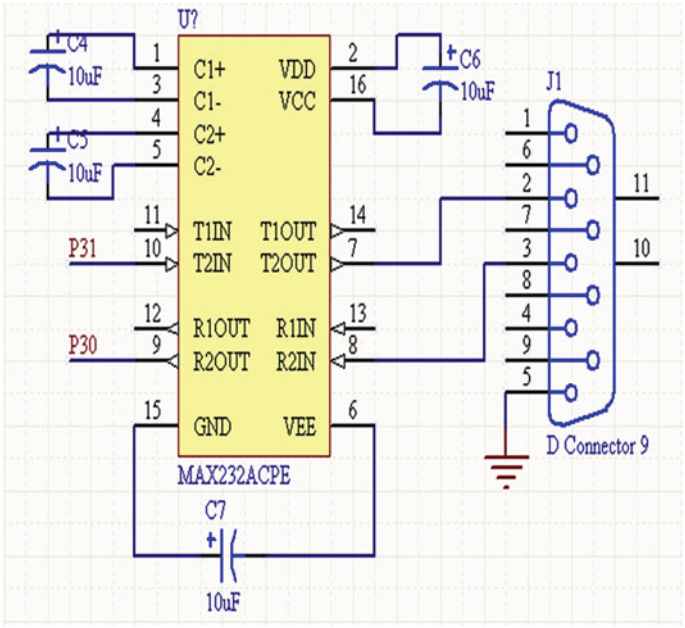


Fig. 20.4 Level conversion circuit of serial communication

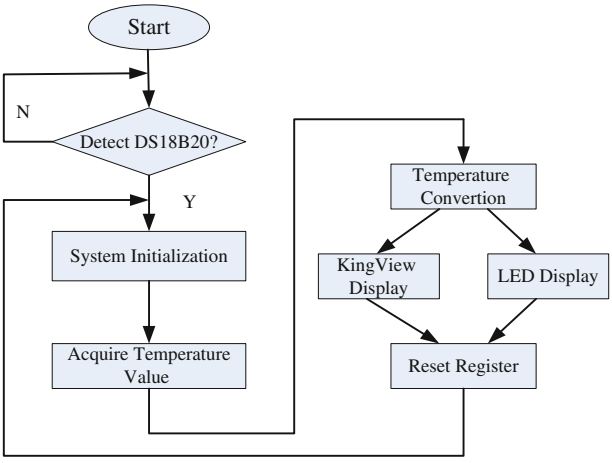


Fig. 20.5 Main program flowchart of system

Three steps have to be completed in temperature conversion of DS18B20 under the control of the SCM. (1) Before every read and write, a reset pulse should be sent by the SCM according to time consequence requirements, and a presence of pulse by DS18B20. (2) One write ROM instruction should be sent after a

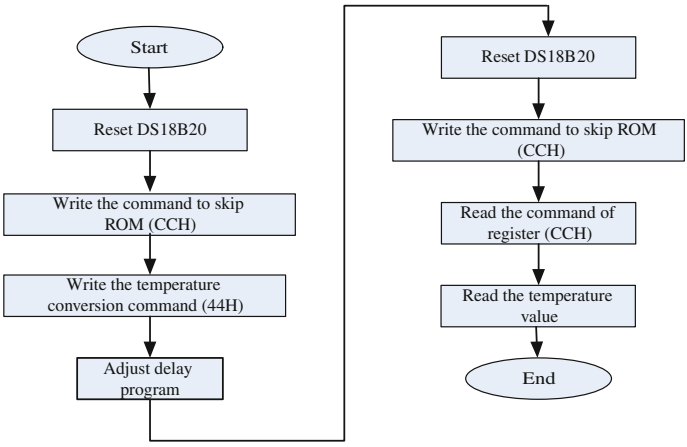


Fig. 20.6 DS18B20 Temperature collecting flowchart

Table 20.1 Time consequence of KingView reading data from SCM

KingView	SCM
Step 1: Send address request	
	Step 2: Respond address
Step 3: Read data request	
	Step 4: Respond and read data
Step 5: If the fourth step responded correctly, end. Otherwise, perform the first step.	
	Step 6: Respond address
Step 7: Read data request	
	Step 8: Respond and read data

successful reset. (3) A memory manipulation instruction should be sent, then scheduled operations to DS18B20 can be done.

20.4.2 Communication Protocol between Kingview and SCM

The protocol provided by Kingview communicating with the SCM can support HEX type and ASCII type, in which HEX type is of efficiency. The time consequence which Kingview read data from the SCM is shown in Table 20.1.

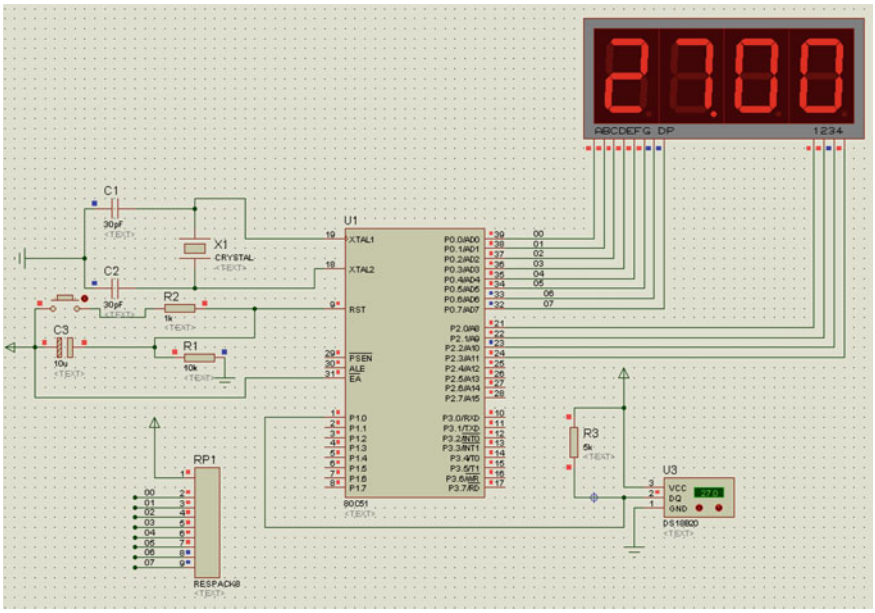


Fig. 20.7 Simulation results of temperature collecting system

20.4.3 Engineering Applications of Kingview

Steps to create a temperature control screen are as follows. First, create the project path. Second, create a configuration screen; then, define the I/O devices. Finally, construct the database. Communication parameters should be set after creating a temperature control screen, such as data bits, stop bits, baud rate and parity. All these communication parameters are decided by the SCM, and the setting of Kingview is the same with the SCM.

20.5 Simulation of System

20.5.1 Simulation of Hardware

After creating a simulation model in the Proteus software, double-click the SCM, load the HEX files and start the simulation. The initial value of the temperature of DS18B20 is 27°C. The simulation results are shown in Fig. 20.7.

When the size of the value of analog temperature is adjusted, the displayed value of the LED will change synchronously, thus the correctness of temperature acquisition program is verified.

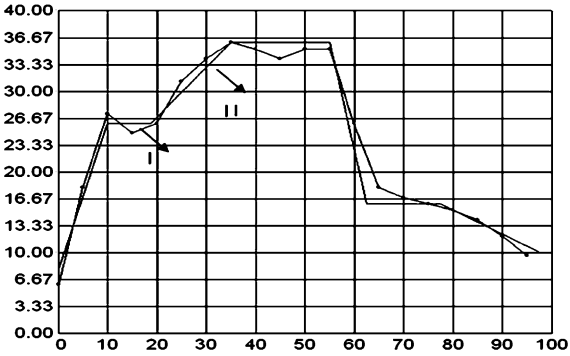


Fig. 20.8 Contrast curve of data collected

20.5.2 Test of Data Collection

Start the software and hardware systems and record displayed value of the LED and the time. Temperature curve of Kingview can be seen at the same time and should be saved, then put the data from LED into it, as shown in Fig. 20.8.

The vertical axis is temperature, and the horizontal is the time value in the contrast curve. Curve 1 is the real-time collection of temperature of Kingview and curve 11 is the real-time display of temperature of LED. The value from the LED can be used as an accurate value because of fast time. It can be seen that there are small errors between curves 1 and 11. The analysis shows that the main reason for the error is due to two aspects. One is the communication delay error, the other is communication timeout parameter settings error of Kingview. The communication timeout parameter settings should be adjusted repeatedly and carefully to find the right delay time and to adapt to different real-time requirements.

20.6 Conclusion

The hardware design of single-Bus temperature sensor DS18B20 is simple and reliable. The temperature display is realized by serial communication with PC through the serial interface chip MAX232, which can measure the ambient temperature at real-time and display it. Up and down temperature display can also be set. The purpose of intelligent control can be achieved by adjusting the ambient temperature according to the current display. The system has a strong value of promotion.

References

1. Yongle A, Ziyi F (2008) Digital electronic technology. Electric Power Press of China, Beijing
2. Hehui L, Ning W, Xiaojin Y et al (2009) National undergraduate electronic design contest refined solution to microcomputer application skills. Electronic Industry Press, Beijing

3. Xianjin D (2008) 51 SCM and C language program development instances. Tsinghua University Press, Beijing
4. Yufeng W, Xiaofeng M (2009) Temperature Collection System Design Based on KingView and DS18B20. Mine Automation 11(02):117–119
5. Xuejun Z (2006) Communication program design between SCM and KingView. Indust Control Comput 19(9):35–37
6. Guili T, Shangqing W (2007) Configuration software control technology. Beijing Institute of Technology Press, Beijing

Chapter 21
Design of Real-Time Ultrasound Control System Based on CPLD

Yanhua Zhang, Lu Yang and Pengfei Jin

Abstract Velocity of sound in the material is one of the major parameters of sound, so it is important to get velocity of sound in material accurately for material property study. Based on the ultrasonic velocity measurement method of TRIP, a real-time ultrasound control system with CPLD was designed. Time measurement with CPLD has high speed, high measurement precision and the whole control module has the advantage of simple interface and flexible operation, which fully meet the engineering requirements of real-time scanning applications in velocity of sound measurement.

KeyWords Ultrasonic nondestructive testing · TIRP · CPLD · Real-time control

21.1 Introduction

Ultrasonic nondestructive evaluation develops on the basis of ultrasonic flaw detection and ultrasonic nondestructive testing, which is one of the nondestructive evaluations having widest use. Using certain characteristics of sound wave propagation to characterize the properties of materials is a direction of ultrasonic nondestructive evaluation technology. The propagation velocity of sound, wave shape, peak, and spectrum and other parameters can be used to describe the different properties of material [1–3]. Among them, the material velocity of sound is most closely associated with the material properties. Longitudinal wave velocity of materials meets the following formula:

Y. Zhang (✉) · L. Yang · P. Jin
National Key Laboratory for Electronic Measurement Technology,
North University of China, Taiyuan, Shanxi, China
e-mail: tyzyhzyh@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_21,

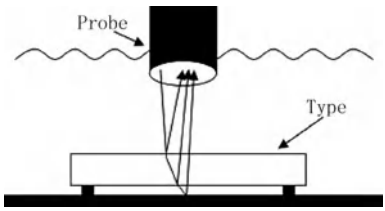


Fig. 21.1 Measurement principle of TIRP

$$v = \sqrt{\frac{E(1-\sigma)}{\rho(1+\sigma)(1-2\sigma)}} \tag{21.1}$$

In the formula, v is longitudinal wave velocity, E is Young's modulus, ρ is density, σ is the ratio of Poisson. From formula (21.1), we can know that longitudinal wave velocity in material is related to Young's modulus, density and the ratio of Poisson, therefore, the material properties can be evaluated by measuring velocity of sound.

21.2 Method and Principle of Velocity of Speed Measurement

In general, there are two methods measuring velocity of speed. The traditional measuring method is using ultrasonic flaw detector and ultrasonic probe as ultrasonic transmitter and receiver, fixing the probe at some point of type material, measuring the thickness of type material d and sound wave propagation time t, then velocity of sound can be calculated according to $v = 2d/t$. In the method, velocity of sound is obtained using known thickness divided by propagation time, therefore, measurement accuracy of speed velocity depends on the accuracy of the material thickness, so it is called measuring method of speed velocity based on the thickness. The method is simple and easy to realize, but its measurement error is comparatively bigger, so it is only used in type measurement with less thickness error. Besides, good coupling is needed between probe and type, so the method has limited use.

Another method is placing an acoustic reflection board under the type, the velocity of sound in the type material can be determined by the sound speed in water and the reflection time between acoustic reflection board and sensor. The measurement principle is as shown in Fig. 21.1, waveform received by probe is shown in Fig. 21.2.

In Fig. 21.2, wave ① is the reflection wave from the top surface of the type, wave ② is the reflection wave from the bottom surface, wave ③ is reflection wave form the acoustic reflection board after type is placed, wave ④ is reflection wave from the acoustic reflection board before type is placed. Suppose time difference between wave ① and wave ② is τ, time difference between wave ③ and wave ④ is Δt, velocity of sound in water is c, the longitudinal wave velocity in material can be got from following formula:

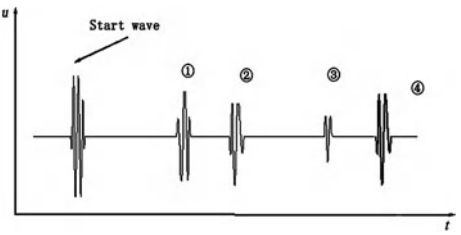


Fig. 21.2 Ultrasonic reflection wave

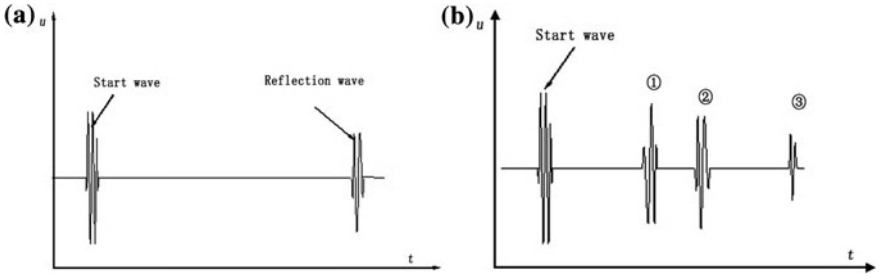


Fig. 21.3 Detection waveform. **a** Probe is not on the top of object. **b** Probe reaches the top of object

$$v = c(\Delta t/2\tau + 1) \tag{21.2}$$

From formula (21.2) we know that velocity of sound in material only relates to velocity of sound in water, time difference between top surface reflection and bottom surface reflection, and the time difference of sound reflection from acoustic reflection board before and after placing the type. It is unrelated to the thickness of type, so the method is named the thickness independent reflector plate (TIRP).

The method can measure velocity of sound without access to material thickness, measurement accuracy depends on the accuracy of c, Δt and 2τ. If the coupling water temperature is controlled strictly, c will be a constant, then the accuracy depends on time measurement.

21.3 Ultrasound Real-time Control System Based on CPLD

21.3.1 Time Measurement Based on CPLD

When the probe is still outside the area of the measured object, the ultrasonic wave reaches the acoustic reflection board directly and is reflected, the waveform is shown in Fig. 21.3a.

Defining the time difference between reflection wave and start wave is t_0. When the probe reaches the top of the measured object, the ultrasound will be reflected at every lay, so we can see three main waves; the chronological relationship is shown

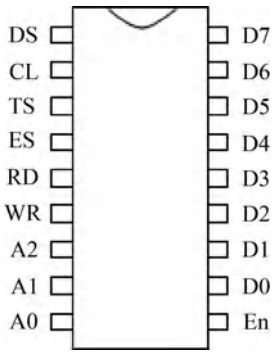


Fig. 21.4 CPLD module

in Fig. 21.3b. Define the time difference between reflection wave ① and start wave is t_1, that between reflection wave ② and start wave is t_2, that between reflection wave ③ and start wave is t_3, then $2\tau = t_2 - t_1$, $\Delta t = t_0 - t_3$. According to formula (21.2) velocity of sound (v) in the material can be obtained.

The key to obtain velocity of sound is obtained the time value of t_1, t_2, t_3. If the demodulation signal is selected by three separate modules, waveform ①, waveform ②, waveform ③ are isolated. Then start wave signal and isolated signals are via edge detection circuit, the count enabled signals can be got, which are high lever from the rising edge of start signal to the rising edge of isolated signals. If the count enabled signals are applied to the counter module, the results of counters are the time value of t_1, t_2, t_3.

21.3.2 Module Function

CPLD used in system can divided into several modules, peripheral interface module, master control module, the waveform strobe modules, counter enabled modules and counter modules (Fig. 21.4).

Peripheral interface module is responsible for data exchange of entire system and PC. It consists of I/O data buses D7–D0, address buses A2, A1, A0, and control buses RD, WR and En. It can send the configuration parameter of every module to master control module via I/O, and can also read the data recorded by counte. Demodulation signal (DS) is the input of demodulation signal, CLK is the input of system clock, Trigger signal (TS) is the input of trigger signal. Edge select (ES) is the selection switch of edge trigger, when in high voltage level it is triggered by rising edge, and triggered by falling edge when in low voltage level. En is the system enabled port, which is effective in low level.

Master control module mainly completes parameter configuration of entire real-time system and data transmission. Parameter configuration includes the configuration of three waveform strobe modules, and recording time data of three counters. Peripheral interface module sets different lever of address buses and

Table 21.1 Instruction code

En	WR	RD	A_2	A_1	A_0	D_7–$D0_0$	Directions
1	×	×	×	×	×	×	Enabled port of system, effective in low level
0	1	0	0	0	0	IN	Send D_7–$D0_0$ to gate position register of the first way
			0	0	1	IN	Send D_7–$D0_0$ to gate width register of the first way
			0	1	0	IN	Send D_7–$D0_0$ to gate position register of the second way
			0	1	1	IN	Send D_7–$D0_0$ to gate width register of the second way
			1	0	0	IN	Send D_7–$D0_0$ to gate position register of the third way
			1	0	1	IN	Send D_7–$D0_0$ to gate width register of the third way
	0	1	0	0	0	OUT	Read the counter of the first way to D_7–$D0_0$
			0	0	1	OUT	Read the counter of the second way to D_7–$D0_0$
			0	1	0	OUT	Read the counter of the third way to D_7–$D0_0$

control buses to realize entire module data configuration. Table 21.1 shows the specific configuration parameters of system modules.

Waveform strobe module is a two-step counter. Master control module can send the data delivered by I/O to the register of each waveform strobe module. When the rising edge of trigger is coming, the first step of counter starts to count, at the same time the module outputs low level. When the count value reaches the preset gate position parameters, the next step of counter starts to count, at the same time set the module output in high level. When the duration of high level reaches the preset gate width parameter, the output is set in low level again, then a complete waveform strobe time sequence is complete. Gate position can be adjusted by setting the parameters of Gate_Position via peripherals, the gate width can be adjusted by setting the numerical value of "Gate_Width", only one waveform in demodulation signals can be selected by setting appropriate gate width parameter, as shown in Fig. 21.5.

Counter enabled module is a signal edge detector. When the rising edge of trigger signal arrives, the count enabled signal terminal of the module outputs high lever. Make demodulation signal and previous gated signal do "And computation", the rising edge of new signal (demodulation signal) is as detection signal of the second rising edge. So when the second rising edge is coming, the count enabled signal terminal outputs low lever. Then enabled signal of the next module is formed. The process is shown in Fig. 21.6.

The accuracy of time recorded of counter module depends on clock frequency of the system. In practical flaw detection, generally a pulsed ultrasound is used whose duration is very short (microsecond order of magnitude), here count frequency of 1 MHz is taken. When master control module detects a falling edge of "Count_En", the count results is latched, sent to the time data register, and read available for peripheral.

For peripherals, the whole control system provides duration time of t_1, t_2, t_3, which is from the rising edge of trigger signal to the rising edge of demodulation signal. Make simple arithmetic with the data read out, Δt and 2τ can be obtained. Then according to formula (21.2) the velocity of sound in the type can be obtained.

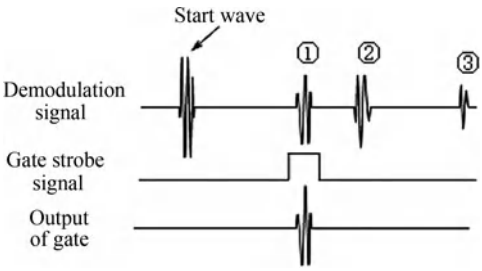


Fig. 21.5 Gate strobe signal

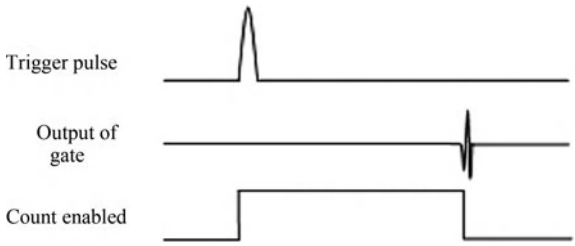


Fig. 21.6 Counter enabled signal

Because the entire count module is implemented with CPLD, which uses hardware to make time counting of whole system, needs a very short time, so it meets requirement of real-time data acquisition in other relevant system.

21.4 Conclusion

The processing of ultrasound waveform and part data computing are accomplished with CPLD in the paper, which meets the real-time requirement of the system. Especially in data measurement, hardware replacing software calculation makes the whole system run faster [4, 5]. In addition, the interface of entire system is simple, flexible, which can be widely used in some other situations required high-speed data acquisition.

References

1. He F (2003) Ultrasonic velocity measurement technology and its applications in materials evaluation. Mater Eng 8:29–39
2. Raišutis R, Kažys R (2007) Application of the ultrasonic characterization methods for highly attenuating plastic materials. NDT E Int 6:324–332
3. Ma HW, Zhang XH, Wei J (2002) Research on an ultrasonic NDT system for complex surface parts. Dep Mech Eng 9:667–670

4. Gang X, Dong-sheng X, Yun X (2010) Design and realization of multi-target pulsed laser range finder on CPLD. Laser Infrared 40(2):152–154
5. Li R (2010) Controller design of electric valve power unit based on CPLD. Micromotor 43(9):106–108

Chapter 22
Study on Smart Energy Storage Technology and Control Strategy in Micro-Grid

Xing-guo Tan and Shan Lu

Abstract Energy storage technology is an indispensable support for reliable operation of micro-grid (MG). Various forms of energy storage system (ESS)'s application in MG are summarized; as energy storage is not easy enough to use, the development of smart energy storage is necessary. The concept of smart energy storage is proposed. The potential topologies of smart energy storage is given and analyzed, and some crucial control technologies of smart storage are also provided, including charging/discharging control method, power flow distribution control method. The principle of a novel designed power charging/discharging control method of ESS is illustrated, and the effectiveness of the method is verified with simulation results in PSIM.

Keywords Micro-grid · Smart energy storage · Topology · Control strategy · Simulation

22.1 Introduction

Micro-grid (MG) combines Distributed Generation (DG), customers/loads, energy storage units together by power electronic technology, to enhance local power stability and flexibility [1]. And can be further defined as a small electric power

X. Tan (✉)
School of Electrical Engineering, Henan Polytechnic University,
Jiaozuo 454000, China
e-mail: tanxingguo1981@126.com

X. Tan
School of Electrical Engineering, Shandong University, Jinan 250061, China

S. Lu
Department of Mechanic and Electrical Engineering, Jiaozuo University,
Jiaozuo 454000, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_22,

system being able to operate physically islanded or interconnected with the utility grid. Because of its small scale, MG is vulnerable to the fluctuation from DG and loads, as more and more renewable source DGs will be incorporated in MG, their inner characters of fluctuation and intermittency make things even worse; at the same time, MG must confirm to strict power quality requirements when being integrated into the utility grid, so as to maintain a high power factor and suppress the harmonics distortion, etc. as the indispensable part of MG, Energy storage system (ESS) provides a relative simple solution to these problems in MG.

22.1.1 Application of ESS in MG

Various ESSs found different applications in MG can be concluded as flows: (1) Providing energy buffer and short time power supply. When power failure or faults are detected in utility grid, MG disconnects utility grid, and works in island mode, the instant power fall can be complemented by ESS, with the buffer of ESS, the seamless transition can be fetched easily in MG; (2) Helping integration of more DGs. More and more DGs, especially renewable energy sources will be incorporated in MG, the fluctuation and intermittence of renewable may crash MG. ESS with quick response, can simply compensate the fluctuation and intermittence; (3) Balance supply and loads economic potential. As power generated by renewable sources and the demand from loads are always not matching in MG, ESS can take abundant energy, then release it when there is a shortage; (4) Economic potential. As time-of-use tariff is widely accepted in many nations, the tariff of peak is several times of off-peak prices, the spread between the peak and off-peak prices represents an opportunity for ESS that can economically store energy for hours or days [2]; (5) Improving power quality. Controlled by power electronics-based interfaces, ESS can work as a power quality regulator to generate specified active or reactive power. Some ESS of instantaneous response, such as super-capacitor or FESS, can quickly absorb or release energy of high density, suitable to tackle transient issues like instantaneous outage caused probably by system failures, sudden voltage swells or sags. Some large capacity ESS like Lead-acid battery can be used to compensate instant power shortage to smooth voltage fluctuations.

22.2 Smart ESS

Today the application of ESS is still not an easy job, as an infrastructure of smart grid, ESS should be even “smarter” in the future. So far very few have given a definite meaning of smart storage. According to the demand of smart grid, a smart ESS in the future may be a unified platform, besides ESS, standard power electronics interfaces, high efficient energy management systems, control systems, human machine interface (HMI) and communication interfaces for interconnection

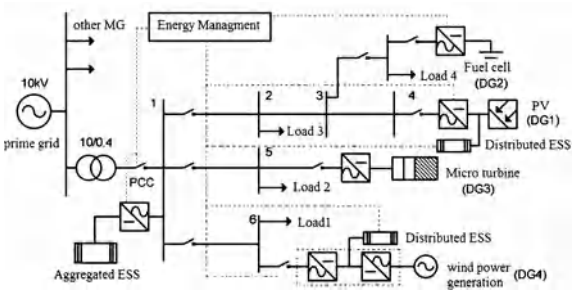


Fig. 22.1 Typical ESS configuration

with smart grid, all should be incorporated. With well-established software and tools, the smart ESS can be used flexibly in future's smart grid just like a plug and play equipment.

A smart ESS may also lie in properly integrating ESSs with other smart grid devices. For example, the integration of an ESS into FACTS devices can provide independent active and reactive power absorption/injection into/from MG, leading to a more economical and/or flexible transmission controller [3], and makes it an ideal candidate for many types of power system applications, including voltage control and oscillation damping.

22.2.1 Topologies of Smart ESS

There are two typical configurations of ESS in MG, the aggregated ESS and distributed ESS, as shown in Fig. 22.1. For the first type, all ESSs are aggregated as one ESS and connected to the MG bus, while the power flow from DGs to the point of common connection (PCC) bus can be kept at a specified value. All the capacity of an aggregated ESS can be used to soothe power flow fluctuations [4], aggregated ESS is partly approved to have a superior performance to the distributed ESS at the same capacity. But if the energy storage capacity increases, both manufacture and control of the ESS become difficult and cost ineffective. It is then possible to use small-scale and distributed energy storage to achieve flexible and efficient power regulation.

For the distributed ESS configuration, ESS units are directly coupled to individual DGs with various interfaces. An advantage of separately connecting each pair of DG and ESS to the grid lies in that, each pair's power electronics interface can be optimized to reduce the cost and maximize the efficiency. These systems tend to be simple since they only have to deal with a single type of source.

For the future MG, smart ESS technology can be used to optimize MG operation. To stabilize MG, smart storage can be plugged into DC bus of MG, large-scale energy storage can connect to smart ESS module with power electronic interface, a proper algorithm can be chosen from its knowledge database, like the aggregated ESS; to optimize respective DG or micro-source, smart ESS can be

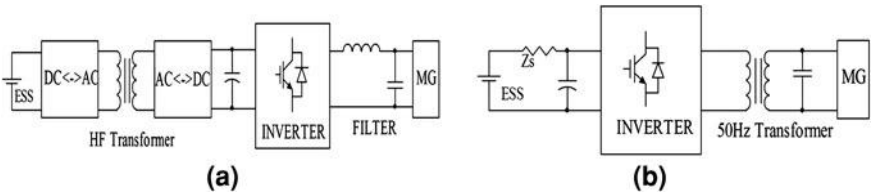
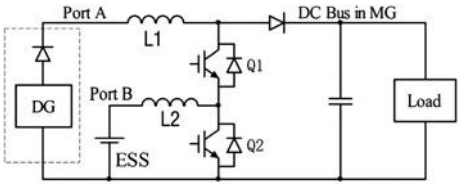


Fig. 22.2 Typical ESS converter

Fig. 22.3 ESS converter without transformer



configured to distributed ESS and optimize DG operation. Two smart ESSs should interconnect with upper controller to coordinate each other.

22.2.2 Smart ESS Converter

The basic requirement of smart ESS converter topologies should be of high efficiency, reliability, and convinced to integrate and expand many different energy storage devices. Advanced power electronics are at the heart of the interface between energy storage and the electrical grid. There are several typical topologies for smart ESS [5]; the first case is shown in Fig. 22.2a, where a bidirectional DC/DC converter accomplishes the power flow between the ESS and the DC bus with a high frequency transformer. This configuration is much complex and cost much, but has a high efficiency.

The second configuration is shown in Fig. 22.2b, where the ESS is directly connected to the inverters' DC bus while a line frequency (50 Hz) transformer is used to boost up the required voltage level. The main advantage lies in simplicity of the configuration; however, high cost and low efficiency exist because of the low frequency transformer. The third configuration connects ESS and DC bus of MG, without transformer. Figure 22.3 is one possible configuration. The ESS can simply be connected to the common DC bus through separate DC/DC converters. This system is much easier but has no electrical isolation. The use of combined multiport converter presents advantages of low cost, high efficiency and reliability than the original separate converter system.

In smart storage, the capacity of ESS should be adjusted, so the topologies are open to expansion. More energy storage device can be incorporated in DC link, or high frequency link, multi-port DC–DC converter may be a better solution.

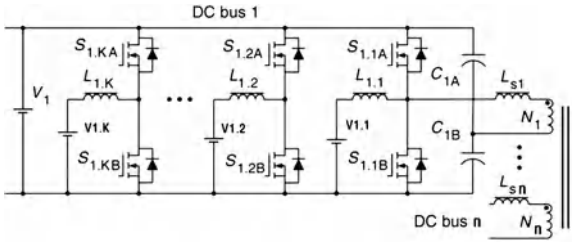


Fig. 22.4 Multi-port ESS converter [6]

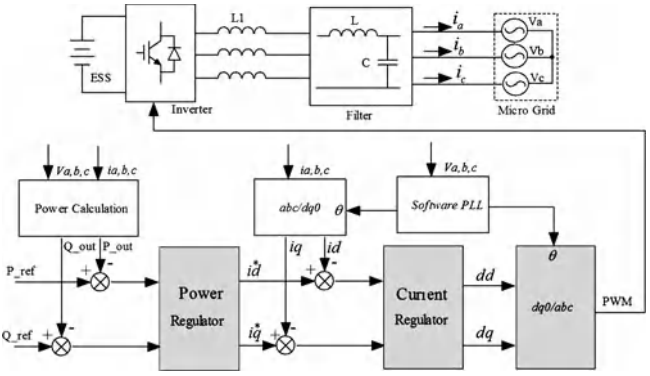


Fig. 22.5 The principle of designed power charge/discharge control of ESS

In Figure 22.4, DC bus joins multi ESS of similar voltage together, then various DC buses are converted to high frequency AC form, and combined together to connect high frequency link of MG, which were then changed to regulated form. This system is open to expansion, and many common parts of different converters are shared, so the cost is acceptable.

22.2.3 Control Strategy of Smart Storage

For smart storage application, three parts should be considered, including power management, real power and reactive power control and its charge/discharge control. In this essay, as charge/discharge control is the basis operation of smart storage, a designed power method is proposed.

Constant current charging and constant voltage charging are two classic charge modes widely used in batteries and super capacitors [7]. But actually, these methods are suitable for battery reserving energy for later usage, for example in electric bicycle or EV. When ESS is used to compensate the fluctuation of renewable source

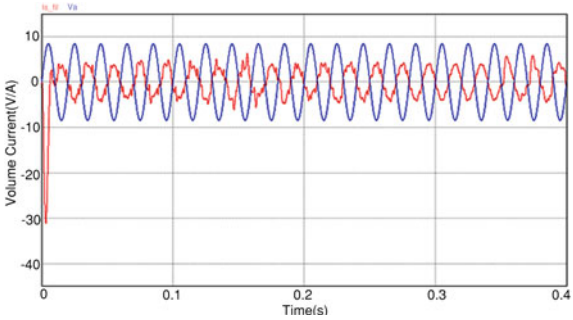


Fig. 22.6 Designed active power charge control

generation or to solve the power quality problem, the battery ESS may often work at partial state of charge (PSOC). PSOC operation of low efficiency occurs when a battery is less than fully discharged and then less than fully recharged before being discharged again. On the other hand, these method of charging or discharging produce much harmonic on PCC, it is not suitable for large capacity ESS. So designed power charge–discharge method is a good alternative. Figure 22.5 is the principle of designed power charge/discharge control of ESS.

According to instantaneous power theory, under the condition of three balanced voltage, active power and reactive power transition between ESS and MG are as follows:

$$p = u_a i_a + u_b i_b + u_c i_c, q = \frac{1}{\sqrt{3}}[(u_a - u_b)i_c + (u_c - u_a)i_b + (u_b - u_c)i_a] \quad (22.1)$$

When transformed to dq0 rotating coordinate system, the instantaneous active power and reactive power are:

$$p = u_{sd} i_d + u_{sq} i_q, q = u_{sq} i_d - u_{sd} i_q \quad (22.2)$$

So the active current i_d and reactive current i_q are:

$$\begin{aligned} i_d &= \frac{u_{sd}}{u_{sd}^2 + u_{sq}^2} p + \frac{u_{sq}}{u_{sd}^2 + u_{sq}^2} q \\ i_q &= \frac{u_{sq}}{u_{sd}^2 + u_{sq}^2} p - \frac{u_{sd}}{u_{sd}^2 + u_{sq}^2} q \end{aligned} \quad (22.3)$$

Once reference given active power and reactive power are given, i_d and i_q references can be calculated separately, by comparing with measured current, so the control variables can be fetched to control the inverters. A simulation model is established in PSIM to test the feasible of designed power control.

1. Designed active power control. This means the reactive power is zero, when active power reference is negative, ESS absorbs energy and being charged, while active power reference is positive, ESS will be discharged and release energy. In Fig. 22.6, active power reference is set to be −50 w, in Fig. 22.5 it turns out to be 50 w.

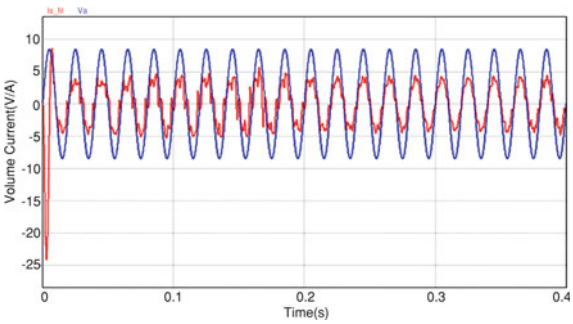


Fig. 22.7 Designed active power discharge control

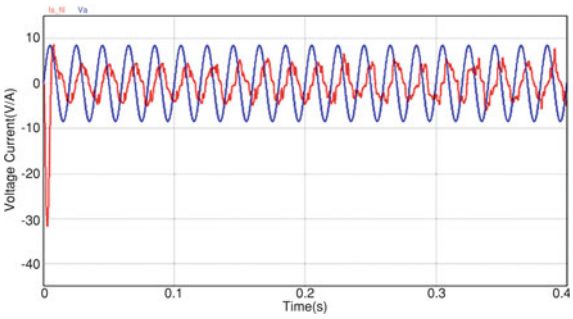


Fig. 22.8 Reactive power control of charge/discharge process

In Fig. 22.6, MG voltage and current are opposite in phase, so ESS is charged by MG; while in Fig. 22.7, MG voltage and current are of the same phase, so ESS discharges itself to provide energy to MG. we noticed that the power factor is about unit 1 in this mode.

2. Designed reactive power control: Some other time, in order to regulate the voltage at ac bus, reactive power balance is indispensable, so designed reactive power is expected from ESS, in Fig. 22.8, reactive power reference is set to be 50 Var. we can see MG current lags behind its voltage by 90°, so only reactive power is provided by ESS. Similarly, when reactive power reference is set to be negative, ESS will absorb reactive power according to MG instructions.

So by simulation, designed power charge–discharge control method for ESS is prove to be feasible, which has a high efficiency and power factor (in active power control), and its main circuit are three phase bridge structure which can provide much high capacity for ESS application.

22.3 Conclusion

For smart ESS, there are still many problems to be solved, for example, how to optimize configuration of ESS, find advanced but cheap energy storage technology, and manage different types of hybrid ESS etc. This essay briefly describes the

meaning and effects of smart ESS, topologies and control methods of ESS is then summarized. To fully utilize ESS, a designed power charge–discharge control method for ESS is analyzed, and in PSIM, this method is modeled and tested for feasibility and accuracy.

References

1. Kyebyung L, Son KM, Gilsoo J (2009) Smart storage system for seamless transition of customers with intermittent renewable energy sources into microgrid. In: Proceedings of thirty-first tele. energy conference, pp 1–5
2. David A, Chris B, John H, Vinh L, Robert S (2010) Energy storage in the New York electricity markets [R]. A New York ISO White Paper, pp 1–14
3. Ribeiro PF, Johnson BK, Crow ML, Arsoy A, Liu Y (2001) Energy storage systems for advanced power applications. In: Proc IEEE 89(12):1744–1756
4. Wei L, Géza J (2007) Comparison of energy storage system technologies and configurations in a wind farm. In: Proceedings of IEEE power electronics specialist conference, pp. 1280–1285
5. Serban I, Marinescu C (2010) A look at the role and main topologies of battery energy storage systems for integration in autonomous microgrids. In: Proceedings of 12th optimization of electrical and electronic equipment, pp 1186–119
6. Tao H, Kotsopoulos A, Duarte JL, Hendrix MAM (2006) Family of multiport bidirectional DC–DC converters. Proc IEEE Electr Power Appl 153(3):451–458
7. Jinhui X, Zhongdong Y, Bingbing W, Ziping W, Jun L (2009) Technology research of novel energy storage control for the PV generation system. In: Proceedings of power and energy engine conference, pp 1–4

Chapter 23
Study of Data Fusion Method in Distributed Detection System

Su-yan Zhang and Jing-feng Zhang

Abstract The datum is disposed by rough set theory in the data fusion center of a distributed detection system. Its basic theoric scheme is explained while its typical applications are discussed. The structure of distributed detection system and crisp fusion method in data fusion center are presented. In crisp method, redundant condition attributes and redundant condition attribute values are removed respectively by the condition attribute reduction, the value reduction of the condition attributes. Decision rules are deduced.

Keywords Distributed detection system · Data fusion · Rough sets theory · Classification

23.1 Introduction

Data fusion [1, 2] has been applied widely in various fields as a comprehensive technology. There are multi-information source in distributed detection system, and the data from every information source may be imperfect. The technique of data fusion can deal with the imperfect data to classify accurately. Data fusion-based rough sets theory has obtained attention by experts and researchers. Rough sets theory [3] was put forward by Poland scientist Pawkak in 1982. Rough sets

S. Zhang (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: susu_yang@163.com

J. Zhang
North Automatic Control Technology Research Institute,
Taiyuan 030006, China
e-mail: zjfvic@yeah.net

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_23,

theory need not have any attributes features or mathematical description of the object, no priori knowledge, classification of the knowledge system of the object, find indiscernibility relations, reduce knowledge system, and deduce decision rules to unsolved problems. In the distributed detection system, the data fusion method-based rough sets theory includes complete and incomplete fusion method for complete and incomplete data [4–6].

23.2 Distributed Detection System and Rough Sets Theory

23.2.1 The Structure of Distributed Detection System

The distributed detection system includes the detected object, sensors, preprocessors and data fusion center (see Fig. 23.1). Such sensors as sensor 1, sensor 2,…, sensor n, collect different datum from the detected objected. Every preprocessor corresponding to every sensor filter disturbance of collected data, and discretize continuous datum are collected. When preprocessed data is complete datum, data fusion center extract and output decision rules applying complete fusion method. When preprocessed data is incomplete datum, data fusion center extract and output decision rules applying incomplete fusion method. In the paper, complete fusion method is applied to solve problem.

23.2.2 Basic Concept of Rough Sets Theory

The knowledge table of the detected object in data fusion center (see Table 23.1) contains the universe U, condition attribute A and decision attribute D where $U = \{x_1, x_2, \ldots, x_k\}$ is the detection times; condition attributes $A = \{a_1, a_2, \ldots, a_n\}$ is composed of all collected data's attribute; decision attribute $D = \{d\}$ is general decision. The value of $A = \{a_1, a_2, \ldots, a_n\}$ and $D = \{d\}$ is u_{ij} and d_i respectively, where $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n$.

The upper approximations $\bar{P}(X)$, the lower approximations $\underline{P}(X)$ and the boundary region $BN(X)$ are defined respectively as follows:

$$\bar{P}(X) = \left\{ x \in U, [x]_Q \cap X \neq \Phi \right\} \tag{23.1}$$

$$\underline{P}(X) = \left\{ x \in U, [x]_Q \subseteq X \right\} \tag{23.2}$$

$$BN(X) = \bar{P}(X) - \underline{P}(X) \tag{23.3}$$

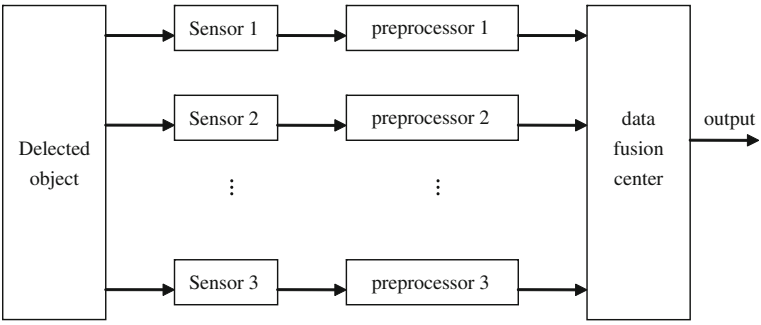


Fig. 23.1 The structure of distributed detection system

Table 23.1 Knowledge table of the detected object

U	a_1	a_1	…	a_n	d
x_1	u_{11}	u_{12}	…	u_{1n}	d_1
x_2	u_{21}	u_{22}	…	u_{2n}	d_2
…	…	…	…	…	…
x_k	u_{k1}	u_{k2}	…	u_{kn}	d_k

Where X denote the subset of elements of the universe U, $[x]_Q$ is the set of the indiscernibility objects x_i. The upper approximations refer to the elements of all possibly and doubtlessly belong to the set X, and the lower approximations refer to the elements that doubtlessly belong to the set X, the boundary region refer to the elements that possibly belong to the set X.

An accuracy measure of X is defined as:

$$\beta_Q(X) = \mathrm{card}(\underline{P}(X)) \Big/ \mathrm{card}(\bar{P}(X)) \tag{23.4}$$

When $\bar{P}(X) = \underline{P}(X)$, X is definable in U, and $\beta_Q(X) = 1$. When $\bar{P}(X) \neq \underline{P}(X)$, X is undefinable in U, and $\beta_Q(X) < 1$. As to definable set X, data fusion center apply complete fusion method. On the contrary, data fusion center apply incomplete fusion method. Where $card(\bullet)$ means the objects' number.

23.2.3 Complete Fusion Method

Complete fusion method to solve problem is illustrated on a knowledge table of a local distributed detection system (see Table 23.2).

Step 1. Construction of elementary sets in D space, that is, classify the universe U on decision attribute D: set 1 $\{x_1, x_3\}$, set 2 $\{x_2, x_4\}$, set 3 $\{x_5, x_6\}$. Decision

Table 23.2 Knowledge table of a example

U	a_1	a_1	a_3	d
x_1	2	1	3	1
x_2	3	2	1	2
x_3	2	1	3	1
x_4	2	2	3	2
x_5	1	1	4	3
x_6	1	1	2	3

Table 23.3 Relevant calculation of classification

Class number	Number of objects	Lower approximation set	Upper approximation set	Measure accuracy
1	2	$\{x_1, x_3\}$	$\{x_1, x_3\}$	1.0
2	2	$\{x_2, x_4\}$	$\{x_2, x_4\}$	1.0
3	2	$\{x_5, x_6\}$	$\{x_5, x_6\}$	1.0

attribute D has a value of 1 for all objects belonging to the elementary set 1, value of 2 for the objects belonging to the elementary set 2, value of 3 for the objects belonging to the elementary set 3.

Step 2. Calculation of upper approximation, lower approximation and classification accuracy of the elementary sets in D according to Eq. (23.1), Eq. (23.2) and Eq. (23.4) (see Table 23.3).

The results presented in Table 23.3 indicate that all elementary sets in D can be properly described in terms of the condition attribute $A = \{a_1, a_2, \ldots, a_n\}$. It means complete fusion method should be adopted in the example.

Step 3. Finding D-core and D-reducts of A to delete redundancy condition attribute. D-core is the dependent attribute set that can be interested in finding all possible minimal subsets of attributes. Redundant condition attribute can be wiped off from decision table by D-reducts, to get D-core. For this, the D-discernibility matrix is constructed, and elements of which discern objects from different groups in decision attribute D (see Table 23.4). Where a_{123} indicates a_1, a_2, a_3; and a_{13} indicates a_1, a_3.

According to the D-discernibility matrix, discernibility function $f_A(D)$ has the following form:

$$\begin{aligned} f_A(D) = {} & (a_1 + a_2 + a_3)a_2(a_1 + a_3)(a_1 + a_3)(a_1 + a_2 + a_3) \\ & \times (a_1 + a_2 + a_3)(a_1 + a_2 + a_3)a_2(a_1 + a_3)(a_1 + a_3)(a_1 + a_2 + a_3) \\ & \times (a_1 + a_2 + a_3) = a_1a_2 + a_2a_3 \end{aligned}$$

Discernibility function $f_A(D)$ derives the D-core a_1 and two D-reducts of knowledge table of a example in Table 23.2 (see Tables 23.5, 23.6).

Step 4. For every D-reduct in Tables 23.5 and 23.6, data fusion center finds D-core and D-reducts of A attribute values. For this, two relative D-discernibility

Table 23.4 The *D*-discernibility matrix

U	1	2	3	4	5	6
1	–					
2	a_{123}	–				
3	–	a_{123}	–			
4	a_2	–	a_2	–		
5	a_{13}	a_{123}	a_{13}	a_{123}	–	
6	a_{13}	a_{123}	a_{13}	a_{123}	–	–

Table 23.5 One *D*-reduct

U	a_1	a_1	d
x_1	2	1	1
x_2	3	2	2
x_3	2	1	1
x_4	2	2	2
x_5	1	1	3
x_6	1	1	3

Table 23.6 Another *D*-reduct

U	a_1	a_3	d
x_1	2	3	1
x_2	3	1	2
x_3	2	3	1
x_4	2	3	2
x_5	1	4	3
x_6	1	2	3

Table 23.7 The relative *D*-discernibility matrix of Table 23.5

U	1	2	3	4	5	6
1	–	a_{12}	–	a_2	a_1	a_1
2	a_{12}	–	a_{12}	–	a_{12}	a_{12}
3	–	a_{12}	–	a_2	a_1	a_1
4	a_2	–	a_2	–	a_{12}	a_{12}
5	a_1	a_{12}	a_1	a_{12}	–	–
6	a_1	a_{12}	a_1	a_{12}	–	–

matrixes must been constructed in terms of foregoing two *D*-reducts. Where, the relative *D*-discernibility matrixes of the *D*-reduct in Table 23.5 is illustrated (see Table 23.7).

According to the relative *D*-discernibility matrix of Table 23.5, the relative discernibility functions as following:

$$f_1(D) = (a_1 + a_2)a_2a_1a_1 = a_1a_2;$$
$$f_2(D) = (a_1 + a_2)(a_1 + a_2)(a_1 + a_2)(a_1 + a_2) = a_1 + a_2;$$

Table 23.8 One minimization *D*-reduct

U	a_1	a_1	d
x_1	2	1	1
x_2	*	2	2
x_3	2	1	1
x_4	*	2	2
x_5	1	*	3
x_6	1	*	3

$f_3(D) = (a_1 + a_2)a_2a_1a_1 = a_1a_2;$
$f_4(D) = a_2a_2(a_1 + a_2)(a_1 + a_2) = a_2;$
$f_5(D) = a_1(a_1 + a_2)a_1(a_1 + a_2) = a_1;$
$f_6(D) = a_1(a_1 + a_2)a_1(a_1 + a_2) = a_1.$

Step 5. After step 4, the *D*-reduct in Table 23.5 is changed into the minimization *D*-reduct (see Table 23.8).

The decision rules can be obtained from Table 23.8:

$a_1 = 2, a_2 = 1 \Rightarrow d = 1; a_2 = 2 \Rightarrow d = 2;\ a_1 = 1 \Rightarrow d = 3.$

Adopting the same way, data fusion center derives the decision rules of the *D*-reduct in Table 23.6. In sum, the decision rules to solving the problem of the distributed detection system in Table 23.2.

23.3 Discussion

Rough sets theory application demonstrated its usefulness in the cognitive process of data fusion. In particular, rough sets theory has provided series of tools for conceptualization, classification and analysis of the various types of data, especially, in the course of dealing with inexact, uncertain or vague knowledge, and in the course of discovering hidden patterns and regularities in applications about information system and artificial intelligence system.

References

1. Li Y (2002) Research on information fusion and intelligent process. Zhejiang University Publications, China
2. Zhang Z, Sun S (2001) Rough neural network and its application in multisensor data fusion. Control and Decision 16:76–78
3. Zeng H (1996) Rough theory and application. Chongqing University Publications, China
4. Pawlak Z (2002) Rough sets and intelligent data analysis. Inf Sci 147:1–12
5. Francis EHT, Shen LX (2002) Economic and financial prediction using rough sets model. Eur J Oper Res 141:641–659
6. Marzena K (1998) Rough set approach to incomplete information systems. Inf Sci 112:39–49

Chapter 24
Exponential Stability Control for T-S Fuzzy Nonlinear Networked Control Systems with Stochastic Time Delays

Hejun Yao and Fushun Yuan

Abstract To investigate the exponential stability control problem for a class of nonlinear networked control systems, a T-S fuzzy model is employed to represent the nonlinear controlled plant with network-induced stochastic delay. Based on the Lyapunov stability theorem, the exponential stability condition and the state feedback fuzzy controller design methods are given in terms of linear matrix inequality. A numerical example is given to demonstrate the effectiveness of the proposed method.

Keywords Networked control systems · T-S model · Stochastic delays · Linear matrix inequality (LMI)

24.1 Introduction

Feedback control systems in which the control loops are closed through a real-time network are called networked control systems (NCSs). NCSs have many advantages over a traditional point-to-point design including low cost of installation, ease of maintenance, low cost and greater flexibility. For these reasons, modeling, analysis and control of NCSs have recently emerged as a topic of much interest to the community [1–4].

The fuzzy control is a useful approach to solve the control problems of nonlinear systems. Based on the T-S fuzzy model with time delay term, many

H. Yao (✉) · F. Yuan
School of Mathematics and Statistics, Anyang Normal University,
Anyang 455002, Henan, China
e-mail: yaohejun@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_24,

researches have been presented to deal with the stability and the stabilization problem of the nonlinear NCSs with time delays [5, 6]. Walsh proposed a two-step design approach to use standard control methodologies and to choose the network protocol and bandwidth in order to ensure that important closed-loop properties are preserved when a computer network is inserted into the feedback loop [7]. Nesic and Teel addressed the input-to-state stability and L_p stability of nonlinear NCSs [8]. Based on this fact, some new methods and approaches should be developed for designing controllers for nonlinear NCSs, which motivates this paper.

This paper considers the exponential stability control problem for a class of nonlinear networked control systems (NCSs). Based on the Lyapunov stability theorem, the exponential stability condition and the state feedback fuzzy controller design methods are obtained.

24.2 Problem Formulation

Consider the following nonlinear control system with delay

Rule i:

IF $z_1(t)$ is M_1^i and $z_2(t)$ is M_2^i,..., and $z_n(t)$ is M_n^i

THEN

$$\begin{aligned} \dot{x}(t) &= (A_i + \Delta A_i(t))x(t) + (A_{\mathrm{di}} + \Delta A_{\mathrm{di}}(t))x(t-d) + (B_i + \Delta B_i(t))u(t) \\ x(t) &= \phi(t) \quad t \in [-d,\, 0] \end{aligned} \tag{24.1}$$

where $z(t) = [z_1(t)\, z_2(t)\, \cdots\, z_n(t)]^T$ is the premise variable, $x(t) \in R^n$ is the systems state vector, $u(t) \in R^m$ is the controlled input vector, $y(t) \in R^l$ is the output vector, $M_k^i (i = 1, 2, \ldots, r;\, k = 1, 2, \ldots, n)$ are fuzzy sets. A, $A_{\mathrm{di}} \in R^{n\times n}$ are known constant matrices, $B_i \in R^{n\times m}$ is input matrix, $\phi(t) = [\phi_1(t)\, \phi_2(t)\, \cdots\, \phi_n(t)]^T \in R^n$ is the given initial state on $[-d, 0]$, d is state delay, q is the number of IF-THEN rules. $\Delta A_i(t), \Delta A_{\mathrm{di}}(t) \in R^{n\times n}$ representing the uncertainties satisfying:

$$[\,\Delta A_i(t) \quad \Delta A_{\mathrm{di}}(t)\, \Delta B_i(t)\,] = DF(t)[\,E_{i1}\, E_{i2} \quad E_{i3}\,]$$

where $D, E_{i1}, E_{i2}, E_{i3}$ are constant matrices with appropriate dimensions, $F(t)$ is a matrix with appropriate dimensions satisfying $F^T(t)F(t) \le I$. By using a center average defuzzifier, product inference, and a singleton fuzzifier, the global dynamics of the T-S fuzzy systems are described by

$$\begin{aligned} \dot{x}(t) &= \sum_i^r \mu_i(z(t))[(A_i + \Delta A_i(t))x(t) + (A_{\mathrm{di}} + \Delta A_{\mathrm{di}}(t))x(t-d) + (B_i + \Delta B_i(t))u(t)] \\ x(t) &= \phi(t) \quad t \in [-d,\, 0] \end{aligned} \tag{24.2}$$

where $\mu_i(z(t))$ satisfying

$$\mu_i(z(t)) \geq 0, \sum_{i=1}^{q} \mu_i(z(t)) > 0, \quad i = 1, 2, \ldots, r$$

Throughout this note, we suppose that all the system's states are available for a state feedback control. In the presence of the control network, data transfers between the controller and the remote system, e.g., sensors and actuators in a distributed control system will induce network delay in addition to the controller proceeding delay. We introduce stochastic delay $\tau(t)$ to denote the network-induced delay. In this note we make the following assumptions:

Assumption 1: Sensor and controller are clock-driven;

Assumption 2: Actuator is event-driven.

We will design the state feedback fuzzy controller

$$u(t) = \sum_{i}^{r} \mu_i(z(t)) K_i x(t - \tau(t)) \tag{24.3}$$

where $\tau(t)$ is the stochastic network-induced delay satisfying $\tau(t) \in [0, \tau]$ Inserting the controller (24.3) into system (24.2), we obtain the closed system:

$$\begin{aligned} \dot{x}(t) &= \sum_{i}^{r} \sum_{j}^{r} \mu_i(z(t)) \mu_j(z(t)) \bar{A}_{ij} \xi(t) \\ x(t) &= \psi(t) \quad t \in [-\bar{d}, 0] \end{aligned} \tag{24.4}$$

where

$$\begin{aligned} \bar{A}_{ij} &= [\bar{A}_i \, \bar{A}_{\mathrm{di}} \, \bar{B}_i K_j] \quad \bar{A}_i = A_i + \Delta A_i(t) \\ \bar{A}_{\mathrm{di}} &= A_{\mathrm{di}} + \Delta A_{\mathrm{di}}(t) \quad \bar{B}_i = B_i + \Delta B_i(t) \\ \xi(t) &= [x^T(t), x^T(t-d), \quad x^T(t - \tau(t))]^T \end{aligned}$$

The initial condition of the state is supplemented as $x(t) = \psi(t)$, where $\psi(t)$ is a smooth function on $[-\bar{d}, 0]$, $\bar{d} = \max\{\tau, d\}$. Therefore, there exists a positive constant $\bar{\psi}$ satisfying

$$||\dot{\psi}(t)|| \leq \bar{\psi} \quad t \in [-\bar{d}, 0]$$

24.3 Main Results

Lemma 1 [2] *For any vectors a, b and matrices N, X, Y, Z with appropriate dimensions, if the following matrix inequality holds*

$$\begin{bmatrix} X & Y \\ Y^T & Z \end{bmatrix} \geq 0$$

ihen we have

$$-2a^T Nb \leq \inf_{X,Y,Z} \begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} X & Y-N \\ Y^T - N^T & Z \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Lemma 2 [8] *For matrices* $X_i, Y_i (1 \leq i \leq r)$ *and matrix* $S > 0$ *with appropriate dimensions, the following inequality is hold*

$$2\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{p=1}^{r}\sum_{l=1}^{r}\mu_i\mu_j\mu_p\mu_l X_{ij}^T S Y_{pl} \leq \sum_{i=1}^{r}\sum_{j=1}^{r}\mu_i\mu_j(X_{ij}^T S X_{ij} + Y_{ij}^T S Y_{ij})$$

where $\mu_i (1 \leq i \leq r)$ *denotes* $\mu_i(z(t)) \geq 0, \sum_{i=1}^{r}\mu_i(z(t)) = 1.$

Lemma 3 [4] *The LMI* $\begin{bmatrix} Y(x) & W(x) \\ * & R(x) \end{bmatrix} > 0$ *is equivalent to*

$$R(x) > 0,\ Y(x) - W(x)R^{-1}(x)W^T(x) > 0$$

where $Y(x) = Y^T(x), R(x) = R^T(x)$ *depend on* x.

Lemma 4 [6] *For constant* $\varepsilon > 0$ *and matrices* D, E, F, *satisfying* $F^T F \leq I$, *then the following inequality holds*

$$DEF + E^T F^T D^T \leq \varepsilon DD^T + \varepsilon^{-1} E^T E$$

Theorem 1 *For the given constants* $\alpha > 0, 1 \geq \beta \geq 0$ *and* $i, j = 1, 2, \ldots, r$, *if there exist positive-definite matrices* $P, Q, R \in R^{n\times n}$ *and matrices* $K_j \in R^{m\times n}$ *and* X_{ij}, Y_i, *with appropriate dimensions, such that the following matrix inequalities hold*

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ * & \Theta_{22} & \Theta_{23} \\ * & * & \Theta_{33} \end{bmatrix} < 0 \tag{24.5}$$

where

$$\Theta_{11} = P\bar{A}_i + \bar{A}_i^T P + Q + 2\alpha P + \tau X_{11} + \tau \bar{A}_i^T R \bar{A}_i + Y_1 + Y_1^T$$
$$\Theta_{12} = P\bar{A}_{\mathrm{di}} + \tau X_{12} + \tau \bar{A}_i^T R \bar{A}_{\mathrm{di}} + Y_2^T$$
$$\Theta_{13} = P\bar{B}_i K_j + \tau X_{13} + \tau \bar{A}_i^T R \bar{B}_i K_j - Y_1 + Y_3^T$$
$$\Theta_{22} = -e^{-2\alpha d} Q + \tau X_{22} + \tau \bar{A}_{\mathrm{di}}^T R \bar{A}_{\mathrm{di}}$$
$$\Theta_{23} = \tau \bar{A}_{\mathrm{di}}^T R \bar{B}_i K_j + \tau X_{23} - Y_2$$
$$\Theta_{33} = \tau K_j^T \bar{B}_i^T R \bar{B}_i K_j + \tau X_{33} - Y_3 - Y_3^T$$

with the controller (24.3), *the network control systems* (24.4) *is exponentially stable*.

Proof Choose a Lyapunov functional candidate for the system (24.4) as follows:

$$V(t) = x^T(t)Px(t) + \int_{t-d}^{t} x^T(s)Qe^{2\alpha(s-t)}x(s)ds + \int_{-\tau}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)\mathrm{d}s\mathrm{d}\theta$$

where P, Q, R positive-definite matrices in Theorem 1. Then, along the solution of system (24.4) we have

$$\begin{aligned}
&\dot{V}(t) + 2\alpha V(t) \\
&\quad = 2x^T(t)P\dot{x}(t) + x^T(t)Qx(t) - x^T(t-d)Qe^{-2\alpha d}x(t-d) + \tau\dot{x}^T(t)R\dot{x}(t) \\
&\qquad + 2\alpha x^T(t)Px(t) - \int_{t-\tau}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)ds
\end{aligned} \tag{24.6}$$

With

$$x(t) - x(t-\tau(t)) - \int_{t-\tau(t)}^{t} \dot{x}(s)\mathrm{d}s = 0$$

For any $4n \times n_{\mathrm{matrix}}$ $N = \begin{bmatrix} N_1^T & N_2^T & N_3^T \end{bmatrix}^T$, we know

$$0 = \xi^T(t)N[x(t) - x(t-\tau(t)) - \int_{t-\tau(t)}^{t} \dot{x}(s)\mathrm{d}s] \tag{24.7}$$

With Lemma 1 and (24.7), we obtain

$$0 \le 2\xi^T(t)Y[x(t) - x(t-\tau(t))] + \tau\xi^T(t)X\xi(t) + \int_{t-\tau}^{t} \dot{x}^T(s)Re^{2\alpha(s-t)}\dot{x}(s)\mathrm{d}s \tag{24.8}$$

Inserting (24.8) into (24.6), we have:

$$\begin{aligned}\dot{V}(t)+2\alpha V(t)\le&\sum_{i}^{r}\sum_{j}^{r}\mu_i(z(t))\mu_j(z(t))\{x^T(t)[P\bar{A}_i+\bar{A}_i^TP+Q+2\alpha P]x(t)\\&+2x^T(t)P\bar{A}_{\mathrm{di}}x(t-d)+2x^T(t)P\bar{B}_iK_jx(t-\tau(t))\\&-x^T(t-d)Qe^{-2\alpha d}x(t-d)+2\xi^T(t)Y[I\quad 0\quad -I]\xi(t)\\&+\tau\xi^T(t)X\xi(t)+\tau\dot{x}^T(t)R\dot{x}(t)\end{aligned}\tag{24.9}$$

With Lemma 2, we have

$$\begin{aligned}\tau\dot{x}^T(t)R\dot{x}(t)\le&\tau\sum_{i}^{r}\sum_{j}^{r}\mu_i(z(t))\mu_j(z(t))\xi^T(t)\\&\times\begin{bmatrix}\bar{A}_i^TR\bar{A}_i&\bar{A}_i^TR\bar{A}_{\mathrm{di}}&\bar{A}_i^TR\bar{B}_iK_j\\*&\bar{A}_{\mathrm{di}}^TR\bar{A}_{\mathrm{di}}&\bar{A}_{\mathrm{di}}^TR\bar{B}_iK_j\\*&*&K_j^T\bar{B}_i^TR\bar{B}_iK_j\end{bmatrix}\xi(t)\end{aligned}\tag{24.10}$$

Obviously

$$2\xi^T(t)Y[I\quad 0\quad -I]\xi(t)=\xi^T(t)\begin{bmatrix}Y_1+Y_1^T&Y_2^T&-Y_1+Y_3^T\\*&0&-Y_2\\*&*&-Y_3+Y_3^T\end{bmatrix}\xi(t)\tag{24.11}$$

Inserting (24.10)–(24.11) into (24.9), we obtain

$$\dot{V}(t)+2\alpha V(t)\le\sum_{i}^{r}\sum_{j}^{r}\mu_i(z(t))\mu_j(z(t))\xi^T(t)\Theta\xi(t)$$

With matrix inequality (24.5), we know

$$\dot{V}(t)<-2\alpha V(t)$$

therefore

$$V(t)<V(0)\mathrm{e}^{-2\alpha t}\le[\lambda_{\max}(P)+d\lambda_{\max}(Q)+\tau\lambda_{\max}(R)\bar{\psi}^2]||\psi(t)||^2\mathrm{e}^{-2\alpha t}\tag{24.12}$$

Obviously

$$V(t)\ge\lambda_{\min}(P)||x(t)||^2\tag{24.13}$$

From (24.12)–(24.13), we obtain

$$||x(t)||<\sqrt{\frac{\lambda_{\max}(P)+d\lambda_{\max}(Q)+\tau\lambda_{\max}(R)\bar{\psi}^2}{\lambda_{\min}(P)}}||\psi(t)||\mathrm{e}^{-\alpha t}$$

With the Lyapunov stability theorem and the above inequality, we know that the system (24.4) is exponentially stable.

Theorem 2 *For the given constants $\alpha > 0$ and $i,j = 1, 2, \ldots, q$, if there exist positive-definite matrices $\bar{P}, \bar{Q}, \bar{R} \in R^{n\times n}$ and matrices $\bar{K}_j \in R^{m\times n}$, $\bar{X}_{ij}, \bar{Y}_i$ with appropriate dimensions, such that the following linear matrix inequalities hold*

$$\Xi = \begin{bmatrix} \Xi_{11} & \Xi_{12} \\ * & \Xi_{22} \end{bmatrix} < 0 \tag{24.14}$$

where

$$\Xi_{11} = \begin{bmatrix} A_i\bar{P} + \bar{P}A_i^T + \bar{Q} + 2\alpha\bar{P} \\ +\tau\bar{X}_{11} + \bar{Y}_1 + \bar{Y}_1^T + \varepsilon_1 DD^T & A_{\mathrm{di}}\bar{P} + \tau\bar{X}_{12} + \bar{Y}_2^T & B_i\bar{K}_j + \tau\bar{X}_{13} - \bar{Y}_1 + \bar{Y}_3^T \\ * & -e^{-2\alpha d}\bar{Q} + \tau\bar{X}_{22} & \tau\bar{X}_{23} - \bar{Y}_2 \\ * & * & \tau\bar{X}_{33} - \bar{Y}_3 - \bar{Y}_3^T \end{bmatrix}$$

$$\Xi_{12} = \begin{bmatrix} \tau\bar{P}A_i^T & \bar{P}E_{i1}^T & \tau\bar{P}E_{i1}^T \\ \tau\bar{P}A_{\mathrm{di}}^T & \bar{P}E_{i2}^T & \tau\bar{P}E_{i2}^T \\ \tau\bar{K}_j^T B_i^T & \bar{K}_j^T E_{i3}^T & \tau\bar{K}_j^T E_{i3}^T \end{bmatrix} \Xi_{22} = \begin{bmatrix} -\tau\bar{R} + \varepsilon_2 DD^T & 0 & 0 \\ * & -\varepsilon_1 I & 0 \\ * & * & -\varepsilon_2 I \end{bmatrix} \quad \textit{with}$$

the controller $u(t) = \sum_i^r \mu_i(z(t))\bar{K}_i\bar{P}^{-1}x(t-\tau(t))$, the systems (24.4) is exponentially stable.

Proof

$$\Theta = \Theta_0 + \alpha^T \frac{1}{\tau} R^{-1} \alpha$$

where

$$\Theta_0 = \begin{bmatrix} P\bar{A}_i + \bar{A}_i^T P + Q + 2\alpha P \\ +\tau X_{11} + Y_1 + Y_1^T & P\bar{A}_{\mathrm{di}} + \tau X_{12} + Y_2^T & P\bar{B}_i K_j + \tau X_{13} - Y_1 + Y_3^T \\ * & -e^{-2\alpha d}Q + \tau X_{22} & \tau X_{23} - Y_2 \\ * & * & \tau X_{33} - Y_3 - Y_3^T \end{bmatrix}$$

$$\alpha = \begin{bmatrix} \tau R\bar{A}_i & \tau R\bar{A}_{\mathrm{di}} & \tau R\bar{B}_i K_j \end{bmatrix}$$

With Lemma 3, we know that inequality $\Theta < 0$ is equivalent to

$$\Sigma = \begin{bmatrix} P\bar{A}_i + \bar{A}_i^T P + Q + 2\alpha P \\ +\tau X_{11} + Y_1 + Y_1^T & P\bar{A}_{\mathrm{di}} + \tau X_{12} + Y_2^T & P\bar{B}_i K_j + \tau X_{13} - Y_1 + Y_3^T & \tau\bar{A}_i^T R \\ & -e^{-2\alpha d}Q + \tau X_{22} & {}_{23} - Y_2 & \tau\bar{A}_{\mathrm{di}}^T R \\ & & \tau X_{33} - Y_3 - Y_3^T & {}_j^T\bar{B}_i^T R \\ & & & -\tau R \end{bmatrix} < 0$$

With Lemma 4, we know that inequality $\Sigma < 0$ is equivalent to

$$\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ * & \Delta_{22} \end{bmatrix} < 0 \tag{24.15}$$

where

$$\Delta_{11} = \begin{bmatrix} PA_i + A_i^T P + Q + 2\alpha P \\ +\tau X_{11} + Y_1 + Y_1^T + \varepsilon_1 PDD^T P \\ * \\ * \end{bmatrix} \begin{bmatrix} PA_{\mathrm{di}} + \tau X_{12} + Y_2^T & PB_i K_j + \tau X_{13} - Y_1 + Y_3^T \\ -e^{-2\alpha d} Q + \tau X_{22} & \tau X_{23} - Y_2 \\ * & \tau X_{33} - Y_3 - Y_3^T \end{bmatrix}$$

$$\Delta_{12} = \begin{bmatrix} \tau A_i^T R & E_{i1}^T & \tau E_{i1}^T \\ \tau A_{\mathrm{di}}^T R & E_{i2}^T & \tau E_{i2}^T \\ \tau K_j^T B_i^T R & K_j^T E_{i3}^T & \tau K_j^T E_{i3}^T \end{bmatrix} \quad \Delta_{22} = \begin{bmatrix} -\tau R + \varepsilon_2 RDD^T R & 0 & 0 \\ * & -\varepsilon_1 I & 0 \\ * & * & -\varepsilon_2 I \end{bmatrix}$$

Pre- and Post-multiplying the inequality (24.15) by
diag$\{P^{-1}\ P^{-1}\ P^{-1}\ R^{-1}\ I\ I\}$,
and giving some transformations:

$$\bar{P} = P^{-1}\ \bar{Q} = P^{-1} Q P^{-1}\ \bar{K}_j = K_j P^{-1}\ \bar{X}_{ij} = P^{-1} X_{ij} P^{-1}\ \bar{Y}_i = P^{-1} Y_i P^{-1}\ \bar{R} = R^{-1}$$

We know that inequality $\Delta < 0$ is equivalent to (24.14). Therefore, the linear matrix inequality (24.14) is equivalent to (24.5).With Theorem 1, the systems (24.4) is exponentially stable.

24.4 Simulation

Consider the networked control systems in the form of (24.4), where

$$A_1 = \begin{bmatrix} 3 & -12 \\ 1 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix},\ A_{d1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix},\ A_{d2} = \begin{bmatrix} 0 & 0.1 \\ 0.2 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ B_2 = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix},\ D = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix},\ E_{11} = E_{12} = E_{13} = [0.1 \quad 0.2],$$

$$E_{21} = E_{22} = E_{23} = [0.01 \quad 0.1],\ E_{31} = E_{32} = E_{33} = [0.2 \quad 0.01]\ C_1 = [1 \quad 0]$$
$$C_2 = [1 \quad 0],\ F(t) = 0.1 \sin t,\ \tau = 0.5,\ \alpha = 0.1,\ d = 0.1$$

Solving the linear matrix inequality (24.14), we can obtain the gain matrix

$$K_1 = \bar{K}_1 \bar{P}^{-1} = [-6.8728\ 3.3598]$$
$$K_2 = \bar{K}_2 \bar{P}^{-1} = [11.2498\ -6.7263]$$

From the Theorem 2, we know that the systems (24.4) is mean-square exponentially stable.

24.5 Conclusion

This paper considers the exponential stability control problem for a class of nonlinear networked control systems (NCSs) with stochastic network-induced delay. A T-S fuzzy model is employed to represent the nonlinear controlled plant in the NCSs. Based on the Lyapunov stability theorem, the exponential stability condition and the state feedback fuzzy controller design method are given in terms of LMI.

Acknowledgments The author thanks the associate editor and the anonymous reviewers for their constructive comments and suggestions to improve the quality and the presentation of this paper. This work was supported by the National Nature Science Foundation under Project Code 61073065; Henan Province Nature Science Foundation under Project Code 092300410145.

References

1. Gao H, Chen T (2008) A new delay system approach to network-based control. Automatica 44:39–52
2. Liu G, Xia Y (2007) Design and stability criteria of networked predictive control systems with random network delay in the feedback channel. IEEE Trans Syst Man Cybernet Part C Appl Rev 37:173–184
3. Seiler P, Sengupta R (2007) An H_∞ approach to networked control. IEEE Trans Autom Control 50:356–364
4. Nikolakopoulos G, Panousopoulou A, Tzes A (2005) Multi-hopping induced gain scheduling for wireless networked controlled systems. In: Proceedings of the 44th IEEE conference on decision and control, and the European control conference, Seville, Spain, pp 470–475
5. Xia X, Zhang D, Zheng L, Zheng N (2008) Modeling and stabilization for a class of nonlinear networked control systems, a T-S fuzzy approach. Progr Nat Sci 18:1031–1037
6. Zhang H, Yang J, Su C (2007) T-S fuzzy model based robust H_∞ design for networked control systems with uncertainties. IEEE Trans Indust Inform 3:289–301
7. Walsh GC, Beldiman O, Bushnell LG (2001) Asymptotic behavior of nonlinear networked control systems. IEEE Trans Autom Control 46:1093–1097
8. Nesic D, Teel AR (2004) Input-to-state stability of networked control systems. Automatica 40:2121–2128

Chapter 25
A Nonlinear-Control Approach Study on Double-Zero Aluminum Foil Mill Profile

Taihua Wang, Yu Guo and Jingcheng Qu

Abstract Double zero aluminum foil rolling is a kind of technique that crowns almost all of the first-class technology of the current automatic control, which includes control of indexes such as speed, shapes and thickness in the process of double zero aluminum foil rolling. Among these, Profile control is the core technology of aluminum foil rolling, the key operation to enhance the finished product rate of aluminum foil and quality, and the fundamental condition to realize high speed rolling. Based on analyzing the traits of the foil profile, the author makes detailed probe into the control of flatness and thickness, which is the key technique of aluminum foil rolling, and analyzes and studies the nonlinear control of different actuators in the process of rolling.

Keywords Double-zero aluminum foil · Nonlinear control · Flatness · Profile control

25.1 Introduction

With the development of economy and technology, total improvement of the market environment and enhancement of people's living standards, both the society and individuals put forward higher demands on products' quality such as packaging and tightness. With qualities of high barrier, lightweight, good tightness

T. Wang (✉) · Y. Guo
School of Electrical Engineering and Automation, Henan Polytechnic University, 454003, Jiaozuo, China
e-mail: taihua@hpu.edu.cn

J. Qu
Shanhai ShenHuo Aluminium Foil Company Limited, 201319, Shanghai, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_25,

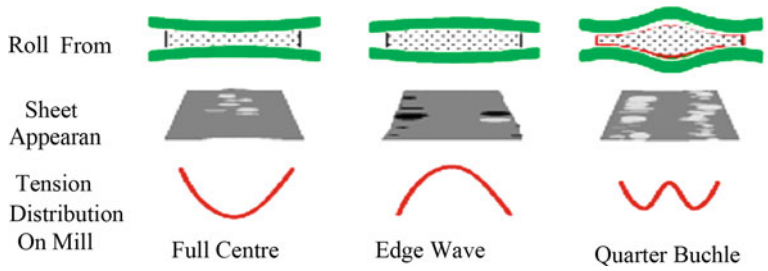


Fig. 25.1 Flatness defects

and uvioresistant, aluminum foil has become a new type of packaging material with market potential in the packing industry. While the users' requirements on the quality of aluminum foil has become increasingly higher, they hope this material becomes thinner and has better mechanics traits and surface quality [1, 2].

25.2 Analysis of the Traits of Aluminum Foil Profile

Profile usually refers to the flatness of the rolled material and whether partial waves, buckle folding, lateral bending and so on, appear on the transverse places of the foil material. The profile's quality depends on whether the extension beyond board band in the direction of width is equal.

25.2.1 The Appearance of Foil Profile's Defects

When aluminum foil is deformed unevenly, the tension in deformable body also distributes unevenly, which leads to the occurrence of additional tension, which remains in deformable body and becomes remaining tension after the deformation [3]. When interactions of the remaining stresses in the deformable body cannot counteract, and exceeds the amount of tension which can make the foil material keep foil surface rigid equilibrium, the aluminum foil under rolling will be in unstable shapes, shape defects such as full centre, edge wave, quarter buckle and so on appearing, and the flatness of the foil surface destroyed, as shown in Fig. 25.1.

Because the deformation in the milling deformed area is complex, the deformation of the rolled material is uneven and also nonlinear.

25.2.2 Flatness Model

The special data of the rolling machine are put into an offline model to predict flatness changing features of various controlling and actuators. The element varying stress of different zones on the profile instrument are all arrayed through

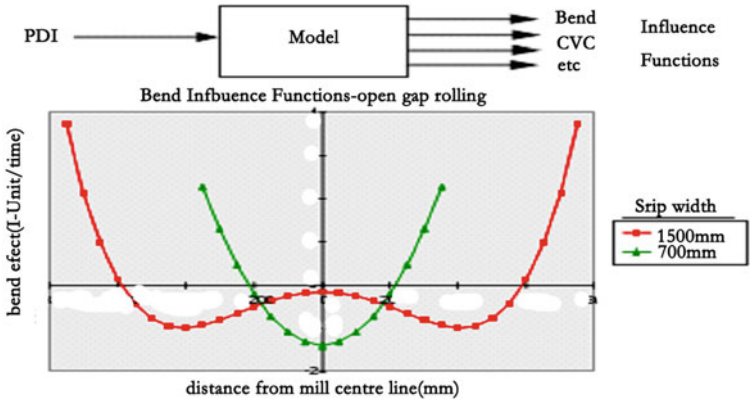


Fig. 25.2 Bending influence

these data, so as to form a suit of features for the width and thickness ranges of all products. This is a flatness model. It provides a suit of "influence function", which defines an element variable quantity of the flatness effects in different actuators. Figure 25.2 shows two kinds of typical bending roll influence functions of strip width and also emphasizes the significance of model function in providing prime control.

The effects that bending roll makes on wider strip and strips are quite different, so the bending roll effects cannot simply be evaluated as the shape of a parabola [4, 5].

25.2.3 Flatness Errors

In the process of cold rolling, profile or flatness errors occur if band shapes (camber) which enter rolling machine do not accord with that of roll gap. Flatness errors are decided by flatness and target flatness measured by profile instrument. How the target flatness is decided is shown in Fig. 25.3. The engineers define a set of profiles according to their experience, then use them to ascertain patterns of flatness target (shapes). All the target flatness should be defined under the condition of maximal band width and then processed based on the present band width so as to keep the same geometry shapes [6, 7].

25.3 The Control of Actuators

Most of the rolling pass of the products made by aluminum foil rolling machines are processed under no gap, which means high standards of products' thickness and flatness, so the technical control usually involves the control of thickness and shapes [8].

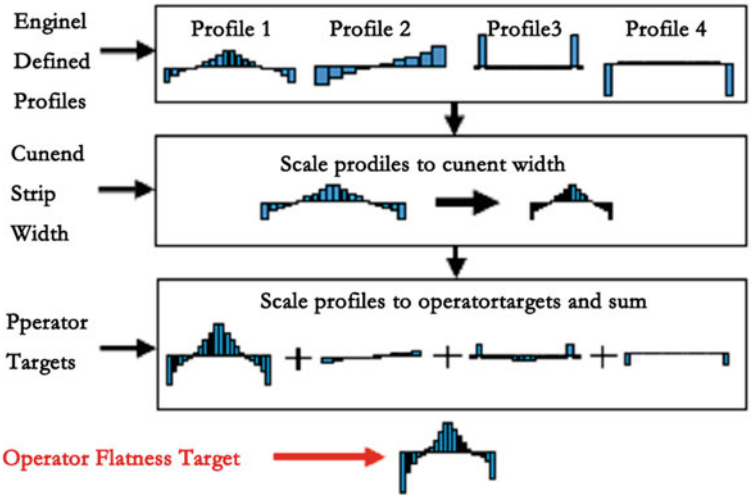
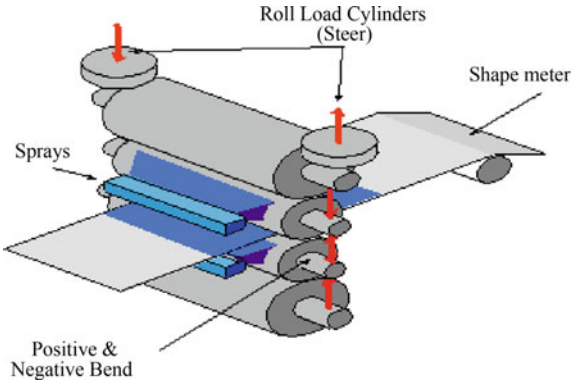


Fig. 25.3 Target generation

Fig. 25.4 Flatness control actuators



25.3.1 The Actuators of Profile

Actually, profile control is to control the transverse thickness differences of outlet aluminum foil Δh_b. The transverse thickness differences are dependant on the roll shape and roll forming of loaded roll gap when rolling. Figure 25.4 displays an operational institution which controls flatness. The main factors that influence roll shape are bending roll and sloping roll, and profile defects formed by the shrinkage because of the cold air are can be remedied by the hot spraying system, as indicated in Fig. 25.5.

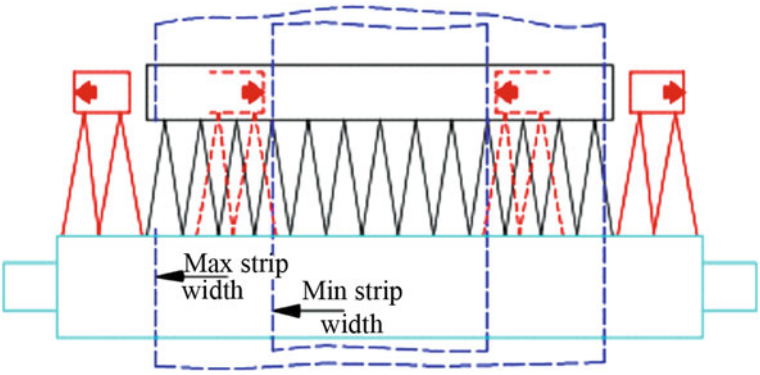
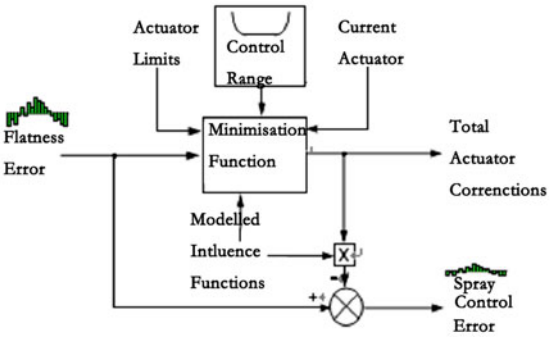


Fig. 25.5 Hot spraying

Fig. 25.6 Optimisation and spill over



25.3.2 Profile Control Strategies

Flatness is very sensitive to changes of sizes between two roll gap. In the process of flatness automatic control, the tension in the direction of width of the whole strip is measured, and its changes reflect the degree of tightness on strip section. It is explained as flatness, which is realized by creating emendation signals for actuators of roll gap to gain target flatness. Profile roll measures the tension model of strip in the process of rolling and sends tension signals to control system.

It is the controlling system that decides rolling machines' actuators and the direction of control movements which need to be controlled.

(1) **Optimisation and Spill over**

Optimization and spill over is the kernel of flatness controllers, which adjust rolling machines' actuators according to demands (general adjustment), so as to minimize flatness errors. Figure 25.6 displays the optimal scheme.

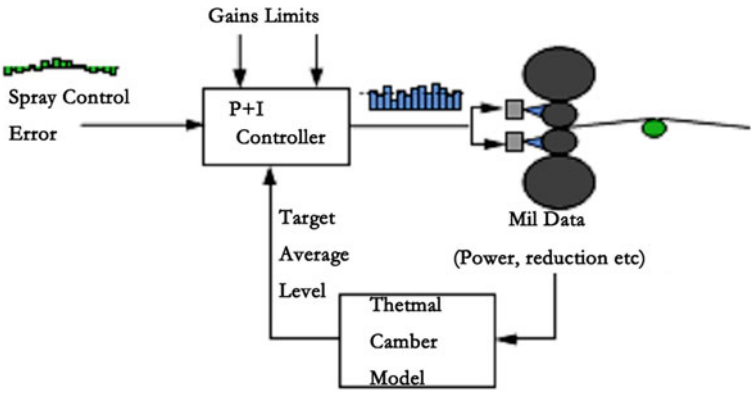


Fig.25.7 Spray control

Although two or more actuators have very similar flatness effects, this kind of control method must be kept "stable". This is the peculiar problem of rolling machines with multi-actuators.

(2) **Spray Control**

Cooling spray of roll plays an important role in controlling flatness, but this kind of control includes two cases, one solving residual local flatness error which cannot be corrected by mechanical actuator, the other controlling the hot camber which has material influence to flatness of strip edge. This kind of control strategy is shown in Fig. 25.7.

This control scheme includes a hot-camber model, which can ascertain the present camber according to rolling machines' data that is measured. Then this model decides the average cooling liquid level according to these data, so as to get the best quality under present rolling conditions. The spraying level can be regulated through a proportional addition integral controller which acts on the remaining flatness errors, and meanwhile makes sure to keep the target average level decided by the model. In addition, controllers' gain and spraying limits of every muzzle can be set, respectively, to flexibly realize the flatness control strategies of sapwood.

(3) **Feedback and Feed-Forward Control**

The feedback flatness control is realized by making use of the existing control and actuators of mechanical flatness and the cooling-spraying system of working roll to revise flatness errors that are measured.

While feedforward flatness control means to minimize the flatness errors caused by changes of the total rolling force (which usually results from controlling movements of strip's thickness) or changes of the thickness of strip at the entry. These flatness errors can be predicted from the measured rolling load, and the symmetrical actuators can be adjusted properly accordingly to the rolling force.

The automatic control of flatness can not only gain relatively high average rolling speed, but also reduce the amount of cooling liquid and the manual tension adjustment in the rolling process. What is most important is that it can improve the quality of products so as to gain better evenness and finished product rate.

25.4 Conclusion

Profile control is the core technology of aluminum foil rolling, the key operation to enhance the finished product rate of aluminum foil and quality, and the fundamental condition to realize high speed rolling. In the process of aluminum foil rolling, influencing factors of profile control are billet material, cross thickness difference, surface roughness of roller, hot camber of roller, elastic bucking and bending of roller and so on. The thickness of profile is a nonlinear change with rolling speed, stretching force and reduction in pass. So there are different nonlinear control methods in production processes such as the transverse thickness difference control, spray control, feedback and feedforward control etc. In this chapter, these control methods are discussed.

References

1. Chen J-X, Fei S-C, Zhang Q-F (2007) Discuss on aluminum foil flatness control method in rolling. Alum Fabr 14(6):31–34
2. Xiang Z-L (2007) Aluminium foil flatness control in rolling. Light Alloy Fabr Technol 35(4):18–21
3. Liu Y, Fu Q-D (2006) Usual sheet shape defects and control in rolling process of Φ670/Φ254*1700 mm aluminum foil mill. Sichuan Nonferrous Metals 26(2):35–39
4. Zhao L-J, Liu X, Liu J (2006) The application research of bending roll system in the gap control of four-roll rolling mill. Mod Manufact Eng 33(11):70–73
5. Liu Y, Fu Q-D, Chen Y (2005) Flatness characteristics and flatness controlling for 1700 mm aluminum foil mill. Alum Fabr 30(1):31–34
6. Ben H-Y, Xin D-F (2006) The influence of foil mill accuracy on strip flatness. Light Alloy Fabr Technol 34(10):25–29
7. Xiao Y-Q (2005) Application manual of aluminum fabrication technology. Metallurgical Industry Press, Beijing
8. Zhong M (2003) Analysis and study on foil roller cooling and the thermal crown adjustment. Alum Fabr 28(3):38–41

Chapter 26
Multilinear Model-Based PI Control of Block-Oriented Nonlinear Systems

Jingjing Du, Xinliang Zhang and Chunyue Song

Abstract In order to overcome the drawbacks of the conventional nonlinearity inversion control method for block-oriented systems, a Multi-PI control method is proposed. The virtue of the proposed method is that the classic PI control algorithm is applied to complex nonlinear systems, which largely simplifies the control problems and improves control performances. Simulations demonstrate the effectiveness of the proposed Multi-PI control method for block-oriented systems with strong nonlinearity.

Keywords Block-oriented systems · Included angle dividing method · Multi-PI control · Nonlinearity inversion control method

26.1 Introduction

Block-oriented nonlinear model structures, chiefly the Hammerstein and Wiener model structures consist of the cascade connection of a linear dynamic block and a nonlinear static block [1], which makes the system analysis and control design

J. Du (✉) · X. Zhang
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: jjdu@hpu.edu.cn

X. Zhang
e-mail: zxldq@hpu.edu.cn

C. Song
State Key Laboratory of Industrial Control Technology,
Institute of Industrial Process Control, Zhejiang University,
Hangzhou 310027, China
e-mail: cysong@iipc.zju.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_26,

easier. The conventional control strategy of Hammerstein models is the nonlinearity inversion method [2, 3], which makes full use of advantages of the block structure, and is easy and effective in certain cases. However, this method needs the static nonlinear element to be invertible, which is generally not the case, especially when the nonlinearity systems exhibit input or output multiplicity. Moreover, nonlinear inversion may lead to performance degradation [4]. So other ways have been tried to overcome this disadvantage. In Ref. [4], the static input nonlinearity is transformed into a polytopic description, and then a linear MPC constrained to LMIs is designed. Whereas, the method also needs the nonlinear element to be invertible. In Ref. [3], an NMPC based on sensitivity analysis is proposed for block-oriented system. However, the sensitivity calculation involves solution of a series of partial differential equations, which makes the whole method complicated.

In order to overcome the shortcomings of the nonlinear inversion method, a Multi-PI control method is proposed for block-oriented systems in this paper. In the proposed method, a complex nonlinear control problem is decomposed into a set of simple linear control problems, which largely reduces the difficulty of the control problem, solves the control of systems with I/O multiplicity, and avoids the performance degradation in the nonlinear inversion method.

26.2 Block-Oriented Systems

Block-oriented model structures, basically Hammerstein models and Wiener models, are popular nonlinear empirical modeling structures. As depicted in Fig. 26.1, a Hammerstein model is a block-oriented nonlinear model, which consists of the cascade structure of a static nonlinear function $f(\cdot)$ followed by a linear dynamic block $H(z)$, whereas a Wiener model contains the same elements in the reverse order [1].

Hammerstein model and Wiener model are described by Eqs. (26.1) and (26.2).

$$\begin{cases} x(k) = f(u(k)) \\ y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{i=1}^{m} b_i x(k-i) \end{cases} \tag{26.1}$$

$$\begin{cases} w(k) = \sum_{i=1}^{n} a_i w(k-i) + \sum_{i=1}^{m} b_i u(k-i) \\ y(k) = f(w(k)) \end{cases} \tag{26.2}$$

where, x and w are intermediate variables; and $f(\cdot)$ is analytic and not necessarily invertible in this paper.

Hammerstein and Wiener model structures can effectively represent and approximate many industrial processes, as they may account for nonlinear effects encountered in most chemical processes [5]. For example, pH neutralization

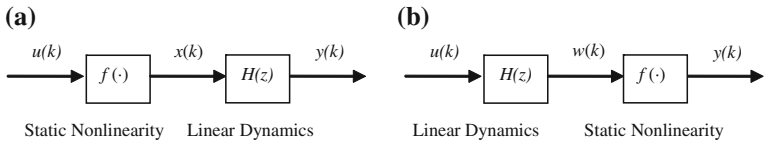


Fig. 26.1 **a** The structure of a Hammerstein model. **b** The structure of a Wiener model

processes, distillation columns, heat exchangers, polymerization reactor, and dryer processes have been modeled by Hammerstein or Wiener models [6].

26.3 Multi-PI Control of Block-Oriented Systems

26.3.1 Multilinear Models of Block-Oriented Systems

The multi-linear model approach has proved to be useful in dealing with nonlinear control problems, and has attracted a lot of attention and studied extensively in the past years. The major motivation for the multi-linear modeling approach is that local linear modeling is simpler than global modeling because locally there are less relevant phenomena, and interactions are simpler, and that the classic control techniques can be used to simplify the nonlinear control problems. Obviously, it is applicable to block-oriented systems and may have potential advantages over other control methods, such as the nonlinear inversion method. Recently, a dividing method is proposed for SISO block-oriented systems, which can effectively divide a nonlinear system into a set of well-approximating linear sub-models [7]. The key point of the dividing method is to divide the nonlinear system according to the slope variation of its static I/O curve. Then multi-linear controllers can be designed to avoid the shortcomings in the common nonlinearity inversion method.

The Hammerstein model (26.1) can be rewritten in the following NARMAX form [1, 7].

$$x(k) = f(u(k)) \tag{26.3}$$

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + \ldots + a_n y(k-n) + b_1 x(k-1) + b_m x(k-m) \tag{26.4}$$

where, $f(\cdot)$ is analytic and not necessarily invertible.

The intermediate variable x can be removed from Eqs. (26.3)–(26.4). Substituting Eq. (26.3) into Eq. (26.4), we get the following NARMAX equation.

$$\begin{aligned} y(k) = {} & a_1 y(k-1) + a_2 y(k-2) \\ & + \ldots + a_n y(k-n) + b_1 f(u(k-1)) + b_2 f(u(k-2)) + \ldots + b_m f(u(k-m)) \end{aligned} \tag{26.5}$$

Then the included angle dividing method is employed to divide the operating space Ω of the Hammerstein system (26.5) into a set of subspaces Ω_i, $i = 1,2, \ldots, m$. Each subspace has an operating point (u_{0i}, y_{0i}) in it.

$$\begin{aligned} y(k) - a_1 y(k-1) - \ldots - a_n y(k-n) &= b_1 f(u_{0i}(k-1)) + b_2 f(u_{0i}(k-2)) \\ &+ \ldots + b_m f(u_{0i}(k-m)) + b_1 f'(u_{0i}(k-1))\Delta u(k-1) \\ &+ b_2 f'(u_{0i}(k-2))\Delta u(k-2) + \ldots + b_m f'(u_{0i}(k-m))\Delta u(k-m) \end{aligned} \tag{26.6}$$

At operating point (u_{0i}, y_{0i}), the following three equations exist.

$$\begin{aligned} &y_{0i}(k) = y_{0i}(k-1) = \ldots = y_{0i}(k-n), \\ &u_{0i}(k) = u_{0i}(k-1) = \ldots = u_{0i}(k-m), \\ &y_{0i}(k) = a_1 y_{0i}(k-1) + \ldots a_n y_{0i}(k-n) + \ldots b_1 f(u_{0i}(k-1)) \\ &+ \ldots + b_m f(u_{0i}(k-m)), \end{aligned} \tag{26.7}$$

Substituting Eq. (26.7) into Eq. (26.6), we get:

$$\begin{aligned} &\Delta y(k) - a_1 \Delta y(k-1) - \ldots - a_n \Delta y(k-n) \\ &\quad = b_1 f'(u_{0i}(k-1))\Delta u(k-1) + \ldots + b_m f'(u_{0i}(k-m))\Delta u(k-m) \end{aligned} \tag{26.8}$$

So we obtain the transfer function of Eq. (26.8) as follows:

$$G_{\mathrm{Hi}}(z) = \frac{b_1 z^{-1} + \ldots + b_m z^{-m}}{1 - a_1 z^{-1} - \ldots - a_n z^{-n}} f'(u_{0i}(k)),\ i = 1, 2, \ldots, m. \tag{26.9}$$

In the same way, we get the submodels of Wiener system (26.2) as Eq. (26.10):

$$G_{\mathrm{Wi}} = \frac{b_1 z^{-1} + b_2 z^{-2} + \ldots b_3 z^{-m}}{(1 - a_1 z^{-1} - a_2 z^{-2} - \ldots - a_n z^{-n})} \times \frac{1}{g'(y_0(k))},\ i = 1, 2, \ldots, m. \tag{26.10}$$

where, without of loss of generality, we suppose $f(\cdot)$ is invertible, and its inverse function is $g(\cdot)$.

26.3.2 Multi-PI Control Algorithm for Block-Oriented Systems

After dividing and linearization, we get a set of linear models Eq. (26.9) or (26.10), which approximate the nonlinear system (26.1) or (26.2). The control problem of a nonlinear system is transformed into a series of linear subsystems, which can be easily and effectively solved by traditional control techniques.

For each linear model in Eq. (26.9) or (26.10), a PI controller is designed as in Eq. (26.11)–(26.12).

$$u_i(k) = u_{0i} + \Delta u_i(k),\ i = 1, 2, \ldots, m \tag{26.11}$$

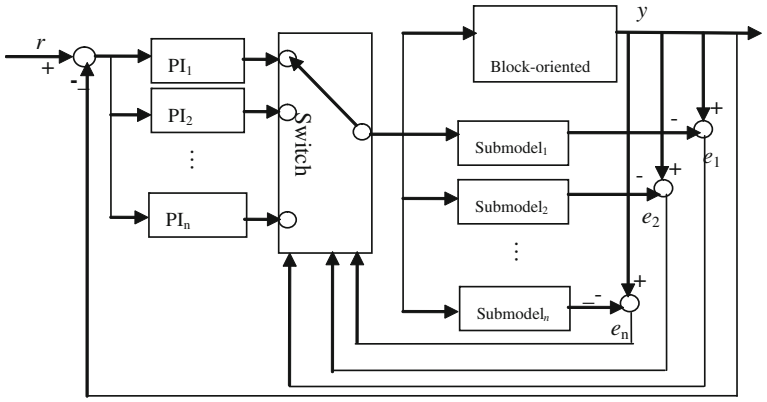


Fig. 26.2 Multi-PI control of a block-oriented system

$$\Delta u_i(k) = K_{\mathrm{Pi}} e(k) + K_{\mathrm{Ii}} \sum_{j=1}^{k} e(j),\ i = 1, 2, \ldots, m \tag{26.12}$$

where, k is the time step; K_{Pi}, K_{Ii} are the proportional, integral coefficients for the ith PI controller, respectively. The m PI controllers are scheduled by hard switching according to the operating conditions, as shown in Fig. 26.2.

26.4 Simulations

26.4.1 pH Process

The pH process [2] is a typical Wiener type system where the material balance differential equations are almost linear and the equilibrium equation (titration curve) is a strong nonlinear static function. Consider the pH process in Ref. [2], which exhibits strong static nonlinearity (output nonlinearity), and the nonlinearity inversion control method is inappropriate. Here the Multi-PI controller is designed. Figures 26.3 and 26.4 show the good performances of the Multi-PI controller, which validates its effectiveness for block-oriented systems with output multiplicity.

26.4.2 Heat-Exchanger

Consider a Heat Exchanger process [8] modeled by a Hammerstein model which exhibits strong static nonlinearity—input multiplicity. Thus, the usual nonlinearity inversion control method is not applicable to it. In this paper, the Multi-PI control

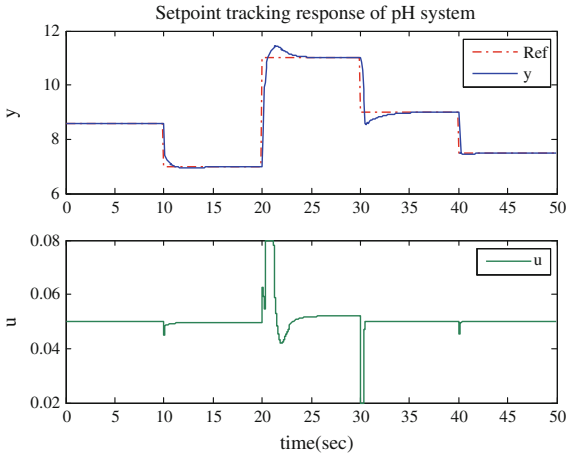


Fig. 26.3 Setpoint tracking of pH process

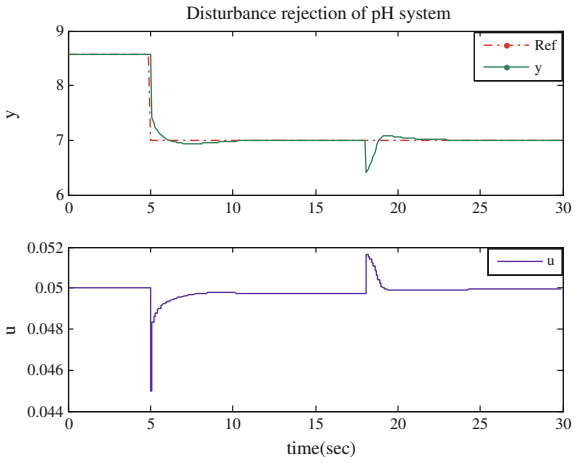


Fig. 26.4 Disturbance rejecting of pH process

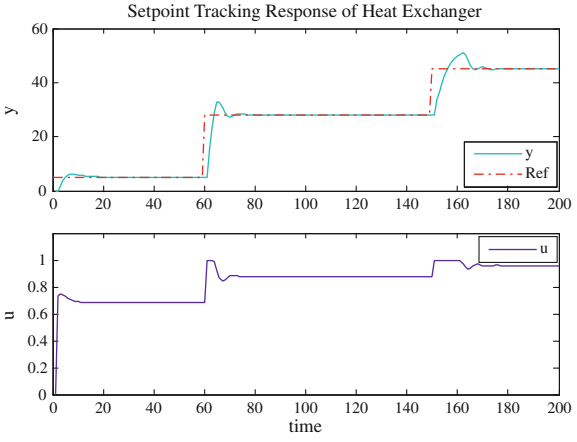


Fig. 26.5 Set-point tracking of heat exchanger

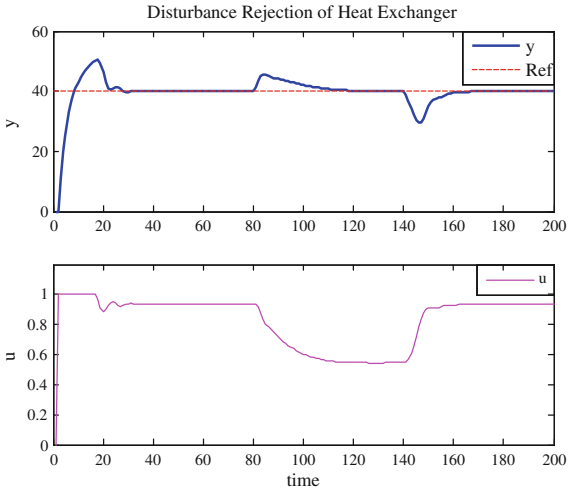


Fig. 26.6 Disturbance rejecting of heat exchanger

is applied to it, and the control performances are displayed in Figs. 26.5 and 26.6, which illustrate the effectiveness of the Multi-PI controller for the Heat Exchanger process with input multiplicity.

26.5 Conclusions

Block-oriented models chiefly denote Hammerstein and Wiener models, which have been successfully applied to modeling of many industrial processes. However, the general nonlinearity inversion control method for block-oriented models needs the static element to be invertible, which mostly is not satisfied, especially for systems with I/O multiplicity. Moreover, the nonlinearity inversion method may lead to a degraded performance. To improve the situation, a special purpose Multi-PI control method is proposed for block-oriented systems in this paper. In our method, a nonlinear control problem is decomposed into a set of linear control ones, to which the conventional PI control strategy can be easily applied. Thus, a complex problem is solved in an easy way. Simulations illustrate the effectiveness of the Multi-PI controller.

Acknowledgments This work is supported partially by the NSF (60974023) of China, partially by the Doctors' Funds (B2011-007, B2010-88) of Henan Polytechnic University, and partially by the Fundamental Research Funds for the Central Universities.

References

1. Pearson RK (2006) Nonlinear empirical modeling techniques. Comput Chem Eng 30:1513–1528
2. Sung S, Lee J (2004) Modeling and control of Wiener-type processes. Chem Eng Sci 59:1515–1521

3. Harnischmacher G, Marquardt W (2007) Nonlinear model predictive control of multivariable processes using block-structured models. Control Eng Pract 15:1238–1256
4. Bloemen HHJ, van den Boom TJJ, Verbruggen HB (2000) Model-based predictive control for Hammerstein systems. In: Proceedings of 39th IEEE conference on decision and control, Sydney, Australia, pp 4963–4966
5. Fruzzetti KP, Palazoglu A, McDonald KA (1997) Nonlinear model predictive control using Hammerstein models. J Process Control 7:31–41
6. Jia L, Chiu MS, Ge SS (2005) A noniterative neruo-fuzzy based identification method for Hammerstein processes. J Process Control 15:749–761
7. Du J, Song C, Li P (2009) Multilinear model control of Hammerstein-like systems based on included angle dividing method and the MLD-MPC strategy. Ind Eng Chem Res 48: 3934–3943
8. Eskinat E, Johnson SH (1991) Use of Hammerstein models in identification of nonlinear systems. AIChE J 37:255–266

Chapter 27
Robust Stability of Stochastic Systems with Time-Delay and Nonlinear Uncertainties

Wei Qian and Lin Chen

Abstract This paper is concerned with the robust exponential stability in mean square for stochastic systems with time delay and nonlinear uncertainties. Based on Ito calculus rules, a more general Lyapunov–Krasovskii functional is constructed and a novel delay-dependent stability criteria is obtained in terms of linear matrix inequalities. Numerical examples are given to demonstrate that the proposed methods are less conservative than the existing ones.

Keywords Stochastic systems · Time-delay · Nonlinear uncertainties · Exponential stability · Linear matrix inequality (LMI)

27.1 Introduction

Robust stability of uncertain time-delay systems has attracted a great deal of attention over the years. Different methods have been used to reduce the conservatism of the corresponding results and many delay-dependent stability criteria have been reported in the literature [1–5].

Recently, considerable attention has been devoted to the study of stochastic systems due to the fact that stochastic modeling has played an important role in science and industry. Many fundamental results for deterministic systems have

W. Qian (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: qwei@hpu.edu.cn

L. Chen
Department of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: chenlin@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_27,

been extended to stochastic systems. In the aspect of robust stability of stochastic systems, in [6], the robust exponential stability of stochastic differential delay equations was discussed based on Lyapunov's direct method. In [7], in terms of linear matrix inequality, a less conservative robust stability condition was obtained for stochastic delay systems with nonlinear uncertainties. By using a descriptor model transformation of the system and by applying Moon's inequality, exponential stability of uncertain stochastic systems with multiple delays was studied in [8], and a delay-dependent result was proposed. By introducing some relaxation matrices, the exponential stability of stochastic systems with time-varying delay, nonlinearity, and markovian switching was investigated in [9].

This note studies the exponential stability of stochastic systems with time delay and nonlinear uncertainties. Based on Ito calculus rules for stochastic systems, a more general Lyapunov–Krasovskii functional is constructed, and as a result, a novel delay-dependent sufficient condition for exponential stability in mean square is derived in terms of LMIs. Examples are also given to illustrate the effectiveness of the method.

27.2 Problem Formulation

Consider the following stochastic systems with time delay and nonlinear uncertainties

$$\begin{aligned} \mathrm{d}x(t) &= [Ax(t) + A_1x(t-\tau) + f(t,x(t),x(t-\tau))]\mathrm{d}t + g(t,x(t),x(t-\tau))\mathrm{d}w(t) \\ x(t) &= \phi(t), t \in [-\tau, 0] \end{aligned} \tag{27.1}$$

where $x(t) \in \Re^n$ is the state, A and A_1 are known constant matrices, τ is the delay. $w(t)$ is the one-dimensional standard Brownian motion satisfying:

$$\boldsymbol{E}\{dw(t)\} = 0, \quad \mathrm{E}\left\{dw(t)^2\right\} = dt$$

$f : \Re_+ \times \Re^n \times \Re^n \to \Re^n$ and $g : \Re_+ \times \Re^n \times \Re^n \to \Re^{n\times m}$ denote the nonlinear uncertainties satisfying

$$\|f(t,x(t),x(t-\tau))\| \le \|G_1x(t)\| + \|G_2x(t-\tau)\| \tag{27.2}$$

$$\mathrm{Trace}\left[g^T(t,x(t),x(t-\tau))g(t,x(t),x(t-\tau))\right] \le \|G_3x(t)\|^2 + \|G_4x(t-\tau)\|^2 \tag{27.3}$$

where G_1, G_2, G_3, G_4 are constant matrices. Initial function $\varphi(\cdot) \in L^2_{F0}([-\tau, 0]; \Re^n)$.

For a given function V, its infinitesimal operator L is defined as

$$LV(x(t),t) = \lim_{\Delta\to 0^+} \frac{1}{\Delta}[E(V(x(t+\Delta),t+\Delta)) - V(x(t),t)] \tag{27.4}$$

The aim of this note is to develop a delay-dependent stability condition in the mean square for system (27.1). For this purpose, the following definition and lemma are introduced.

Definition 1 System (27.1) is said to be exponentially stable in mean square if there exists a positive constant λ such that

$$\lim_{t\to\infty} \sup \frac{1}{t} \log\left[\mathrm{E}\|x(t)\|^2\right] \leq -\lambda$$

Lemma 1 [10] *For any scalars α, β with $\alpha > \beta$, symmetric positive-definite constant matrix M, and vector-valued function $w : [\beta, \alpha] \to \Re^n$ such that the integrations concerned are well-defined, then*

$$\left(\int_{\beta}^{\alpha} w(\sigma)\mathrm{d}\sigma\right)' M \left(\int_{\beta}^{\alpha} w(\sigma)\mathrm{d}\sigma\right) \leq (\alpha - \beta) \int_{\beta}^{\alpha} w'(\sigma) M w(\sigma)\mathrm{d}\sigma$$

27.3 Main Results

Consider a Lyapunov–Krasovskii functional candidate as

$$V(x(t), t) = \sum_{i=1}^{5} V_i(x(t), t) \tag{27.5}$$

where

$$V_1(x(t), t) = x^T(t) P x(t) \tag{27.6}$$

$$V_2(x(t), t) = \int_{t-\tau}^{t} x^T(s) R_1 x(s) ds \tag{27.7}$$

$$V_3(x(t)) = \int_{t-\tau}^{t} \int_{s}^{t} x^T(u) R_2 x(u) \mathrm{d}u \mathrm{d}s \tag{27.8}$$

$$V_4(x(t), t) = 2x^T(t) Q_1 \int_{t-\tau}^{t} x(s) ds \tag{27.9}$$

$$V_5(x(t), t) = \int_{-\tau}^{0} ds \int_{-\tau}^{0} x^T(t+s) Q_2 x(t+\theta) d\theta \tag{27.10}$$

Theorem 1 *If matrices* $P = P^T, R_1 = R_1^T, R_2 = R_2^T, Q_1, Q_2 = Q_2^T \in \Re^{n\times n}$ *satisfies the following LMIs*

$$R_1 > 0 R_2 > 0 \begin{bmatrix} P & Q_1 \\ Q_1^T & \tau^{-1}R_1 + Q_2 \end{bmatrix} > 0 \tag{27.11}$$

then the Lyapunov–Krasovskii functional (27.5) *is positive definite.*

Proof Let $y(t,\tau) := \int_{t-\tau}^{t} x(s)ds$, then we can get:

$$V_4(x(t),t) = 2x^T(t)Q_1 y(t,\tau)$$

$$V_5(x(t),t) = y^T(t,\tau)Q_2 y(t,\tau)$$

From Lemma 1, we have:

$$V_2(x(t),t) \geq \tau^{-1}\left[\int_{t-\tau}^{t} x(s)\mathrm{d}s\right]^T R_1 \left[\int_{t-\tau}^{t} x(s)\mathrm{d}s\right] = \tau^{-1}y^T(t,\tau)R_1 y(t,\tau)$$

Noting (27.6) and (27.8), we know that if the LMIs (27.11) hold, the Lyapunov–Krasovskii functional (27.5) is positive definite. This completes the proof.

Theorem 2 *For any given* τ, *the stochastic system* (27.1) *is exponentially stable in the mean square. If there exist positive scalars* $\rho > 0$ $\varepsilon_i > 0, i = 1,2,3,4$ *and matrices* $P = P^T, R_1 = R_1^T, R_2 = R_2^T, Q_1, Q_2 = Q_2^T \in \Re^{n\times n}$ *satisfying Theorem* 1, *such that the following LMIs hold*

$$\begin{bmatrix} \Pi & PA_1 - Q_1 & \tau A^T Q_1 + \tau Q_2 & P \\ * & \rho G_4^T G_4 + \varepsilon G_2^T G_2 - R_1 & \tau A_1^T Q_1 - \tau Q_2 & 0 \\ * & * & -\tau R_2 & \tau Q_1^T \\ * & * & * & -\varepsilon I \end{bmatrix} < 0 \tag{27.12}$$

$$P \leq \rho I \tag{27.13}$$

where $\Pi = PA + A^T P + R_1 + \tau R_2 + Q_1 + Q_1^T + \varepsilon G_1^T G_1 + \rho G_3^T G_3$.

Proof Choosing the Lyapunov–Krasovskii functional (27.5), by Ito calculus rules, the infinitesimal operator $LV(x(t),t)$ of the stochastic process $\{x_t, t \geq 0\}$ along the trajectory of system (27.1) are given by

$$\begin{aligned} LV_1(x(t),t) &= \lim_{\Delta\to 0^+}\frac{1}{\Delta}[E(V_1(x(t+\Delta),t+\Delta)) - V_1(x(t),t)] \\ &= 2x^T(t)P[Ax(t) + A_1x(t-\tau) + f(t,x(t),x(t-\tau))] \\ &\quad + \text{Trace}\left[g^T(t,x(t),x(t-\tau))Pg(t,x(t),x(t-\tau))\right] \end{aligned}$$

By (27.3) and (27.13), we get

$$\begin{aligned} LV_1(x(t),t) \le\ & 2x^T(t)PAx(t) + 2x^T(t)PA_1x(t-\tau) + 2x^T(t)Pf(t,x(t),x(t-\tau)) \\ & +\rho\left[x^T(t)G_3^TG_3x(t) + x^T(t-\tau)G_4^TG_4x(t-\tau)\right] \end{aligned} \tag{27.14}$$

$$LV_2(x(t),t) = x^T(t)R_1x(t) - x^T(t-\tau)R_1x(t-\tau) \tag{27.15}$$

$$LV_3(x(t),t) = \tau x^T(t)R_2x(t) - \int_{t-\tau}^{t} x^T(s)R_2x(s)\mathrm{d}s \tag{27.16}$$

$$\begin{aligned} LV_4(x(t),t) &= \lim_{\Delta\to 0^+}\frac{1}{\Delta}[E(V_4(x(t+\Delta),t+\Delta)) - V_4(x(t),t)] \\ &= 2x^T(t)Q_1x(t) - 2x^T(t)Q_1x(t-\tau) + 2x^T(t)A^TQ_1\int_{t-\tau}^{t} x(s)\mathrm{d}s \\ &\quad + 2x^T(t-\tau)A_1^TQ_1\int_{t-\tau}^{t} x(s)\mathrm{d}s + 2f^T(t,x(t),x(t-\tau))Q_1\int_{t-\tau}^{t} x(s)\mathrm{d}s \end{aligned} \tag{27.17}$$

$$\begin{aligned} LV_5(x(t),t) &= 2[x(t) - x(t-\tau)]^TQ_2\int_{t-\tau}^{t} x(s)ds \\ &= 2x^T(t)Q_2\int_{t-\tau}^{t} x(s)ds - 2x^T(t-\tau)Q_2\int_{t-\tau}^{t} x(s)ds \end{aligned} \tag{27.18}$$

Noting (27.2), we can calculate that, for any $\varepsilon > 0$, it has

$$0 \le -\varepsilon f^T(t)f(t) + \varepsilon x^T(t)G_1^TG_1x(t) + \varepsilon x^T(t-\tau)G_2^TG_2x(t-\tau) \tag{27.19}$$

Combining (27.14)–(27.19), we have

$$LV(x(t),t) = \sum_{i=1}^{5} LV_i(x(t),t) \le \frac{1}{\tau}\int_{t-\tau}^{t} \eta^T(t,s)\Omega\eta(t,s)\mathrm{d}s \tag{27.20}$$

where

$$\eta^T(t,s) = \left[x^T(t) \quad x^T(t-\tau) \quad x^T(s) \quad f^T(t,x(t),x(t-\tau)) \right]$$

We know that LMI (27.12) is equivalent to $\Omega < 0$. Moreover, it is easy to see that there exists a scalar $\mu > 0$ such that

$$LV(x(t),t) \leq -\mu \|x(t)\|^2 \tag{27.21}$$

Therefore, by Dynkin formula, we have

$$EV(x(t),t) = EV(x(0),0) + E\int_0^t LV(x(s),s)\mathrm{d}s \tag{27.22}$$

It can also be observed that

$$V_2(x(t),t) \leq \lambda_{\max}(R_1) \int_{-\tau}^{0} \|x(t+\theta)\|^2 \mathrm{d}\theta$$

$$V_3(x(t)) \leq \tau\lambda_{\max}(R_2) \int_{-\tau}^{0} \|x(t+\theta)\|^2 \mathrm{d}\theta$$

$$V_4(x(t),t) \leq \lambda_{\max}(Q_1 Q_1^T)\|x(t)\|^2 + \int_{-\tau}^{0} \|x(t+\theta)\|^2 \mathrm{d}\theta$$

$$V_5(x(t),t) \leq \tau\lambda_{\max}(Q_2) \int_{-\tau}^{0} \|x(t+\theta)\|^2 \mathrm{d}\theta$$

So we can deduce that

$$EV(x(0),0) \leq (\alpha + \tau\beta) \sup_{-\tau \leq \theta \leq 0} E\|\xi(\theta)\|^2 \tag{27.23}$$

$$EV(x(t),t) \geq \lambda_{\min}(P)\|x(t)\|^2 \tag{27.24}$$

where $\alpha = \lambda_{\max}(P) + \lambda_{\max}\left(Q_1 Q_1^T\right)$, $\beta = 1 + \lambda_{\max}(R_1) + \tau\lambda_{\max}(R_2) + \tau\lambda_{\max}(Q_2)$.

By Gronwell Inequality, we can get

$$E\|x(t)\|^2 \leq \frac{(\alpha + \tau\beta)}{\lambda_{\min}(P)} \sup_{-\tau \leq \theta \leq 0} E\|\xi(\theta)\|^2 \exp\left(-\frac{\mu}{\lambda_{\min}(P)} t\right) \tag{27.25}$$

This completes the proof.

Table 27.1 Comparison of delay-dependent stability conditions of Example 1

Methods	Maximum τ allowed
Mao [6]	0.175
Yue and Won [7]	0.8635
Chen and Guan [8]	1.1997
Theorem 2	2.0859

Remark 1 $LV(x(t),t)$ is obtained based on Ito calculus rules and the independent increments property of Brownian motion, which is different from the calculus rules in deterministic systems. It is the difference that makes the difficulty in generalizing the analysis methods for time-delay systems to the stochastic case.

Remark 2 When both stochastic mode and the nonlinear uncertainties of (27.1) are removed, system (27.1) becomes the traditional linear time-delay system, by the same method, the stability and robust stability criterion can be obtained easily.

Corollary 1 *Consider the following linear time-delay system*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + A_1 x(t-\tau) \\ x(t) &= \varphi(t), t \in [-\tau, 0] \end{aligned} \tag{27.26}$$

For any given τ, *the linear time-delay system is exponentially stable, if there exist positive scalars* $\varepsilon > 0$ *and matrices* $P = P^T, R_1 = R_1^T, R_2 = R_2^T, Q_1, Q_2 = Q_2^T \in \Re^{n\times n}$ *satisfying Theorem 1, such that the following LMI hold*

$$\begin{bmatrix} PA + A^T P + R_1 + \tau R_2 + Q_1 + Q_1^T & PA_1 - Q_1 & \tau A^T Q_1 + \tau Q_2 \\ * & -R_1 & \bar{\tau} A_1^T Q_1 - \tau Q_2 \\ * & * & -\tau R_2 \end{bmatrix} < 0$$

27.4 Numerical Examples

Example 1 Consider the stochastic delay system (27.1) with

$$A = \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix}, A_1 = \begin{bmatrix} -1 & 0 \\ -0.5 & -1 \end{bmatrix}, \|\Delta A(t)\| \le 0.1, \|\Delta A_1(t)\| \le 0.1$$

$$\text{Trace}\left[g^T g\right] \le 0.1\|x(t)\|^2 + 0.1\|x(t-\tau)\|^2$$

Table 27.2 Comparison of delay-dependent stability conditions of Example 2

Methods	Maximumτ allowed
Moon et al. [3]	4.3588
Fridman and Shaked [1]	4.47
Xu and Lam [4]	4.4721
Corollary 1	4.47213

The compared results are listed in Table 27.1. It can be seen that the delay-dependent exponential stability criteria in this note is less conservative in the sense of the computed maximum delay bound.

Example 2 Consider the time-delay system (27.26) with

$$A = \begin{bmatrix} -2 & 0 \\ 0 & -0.9 \end{bmatrix}, \quad A_1 = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix}$$

For this deterministic system, we compared the results in Table 27.2. The results have demonstrated that the delay-dependent stability condition in this note is less conservative.

27.5 Conclusions

The exponential stability for stochastic delay systems with nonlinear uncertainties is investigated. Based on Ito calculus rules, a general type of Lyapunov–Krasovskii functional is introduced, and a novel stability condition is obtained in terms of linear matrix inequalities. Numerical examples are given to show that the criteria perform much better than the existing stability results.

Acknowledgments This work is supported by the Doctoral Foundation Project of Henan Polytechnic University under Grant B2010-50, the Natural Science Project of Henan Province Office of Education under Grant 2008B510008 and Henan Provincial Open Laboratory for Control Engineering Key Disciplines under Grant KG2009-07.

References

1. Fridman E, Shaked U (2002) A descriptor system approach to H control of linear time-delay systems. IEEE Trans Automat Control 47:253–270
2. He Y, Wang QG, Lin C, Wu M (2005) Augmented Lyapunov functional and delay-dependent stability criteria for neutral systems. Int J Robust Nonlin Control 15:923–933
3. Moon YS, Park P, Kwon WH, Lee YS (2001) Delay-dependent robust stabilization of uncertain state-delayed systems. Int J Control 74:1447–1455

4. Xu S, Lam J (2005) Improved delay-dependent stability criteria for time-delay systems. IEEE Trans Autom Control 50:384–387
5. Qian W, Cong S, Sun YX, Fei SM (2009) Novel robust stability criteria for uncertain systems with time-varying delay. Appl Math Comput 215:866–872
6. Mao X (1996) Robustness of exponential stability of stochastic differential delay equations. IEEE Trans Autom Control 41:442–447
7. Yue D, Won S (2001) Delay-dependent robust stability of stochastic systems with the time delay and nonlinear uncertainties. IEEE Electron Lett 37:992–993
8. Chen W, Guan Z, Lu X (2005) Delay-dependent exponential stability of uncertain stochastic systems with multiple delays: an LMI approach. Syst Control Lett 54:547–555
9. Yue D, Han Q (2005) Delay-dependent exponential stability of stochastic systems with time-varying delay, nonlinearity, and markovian switching. IEEE Trans Autom Control 50: 217–222
10. Gu K (2000) An integral inequality in the stability problem of time-delay systems. In: Proceedings of 39th IEEE conference on decision control, Sydney, Australia, pp 2805–2810

Chapter 28
Iterative Feedback Tunning for Boiler–Turbine Systems

Xuhui Bu, Fashan Yu and Fuzhong Wang

Abstract This paper presents an application of iterative feedback tuning (IFT) based on PID controller to the boiler–turbine system. We first transfer the nonlinear boiler–turbine system into a linear system, and then the PID controller can be designed to solve the regulation problem. The PID controller parameters are online tuned based on IFT method. Compare to the fixed parameter PID controller, the IFT-tuned PID controllers can obtain better performance. Simulation shows the effective of the proposed approach.

Keywords Iterative feedback tunning · Boiler–turbine system · PID controller

28.1 Introduction

Due to increasing fuel costs, improving the performance of a boiler–turbine is of great interest to the energy industry. A boiler–turbine system is an energy conversion device which consists of a steam boiler and turbine. The aim of a steam boiler is to transfer the input chemical energy of fuel into the thermal energy which is directly fed into a turbine. The boiler–turbine systems are the preferred electricity generation system, since they are able to meet electricity demands much faster than header systems. The requirements of a typical boiler–turbine control system are as follows:

1. Electric power output must meet the load demand.

X. Bu (✉) · F. Yu · F. Wang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454001, Henan, China
e-mail: buxuhui@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_28,

2. The drum pressure must be maintained within some tolerances despite the load's variations.
3. The water level in the steam drum of the boiler must be maintained at the desired level to prevent overheating or flooding.
4. The steam temperature must be maintained at the desired level to prevent overheating or leaking wet steam to turbines.
5. The constraints of input and output are imposed by physical limits such as the magnitude and saturation of the internal control valves actuator, magnitude and valves tuning rates must be taken into consideration.

Typically, the steam temperature and the drum water level are regulated tightly around a desired level over the full operating range. Therefore, the above third and fourth requirements can be treated as one. The boiler–turbine control system is a multi-input multi-output (MIMO) nonlinear system [1] and its state variables are strongly coupled. The boiler–turbine unit has been widely studied in the literature using various control techniques. In [2–4], an H_∞ controller is designed for a linearized boiler–turbine system. In [5], a genetic algorithm is used to find the optimal proportional-integral gains and state feedback gains on a linearized boiler–turbine system. In [6] an online self-organizing fuzzy logic controller to a boiler–turbine system of fossil power plant has been proposed. In [7], a gain scheduling control has been proposed for the nonlinear boiler–turbine dynamics which is represented as a linear parameter varying (LPV) form. In [8, 9], the predictive control has been introduced into the boiler–turbine system. The aforementioned approaches provide a design methodology to construct controllers for boiler–turbine systems which guarantee the closed-loop stability. However, these approaches only give the theoretical results, and they are difficult to be used in practice. The most widely used in practical boiler–turbine system is also PID controller [10]. However, the PID controller parameters are difficult to be tuned, especially for the MIMO boiler–turbine system.

The iterative feedback tuning (IFT) method is a model-free technique for the optimization of the parameters of a controller of fixed structure using only signal information on the closed-loop system [11–15]. The controller parameters are selected to optimize a quadratic cost criterion. The important contribution of the IFT methodology is that the gradient of the cost criterion with respect to the controller parameters can be obtained entirely from input–output data collected by performing specific experiments on the actual closed loop system. In [16], the IFT method is introduced to tune PID parameters in applications where the objective is to achieve a fast response to set point changes. Comparing the performance of these IFT-tuned PID controllers with the performance achieved by four classical PID tuning schemes that are widely used in the industry, the IFT always achieves a performance that is at least as good as that of the classical PID tuning schemes, and often dramatically better: faster settling time and less overshoot.

In this paper, we develop the IFT-tuned PID controllers for the boiler–turbine system. We first transfer the nonlinear boiler–turbine system into a linear system, and then the PID controller can be designed to solve the regulation problem, and

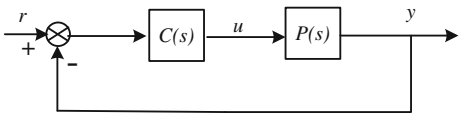


Fig. 28.1 Conventional feedback system

the PID controller is tuned by the IFT method. Compared to the fixed parameter PID controller, the IFT-tuned PID controllers can obtain better performance. This paper is organized as follows. Section 28.2 presents a brief overview of IFT method. In Sect. 28.3, the model of the boiler–turbine system is given. The design of IFT-tuned PID controller for the constrained boiler–turbine system is given in Sect. 28.4. Section 28.5 gives the simulation results and conclusions are given in Sect. 28.6.

28.2 Iterative Feedback Tunning

Detail of IFT algorithm is given in the literature [11–15]. In this section, only those equations that are necessary for the development of this paper are given. As shown in Fig. 28.1, $C(s)$ and $P(s)$ are the controller and plant, respectively. The signals r, u, y are the reference, control and plant output, respectively.

A user-specified desired output y_d and a quadratic criterion

$$J(\rho) = \frac{1}{2N}\sum_{t=1}^{N}[\tilde{y}_t(\rho)]^2, \tag{28.1}$$

are defined, where ρ, t, N are the controller parameter, discrete time index and the number of samples considered, respectively, and $\tilde{y}_t(\rho) = y_t(\rho) - y_t^d$. A solution for ρ to the equation

$$\frac{\partial J(\rho)}{\partial \rho} = 0, \tag{28.2}$$

where

$$\begin{aligned}\frac{\partial J}{\partial \rho}(\rho) &= \frac{1}{N}\sum_{t=1}^{N}\left(\tilde{y}_t(\rho)\frac{\partial \tilde{y}_t}{\partial \rho}(\rho)\right)\\ &= \frac{1}{N}\sum_{t=1}^{N}\left(\tilde{y}_t(\rho)\frac{\partial \tilde{y}}{\partial \rho}(\rho)\right),\end{aligned} \tag{28.3}$$

would give the minimum of J. If the gradient $\partial J/\partial \rho$ could be computed, then the solution of Eq. (28.2) could be obtained by the following iterative algorithm

$$\rho_{i+1} = \rho_i - \gamma_i R_i^{-1}\frac{\partial J}{\partial \rho}(\rho_i). \tag{28.4}$$

Here R_i is some appropriate positive definite matrix

$$\begin{aligned} R_i &= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial \tilde{y}_t(\rho_i)}{\partial \rho}\left[\frac{\partial \tilde{y}_t(\rho_i)}{\partial \rho}\right]^T\right) \\ &= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial y_t(\rho_i)}{\partial \rho}\left[\frac{\partial y_t(\rho_i)}{\partial \rho}\right]^T\right), \end{aligned} \tag{28.5}$$

typically, a Gauss–Newton approximation of the Hessian of J, while γ_i is a positive real scalar that determines the step size.

From

$$y(\rho) = \frac{C(\rho)P}{1 + C(\rho)P}r,$$

we have

$$\frac{\partial y(\rho)}{\partial \rho} = \frac{1}{C(\rho)}\frac{\partial C(\rho)}{\partial \rho}\left[\frac{C(\rho)P}{1 + C(\rho)P}(r - y(\rho))\right].$$

The term in the square bracket can be obtained by subtracting the plant output of one experiment on the closed-loop system from the reference, and using this signal as the reference signal in a new experiment. The term in the square bracket is the plant output of the new experiment. This leads to the experiments in IFT: for each iteration i of the controller parameter ρ_i, two experiments are conducted. For a chosen N-length signal, r_s, the first experiment consists of setting the reference $r = r_s$ and collecting the corresponding N samples of plant output denoted as y_1. The second experiment consists of setting the reference $r = r_s - y_1$ and collecting the corresponding N samples of plant output denoted as y_2. The derivative required in Eq. (28.3) is then computed as

$$\frac{\partial y(\rho_i)}{\partial \rho} = \frac{1}{C(\rho_i)}\frac{\partial C(\rho_i)}{\partial \rho}y_2(\rho_i), \tag{28.6}$$

and $\partial J(\rho_i)/\partial \rho$ in Eqs. (28.3) and (28.4) can be computed.

28.3 The Boiler–Turbine Unit Model

A 160 MW oil-fired electrical power plant model is used in this article constituted with a drum-type boiler and a turbine. The model is based on the P16/G16 at the Sydvenska Kraft AB plant in Malmo, Sweden. The boiler dynamic model is provided by both physical and empirical methods based on the data acquired from a series of experiments and identification which captures all the relevant

characteristics of the process. The boiler–turbine dynamic nonlinear model is given in the following equations [1]:

$$\begin{cases} \dot{x}_1 = -0.0018u_2x_1^{9/8} + 0.9u_1 - 0.15u_3, \\ \dot{x}_2 = (0.073u_2 - 0.016)x_1^{9/8} - 0.1x_2, \\ \dot{x}_3 = (141u_3 - (1.1u_2 - 0.19)x_1)/85, \\ y_1 = x_1, \\ y_2 = x_2, \\ y_3 = 0.05(0.13073x_3 + 100a_{cs} + q_e/9 - 67.975), \end{cases} \tag{28.7}$$

where state variables x_1, x_2, and x_3 denote drum pressure (kg/cm^2), electric output (MW), and fluid density (kg/m^3), respectively. The inputs, u_1, u_2, and u_3 are the valve positions for fuel flow, steam control, and feedwater flow, respectively. The output y_3 is the drum water level (m) and a_{cs} and q_e are steam quality and evaporation rate (kg/s), respectively and are given by

$$\begin{aligned} a_{cs} &= \frac{(1 - 0.001538x_3)(0.8x_1 - 25/6)}{x_3(1.0394 - 0.0012304x_1)}, \\ q_e &= (0.854u_2 - 0.147)x_1 + 45.59u_1 - 2.514u_3 - 2.096. \end{aligned} \tag{28.8}$$

Due to actuator limitations, the control inputs are subject to the following constraints:

$$\begin{aligned} 0 \le u_i \le 1 (i = 1, 2, 3), \\ -0.007 \le \dot{u}_1 \le 0.007, \\ -2 \le \dot{u}_2 \le 0.02, \\ -0.05 \le \dot{u}_3 \le 0.05. \end{aligned} \tag{28.9}$$

28.4 Design of IFT-Tuned PID Controller

This section focuses on the design of a PID controller which is tuned by iterative feedback tuning method. In the Sect. 28.3, the IFT is proposed for the linear system. Hence, we should transfer the nonlinear boiler– turbine system into a linear system. From the system (28.7), we can give the nominal operating point as follows:

$$\begin{aligned} x_0 &= [108\ 66.65\ 428]^T \\ u_0 &= [0.34\ 0.69\ 0.436]^T \\ y_0 &= [108\ 66.65\ 0]^T \end{aligned}$$

At these values, the following linearized model can be obtained

$$\begin{cases} \bar{x} = A\bar{x} + B\bar{u}, \\ \bar{y} = C\bar{x} + D\bar{u}, \end{cases} \tag{28.10}$$

where $\bar{x} = [x_1\ x_2\ x_3]^T, \bar{u} = [u_1\ u_2\ u_3]^T, \bar{y} = [y_1\ y_2\ y_3]^T$ and the system matrices are given by

$$A = \begin{bmatrix} -2.509e-3 & 0 & 0 \\ 6.94e-2 & -0.1 & 0 \\ -6.690e-3 & 0 & 0 \end{bmatrix}, \ B = \begin{bmatrix} -0.9 & -0.349 & -0.15 \\ 0 & 14.155 & 0 \\ 0 & -1.398 & 1.695 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 6.34e-3 & 0 & 4.71e-3 \end{bmatrix}, \ D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.253 & 0.512 & -0.014 \end{bmatrix}.$$

Now, we can design the PID controller for the linear system Eq. (28.10) with the following expression

$$\bar{u} = K_p e(k) + K_i \sum_{j=0}^{k} e(j) + K_d(e(k) - e(k-1))$$

where $e(k) = y^* - y(k)$, y^* is the desired output.

Let $\rho = [K_p, K_i, K_d]^T$, then the PID controller can be tuned using IFT method, and the tuning process can be given as follows:

Preparation: Select the controller parameters to stabilize the plant.

Test: Perform two experiment on the closed loop, the first experiment consists of setting the reference $r = r_s$ and collecting the corresponding N samples of plant output denoted as y_1. The second experiment consists of setting the reference $r = r_s - y_1$ and collecting the corresponding N samples of plant output denoted as y_2.

Parameter update: Calculate the matrix R and the gradient of the cost function $\partial J(\rho)/\partial\rho$, Revises the parameter ρ in accordance with the equation

$$\rho_{i+1} = \rho_i - \gamma_i R_i^{-1} \frac{\partial J}{\partial \rho}(\rho_i).$$

28.5 Simulation Result

The control systems for the boiler–turbine model are developed in MATLAB.

Simulations presented here illustrate the performances of the IFT-tuned PID controllers for various kinds of electric loads and drum pressure demands. In this section, the following desired output is considered

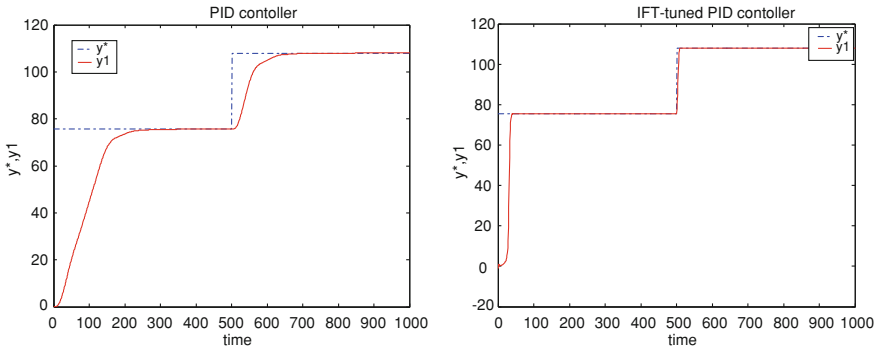


Fig. 28.2 The system output y_1 for different controllers

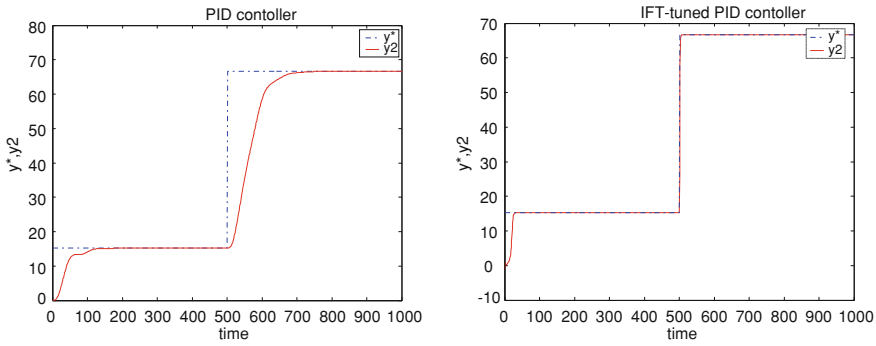


Fig. 28.3 The system output y_2 for different controllers

$$\begin{cases} y_1^* = 75.60, y_2^* = 15.27, y_3^* = 0, & 0 \le t \le 500 \\ y_1^* = 108, y_2^* = 66.65, y_3^* = 0, & 500 \le t \le 1000 \end{cases}$$

The PID controller is designed with $K_p = 0.15$, $K_i = 0.12, K_d = 0.04$, and the parameters are tuned by IFT method in the control process. Simulation results are shown in Figs. 28.2 and 28.3 The figure shows that outputs of the IFT-tuned PID control follow the reference signal faster than the PID control with fixed parameters. It is obvious that the proposed method can obtain better output performance.

28.6 Conclusions

In this paper, we introduce the iterative feedback tuning method to the boiler–turbine system. We first transfer the nonlinear boiler–turbine system into a linear system, and then the PID controller can be designed to solve the regulation

problem. The PID controller parameters are online tuned based on iterative feedback tuning. The method can optimize the parameters of PID controller using only signal information on the closed-loop system. Simulation results show that the IFT-tuned PID controllers can obtain better performance than the fixed parameter PID controller.

Acknowledgments This work was supported by Program for Science & Technology Innovation Talents of Henan Province (No. 104200510021).

References

1. Tan W, Marquez HJ, Chen T, Liu J (2005) Analysis and control of a nonlinear boiler–turbine unit. J Process Control 15(8):883–891
2. Kwon WH, Kim SW, Park PG (1989) On the multivariable robust control of a boiler turbine system. In: Proceedings of IFAC symposium on power systems and power plant control, pp 219–223
3. Tan W, Niu YG, Liu J (1999) H∞ control for a boiler–turbine unit. In: Proceedings of IEEE conference on control applications, pp 807–810
4. Wu J, Nguang SK, Shen J et al (2010) Robust H∞ tracking control of boiler–turbine systems. ISA Trans 49(3):369–375
5. Dimeo R, Lee KY (1995) Boiler–turbine control system design using a genetic algorithm. IEEE Trans Energy Convers 10(4):752–759
6. Moon UC, Lee KY (2003) A boiler–turbine system control using a fuzzy autoregressive moving average (FARMA) model. IEEE Trans Energy Convers 18(1):142–148
7. Chen PC, Shamma JS (2004) Gain-scheduled L1-optimal control for boiler–turbine dynamics with actuator saturation. J Process Control 14(3):263–277
8. Poncia G, Bittanti S (2001) Multivariable model predictive control of a thermal power plant with built-in classical regulation. Int J Control 74(1):1118–1130
9. Peng H, Ozaki T, Toyoda Y, Oda K (2001) Exponential ARX model-based long range predictive control strategy for power plants. Control Eng Pract 9(12):1353–1360
10. Tan W, Liu JZ, Fang F, Chen YQ (2004) Tuning of PID controllers for boiler–turbine units. ISA Trans 43(3):571–583
11. Hjalmarsson H, Gunnarsson S, Gevers M (1994) A convergent iterative restricted complexity control design scheme [C]. In: Proceedings of the 33rd IEEE conference on decision and control. IEEE, Orlando, U S A, pp 1735–1740
12. Hjalmarsson H (2002) Iterative feedback tuning-an overview [J]. Int J Adapt Control Signal Process 16(5):373–395
13. Hjalmarsson H (1998) Control of nonlinear systems using iterative feedback tuning [C]. In: Proceedings of the 1998 IEEE American control conference. IEEE, Philadelphia, pp 2083–2087
14. Hjalmarsson H (1999) Efficient tuning of linear multivariable controllers using iterative feedback tuning [J]. Int J Adapt Control Signal Process 13(7):553–572
15. Hjalmarsson H, Gevers M, Gunnarsson S, Lequin O (1998) Iterative feedback tuning: theory and applications [J]. IEEE Control Syst Mag 18(4):26–41
16. Olivier L, Michel G, Magnus M (2003) Iterative feedback tuning of PID parameters: comparison with classical tuning rules. Control Eng Practice 11(9):1023–1033

Chapter 29
New Exact Solutions for the (2 + 1)-Dimensional General Nizhnik-Novikov-Veselov System

Song-Hua Ma

Abstract The mapping method is a kind of classic, efficient and well-developed method to solve nonlinear evolution equations, the remarkable characteristic of which is that we can have many different ansatzs and therefore, a large number of solutions. In this paper, via the improved mapping method and a variable separation method, a series of variable separation solutions (including solitary wave solutions, periodic wave solutions and rational function solutions) to the (2 + 1)-dimensional general Nizhnik-Novikov-Veselov (GNNV) system is derived.

Keywords Improved mapping method · Variable separation method · General Nizhnik-Novikov-Veselov system · Variable separation solutions

29.1 Introduction

Seeking exact solutions to nonlinear evolution equation is one of the most important problems and it is always an interesting topic. The (1 + 1)-dimensional solitons and solitary wave solutions have been studied quite widely both in theoretical and experimental aspects. In (2 + 1)-dimensions, some significant integrable models have also been established in nonlinear physics. These (2 + 1)-dimensional nonlinear evolution equations possess richer localized solutions than the (1 + 1)-dimensional models do. In order to find the localized structures of (2 + 1)-dimensional nonlinear evolution equations, a wealth of effective methods have been set up, such as a bilinear method, symmetry reductions, standard and

S.-H. Ma (✉)
Department of Physics, Zhejiang Lishui University,
Lishui 323000, China
e-mail: msh6209@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_29,

extended Painlevé truncated expansions, the method of "coalescence of eigenvalue" or "wavenumbers", the homogenous balance method, the "variable separation approach" and the "mapping method" [1–12], etc. From these studies of (2 + 1)-dimensional models, one can see that there exist more abundant coherent structures than in lower dimensions. The fact hints us that the coherent structures of other (2 + 1)-dimensional models may have similar or new structures that have not yet been revealed. In this paper, we study the following (2 + 1)-dimensional general Nizhnik-Novikov-Veselov system.

$$\begin{aligned} &u_t + mu_{xxx} + nu_{yyy} - 3m(uv)_x - 3n(uQ)_y = 0, \\ &u_x = v_y, \quad u_y = Q_x \end{aligned} \tag{29.1}$$

where m and n are real parameters. This model is an isotropic Lax extension of the classical (1 + 1)-dimensional shallow water wave KdV system. Many authors have studied the solution of the GNNV system. For instance, the integrability of the above GNNV system was established and shown that the GNNV system was Hamiltonian. By certain transformation the sin-Gordon equation or sinh-Gordon equation can be derived from the GNNV system and Boiti-Leon-Pempinelli model. These equations arise in many branches of mathematical physics and have been widely applied in many realistic problems of atomic physics, molecular physics, particle physics, nuclear physics and shallow water theory [13]. Some fractal and chaotic localized solutions for this system have been discussed by Zheng and chen [14].

29.2 New Exact Solutions to the (2 + 1)-Dimensional GNNV System

As it is well known, to search for the solitary wave solutions to a nonlinear physical model, we can apply different approaches. One of the most efficient methods to find soliton excitations of a physical model is the so-called improved mapping approach. The basic ideal of the algorithm is as follows. For a given nonlinear partial differential equation (NPDE) with the independent variables $x = (x_0 = t, x_1, x_2,\ldots,x_m)$, and the dependent variable u, in the form

$$P(u, u_t, u_{x_i}, u_{x_i x_j}, \cdots) = 0 \tag{29.2}$$

where P is in general a polynomial function of its arguments, and the subscripts denote the partial derivatives, the solution can be assumed to be in the form,

$$u = A(x) + \sum_{i=1}^{n} B_i(x)\phi^i q(x) + C_i(x)\phi^{i-1} q(x)\sqrt{\sigma + \phi^2 q(x)} \tag{29.3}$$

with

$$\phi' = \sigma + \phi^2, \tag{29.4}$$

where σ is a constant and the prime denotes the differentiation with respect to q. To determine u explicitly, one may substitute (29.3) and (29.4) into the given NPDE and collect coefficients of polynomials of ϕ, then eliminate each coefficient to derive a set of partial differential equations of A, B_i, C_i and q, and solve the system of partial differential equations to obtain A, B_i, C_i and q. Finally, as (29.4) possesses the general solution, (the tanh-type and tan-type solutions are neglected here),

$$\begin{array}{lll} \text{(a)} & \phi = -\sqrt{-\sigma}\,\coth\left(\sqrt{-\sigma}q\right), & \text{when } \sigma<0, \\ \text{(b)} & \phi = -\sqrt{\sigma}\,\coth\left(\sqrt{\sigma}q\right), & \text{when } \sigma>0, \\ \text{(c)} & \phi = \frac{-1}{q}, & \text{when } \sigma=0. \end{array} \tag{29.5}$$

Substituting A, B_i, C_i, q and (29.5) into (29.3), one can obtain the exact solutions to the given NPDE. Now we apply the improved mapping approach to (29.1). By the balancing procedure, ansatz (29.3) becomes

$$\begin{aligned} u &= f + g\phi(q) + h\phi^2(q) + A\sqrt{\sigma+\phi^2(q)} + B\phi(q)\sqrt{\sigma+\phi^2(q)}, \\ v &= F + G\phi(q) + H\phi^2(q) + C\sqrt{\sigma+\phi^2(q)} + D\phi(q)\sqrt{\sigma+\phi^2(q)}, \\ Q &= a + b\phi(q) + c\phi^2(q) + d\sqrt{\sigma+\phi^2(q)} + e\phi(q)\sqrt{\sigma+\phi^2(q)}, \end{aligned} \tag{29.6}$$

where f, g, h, A, B, F, G, H, C, D, a, b, c, d, e and q are the functions of (x, y, t) to be determined. Substituting (29.6) and (29.4) for (29.1) and collecting the coefficients of the polynomials of ϕ, (the $\sqrt{\sigma+\phi^2(q)}$ is contained here. Letting $\mathrm{K} = \sqrt{\sigma+\phi^2(q)}$, then collecting the coefficients of K), then setting each coefficient to zero, yields

$$cq_x - hq_y = 0, \tag{29.7}$$

$$gq_y - c_x - bq_x + h_y = 0, \tag{29.8}$$

$$Bq_b - eq_x = 0, \tag{29.9}$$

$$2hq_y\sigma - b_x - 2cq_x\sigma + g_y = 0, \tag{29.10}$$

$$B_y - e_x - dq_x + Aq_y = 0, \tag{29.11}$$

$$gq_y\sigma - a_x - bq_x\sigma = 0, \tag{29.12}$$

$$2hq_x - Hq_y = 0, \tag{29.13}$$

$$h_x - H_y + gq_x - Gq_y = 0, \tag{29.14}$$

$$Bq_x - Dq_y = 0, \tag{29.15}$$

$$g_x - G_y + 2hq_x\sigma - 2Hq_y\sigma = 0, \tag{29.16}$$

$$B_x + Aq_x - Cq_y - D_y = 0, \tag{29.17}$$

$$f_x + gq_x\sigma - Gq_y\sigma = 0, \tag{29.18}$$

$$2mBq_y^3 + 2nBq_y^3 - nBq_y - mhq_xD - mBq_xH - nhq_ye = 0, \tag{29.19}$$

$$\begin{aligned} &18mBq_{xx} - 3mh_xD - 3mB_xH - 3nBc_y + 18nBq_{yy}q_y - 9mAq_xH \\ &- 3mhD_x + 18B_yq_y^2 - 9mhq_xC - 9nAq_yc - 9mgq_xD = 0, \end{aligned} \tag{29.20}$$

$$\begin{aligned} &2mhq_x^3 + nhq_y^3 - nhq_yc - mhq_xH - mhq_xH - mBq_xH \\ &- mBq_xD - nBq_ye = 0, \end{aligned} \tag{29.21}$$

$$\begin{aligned} &2mBq_{xxx} - 3mA_xH - 3mg_xD - 3mhC_x \\ &- 3nh_xC - 3mh_xC - 3nBb_y - 3nhd_y - 3B_yb \\ &- 33nhq\,e\sigma - 6mAq_xG - 6nfeq_y - 6ngq_yd \\ &- 6mfDq_x - 6mBq_xF + 76nBq_y^3\sigma \\ &+ 76mBq_x^3\sigma - 6nAq_yb - 33mBq_xHc - 3mBG_x \\ &- 3mB_xG - 6nBq_ya + 2Bq_t + 6mBq_xF + 76nBq_y^3\sigma \\ &+ 6nAq_y^2 - 3mgD_x - 3nAc_y - 3ng_ye \\ &- 3mAH_x - 6nBq_ya + 2Bq_t + 6mB_{xx}q_x + 2nBq_{yyy} \\ &+ 6nByyq_x + 6mB_xq_{xx} + 6Aq_{xx}q_x + 6naq_y + 6nB_yq_{yy} = 0, \end{aligned} \tag{29.22}$$

$$\begin{aligned} &18mh_xq_x^2 - 9nhq_yb + 18nh_yq_y^2 - 9ngg_yc \\ &- 9nBq_yd - 9mAq_xD - 3mB_xD + 6mgq_x^3 + 2nBq_{yyy} \\ &- 9mBq_xC - 9mgq_xH + 18nhq_ygq_{yy} - 3nh_yc - 3nB_ye = 0, \end{aligned} \tag{29.23}$$

$$\begin{aligned} &3nAq_{xxx} - 3mA_xG - 3mB_xF - 3nB_ya - 3nBa_y - 3mfD_x - 3nfdq_y \\ &- 6nhe_y\sigma - 3ng_yq_y + 3nA_yq_{yy} + 51mBq_{xx}q_x\sigma + 51nBq_{yy} \\ &+ mAq_{xxx} + 3mA_{xx}q_x + nAq_{yyy} + nB_{yyy} + 3mA_xq_{xx} = 0, \end{aligned} \tag{29.24}$$

$$\begin{aligned} &- 3mgH_x - 3B_xC - 24nhq_yc\sigma + 6ng_yq_y^2 - 3mh_yb - 30nBq_ye\,\sigma \\ &- 24mhq_xH\sigma - 30mBq_xD\sigma - 3mAD\sigma - 6mfHq_x - 3ngc_y \\ &- 3nAe_y + 2nhq_{yyy} + 6mh_{xx}q_x \\ &- 3nB_yd + 2hqt - 6nfcq_y + 6mgq_xq_{xx} + 6ngq_yq_{yy} = 0, \end{aligned} \tag{29.25}$$

$$
\begin{aligned}
&15nAq_{yy}q_y\sigma + 15B_{yy}q_y\sigma + 5mBq_{yyy}\sigma + 15mAq_{xx}q_x\sigma \\
&\quad - 6ng_ye\sigma - 15nBq_ya\sigma - 30nBq_yc\sigma^2 - 15ngq_yd\sigma - 6ng_ye\sigma \\
&\quad - 15mBq_xF\sigma - 30mhq_xD\sigma^2 - 15nAq_yb\sigma - 15mAq_xG\sigma \\
&\quad - 15nfq_y\sigma - 30mBq_xH\sigma - 6nA_yc\sigma - 6mgD_x\sigma - 6mh_xC\sigma \\
&\quad - 6nAc_y\sigma + 85nBq_y^3\sigma^2 + 15mA_xq_x^2 - 3mAF_x + A_t = 0,
\end{aligned} \tag{29.26}
$$

Based on the above equations, we have

$$
\begin{aligned}
&f = \int \left(q_{xx}q_y - q_{xy}q_x\right)\sigma dx, \quad g = q_{xy},\ h = q_xq_y,\ A = -q_{xy},\ B = -q_xq_y, \\
&F = \Bigg[nq_{xyy}q_y^2 - \left(nq_y^2q_x\sigma + mq_y^3\sigma + mq_x^3\sigma\right)\int \left(q_{xx}q_y - q_{xy}q_x\right)dx \\
&\quad + \frac{1}{3}mq_xq_yq_{xxx} - mq_{xy}q_{xx}q_x + mq_{xxx}q_x^2 - nq_{xy}q_{yy}q_y + \frac{1}{3}nq_xq_yq_{yy} \\
&\quad + \frac{5}{3}mq_yq_x^4\sigma + \frac{5}{3}nq_y^4q_x\sigma + \frac{1}{3}q_yq_xq_t\Bigg] \Big/ q_x^2q_ym, \\
&G = q_{xx},\ H = q_x^2,\ C = -q_{xx},\ D = -q_x^2,\ a = \int \left(-q_{yy}q_x + q_{xy}q_y\right)\sigma dx, \\
&b = q_{yy}, c = q_y^2, d = -q_{yy}, e = -q_y^2
\end{aligned} \tag{29.27}
$$

Substituting (29.27) for (29.7)–(29.26) yields the following equations:

$$
\begin{aligned}
&q_x^2q_ymq_{xx}q_{xxx} - 15q_x^3q_y^3nq_y^3nq_{yy}\sigma - 3q_xq_y^3nq_{yy}\sigma + 17q_x^2q_y^4nq_{xy}\sigma \\
&\quad - 3q_x^2q_y^2nq_{xx}\sigma \int \left(q_{xy}q_y - q_{yy}q_x\right)dx + 6q_xq_y^2nq_{xy}q_{xyy} + 3q_{xx}q_y^2nq_{xy}nq_{yy} \\
&\quad + 3n\sigma q_x^3q_yq_{yy} \int \left(q_{xy}q_y - q_{xy}q_x\right)dx + 3nq_x^3q_y^2\sigma \int \left(q_{xyy}q_y - q_{yyy}q_x\right)dx \\
&\quad + 3q_xq_y^3nq_{xxyy} + q_x^2q_yq_{xy}q_t - 3q_x^2q_{xy}^2mq_{xx} - 3q_x^2q_y^2nq_{xyy} = 0,
\end{aligned} \tag{29.28}
$$

$$
\begin{aligned}
&q_x^2q_ymq_{xy}q_{xxx} - 15q_x^3q_y^3nq_{yy}\sigma - 3q_xq_x^5nq_{xy} - 3q_x^2q_y^2n \int \left(q_{xy}q_y - q_{yy}q_x\right)dx \\
&- 3q_x^4q_ym62mq_{xxy}\sigma + 6q_xq_y^2nq_{xy}q_{xyy} - 12q_xq_y^3nq_{xy}\sigma \int \left(q_{xx}q_y - q_{xy}q_x\right)dx \\
&\quad + 17q_x^2q_y^2nq_{xy}\sigma + 3q_{xx}q_y^2nq_{xy}q_{yy} + 4q_x^2q_yq_{yyy} - 6q_xq_ynq_{xy}^2q_{yy} \\
&\quad + 3q_x^3q_{xy}mq_{xxy} + q_x^2q_yq_{xy}q_t - nq_xq_x^3q_{yyy} - 3q_{yt}q_x^3q_y - 3q_xq_{xy}^3mq_{xx} \\
&\quad - 3q_x^2q_y^2nq_{xyyy} - q_x^3q_ymq_{xxxy} - 9q_xq_y^2nq_{xxxy}q_{yy} = 0,
\end{aligned} \tag{29.29}
$$

$$
\begin{aligned}
&3q_x^4 q_{xy} m\sigma \int (q_{xy}q_y - q_{xy}q_x)dx + 3nq_x^3 q_y^2 \sigma \int (q_{xyy}q_y - g_{yyy}q_x)dx \\
&+ 3q_x^3 q_{xy} m q_{xxy} + q_x^2 q_y q_{xy} q_t - nq_x^3 q_y q_{yyy} - 3q_{xx} q_y^2 n q_{xyy} + 3q_x q_y^3 n q_{xxyy} \\
&+ 3q_x q_y^3 n q_{xy}^3 m - 3q_x^2 q_y^2 n q_{xyyy} - q_x^3 q_y m q_{xxxy} - 6q_x q_y n q_{xy}^2 q_{yy} \\
&- 3q_x q_y^5 n q_{yy}\sigma - 3q_x q_y^2 n q_y^2 n q_{xxy} q_{yy} + q_x^2 q_y m q_{xy} q_{xxx} - q_x^3 q_{yt} q_t = 0,
\end{aligned} \tag{29.30}
$$

Obviously, it is very difficult to obtain the general solutions to the above equations. Fortunately, in this special case, one of the special solutions can be expressed with the function q in the following variable separation form

$$
q = \chi(x) + \tau(t) + \varphi(y), \text{ and } \varphi(y) = ky, \tag{29.31}
$$

where k is an arbitrary constant. Based on the solutions of (29.4), one thus obtains an explicit solution of (29.1).

Case 1. For $\sigma = -1$, we can derive the following solitary wave solutions of (29.1):

$$
u_1 = k\chi_x[\operatorname{csc} h(\chi + \tau + ky)^2 + \cot h\ (\chi + \tau + ky) \operatorname{csc} h(\chi + \tau + ky)], \tag{29.32}
$$

$$
\begin{aligned}
v_1 = &\frac{1}{3}\frac{m\chi_{xxx} - 2nk^3 + \chi_t + \tau_t}{m\chi_x} - \chi_{xx}[\cot h(\chi + \tau + ky) + \operatorname{csc} h(\chi + \tau + ky)] \\
&+ \chi_x^2 \cot h\ (\chi + \tau + ky) - \frac{2}{3} + \cot h(\varphi + \tau + ky) \operatorname{csc} h(\chi + \tau + ky),
\end{aligned} \tag{29.33}
$$

$$
Q_1 = k^2[\cot h(\chi + \tau + ky)^2 + \cot h(\chi + \tau + ky) \operatorname{csc}\ h(\varphi + \tau + ky)] \tag{29.34}
$$

Case 2. For $\sigma = 1$, we obtain the following periodic wave solutions of (29.1):

$$
u_2 = k\chi_x[\operatorname{csc}(\chi + \tau + ky)^2 + \cot(\chi + \tau + ky) \operatorname{csc}(\chi + \tau + ky)], \tag{29.35}
$$

$$
\begin{aligned}
v_2 = &\frac{1}{3}\frac{m\chi_{xxx} - 2nk^3 + \chi_t + \tau_t}{m\chi_x} - \chi_{xx}[\cot(\chi + \tau + ky) + \operatorname{csc}(\chi + \tau + ky) \\
&+ \chi_x^2 \cot(\chi + \tau + ky) + \frac{2}{3} + \cot(\chi + \tau + ky) \operatorname{csc}(\chi + \tau + ky)],
\end{aligned} \tag{29.36}
$$

$$Q_2 = k^2[\cot(\chi+\tau+ky)^2 + \cot(\chi+\tau+ky)\csc(\chi+\tau+ky)] \tag{29.37}$$

Case 3. For $\sigma = 0$, we find the following variable separated solution of (29.1):

$$u_3 = 2\frac{k\chi_x}{(\chi+\tau+ky)^2} \tag{29.38}$$

$$v_3 = \frac{1}{3}\frac{m\chi_{xxx}+\chi_t+\tau_t}{m\chi_x} - 2\frac{\chi_{xx}}{\chi+\tau+ky} + 2\frac{\chi_x^2}{(\chi+\tau+ky)^2} \tag{29.39}$$

$$Q_3 = 2\frac{k^2}{(\chi+\tau+ky)^2} \tag{29.40}$$

29.3 Summary and Discussion

In summary, via the improved mapping approach and a variable separation approach, we have found a new family of variable separation solutions (including solitary wave solutions, periodic wave solutions and rational function solutions) to the (2 + 1)-dimensional general Nizhnik-Novikov-Veselov system. The improved mapping method in nonlinear physics is a very powerful method. With the method, we have already solved many equations of (2 + 1)-dimensional system, such as (2 + 1)-dimensional Boiti-Leon-Pempinelli system, (2 + 1)-dimensional Broer-Kaup-Kupershmidt system, (2 + 1)-dimensional Generalized Broer-Kaup system and the like. We expect that this method may also be applied to other (2 + 1)-dimensional models and/or even (3 + 1)-dimensional models.

Acknowledgments The authors would like to thank professor J.F. Zhang for fruitful and helpful suggestions. This work was supported by the Natural Science Foundation of Zhejiang Province (Grant Nos. Y6100257, Y6110140).

References

1. Zheng CL, Zhang JF (2002) General solution and fractal structures for the (2 + 1)-dimensional generalized Ablowitznm-Kaup-Newell-Segur system. Chin Phys Lett 19(10):1399–1402
2. Zheng CL, Zhang JF (2003) Exit solutions and abundant coherent soliton structures of (2 + 1)-dimensions Ablowitznm-Kaup-Newell-Segur system. Commun Theor Phys 39(1):9–14
3. Zheng CL, Chen LQ (2004) A general mapping approach and new travelling wave solutions to (2 + 1)-dimensional Boussinesq Equation. Commun Theor Phys 41(5):671–674
4. Ma ZY, Zheng CL (2006) Two classes of fractal structures for the (2 + 1)-dimensional dispersive long wave system. Chin Phys B 15(1):45–52

5. Zhu HP, Zheng CL, Fang JP (2006) Fractal and chaotic patterns of Nozhnik-Novikov-Veselov system derived from a periodic wave solution. Phys Lett A 355(12):39–46
6. Zheng CL (2002) Coherent soliton structures with chaotic and fractal behaviours in a generalized (2 + 1)-dimensional korteweg de-vries system. Chin J Phys 41(5):442–453
7. Ma SH, Fang JP, Lv ZJ (2009) Folded solitary wave excitations for (2 + 1)-dimensional Nozhnik-Novikov-Veselov system. Commun Theor Phys 51(3):479–484
8. Fang JP, Zheng CL, Liu Q (2005) Non-propagating soliton in (2 + 1)-dimensional dispersive long-water system. Commun Theor Phys 43(2):245–250
9. Fang JP, Zheng CL, Zhu JM, Ren QB (2005) New family of exit solutions and chaotic solitons of generalized Breor-Kaup system in (2 + 1)-dimensions via an extended mapping approach. Commun Theor Phys 44(2):203–208
10. Ma SH, Fang JP, Zheng CL (2008) Complex wave excitations and chaotic patterns for a generalized (2 + 1)-dimensional korteweg de-vries system. Chin Phys B 17(8):2767–2773
11. Ma SH, Wu XH, Fang JP, Zheng CL (2006) Chaotic solitons for the (2 + 1)-dimensional modified dispersive water-wave system. Z Naturforsch 61a:249–252
12. Ma SH, Wu XH, Fang JP, Zheng CL (2009) New exact solutions for the (3 + 1)-dimensional Jimbo-Miwa system. Chaos Solitons Fractals 40:1352–1355
13. Zhang JF, Zheng CL (2003) Abundant localized coherent structures of the (2 + 1)-dimensions generalized Nozhnik-Novikov-Veselov system. Chin J Phys 41(3):242–254
14. Zheng CL, Chen LQ (2005) New localized excitations in a (2 + 1)-dimensional generalized Nozhnik-Novikov-Veselov system. Chin J Phys 43(3):393–399

Chapter 30
A Role-Based Feature Model Componentization Method

Jun Zhang

Abstract Feature model is a requirement model in software product line which represents a family of products in a uniform form. To solve feature model's chaos and entanglement in model description and configuration, we design a feature model componentization method. The method introduces the concept of role and reference role, and implements an algorithm called feature-role-component which maps product line features generated by requirement elicitation and analysis in domain engineering phase to different model components. As the intermediary, the role and reference role decouples the feature and component, which makes convenient selection and composition between feature variants and enhances the componentization level of the system. The method decomposes the system functions in the early phase of software lifecycle, improves the flexibility and adaptability, which makes a stable foundation for higher quality product.

Keywords Role · Reference role · Model componentization · Component

30.1 Introduction

One important problem of software engineering is the non standardization in the development. The whole process depends excessively on artistic creativity of developer himself other than a series of executable standards, which makes it unpredictable and uncontrollable. The lack of experience and the fault of decision inevitably influence the quality, flexibility and extensibility of software, increase the difficulties in product maintenance and evolution. To make the software process

J. Zhang (✉)
College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zhangjun@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_30,

controllable and automation and improve software quality, a series of technologies and theories such as OO, design patterns, component, software architecture emerge as the times require. As a paradigm of developing product family, software product line (SPL) attracts a lot of attention due to its high efficiency and low cost.

SPL is a set of software intensive systems sharing a common, managed set of features that satisfy the specific needs of a particular market segment or mission that are developed from a common set of core assets in a prescribed way [1]. The core idea is mapping a feature model satisfying all function requirements and constraints for entire domain products according to domain requirement specification, which is called domain modeling. And then through selecting and combining the feature and variants according to the product requirement, which is called feature customization, the concrete product is generated automatically; this process is called requirement engineering. The feature model provides a solution to the development of high customized and reusable software.

At present, there were some problems in domain feature modeling as follows: first, the inherent coupling between domain modeling and requirement engineering make the feature model variants entangle and interleave together and all variants must be found out in advance; second, present method scarcely support system evolution caused be no-existing feature adding. To solve all these problems, this chapter puts forward a method mapping feature to component. Through introducing the concept of role and reference role, the method analyzes and elicits knowledge from domain, and breaks the feature into simple and reusable roles. Making use of the intermediary role, we design an algorithm called feature-role-component, the algorithm realizes the transformation from feature model to component model, decouples feature model and requirement model, solves the entanglement problem and make the feature selected and combined freely, which solve software evolution problem and to some extent The method decomposes the system functions in the early phase of software lifecycle, improves the flexibility and adaptability, which makes a stable foundation for higher quality product.

The remainder of this chapter is organized as follows. In Sect. 30.2, we give a general description related to our work and method, analyzes the advantages and shortcomings of current researches. Section 30.3 depicts the feature model and Sect. 30.4 introduces the concepts of role and reference role. In Sect. 30.5, we give a detailed description of our method and the implementation of the feature-role-component mapping algorithm. This chapter closes with some conclusions and a future outlook in Sect. 30.6.

30.2 Related Works

As a potential paradigm for industrialized software family production, SPL attracts a lot of attention due to its high efficiency and low cost.

Kang etc. first put forward the concept of feature modeling and related FODA method in the literature [2], and present FORM method in literature [3] which

extends FODA. The FORM divides the SPL to the phase of domain engineering and application engineering. Czarnecki represents SPL as feature model and model template in the literature [4]. Model template is the set of all available template instances and the system model is generated by template tailoring with the evaluation of existing conditions and meta expressions.This method expresses the commonality and variability intuitively and traces the feature and its implementation well. However, for large scale application, the template is of huge scale, and it requires the designers and maintainers to have a comprehensive and painstaking knowledge of the domain and modeling all the requirement variants by one time, which brings extremely huge amount of workload and error inclination. Guelfi and Perrouin [5] present a method of requirement analysis based on model transformation. The method uses FIDJI analysis model and UML, OCL and textual use case languages to realize the mapping from requirement to model. The method is focus on requirement elicitation other than combination so there is little support for model function combination. In Chinese literature [6–8], they also presents some method for feature componentization, but all their goal is to solve how to design a component other than the feature model mapping itself, their concerns is which features the component implements but not which components the feature selects and combines, so all these method can solve the problem of free feature selection and combination.

30.3 Feature Model

Feature is the cohesive description of single requirement, to users, it is the visible properties of software system. Feature model can be considered as a series of features elicited and abstracted from the domain and their associations and constraints. The associations of feature describe the relationships of features which can be subdivided to specialization, decomposition and so on. Constraints reflect the static binding relationship between features, which support the verification of requirement customization in application engineering phase. Basic constraints include requires, which means if feature A is selected then feature B must be selected, and excludes, which means if feature A is selected then feature B can not be selected.

As shown in Fig. 30.1, the feature model of Air Traffic Control domain includes a lot of features such as flight-plan-design, radar-management, flight-plan-management, area-management and human–computer interface and so on. And the feature flight-plan-management can be further divided to sub features of trace-generation, conflict-management and conflict-detect. So the creation of feature model is the process of decomposition with features in different abstract levels.

In Fig. 30.1, we can find specialization association in human–computer interface feature, which means the feature can be implements with different technologies. However in our method, we pay little attention to this association, it can be implemented back to the phase of transformation from PIM to PSM [9].

ATC feature model has a requires constraint between the feature conflict-detect and conflict-management, that is requires (conflict-detect,conflict-management).

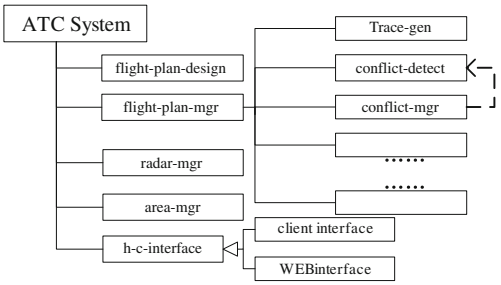


Fig. 30.1 Feature model of ATC domain

In our method, constraints like this have no influence to the method because of the realization of free feature selection and combination.

30.4 Basic Concepts

To show our method in detail, we introduce the concepts of role and reference role [10].

Definition 1 Role is a quaternion R = (ID, Name, Des, OpSet),where ID is the identifier of R, Name is a general description of R, Des is a description of R's function, in either formal or informal language, OpSet is the set of R's tasks, in which stores R's operations interacting with environment or other roles.

According to Definition 1, role can be used to define the responsibility or protocol of feature; it describes the specific task of feature. OpSet obtained by feature operationalization is the interacting interface of component model, so role is a media for feature-component mapping.

After introducing the concept of role, the meta-model of feature operationalization depicts in Fig. 30.2.

From the figure, we can see that a role does not belong to a single feature but to all features in domain product family; it can be used or referred by all features. However, for a certain role, it plays different part in different role analysis model. So to take into account both reusage and operationalization, we introduce the concept of Reference Role.

Definition 2 a Reference Role is a pair RR = (R, type), where R represents the role RR refers to, and type represents the function RR plays in the role analysis model. The value of type is one of followings:

Required: indicates the function of R is not implemented by the feature contains it but others,

Provided: the function of R is implemented by the feature refer to RR and the role is one part of communication interface to other features,

Internal: the function of R is implemented by the feature refer to RR, but R is used in the feature itself other than provide service to other features.

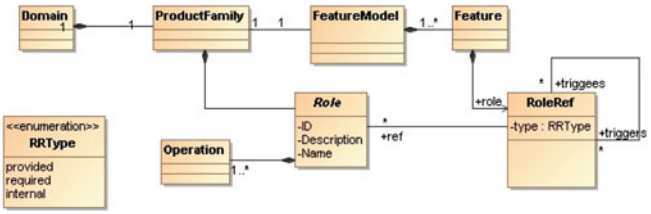


Fig. 30.2 Meta-model for feature operationalization

From the definition, we can see that reference role is just a local copy of role in role analysis model which adds a type attribute to the role.

Figures 30.3, 30.4 illustrates the feature-component mapping change after introducing the concepts of role and reference role. Without the role and reference role, the features and components hold a high-coupled many-to-many relation, while with the media of role and reference role, the many-to-many relation is decomposed to two one-to-many relation, that is, a feature can be represent by the interaction between many roles and the component can be considered as the implementation of many roles.

30.5 Componentization Method

The method present includes three main phases:

Feature operationalization.

Feature operationalization is the process that breaks features into compute-oriented roles according to specifications. The process can be accomplished by extracting functional nouns. For example, we can extract two nouns flight-plan and flight-trace from trace-generation, so there are two corresponding roles.

Role analysis model creation.

This process of operationalization makes the feature represented by different roles. In order to map the feature to component, we also need to know the relationships of these roles, which is called role analysis process. The process has two tasks:

Domain expert and modeler should analyze the function of roles and assign a type according to it.

Analyze the dependency of interaction among roles, create role analysis model.

Role analysis process is accomplished by extracting the verbs and adverbs from the specifications. In the process, we first confirm the function provided by the feature. E.g. the function of trace-generation feature is to provide flight-trace, so the type of flight-trace algorithm should be provided; and then we should find verb or adverb around the noun represent the role, e.g., we "find generate… according to according to preinstalled flight-plan, and the noun(role) follows the adverb appears to the required/internal role. The difference between the type required and internal is whether the feature implements the function, if it implements the

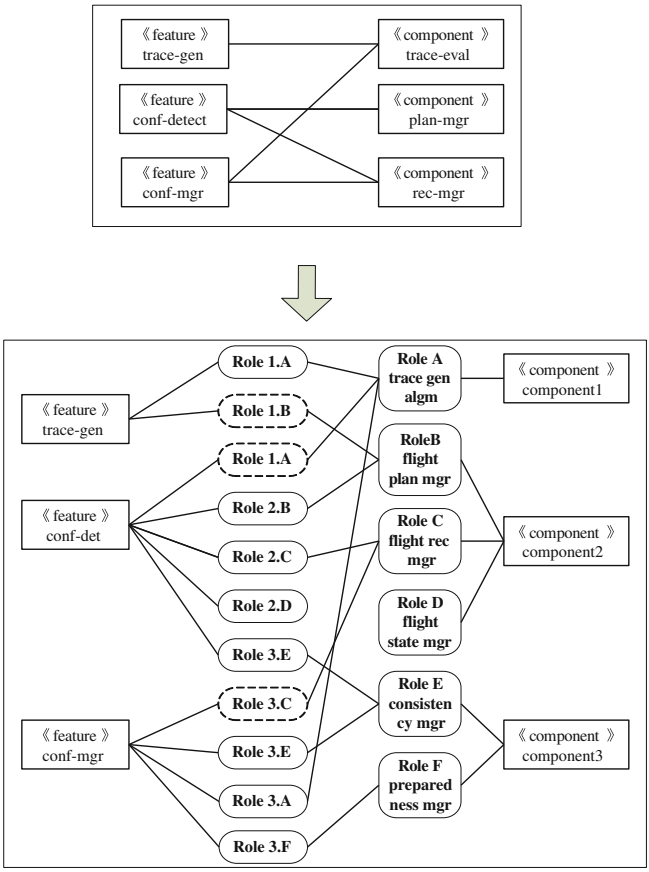


Fig. 30.3 The mapping among features and components

function, then the type is internal, else the type is required. For some complex specification, if the modeler can determine whether the feature implements the function of role, he can assign the type with the value internal, and the algorithm introduced in following section will amend the type according to the result of role analysis model".

Feature-role-component mapping.

When role analysis model creation is finished, feature-role-component mapping algorithm will be carried out to generate component automatically. The main idea of the algorithm is to compute the reusage ratio of roles and classify them to different atom components according the ratio, and then assemble atom components to composite component according the feature specification. The flow diagram of the algorithm shows in Fig. 30.5.

(1) The first step of algorithm is to compute certain feature sets of all roles and amend the type of internal role when necessary.

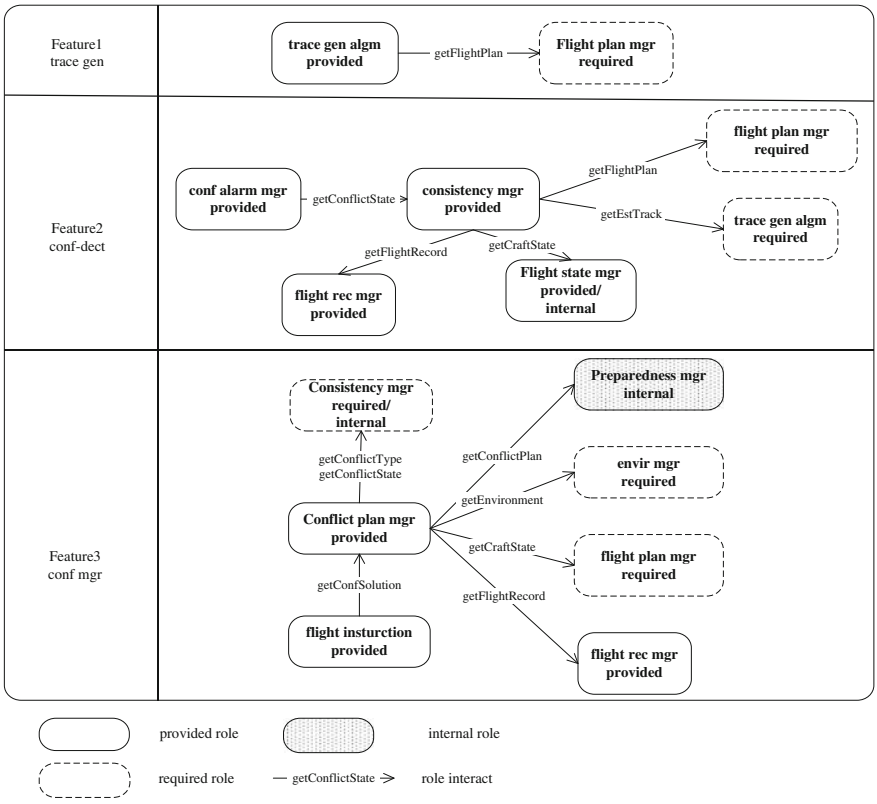


Fig. 30.4 Role analysis model of ACT domain

The definition of the feature sets give as follows:

Definition 3 Feature Set of Role r RFS(r) is a set of features, where every feature in the set contains reference of role r, that is

$$\mathrm{RFS}(r) = \bigcup_{i=1}^{|r.\mathrm{ref}|} \{f | f = r.\mathrm{ref}_i.ftr\} \tag{30.1}$$

In formula (30.1), |r.ref| represents dimension of role r's reference, $\mathrm{r.ref}_i$ is *i*th reference of r, $\mathrm{r.ref}_i$.ftr is the feature contains $\mathrm{r.ref}_i$. According to r's type, we divide RFS(r) to three subsets, whose definitions are listed below:

Definition 4 Positive Feature Set of Role *r* PoRFS(*r*) is a subset of RFS(r) where reference type of *r* is provided, that is,

$$\mathrm{PoRFS}(r) = \bigcup_{i=1}^{|r.\mathrm{ref}|} \{f | f = r.\mathrm{ref}_i.ftr \wedge r.ref_i.\mathrm{type} = \} \tag{30.2}$$

Fig. 30.5 Feature-role-component mapping process

Definition 5 Passive Feature Set of Role r PaPFS(r) is also a subset of RFS(r) but the reference type of r is required, that is,

$$\text{PaRFS}(r) = \bigcup_{i=1}^{|r.\text{ref}|} \{f | f = r.\text{ref}_i.ftr \wedge r.ref_i.\text{type} = \text{required}\} \quad (30.3)$$

Definition 6 Similar to PoRFS(r) and PaRFS(r), Internal Feature Set of Role r InPFS(r) is a subset of RFS(r) and the reference type of r is internal, that is

$$\text{InRFS}(r) = \bigcup_{i=1}^{|r.\text{ref}|} \{f | f = r.\text{ref}_i.ftr \wedge r.ref_i.type = \text{internal}\} \quad (30.4)$$

As depicts in role analysis phase, when the type of a role can been specified determinately, the value will be assigned to internal temporarily. But before we carry out the algorithm, we should confirm or amend the value. The main idea of the algorithm is that if the role's function is provided by other features, the type should be assigned to required, if there are no other feature implement this role but some feature require this function, the role type should be reassigned to provided, The algorithm for amending lists in detail below:

```
Role type amending Algorithm:
a.  Initialize role set RS
b.  IF RS = Ø  THEN return; ENDIF
c.  For first role r in RS, compute DS(r) = RFS(r)-PoRFS(r)-PaRFS(r).
d.  IF DS(r)= Ø THEN //there is no internal role.
      IF RFS(r) = PaRFS(r) THEN warning("role r without implementation") ENDIF.
e.  ELSE IF PoRFS(r) ≠ Ø THEN
      FOR i := 0 to |DS(r)|-1 do.
        DS(r)[i].rtype := required ; PaRFS(r):= PaRFS(r) ∪ {DS(r)[i]} ;
        InRFS(r):=InRFS(r) - {DS(r)[i]}.
      ENDFOR.
f.  ELSE IF PaRFS(r) ≠ Ø THEN
      FOR i := 0 to |DS(r)|-1 do.
          DS(r)[i].rtype := provided; PoRFS(r):= PoRFS(r) ∪ {DS(r)[i]};
          InRFS(r):=InRFS(r) - {DS(r)[i]}.
      ENDFOR.
    ENDIF.
g.  RS := RS-{r}, goto b.
```

(2) This step is to create model component according to feature sets of roles

Definition 7 Model Component is a unit of System Models; it consists of interface and implementation. Formally, we define it as an octo-tuple MComp = {MCID, PIntf, RIntf, PRole, RRole, IRole, Impl, Init}, where CID is

the identifier of MComp, PIntf is the provided interface set, RIntfis the required interface set, PRole represents provided role set, RRole represents required role set, IRole represents internal role set, Impl is the implantation of MComp and Init gives the initial state of MComp.

The creation of model component should maximize the reusage, that is, the role appears in different features should be extracted out to be a single component for invocation. Based on this basic idea, the algorithm createMComponentSet is designed. According to the type of role, all role defined in domain can be classified into two categories, one is provided, these kind of roles provided services to others, so they can be used as interact interface by the feature, and the component is created based on the value of PoRFS(r); the other is internal roles, these roles are invoked by the roles within the feature refer to it, so they are assigned to IRole set of component, which is the most cases. But in few cases, the internal role may be invoked by roles referred by different features, and we should create a new component for the internal role itself. The details of algorithm depicts below:

```
createMComponentSet;.
a.   Obtain the subset PRS of RS,satisfy ∀r ∈ PRS, PoPRS(r) ≠ ∅ ;.
     Obtain the subset IRS of RS, satisfy ∀r ∈ IRS, InPRS(r) ≠ ∅ ;.
b.   IF PRS = Ø THEN goto g; ENDIF.
c.   r := PRS[0], CRS := {r}.   PRS := PRS –{r}.
d.   FOR i := 0 to |PRS|-1 DO   .
       IF PoRFS(PRS[i]) = PoRFS(r) THENCRS := CRS   U {PRS[i]};ENDIF.
     ENDFOR.
e.   MC := createNewMComponent(); MC.PRole := CRS;  .
     FOR i := 0 to |MC.PRole|-1 DO    .
        pr := MC.PRole[i];.
        FOR I := 0 to |PoRFS(pr)|-1 DO.
          RRS := RRS ∪PoRFS(pr)[i].rolepr.triggee;.
        ENDFOR.
     ENDFOR.
     RRS := RRS – MC.PRole; MC.RRole := RRS;.
f.   MCS := MCS ∪ MC;PRS := PRS – CRS; goto b;.
g.   WHILE IRS ≠ Ø DO.
        ir := IRS[0]; RTrS:= Ø; RTeS := Ø;  .
        FOR i:= 0 to |InRFS(ir)|-1 DO  .
          RTrS := RTrS∪ InRFS(ir)[i].roleir.trigger.
          RTeS := RTeS ∪InRFS(ir)[i].roleir.triggee.
        ENDFOR.
          IF isInSameMComponent(RTrS) THEN   .
            mc := getMComponent(RTrS); mc.IRole := mc.IRole   ∪ {ir}.
          mc.RRole := (mc.RRole–mc.Irole)∪RTeS.
        ELSE          .
          MC := createNewMComponent(); MC.PRole := {ir};.
          MC.RRole := RTeS; MCS := MCS ∪ MC;  .
     ENDIF.
  ENDWHILE.
```

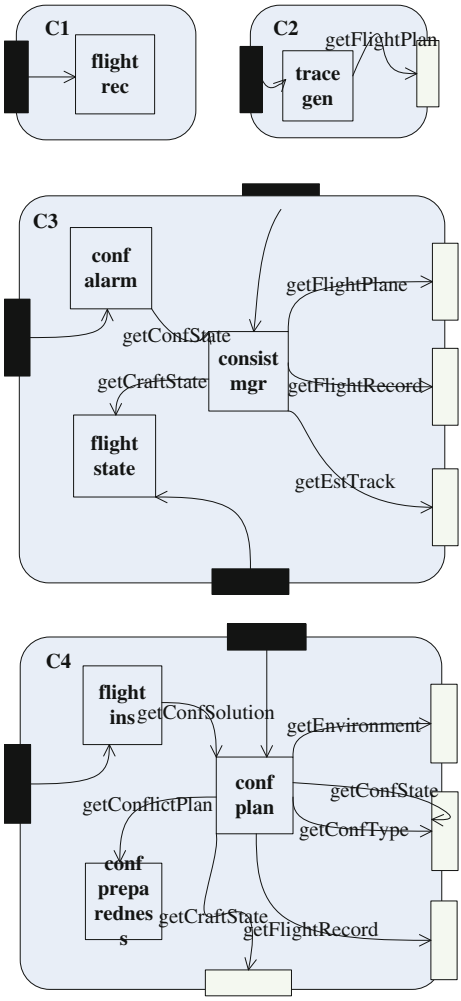


Fig. 30.6 Atom component of ATC segment

(3) This step is to compute the required and provided interface of the component

According to the definition, the interface set of atom component is union of operations in corresponding role set, so, we can obtain the provided and required interfaces using following formulas:

$$\mathrm{PIntf}(c) \;=\; \{op | \exists i \cdot op \in \mathrm{PRole}_i(c).OPS\} \tag{30.5}$$

$$\mathrm{RIntf}(c) \;=\; \{op | \exists i \cdot op \in \mathrm{RRole}_i(c).OPS\} \tag{30.6}$$

(4) Construct the mapping between feature and component.

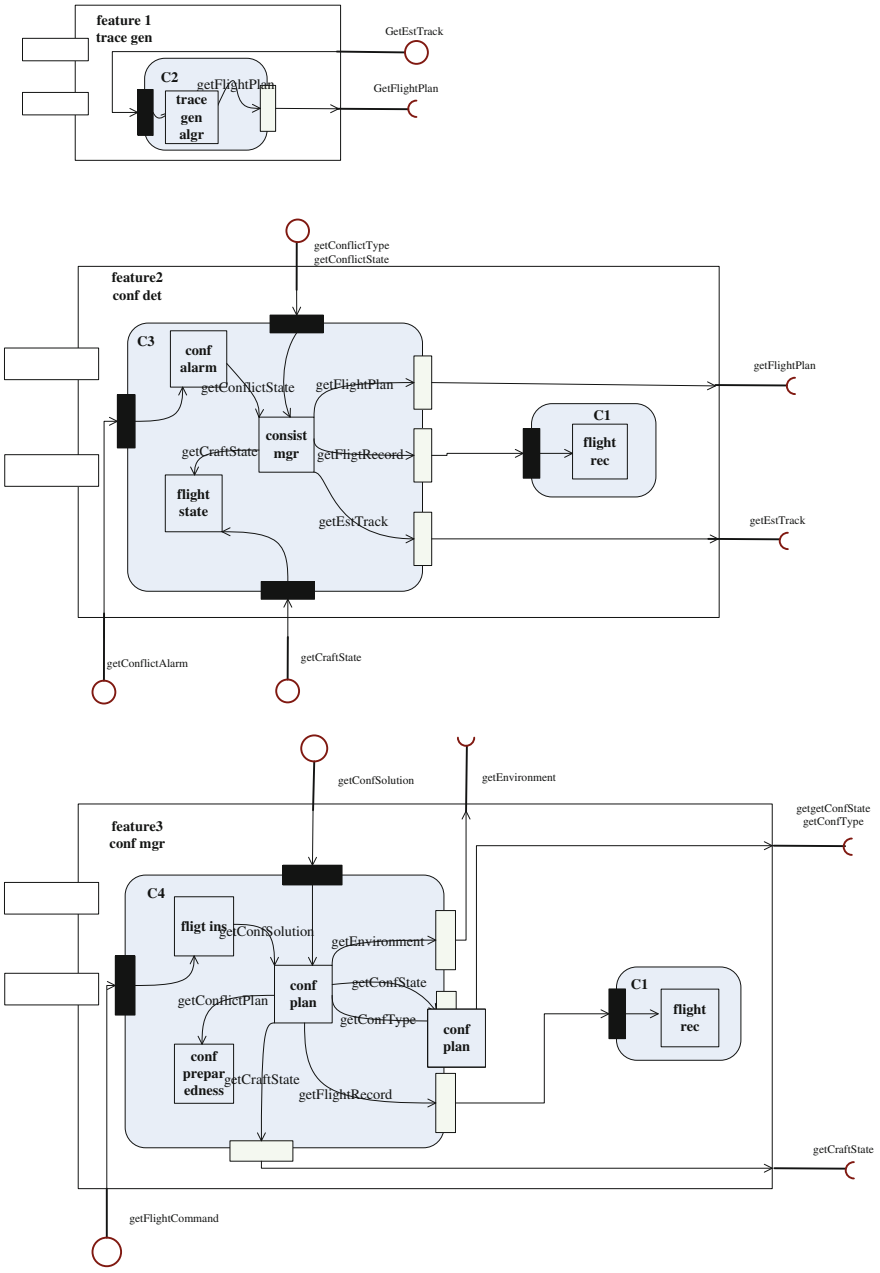


Fig. 30.7 Mapping from feature to component

In this step, we construct the (composited) component according to the role type and atom component interface. The sub algorithm shows as follows:

First, obtain the set of atom component (FCS) that constitute the (composited) component according to formula (30.7)

$$FCS(f) = \{c | \exists i \cdot f.role_i \in PRole(c)\} \tag{30.7}$$

And then, compute the set of provided interface and required interface, the set of provided role and required role according to the formula (30.8), (30.9), (30.10) and (30.11) respectively

$$\text{PIntf}(c_f) = \cup \text{PIntf}(FCS_i(f)) \tag{30.8}$$

$$\text{RIntf}(c_f) = \cup \text{RIntf}(FCS_i(f)) - \cup \text{PIntf}(FCS_i(f)) \tag{30.9}$$

$$PRole(c_f) = \cup PRole(FCS_i(f)) \tag{30.10}$$

$$RRole(c_f) = URRole(FCS_i(f)) \tag{30.11}$$

Finally, construct the composited component using all these set.

After the mapping, system designer implement the component model according to the role specification to finish the modeling. Because the model is interface-based, there is no limitation or special requirement in feature composition. Whenever the interface is matched, the features can composited together for free. It decouples the requirement and domain engineering, make the system more extensible and evolutional (Figs. 30.6, 30.7).

30.6 Conclusion and Future Work

This paper discusses deeply on the SPL and feature model technology and presents a role-based method mapping feature to component models; the method overcomes the shortcoming of no freely feature selection and combination. As the intermediary, the role and reference role decouples the feature and component, which makes convenient selection and composition between feature variants and enhances the componentization level of the system.

The presentation of the role increase componentization level of feature and model, but there may be a problem of over refinement, there may be a lot of tiny component without any semantics. One potential solution to the problem is to add semantics to the role, and then we can analyze the semantics of tiny component and combine components with same or similar semantics to a composite component. In addition, role evolution is also a research direction in the future.

References

1. Clements P, Northrop L (2002) Software product line: practices and patterns. Addison Wesley, Reading
2. Kang K, Cohen S, Hess J, Novak W, Peterson S (1990) Feature-oriented domain analysis (FODA) feasibility study, Technical Report CMU/SEI-90-TR-21, Software Engineering Institute, Pittsburgh
3. Kang K, Kim S, Lee J, Kim K, Shin E, Hug M (1998) Form a feature-oriented reuse method with domain-specific reference architecuture. Ann Softw Eng 5
4. Czarnecki K, AntKiewicz M (2005) Mapping features to models: a template approach based on superimposed variants. In: GPCE'05: proceedings of the fourth international conference on generative programming and component engineering
5. Guelfi N, Perrouin G (2007) A flexible requirements analysis aproach for software product lines. In: proceedings of the 13th international working conference on requirements engineering: foundation for software quality
6. Wang X, Liu S, Zhang J (2009) A framework based on domain model and component composition. Acta Electronica Sinica 37(3):540–545 in Chinese
7. Zhong-jie W, Xiao-fei X, De-chen Z (2006) Feature-based component model and normalized design process. J Softw 17(1):39–47 in Chinese
8. Xin P, Wen-geng Z, Yong-le Q (2006) Semantic representation and composition of business components based on domain feature ontology. Acta Electronica Sinica 34(12):2473–2477 in Chinese
9. Avila-García O, Estévez García A, Sánchez Rebull EV (2007) Using software product lines to manage model families in model-driven engineering. In: Proceedings of The 22nd annual ACM symposium on applied computing—model transformation track, ACM Press
10. Jun Z, Shu-fen L, Zhi-lin Y (2011) A feature model componentization method based on role. Acta Electronica Sinica 39(2):304–308

Chapter 31
The Design and Implement of Wireless Network of Intelligent Lighting in Classroom

Fashan Yu, Mingjie Zong and Chaochao Han

Abstract After introducing the traditional lighting instrument and its control character in this paper, I use the wireless communication technology—ZigBee to realize the intelligent lighting in classroom. Network formation of cluster-like tree structure layout is adopted in the system, CC2430 is used as the hardware core. At the end of the paper, the system flow and the function implement is stated. The main body consists of system flow, system initialization, network formation binding, data transmission and dimming with PWM.

Keywords Intelligent lighting · ZigBee · Cluster-like tree structure · CC2430

31.1 Introduction

In the traditional control method, the equipment is connected with the controller in the control room by wire. Therefore, more power cable is used, the cost increases, the radiated interference increases and electromagnetic environment is polluted. At present, energy conservation of lighting system in schools is far from expectation. So, economic losses occur. In order to solve the issue above, we put LED lighting with high luminance in use. The LED lighting has advantages such as low

F. Yu · M. Zong (✉) · C. Han
Henan Polytechnic University, Jiao zuo 454000, Henan, China
e-mail: zongmingjie2011@163.com

F. Yu
e-mail: yufs@hpu.com

C. Han
e-mail: 542228195@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_31,

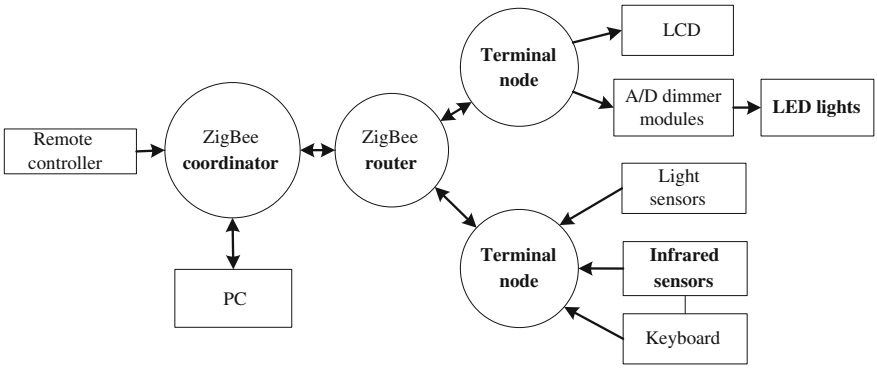


Fig. 31.1 Layout of classrooms and wireless networking

supply voltage, low power consumption, long-life and ease to control. Compared with filament lamps, 80% of the power can be saved. By connecting the lighting control system based on ZigBee with the host computer, an entirely distributed lighting control system based on microcomputer is built. It can be used for centralized control and management [1].

31.2 System Planning and Layout

The whole network using cluster-like tree structure, with each classroom as a unit, all LED lighting as an output terminal equipment, or LED lights inside the classroom is divided into several groups (group numbers can be real light on the number of classrooms) as the number of output devices. Classroom light sensors, infrared sensors and the human body are keyboard input devices. Each classroom sets up a router node, and in the entire school building is located a coordinator; the router can communicate with the coordinator, the coordinator may be PC to PC for data transfer and communication [2]. To simplify our experimental model input and output terminals are treated as terminal nodes. The whole layout of classrooms and wireless networking is shown in Fig. 31.1.

Because school classrooms have different functions, so it can be different depending on the circumstances of the classroom for the functionality of terminal equipment set. In general, all classrooms require for all the terminal devices switch control of LED lamps and brightness adjustment.

In order to facilitate quick adjustment, a remote control panel must be added for brightness settings and switch control, which is set to dim the brightness level and vice versa. The system must also have the following functions:

a. Intelligent light: the light intensity will gradually brighten; lights out when there is a growing dim light brightness. Protects vision and extends the lamp life.

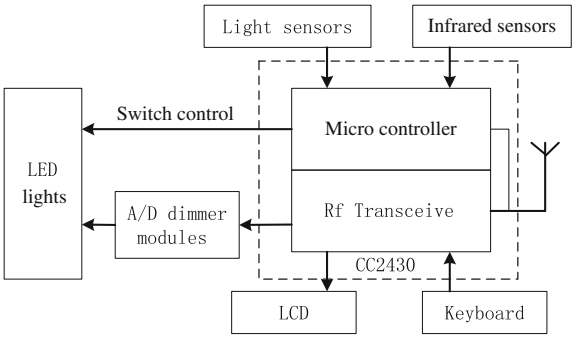


Fig. 31.2 Hardware structure of the terminal node

b. Delay function: press the edge switch or send the delay instruction, all the lights will turn off after a certain period of time.
c. Full off: a key to achieve full light of all relevant functions.

31.3 Design of Node Hardware

Because ZigBee coordinator nodes and router nodes have off-the-shelf modules that can be conveniety purchased, and product technology is mature, the price is cheap, and we will not have to design that which can be bought. According to the specific circumstances of the classroom, I want to emphasize the design of terminal equipment node stated before. To be simple, we design the model with the whole classroom in mind as a terminal equipment, with all the sensor and input/output devices integrated in a node. The node can make communication with the ZigBee router, and the hardware structure of the terminal node is shown in Fig. 31.2.

Here we use CC2430 chips as the core of the terminal node circuit, and it is a single-chip microcomputer which is accord with the standard of technology ZigBee system. The chip is integrated with the 8051 micro controller enhanced core and 2.4 GHz rf transceive that complies with the IEEE802.15.4. So it has excellent sensitivity to wireless acceptance and strong anti-jamming performance as well as low consumption. Its support from hardware CSMA/CA mechanism, integrated with ADC, AES safety coprocessor, USART and rich I/O interfaces. Simply a few components such as crystals can complete a node design [3]. The terminals of the node has 44 keyboard used to set the parameters of the node as well as send instructions and 12864 dot matrix LCD module used to display network state as well as the light of the information. Micro controller output switch direct finish

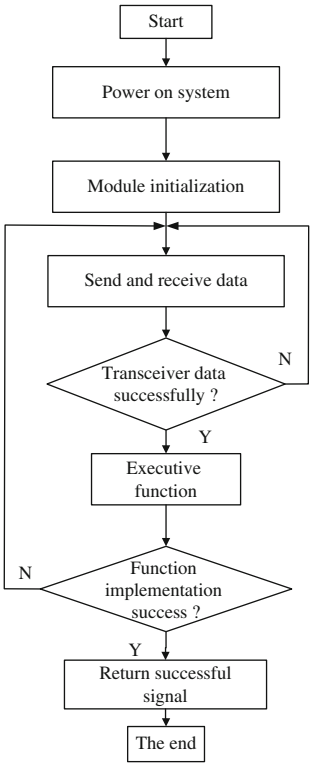


Fig. 31.3 System process

lamp switch control, and micro controller realize the 256-levels light control of the lamp by 8 bits of A/D converter. The nodes has also light sensors and infrared sensors. The optical sensors are used for the real-time measuring of indoor light intensity, and then converted to digital signals to the microprocessor, so as to produce a suitable PWM signal, and then the PWM signal is transmitted to light module to adjust the brightness of the light. Infrared sensors are used to monitor, and according to the results it can directly determine lamp switch save power.

31.4 Working Process and Functions Realization

31.4.1 System Process

The whole system software flow chart is shown in Fig. 31.3. Explain the function realized by the system as a whole, which sends instructions through transmitter module and then decodes it, so as to control the LED light and closing.

31.4.2 Power on the Whole System

After powering on system, at the first place operate a series of initialization, including liquid crystal display, buttons, CC2430 module, timer and UART interface initialization. After the initialization, the system will start a self-acting network function according to the algorithm, including sending access request and node net request, our model is a terminal node request, and then set up network. First the tuner sends the net request signal, the router will receive it, coordinator and router were assigned a related ID address seperately, and then the router will send a StartZBNwk function to write instructions to CC2430 to start a ZigBee protocol stack. After launching the protocol stack, the system will use GetDeviceInfo function to get the equipment information CC2430 send to ARM through the UAR. After CC2430 get a 16 bit short address, the system will display network status information on the display screen and upper machine, so as to support on completion of the whole electricity process [4].

31.4.3 Network Binding

After the Electric, system get into the network binding condition. First the system call the ZB_APP_REGISTER_REQUST() function to write configuration in the ZigBee registers, then call ZgBeeAllowBind() function to send allow binding orders. After that, the router will call menudis() function for binding, the specific operation ias follows: First scan button in order to get binding coordinator address required by the user input, then call ZgBeeZD-OFindIEEE function to find the physical address in the network. After getting the adress, the system will call ZgBeeBind function to bind and show that binding results out, thus completing network binding.

31.4.4 Data Transmission and Resizing Light

The system can complete the data transmission network and resize light. System uses ZgBeeSendData() function for data transmission, the main parameters of the function includes the destination address, command type, and data, but the destination address and user data are to be written register through the keyboard scan users.

We use PWM signal modulation to resize light, concrete implementation: using a while loop which embedded a case to implement, while loop is responsible for checking whether the arrival of the external command, each branch of the case produce the corresponding PWM signal, When it finds the external command coming, the program jumps out of the branch and come into the appropriate branch

to generate new profile [5]. For the PWM signal, we use the implementation of timing cycle, that is, outputting 0 and 1 through different length of delay achieve different dutyfactor square wave. The basic calculation as follows: If unit delay is 1/12 of the crystal, the system output 1after [(square wave cycles/unit time delay) × duty cycle] delays, and output 0 after [(square wave cycles/unit time delay) × (1 − duty cycle)] delays, then repeat cycle can be realized PWM function.

31.5 Conclusion

Intelligent lighting control system and its applications have made some progress. Also a lot of good products have appeared, compared to previous times with greatly improved technology to hardware. Some mature products have been put into use in industrial production. Intelligent lighting control system technology applied to schools is still relatively rare, though great development is going on. For this issue in the research content, it now has primary application on other systems and the initial data. According to the characteristics of the system, the theory and analysis methods available in this system, it passes our model experiments, has significant effect, Improves information, intelligent lighting system, has a great role in saving energy, and complies with the national energy-saving and emission-reduction strategy.

References

1. Zhang D (2003) The intelligent controller for classroom lighting [J]. Guizhou Industrial University (Natural Science)
2. Wang D (2006) Wireless sensor network system design and application. Doctoral dissertation, Chongqing University
3. Zhu R (2004) Field bus environmental multi-sensor data fusion. Master's thesis, Shanghai University
4. Jiang X (2005) ZigBee technology and its applications. J Low Volt Electr 7:27–33
5. Pan X, Wang Y (2002) Micro-computer control technology [M]. Electronic Industry Press, Beijing

Chapter 32
Application of Hydraulic System in Wave Energy Converter

Yonggang Lin, Wei Huang, Dahai Zhang, Wei Li and Jingwei Bao

Abstract Energy and environment have always been the two main issues in today's world. As green renewable energy is abundant, wave energy has become a hotspot in the area of renewable energy. The state of current research in the wave energy converter is described in this paper, especially, the application of hydraulic systems. The stability of output power, the output frequency of hydraulic system in wave energy converter and hydraulic system of manifolding wave energy converters in the wave energy converter are also discussed in this paper.

Keywords Wave energy converter · Hydraulic system · The stability of output power · The stability of output frequency · Manifolding wave energy converters

32.1 Introduction

Because of the oil crisis in the 1970s and the shortage of oil today, humans focus their attention on clean renewable energy. Thus, solar energy, wind energy, geothermal energy, ocean energy and other renewable green energy are becoming more and more important around the world. According to the data published by IPCC Scoping Conference on Renewable Energy 2008: the global wave energy reserves is

Y. Lin (✉) · W. Huang · D. Zhang · W. Li · J. Bao
State Key Laboratory of Fluid Power and Control,
Zhejiang University, Zheda Road. 37, Hangzhou 310027, China
e-mail: yglin@zju.edu.cn

W. Huang
e-mail: zuosi1314@163.com

D. Zhang
e-mail: zhangdahai1027@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_32,

8,000–80,000 TWH/year or more in theory [1], that is to say, if only 50% of wave energy resources can be used, the global demand for electricity will be met.

Wave energy converter (WEC) programs can be divided into pneumatic, mechanical and hydraulic by energy transmission. In the Pneumatic WEC, conversion efficiency is low, typically only 10–30%. Mechanical WEC has a high failure rate, the cost of maintenance and repair is high, and the mechanical structure will greatly increase the volume of the device which makes it hard to be installed. At present, most WECs use hydraulic energy transmission, so the hydraulic system has great meaning and prospect in the research of WEC [2].

32.2 Application of Hydraulic System in WEC

Today, WEC is usually divided into attenuators, point absorbers and terminators by its power generation type at home and abroad.

32.2.1 Application of the Hydraulic Technology in Attenuators

In the attenuators, a number of hinged rafts are floating on the water to capture wave energy, then the hydraulic system drives generator to generate electricity. The researchers find that when the system's natural frequency equals the wave frequency, the converter can reach the maximum output efficiency [3].

At present, attenuators are seldom used in the commercial; Pelamis designed by Ocean Power Delivery in the UK is the first commercial operation and the offshore wave energy demonstration project in the world [4].

Pelamis allows angular displacement in two directions, so the wave resistance ability is greatly improved compared to traditional attenuators which can have angular displacement in only one direction. Pelamis P-750 machines are capable of producing 750 kw, the total length is 150 m, and is placed in the depth of 50–60 m under the sea [5].

Pelamis uses the hydraulic system. The hydraulic principle is shown in Fig. 32.1; the most critical technology is changing unstable wave energy to stable hydraulic energy by high-pressure accumulator and then generating electricity by driving the generator.

32.2.2 Application of the Hydraulic Technology in Point Absorbers

Point absorbers WECs have developed rapidly in recent years; they capture wave energy by the buoy. Professors in Lancaster University proved that the point absorbers capture more wave energy compared to other devices in the mathematical theory [6].

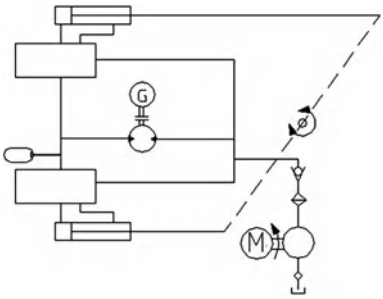


Fig. 32.1 Hydraulic principle of Pelamis

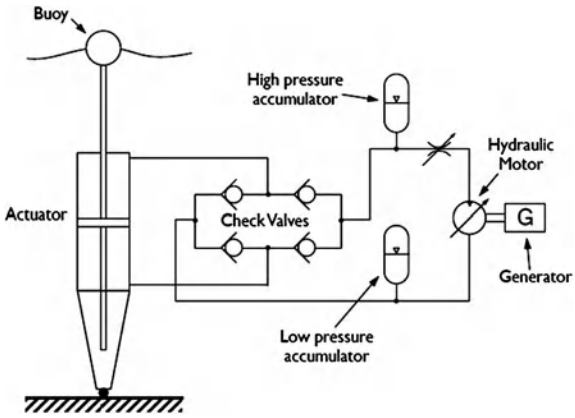


Fig. 32.2 Hydraulic principle of point absorbers

A simple hydraulic principle of point absorbers is shown in Fig. 32.2, the buoy floats up and down along with the waves, it makes the oil flow out from the hydraulic cylinder to drive the hydraulic motor. As energy is fluctuating because of the instability of wave, a group of high pressure accumulators and low pressure accumulators are used to store and release the excess energy in the system. The gas used in accumulators is usually nitrogen, which can reach 50–60 Mpa [7].

WRASPA designed by Lancaster University, and Rexroth designed by Bosch, are successful cases at present.

32.2.3 Application of the Hydraulic Technology in Terminators

Terminators WECs utilize their own geometry to avoid the wave radiating back and reduce the wave-making resistance, such as Salter duck and Pendulum devices.The principle of Pendulum WEC is first proposed by the Professor WATABE Tomiji in Muroran Institute of Technology [8, 9].

WaveRoller is a pendulum WEC designed by AW-Energy Company in Finland [10]: under the push of wave, the pendulum rotates around the axis, and the

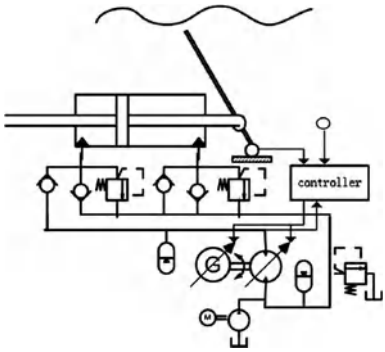


Fig. 32.3 Hydraulic principle of WaveRoller

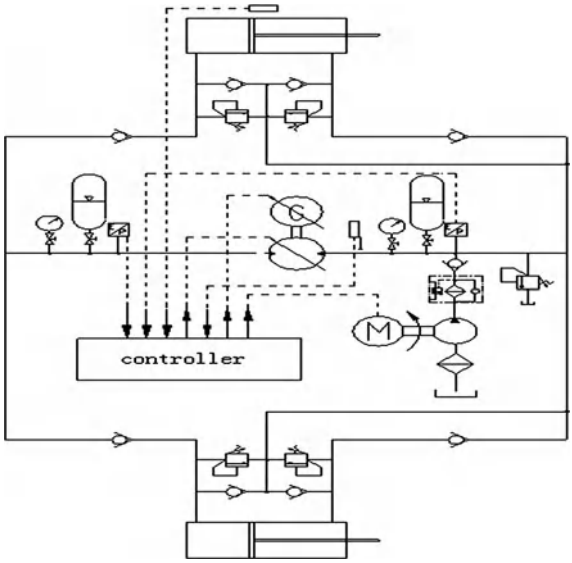


Fig. 32.4 Hydraulic principle of inverse pendulum

hydraulic cylinders collect the kinetic energy generated by the pendulum, and then the generators set on the shore convert it to electricity, the average power of the device is up to 13 KW. The hydraulic principle is shown in Fig. 32.3; it can be seen from the figure that the hydraulic system only acts when the pendulum swings clockwise.

State Key Laboratory of Fluid Power and Control in Zhejiang University has designed the 10 kW wave energy converter of inverse pendulum (hydraulic principle is shown in Fig. 32.4). The prototype plant test and the semi-physical simulation test have been done. And the operating principle and the feasibility of the wave energy converter of inverse pendulum have been verified. Compared to WaveRoller, it can capture more wave energy in the same time.

32.3 Key Technologies of Hydraulic System in Wave Energy Converter

32.3.1 Stability of Output Power

As the wave is random, it is inevitable that the output power of WEC is unstable, so how to solve this problem becomes the key issue in the design process of hydraulic system naturally. The flywheel and accumulator as the energy storage system component to reduce the energy fluctuation are usually used.

The hydraulic principle of the variable hydraulic motor with flywheel is shown in Fig. 32.5, the hydraulic motors and generators connect through the flywheel. When the waves are large, the speed of hydraulic motors is high, and flywheel stores excess energy; and when the waves are small, the speed of hydraulic motor is low, flywheel releases energy. In this way, it can ease the energy fluctuations in the system and stabilize the speed of generator.

However, because the flywheel itself will consume part of the energy, and its installing and maintenance are not easy, the hydraulic system of WEC rarely uses the flywheel as energy storage component. Compared to the flywheel, accumulator has many advantages, for example, it is easy installation, long life, low noise, small fluctuation and can supply hydraulic pressure of the system, keep the safe pressure of the system. So, now the accumulator is generally used in the hydraulic system of WEC, its hydraulic principle is shown in Fig. 32.3, the accumulator plays the same part as the flywheel. The common accumulator hydraulic system sets the working pressure base on a defined range of pressure, but the pressure in the hydraulic system of WEC is between zero and limit pressure, therefore the design of the accumulator in the common hydraulic system is not suit to the hydraulic system of WEC.

Professor António F. de O. Falcão in Lisbon University of Portugal models and simulates the hydraulic system of WEC. The results show that the hydraulic system with accumulators can transform the wave energy with a stable output power better. But there is no perfect solution to stabilize the output power [11]. In this paper, AMESim is used to model and simulate the hydraulic system shown in Fig. 32.4, and the results (Fig. 32.6) show that the hydraulic system with accumulator can make the output power stabilize at around 5.5 kW.

32.3.2 Stability of Output Frequency

Considering that electric energy generated by WEC must be supplied to the grid, the frequency should be the same as the frequency of grid in order to ensure the quality of grid electricity, so the stability of output frequency in WEC relates closely with its commercial value.

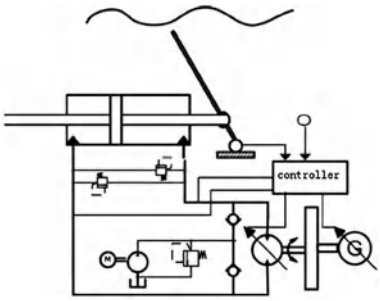


Fig. 32.5 Hydraulic principle

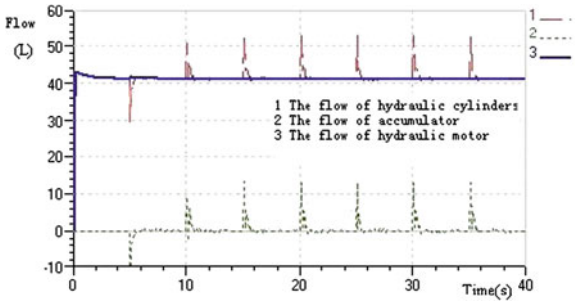


Fig. 32.6 Results of the simulation

The energy changes along with the oscillation of the wave speed, so the pressure and flow are always different in the hydraulic system. Variable displacement motor is commonly used to solve this problem. Formula of the variable displacement motor speed is shown as follows:

$$n = \frac{q_0}{60Q\eta_v}$$

The flow Q in the hydraulic system of WEC always changes. If the displacement q_0 can change along with the flow Q, it can stabilize the rotational speed at a fixed value and will result in a stable output power. Pelamis and WaveRoller both use the variable displacement motor to stabilize the generator's output frequency, it adjusts the displacement in real time according to the wave condition, the generator will achieve stable output rotational speed and export AC with frequency constant. Meanwhile, using the variable displacement motor can make generator speed around the rated speed which improves the efficiency of the generator.

In this paper, the co-simulation experiment of Simulink and AMESim on the hydraulic system shown in Fig. 32.4 is done. This research uses varying displacement motor and variable load to achieve the constant frequency of generator output, and the result is shown in Fig. 32.7. The generator speed can be stabilized in the 750 rpm which is the generator's rated speed. This indicates that both the control of variable load and the adjustment of variable displacement motor can achieve a constant speed of generator output. At the same time, the control of variable load can achieve a better effect than the adjustment of variable displacement motor as shown in the result.

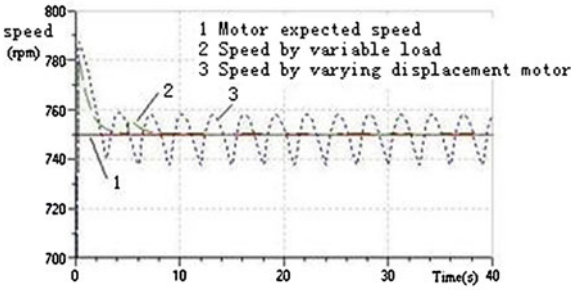


Fig. 32.7 Results of simulation

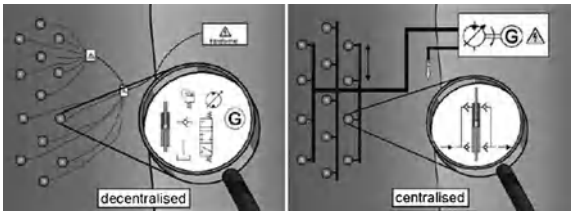


Fig. 32.8 Centralized and decentralized systems

32.3.3 Manifolding Wave Energy Converters

Because one WEC can only generate a little electricity, so the research of the manifolding WECs is the only way to use the wave energy in the commercial.

To reach this purpose, there are two different approaches: decentralized and centralized, which mainly differ in the point where the mechanical energy is ideally converted into electricity. If wave energy is converted into electricity in each WEC, the system refers to a decentralized system. The opposite is the transportation of fluid power energy to a remote location, for example with a pipeline, and converts it into electricity inside a station. This method will be referred to as a centralized system. Both approaches have their justification, but for many cases the advantages of the centralized are on hand: many converters can be used to feed a pipeline, which transmits the absorbed wave power through a fluid onshore. There, one or more turbines are used to generate electricity. This strategy, also called manifolding, can reduce the mechanical and electrical complexity of the WEC and hence contributes to the systems' stability and maintainability since vulnerable components are placed safely onshore. Figure 32.8 illustrates the differences of centralized and decentralized systems and clarifies their complexity as well [12].

In a centralized program, several WECs output hydraulic energy to drive a hydraulic motor. In the intersection of the pipelines, the pressures in different pipes are not equal, the output power curve of hydraulic cylinder is similar to sine curve because of the fluctuations of the wave energy. If the hydraulic energy in different pipes converges, they must interference each other, so before they come together, we must adjust them to match each other. The principle of the grid can supply

reference to solve the interference between the different plants, but they are completely different in nature. The key technologies and difficulties in the research of manifolding wave energy converters is how to conduct a reasonable design.

Waveroller 300 kw is modular and there is no natural upper limit to the number of plates that can be used and can offer a high level of scalability [13]. Onshore computers monitor the waves and now in response to each incoming wave, they are able to actively tune the hydraulic circuit. Every 20 min, they can control the buoyancy of the flap by moving water around inside the swinging "door" part [14].

32.4 Conclusion

This paper describes the state of current research of WEC, especially, the application of hydraulic system in it. Wave energy technology is a useful solution to solve the energy and environmental issues, and the hydraulic transmission is an important component. Deeper research should be done on how to improve the efficiency, power stability, frequency stability, life and other issues in the future.

Acknowledgments This work is supported by the State Key Program of National Natural Science of China (Grant No. 50735004) and the Program of National Natural Science of China (Grant No. 50975253).

References

1. Soerensen HC, Weinstein A (2008) Ocean energy: position chapter for IPCC. In: IPCC scoping conference on renewable energy, vol 837, IPCC, Germany, pp 1–2
2. Zhang D, Li W, Lin Y (2009) Wave energy in China: current status and perspectives. Renewable energy, vol 34. Elsevier, England, pp 2089–2092
3. Haren P (1978) Optimal design of Haren–Cockerell raft, vol 13. Massachusetts Institute of Technology, Cambridge, pp 23–26
4. Henderson R (2006) Design, simulation, and testing of a novel hydraulic power take-off system for the Pelamis wave energy converter. Renewable energy, vol 31. Elsevier, Edinburgh, pp 271–283
5. http://www.pelamiswave.com/content.php?id=142
6. http://www.engineering.lancs.ac.uk/lureg/group_research/wave_energy_research
7. Antonio F, de Falcao O (2010) Wave energy utilization: a review of the technologies. Renewable and sustainable energy reviews, vol 14. Elsevier, Lisbon, pp 899–918
8. Thorpe TW (1999) A brief review of wave energy, 1st edn. UK Department of Trade and Industry, UK
9. Thorpe TW (1999) An overview of wave energy technologies: status, performances and costs. Wave power: moving towards commercial viability, vol 26. UK Department of Trade and Industry, London, pp 50–120
10. Koivusaari R (2003) A process and an apparatus for utilising wave energy. Finland: WO/2003/036081
11. Virvalo T (2009) Hydraulic systems in wave energy application, 1st edn. World Publishing Corporation, China, pp 56–60

12. Kamizuru Y, Liermann M, Murrenhoff H (2010) Simulation of an ocean wave energy converter using hydraulic transmission. In: 7th international fluid power conference, vol 7. Fluid Power Institute, Germany, pp 108–117
13. http://www.aw-energy.com/concept.html
14. http://cleantechnica.com/2010/06/20/next-gen-waveroller-is-not-a-stupid-device/#

Chapter 33
Dynamical Output Feedback Control for Distributed Delay Systems

Juan Liu

Abstract The dynamical output feedback controller design for systems with distributed delay is discussed in this chapter. By neutral transformation and constructing proper Lyapunov-Krsovskii functional, the stability criterion is obtained which guarantees the asymptotical stability of closed-loop system. Then the parameterization of controller is used and the design condition of the desired controller is established in terms of Linear matrix inequality. Finally, a simulation is given to show the effectiveness of the proposed method.

Keywords Distributed delay · Dynamical output feedback · Lyapunov-Krsovskii functional · Linear matrix inequality (LMI)

33.1 Introduction

Distributed time-delay is commonly encountered in various physical and engineering systems. It has been shown that the existence of delay is the source of instability and poor performance of control systems. Therefore, stability analysis and controller design of linear systems with distributed delay have attracted much attention.

To stabilise the analysis of time-delay systems, in order to reduce the conservatism of stability criteria, many researchers developed different approaches, such as model transformation [1], free matrices [2, 3] and constructing proper L-K functional [4, 5]. However, though the controller design depends on the stability

J. Liu (✉)
Department of Mathematics and Information Science, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: liujuan@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_33,

results, many stability results cannot be used for controller design. Therefore, for the control of time-delay, solving the parameters of controller is important as well as reducing the conservatism. In Ref. [6], the state feedback control of distributed delay system was investigated. In Ref. [7], the controller design of linear system with distributed delay is studied via static output feedback. In Refs. [8, 9], the dynamical output feedback controllers were designed for distributed delay systems. To the best of our knowledge, few results have been achieved on delay-dependent dynamical feedback stabilization for linear systems with distributed delay.

Based on the above discussion, in this note, we study the stabilization problem for distributed delay system via dynamical output feedback. Through neutral model transformation of closed-loop system and constructing proper L-K functional, a delay-dependent stability criterion is obtained. Then by using parameterization of controller, the design condition of the controller, which guarantees the asymptotical stability of closed-loop system is established in terms of Linear matrix inequality (LMI), and an example is also given to illustrate the effectiveness of the proposed method.

33.2 Problem Formulation

Consider the following system with distributed delay:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + A_1x(t-h) + A_2\int_{t-d}^{t} x(\sigma)d\sigma + Bu(t) \\ y(t) &= Cx(t), x(t) = \varphi(t), t \in [-\gamma, 0] \end{aligned} \tag{33.1}$$

where $x \in \Re^n, u \in \Re^m, y \in \Re^h$ is state, control input and output respectively; $h, d > 0$ are constant delays $\gamma = \max\{h, d\}$, A, A_1, A_2, B, C are matrices with appropriate dimension.

The aim of the note is to design the following dynamical output feedback controller:

$$\begin{aligned} \dot{\hat{x}}(t) &= A_c\hat{x}(t) + B_cy(t) \\ u(t) &= C_c\hat{x}(t) + D_cy(t), t \geq 0 \end{aligned} \tag{33.2}$$

which guaranteed the asymptotical stability of the closed-loop system.

$$\dot{\xi}(t) = \bar{A}\xi(t) + \bar{A}_1\xi(t-h) + \bar{A}_2\int_{t-d}^{t} \xi(\sigma)d\sigma, t \geq 0 \tag{33.3}$$

where

$$\xi = \begin{bmatrix} x \\ \hat{x} \end{bmatrix}, \bar{A} = \begin{bmatrix} A + BD_cC & BC_c \\ B_cC & A_c \end{bmatrix}, \bar{A}_1 = \begin{bmatrix} I_n \\ 0 \end{bmatrix}[A_1 \quad 0], \bar{A}_2 = \begin{bmatrix} I_n \\ 0 \end{bmatrix}[A_2 \quad 0].$$

For closed-loop system (33.3), let

$$G\xi_t = \xi(t) + \bar{A}_1 \int_{t-h}^{t} \xi(\sigma)d\sigma + \bar{A}_2 \int_{t-d}^{t} (\sigma - t + d)\xi(\sigma)d\sigma$$

then we have

$$\frac{d}{dt}(G\xi_t) = \hat{A}\xi(t), t \geq 0 \tag{33.4}$$

where $\hat{A} = \bar{A} + \bar{A}_1 + d\bar{A}_2$.

In order to get the main result, we need the following lemmas:

Lemma 1 [3] *If there exist matrix* $M > 0$ *and scalars* $\alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 < 1$ *satisfying the following constraints*:

$$\begin{bmatrix} h^2\bar{A}_1^T M\bar{A}_1 - \alpha_1 M & dh\bar{A}_1^T M\bar{A}_2 \\ dh\bar{A}_2^T M\bar{A}_1 & d^2\bar{A}_2^T M\bar{A}_2 - \frac{3\alpha_2 M}{d^2} \end{bmatrix} < 0 \tag{33.5}$$

then the operator G is stable.

Lemma 2 [10] *Given scalars* α, β *and matrix* $S > 0$, *if* $\alpha > \beta$, *then*

$$\left(\int_{\beta}^{\alpha} w(\sigma)d\sigma\right)^T S\left(\int_{\beta}^{\alpha} w(\sigma)d\sigma\right) \leq (\alpha - \beta)\int_{\beta}^{\alpha} w^T(\sigma)Sw(\sigma)d\sigma$$

33.3 Main Results

Theorem 1 *If Lemma 1 holds and there exist symmetric positive-define matrices* $P = \begin{bmatrix} P_1 & P_3 \\ P_3' & P_2 \end{bmatrix} \in \Re^{2n\times 2n}; S_i \in \Re^{n\times n}, i = 1, 2$ *satisfying the following inequality*:

$$\begin{bmatrix} \hat{A}^T P + P\hat{A} & h\hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} & \hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} & h\begin{bmatrix} A_1^T \\ 0 \end{bmatrix} & d^2\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} \\ * & -hS_1 & 0 & 0 & 0 \\ * & * & -S_2 & 0 & 0 \\ * & * & * & -hS_1^{-1} & 0 \\ * & * & * & * & -2S_2^{-1} \end{bmatrix} < 0 \tag{33.6}$$

then the closed-loop system (33.3) is asymptotically stable.

Proof Constructing the Lyapunov-Krasovskii functional:

$$\begin{aligned} V(\xi_t) &= V_1(\xi_t) + V_2(\xi_t) + V_3(\xi_t) \\ &= (G\xi_t)^T P(G\xi_t) + \int_{-h}^{0}\int_{t+\theta}^{t} \xi^T(\sigma)\begin{bmatrix} A_1^T \\ 0 \end{bmatrix} S_1 [A_1 \quad 0]\xi(\sigma)d\sigma d\theta \\ &\quad + d^2 \int_{0}^{d}\int_{t-\theta}^{t} (\sigma - t + \theta)\xi^T(\sigma)\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} S_2 [A_2 \quad 0]\xi(\sigma)d\sigma d\theta \end{aligned} \tag{33.7}$$

then the derivative of $V_1(\xi_t)$ along the trajectory of the closed-loop system (33.3) is:

$$\dot{V}_1(\xi_t) = 2\left(\frac{d(G\xi_t)}{dt}\right)^T P(G\xi_t) = \xi^T(t)\left(\hat{A}^T P + P\hat{A}\right)\xi(t) + \eta_1 + \eta_2 \tag{33.8}$$

where

$$\begin{aligned} \eta_1 \le\, & h\xi^T(t)\hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} S_1^{-1} [P_1 \quad P_3]\hat{A}\xi(t) \\ & + \frac{1}{h}\int_{t-h}^{t} \xi^T(\sigma)\begin{bmatrix} A_1^T \\ 0 \end{bmatrix} d\sigma S_1 \int_{t-h}^{t} [A_1 \quad 0]\xi(\sigma)d\sigma \end{aligned} \tag{33.9}$$

$$\begin{aligned} \eta_2 \le\, & \xi^T(t)\hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} S_2^{-1} [P_1 \quad P_3]\hat{A}\xi(t) \\ & + d^2 \int_{t-d}^{t} (\sigma - t + d)\xi^T(\sigma)\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} S_2 [A_2 \quad 0]\xi(\sigma)d\sigma \end{aligned} \tag{33.10}$$

by Lemma 2, we also have:

$$\dot{V}_2(\xi_t) \le h\xi^T(t)\begin{bmatrix} A_1^T \\ 0 \end{bmatrix} S_1 [A_1 \quad 0]\xi(t) - \frac{1}{h}\int_{t-h}^{t} \xi^T(\theta)\begin{bmatrix} A_1^T \\ 0 \end{bmatrix} d\theta S_1 \int_{t-h}^{t} [A_1 \quad 0]\xi(\theta)d\theta \tag{33.11}$$

$$\begin{aligned} \dot{V}_3(\xi_t) = & \frac{d^4}{2}\xi^T(t)\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} S_2 [A_2 \quad 0]\xi(t) \\ & - d^2 \int_{t-d}^{t} (\sigma - t + d)\xi^T(\sigma)\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} S_2 [A_2 \quad 0]\xi(\sigma)d\sigma \end{aligned} \tag{33.12}$$

from (33.7)–(33.12), we can get:

$$\dot{V}(\xi_t) \le \xi^T(t)\left[\hat{A}^T P + P\hat{A} + \hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix}\left(hS_1^{-1} + S_2^{-1}\right)[P_1 \quad P_3]\hat{A} + h\begin{bmatrix} A_1^T \\ 0 \end{bmatrix}\right.$$
$$\left. S_1[A_1 \quad 0] + \frac{d^4}{2}\begin{bmatrix} A_2^T \\ 0 \end{bmatrix} S_2[A_2 \quad 0]\right]\xi(t) \tag{33.13}$$

Therefore, by the Schur complement, if the matrix inequality holds, then $\dot{V}(\xi_t) < 0$, which means the closed-loop system (33.3) is asymptotically stable.

In order to solve functional parameters, the cone complement linearization iterative algorithm can be used. Here, we use another method. Introducing scalars $\lambda_i > 0, i = 1, 2$ and the constrains:

$$S_i \le \lambda_i I_n, i = 1, 2 \tag{33.14}$$

Then the matrix inequality (33.6) can be transformed to the following LMI:

$$\begin{bmatrix} \hat{A}^T P + P\hat{A} & h\hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} & \hat{A}^T \begin{bmatrix} P_1 \\ P_3^T \end{bmatrix} & h\lambda_1 \begin{bmatrix} A_1^T \\ 0 \end{bmatrix} & d^2\lambda_2 \begin{bmatrix} A_2^T \\ 0 \end{bmatrix} \\ * & -hS_1 & 0 & 0 & 0 \\ * & * & -S_2 & 0 & 0 \\ * & * & * & -hS_1 & 0 \\ * & * & * & * & -2S_2 \end{bmatrix} < 0 \tag{33.15}$$

To get the parameters of dynamical controller from (33.15), we introduce the following parameter set:

$$\Phi := \left\{X > 0 \in \Re^{n\times n}, Y > 0 \in \Re^{n\times n}, R \in \Re^{n\times n}, U \in \Re^{m\times n}, V \in \Re^{n\times h}, W \in \Re^{m\times h}\right\} \tag{33.16}$$

Let $\tilde{A} = A + A_1 + dA_2$ and $Z = X - Y^{-1}$, then we can get:

$$\begin{bmatrix} D_c & C_c \\ B_c & A_c \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ B & -Y^{-1} \end{bmatrix}\begin{bmatrix} W & U \\ V & R - Y\tilde{A}X \end{bmatrix}\begin{bmatrix} I_h & -CXZ^{-1} \\ 0 & Z^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} W & (-WCX + U)Z^{-1} \\ BW - Y^{-1}V & \left(-BWCX + BU + Y^{-1}VCX - Y^{-1}R + \tilde{A}X\right)Z^{-1} \end{bmatrix} \tag{33.17}$$

$$P^{-1} = Q = \begin{bmatrix} X & Z \\ Z & Z \end{bmatrix} \tag{33.18}$$

that is $P = \begin{bmatrix} Y & -Y \\ -Y & Z^{-1}XY \end{bmatrix}$. Substituting (33.17) into (33.3), we can get:

$$\hat{A} = \begin{bmatrix} \tilde{A} + BWC & (-BWCX + BU)Z^{-1} \\ BWC - Y^{-1}VC & \left(-BWCX + BU + Y^{-1}VCX - Y^{-1}R + \tilde{A}X\right)Z^{-1} \end{bmatrix} \tag{33.19}$$

From the above deduction, the main result can be obtained as follows.

Theorem 2 *If Lemma 1 holds, and if there exist scalars* $\lambda_i > 0, i = 1, 2$ *and symmetric positive-define matrices* $S_i, i = 1, 2$ *satisfying the LMIs*:

$$\begin{bmatrix} \Gamma_1 + \Gamma_1^T & h\Gamma_2^T & \Gamma_2^T & h\Gamma_3^T & d^2\Gamma_4^T \\ * & -hS_1 & 0 & 0 & 0 \\ * & * & -S_2 & 0 & 0 \\ * & * & * & -hS_1 & 0 \\ * & * & * & * & -2S_2 \end{bmatrix} < 0 \tag{33.20a}$$

$$\begin{bmatrix} X & I_n \\ I_n & Y \end{bmatrix} > 0 \tag{33.20b}$$

$$S_i \le \lambda_i I_n, i = 1, 2 \tag{33.20c}$$

then system (33.3) is asymptotically stabl,. where $\Gamma_1 = \begin{bmatrix} \tilde{A}X + BU & \tilde{A} + BWC \\ R & Y\tilde{A} + VC \end{bmatrix}$,

$$\Gamma_2 = \begin{bmatrix} R & Y\tilde{A} + VC \end{bmatrix}, \Gamma_3 = \lambda_1 \begin{bmatrix} A_1 X & A_1 \end{bmatrix}, \Gamma_4 = \lambda_2 \begin{bmatrix} A_2 X & A_2 \end{bmatrix}$$

Proof Let $T = \begin{bmatrix} I_n & Y \\ 0 & -Y \end{bmatrix}$, by (33.18), we know that the positive-definiteness of functional parameters is equivalent to $Q > 0$, so from the following $T^T QT = \begin{bmatrix} I_n & 0 \\ Y & -Y \end{bmatrix} \begin{bmatrix} X & Z \\ Z & Z \end{bmatrix} \begin{bmatrix} I_n & Y \\ 0 & -Y \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ Y & -Y \end{bmatrix} \begin{bmatrix} X & I_n \\ Z & 0 \end{bmatrix} = \begin{bmatrix} X & I_n \\ I_n & Y \end{bmatrix}$, we know (33.20b) is equivalent to the positive-definiteness of (33.18).

Then substituting (33.18) and (33.19) into (33.15), we can get:

$$\begin{bmatrix} \hat{A}^T P + P\hat{A} & h\hat{A}^T \begin{bmatrix} Y \\ -Y \end{bmatrix} & \hat{A}^T \begin{bmatrix} Y \\ -Y \end{bmatrix} & h\lambda_1 \begin{bmatrix} A_1^T \\ 0 \end{bmatrix} & d^2\lambda_2 \begin{bmatrix} A_2^T \\ 0 \end{bmatrix} \\ * & -hS_1 & 0 & 0 & 0 \\ * & * & -S_2 & 0 & 0 \\ * & * & * & -hS_1 & 0 \\ * & * & * & * & -2S_2 \end{bmatrix} < 0 \tag{33.21}$$

Constructing the similarity transformation matrix $diag\{TQ \quad I_{2n} \quad I_{2n} \quad I_{2n} \quad I_{2n}\}$, and Pre- and post-multiplying both sides of inequality (33.21), we can get (33.20a).

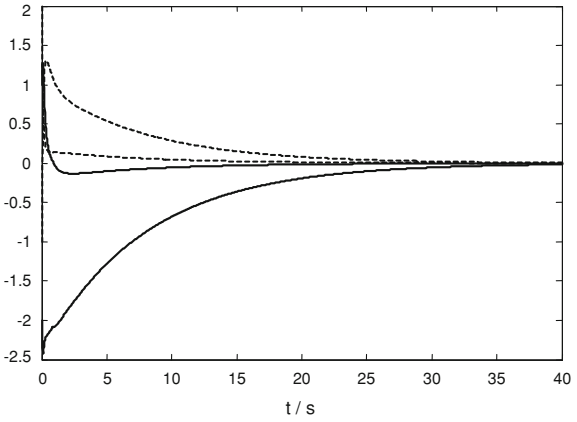


Fig. 33.1 State trajectory of closed-loop system

33.4 Simulation

Example Consider the following time-delay system:

$$\dot{x}(t) = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x(t-h) + \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} \int_{t-d}^{t} x(\theta) d\theta + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(t)$$

$$y(t) = [-1 \quad 1] x(t), t \geq 0$$

Choosing $\alpha_1 = 0.5, \alpha_2 = 0.2; \lambda_1 = 5, \lambda_2 = 5$ in equalities (33.5) and (33.20a, b, c), when $d = 0.3$, by Theorem 2 and using Matlab LMI Toolbox, we can get the allowable time delay is $\bar{h} = 0.95$ and

$$X = \begin{bmatrix} 36.45 & -1.06 \\ -1.06 & 1.66 \end{bmatrix}, Y = \begin{bmatrix} 11.24 & 14.74 \\ 14.74 & 19.35 \end{bmatrix}, R = \begin{bmatrix} -1.57 & 1.03 \\ 1.25 & -1.35 \end{bmatrix}$$

$$U = [-17.38 \quad 21.30], V = \begin{bmatrix} -7.24 \\ -18.42 \end{bmatrix}, W = 4.02$$

and the parameters of the dynamical output feedback controller is

$$\begin{bmatrix} D_c & C_c \\ B_c & A_c \end{bmatrix} = \begin{bmatrix} 6.32 & 7.01 & 15.33 \\ 2.80 & 2.12 & 18.23 \\ -3.08 & -3.57 & -19.15 \end{bmatrix}$$

Let initial condition $x'(t) = [1 \quad 2 \quad -1 \quad -2]$, the state trajectory of the closed-loop system is showed in Fig. 33.1, from which we can see the closed-loop system is asymptotically stable with the dynamical output feedback controller.

33.5 Conclusion

In this note, the design of dynamical output feedback controller for linear systems with distributed delay is investigated. By the neutral transformation and controller parameterization, the sufficient condition, which guarantee the asymptotical stability of the closed-loop system, is established in terms of LMI and simulation illustrate the effectiveness of the proposed method.

Acknowledgments This work is supported by the Doctoral Foundation Project of Henan Polytechnic University under Grant B2010-50, the Natural Science Project of Henan Province Office of Education under Grant 2008B510008 and Henan Provincial Open Laboratory for Control Engineering Key Disciplines under Grant KG2009-07.

References

1. Fridman E, Shaked U (2002) A descriptor system approach to H control of linear time-delay systems. IEEE Trans Autom Control 47:253–270
2. He Y, Wu M, She GP, Liu JH (2004) Parameter-dependent Lyapunov functional for stability of time-delay systems with polytotic-type uncertainties. IEEE Trans Autom Control 49:828–832
3. Yue D, Won S, Kwon O (2003) Delay dependent stability of neutral systems with time delay: an LMI approach. IEEE Proc Control Theory Appl 150:23–27
4. Qian W, Cong S, Sun YX, Fei SM (2009) Novel robust stability criteria for uncertain systems with time-varying delay. Appl Math Comput 215:866–872
5. Qian W, Liu J, Sun YX, Fei SM (2010) A less conservative robust stability criteria for uncertain neutral systems with mixed delays. Math Comput Simul 80:1007–1017
6. Xie L, Fridman E, Shaked U (2001) Robust H-infinity control of distributed delay systems with application to combustion control. IEEE Trans Autom Control 46:1930–1935
7. Choi HH, Chung MJ (1995) Memoryless H-infinity controller design for linear systems with delayed state and control. Automatica 31:919–971
8. Xu S, Chu YM, Lu JW, Zou Y (2006) Exponential dynamic output feedback controller design for stochastic neutral systems with distributed delays. IEEE Trans Syst Man Cybern Part A: Syst Hum 36:540–548
9. Xue X, Qiu D (2000) Robust H-infinity compensator design for time-delay systems with norm-bounded time-varying uncertainties. IEEE Trans Autom Control 45:1363–1369
10. Gu, K (2000) An integral inequality in the stability problem of time-delay systems. In: Proceedings of 39th IEEE conference decision control, Sydney, Australia, pp 2805–2810

Part III
Electric Machines and Electrical Apparatus

Chapter 34
Novel Modulation of Enhanced Z-Source Inverter to Minimize Switching Frequency and Volume of the Z-Source Passive Components

Chunwei Cai, Yanbin Qu and Kuang Sheng

Abstract Z-source inverters are proposed as single-stage power converters with buck-boost energy conversion ability. But their voltage boost inversion capability is limited. This paper proposes an enhanced Z-source inverter. Compared to the traditional Z-source inverter, it can obtain high voltage conversion factor with a short shoot-through duration, and can decrease Z-source capacitor voltage stress significantly. Moreover, a novel modulation of proposed Z-source inverter is presented to obtain same voltage gain as maximum constant boost control retaining the switching frequency unchanged. And it can greatly reduce the L and C volume in the Z-network under the six shoot-through intervals per carrier period. Simulation and experimental results have demonstrated the merits of the proposed method.

Keywords Z-source inverter · Boost ability · Maximum constant boost control · Shoot-through states · Voltage gain

34.1 Introduction

Recently, a Z-source inverter [1] which possess both voltage-buck and boost capabilities in one single stage has been used in distributed generation (DG) systems. The ZSI can boost the dc voltage and produce a desired output which is higher than dc bus voltage. In addition it can greatly reduce the output distortion

C. Cai (✉) · Y. Qu · K. Sheng
School of Information Science and Engineering,
Harbin Institute of Technology, Weihai, People's Republic of China
e-mail: Caichunwei@hit.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_34,

and improve the reliability since dead time is no longer needed and a short circuit across any phase leg is allowed. So, Z-source inverters have been developed in different directions [2–9]. Unfortunatly, the traditional voltage fed ZSI has some significant drawbacks such as the discontinuous input current in the boost mode and a high sustaining voltage of the capacitors. Authors in [2, 3] presented the improved Z-source inverters to overcome the drawbacks of traditional topology, but all these topologies cannot achieve infinite gain.

In order to produce the shoot-through in zero states the modified Pulse-width-modulation (PWM) controls for the Z-source inverter were investigated in [4–9]. The simple control method was applied to ZSI for the first time. A maximum boost control was presented to produce the maximum dc voltage at any given modulation index [4]. However, careful analysis has shown that it introduces adverse low-frequency current ripple in inductor. This drawback of the maximum boost control has been overcome by the maximum constant control [5], which could obtain the maximum voltage boost gain while maintain a constant shoot-through duty ratio. But all the control methods above are edge-insertion (EI) PWM, switching frequency is adversely doubled, and additional cyclic logic is needed for realization.

In this paper, we will propose an enhanced Z-source inverter topology with high voltage boost inversion ability. And a novel modulation of proposed Z-source inverter will be presented to achieve maximum constant voltage boost gain without increasing the switching frequency. Moreover, the component volume will be decreased by increasing the shoot-through states per carrier period. The constrains of the shoot-through time in SVPWM will be derived, therefore the maximum shoot-through duty radio can be deduced. Afterwards the operation principle and the merits of the proposed control method will be investigated and verified by simulation and experiments.

34.2 Enhanced Z-source Inverter

The Enhanced Z-source inverter is shown in Fig. 34.1. The difference between Z-source Inverter and enhanced Z-source Inverter is inductors. The inductors in enhanced topology are replaced by the switched inductor branches. The first branch is consisted by the two inductors (L_1, L_3) and three diodes (D_1, D_2, D_3), and the second branch is consisted by the two inductors (L_2, L_4) and three diodes (D_4, D_5, D_6). The energy is stored and transferred by both switched inductors from the capacitors to the dc bus in the switching duration.

Assuming that $L_1 = L_2 = L_3 = L_4 = L$, $C_1 = C_2 = C$, so we can get

$$v_{L1} = v_{L2} = v_{L3} = v_{L4} = v_L,\ v_{C1} = v_{C2} = v_C \tag{34.1}$$

When in the non-shoot-through state (including the active and null states), the inverter side can be simplified to be an equivalent current source, one can derive

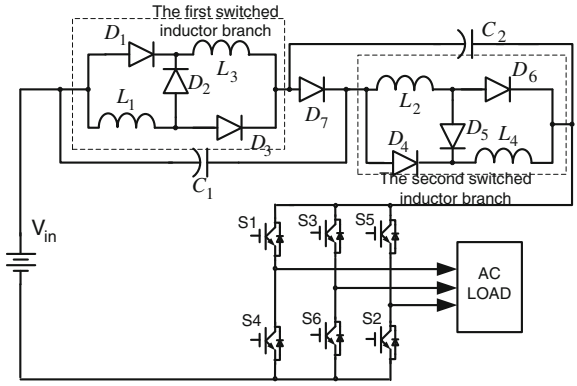


Fig. 34.1 Enhanced Z-source inverter

$$2v_L = -v_C,\ v_{\mathrm{PN}} = 2v_c + V_{\mathrm{in}} \tag{34.2}$$

When in the shoot-through state, the inverter side is shorted, one can get

$$v_L = V_{\mathrm{in}} + v_C,\ v_{\mathrm{PN}} = 0 \tag{34.3}$$

Because the average voltage across the inductors over one switching period is zero, we can deduce

$$V_c = \frac{2D_{\mathrm{sh}}}{1 - 3D_{\mathrm{sh}}} V_{\mathrm{in}} \tag{34.4}$$

where D_{sh} is the shoot-through duty ratio.

The peak dc-link voltage across the inverter phase legs V_{PN} and peak output phase voltage V_p can be expressed as

$$V_{\mathrm{PN}} = 2V_C + V_{\mathrm{in}} = \frac{1 + D_{\mathrm{sh}}}{1 - 3D_{\mathrm{sh}}} V_{\mathrm{in}} = BV_{\mathrm{in}},\quad V_p = M\frac{V_{\mathrm{PN}}}{2} = MB\frac{V_{\mathrm{in}}}{2} \tag{34.5}$$

where B is the boost factor and M is the modulation ratio.

34.3 The Modified Maximum Constant Boost Control

34.3.1 The Maximum Constant Boost Control

The maximum constant boost control [5] is shown in Fig. 34.2, which can obtain the maximum voltage gain at any given modulation index and overcome the low-frequency current ripple of the maximum boost control. There are five curves

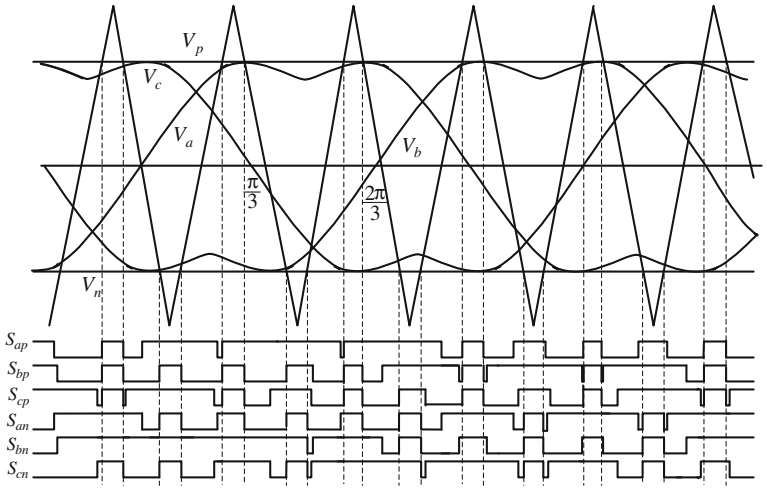


Fig. 34.2 Sketch map of maximum constant boost control

in this control method: three reference signals V_a, V_b, V_c, and two shoot-through envelope signals V_p and V_n.

As shown in Fig. 34.2, V_a reaches its peak value $(\sqrt{3}/2)M$ (V_p) while V_b is at its minimum value $-(\sqrt{3}/2)M$ (V_n) at $\pi/3$. So the upper and lower envelope curves are easily obtained. The shoot-through zero states happen when the carrier triangle wave is higher than the upper shoot-through envelope V_p or lower than the bottom shoot-through envelope V_n.

Obviously, under a given modulation index M, the distance between these two curves is kept as a constant $\sqrt{3}M$. This indicates that the shoot-through duty ratio is constant and can be expressed as:

$$D_{\text{sh}} = \frac{T_0}{T} = \frac{2 - \sqrt{3}M}{2} = 1 - \frac{\sqrt{3}M}{2} \tag{34.6}$$

where T_0 is the shoot-through time, T is the switching cycle.

The boost factor B and the voltage gain G can be calculated as follows:

$$B = \frac{1 + D_{\text{sh}}}{1 - 3D_{\text{sh}}} = \frac{4 - \sqrt{3}M}{3\sqrt{3}M - 4}, \quad G = MB = \frac{4 - \sqrt{3}M}{3\sqrt{3}M - 4}M \tag{34.7}$$

Through analyses above, the improved maximum constant boost method will be proposed basing on SVPWM and employing advantages from both the traditional maximum constant boost method and the modified reference PWM.

34.3.2 Zero States Time of SVPWM

In sector I, time for active voltage vector is

$$T_1 = \sqrt{3}\frac{V_{\text{ref}}}{V_{\text{dc}}}T_s \sin(\frac{\pi}{3} - \theta), \quad T_2 = \sqrt{3}\frac{V_{\text{ref}}}{V_{\text{dc}}}T_s \sin\theta \quad (0<\theta<\frac{\pi}{3}) \tag{34.8}$$

The null-states time T_N and the minimum null-states time $T_{N\min}$ can be calculated as

$$\begin{aligned} T_N &= T_s - T_1 - T_2 = T_s - \sqrt{3}\frac{V_{\text{ref}}}{V_{\text{dc}}}T_s \sin\left(\frac{\pi}{3} + \theta\right), \\ T_{N\min} &= T_s - \sqrt{3}\frac{V_{\text{ref}}}{V_{\text{dc}}}T_s, \ \left(\theta = \frac{\pi}{6}\right) \end{aligned} \tag{34.9}$$

34.3.3 The Improved Maximum Constant Boost Control

Shoot-through states are inserted in null intervals, so under the prerequisite of maintaining active intervals constant, the maximum shoot-through intervals is $T_{N\min}$, if the chosen shoot-through duty time is just equal to $T_{N\min}$, we will get the same effect as the maximum constant control. The shoot-through time T_0 and ratio D_{sh} is

$$T_0 = (1 - \frac{\sqrt{3}}{2}m)T, \quad D_{\text{sh}} = D_{\max} = 1 - \frac{\sqrt{3}}{2}m \tag{34.10}$$

The boost factor B and the voltage gain G can be calculated as follows

$$B = \frac{1 + D_{\text{sh}}}{1 - 3D_{\text{sh}}} = \frac{4 - \sqrt{3}M}{3\sqrt{3}M - 4}, \quad G = \frac{U_p}{V_0/2} = MB = \frac{4 - \sqrt{3}M}{3\sqrt{3}M - 4}M \tag{34.11}$$

where U_p is the output peak phase voltage, Eq. (34.11) shows that the voltage gain is determined by the modulation index. Obviously the proposed improved modulation strategy can achieve maximum constant effects for Eq. (34.11) is constant with Eq. (34.7). With (34.10), (34.11), we can get

$$\begin{aligned} D_{\text{sh}} &= \frac{3\sqrt{3}G - \sqrt{27G^2 - 8\sqrt{3}G + 16}}{4}, \ M \\ &= \frac{4 - 3\sqrt{3}G + \sqrt{27G^2 - 8\sqrt{3}G + 16}}{2\sqrt{3}} \end{aligned} \tag{34.12}$$

with Eqs. (34.9)–(34.11), the shoot-through ratio and modulation index can be calculated by input voltage V_0 and output peak phase voltage U_p.

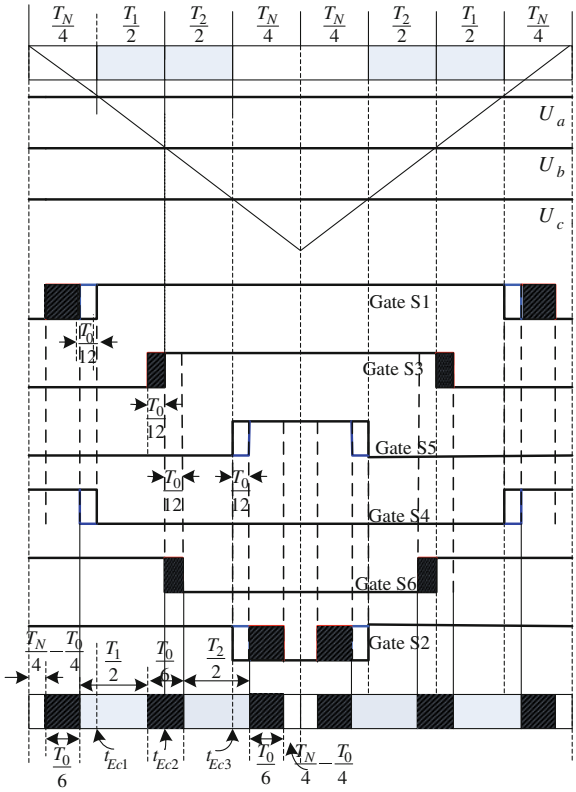


Fig. 34.3 Improved maximum constant control in section I

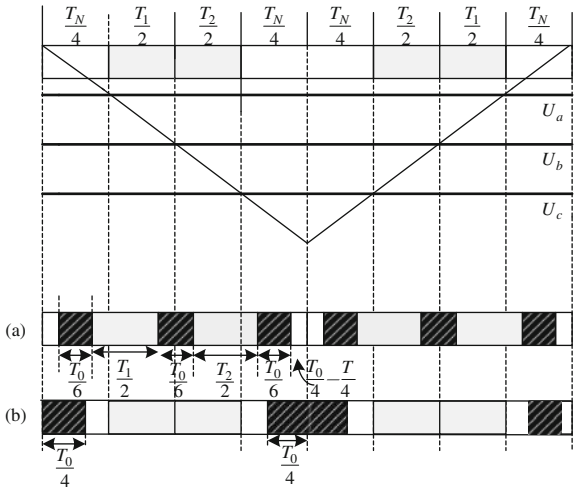


Fig. 34.4 **a** The switching pattern for improved maximum constant control. **b** Sketch map of traditional maximum constant boost control

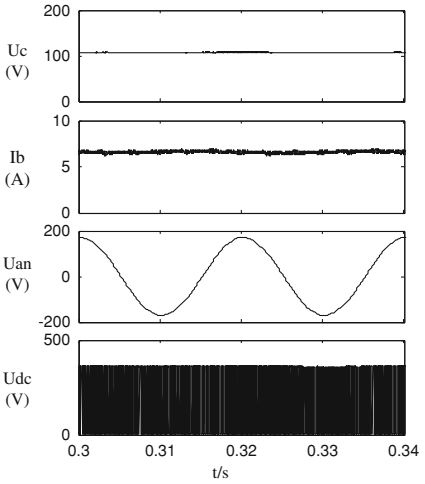


Fig. 34.5 Simulation results with $M = 0.94$ and input voltage 145 Vdc

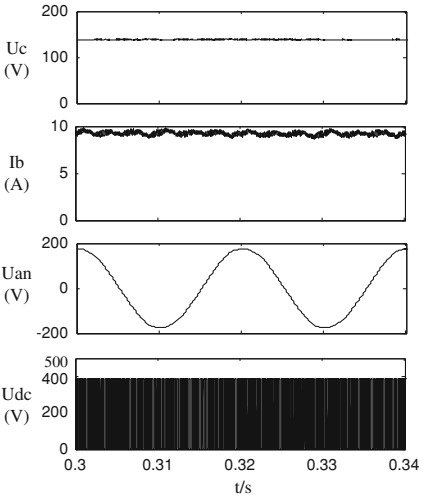


Fig. 34.6 Simulation results with $M = 0.91$ and input voltage 110 Vdc

Figure 34.3 shows the proposed modified SVPWM for ZSI in section I. The first inserting shoot-through state occurs when the upper switch S1 is turned on at $(t_{Ec1} - T_0/4)$, while the shoot-through is ended by the lower switch S4 turned off at $(t_{Ec1} - T_0/12)$. Here, the $T_0/12$ motion is to make sure that the third insertion has equal utilization time as the first one, and then a $T_0/6$ shoot-through time is inserted. The second one starts at the time of the upper switch S3 is turned on at $(t_{Ec2} - T_0/12)$, and ends by S4 turned off at $(t_{Ec2} + T_0/12)$, then a $T_0/6$ shoot-through time is created. And the third one happens when the upper switch S5 turns on at $(t_{Ec3} + T_0/12)$ and ends after S6 turning off at $(t_{Ec3} + T_0/4)$. In this part, the $T_0/12$ motion is to keep the active state as before and the residual part is the shoot-through time.

Table 34.1 Capacitor voltage, theoretical voltage stress and output voltage under different conditions

Operation condition	Capacitor voltage (V)	Voltage stress (V)	Output voltage (V)
M = 0.94, V0 = 145 V	106	350	170
M = 0.91, V0 = 110 V	142	390	170

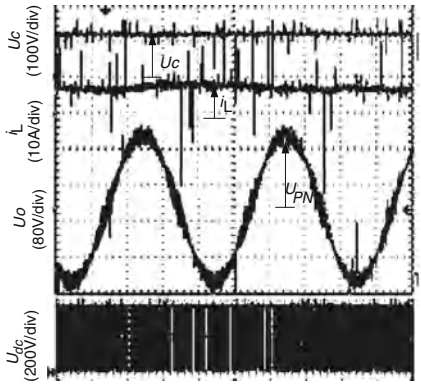


Fig. 34.7 Experimental results with $M = 0.94$ and input voltage 145 Vdc

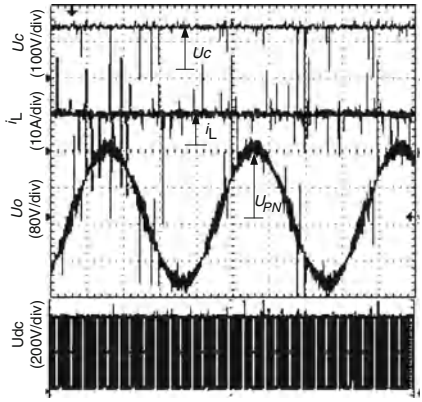


Fig. 34.8 Experimental results with $M = 0.91$ and input voltage 110 Vdc

34.3.4 The Switching Frequency

The switching frequency of the modified constant boost control is only half in comparison with the traditional constant boost control because every switch in Fig. 34.3 switches once per half carrier period, while the switch in Fig. 34.2 switches two times. The switching frequency of this method is same as the modulation of traditional VSI, therefore the Z-source inverter is applicable to circumstances with higher carrier frequency without constrains set by switching frequency limitation.

34.3.5 The Inductor Current Ripple

The inductor current ripple can be expressed

$$\Delta i_L = \frac{D_0 T V_{C1}}{L} = \frac{D_0(1 - D_0)V_{\text{in}}}{(1 - 2D_0)fL} \tag{34.13}$$

where, f is the charging frequency. In Fig. 34.4, there are six shoot-through states in (a) while there are two shoot-through states in (b). Therefore the charging frequency of inductor in (a) is three times higher than that of (b). So from (34.13), we can conclude that the inductor current ripple (a) is much smaller than (b) in the same conditions.

34.4 Simulation Results

To verify the control strategies with reduced voltage stress cross devices, simulations were performed in Matlab/Simulink. The parameters are given:

(1) *Z-source network* $L_1 = L_2 = 1.2$ mH, $C_1 = C_2 = 1000$ μF
(2) *Output filter* $L_2 = 3.5$ mH, $C_f = 10$ μF
(3) *Switch frequency* 7.2 kHz.
(4) *Load* three-phase resistance load $R = 43\ \Omega$.

The purpose of the system is to compare the performance of the Z-source inverter under different input voltage and similar output three-phase voltage of around 208 V rms. The simulation results at input voltage level 110 and 145 V are shown in Figs. 34.5 and 34.6, respectively. The corresponding modulation indexes are $M = 0.91$, $M = 0.94$ respectively.

In the simulation results, V_c is the capacitor voltage, V_{dc} is the dc bus voltage, which is also the voltage stress V_s, and V_{an} is the output voltage after the filter. Based on the analysis above, the theoretical voltage stress and output line to phase peak voltage are listed in Table 34.1. From Figs. 34.5 and 34.6, the capacitor voltage is only about 110 and 150 V, respectively. And the voltage stress in Fig. 34.6 is higher than the one in Fig. 34.5. So the results are quite consistent with the theoretical analysis, which verifies the above analysis and the control concept.

34.5 Experimental Verification

The parameters used in the experiment and simulations are identical. The prototype is also used to verify the design issues discussed above. The experimental results with the same conditions are shown in Figs. 34.7 and 34.8, respectively. The voltage of capacitor is below 200 V and the voltage stress is same to that of

the maximum constant boost control. And the results agree well with simulation results.

34.6 Conclusions

This paper has presented a novel Z-source inverter topology. Comparing with the previous Z-source inverters, the enhanced Z-source has the advantage of high voltage boost inversion ability and low Z-source capacitor voltage stress. A modified maximum constant boost method has also been presented. Compared with the previous maximum constant boost method, the switching frequency is reduced resulting in lower turn off loss in the same voltage boost gain condition. And it can greatly reduce the L and C volume in the Z-network. In the end, simulation and experimental results have verified the merits of the proposed method.

Acknowledgments This work was supported by the Natural Science Foundation of Shandong Province of China (ZR2010EM065). The authors thanks for the financial support.

References

1. Peng FZ (2003) Z-source inverter. IEEE Trans Ind Appl 39(2):504–510
2. Anderson J, Peng FZ (2008) Four quasi-Z-source inverters. 39th IEEE annual power electronics specialists conference, Rhodes, Greece pp 2743–2749
3. Tang Y, Xie SJ, Zhang CH (2009) Improved Z-source inverter with reduced Z-source capacitor voltage stress and soft-start capability. IEEE Trans Power Electron 24(2):409–415
4. Peng FZ, Shen MS, Qian ZM (2005) Maximum boost control of the Z-source inverter. IEEE Trans Power Electron 20(4):833–838
5. Shen MS, Wang J, Joseph A, Peng FZ, Tolber LM, Adams DJ (2006) Constant boost control of the Z-source inverter to minimize current ripple and voltage stress. IEEE Trans Ind Appl 42(3):770–778
6. Jung JW, Keyhani A (2007) Control of a fuel cell based Z-source converter. IEEE Trans Energy Convers 22(2):467–476
7. Liu Jb, Hu JG, Xu LY (2005) A modified space vector PWM for Z-source inverter-modeling and design. Electr Machin Syst ICEMS 2:1242–1247
8. Zimmermann MV, Lechler M, Piepenbreier B (2009) Z-source drive inverter using modified SVPWM for low output voltage and regenerating operation. Power Electr Appl EPE'09:1–10
9. Loh PC, Vilathgamuwa DM, Li YS, Chua GT, Li YW (2005) Pulse-width modulation of Z-source inverters [J]. IEEE Trans Power Electron 20(6):1346–1355

Chapter 35
Study of Restraining DC-Bus Voltage Fluctuation in Dual-PWM Inverter

Haijun Tao, Ben Xu and Xiumei Xue

Abstract In a system of dual-PWM frequency conversion and speed regulation, DC-bus voltage fluctuates as the load disturbance, which makes it impossible for the system to run safely and stably. In order to restrain DC-bus voltage fluctuation, the principle of DC-bus voltage fluctuation was analyzed by small-signal model, and a kind of feedforward control strategy which can improve the dynamic response of DC-bus voltage was pointed out. The simulation results verify the effectiveness of the control strategy and the possibility of reducing capacitance capacity.

Keywords Dual-PWM · DC-bus voltage · Small-signal model · Feed-forward control

35.1 Introduction

Dual-PWM inverter (as shown in Fig. 35.1), on the basis of general inverter, introduces reversible PWM rectifier to replace uncontrolled rectifier. It can actively eliminate electric-network harmonic pollution from the frequency conversion device, improve the power factor of device, realize energy bi-directional flow, save the energy, and make motor working on four quadrants more convenient [1, 2]. Now it is widely used in mechanical, electric power, coal and other industries.

H. Tao (✉) · B. Xu · X. Xue
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: Xuben0108@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_35,

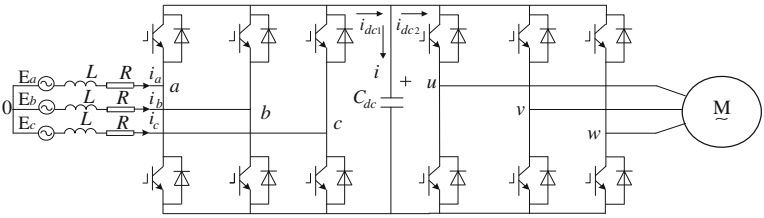


Fig. 35.1 Schematic diagram of dual-PWM inverter

DC-bus voltage will produce sudden change when the motor working states abruptly change, especially when the motor feedbacks energy to the power grid. The energy of the motor feedback is stored on the capacitance of DC-bus if improper control. This kind of situation causes a sharp rise of DC-bus voltage which leads to power switching device breakdown due to over-voltage. In order to ensure the stable operation of the dual-PWM frequency conversion and speed regulation system, it is important to restrain DC-bus voltage fluctuation [3].

This paper adopts an analytical approach of small-signal model, and discusses feedforward control strategy of how to restrain DC-bus voltage fluctuation. Then the conditions to maintain a constant DC-bus voltage is analyzed. Under the condition of a given value of DC-bus voltage, the level of restrain DC-bus voltage fluctuation by the proposed control strategy is given out. With simulation results presented to prove the proposed control strategy it effectively restrains DC-bus voltage fluctuation. Hence, it provides the possibility of cost saving and reduces the capacitance.

35.2 Analysis of DC-Bus Voltage Fluctuation

The rectifier side adopts double closed-loop control strategy with the outer voltage loop as voltage regulator and the inner current loop as current regulator. It is assumed that DC-bus voltage fluctuation is largely influenced by load disturbance. The loss of the power components on and off is ignored.

Therefore, according to the power balance principle and kirchhoff's current law, the equality is described as [4]

$$\begin{cases} i_{\mathrm{dc1}} = C_{\mathrm{dc}} \frac{dv_{\mathrm{dc}}}{dt} + i_{\mathrm{dc2}} \\ 3vi = v_{\mathrm{dc}} i_{\mathrm{dc1}} \\ i = Ki_{\mathrm{ref}} \end{cases} \tag{35.1}$$

where v, i, i_{dc1}, and i_{dc2} denote the effective value of voltage and current, the current from the converter to the dc link and the current from the dc link to the inverter, respectively. The DC-bus voltage and the DC-bus capacitance is denoted

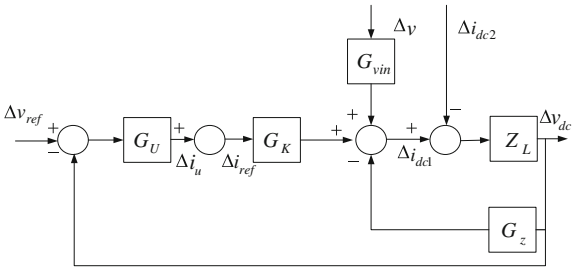


Fig. 35.2 The block diagrams of small-signal control without feedforward

by v_{dc} and C_{dc}, respectively. The i_{ref} represents the reference of inner current loop (the reference of d-axis active component). $K = 1/\sqrt{2}$.

Using small-signal linear analysis method according to the first equality

$$\begin{cases} C_{\mathrm{dc}} \frac{d(V_{\mathrm{dc}}+\Delta v_{\mathrm{dc}})}{dt} = (I_{\mathrm{dc1}} - \Delta i_{\mathrm{dc1}}) - (I_{\mathrm{dc2}} - \Delta i_{\mathrm{dc2}}) \\ 3(V+\Delta v)(I+\Delta i) = (V_{\mathrm{dc}} + \Delta v_{\mathrm{dc}})(I_{\mathrm{dc1}} + \Delta i_{\mathrm{dc1}}) \\ I + \Delta i = K(I_{\mathrm{ref}} + \Delta i_{\mathrm{ref}}) \end{cases} \tag{35.2}$$

Here, the higher orders are ignored ($\Delta v \Delta i = \Delta v_{\mathrm{dc}} \Delta i_{\mathrm{dc1}} = 0$), steady-state equality (35.3) and transient equality (35.4) are given by

$$\begin{cases} 3KVI_{\mathrm{ref}} = V_{\mathrm{dc}} I_{\mathrm{dc1}} \\ I_{\mathrm{dc1}} = I_{\mathrm{dc2}} \end{cases} \tag{35.3}$$

$$\begin{cases} \Delta i_{\mathrm{dc1}} = \frac{3KV}{V_{\mathrm{dc}}} \Delta i_{\mathrm{ref}} + \frac{3KI_{\mathrm{ref}}}{V_{\mathrm{dc}}} \Delta v - \frac{I_{\mathrm{dc1}}}{V_{\mathrm{dc}}} \Delta v_{\mathrm{dc}} \\ C_{\mathrm{dc}} \frac{d\Delta v_{\mathrm{dc}}}{dt} = \Delta i_{\mathrm{dc1}} - \Delta i_{\mathrm{dc2}} \end{cases} \tag{35.4}$$

Figure 35.2 shows the block diagrams of small-signal control without feedforward. Besides,

$$G_K = \frac{3KV}{V_{\mathrm{dc}}},\ Z_L(S) = \frac{1}{SC_{\mathrm{dc}}},\ G_z = \frac{I_{\mathrm{dc1}}}{V_{\mathrm{dc}}},\ G_{\mathrm{vin}} = \frac{3KI_{\mathrm{ref}}}{V_{\mathrm{dc}}},\ G_U = K_p + \frac{K_i}{S}$$

By Fig. 35.2 the expression of the DC-bus voltage fluctuation can be deduced.

$$\Delta v_{\mathrm{dc}} = T_{\mathrm{ref}} \Delta v_{\mathrm{ref}} + T_v \Delta v + Z_0 \Delta i_{\mathrm{dc2}} \tag{35.5}$$

Besides

$$T_{\mathrm{ref}} = \frac{\Delta v_{\mathrm{dc}}}{\Delta v_{\mathrm{ref}}} = \frac{G_K Z_L G_U}{\Delta},\ T_v = \frac{\Delta v_{\mathrm{dc}}}{\Delta v} = \frac{Z_L G_{vin}}{\Delta},\ Z_0 = \frac{\Delta v_{\mathrm{dc}}}{\Delta i_{\mathrm{dc2}}} = \frac{-Z_L}{\Delta},$$
$$\Delta = 1 + Z_L(G_K G_U + G_z)$$

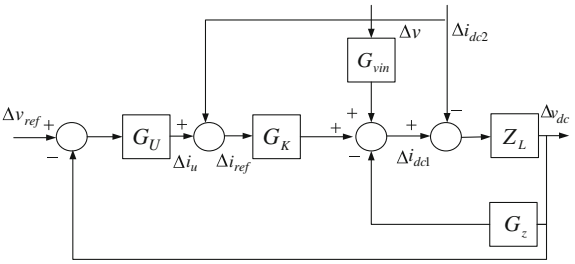


Fig. 35.3 The block diagrams of small-signal control with feedforward

Therefore, DC-bus voltage fluctuation is influenced by load disturbance and power grid fluctuation due to the existence of the T_v and Z_0. However, if T_v and Z_0 equal to zero, theoretically guarantee that DC-bus voltage fluctuation is not influenced by load disturbance and power grid fluctuation. Feedforward control strategy is designed according to this principle.

35.3 Feedforward Control Strategy

The above analysis shows that DC-bus voltage fluctuations are directly related with load current changes. Hence, load current is an external disturbance signal for the whole control system. If load current changes, the DC-bus voltage is first affected and deviated from the set value which lead to increase in deviation of the set value and DC-bus voltage. The adjust action of voltage regulator can reduce gradually and even eliminate the deviation until the system enters steady state. But, load current increases suddenly in a period of time and rectifier still cannot provide load for energy consumption because the speed adjustment of voltage regulation ring is slow. At this time, capacitance will release its store energy, and provide the energy to load with the help of rectifier. If load current decreases suddenly, the energy rectifier provides beyond the energy load consumption. The redundant energy flows to the DC-bus capacitance and takes charge of it. Thanks to the imbalance of input and output energy in the dynamic process, DC-bus voltage fluctuates. According to the control theory knowledge, the feedforward control can eliminate the influence of the load disturbance on the system, which can improve the dynamic performance of the system [5].

Feedforward control strategy puts load current changes of inverter side directly reflecting the reference of inner current loop of rectifier side, which can reduce the transient current change of capacitance and maintain DC-bus voltage constant.

Figure 35.3 shows the block diagrams of small-signal control with feedforward.

By Fig. 35.3 the expression of the DC-bus voltage fluctuation can be deduced.

$$\Delta v_{\mathrm{dc}} = T_{\mathrm{ref}} \Delta v_{\mathrm{ref}} + T_v \Delta v + Z_0{}' \Delta i_{\mathrm{dc2}} \tag{35.6}$$

Besides,

$Z_0' = \frac{\Delta v_{\mathrm{dc}}}{\Delta i_{\mathrm{dc2}}} = \frac{-(1-G_K)}{\Delta} Z_L = (1 - G_K) Z_0.$

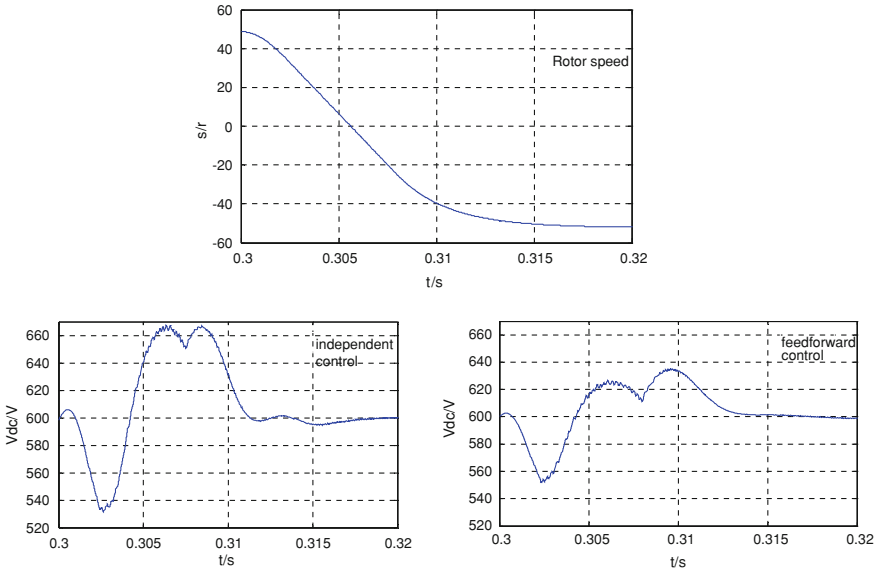


Fig. 35.4 Comparison of DC-bus voltage fluctuation when speed changes

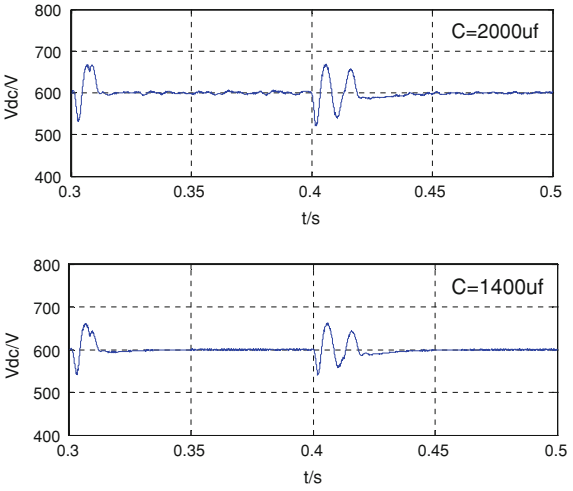


Fig. 35.5 Comparison of the control effect with different capacitance

If the feedforward control strategy can make $(1 - G_K) = 0$, DC-bus voltage will not fluctuate. So, make the $(1 - G_K) = 0$, the equality is given by

$$1 - G_K = 1 - \frac{3KV}{V_{\text{dc}}} = 1 - \frac{3V}{\sqrt{2}V_{\text{dc}}} = 0$$

where $V_{dc} = 3V/\sqrt{2}$. This equality $(1 - G_K) = 0$ is established. However, in order to guarantee that the system works in linear modulation zone, the DC-bus voltage $V_{dc} \geq \sqrt{6}V$. Obviously, $(1 - G_K) = 0$ cannot be established. Therefore the feedforward control strategy cannot completely eliminate the impact of load change.

$$G_K = \frac{3V}{\sqrt{2}V_{dc}} \leq \frac{3V}{\sqrt{2} \times \sqrt{6}V} = \frac{\sqrt{3}}{2} \Rightarrow 1 - \frac{\sqrt{3}}{2} \leq 1 - G_K < 1$$

That is, the feedforward control strategy can restrain DC-bus voltage fluctuation due to load changes in a certain extent. It seems that, $V_{dc} = \sqrt{6}V\ (1 - G_K)$ will be equal to 0.13. It means that the influence of load changes on DC-bus voltage can be reduced to 13%.

35.4 Simulation Results

In order to assess the feasibility of proposed control strategy to suppress the DC-bus voltage fluctuation, the simulation model of dual-PWM frequency conversion and speed regulation system is built by Matlab software and the simulation results are given out. The rectifier side adopts space voltage vector control strategy and the inverter side adopts rotor field-oriented vector control strategy.

The parameters of simulation are shown as follows:

Line voltage: 380 V Inlet inductance: 1.1 mH DC-bus capacitance: 2000 μF DC-bus voltage: 600 V

Figure 35.4 shows the simulation results when the motor speed changes from +50 rad/s to −50 rad/s at 0.3 s.

The second simulation adopts independent control. The maximum wave amplitude of DC-bus voltage fluctuation when speed changes is 68 V. However, when the third simulation adopts feedforward control, the maximum wave amplitude is suppressed for 26 V, only 39% of the former. The recovery time of former is 0.02s. However, the recovery time of latter is 0.014s, only 70% of the former.

Figure 35.5 shows the comparison of DC-bus voltage fluctuation in the condition of independent control with capacitance 2000 μF and feedforward control with capacitance 1400 μF.

Ensuring that the conditions are of almost the same control effect, the feedforward control strategy can largely reduce the DC-bus capacitance.

35.5 Conclusion

The principle of DC-bus voltage fluctuation was analyzed using the technology of small-signal model, and a kind of feedforward control strategy which can improve the dynamic response of DC-bus voltage was pointed out. The simulation results

show that the feedforward control can effectively restrain the DC-bus voltage fluctuation compared with independent control, and reduce the DC-bus capacitance capacity.

References

1. Zheng Z, Zhao H (2007) Research on dual-PWM inverter based on power feedback integrared control strategy. Control Eng 7(14):182–184
2. Chen B, Chen M (2005) Ac speed regulation system. Machinery Industry Press, Beijing
3. Gu B, Kwanghee AN (2002) DC link capacitor minimization method through direct capacitor current control. IEEE Trans Ind Appl 42(2):573–574
4. Chongwei Zhang, Xing Zhang (2003) PWM rectifier and its control. Machinery Industry Press, Beijing
5. Gu B, Kwanghee N (2005) Theoretical minimum DC-link capacitance in PWM converter–inverter systems. IEEE Proc Electr Power Appl 152(1):81–83

Chapter 36
PCSS-Based Nanosecond High Voltage Pulse Generator for Biological and Biomedical Application

Yafang Tan, Hongchun Yang, Jun Xu and Gang Zeng

Abstract Benefit from photoconductive semiconductor switch (PCSS)'s fast response, high power, high stability and long lifetime properties, design and operation of high voltage nanosecond pulse generator based on PCSS for biological and medical application is presented. The generator comprises pulse power source, photoelectric synchronization control apparatus, PCSS based on Blumlein transmission line and laser diode circuit. It produces 1 kHz repetition frequency rate (RFR), 2-ns-wide, 8 kV amplitude pulses, which can be used in cancer treatment.

Keywords Photoconductive semiconductor switch · Blumlein transmission line · Pulse generator

36.1 Introduction

Biological experiments indicated that PEF with appropriate parameters, such as pulse duration, electric strength and RFR, etc., will induce reversible or irreversible nanoscale electroporation in the cell membrane, organelle membrane

Y. Tan (✉) · H. Yang · J. Xu · G. Zeng
School of Physical Electronics, University of Electronic Science and Technology of China, Chengdu 610054, Sichuan, China
e-mail: Yafang-tan@163.com

H. Yang
e-mail: Hongchun.Yang@163.com

J. Xu
e-mail: Jun.Xu@163.com

G. Zeng
e-mail: Gang.Zeng@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_36,

or nuclear envelope, induce endoplasmic reticulum Ca^{2+} imbalance, cytochrome C release, caspase activation, cell volume shrinkage and a series of bio-medical phenomena [1–4]. These phenomena will induce apoptosis or necrosis of cells. Concretely, the cell membrane electroporation occurs while the inner membranes do not lose the integrity (nuclear envelope and mitochondrial membrane are not perforated) under lower electric field intensity and longer pulse duration (typical parameters: 10^2–10^3 ns, 1 kV/cm). This effect is called the extra cellular processing (ECP) [5, 6]. On the hand, the inner membranes electroporation occurs while the cell membrane does not loss the integrity under higher electric field intensity and shorter pulse duration (typical parameters: 10^{-1}–10^1 ns, 10 kV/cm). This effect is called intracellular electromanipulation (IEM) [7–9]. The ECP potentates the cytotoxicity of chemotherapeutic drugs by permeabilizing the cell membrane at the site of electric pulses application, allowing for increasing transmembrane transport of non-permeant or poorly permeant anticancer drugs. It has been used in cancer treatment [10, 11]. IEM will induce apoptosis of cells [12, 13] and inhibits the tumor propagation channels. Moreover, IEM has no chemotherapy damage to the health of patients, or adverse reactions. So it has been considered as a potential treatment means of cancers and is becoming a hot topic [14]. Therefore, design and operation of high-voltage nanosecond pulse generator used in cancer treatment is very important.

Existing high-voltage pulse generators are based on spark gap switched transmission lines [15], or radiofrequency MOSFET switched capacitors [16]. The drawbacks of the pulse generators based on spark gap include large size, low RFR, short lifetime, and erratic, high jitter triggers [15]. The MOSFET switched capacitor can't generate faster or narrower pulses than 15–20 ns due to complications of the MOSFET driving circuit and inherent limitations of the MOSFET device [16].

Design and operation of high-voltage nanosecond electric pulse generator based on PCSS is presented. To benefit from PCSS's fast response, high power, high stability and long lifetime properties, the generator gets over the disadvantages of generators based on spark gap switched transmission lines and radiofrequency MOSFET switched capacitors, and produces RFR 1 kHz, duration 2 ns, amplitude 8 kV pulses with sharp rising and falling edges.

36.1.1 Design and Operation

The pulse generator consists of several parts: pulse power source, photoelectric synchronization control apparatus, PCSS based on Blumlein transmission line and laser diode circuit (Fig. 36.1)

The voltage regulator supplies continuous variable voltage to pulse power source, which produces high-voltage *us* Gaussian pulses triggered by the photoelectric synchronization control apparatus. The *us* pulses furnish PCSS with electricity by energy storage Blumlein transmission line. The laser diode produces

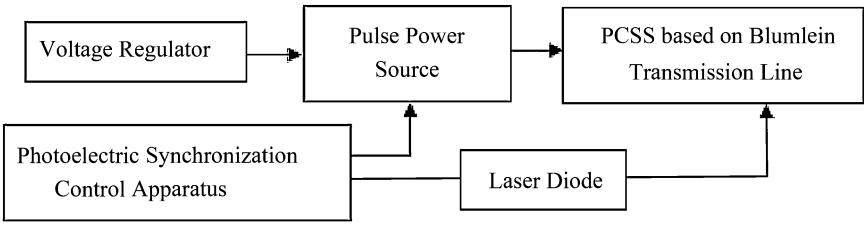


Fig. 36.1 The diagram of high-voltage nanosecond pulse generator

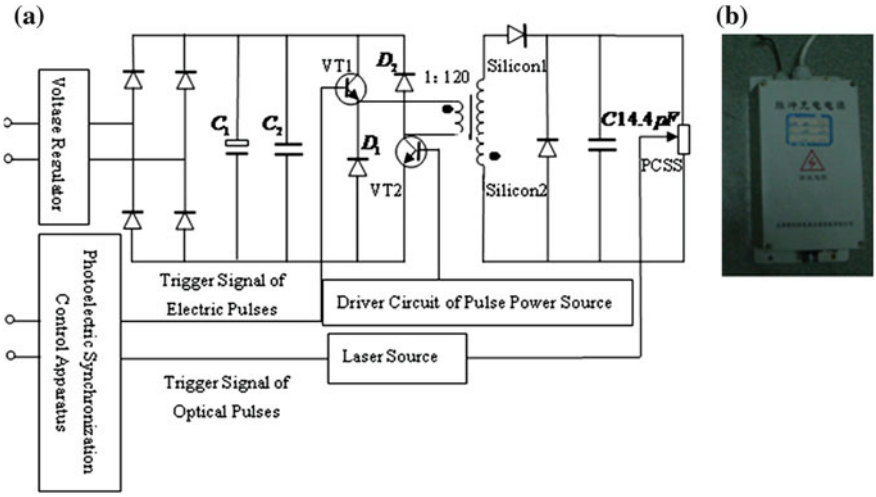


Fig. 36.2 The circuit diagram and photograph of the pulse power source. **a** The circuit diagram of the pulse power source. **b** The Photograph of the pulse power source

optical pulses triggered by photoelectric synchronization control apparatus, which turn on the PCSS under high bias voltage. The PCSS releases the energy stored in Blumlein line rapidly, and produces high-voltage *ns* pulse.

36.1.2 Design of Pulse Power Source

Main considerations in design for the pulse power source are output energy and waveform parameters of external trigger signal. The circuit diagram and photograph are shown in Fig. 36.2.

The design of pulse power source is based on flyback transformer. The work process is as follows: The DC U_i is got by boosting 220 V/50 Hz AC voltages with transformer and then converting voltage with rectifier filter. The primary voltage of the transformer is approximately U_i, and no current flows in the sub-side for the

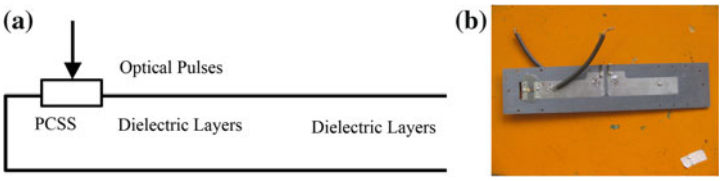


Fig. 36.3 The plannar structure and photograph of the Blumlein line and PCSS integrated design. **a** The plannar structure of integrated design. **b** The photograph of integrated design

silicon 1's reverse bias cut-off, when the switches VT1 and VT2 turn on. Primary coil of the transformer stores the magnetic energy converted from the electric energy. The voltage of the secondary coil is reversed, and silicon 2 turns on, when the switches VT1 and VT2 turn off. The secondary coil gets the energy stored by primary coil, and releases energy to load by silicon 1.

C is the capacitance of Blumlein transmission line as an energy storage capacitor, which is determined by the rule: the energy stored by the capacitor is greater than the energy produced by PCSS. L_2, the inductance of secondary coil, is determined by the charging time of LC circuit and the technological level of flyback transformers. L_1, the inductance of primary coil, is determined by its energy storage, change proportion to the secondary coil and the peak current it can endure. C_1 and C_2 are approximately dozens of pF and nF, respectively. Silicon 1, 2 are several diodes in series.

36.1.3 Blumlein Transmission Line and PCSS Integrated Design

The planar structure is developed in the Blumlein line and PCSS integrated design, as shown in Fig. 36.3a. The photograph is show in Fig. 36.3b.

C is obtained in design for pulse power source. According to the formula of parallel plate capacitors, $C = \varepsilon S/d$, thickness of dielectric substrate and choice of dielectric constant, the area S for energy storage miscrostrip in Blumlein line, is determined. With the formula, $l = \frac{Tc}{2\sqrt{\varepsilon_r}}$, the Blumlein line length l is determined, where T is output pulse duration, c is velocity of light, ε_r is dielectric constant of dielectric layers. The formulas for calculating the Blumlein line width w are presented as follows:

$$\begin{cases} Z_0 = \dfrac{377}{\sqrt{\varepsilon_r}}\left\{\dfrac{w}{h} + 0.883 + 0.165\dfrac{\varepsilon_r - 1}{\varepsilon_r^2} + \dfrac{\varepsilon_r + 1}{\pi\varepsilon_r}\left[\ln\left(\dfrac{w}{h} + 1.88\right) + 0.758\right]\right\}^{-1} & w/h > 1 \\ Z_0 = \dfrac{120}{\sqrt{2(\varepsilon_r + 1)}}\left[\ln\dfrac{8h}{w} + \dfrac{1}{32}\left(\dfrac{w}{h}\right)^2 - \dfrac{\varepsilon_r - 1}{\varepsilon_r + 1}\left(0.2258 + \dfrac{0.1208}{\varepsilon_r}\right)\right] & w/h \le 1 \end{cases} \tag{36.1}$$

Z_0 is input impedance of Blumlein transmission line.

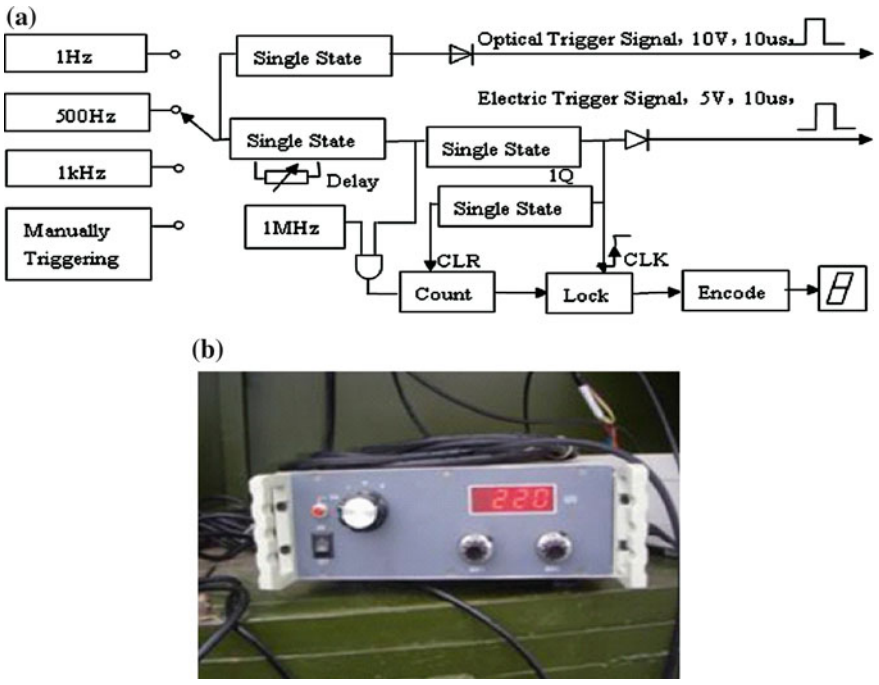


Fig. 36.4 The principle diagram and photograph of the photoelectric synchronization control apparatus. **a** The principle diagram of the photoelectric synchronization control apparatus. **b** The photograph of the photoelectric synchronization control apparatus

36.1.4 Photoelectric Synchronization Control Apparatus

The time of the optical and electric pulses exciting PCSS must be adjusted accurately. In this way the optical pulse could timely excite PCSS, when the electric pulse reaches the peak level. The time adjustment of the optical and electric pulses is done by photoelectric synchronization control apparatus, which ensures PCSS being triggered under high bias voltage. The principle diagram of the photoelectric synchronization control apparatus is shown in Fig. 36.4a.

The crucial issue in the design is time delay between the optical and electric pulses, and it can be determined by experiments.

36.2 Output Pulses

The generator produces 2-ns-wide, 8 kV amplitude pulses. The output pulse is shown in Fig. 36.5. The pulse can be used in cancer treatment.

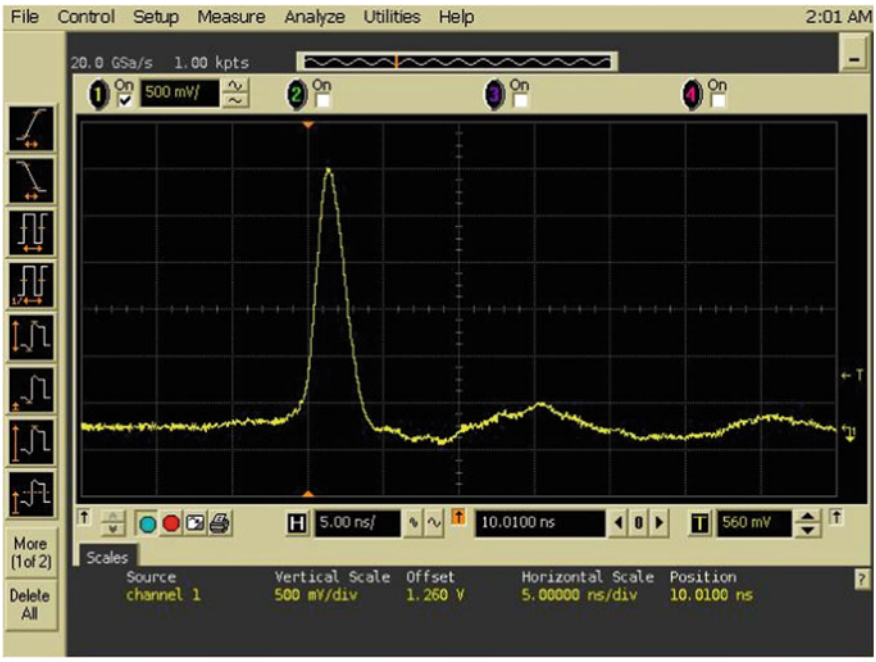


Fig. 36.5 The output pulse

36.3 Conclusion

To benefit from PCSS's fast response, high power, high stability and long lifetime properties, design and operation of high-voltage nanosecond electric pulse generator based on it is presented. The generator can produce RFR 1 kHz, duration 2 ns, amplitude 8 kV pulses with sharp rising and falling edges, which are applied in the treatment of cancers.

References

1. Craviso GL, Chatterjee P, Maalouf G, Cerjanic A, Yoon J, Chatterjee I et al (2009) Nanosecond electric pulse-induced increase in intracellular calcium in adrenal chromaffin cells triggers calcium-dependent catecholamine release. IEEE T Dielect El In 16:1294–1301
2. Mi Y, Yao CG, Li CX, Sun CX, Tang JY, Yang FL et al (2010) Caspase-3 activation by exponential decay nanosecond pulsed electric fields on tumor-bearing BALB/c nude mice in vivo. IEEE T Plasma Sci 38:1963–1971
3. Gregory BW, Osgood CJ, Schoenbach KH, Kolb JF (2010) Membrane permeability and cell survival after nanosecond pulsed-electric-field exposure–significance of exposure-media composition. IEEE T Plasma Sci 38:2948–2953

4. Beebe SJ, Fox PM, Rec LJ, Somers K, Stark RH, Schoenbach KH (2002) Nanosecond pulsed electric field(nsPEF) effects on cells and tissues: apoptosis induction and tumor growth inhibition. IEEE T Plasma Sci 30:286–292
5. Weaver JC, Chizmadzhev YA (1996) Theory of electroporation: a review. Bioelectrochem Bioenerg 41:135–160
6. Tien HT, Ottova A (2003) The bilayer lipid membrane (BLM) under electrical fields. IEEE T Dielect El In 10:717–723
7. Schoenbaeh KH, Beebe SJ, Buescher ES (2001) Intracellular effect of ultrashort electrical pulses. Bioelectromagnetics 22:440–448
8. Schoenbaeh KH, Joshi RP, Kolb JF, Chen NY, Stacey M, Blackmore PF et al (2004) Ultrashort electrical pulses open a new gateway into biological cells. P IEEE 92:1122–1137
9. Chen NY, Schoenbach KH, Kolb JF, Swanson RJ, Garner AL, Yang J et al (2004) Leukemic cell intracellular responses to nanosecond electric fields. Biochem Bioph Res Co 317: 421–427
10. Sersa G (2006) The state-of-the-art of electrochemotherapy before the ESOPE study: advantages and clinical uses. Eur J Cancer 4:52–59
11. Byrne CM, Thompson JF (2006) Role of electrochemotherapy in the treatment of metastatic melanoma and other metastatic and primary skin tumors. Exp Rev Anticanc 5:671–678
12. Liu X, Kim CN, Yang J, Jemmerson R, Wang X (1996) Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome C. Cell 86:147–157
13. Latta M, Künstle G, Leist M, Wendel A (1999) Metabolic depletion of ATP by fructose inversely controls CD95- and tumor necrosis factor receptor1-mediated hepatic apoptosis. J Cell Bio 191:1975–1986
14. Nuccitelli R, Pliquett U, Chen XH, Ford W, Swanson J, Beebe SJ et al (2006) Nanosecond pulsed electric fields cause melanomas to self-destruct. Biochem Bioph Res Co 343:351–360
15. Behrend M, Kuthi A, Gu X, Vernier PT, Marcu L, Craft CM, Gundersen MA (2003) Pulse generators for pulsed electric field exposure of biological cells and tissues. IEEE T Dielect El In 10:820–825
16. Behrend M, Kuthi A, Vernier PT, Marcu L, Craft C, Gundersen M (2002) Micropulser for real time microscopy of cell electroperturbation. In: Proceedings of 25th International IEEE Power Modulator Symposium pp 358–361

Chapter 37
Predictive Current Deadbeat Control for Permanent Magnet Synchronous Machine Based on Space Vector Moduation

Ke Zhu and Xuanfeng Shangguan

Abstract In this paper, a kind of predictive current deadbeat control for permanent magnet synchronous motor is proposed, which is deduced from the PI control strategy and conventional rotor field-oriented vector control. The expected reference voltage veclor is calculated from the current error and the rotating voltage and realized by space vector modulation. This new method realizes the dynamic decoupling effect of *dq* axis current and improves the transient performance by eliminating time delay caused by the current loop. The predictive current value is equal to the reference current value and in this way the expected voltage vector can be predicted one sample period ahead, which eliminates the time delay caused by calculation. The experimental results prove the validity of this method.

Keywords Deadbeat control · Space vector modulation · Predictive current · Decoupling control

37.1 Introduction

Rotor field-oriented vector control strategy is a commonly used control method, and its basic principle is aimed to realize the decoupling of field current and torque current in the rotor reference frame, through the coordinate transformation. The traditional PI regulation method just calculates the reference voltage vector

K. Zhu (✉) · X. Shangguan
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zhuke111@163.com

X. Shangguan
e-mail: gxf@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_37,

directly using d, q axis current error, does not completely realize the current decoupling control and has a bad torque response and a time delay. This paper designed a kind of predictive current deadbeat controller based on dynamic voltage model of permanent magnet synchronous motor, and the simulation results in Matlab/simulink verify the correctness and excellent static dynamic performances of this method.

37.2 Voltage Equations of PMSM

The dynamic $d-q$ axis model for PMSM is used for rotor field-oriented vector control. It is done by converting the three phase voltages and currents variables into $d-q-0$ variables in rotor reference frame by using Parks transformation. Then the following voltage equations of PMSM are obtained [1].

$$\begin{cases} u_d = i_d R_s + L_d p i_d - \omega L_q i_q \\ u_q = i_q R_s + L_q p i_q + \omega L_d i_d + \omega \psi_f \end{cases} \tag{37.1}$$

where $u_d, u_q, i_d, i_q, L_d, L_q$ are voltages, real-time sampling current and winding inductance of d, q axis respectively in $d-q$ reference frame, ω is the rotor electrical angular speed. Since the rotor has no winding but only magnets and the rotor magnetic flux is along d axis, this flux can be modeled as ψ_f. p is named differential operator.

The conventional field-oriented vector control just realizes the decoupling of field current and torque current in synchronous rotating $d-q$ frame. From Eq (37.1), we know d axis voltage vector is not only determined by i_d but also by i_q, and with the increase of ω, the influence of rotating voltage section is change largely. So, conventional PI control method does not realize the dynamic decoupling.

37.3 The Principle and Simulation Model for Predictive Current Deadbeat Control

Equation (37.1) equivalent to:

$$\begin{cases} p i_d = \dfrac{i_d^* - i_d}{\Delta T} = \dfrac{R_s}{L_d} i_d - \dfrac{\omega L_q}{L_d} i_q - \dfrac{1}{L_d} u_d \\ p i_q = \dfrac{i_q^* - i_q}{\Delta T} = \dfrac{R_s}{L_q} i_q + \dfrac{\omega L_d}{L_q} i_d - \dfrac{1}{L_q} u_q + \dfrac{\omega \psi_f}{L_q} \end{cases} \tag{37.2}$$

Here, i_d^*, i_q^* are predictive current, ΔT is one sample period. We can name e current error, then:

$$e = (e_d, e_q)^T = (i_d^* - i_d, i_q^* - i_q)^T$$

In order to obtain suitable voltage vector and make e close to zero, we can structure lyapunov function:

$$V(e) = \frac{1}{2} e^T e > 0 \tag{37.3}$$

The derivation of Eq. (37.3) is:

$$(V(e))' = \frac{1}{2}((e^T)'e + e^T(e)') \tag{37.4}$$

While e is close to zero, system is asymptotically stable, then, $(V(e))' = \frac{1}{2}((e^T)'e + e^T(e)') < 0$, we can get Eq. (37.5),

$$\begin{cases} u_d = i_d R_s + L_d p i_d - \omega L_q i_q + K L_d e_d \\ u_q = i_q R_s + L_q p i_q + \omega L_d i_d + \omega \psi_f + K L_q e_q \end{cases} \tag{37.5}$$

Here $k > 0$, Merger Eqs. (37.1) and (37.5), we can obtain:

$$(e)' = -Ke. \tag{37.6}$$

Merger Eqs. (37.6) and (37.7), we can obtain:

$$(V(e))' = -\frac{1}{2}(e^T e + e^T e) = -Ke^T e < 0 \tag{37.7}$$

Through the above reasoning results, we know that the control system based on Eq. (37.2) is stable. Deadbeat control is aimed to make the error of predictive current and real-time sampling current close to zero [2]. So, we can think that i_d is equal to i_d^* axis in one sampling period, the same as q axis and ignore the influence of the resistance pressure drop, Eq. (37.5) can be simplified as:

$$\begin{cases} u_d = -\omega L_q i_q^* + K L_d e_d \\ u_q = \omega L_d i_d^* + \omega \psi_f + K L_q e_q \end{cases} \tag{37.8}$$

From Eq. (37.8), we know rotating voltage is directly related to the predictive current, realizing the compensation to reference voltage vector, and it also get the current dynamic decoupling control effect. This compensation is obtained from feedforward current but not feedback sampling current and aimed to make the error of predictive current and real-time sampling current close to zero, so we call it predictive current feedforward compensation deadbeat control. Figure 37.1 is the simulation model of this method [3].

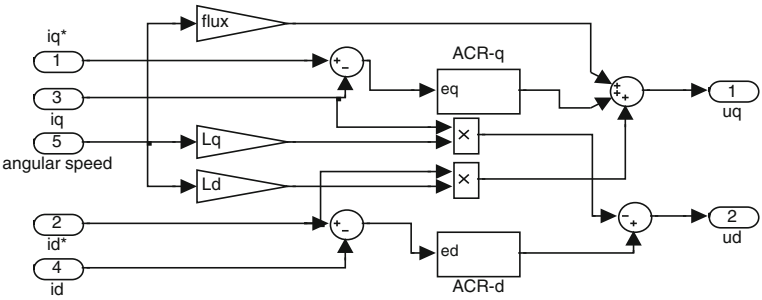


Fig. 37.1 Simulation model of predictive current deadbeat control

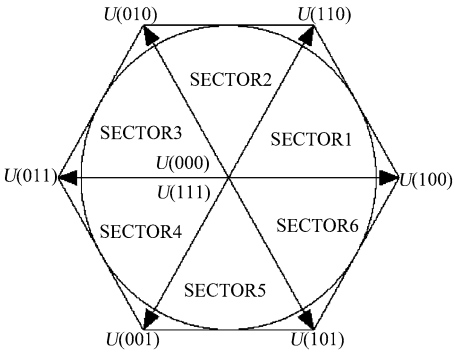


Fig. 37.2 Voltage space vector diagram

37.4 The Principle and Simulation Model for Space Vector Pulse Width Modulation (SVPWM)

The Space Vector Pulse Width Modulation (SVPWM) technique proposed in the recent years, refers to a special way to determine switching sequence of the three-phase voltage source inverters. It uses basic space vectors to generate the output voltages to the motor. The different on and off states of the inverters compose eight possible vectors. The eight vectors are called the Basic Space Vectors. They are denoted by u1(001), u2(010), u3(011), u4(100), u5(101), u6(110), u0(000) and u7(111).

There are six non-zero vectors and two zero vectors. The six non-zero vectors form a hexagon. The angle between any two adjacent non-zero vectors is 60 degrees [4]. The two zero vectors are at the original point. In the basic space vectors diagram, as shown Fig. 37.2, the rotating reference output voltage, denoted Uout, can be made through a combination of any two adjacent space vectors for any short amount of time. The combination can be used to instantaneously represent the reference voltage as long as this period of time is much smaller than the reference voltage rotating speed.

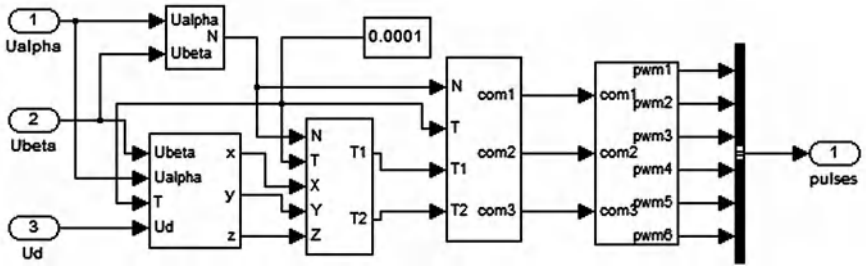


Fig. 37.3 Simulation model of SVPWM

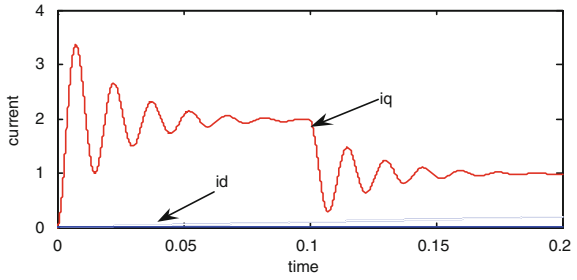


Fig. 37.4 Current waveform of traditional PI adjuster

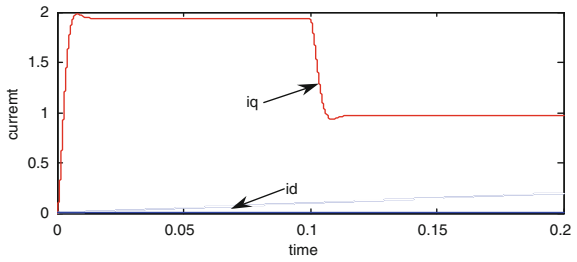


Fig. 37.5 Current waveform of deadbeat control

In order to realize the SVPWM, we need to solve with such problems: the judgment of the sector, calculation and allocation of working voltage vector effect time and determine the vector switch points, etc. If we know the voltage vector u_α, u_β in the static coordinate system corresponding to reference voltage sector u_d, u_q, and PWM period T (PWM cycle for half), then we can obtain the simulation model as shown in Fig. 37.3 based on the above principle of SVPWM [5].

37.5 Simulation Results and Discussion

The simulation of the proposed control method has been carried out using Matlab/simulink software, and the simulation results are shown in Fig. 37.4. For a comparison, a conventional PI controller was designed and simulated, and the results are shown in Fig. 37.5. From the results, we know it just needs 0.01 unit of time in the deadbeat control, and 0.1 unit of time in the conventional PI method. Now we verify the effective of predictive current deadbeat control.

References

1. Yongdong L (2005) Ac motor digital control system. Mechanical industry press, Beijing
2. Lee J (2002) A dead-beat type digital for the direct torque control of an induction motor. IEEE Trans Power Electron 17(5):739–746
3. Naigang H (2010) Power electronics, modeling and simulation of electrical control system. Mechanical industry press, Beijing, p 1
4. Chongjian L (2006) Ac synchronous motor speed control system. Science press, Beijing, pp 266–268
5. Ruan Y, Chen B (2009) Control systems of electric drives: motion control systems. Mechanical industry press, Beijing, p 8

Chapter 38
Design for Charge–Discharge System of Battery Based on the Three-Phase PWM Rectifier

Zheng Zheng and Wenbin Zhou

Abstract To solve the problem of low power factor and harmonic pollution in traditional charge–discharge device, a novel configuration which consists of a three-phase voltage source PWM rectifier and Bi-directional DC–DC converter was proposed. It charges battery in rectifier mode and allows energy which is discharged by battery to feedback to power grid. A charge–discharge simulation model based on constant charge/discharge is designed. Simulation results show that the system can achieve sinusoidal input currents and reduce harmonic pollution, and at the same time achieve discharge process of energy to feedback to power gird and save energy.

Keywords Battery · PWM rectifier · Charge and discharge

38.1 Introduction

Battery, as an energy storage device, which is widespread and used by various departments of the national economy, is an indispensable part of the social production and management. In recent years, environmental protection and energy saving have become the focus of attention, electric vehicle (EV) is becoming more and more concerned by people [1–3], However, the battery acts as power source

Z. Zheng (✉) · W. Zhou
Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zhengzh@hpu.edu.cn

W. Zhou
e-mail: hwbin@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_38,

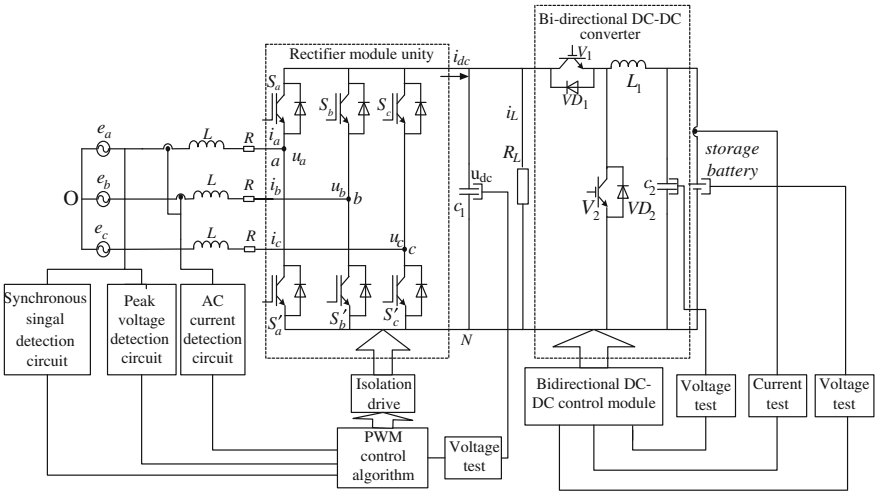


Fig. 38.1 Block diagram of charge–discharge device

for EV, it directly influences and restricts the development of the industry. Charge–discharge performance of the system directly affects the battery technical status, service life, and determines the discharge of the pollution to the power grid.

At present, the charge–discharge device widely used is uncontrollable diode rectifier and silicon controlled rectifier (SCR), the common problems are that both input power factor is low and the power grid is heavily polluted under battery charging control. On battery discharging, because the devices only allow energy to flow in one direction, charge–discharge device usually use resistance as load, which cause electric energy is transformed into heat and wasted. Owing to VSR's advanced features including bi-directional power flow, sinusoidal input current at unit power factor and controllable dc-link voltage etc., VSR technology is applied to charge–discharge system and there are great industrial application prospects.

According to the battery charge/discharge requirements, charge–discharge device adopt VSR, and bi-directional DC–DC converter structure, Simulation results show that the system can achieve sinusoidal input currents and reduce harmonic pollution, and at the same time achieve discharge process of energy to feedback to power gird and save energy.

38.2 The Overall Scheme Design of Charge–Discharge Device

Block diagram of charge–discharge device is showed in Fig. 38.1; the main circuit of charge–discharge device consists of four parts [4, 5]: the three-phase power supply, rectifier module unity, Bi-directional DC–DC converter and storage battery. In order to improve the dynamic performance and reduce the harmonic

content, the device's main circuit adopts voltage type rectifier. C_1 denotes the dc-link capacitance which inhibits dc-link harmonic voltage and stabilizes dc voltage, the role of VSR is to provide a stable DC voltage for post-stage DC–DC converter and achieve bi-directional transmission of energy. Therefore, VSR adopts double closed-loop that is an outer dc-link voltage control loop and an inner AC input current control loop. The dc-ink voltage outer loop is designed for stabilizing dc–link voltage and the inner current control loop tracks current command which is provided by dc-link voltage outer loop. The current control loop is to realize unity power factor sinusoidal current control.

Bi-directional DC–DC converter is to achieve charge or discharge to the battery. C_2 is to store energy. When DC–DC converter operates at buck mode, switch V_1 is closed and switch V_2 is opened, battery is recharged via the charge–discharge device, electric energy transfers from the power grid to the battery. When DC–DC converter operates at boost mode, switch V_1 is opened and switch V_2 is closed, the low-voltage side battery will discharge and provide power for dc-link. The energy of battery is transmitted to the power grid through the PWM rectifier and the purpose of saving energy is achieved.

38.3 Control of Charge–Discharge System

38.3.1 Vector Control Scheme Based on VSR

According to [6–9], VSR current control is divided into direct current and indirect current control; indirect current control has a series of shortcomings such as slow dynamic response of ac side current, sensitivity to the variation of system parameters and so on. It has been gradually replaced by direct current control. In this paper, direction current control based on space vector pulse width modulation (SVPWM) control strategy is proposed; it not only improves the voltage utilization and reduces switching losses, but also achieves sinusoidal input currents and fast dynamic response. Direct current control adopts double closed–loop structure. The feed-forward decoupling control scheme is introduced. We can obtain the decoupling equations of u_d^* and u_q^* are as follow.

$$\begin{bmatrix} u_d^* \\ u_q^* \end{bmatrix} = \begin{bmatrix} -(K_{ip} + K_{il}/s)(i_d^* - i_d) + \omega L i_q + e_d \\ -(K_{ip} + K_{il}/s)(i_q^* - i_q) - \omega L i_d + e_q \end{bmatrix} \tag{38.1}$$

Obviously, from (38.1), voltage commands have achieved decoupling control. VSR adopts double closed-loop control strategy. Block diagram of three-phase PWM rectifier is showed in Fig. 38.2. Its working principle is as follow, Design on anti-interference for three-phase voltage source PWM rectifier. Electric Drive unpublished: comparison values that voltage command minus voltage sampling data is sent to voltage PI controller and turned into current command. Because

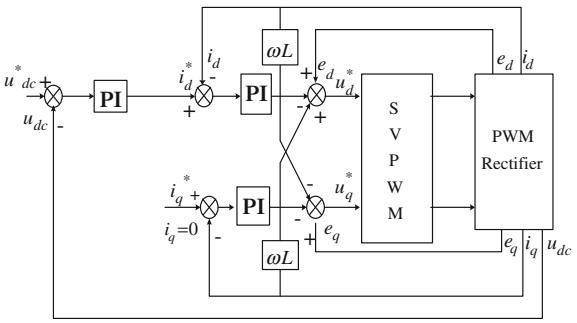


Fig. 38.2 Control block diagram of three-phase PWM rectifier

rectifier works in unit power factor in charge–discharge device, we regulate i_q to be zero. Comparison values that current command minus current sampling data is send to current PI controller and turned into voltage signal, and this voltage is sent into PWM pulse unity through coordinate transformation to complete current closed-loop control. Vector control module generates the PWM wave by vector operation, controls bi-directional PWM converter and achieves sinusoidal input current and output voltage stability.

38.3.2 Bi-Directional DC–DC Converter Control

As can be shown in Fig. 38.1, DC–DC converter operates in buck and boost mode, respectively, to achieve the battery charge and discharge. In order to meet different charge and discharge requirements, bi-directional DC–DC converter adopts current mode control that is an outer output voltage control loop and an inner averaged inductor current control loop [10]. The system having double closed-loop can achieve precise control of output voltage and output current. Control block diagram of bi-directional DC–DC converter is showed in Fig. 38.3. Current command is provided by voltage PI controller or the function control unit. By controlling switch states of s_1 and s_2, the system can change closed-loop structure, which switches between single-loop and double-loop; and the function control unit controls the voltage settings U_{set} and current settings I_{set}, so the DC–DC converter can achieve different ways to meet charge and discharge requirements, such as constant current charge, pulse charge, and constant current discharge and so on. Where VR and CR denote the voltage controller and current controller, respectively, K_{PWM} is scale factor.

38.3.3 The Battery Model

The battery is modeled using a simple controlled voltage source in series with a constant resistance, as shown in Fig. 38.4. This model assumes the same characteristics for the charge and the discharge cycles and uses only state of charge (SOC) as a state variable in order to accurately reproduce the manufacturer's

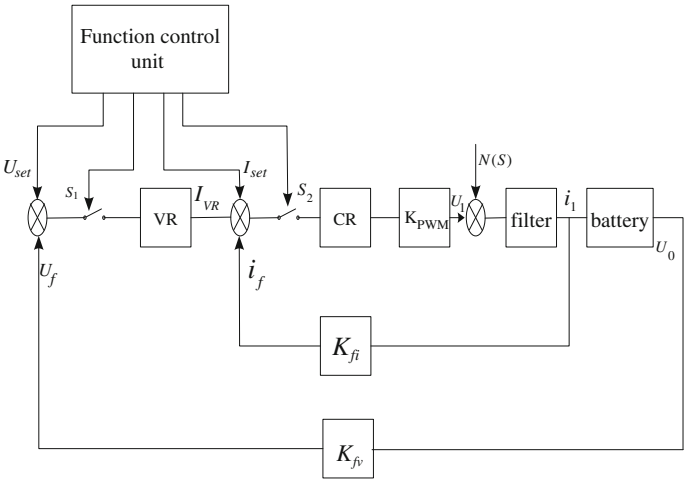


Fig. 38.3 Control block diagram of bi-directional DC–DC converter

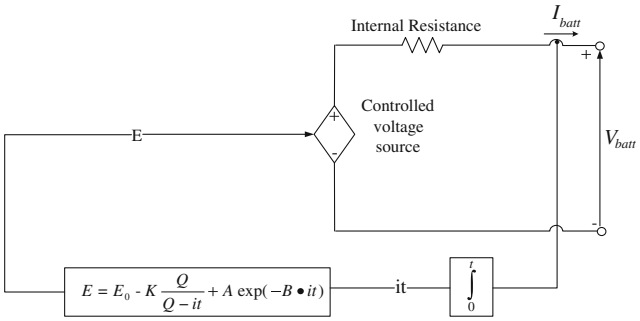


Fig. 38.4 Non-liner battery model

curves for the battery chemistries. The open voltage source is calculated with a non-linear equation based on the actual SOC of the battery. The model parameters can be deduced from a manufacturer's discharge curve [11].

The controlled voltage source is described by (38.2) and (38.3):

$$E = E_0 - K\frac{Q}{Q - it} + A\exp(-B \times it) \tag{38.2}$$

$$V_{\text{batt}} = E - R \times i \tag{38.3}$$

Where E and E_0 denote no-load voltage battery and constant voltage respectively, K and Q the polarization voltage and battery capacity (Ah), $\int idt$ and V_{batt} the actual battery charge (Ah) and battery voltage, A and B the exponential zone amplitude and exponential zone time constant inverse, R denotes internal resistance and i is battery current.

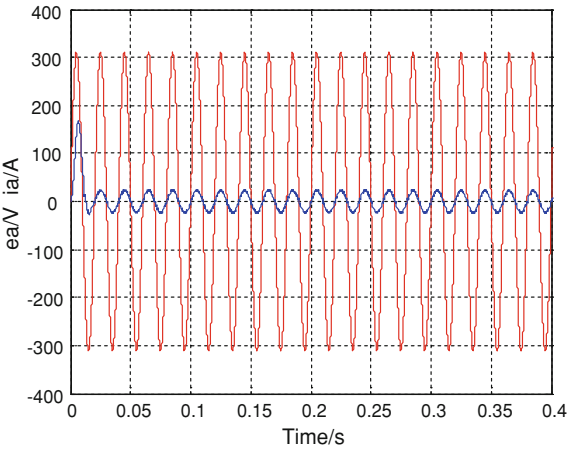


Fig. 38.5 Waveforms of source voltage and current in charge mode

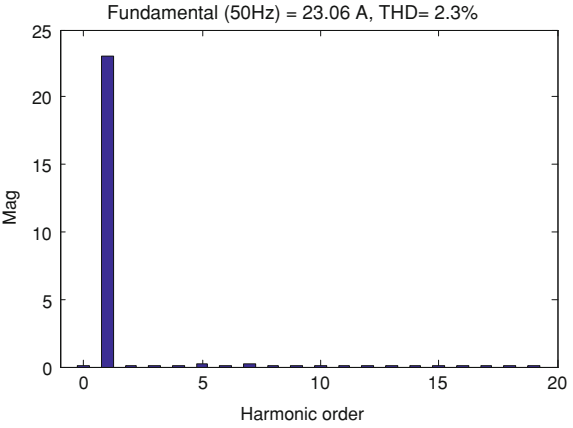


Fig. 38.6 Amplitude spectrum of input current in charge mode

38.4 Simulation Results and Analysis

In order to verify the feasibility and effectiveness of this control strategy, a charge–discharge simulation model based on constant charge/discharge is established. Select a group of dedicated lithium-ion battery of electric vehicles as the research object, its rated voltage is 330 V and rated capacity is 50 Ah. Simulation studies were carried out under the following condition: $e_a = 220\mathrm{V_{rms}}\, C_1 = 2000\mu f$ $C_2 = 6400\mu f, L = 7\mathrm{mH}, L_1 = 50\mathrm{mH}\, u_{dc} = 600\mathrm{V}, f_1 = 5\mathrm{KHZ}(VSR), f_2 = 20\mathrm{KHZ}$ (DC–DC Converter).

When the battery is charged, bi-directional DC–DC converter operates in buck mode, the initial status of battery is set to 90% SOC (charge state). The charging current is constant which is 0.2*C = 10A. Figures 38.5 and 38.6 show simulation results of source voltage and current, and amplitude spectrum of input current in charge mode. As can be seen from the figures, the source voltage and current have

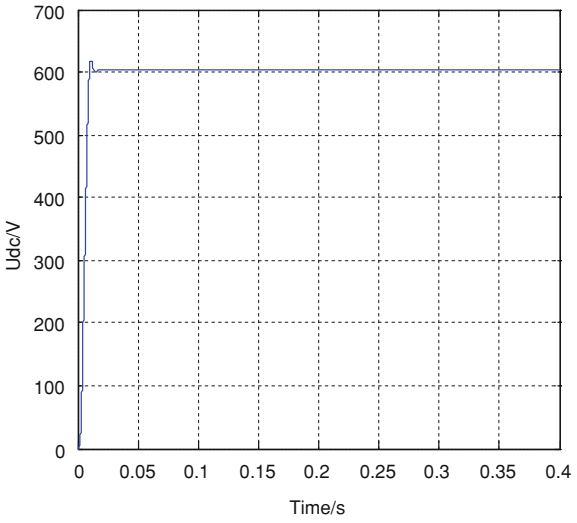


Fig. 38.7 Waveforms of dc-link voltage in charge mode

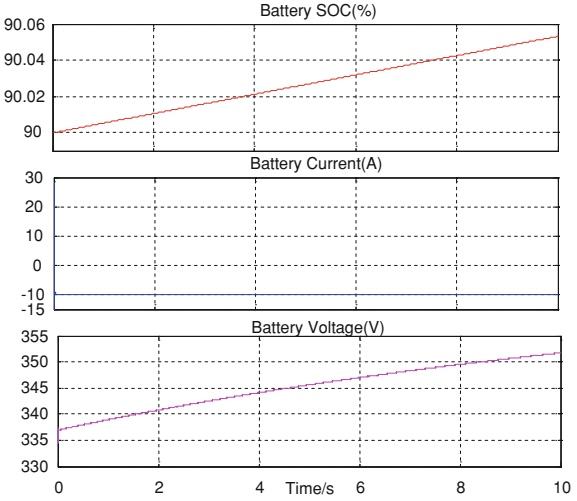


Fig. 38.8 Battery simulation results in charge mode

the same phase and system can run quickly under unity power factor. The total harmonic current of charge–discharge should be less than 5% of the rated value; the total harmonic distortion (THD) of input current is only 2.3%, which meets this standard. Figure 38.7 shows simulation result of dc-link voltage in charge mode, Dc-link voltage has a small overshoot and fast dynamic response. Figure 38.8 shows the result of battery current, voltage and SOC. The response speed of charge current is fast and battery voltage rises steadily. This means that the proposed scheme can reach good result.

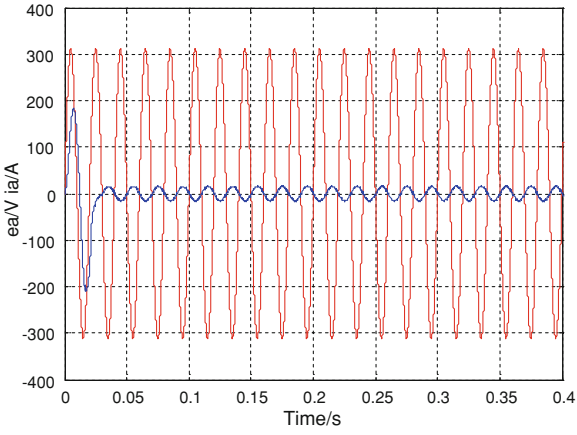


Fig. 38.9 Waveforms of source voltage and current in discharge mode

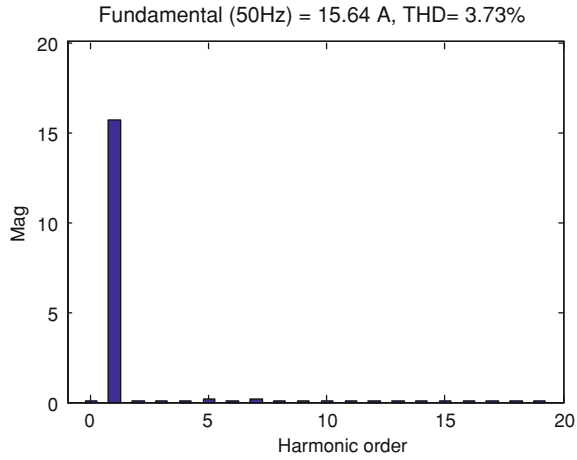


Fig. 38.10 Amplitude spectrum of input current in discharge mode

When the battery discharging, bi-directional DC–DC converter operates in boost mode. The initial status of battery is set to 100% SOC (discharge state). The discharging current is constant which is 1*C = 50A. Figures 38.9 and 38.10 show simulation results of source voltage and current, and amplitude spectrum of input current in discharge mode. The source voltage and current become reverse after the adjustment of 1.5 frequency cycle and power factor is nearly−1. THD of input current is small and only 3.73% in discharge mode, this reduces power gird harmonic, effectively. Figure 38.11 shows simulation result of dc-link voltage, the dc-link voltage is restored to the steady-state value in short time and the dynamic process is non-oscillatory. Figure 38.12 shows the result of batter current, voltage and SOC. The discharge current has a small overshoot in the initial state and can achieve a given current after a brief adjustment. This meets the discharge requirements of battery.

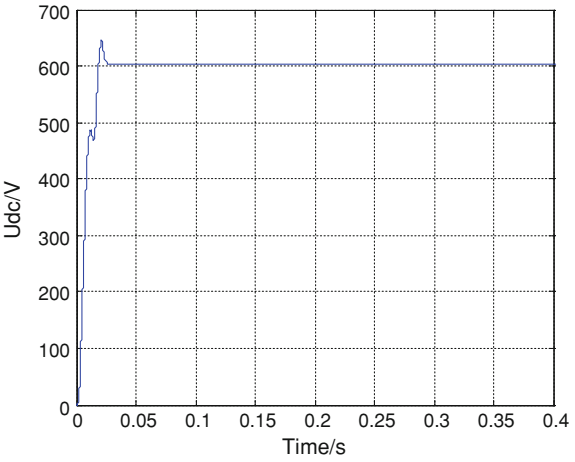


Fig. 38.11 Waveforms of dc-link voltage in discharge mode

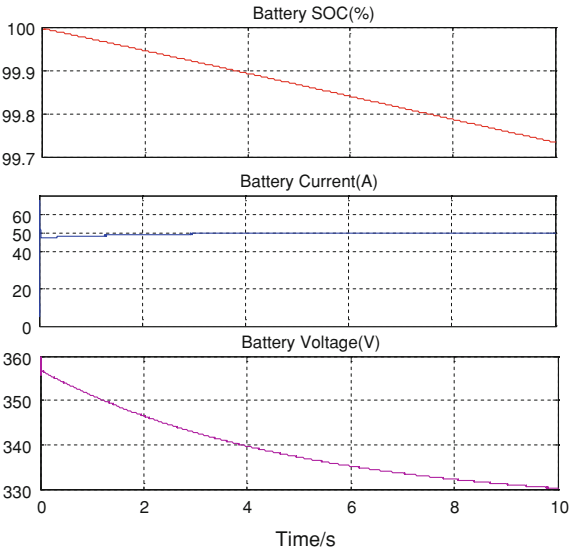


Fig. 38.12 Battery simulation results in discharge mode

38.5 Conclusion

The novel charge–discharge device consists of VSR and bi-directional DC–DC converter; it can not only work in the charge or discharge mode, but also has quick responsiveness and good stability. Simulation results show that the system can realize sinusoidal input currents and bi-directional power flow. When PWM technology is applied to charge–discharge device, this effectively improves power supply quality, increases the utilization rate of electric energy and suppresses harmonic content. Therefore, charge–discharge system based on three-phase reversible rectifier has a great industrial application value and prospect.

Acknowledgments This work is supported by the National Natural Science Foundation of China (Grant No 51077125), Hennan Science and Technology key project (Grant No 082102240008) and Educational Commission of Henan Province of china (Grant No 2008A470004).

References

1. Yu L, Yue Y, Liu Q, Yu C (2009) Research on the DC power supply of the Urban rail transit. Power Electron 43:69–71
2. Zhang W (2009) Research on locomotive battery charging and discharging device based on PWM rectifier. Power Electron 43:76–78
3. Pei X, Jiang J, Feng T (2008) Implementation of battery charge and discharge system in electric vehicle. Power Electron 42:17–19
4. Qu L, Yang Zhaohua, Yang Zhenkun, Qin Y (2004) New control device of battery charge and discharge for power system. East China Electr Power 32:48–50
5. Peng Z, Liang J, Jin X, Tong Y (2008) Research on three-phase PWM rectifier in charging-discharging device of battery. Power Supply Technol Appl 11:14–17
6. Zheng Z, Tao H (2005) Fuzzy self-tuning adaptive PI adjustment in three-phase rectifier. Electro Tech Appl 24:65–68
7. Zheng Z (2007) Fast space vector algorithm of the pulse width modulation rectifier. In: IEEE international conference on control and automation 4:2183–2185
8. Zheng Z, Jing X (2010) A simplified algorithm for three-phase PWM rectifier based on SVPWM. Electr Drive Autom 32:13–17
9. Zhang C, Zhang X (2003) PWM rectifier and its control. China Machine Press, Beijing
10. Xu H, Wen X, Li K (2003) Analysis and design of digitally controlled bi-directional DC/DC converter. Power Electr 37:13–16
11. Remblay O, Dessaint LA, Dekkiche AI (2007) A generic battery model for the dynamic simulation of hybrid electric vehicles. In: IEEE Vehicle power and propulsion conference (VPPC):284–289 9–12 Sept 2007

Chapter 39
Base on One-Cycle Control for DC Side Active Power Filter and Design PI Regulator

Donglei Xie, Liming Ji and Xiansheng Chen

Abstract The basic working principle, energy flow and output voltage ripple of DC side Active Power Filter (APF) are analyzed theoretically, the corresponding one-cycle control equation is deprived, the PI regulator for control system is designed. Correcting control system to be the typical type II system, which eliminates the steady-state error and improves the anti-jamming performance. At the same time, the new topology simplifies the control and drive circuit greatly, which has important theoretical value on power factor correction and harmonic treatment of single-phase rectifier load and provides theoretical basis for system synthesis design, too. Simulation results show that the theory is correct and reliable.

Keywords One-cycle control · DC side · II system · PI regulator

39.1 Introduction

With the widely use of power electronics devices especially large capacity variable current equipments, the power grid is polluted by the harmonic and reactive current more and more. The AC side active power filter has been used to inhibit the

D. Xie (✉)
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: xdleicqu@hpu.edu.cn

L. Ji
Power Supply Bureau, Qu Zhou, Zhejiang 324000, China
e-mail: 78075157@qq.com

X. Chen
Power Supply Bureau, Tonglu County, Tonglu, Zhejiang 311500, China
e-mail: 414142639@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_39,

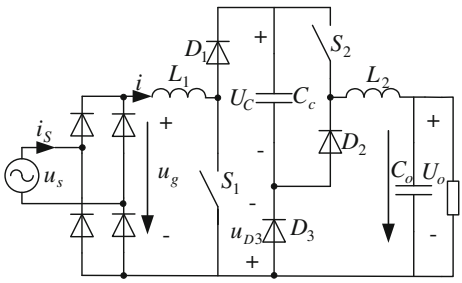


Fig. 39.1 The main circuit of the new DC side APF

harmonics and reactive current of rectifier first [1, 2]. In order to reduce the number of active switch and loss further, Ref. [3] puts forward the DC side APF in parallel. Because of the influence of the characteristics of the harmonic source, parallel type APF cannot be used to compensate for current type harmonics source [2–4]. Reference [5] puts forward a new DC side APF for the voltage type harmonic source. The new DC side APF adapts frequency control based on one-cycle control, which simplifies the control and drive circuit greatly. In order to eliminate the steady-state error and improve the anti-jamming performance, the control system is corrected to be typical type II system.

39.2 Working Principle and Energy Flow Diagram of the New DC Side APF

The main circuit of the new DC side APF is shown in Fig. 39.1. Switch tube S1, S2 are driven by the same high frequency signal. What's more, the switching frequency of S1, S2 is far higher than the frequency of power. The period of switching is TS, the duty-cycle is d. Due to the capacity of capacitor is large enough, the value of capacitor voltage U_C is approximate to the steady-state value. The compensation effect is shown in Fig. 39.2. If the diode is supposed to be ideal, the equivalent circuit is shown in Fig. 39.3.

The energy flow diagram of the new DC side APF mainly includes the charge and discharge energy of the capacitor. When the two active switches are off, energy flow is shown in Fig. 39.4a, power grid provides energy to the capacitor directly, capacitor reserves energy, capacitor voltage rises, at the same time, power grid provides energy to the load, too. When the active switches close, energy flow is shown in Fig. 39.4b, the energy of the capacitor feedbacks to power grid, and at the same time, the capacitor provides energy to the load, too.

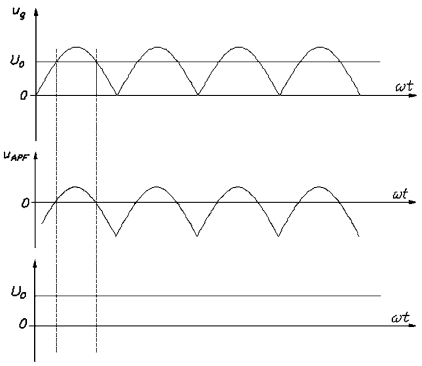


Fig. 39.2 The compensation effect diagram of new DC side APF

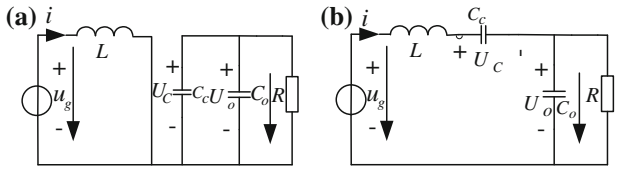


Fig. 39.3 The equivalent circuit of new DC side APF. **a** $0 < t < dT_S$ equivalent circuit. **b** $dT_S < t < T_S$ equivalent circuit

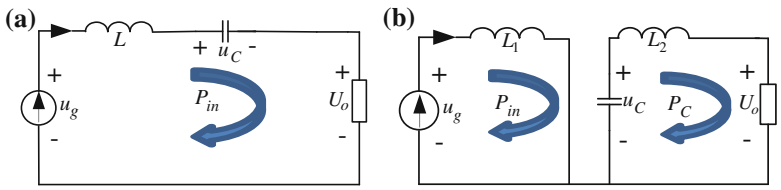


Fig. 39.4 **a** The power state flow diagram of charge of capacitor energy. **b** The power state flow diagram of discharge of capacitor energy

39.3 The Control Equation of DC Side APF and One-Cycle Control

According to the equivalent circuit and the voltage-second balance characteristics of inductor, we should obtain the following formula

$$u_g = u_c(1 - 2d) + U_0 \tag{39.1}$$

According to the Kirchhoff's voltage law, we can obtain the following formula

$$u_{L1} = u_g - u_1 = -2u_c d + \mathrm{sgn}(x)(u_c + U_0) \tag{39.2}$$

$$\mathrm{sgn}(x) = \begin{cases} 1, & S_1 \ is \ off \\ 0, & S_1 \ is \ on \end{cases}$$

where

It is known that formula (39.2) has four variables. As long as any of the three variables are controlled effectively, the remaining will also be controlled effectively.

One-cycle control is a new nonlinear control mode [6–9]. In order to satisfy control target, the average value of the switch variables is forced to be equal or proportional to the control reference parameters. In each switching cycle, the control target of main circuit is to force the phase of power current and voltage to be in-phase by controlling APF. From DC side of the rectifier bridge, the APF and the harmonic source load are equivalent to a pure resistance R_e together [5]. The control target can be expressed as the follow formula

$$u_g = R_e i \tag{39.3}$$

According to the formula (39.1), we can obtain the follow formula

$$V_m + kU_0 - R_S i = \frac{2}{T_S} \int_O^T dV_m d(t) \tag{39.4}$$

where $k = R_S/R_e, V_m = \frac{R_S}{R_e} U_C = kU_C \ R_s$ is the sample rate of current i. T_S is the switching cycle.

Formula (39.4) can be realized through one-cycle control. The control circuit is shown in Fig. 39.5. It is made of an integrator with reset function, a comparison, a RS flip-flops, a clock signal etc.

39.4 Analysis of the Ripple of Output Voltage of DC Side APF

In order to inhibit the harmonic wave in the power grid, the purpose of control is to make the current of power grid track the network voltage u_s real-time, the power $P_{\mathrm{IN}}(t)$of power grid can be expressed as the following formula

$$p_{\mathrm{in}}(t) = u_s i_s = U_S I_S \sin^2 \omega t = U_S I_S - U_S I_S \cos(2\omega t) = P_{\mathrm{in}} + \hat{p}_{\mathrm{in}}(t) \tag{39.5}$$

If the defined instantaneous power of capacitor is $\hat{p}_o$, average power of load is P_o, the output power can be expressed as the following formula

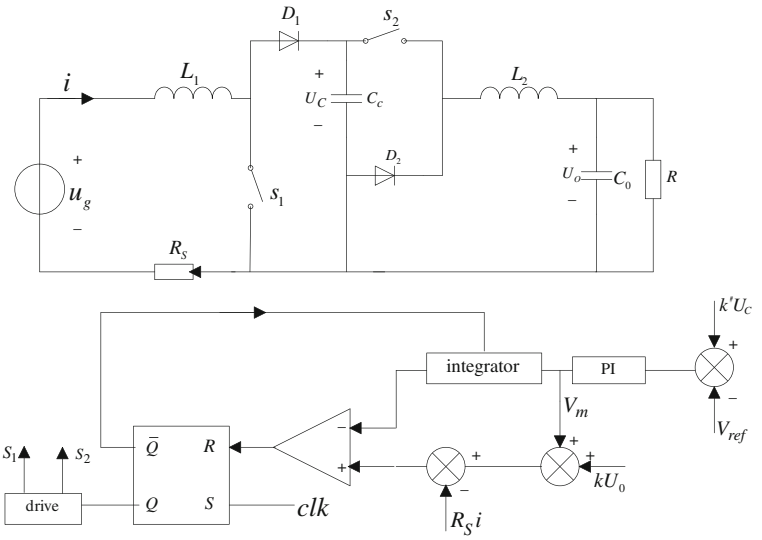


Fig. 39.5 Model of new APF with one-cycle control

$$p_0(t) = u_0^2/R + \hat{p}_0 = (U_0 + \hat{u}_0)/R + C_0(U_0 + \hat{u}_0)d(\hat{u}_0)/dt \tag{39.6}$$

where $\hat{u}_0$ is load voltage contains ripple.

Due to $\hat{u}_o \ll U_o$, the oscillatory power of resistance can be ignored, the formula (39.6) can be expressed as the following formula

$$p_0(t) \approx P_0 + 2\omega C_0 U_0 \hat{U}_0 \cos(2\omega t) = P_0 + \hat{P}_0(t) \tag{39.7}$$

where $\hat{U}_0$ is the ripple amplitude of load.

When the ripple is considered and the instantaneous power is ignored, according to power conservation law, we can obtain the follow formula

$$\hat{U}_0 \approx -\frac{p_{\text{in}}}{2\omega C_0 U_0}. \tag{39.8}$$

So the instantaneous of output ripple can be expressed as the following formula

$$\hat{u}_0 \approx -\frac{p_{\text{in}}}{2\omega C_0 U_0}\sin(2\omega t) \tag{39.9}$$

From the formula (39.9), we can draw a conclusion that factors such as input power pin, angular frequency ω, filter capacitor, DC voltage of output of the load and so on, will affect the ripple of output voltage of the load, when the circuit is optimized, the factors must be considered.

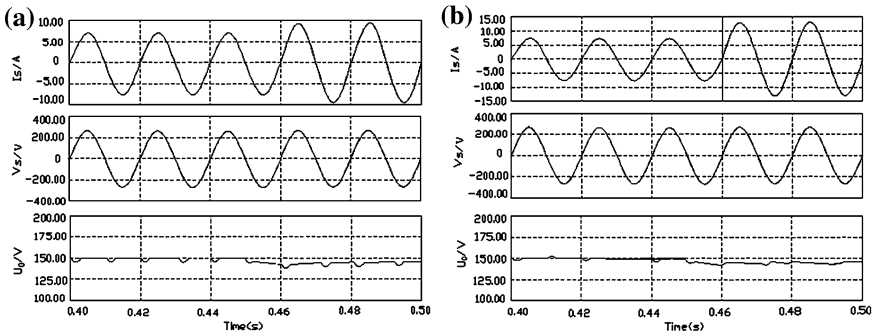


Fig. 39.6 The simulation waveforms when output voltage is 150 V. **a** Waveforms without PI regulator. **b** Waveforms with PI regulator

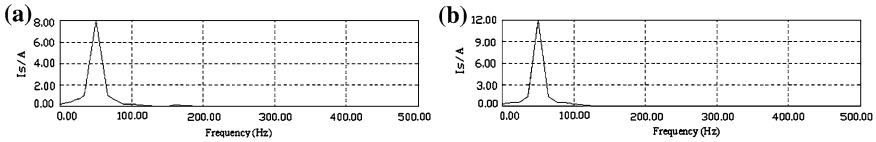


Fig. 39.7 The frequency spectrums of source current when output voltage is 150 V. **a** Waveforms without PI regulator. **b** Waveforms with PI regulator

39.5 The PI Regulator Design and Simulation

In order to keep system stable, one-cycle control APF based on the traditional PI regulator is used to maintain a constant voltage of capacitor, so the parameters design of PI regulator is key to improve the steady-state stability and dynamic performance of system. While ignoring the influence of load power and the fluctuation of power voltage ($\hat{u} = 0$), at the same time, the average of capacitor current is supposed to be 0 ($\bar{I}_C = 0$), we can obtain the following formula

$$\frac{\hat{u}_c(s)}{\hat{v}_{gg}(s)} = \frac{U^2}{R_S U_c^2} \cdot \frac{1}{CS} \tag{39.10}$$

According to the formula (39.10), we can draw a conclusion that it is a first order system. In order to improve the anti-interference ability of the system, the system can be corrected to be standard type II system. The transfer function of regulator can be expressed as the following formula

$$G(s) = \frac{1 + \tau_1 s}{\tau s(1 + \tau_2 s)} \tag{39.11}$$

The optimum design principle of typical type II system will be adapted, the corresponding parameters can be determined by Ref. [10].

The circuit parameters of simulation are expressed as follows: the effective value of input voltage is set to be 220 V, and the frequency is set to be 50 Hz,

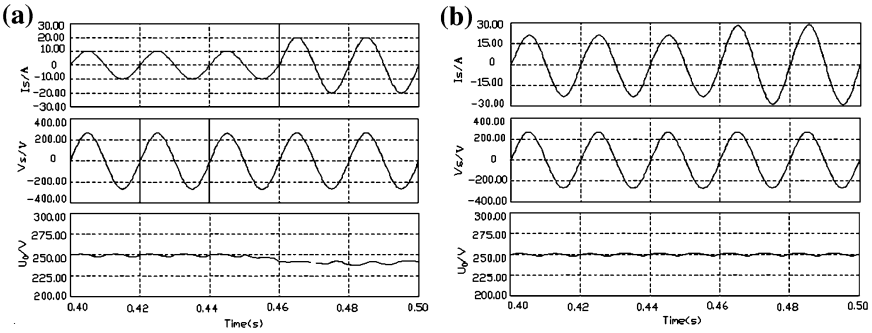


Fig. 39.8 The simulation waveforms when output voltage is 250 V. **a** Waveforms without PI regulator. **b** Waveforms with PI regulator

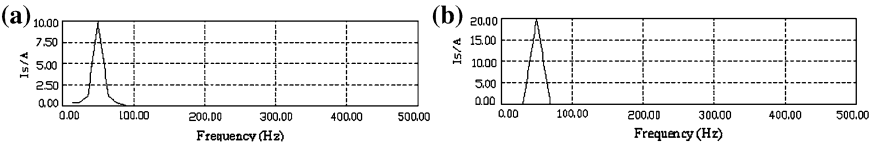


Fig. 39.9 The frequency spectrums of source current when output voltage is 250 V. **a** Waveforms without PI regulator. **b** Waveforms with PI regulator

filtering inductance L1 is set to be 1 mH, output inductance L2 is set to be 1.35 mH, capacitance C is set to be 2 mF, the rating value of load resistance is set to be 40 Ω. The frequency of clock is set to be 25 kHz. The resistance load is parallel at 0.45 S. The simulation results are shown in Figs. 39.6, 39.7, 39.8, 39.9.

According to Figs. 39.6 and 39.7, we can obtain that the system has good dynamic and static characteristics. According to Figs. 39.8 and 39.9, we can draw a conclusion that the steady-state error is eliminated effectively, the anti-jamming performance is improved, the harmonic content of output power current is reduced after the original type I system is corrected to be typical type II system.

The new topology which has been used to restraint the harmonic of rectifier load. Simple control circuit, good compensation performance, quick dynamic response is the advantages of the topology.

Acknowledgments This work is supported by the Project of the Natural Science Research of Education Department, Henan, China (No. 2001A470004).

References

1. Qiu D-Y, Yip Henry S-C, Chung SH et al (2002) Single current sensor control for single-phase active power factor correction. IEEE Trans Power Electron 17(5):623–632
2. Wang Q, Yao W, Liu J et al (2001) Harmonic source and the compensation characteristic of active power filter. J Chin Electr Eng 21(2):16–20

3. Xie P, Du X, Zhou L (2003) One cycle controlled DC side single phase active power filter. Trans China Electrotech Soc 18(4):51–55 (in Chinese)
4. Wang Q, Yao W, Liu J (2000) Voltage type harmonic source and series active power filter. Autom Electr Power Syst 24(7):30–35 (in Chinese)
5. Hou S, Zheng H et al (2008) A novel DC series active power filter. Power Syst Technol 31(15):36–40
6. Smedley KM, Cuk S (1995) One-cycle control of switching converters. IEEE Trans Power Electron 10(6):625–633
7. Zhou L, Smedly KM (2000) Unified constant-frequency integration control of active power filters. APEC, New Orleans, pp 406–412
8. Qiao C, Smedley KM, Maddaleno F (2004) A signal-phase active power with one cycle control under unipolar operation. IEEE Trans Circuits Syst 51(8):1623–1630
9. Li C, Zou YP (2004) One-cycle control active power filter for three-phase four-wire systems. In: Proceeding of IEEE on power electronics and application, vol 1. Hong Kong, China, pp 61–65
10. Zheng HB (2009) The research of dc side series active power filter based on one-cycle control. Master's degree thesis of Chongqing University

Chapter 40
Numerical Simulation of New Radiator in Electronic Device

Shu Xu

Abstract An experimental apparatus is set up to test the heat transfer performance of the new radiator, which mainly includes vacuum-pumping and Central Processing Unit tunnel system, heating system and measurement system. The heat dissipation of the Central Processing Unit chip is imitated by the copper bar which is heated by thermal resistance wire. The experimental results showed that the temperature of the Central Processing Unit chip decreased with growth of cooling air velocity, but the temperature drop tendency became gentle when it continued to increase. Higher heat dissipation power resulted in higher CPU chip temperature, but oversized heat flux density made the phase changing process worse.

Keywords Experiment · Electronic devices · Radiator

40.1 Introduction

In recent years, numerical simulation technology has been become one of the fastest, most economical and effective kinds of new research methods. In the radiator thermal performance numerical simulation calculation, many domestic and foreign scholars have done a lot of research work. At present, the commonly used commercial CFD software are: CFX, PHOENICS, FLUENT and so on. Among them, the CFD FLUENT software launched by the United States FLUENT company is currently the most comprehensive, and the most widely used software [1].

S. Xu (✉)
School of Mechanical Engineering, Huaihai Institute of Technology, Lianyungang, China
e-mail: gladxushu@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_40,

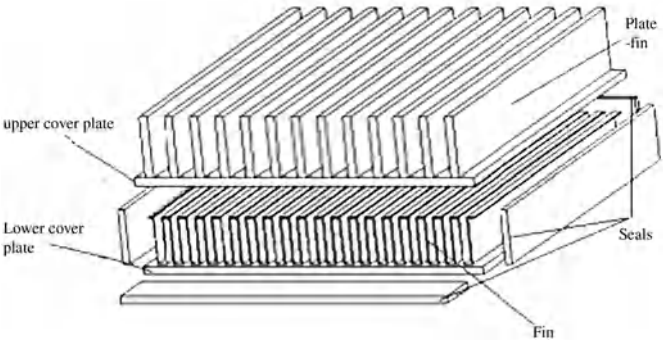


Fig. 40.1 New radiator structure diagram

In this paper software FLUENT on heat transfer and flow of new radiator and performance analysis of heat flux density will be used, air flow and other parameters on the performance of new radiator are investigated.

40.2 Design Structure of New Radiator

In this paper on the surface of a new radiator is designed as rectangular and a new radiator mainly composed of three parts:

(1) The upper and the lower cover plates: materials of the upper and lower cover plates are aluminum alloy. They are smooth rectangular planes, and their thickness is 5 mm.
(2) Core of fin: porous straight fin is used and facilitates the filling liquid refrigerant in the core body flow uniformly. It is equivalent to the wick. Fin layer, upper baffles, lower baffles and seals are connected as a whole by the vacuum diffusion welding method, forming a sealing steam chamber with fin support structure like honeycomb; the bearing capacity of refrigerant in the cavity is greatly improved. Refrigerant in the cavity can flow, evaporate and condense.
(3) The straight rib: the role of rib structure is the augmentation of convection. Heat on the surface through the air forced convection is taken away. This improves steam refrigerant condensation speed and makes it reflux to evaporation–condensation cycle (Fig. 40.1).

40.3 CFD Calculation Model

Multiphase flow heat transfer and flow of work situation of new radiator is very complex. There is heat transfer in the evaporation end, and condensation reflux is existed in the condensing surface. This is a very complex solid–liquid coupling

problem. In addition, the phase change heat transfer is also exists within the closed space. This makes study more difficult. Therefore, the following assumptions are used in the establishment of the CFD numerical model.

(1) The equivalent thermal resistance method used in thermal fin plate steam chamber within the complex phase change heat transfer equivalent heat conduction phenomenon, will be converted into steam chamber equivalent thermal resistance, thus simplifying the model [2].
(2) Fluid is defined as an incompressible Newtonian fluid.
(3) Physical parameters of air is set as a constant.
(4) Thermal radiation is neglected, and the influences of thermal radiation are not considered [3].

40.3.1 Control Equation

According to the law of conservation of mass and momentum conservation law and the law of conservation of energy [4]. Viscous fluid flow satisfies the continuity equation, momentum equation, k–ε equation and energy equation of basic control equation in the three-dimensional orthogonal cartesian coordinate system [5].

40.3.1.1 Continuity Equation

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t}+\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{40.1}$$

ρ—fluid density, kg/m^3,
u, v, w—the direction of the velocity for x, y, z respectively, m/s

40.3.1.2 Momentum Equation

$$\frac{\partial(\rho u)}{\partial t}+\nabla(\rho\vec{u}u)=\nabla(\mu g\mathrm{rad}u)-\frac{\partial p}{\partial x}+S_u \tag{40.2}$$

$$\frac{\partial(\rho v)}{\partial t}+\nabla(\rho\vec{u}v)=\nabla(\mu g\mathrm{rad}v)-\frac{\partial p}{\partial y}+S_v \tag{40.3}$$

$$\frac{\partial(\rho w)}{\partial t}+\nabla(\rho\vec{u}w)=\nabla(\mu g\mathrm{rad}w)-\frac{\partial p}{\partial z}+S_w \tag{40.4}$$

P—pressure, Pa
M—Fluid viscosity

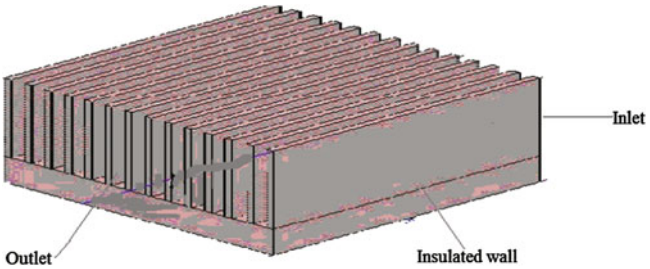


Fig. 40.2 Model boundary condition

40.3.1.3 *k–ε* turbulence Equations

K equation :

$$\frac{\partial(\rho\kappa)}{\partial t}+\nabla(\rho\vec{u}\kappa)=\nabla\left[(\mu+\frac{\mu_t}{\sigma\kappa})\mathrm{grad}\kappa\right]+\mu_t\frac{\partial u_i}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_i}{\partial x_j}\right)+\rho\varepsilon \tag{40.5}$$

ε equation :

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t}+\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{40.6}$$

40.3.1.4 Energy Equation

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t}+\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{40.7}$$

C_P—the specific heat at constant pressure
T—Temperature, *K*
t—Time, *s*

40.3.2 Boundary Condition

Whether Selection of boundary condition is appropriate is an important factor to draw the correct numerical results. The inlet conditions of Model is the velocity boundary, given the mass flow of air and inlet temperature; pressure outlet is set for exit boundary. Uniform heat flux is set for bottom center area load of model; the rest is the adiabatic boundary [6] (Fig. 40.2).

Fig. 40.3 Part of model mesh

40.3.3 Grid Division

Unstructured hexahedral meshes is used in Model grid partition (see Fig. 40.3). High quality mesh can be generated, and the grid system and flow have same direction reducing the false diffusion. It is very advantageous for the convergence and calculation speed [7]. The local encryption processing on the key parts of, grid is a total of 132862. This can not only ensure the grid quality but also accelerate the calculation speed.

40.4 The Result of Calculation and Analysis

The finite volume method is used to solve the control equations, the SIMPLE is used in pressure and velocity coupling algorithm, two-order upwind scheme are used in convection, using a coupled, steady-state implicit scheme for solving. When the variable error of continuity equation, momentum equation and energy equation is less than a predetermined value, the calculated value is changed, that is, the calculation has been converged.

Figure 40.4 is a distribution map hot of fin air channel pressure field; Fig. 40.5 is the plane distribution map of velocity field. Speed is higher near the inlet boundary and highest at the fin root. Due to straight fins form being used for the air channel, the resistance air flow encountered is little and so the pressure drop is very small.

40.4.1 Flux Effect on Heat Transfer Performance

Figure 40.6 is temperature distribution of CPU chip surface under different heat flux and the same air flow. It is visible that, along with the increase in heat flux and temperature of the chip surface, temperature different of the chip surface under two adjacent heating power increases as the heat flux density continues to increase. This shows that when the heat flux density increases to a certain extent,

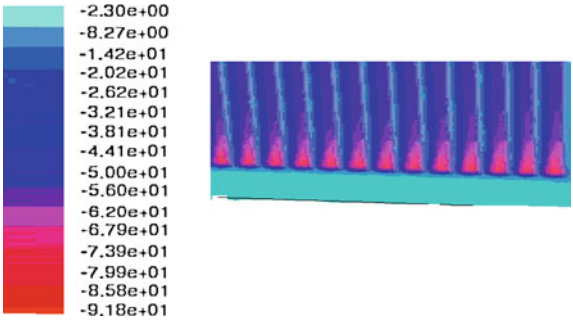


Fig. 40.4 Distribution of the pressure field

evaporation–condensation cycle to adapt the heat flux density is not completed by steam of the working fluid in chamber, and working status begins to get gradually worse.

40.4.2 Air Flow Effect on Heat Transfer Performance

Figure 40.7 is temperature distribution of different air flow radiator surface under the same flux density. It can be seen by the graph that the surface temperature of the chip is reduced with air flow increasing; but when the volume continues to increase, the temperature change becomes flat, that is, the cooling effect of the increased amount weakens. As the air flow increases, the pressure loss is increased and energy loss enters into the system in the form of heat. Eventually, the increased amount of the cooling effect gradually weakens.

In this paper transfer performance of radiator is studied by numerical simulation. The steam cavity is a closed space, and the internal phase change heat transfer model is established. In refrigerant of heat transfer there would be equivalence to conducting body of a big heat transfer coefficient heat; the following conclusion is got:

(1) With heat flux density increasing, CPU chip surface temperature increases, and as the heat flux density continues to increase, temperature difference of the chip surface under two adjacent heating power increases as the heat flux density continues to increase due to hot plate working condition becoming worse;
(2) With the air flow rate increasing, temperature of radiator surface is reduced; but when the volume continues to increase, the temperature change will flatten;
(3) With the increasing of flow temperature, the temperature of radiator surface increases proportionally.

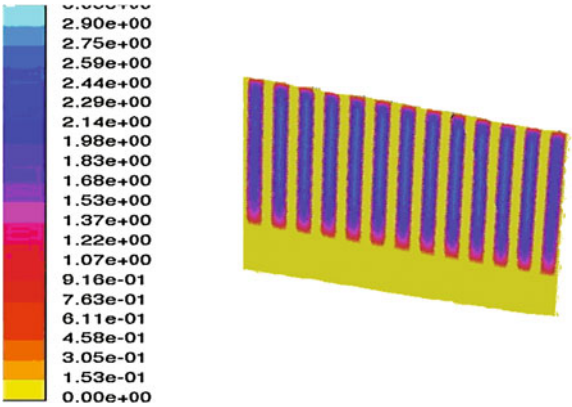


Fig. 40.5 Distribution of the velocity field

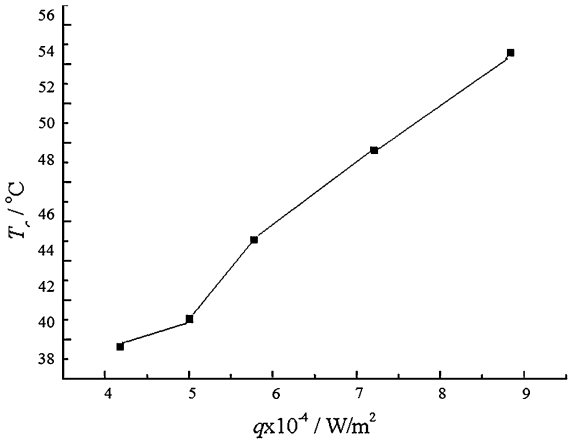


Fig. 40.6 Surface temperature of radiator under different heat fluxes

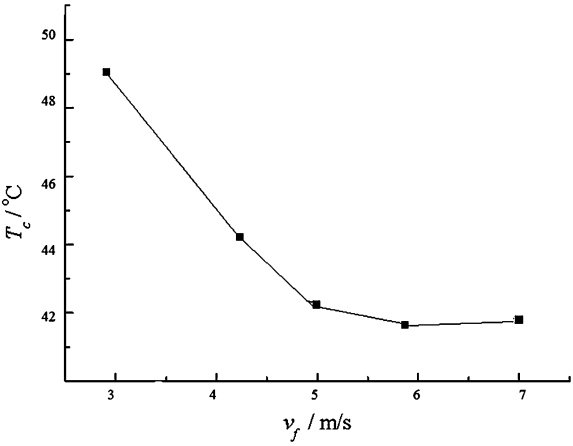


Fig. 40.7 Temperature distribution of radiator surface under different air velocities

References

1. Adami M, Yimer B (2009) Analysis of flat heat pipe with capillary grooves. Int Heat Pipe Transf Conf South Korea 10(3):38–47
2. Wang Y, Vafai K (2010) An experimental investigation of the thermal performance of an asymmetrical flat plate heat pipe. Int J Heat Mass Transf 43(16):2657–2668
3. Hu X, Tang D (2010) Experimental investigation on flow and thermal characteristics of a micro phase-change cooling system with a microgroove evaporator. Int J Therm Sci 46(7):1163–1171
4. Joung W, Yu T, Lee J (2008) Experimental study on the loop heat pipe with a planar bifacial wich structure. Int J Heat Mass Transf 51(3):1573–1581
5. Howard AH, Peterson GP (2009) Investigation of a heat pipe array for convective cooling. J Electron Package 117(9):208–214
6. Esarte J (2003) Experimental analysis of a flat heat pipe working against gravity. Appl Therm Eng 23(5):1619–1627
7. Nelson K (2009) Experimental evaluation of micro heat exchangers fabricated by silicon. Nat Heat Transf Conf 3(1):839–836

Chapter 41
Starting Performance Analysis of Asynchronous Starting Permanent Magnet Synchronous Motor

Xudong Wang, Zan Zhang, Haichao Feng and Xiaozhuo Xu

Abstract Due to the existing magnets on the rotor of Asynchronous Start Permanent Magnet Synchronous Motor and the owning of a cage structure, as well as magnetic asymmetry of the rotor *d*–*q* axis, the starting process is affected by more than one torques, the electromagnetic torque becomes very complicated in the whole journey. This paper analyzes the main starting torque in order to explore ways of improving the motor starting performance.

Keywords Elliptical rotating magnetic field · Permanent magnetic field · Starting performance · ASPMSM

41.1 Introduction

In short, Asynchronous Start Permanent Magnet Synchronous Motor (ASPMSM) add cage-type structure in rotor shaft of common Permanent Magnet Motor, not only own the features of high power index in Permanent Magnet Motor, but also

X. Wang · Z. Zhang (✉) · H. Feng · X. Xu
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zane0391@163.com

X. Wang
e-mail: wangxd@hpu.edu.cn

H. Feng
e-mail: fhc@hpu.edu.cn

X. Xu
e-mail: xxz@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_41,

has the starting capability of self-induction motor. Compared with the induction motor its efficiency is generally higher than 1–2% points, so ASPMSM has strong motivation and application of competitive prospects. But, due to his special structure its starting process is more complex than an ordinary induction motor. For ASPMSM motor design, its starting performance be analyzed is imperative.

41.2 Magnetic Field Analysis

41.2.1 The Oval Rotating Magnetic Field

For clarity, first not considering the effect of permanent magnets, the ASPMSM can be regarded as a three-phase induction motor of rotor magnetic asymmetry [1]. Stator and rotor which are affected by the magnetic field of frequency f generate asynchronous torque T_a. While magnetic field of the speed at $2n_1 - n_o$ in the stator and rotor provides reluctance negative sequence components torque T_b. When the rotor speed $n_1 \leq n_0/2$, the rotation direction of the magnetic field is the same as that of the rotor, providing a positive torque to play the function of driving; When the rotor speed $n_1 > n_0/2$, the rotation direction of the magnetic field is contrary to that of the rotor, providing a brake torque. In addition, because the rotor has permanent magnets, it will generate a rotating magnetic field of the speed n_1, the magnetic field in the armature winding produce the current of frequency $(1 - s)f$ and, thus, generates stator rotating magnetic field of the speed n_1. Its mechanism is equivalent to a synchronous generator, whose magnetic field generates a braking torque T_g. Total average torque T_{av} of ASPMSM during starting procedure is composed by T_a, T_b and T_g:

$$T_{\text{av}} = T_a + T_b + T_g \tag{41.1}$$

41.2.2 Average Torque

Due to the rotor magnetic circuit asymmetry of permanent magnet synchronous asynchronous, T_a and T_b cannot be calculated accurately, in the engineering practice we merge T_a and T_b into T_c:

$$T_c = T_a + T_b = \frac{mpU^2 R_2^{'}/s}{2\pi f[(R_1 + c_1 R_2^{'}/s)^2 + (X_1 + c_1 X_2^{'})^2]} \tag{41.2}$$

In the formula, R_1 and X_1 are respectively stator resistor and leakage reactance, R_2' and X_2' are convert value of rotor resistance and leakage reactance. $c_1 = 1 + \frac{x_1}{x_m}$, x_m is excitation circuit reactance.

Power generation braking torque T_g is calculated as follow:

$$T_g = -\frac{mp}{2\pi f(1-s)} \cdot \frac{R_1 E_0^2 (1-S)^2}{R_1^2 + X_d X_q (1-s)^2} \cdot \frac{R_1^2 + X_q^2 (1-s)^2}{R_1^2 + X_d X_q (1-s)^2} \tag{41.3}$$

41.3 Parameter Analysis

With regard to ASPMSM, Torque subsidence which is caused by the asynchronous torque T_b is relatively smaller, so it is also has less influence on the starting process. However torque subsidence caused by power braking torque T_g is comparatively larger, minimum torque in the starting process is caused by the dynamic braking torque, whose corresponding speed is low, therefore it has a greater impact on the starting performance [2].

41.3.1 Design Parameters on the Effect of Minimum Torque

41.3.1.1 Design Parameters on the Effect of Minimum Torque

Make derivative to s of Eq. (41.3), and supposing $\frac{\mathrm{dT_g}}{\mathrm{d}s} = 0$, the slip ratio s_{gm} corresponding to maximum dynamic braking torque is:

$$s_{\mathrm{gm}} = 1 - \frac{R_1}{X_q} \sqrt{\frac{3}{2}\left(\frac{X_q}{X_d} - 1\right) + \sqrt{\frac{9}{4}\left(\frac{X_q}{X_d} - 1\right)^2 + \frac{X_q}{X_d}}} \tag{41.4}$$

Put (41.4) into (41.3), then:

$$T_{\mathrm{gm}} = -\frac{mpE_0^2}{2\pi f X_q} \cdot \frac{\left(\frac{X_q}{X_d}\right)^2 \sqrt{\frac{3}{2}\left(\frac{X_q}{X_d} - 1\right) + \sqrt{\frac{9}{4}\left(\frac{X_q}{X_d} - 1\right)^2 + \frac{X_q}{X_d}}}}{1 + \frac{3}{2}\left(\frac{X_q}{X_d} - 1\right) + \sqrt{\frac{9}{4}\left(\frac{X_q}{X_d} - 1\right)^2 + \frac{X_q}{X_d}}} \tag{41.5}$$

Some conclusion are supported by Eq. (41.5) finding that: value of T_{gm} has little relation with stator resistance R_1, less EMF E_0 can substantially reduce T_{gm}. Increase X_d, reduce synchronous reactance ratio of d–q axis can also reduce T_{gm} [3].

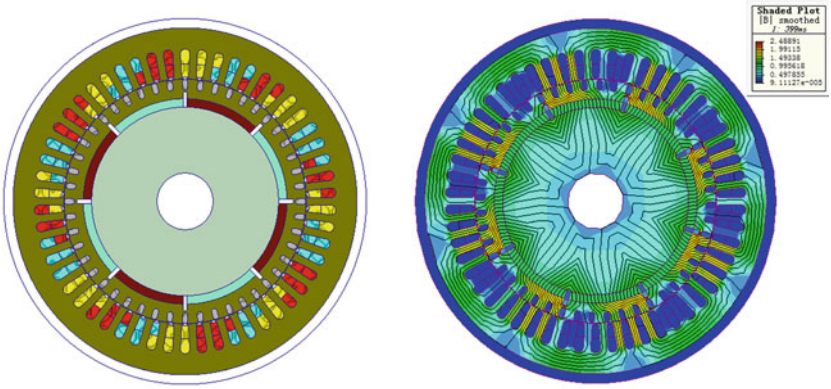


Fig. 41.1 Physic model of the motor and Flux density plot of the model

41.3.1.2 Maximum Asynchronous Torque

Make derivative to s of Eq. (41.3), and supposing $\frac{dT_g}{ds} = 0$, the critical slip ratio s_m is:

$$s_m = \frac{c_1 R_2'/s}{\sqrt{R_1^2 + (X_1 + c_1 X_2')^2}} \tag{41.6}$$

Put s_m into (41.3), then get maximum asynchronous torque T_{cM}:

$$T_{cM} = \frac{mp}{2\pi f} \cdot \frac{U^2}{2c_1 \left[R_1 + \sqrt{R_1^2 + (X_1 + c_1 X_2')^2} \right]} \tag{41.7}$$

Equation (41.7) leads to some conclusions: Increasing the stator resistance, stator leakage reactance, rotor leakage reactance will reduce T_{cM}, however changing the rotor resistance has no impact to T_{cM}.

41.3.1.3 Improve Starting Torque

Make $s = 1$, then:

$$T_{st} = \frac{mp}{2\pi f} \cdot \frac{U^2 R_2'}{(R_1 + c_1 R_2')^2 + (X_1 + c_1 X_2')^2} \tag{41.8}$$

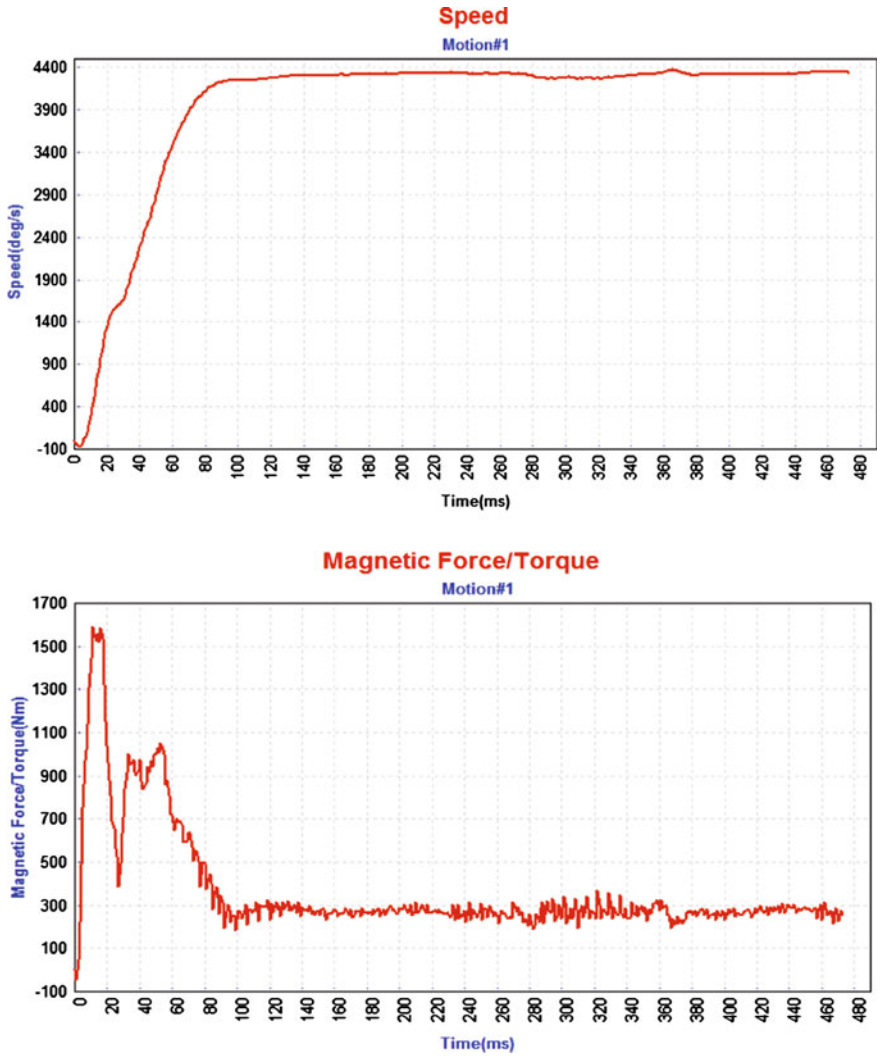


Fig. 41.2 speed–time curve and torque–time curve of original model

From the above equation methods to increase the starting torque of the asynchronous are summarized as follows: (1) Increasing the rotor resistance to narrow and deep the conductivity area, and to make the ratio of groove depth to width between 10–12, takes full use of the rotor tank skin effect; (2) Reduce the stator resistance; (3) Reduce the stator, rotor leakage reactance, change the closed tank for the open slot and so on.

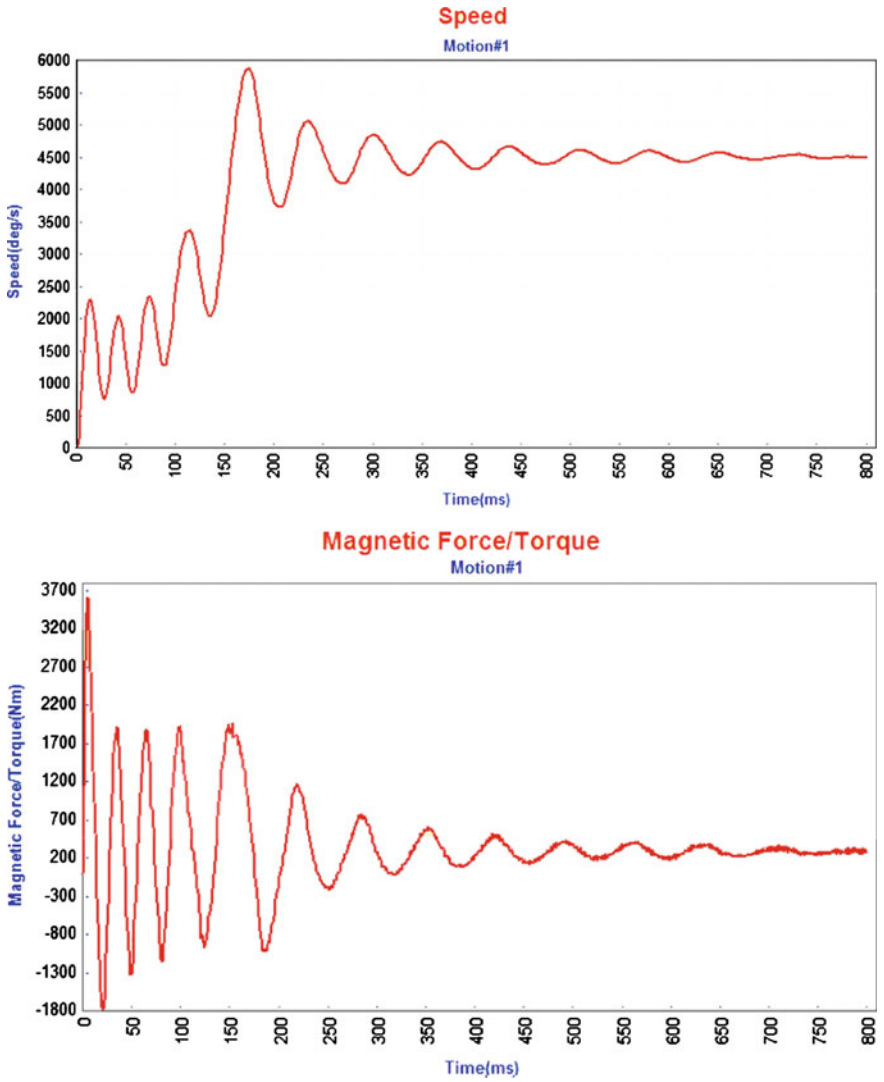


Fig. 41.3 speed–time curve and torque–time curve of original model

41.3.2 The Method to Improve the Minimum Torque

From Eqs. (41.2) and (41.3), known that the dynamic braking torque T_g has noting to do with the resistance value of R_2, however, the larger R_2 can improve the value of the asynchronous torque T_c, hence, appropriately increasing rotor resistance will better improve the motor starting performance [4]. Reducing the size of EMF E_0 can effectively reduce the dynamic braking torque T_g. Reduce ratio of $\frac{X_q}{X_d}$, can also be more effective measures [5].

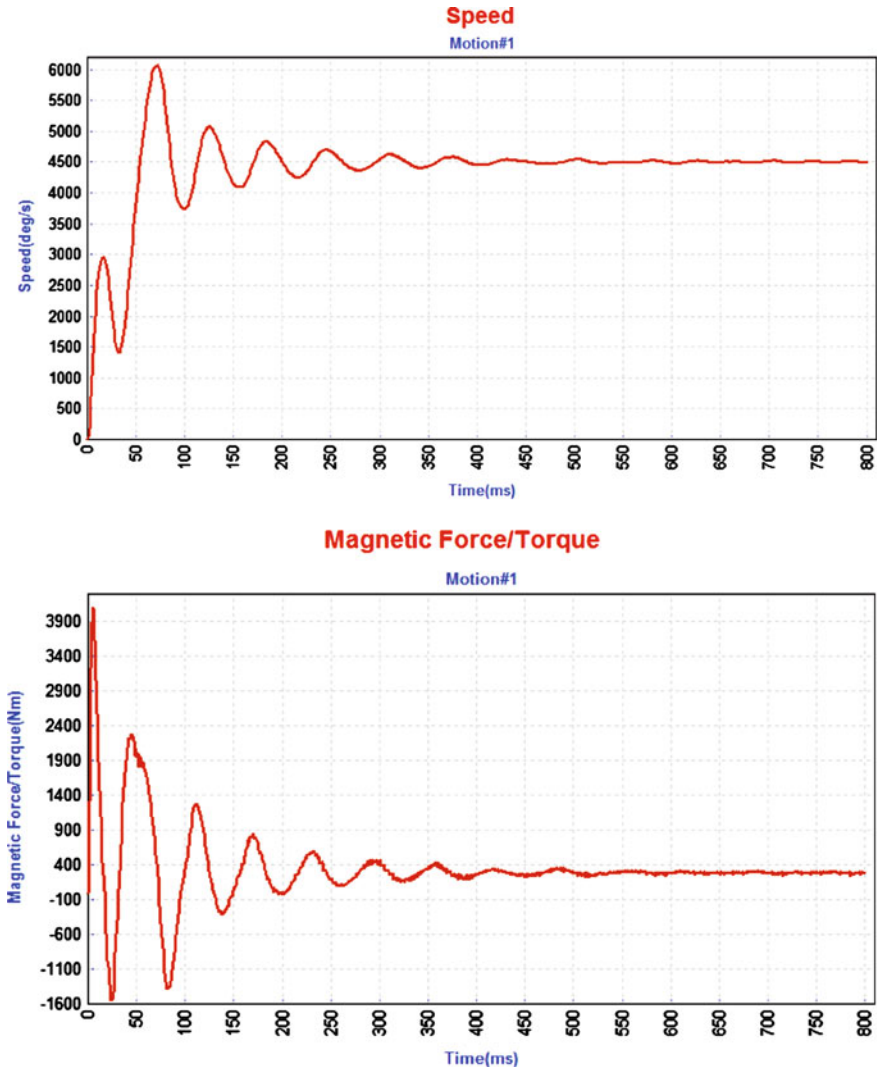


Fig. 41.4 speed–time curve and torque–time curve of original model

41.4 Simulation Verification

In this paper, finite element analysis of electromagnetic analysis software Magnet was used to establishment a model of 22 kw-8 poles ASPMSM, and analysis its starting performance. As it is shown in Fig. 41.1.

Figure 41.2 shows the waveform of speed and torque of original model under rated load. It can be seen that the rotor cannot reach rated speed. The main factor

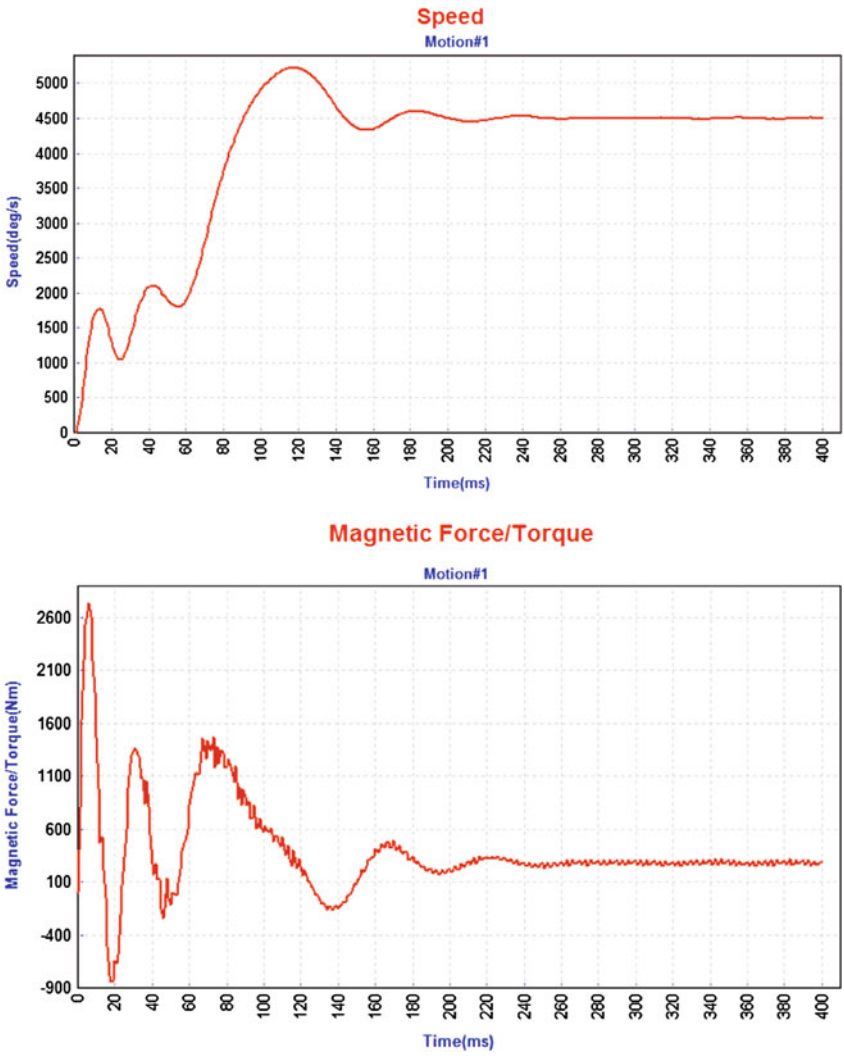


Fig. 41.5 speed–time curve and torque–time curve of original model

of this problem is the lack of rotor synchronous pull ability. The solution method is as follows: reduce the resistance of stator and rotor, and reduce sectional area of rotor slot (Fig. 41.3).

After the rotor slot is improved, motor can reach synchronous speed, but in the asynchronous start-up period speed fluctuation is very large, slip increases suddenly and reduces suddenly, speed shocks severely during a long time before it achieves stable period. By reducing stator resistance to improve synchronous torque, meantime increasing permanent magnet amount to improve motor pull-in capability. Specifically, reduce armature winding turns, thicken the permanent magnets (Fig. 41.4).

Table 41.1 Calculated by analytic method and FAE

	Analytic	FAE
Starting torque multiple	10.3	10.59
Overload torque multiple	3.3	3.24
Starting current multiple	2.98	3
Efficiency	0.95	0.9624
Power factor	0.99	0.9806

After further improvement, the asynchronous motor starting performance has been improved, but synchronous motor torque still needs further improvement. The amount of permanent magnet can be increased further (Fig. 41.5).

It can be seen from the Figure, the motor speed to complete the process in time 150 ms, 300 ms when the state into a stable, successful completion of the start (Table 41.1).

41.5 Conclusion

In conclusion, motor is a contradiction; changing any amount will result in changes in multiple motor indexes. Increasing the rotor resistance can achieve better asynchronous start torque, but too much will decrease the biggest asynchronous torque value (maximum offsets), and the motor synchronous pull-in capability will drop. Reducing back EMF can decrease T_g, meanwhile pull-out torque T_{em}, which makes the motor load ability insufficient. Similarly, too small $\frac{X_q}{X_d}$ ratio will lead to the motor load capacity decline.

Therefore, to get the optimal performance, motor parameters of overall performance should be adjusted properly, instead of one index emphasized.

Acknowledgments This work was supported by the National Natural Science Foundation of China (NO.61074095), and Henan Outstanding Person Plan (NO. 104200510021), and Ministry of Education Scientific Research Foundation for Chinese Overseas Returnees, and Ministry of Education Research Fund for Doctoral Program of Higher Education, and Henan Province Key Project (NO. 092102210359) and Doctoral Fund of Henan Polytechnic University.

References

1. Tang R (2010) Theory and design of modern permanent magnet motor. Machinery Industry Press, Beijing, pp 195–196
2. Wang X (2009) Theory, design and testing of asynchronous line-start permanent magnet synchronous motor. Machinery Industry Press, Beijing, pp 115–116
3. Liu X, Li W, Chen W (2006) Analysis of starting characteristics of three-phase line-start permanent magnet synchronous motor. Electric Mach Control 10(3):269–274
4. Zhao M, Zou J, Hu J et al (2007) Analysis of characteristics of line-start permanent magnet synchronous motor. Trans China Electro Tech Soc 22(7):145–149
5. Zhang Z, Yao G (2005) Discussion on the method of improving efficiency of small-sized motor. Explos Proof Electr Mach 40:15–17

Chapter 42
Lateral Jet Force Model Identification Based on FCM–SVM

Xiaofeng Liu, Yunfeng Dong and Xiaolei Wang

Abstract To describe the lateral jet force model accurately, an identification method based on FCM–SVM is provided. The experimental data is clustered with this method to gain the best partition and separating hyperplanes. Different classes are identified respectively using maximum-likelihood estimation method. The test data is classified through separating hyperplanes. The output of test data can be forecasted according to the expression of corresponding class. Result shows that the precision is increased by 31% compared with traditional identification model. The identification result based on FCM–SVM can supply support for the design of control system.

Keywords Lateral jet · Model · FCM · SVM · Identification

X. Liu (✉) · Y. Dong
School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing, People's Republic of China
e-mail: liuxiaofengshen@163.com

Y. Dong
e-mail: sinosat@buaa.edu.cn

X. Wang
Science and Technology on Space System Simulation Laboratory, Beijing Simulation Center, Beijing, People's Republic of China
e-mail: wangxiaolei_2345@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_42,

42.1 Introduction

To improve the response speed and guidance accuracy, the lateral jet control system was adopted in many missiles' design [1, 2], such as PAC3 missile.

The precision of this compound control system is not only related to the design of control system, but also depends on the accuracy of mathematical models. So it is very important to get high-accuracy identification models of lateral jet force.

Maximum-likelihood estimation (MLE) is the most commonly used technique for identify aerodynamics coefficients [1]. Although MLE is effect for most of the model identification problems, but it ignores statistical characteristics of data set.

This paper proposes a new missile lateral force model identification method based on FCM–SVM. FCM is used for classifying. SVM can get the better separating edge. After clustering, different classes of the data set are identified respectively adopting MLE to get different expressions. The separating hyperplanes are determined by SVM among different classes. The test sample is classified through separating hyperplanes. The lateral jet force can be forecasted according to corresponding expression.

42.2 Model Identification

Missile lateral force is usually provided by the lateral jet engine. In flight segment, the lateral jet and air flow generate a very complex interference flow field, thus leading to the lateral force models are multivariable, coupled and nonlinear models.

Suppose that there are 10 lateral jets uniform distributed in missile XY plane, the symmetry plane is XZ plane. Lateral force and torque is the interference of control system. Interaction force coefficient Kf is association with altitude H, Mach number Ma, attack angle and fire number of jets N.

$$Kf = f(H, \mathrm{Ma}, \alpha, N). \tag{42.1}$$

The data set is from wind tunnel experiment.

42.3 FCM–SVM

42.3.1 FCM

The FCM clustering algorithm is widely used to cluster (whole) object data sets X. FCM uses the membership grade function to determine which class a group data belongs to. Let $X = \{x_1, x_2, \ldots, x_n\}$ is data sets given, c is the number of

classes. The goal of clustering is to seek mapping $f \rightarrow \{1, 2, \ldots, c\}$, which assorts n groups of data into c classes.

If the Euclidean distance is used for similarity measurement, and the distance sum of square is used for clustering criterion, then the object function is

$$\begin{cases} J_f = \min \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^q d_{ij}^2 \\ d_{ij} = \left\| x_i - v_j \right\| \end{cases}, \tag{42.2}$$

where u_{ij} is a element of the membership matrix $U = [u_{ij}]_{c \times n}$, and $0 \le u_{ij} \le 1$, $\forall i, j, \sum_{i=1}^{c} u_{ij} = 1$. v_i is the clustering center of the ith cluster data set, $q \in (1, \infty)$ is fuzzy index.

The best partition method adopts gradient method to solve J_f and gets the point (u, v) that makes J_f the minimum value.

The details of one version of FCM are now given.

At the condition of $\min\{J_f(u, v)\}$, the iterative formula of membership matrix U and clustering center V is

$$u_{ij} = \begin{cases} \sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{jk}} \right)^{\frac{2}{q-1}}, & d_{ij} > 0, \\ 1, & d_{ij} = 0 \end{cases} \quad v_i = \frac{\sum_{j=1}^{n} u_{ij}^q x_j}{\sum_{j=1}^{n} u_{ij}^q}. \tag{42.3}$$

The details of one version of FCM are now given.

Step 1, normalizing the data set X. Fix c, m, and ε satisfying $1 \le c \le n$, m>1, $\varepsilon > 0$. Pick $v^{(0)}$ as the initial clustering center. Set iteration = 1 and appropriate max_iteration

Step 2, calculating formula (42.3) to refresh membership matrix U and clustering center V.

Step 3, if $\left\| \text{Cent}^{(ineration)} - \text{Cent}^{(\text{iteration}+1)} \right\| < \varepsilon$, then algorithm is end, else return Step 2.

42.3.2 The Optimal Clustering Number

Clustering is an unsupervised learning method. The labels of sample are unknown before classifying. The practical data set has optimal clustering number generally. So finding the optimal clustering number is very important.

A well-established hard cluster validity criterion is the separation index H_{cs}, which identifies compact grade comp and separate grade sep. H_{cs} is defined by

$$H_{\text{cs}} = \frac{\text{sep}}{\text{comp}} = \frac{\min_{i \ne j} \left\| v_i - v_j \right\|^2}{\frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij}^m \left\| x_j - v_i \right\|^2} \tag{42.4}$$

sep is the separate grade of the different classes. comp is the compact grade of intra-class. A well clustering result contains two aspects, first is sep should be enough small and second is comp should be enough large.

The method of finding the optimal clustering number is as following. The first step is fitting the maximum class number $c_{\max}$. After that, clustering use FCM algorithm and calculating H_{cs} using above formula when the value of c from 2 to $c_{\max}$, the c of the minimum H_{cs} is the optimal clustering number.

42.3.3 Nonlinear Separable SVM

SVM classifications may be more accurate than the widely used alternatives such as classification by decision tree, and neural network-based approaches. An SVM aims to fit optimal separating hyperplanes (OSH) among classes by focusing on the training samples that lie at the edge of the class distributions, the support vectors.

In SVM classification, we are given training data $\{x_1, x_2, \ldots, x_n\}$ that are vectors in some space $x_i \in R^d$. We are also given the labels $\{y_1, y_2, \ldots, y_n\}$. For nonlinearly separable classes, $\{\xi_i\}_{i=1}^n$ may be introduced with the constraint becoming [3]

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i. \tag{42.5}$$

A penalty term, $C\sum_{i=1}^n \xi_i$, is generally added to penalize solutions for which ξ_i are very large. Under the constraints of (42.5), then the optimization becomes [3]

$$\min\left[\frac{\|w\|^2}{2} + C\sum_{i=1}^n \xi_i\right]. \tag{42.6}$$

This is a convex programing problem, which is dealt with by Lagrangian function [4].

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \phi(x_i^T)\phi(x_j). \tag{42.7}$$

Subject to constraints $\alpha_i \geq 0$, $i = 1, 2, \ldots, n$ and $\sum_{i=1}^n \alpha_i y_i = 0$.

Where, α_i is Lagrange multiplier of the ith sample.

Generally, for getting OSH, the input data may be mapped into a high-dimensional linear space through a nonlinear mapping function ϕ. An input data point x can be represented as $\phi(x)$ in the high-dimensional space [3]. The formula of $(\phi(x) \cdot \phi(x_i))$ in a high-dimensional space is reduced by using a positive definite kernel such as

$$\phi(x) \cdot \phi(x_i) = k(x, x_i). \tag{42.8}$$

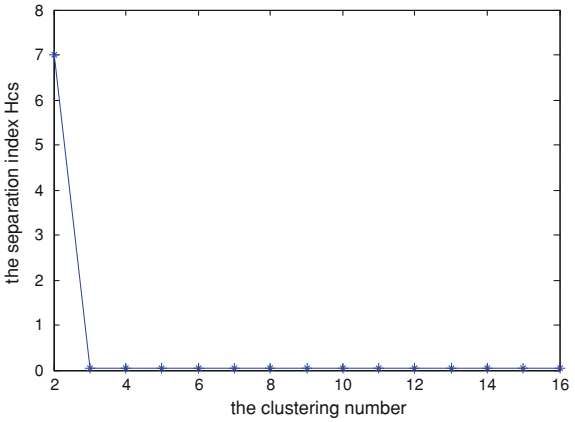


Fig. 42.1 Calculating the optimal clustering number of experiment data

The function K is called the Mercel kernel, which must satisfy the Mercer condition. The Mercel kernel used in this paper is the radial basis function

$$k(x, x_i) = e^{-\alpha \|x - x_i\|^2}, \tag{42.9}$$

where α controls the width of the kernel function.

Thus, by solving the problem, the decision function is mapped into [3, 4, 5].

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i k(x \cdot x_i) + b\right). \tag{42.10}$$

42.4 Identification Result

Fitting $c_{\max} = 16$, the wind tunnel experiment data is clustered by FCM. The separation index H_{cs} is shown in Fig. 42.1.

When c is 2, H_{cs} is the maximum one. So the optimal clustering number is 2. The 270 points are clustered into two classes.

Although which class a group of test data belongs to can be fixed based on the distance from the test data to different clustering centers, the result through this method is not accurate sometimes. SVM can get the better separating edge.

After getting the class label used FCM, we select 260 groups of experiment data as the training sample, and the rest 10 groups as the test sample. Different classes of the training sample are identified, respectively, adopting MLE to get different expressions. Besides, the separating hyperplane is determined by SVM. The test sample is classified through the separating hyperplane. Interaction force coefficient Kf of the

Table 42.1 The result forecasted based on FCM–SVM

Group	Labels	Labels forecasted by SVM	Kf	Kf forecasted
1	1	1	0.1038	0.1046
2	2	2	0.3694	0.3702
3	1	1	0.4582	0.4577
4	1	1	0.8547	0.8468
5	2	2	0.2212	0.2137
6	1	1	0.5841	0.5844
7	1	1	1.3695	1.3712
8	1	1	−0.2485	−0.2496
9	2	2	−0.5418	−0.5423
10	2	2	−0.7637	−0.7629

Table 42.2 The compared result between the identification based on FCM–SVM and the traditional identification

Model	Precision
Model identified directly	0.7295
The 1st class	0.9598
The 2nd class	0.9736

test sample is forecasted according to corresponding expression. The result forecasted is shown in Table 42.1.

The compared result between the identification based on FCM–SVM and the traditional identification is shown in Table 42.2.

42.5 Conclusion

The model identification method based on FCM–SVM mines the category information sufficiently, supplies a new means for identification of the missile lateral force.

The simulation result shows that the identification precision after classified by FCM–SVM is increased by 31% at lease compared with the result identified directly. The identification result based on FCM–SVM can supply reference for the design of control system.

Acknowledgments The authors thanks Beijing Simulation Center for their support and for providing data used in this work. The author also thanks the financial support of science and technology on space system simulation laboratory fund item of china.

References

1. Hirokawa R, Sato K, Manabe S (2001) Autopilot design for a missile with reaction-jet using coefficient diagram method. AIAA–2011–4162
2. Liu Z, Liang X, Cao B et al (2006) An optimal backstepping design for blended aero and reaction-jet missile autopilot. J Appl Sci 6(12):2623–2628
3. Mathur A, Foody GM (2008) Member. Multiclass and binary SVM classification: implications for training and classification users. IEEE Geosci Remote Sens Lett 5:241–244
4. Manikandan J, Venkataramani B (2008) Diminishing learning based SVM classifier with non-linear kernels. In: Proceedings of IEEE 2008 international conference on electronic design
5. Kim Y-G, Jang M-S, Cho K-S et al (2006) Performance comparison between backpropagation, neuro-fuzzy network, and SVM. CSR, LNCS, vol 3967. pp 438–446

Chapter 43
Design and Implementation of an Intelligent COTS System

Ping Zhang and Bin Tao

Abstract COTS system is one of the main software development methods in the future; this paper first analyzes the needed functions of the intelligent COTS system, and then introduces design and implementation methods of COTS system. Intelligent COTS system is a software development environment supporting component reuse, not only to manage effectively and query components, but also to well support the integrated components to develop software system based on the reuse.

Keywords COTS components · COTS system · Intelligent

43.1 Introduction

As the software component technology research and practice continually develops, the current trend is to form the software system by the use of either commercial developed components or useable components assembly in accordance with application requirements. COTS system is defined by COTS components (abbreviated by component below) assembled from the software system, in the ideal situation, the software will be more reliable, better structured, faster developed by using sophisticated COTS components. The requirements of software system is

P. Zhang (✉) · B. Tao
Department of Computer Science and Engineering,
Henan Engineering College, Zhengzhou 451191, China
e-mail: zpings@sina.com

B. Tao
e-mail: tb3190@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_43,

complex, changeable and uncertain that determines software systems which use components demands continuous evolution, accordingly, the need for the replacement, add and delete of components in software system or the add and modify of components service and so on. The establishment of an intelligent COTS system is of decisive significance for software system development. This paper will focus on the design and implementation methods of intelligent COTS system.

43.2 Function Modules of Intelligent COTS System

Intelligent COTS system should have the functions of components which involved classification description, component search and matching and assembly debugging. The existing components management generally only limited to add, deletes, modifies and search to various types of information on the components. If the software system used to retrieve the modeling and integration of COTS components of intelligent system these two parts of the work included in the component database management system components not only can provide the basis for the description of query and help, but also to play better the role of component library management system. According to the above thinking, the intelligent COTS system, will be divided into five functional modules: storage system components, component database maintenance systems, software modeling system, component matching support systems and software systems generator.

(1) Component classification storage system [1–3]: has the large number of components and shapes. In order to make users correctly understand the use of these components, reducing the use of cost, component storage system needs to describe and classify components and storage.
(2) The software modeling system: to provide some system modeling tools and methods, tools and methods through the analysis from the perspective of the field of software need what kind of components, the task of each component, the query conditions and connections between the components.
(3) Component database maintenance system: included to retrieve, modify, remove components in the component library, as well as change all types of information on the components and the code and statistic the number of components of various types, use frequency, and software maturity.
(4) Component matching support system: the components in actual use, the user's understanding and matching of components is often greatly subjective, choice of components matching the program's objective is no scientific basis, thereby increasing the pressure of the user, matching the subjectivity of the program, affecting the efficiency of component reuse. Users need to provide an objective tool to support effective decision making for component matching to improve the efficiency of system development based.

(5) Software system generator [4–6]: the functions to make the retrieved software modeling system components in accordance with the relationship between the files of components in the deployment of information integration as a new component or a complete system. It can be automatically or semi-automatically to modify the cases of components and the type of conversion and other modification work, which is to meet the system requirements. The modified components of the modified system model based on the interrelationship provided by software system module can be assembled into software system.

43.3 LNCS Online

43.3.1 Classification Method of Software Components Storage

Component category is to generate and maintain the components directory information base. The system uses the classification description method that supports a multi-field component and the custom classification model, can expand flexibly on describing the properties of components. Specifically, by the description method, it includes category and property expansion. Classification extension mechanism support keyword, enumeration and faceted classification model classified under the law of custom features, the owner can release, modify and maintain a specific classification through the; properties extension mechanism through custom type for the property-value of a specific category classification model, to achieve the expansion of component properties.

43.3.2 Software Modeling System

Software modeling system establishes four models: dynamic model, organization model, functional model and information model.

(1) Dynamic model: description of the existing systems or future systems with time-varying dynamics.
(2) Organizational model: describes the hierarchical structure of software.
(3) Functional model: description of system features and the information link between functions. Functional model can be used as a tool for system analysis and exchange, not only express requirements for functional system, but also provide an effective communication and exchange between the user and the modeled one. Each entity of the functional model is the needed component in a software system, showing the query conditions in the model.
(4) Information model: describes the various information systems and their interlinks.

Modeling orders in software modeling system is:

(1) Establish the system dynamic model that describes the actual business process system;
(2) Create organization model of system, and describe the organization system;
(3) Generate the functional model system by extract the system features from the dynamic model of the business process;
(4) Generate information entities, access to information entities based on the dynamic model and functional model, derived information entities to add information to the entity's property, access to information through the responsible entity of the message.

43.3.3 Component Database Maintenance System

Component maintenance system database is responsible for the system components to complete the information in the library to add, modify and delete. Maintenance system is also coping with the different types of soft classification of the number of components, reuse, reuse frequency and cost of real-time statistics. In addition, the system's main job that responsible for maintaining is to retrieve the required components in the component database. Besides, it can be divided into two tasks in search of components: the components classification index and query of the required components. Taking into account the shortage of traditional search methods, this paper use an improved algorithm for B-Tree indexing valid query component can be used not only effectively improve the time efficiency of the query components, but also increase the parallel component parallel query degrees [7–9].

43.3.4 Component Matching Support System

In order to enhance the objectivity of matching components, reduce the pressure of user, and improve the efficiency of the software, it adds to a component matching support system based on the component database. The user feedback on the component database of reusable components in application development in the past reuse as well as their basic properties as a case input the system. It consists of two subsystems: feedback subsystem and evaluation subsystem. Feedback subsystem responsible for collecting user feedback component views of reuse of feedback and evaluation; evaluation subsystem put up with the evaluation of components through the historical data analysis of candidate components, including quality information and reuse in the current match the case of decision support information, helping user select components and make matching decision. And it reduces knowledge of the data in component matching support system, and digs out the useful classification rules and their matching support [10–14].

43.3.5 Software System Generator

At present, the existence of different component model, resulting in difficult assembly and cross-platform interoperability. In this paper, it proves the characteristics of the existing component assembly method to compare differences between the mainstream component model, and software system generator with support for heterogeneous component-assembly approach. Software system generator includes five components: component wrapper, packing plants, connectors, assembly plant and the user interface module. Component wrapper component model encapsulates the corresponding components, so that components outside to provide a unified interface, showed a consistent view of the components, shielding component differences. The entire components wrapper constitutes packing plant, including a collection of interfaces and classes; it can be independent of the component mode through the interface, a unified way to access components. Packing plants have also provided the components from the packaging to the specific components of the mapping mechanism. Connector plant, including different types of connectors, generates and instanced different connector according to the diverse connections. Assembly plant undertakes the deployment of components to generate a component instance. User interface module for the component assembly developers a powerful graphical interface, the developer through the user interface, such as the component viewer, property pages and other components to achieve browsing, a graphical representation of components and component assembly visualization and many other functions, facilitate the development of personnel component assembly work.

43.4 Conclusion

Intelligent COTS system is a software development environment supporting component reuse, not only to manage effectively and query components, but also to well support the integrated components to develop software system based on the reuse. The system has the following features:

(1) Making modeling theory into the system queries and integration of soft components, so that the model throughout the whole process of system development, it would be better to meet demand, be flexible and easier to modify.
(2) Using a support custom classification method in the classification of software component.
(3) Enabling users to easily query the software components to meet the needs by taking into account the quality and frequency of reuse and other factors in the search of component;
(4) Enhancing successful rate of matching by using component matching support system;

(5) Software system generator support homogeneous and heterogeneous components assembled.

References

1. Hydalsvi GM Karlsson EA (1991) Object-oriented development with and for reuse. In: Proceedings of the 5th international conference on computing and information, pp 1–13
2. Raggl A, Slany WA (1998) Reusable iteratie optimization software library to solve combinatorial problems with approximate reasoning. Int J Approx Reason, pp l61–l91
3. Batory D, Bait JG (1996) Validating component compositions in software system generators. In: Proceedings of the 4th international conference on software reuse, pp 36–48
4. Thomas A (1998) Enterprise JavaBeans technology server component model for the Java platform. Sun Microsystems Inc
5. Lohr KP (2002) Towards automatic mediation between heterogeneous software components http://www.elsevier.nl/Iocatc/entcs/volume65.html
6. Jiang G (2002) Interoporability technology assessment http://www.elsevier.n1/Iocate/entcs/volume65.html
7. Paul CC (1996) From subroutines to subsystems: compoment-based software development. In: Alan WB (ed) Component-based software engineering: selected papers from the software engineering institute. Los Alamitos, CAUSA: IEEE Computer Society Press, New York, pp 3–6
8. Subrahmanyam A, Cedric B, John D (2001) Professional Java server programming J2EE, 113th edn. Wrox Press, Birmingham, pp 694–646
9. Jacobson I, Booch G, Rumbaugh J (1999) The unified software development process. Addison2Wesley, Boston
10. Taylor RN, Medvidovic N, Anderson K et al (1996) A component and Message-Based architectural style for GUI software. IEEE Trans Softw Eng 22(6):390–406
11. Shaw M, Deline R, Klein D et al (1995) Abstractions for software architecture and tools to support theme. IEEE Trans Softw Eng 21(4):314–335
12. Haddox J (2002) Testing commercial-off-the-shelf software components, Software engineering, p 9
13. Mark RV (2000) System implementation using commercial off-the-shelf (COTS) software, Technical report
14. Mark V (1999) Inspecting COTS based software systems, Technical report

Chapter 44
Study of the Intelligent Monitoring and Analysis System of High-Rise Building Fire Based on Expert System

Tao Li and Hui Li

Abstract To study the ways how to monitor the high-rise building fire timely and analyze it accurately, a kind of intelligent monitoring and analysis system of high-rise building fire based on expert system is designed in this paper. In this system the lower machine is based on single-chip micro-processor and the parameters such as temperature, airflow rate, smoke concentration and so on are detected in the case of unattended by multi-transducers which are linked to the single-chip microprocessor by single-bus technology. The system's structure is simple and its capability is stable, so automatic detection and management without person can be realized by this system, which has great help in the high-rise building fires prevention and reducing fire loss.

Keywords High-rise building · Expert system · Multi-transducers · Fire monitoring

44.1 Introduction

With the rapid development of urban construction, more and more high-rise buildings have been built in modern cities. However, the gap has been obvious between the ability of existing fire-fighting equipment and high-rise buildings of great height. City high-rise building construction in particular, in the event of fire, often cannot be effectively fighting, eventually resulting in significant property damage and

T. Li (✉) · H. Li
School of Electrical Engineering and Automation,
Henan Polytechnic University, 454000 Jiaozuo, Henan, China
e-mail: lthn@hpu.edu.cn

H. Li
e-mail: lihui@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_44,

casualties. So how to detect and deal with high-rise building fire to avoid the development of fire and how safe and timely transfer of high-rise construction workers are important and difficult problems. So the development and promotion of accurate and reliable high-rise building fire monitoring and forecasting system is imperative.

The CAN bus technology has many advantages such as the strong ability of antijamming, the high speed of communication, the large number of detection nodes and so on. So the networks based on it has many advantages just like the antijamming ability, the networks safety and the low building cost, it can also realize the purpose of the distributed and real-time control, especially in the complex environment. According to these characters, an detection and control system is discussed in this paper, the higher computer of this system is PC, the lower computer of it is AT89S52 single-chip microprocessor (SCM), the transducers are used to make real-time monitoring for temperature, wind speed, smoke concentration and other important parameters in high-rise buildings key areas by a single bus. And the real-time test results are transferred to PC. At the same time PC software uses expert systems to accurately determine in the current building whether the fire occurred, fire trends and the direction. Then automatic fire extinguishing system andautomatic alarm are started after confirming the occurrence of fire to issue fire alarm. According to the development trend of fire to provide a safe evacuation route and guide senior staff were safely evacuated, by achieving timely detection, handling fire, safe transfer of personnel to reduce personnel and property loss purposes.

44.2 Overall System Plan

According to features of high-rise building, such as crowded, complex, high floor, long distance transmission and testing a wide range, the system based on CAN bus technology in high-rise building fire monitoring was designed in this paper. The system mainly includes three parts: the higher computer, the core of which is PC; the lower computer, the core of which is the AT89S52 (SCM) and the communication part, the core of which is the CAN bus. Among them, the higher computer is for data analysis, processing, storage and the lower computer control. SCM system is the core of all the key area monitor to complete a variety of data collection work and the hazard warning. CAN bus is used for higher computer and lower computer systems and the networking among the lower machines to be an efficient system of work [1]. Overall system structure is shown in Fig. 44.1.

44.3 System Design

Reliable on-site collection of the important parameter is an important part of high-rise building fire monitoring system directly related to the accuracy of judgments. In this system, the collection of the parameter is mainly accomplished by the low

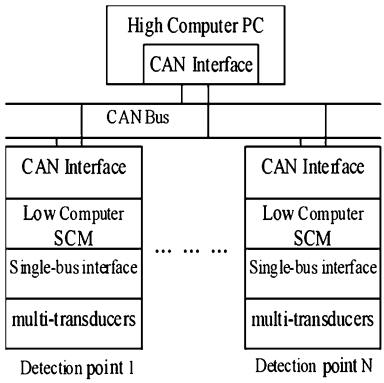


Fig. 44.1 Overall structure of the whole system

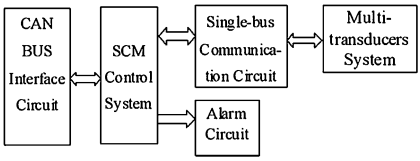


Fig. 44.2 Overall structure of the hardware system

computer system. The core of the lower computer is the AT89s52 SCM and also includes data detection unit, control unit , single-bus communication unit and so on. Overall structure of the hardware system is shown in Fig. 44.2.

44.3.1 CAN Bus Interface Circuit Design

CAN interface circuit is the key system connected with the lower computer and the higher computer for data transfer. The system CAN bus communication interface circuit is mainly composed of the CAN bus controller SJA1000, CAN drive PCA82C250 and optical isolation circuit. In the information transmission it will be to the CAN bus communication interface SJA1000 through P0 interface of the SCM. The channels between CAN bus and the SJA1000 is formed by the TX0 and RX0 of SJA1000 and CAN drives PCA82C250 through the photoelectric coupler 6N137 [2]. The basic structure is shown in Fig. 44.3.

44.3.2 Design of Data Acquisition and Alarm Unit

Data acquisition and alarm unit can be divided into on-site sensor detection, analog/digital conversion unit, communication interface and converting and sound and light alarm. Its main function is to capture real-time and important environmental parameters and transmit the data to the higher computer and in the face of

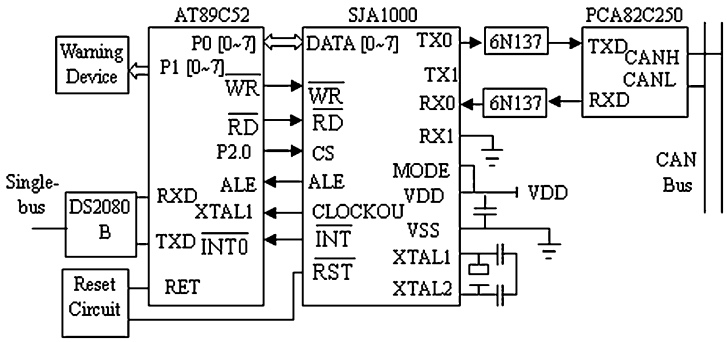


Fig. 44.3 The basic structure of interface circuit between lower computers and the CAN bus

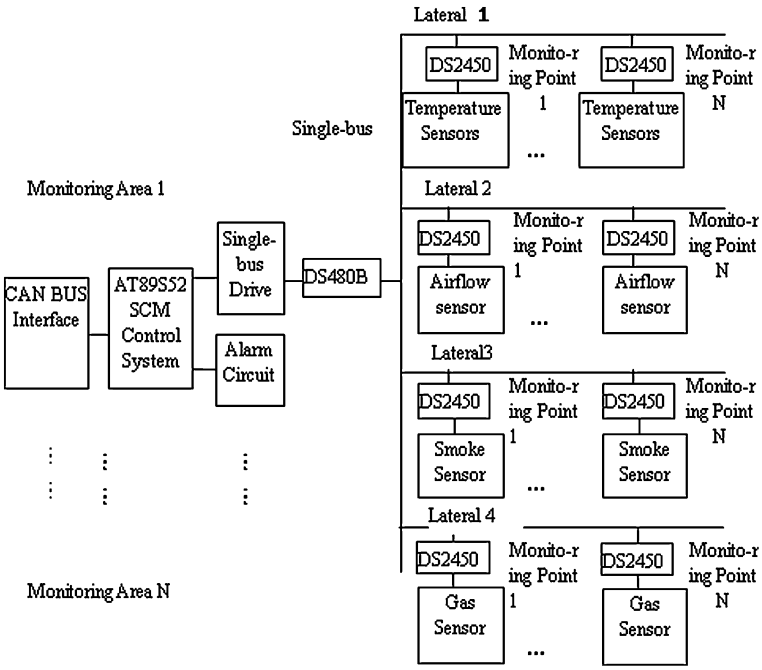


Fig. 44.4 The basic structure of the data acquisition and alarm unit

danger emit automatic alarm at the same time. The basic structure is shown in Fig. 44.4.

As Fig. 44.4 shows, the core of the lower computer is the AT89s52 SCM. It can communicate with the higher computer by CAN bus interface while using a single-bus line to connect with multi-sensor data acquisition system in order to achieve the real-time monitoring of the building temperature, air speed and direction,

smoke concentration and other important information. Single-bus lines are address lines, data lines and control lines together as a signal line that can complete data exchange, control functions etc. Single-bus lines allow the devices which connect with it release the bus that is through the open-drain or tri-state port achieving the other device free to use bus. Every single bus device can be identified based on chip encode unique serial number. So hundreds of detection devices can be connected in a single bus, which is conducive to building a monitoring regional data collection network. At the same time the analog signal can be digitized before transmission by the chip at the detecting point to improve the system's noise immunity and reliability [3].

The basic job is the field data is collected by various types of sensors in the field and it is converted to the digital signal by DS2450 and sent to a single bus. According to master-slave structure principle, only when a single-bus device is called by SCM system, a single-bus device can respond to it. And data is sent to SCM by serial interface single-bus line drive DS2480B. After the finial disposal of information by SCM, all data will be transmitted by CAN bus to the higher computer for subsequent analysis and processing. When the system detects at risk, data is sent to in time the higher computer and at the same time audible warning signal is emitted immediately so that construction personnel can promptly respond.

44.3.3 Fire Prevention System Based on Expert System Design

In order to make fire analyze system accurately judge on if the fire occurred, the trends and direction of fire, current security evacuation routes etc. and then according to Case-Based Reasoning, CBR and Rule-Based Reasoning, RBR, model of high-rise building fire prevention and control based on expert system is established as shown in Fig. 44.5. The basic principle is Case-based reasoning. Then people past experience is saved to be the case base. When new problems arise, a suitable deal with the problem of the case can be found in the case library as a new reference that is experience reuse. If the case is not very appropriate, changes are made to settle current problems. The case will be modified into the case base for the next use which implements self-learning experience. The character is the knowledge expressed by rules can be used to describe the problem and deploy the solution. Then the proven valuable case will be saved into the case base for reasoning in order to achieve a combination of CBR and RBR [4, 5].

Through this self-learning expert system to continuously update and enrich the database of experts for improving the expert system to determine the accuracy of the high-rise building fire, at the same time based on real-time acquisition of critical data on fire, the fire trends and direction can be assessed and judged to make the best evacuation route. Finally by the radio system and the lower computer alarm unit to guide the person in the high-rise building safely evacuating as soon as possible for reducing the fire loss of property and personnel [6].

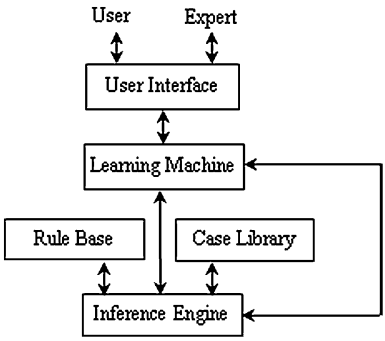


Fig. 44.5 The model of high-rise building fire prevention and control based on expert system

44.4 Conclusion

The supervision control system of high-rise building fire is a kind of integrated detection technology and the key of the high-rise building fire forecasting accuracy is the monitoring result, which has great relations with the property and personal safety of the residents and staff in high-rise buildings. So a kind of high-rise building fire monitoring system based on expert system is designed in this paper. The system is designed to be a distributed bus structure and the building can be real-time monitored. While the CAN-BUS structure can reduce the network's building and maintenance costs, it improves the system's stability, reliability and anti-jamming capabilities and makes the supervision control system of the high-rise building fire develop into digital, network-based, real-time supervision system.

References

1. Fu M, Feng L (2006) Realization and application of the data acquisition system based on CAN bus. Mod Electron Tech 2:71–72,75
2. Li P (2007) Design of the CAN communication in data collecting system. Comput Knowl Technol (Academic Exchange) 13:15–16
3. Huang Y, Zhang L, Xu Y (2007) Discussion on the hardware design solutions for intelligent node based on CAN bus. Instrum Stand Metrol 2:29–32, 44
4. Qin Y, Wang J (2008) Alarm and control linkage for fires based on fuzzy expert system. Control Autom 24:117–119
5. Liu S, Liu W (2009) Fire risk evaluation system based on analysis of ability and frangibility. Sci Technol Rev 22:76–80
6. Zong H, Zong C (2010) Research and application on NFE model of multi-sensor information fusion. J Electron Inf Technol 3:522–527

Chapter 45
Research on Actuator Fault Health Management Method Based on Fuzzy Cluster Analysis

Zhiyi Huang, Weiguo Zhang, Liyuan Sun and Xiaoxiong Liu

Abstract The equipment safety and reliability is improved by used health management technology. Considered the characteristics of the actuator health management, a health management method based on fuzzy cluster analysis is proposed in the flight control system. First, an actuator fault parameters model is created according to the characteristics of electric steering actuator. Then, the fault feature extraction and health situation assessment are analysed by applying fuzzy C-means clustering algorithm. Health prognostic is processed by the development of health situation to divide the health degree and health assessment results to apply the exponential smoothing method. The simulation results demonstrate the proposed algorithm is effective for the system.

Keywords Health management · Fuzzy cluster · Electric steering actuator · Feature extraction

45.1 Introduction

Health management of the equipment is a method which not only detects and monitors health situation of the system, but also takes appropriate measures to maintain the equipment perform or the ability of safe operation. The device is diagnosed and maintained by adopting the corresponding fault diagnosis and prediction method for improving the system reliability [1–3]. The classification method of failure pattern is an effective method to solve the complex device state

Z. Huang (✉) · W. Zhang · L. Sun · X. Liu
College of Automation, Northwestern Polytechnical University, xi'an, China
e-mail: yizhi_lz@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_45,

detection and failure diagnosis. When Health management method based on fuzzy cluster analysis is applied [4, 5], we do not need to establish the accurate mathematical model of the system, and it is based on the statistics of pattern recognition theory and the learning. It has an advantage in turning the data of high-dimensional noise into the low-dimensional information for diagnosis and prognostic. By using the traditional cluster analysis, each object is strictly divided into a class. However, in fact most objects do not have strict properties. So the actual relationship of object and category cannot be really reflected by the classification. This phenomenon can be overcome by the method of fuzzy clustering analysis [4, 5].

In summary, system fault feature extraction and analysis are mostly obtained by the clustering method in the above research achievements. This paper presents a health management method based on fuzzy cluster technology of the actuator in the flight control system. The actuator of fault feature extraction, health situation assessment and health prognosis are analysed in the flight control system. The Simulation results demonstrate the proposed algorithms are effective.

45.2 The Problem Statement

Flight control system actuators are made up of the basic servo motor control system. Its working principal, in the role of voltage controlled by motor, meets the armature voltage balance equation, the electromagnetic torque equation, back-EMF (Electromotive Force) equation and motor torque balance equation. The mathematical model of electric steering gear are constituted by the above four equations and control surface load torque equation. The nonlinear state equations that describe the input voltage u and the servo motor response θ_m are as follows

$$\begin{cases} u = E_b + i_a R_a + L_a \frac{\mathrm{d}i_a}{\mathrm{d}t} \\ E_b = K_E \frac{\mathrm{d}\theta_m}{\mathrm{d}t} \\ T_M = K_T i_a \\ T_B = T_M - T_L = J_G \frac{\mathrm{d}^2\theta_m}{\mathrm{d}t^2} \end{cases} \tag{45.1}$$

Here, L_a is inductance (h); R_a is resistance (Ω); T_M is motor torque $(N \cdot m)$; T_L is load torque converted to motor shaft $(N \cdot m)$; T_B is acceleration torque $(N \cdot m)$; J_G is total moment of inertia converted to motor shaft $(kg \cdot m^2)$; K_E is back electromotive force coefficient $(V \cdot s/\mathrm{rad})$; K_T is torque coefficient $(N \cdot m/A)$.

Through mathematical derivation, in the no-load conditions $T_L = 0$, the transfer function of the servo motor is follows.

$$\frac{\theta_m(s)}{u(s)} = \frac{K_T}{s(J_G L_a s^2 + J_G R_a s + K_T K_E} \tag{45.2}$$

Mathematical model of actuator expresses its physical properties. The change of model parameter directly is related to some certain physical state of actuator.

The fault of actuators performs in the feature parameters of system and feature extraction proceeds according to the input/output data of system. This paper mainly discusses the following types of actuator failure mode: the transmission gear wear, the motor demagnetization, and the motor speed instability and actuator mechanical damage. The above four failure modes are corresponding to these parameters: the motor torque T_M, inductance L_a, acceleration torque, moment of inertia J_G, load torque T_L.

Health management for the equipment is the fault diagnosis and health situation assessment. The good effect is obtained by using the pattern classification algorithm. The so-called pattern recognition model is the right mapping from the pattern space to category membership space. By using the discriminator function, the basic method of pattern recognition is to divide each category in the traditional pattern recognition technology. Normally, the fault pattern recognition is realised by choosing sensitive characteristics to simply compare. It is easy to achieve for the simple system. However, for the complex systems and phenomenon, it involves the failure mode and the problem of the normal pattern identification. Clustering analysis method is an effective algorithm for fault diagnosis and feature selection, which can carry out health monitoring and health management of the equipment.

45.3 Fuzzy C-Means Clustering Algorithm

Fuzzy C-means clustering algorithm (FCM) is a constrained nonlinear programing problem which sums up the clustering and obtains the fuzzy partition of dataset by an optimisation solution. The basic idea is to realise approaching clustering by repeatedly modifying the cluster centre and membership matrix. It divides the n vectors $x_i (i = 1, 2, \ldots, n)$ into C category $(2 \leq c \leq n)$ by means of iteration. This algorithm aims at minimising an objective function by ensuring the cluster centre $V = \{c_1, c_2, \ldots, c_c\}$. The degree that each data point belongs to a cluster is confirmed by adopting the membership $u_{ij} (i = 1, 2, \ldots, c,\ j = 1, 2, \ldots, n)$ in FCM algorithm. $U = (u_{ij})_{c \times n}$ is the membership matrix. Where, n_{ij} presents the degree of membership that the jth sample date is assigned to the ith category and it meets the condition

$$u_{ij} \in [0, 1], \quad \sum_{i=1}^{c} u_{ij} = 1 \tag{45.3}$$

The objective function of FCM is

$$J(U, V) = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{m} d_{ij}^{2} \tag{45.4}$$

where, c_i is the cluster centre of ambiguity group, $d_{ij} = \|ci - xj\|$ is Euclidean distance between the ith cluster centre and the jth data point. $m \in [1, \infty)$ is a fuzzy weighted exponential.

When it meets function (45.3), FCM algorithm is the cluster results, which makes the objective function has a minimum J. The conditional extremum J can be given by Lagrange multiplier method, and the iterative formula of cluster centre and the membership follows.

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{45.5}$$

$$\begin{cases} u_{ij} = \dfrac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{2/(m-1)}} & \text{when} \quad I_j = \phi \\ \forall i \in \overline{I}_j, u_{ij} = 0, \sum\limits_{i \in I_j} u_{ij} = 1 & \text{when} \quad I_j \neq \phi \end{cases} \tag{45.6}$$

FCM algorithm aims at looking for the minimum extremum point through the iteration. So, the main idea of the FCM clustering algorithm is to adjust the (U, V) value for minimizing the value J of object function.

45.4 The Actuator Health Management

According to the feature of actuator researched in this paper, the health management method based on fuzzy clustering technique contains four basic processes which are data acquisition, health feature extraction, health assessment, the remaining life prediction. Its basic design is composed of the following steps:

45.4.1 Data Acquisition

For designing the health management system based on data driven, a mount of dataset related to the actuator in the flight control system are required. The dataset includes the data of actuator normally working during the flight course and the operation date of actuator in the different level of relegation. This paper builds the simulation model of the actuator and adds different levels of failure according to preset failure modes, then the simulation data under different levels of failure can be accessed. According to the description of the system failure in above paragraphs, the four patterns are assigned respectively, which are the system working normally, piston mechanical damage, hydraulic cylinder oil leakage, and servo value failure. Its corresponding fault feature parameter is $F_{\text{fault}} = (T_M, L_a, T_B, J_G, T_L)$. Collect and discuss the data of four patterns. In terms of the formula (45.6), all the data of a data vector are averaged by its mean value

for pretreating the collected data, and then the new data vector with unified dimension is obtained.

$$x_i' = \frac{x_i}{\bar{x}_i} = \left\{ \frac{x_{i1}}{\bar{x}_i}, \frac{x_{i2}}{\bar{x}_i}, \ldots, \frac{x_{in}}{\bar{x}_i} \right\} \tag{45.7}$$

where, $\bar{x}$ is the mean value of data vector x, $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} |xi|$.

45.4.2 Fault Feature Extraction

Feature extraction is a general step in the system monitoring and health management. It is the core of forecast reasoning technique based on data driven. After obtaining the system state characteristics, the extracted eigenvalues map to the corresponding description of system health situation by adopting the clustering algorithm. In this paper, the feature data of known history fault are $F_{\text{fault}} = (T_M, L_a, T_B, J_G, T_L)$. They are the detected physical quantity, respectively. The feature sets of data clustering are composed of these feature data and the corresponding fault types and levels.

Considering the size of processed data in the health management, fault feature extraction is made by adopting dynamic k-means clustering algorithm which can cluster the system fault feature preferably in this paper. The concrete operational steps are shown.

Step 1: choose the number of clusters c, the weighted exponential m, iterations n, the error threshold ε and initialized clustering centre c_i;
Step 2: calculate the membership value according to formula (45.6);
Step 3: update the clustering centre value according to formula (45.5);
Step 4: calculate the error value $e = \sum_{i=1}^{k} \left\| c_{i,n+1} - c_{i,n} \right\|$. If $e < \varepsilon$, stop the iteration and output the result. Or else $n = n + 1$, turn to step 2;
Step 5: according to the acquired clustering centre and similarity function value, test the test-data to judge the collected data belonging to some category. That is the classification of failure pattern.

45.4.3 Health Assessment

It is known that the data vector belongs to that fault type according to the results of feature extraction. But if we want to make the health assessment in detail, the further analytical investigation is needed for the data.

First, the distributions of health feature in its characteristic parameters space are ensured which correspond with the developing system process under the normal situation and each known failure F_D (for example, 10% failure develops to 100%

complete failure). And these distributions are preconfirmed by the empirical data. Then, the collected data and empirical data are compared to ensure the health situation in the present moment system. So the relationship between each feature of learning algorithm and the system health situation must be applied. A type of attached notion should be adopted owning to the continuous process in which the failure takes place and develops. Health degree of the system is divided correctly and then the health assessment to the system is accomplished.

For the corresponding design in this paper, health assessment is the assessments which take the data of the known failure development process as sample characteristics, and then do the fuzzy C-means cluster analysis using the collected data and empirical data to get the membership function. Finally, the health status assessment is measured by adopting the principle of membership.

45.4.4 Health Prognostic

The purpose of health prognostic is to predict future health status of the system, and to ensure the time from the current system health status to the completely ineffective performance. As the system performance are continuously relegating, the similarity between the distribution characteristics in the current system status and failure distribution area (it indicates the system performance failure) is more and more higher. In the system feature space, the time form the current system health status to completely ineffective performance can be ensured. First, the statistical trend analysis method is used and the changing path and tendency of each characteristic parameter are tracked and extended. Then, the changing tendency of each characteristic are merged in the feature space.

When the system health prognostic is processed, the exponential smoothing method can be adopted to track and predict the future process of system characteristic parameters in the feature space [6]. A weighted average index function is used in Double exponential smoothing method. In the function, predict the future value on the grounds of the past prediction which is weighted by a exponential function. The basic expression of double exponential smoothing method is as follows.

$$S_t = \alpha y_t + (1-\alpha)(S_{t-1} + b_{t-1}) \quad 0 \le \alpha \le 1 \tag{45.8}$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \quad 0 \le \gamma \le 1 \tag{45.9}$$

$$F_{t+m} = S_t + mb_t \tag{45.10}$$

where, α is smoothing index, S_t is the smoothing value acquired at the moment of t, b_t is the changing trend value before the moment t, F_{t+1} is the trend value of single step, F_{t+m} is the predicted value in the step of m.

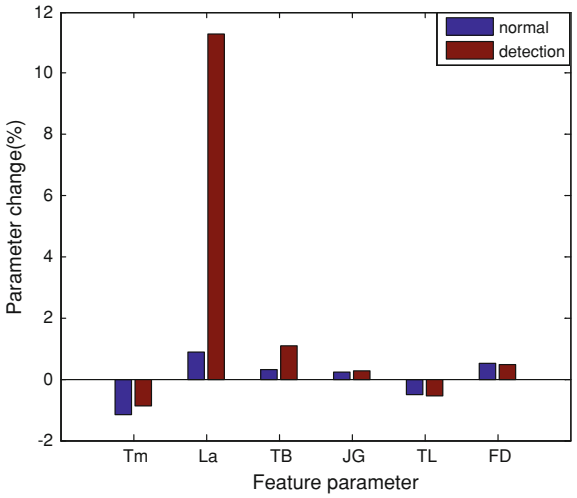


Fig. 45.1 Graph of health assessment result

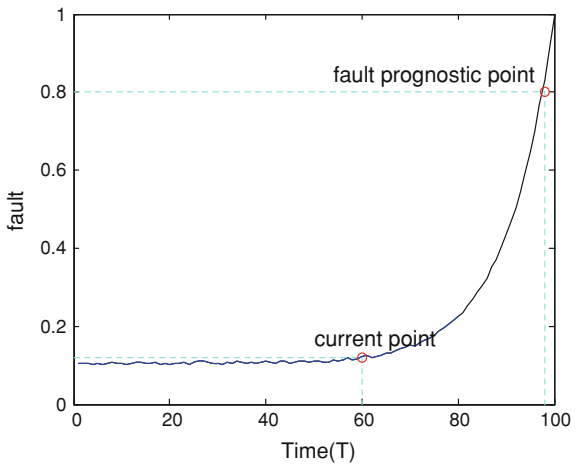


Fig. 45.2 The graph of prognostic result

Suppose the current time is t, and system health feature development after the moment of t is predicted by using Double exponential smoothing method. The steps of development trend are:

Step 1: for characteristic parameter $F_{\text{fault}} = (L_a, T_M, T_B, J_G, T_L)$, set the initial value b_1 and S_1. Set S_1 to the actual value of F_{fault} in the initial time, $b_1 = [(y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3)]/3$. Where, y_j is the actual value of F_{fault} in the jth time;

Step 2: calculate the smoothing value $S_k, (k = 2, 3, \ldots)$ of characteristic parameter F_{fault} at the moment of k, according to the formula (45.8). Its purpose is to eliminate the lag;

Step 3: update the tendency value of characteristic parameter F_{fault} at the moment of K, according to the formula (45.9);
Step 4: calculate the predicted value F_{k+m} of characteristic parameter F_{fault} at the moment of $k + m$;
Step 5: $i = i + 1$, predict the characteristic parameter F_{fault}, and turn to step 1;
Step 6: calculate all the predicted value of characteristic parameter. Then predict the development of characteristic in the feature space composed of characteristic parameter. According to the predicted characteristic trend, apply the health assessment method and calculate the future health status of system until characteristic parameters come to the fault characteristic region. In other words, the system is completely ineffective. The difference value between the final and the tune k is the remaining useful life of system.

45.5 Simulation

According to the proposed algorithm above, we analyse the simulation by using the simulation model. First, the fault type of motor demagnetization is set to simulate. The simulation results are shown in Figs. 45.1 and 45.2. The parameter change of the inductance L_a at monitoring point is maximum as Fig. 45.1. The fault of motor demagnetization can be drawn and the health degree is 0.15. Simultaneously, the residual useful life could be predicted by applying the exponential smoothing method. It is can be seen that the residual useful life of system is about 40 h as shown in Fig. 45.2.

45.6 Conclusion

For improving the actuator reliability in the flight control system, the method of actuators health management is studied by applying fuzzy clustering method. First, the fault feature and health situation parameters of actuators are drawn. Second, failure diagnosis is finished based on health feature and the health degree is divided based on health situation feature. Finally, the remaining life is predicted by adopting the double exponential smoothing method. Simulation results demonstrate the proposed algorithms are effective to realise the system health feature extraction, the situation monitoring and the remaining life prediction.

Acknowledgments This work is supported by the Aeronautical Science Foundation of China (No. 20100753009).

References

1. Johnsoni SB (2011) System health management theory and design strategies, Infotech@Aerospace. St. Louis, Missouri
2. Balaban E (2011) Experimental validation of a prognostic health management system for electro-mechanical actuators, AIAA, Infotech@Aerospace. St. Louis, Missouri
3. Schwabacher MA (2005) A survey of data-driven prognostics, AIAA, Infotech@Aerospace, Arlington, Virginia
4. Amin S, Byington C (2005) Fuzzy inference and fusion for health state diagnosis of hydraulic pumps and motors, Fuzzy Information Processing Society Annual Meeting of the North American
5. Iverson DL (2004) Inductive system health monitoring. In: Proceedings of the international conference on artificial intelligence
6. Matthew W, Carl B (2004) Dynamic modeling and wear-based remaining useful life prediction of high power clutch systems. In: Proceedings of ASME/STLE international joint tribology conference long beach, California, USA

Chapter 46
Design of Intelligent Protection and Control Device for Medium and Small-Scale Asynchronous Motor

Dao-lin Li, Zhi-qiang Wu, Xiang-yang Chen, Jun-gang Li and Yong Wei

Abstract This chapter research the basic operation principle of asynchronous motor, perform an in-depth analysis of the characteristics of asynchronous motors of all kinds of faults occurred and the quantity relationship between the fault reason, to establish judgment of the stator current positive sequence, negative sequence, zero sequence components or their combination primarily, by detecting voltage, thermal resistance, external failure of complementary, multiple protective criteria, comprehensive protection scheme induction motor. And to motor working state of recognition, heat and cooling process unified heat capacity, the calculation formulas of zero sequence of protective grounding prevent abnormal operation measures to deal with the problem are discussed, and proposed a novel and simple and practical solutions.

Keywords Medium small-scale motor · Relay protection and control device · Power system relay protection · Hardware design · Software design

D. Li · X. Chen
Three Gorges Vocational College of Electric Power,
YiChang 443000, Hubei, China

Z. Wu
Shandong Electric Power Staff Skills Training Center,
JiNan 271000, Shandong, china

J. Li · Y. Wei (✉)
Xuji Group Corporation Technical Center,
XuChang 461000, Henan, china
e-mail: yongw@xjgc.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_46,

46.1 Introduction

Motor more in high temperature, high humidity, dusty working environment and its faults form also has a variety of, by and large, can be divided into electric breakdown and mechanical failure two aspects. According to statistics, production is used on three-phase asynchronous motor, the malfunction of electrical failure of winding damaged about 85%, machine shaft and other mechanical failure accounted for about 15%, most electric breakdown and mechanical failure, eventually reflected in motor on each sequence current. This device is mainly based on symmetrical component method, to positive, negative, zero sequence components or their combination according to various protections setting motor.

46.2 Relay Protection Scheme of Asynchronous Motor Comprehensive Protection

Protection principle is the device software and hardware design basis (especially implement software), hardware and software is realizing various protection control functions of means. Clarity, perfecting the principle, software and hardware design view, made only when product can fully functional, reliable performance. Motor microcomputer protection mainly through measuring power (current, voltage and switch state, etc.) to monitor the operation status of motor. Electric power industry standard DL/T 11.2-2001 provisions asynchronous motor microcomputer relay protection device should be installed to the basic protection are: negative sequence current (imbalances) and zero sequence current (single-phase grounding), differential, electric current instantaneous break, positive sequence currents (plugging turn), stator winding over load, thermal overload, pressure loss/under voltage etc. In addition to these standards of basic protection function, in an overall consideration which causes the fault of the main factors, on the basis of this device also set the start time is too long, over-voltage, owe current, external non-electric quantity failure protection, to have embedded motor, still can through to thermostats resistance measurements of internal temperature monitoring motors directly. The main protection as follows:

46.2.1 Negative Sequence Current Protection (Imbalances)

Negative sequence current protection used as motor broken phase, sequence, faulty turn-to-turn short circuit and voltage supply large unbalance protection, to motor asymmetric short-circuit fault also plays backup protection. The traditional so-called “current imbalance protection” is often used to current differential principle, namely the three-phase current amplitude attributes of the difference between

the three and three-phase I train I_{ave} average ratio of train I/I_{ave} as protection movement of criterion. After some domestic subscribers in the operation experience, generally reflected the method setting difficulties, protect performance is poor, often mal operation or refusing action. But with negative sequence current as motor various imbalances fault protection, negative sequence current setting value can according to the operation environment of possible imbalance fault type and fault point position through calculating conveniently [1], can play better protection. Negative sequence current protection must escapes from the scope of the permit of the network voltage caused by an imbalance motor negative sequence current. If the device of external alternate with no input, short-circuit device action time limit to the time limit and switch off the external fault suitably, to ensure the selectivity.

46.2.2 Positive Sequence Over-Current Protection (Plugging Turn)

Motor over-load operation, especially positive happened plugging turn sequence current will increase sharply, easy to cause the motor burnt. This device set time limit positive sequence over-current protection. Current action is obtained commonly take action time limit, shall be escaped the motor start-up time, it is desirable since. Due to the general starting current is normal, can pass setting value items, in starting to set a time delay, make from start to end, start time delay over-current protection function can be set to quit, infinite, namely the starting process sequence over-current protection has been exit, enter only after running status input. With the above tripping set values are similar, positive sequence over-current protection and a warning setting value, a slightly different is: in the whole time, alarm set values of operating current general lesser [2], action time also set a shorter, even to zero, in this case once overload immediately report to the police.

46.2.3 Stator Winding Overload Protection

Stator current protection can be fixed over-load time limit properties, also can be used for inverse time characteristics. In most Chinese products adopt fixed time limit, protect the criterion is usually stator current base wave value, useful also positive sequence over-current protection and replace over load of protection. Foreign product usually adopts the inverse time. In comparison, inverse time can better with motor thermal characteristics of the match, so this device also chooses the latter. This device stored the 12 over-load curve, including four accord with NEMA standards (see appendix). Current over-load multiples smaller, action the change of time, the greater the in order to occupy storage area below certain condition, ensure the accuracy and the action time with current over-load multiples increases, reference point of selecting also by secret to Succoth.

Users can provide the motor manufacturer according to over-load capacity curve data choice appropriate protective curve, it's just below the motor to protect curve damage curve. If motor manufacturer cannot provide the corresponding data, but by plugging turned current and allow plugging turn time to choose protect curve, the method is in the corresponding over-load curve by plugging in trace turn current and allow plugging turn time determined the point, at this point below the closest curve is to compare appropriate protective curve.

46.3 Motor Working Status Identification

Motors during start-up and run-time every protection of fixed value or retreating often have different, so accurate to motor real-time state judge, directly related to the performance of motor protecting. Previous motor protecting device most of the current in a relatively short period from zero to a more than rated mutations k_{IN}, as a result, the starting signal of the beginning of the current descends to near slightly higher than the average rating (rated) as starting over, into the operation state of the sign, this in soft start-up etc. step-down start easily when appear misjudgment, such as variable frequency soft start-up of starting current may have been no more than k_{IN} [3], and as a multi-stage string resistance starter, current may repeatedly over k_{IN}, also may have been no more than k_{IN}. This device fully consider all sorts of start-up mode under the current changes over time characteristics with current and time of combining the judgment method, can real-time accurate detection motor downtime, startup, running three states, thereby ensuring motor various protection of sensitivity and reliability. State criterion is as follows:

Three-phase current I_{ave} less than average power flow up/threshold $I_{stop/start}$ (2%) set IN the time delay t_{stop}—more than stop judging recognition, the stop condition, from the starting to the next stop to monitoring the time for downtime t_{stop}. $T_{stop\text{-}recognition}$ is set in order to avoid starting or operation process of the current of a short suspension, such as star/triangle starting and self coupling decompression starting if adopting open circuit is designed to convert, switch to another level from level 1 when the power is points epigastria connect. Again,

Before a state for downtime, three-phase current average more than $I_{stop/start}$ I_{ave} is starting begin, Device testing to stop or running state is starting over. Starting began to start at the end of the time for starting time t_{start}.

At the same time satisfy the conditions:

1. starting confirm time $t_{start\ recognition}$ has passed;
2. three-phase currents in $I_{stop/start}$ and thermal overload protection movement between current k_{IN}, it is running state. $T_{start\text{-}recognition}$ is to prevent the motor is premature convicted of operation mode of the setting, it from the current more than $I_{stop/start}$ timing starts, the duration of $t_{start\text{-}recognition}$ to ensure that timing is over current has reached the starting process of current peak (Fig. 46.1).

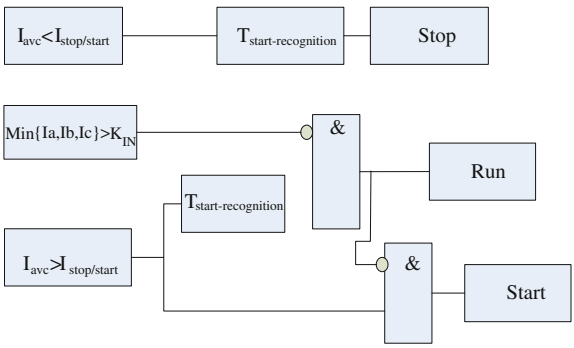


Fig. 46.1 Motor working condition judgment diagram

46.4 Hardware System Design

This device hardware system is mainly composed of power system, DSP smallest system, switch input, switching output, analog inputs, current loop output, NTC thermostats measurement circuit module, communication etc. Respectively at the motherboard, arrangement of power on the board, communication mode out heat resistance measurements plate and transformers relay board on four blocks PCB. Hardware system diagram is shown in Fig. 46.2. In this figure the use of various noted power (power supply system, and discusses) refer to flow direction arrow representative signal.

Power system and power protection by module filter circuits, low pressure drop linear regulator (LDO).quasi DC/DC power components, isolation as shown. This device for the job by ac, DC 220 V power 110 V or exchange control circuit of 220 V provide, the protection filter circuit input module power, module power supply output voltage DC 5 V LDO again from 3.3 V output voltage to 5 V, most of the device on the main board supply voltage is 3.3 V, only DAC, reach linton drive array minority device adopts 5 V power and like communication, current loop output, NTC thermistors measurements all direct and external interface of part, the reliability of the consideration, requirements for internal and external cannot have circuit connected directly, so not only to use all kinds of photoelectric couplers isolation digital and analog signal [4], and inside and outside of the power supply circuit also want isolation, isolating DC/DC converter role lies in. Display module themselves have no power, its power source for controller ontology 485 communication from the isolation of power supply, the same mouth, display module also with a isolation DC/DC converter, which shows that it is the communication and circuit inside the separated. This display modules and control the power cord no direct connection, plug and play, user-friendly.

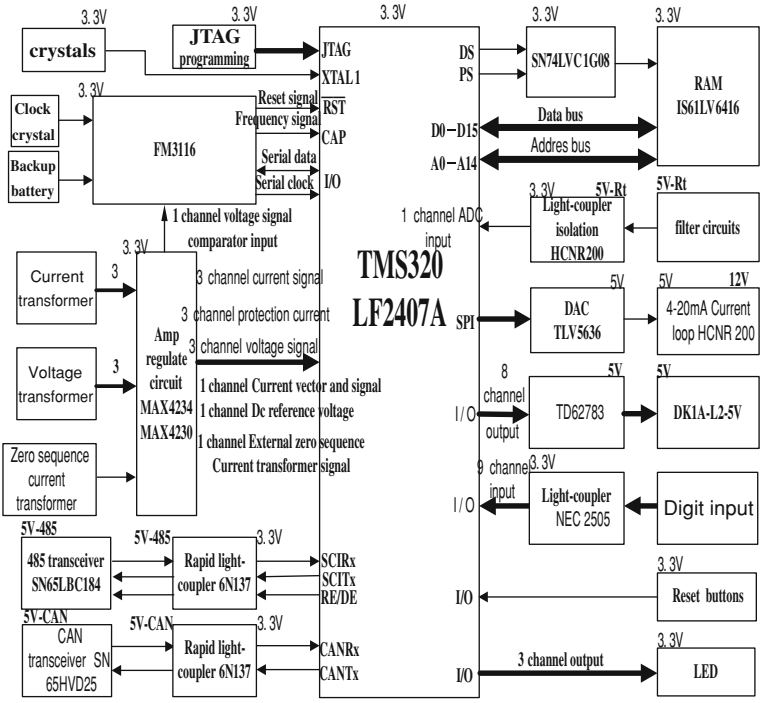


Fig. 46.2 Substation monitoring layer system structure

46.5 Software System Design

Software according to function can be divided into two kinds: one is the execution software, it finishes all kinds of substantive features, such as measurement, calculation, output control, communication, etc.; another kind is monitoring software. It is dedicated to coordinate the implementation module and operator relationship in system software in the middle of a group scheduling role. Implement software design, the emphasis on the efficiency of the algorithm to solve the problem that clear pure, so relatively easy programming. While monitoring software to the overall situation, and more, need to have strict clue of logic (Fig. 46.3).

This device's function modules, among various modules to exchange information and share resources are relatively less, miniaturization request was limited this device hardware resources might not have too much. And device is very strict in real-time, fault diagnosis must accurately and rapidly. This chapter foreground-background system and traditional RTOS their respective advantages, based on the experience in had improved, and designed a kind of simple and applicable for this device's software system, in foreground-background system of introducing RTOS real-time task scheduling mechanism. According to each function of real-time demand different tasks into three categories: the internal and external tasks, task

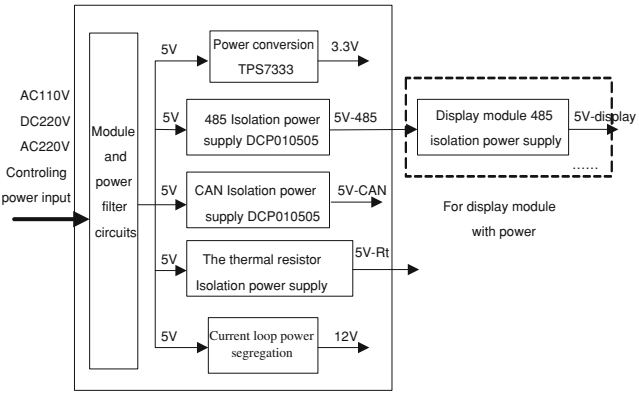


Fig. 46.3 Power system frame

main circulation task. Internal task in interrupt finish, namely the receptionist procedures, real-time demand supreme; the main loop tasks in the system free on, equivalent to traditional foreground-background system background RTOS application, or the lowest priority task, real-time demand lowest, External task in real-time, in-between, can be considered as the traditional front and backend application in real-time demand function, find out a hierarchy of separate constitute the task, such already make each front desk interrupt service routine as brief and shorten the low priority interrupted the response time, also make the interrupt ended, the corresponding real-time operating system would immediately executed, real-time greatly improved. This software system up to 16 level external task, interrupt and the running task to be able to make some external task is ready, different is, if the interruption of task is ready, because be interrupted has saved into the scene [5], and interrupt end can directly task switching, If in the running task in make other tasks ready, then must through the soft interrupt kept the site again task switching. And restore the scene of the work may be finished in two places: if you do not need to task switching, the interrupt end restores the scene, if any task switching is in the external task execution after recovery site (Fig. 46.4).

This device's hardware interrupt, priority from high to low order is as follows: A/D sampling interrupt, panel communication receiving (SCI receive) interrupt, panel communication sent (SCI sent) interrupt, PC communications (CAN email) interrupt, PC communications error (CAN errors) interrupt, hardware interrupt frequency measurement. Besides the one task switching soft interrupt, its lowest priority. Each interrupt after entering first save the site, interrupt nested logo 1, after allowing eraser interrupt as "0" (to ensure timely handling internal task, A/D sampling interrupt not clear INTM), allowing interrupt nested, then carries on the corresponding interrupt handling. After this process shut disruption, the interrupt nested logo minus 1, according to interrupt nested without interruption nested symbol judgment. Existing interrupt nested (interrupt nested logo is not zero) does not do task switching, restore live open after interrupting return low priority

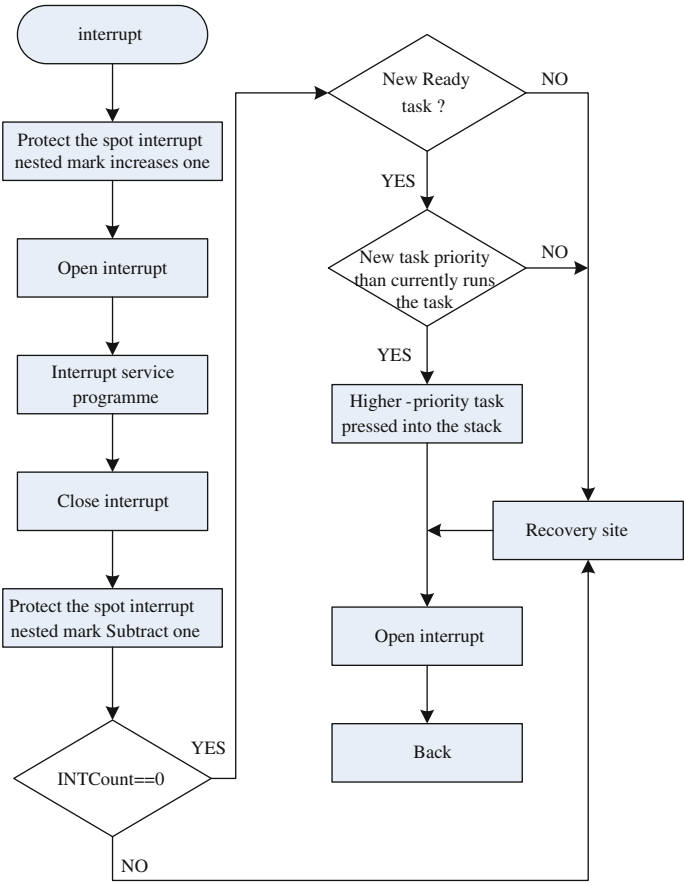


Fig. 46.4 Flow diagram

interrupt [6], Do not exist interrupt nested, then decide whether it is necessary to task switching, if needed on the switch to higher-priority task, do not need to return the interrupted task. Interrupt handling procedures as shown in the diagram.

46.6 Conclusion

This chapter focus on small and medium asynchronous motor protection and control problems on research and development, analyzes the asynchronous motor protection principle, software and hardware design, completed device functions and resisting flexibility test, based on the current motor protecting control device

are analyzed and compared, in microcomputer motor protecting functions are perfect, make motor can get a more comprehensive and reliable protection.

References

1. Xu FR (1999) AC indution motor starting control energy-saving protection control review. IEEE Trans Tech Mag 21(3):1–5
2. Liu LG (2003) Asynchronous motor starting control energy-saving protection control the development situation of [J]. J Coal Tech 22(8):29–30
3. James K (2002) Analysis of the Sallen-Key Architecture (Rev B)[J/OL]. http://focus.ti.com/docs/apps/catalog/resources/appnoteabstract.jhtml?abstractName=sloa024b,2002-9
4. Wang RW, Song ZX, Yang W (2003) Electrical intelligent principle and application. Publishing house of electronics industry, Beijing, pp 45–66 92–107, 141–144
5. Liu ZX, Zhang Z (2001) Yin jooping item root, etc. With vector control principle of asynchronous motor on-line monitoring and fault diagnosis method of principle analysis. Addison Wesley AIAA (7):15–18, 2003-0186, 2003
6. Xu BJ (2002) Asynchronous motor winding fault analysis and detection method. PhD Thesis, Baoding north China electric power electricity engineering university

Chapter 47
EMS Investigation of Compound Power Supply for PEV

Shijing Xu

Abstract It is still much to be desired for battery to power alone as energy storage system in pure electric vehicles (PEV). A compound power supply made up by batteries and ultracapacitors (UCs) is proposed and a rule-based energy management strategy (EMS) is designed in order to improve the working efficiency and extend the cycle-life of the battery. Simulation results from ADVISOR show that PEV with compound power supply is more energy efficient and dynamic performance is better compared with PEV with battery only. In addition, the battery current drops greatly, which will help extend the cycle-life of the battery.

Keywords PEV · Compound power supply · EMS · UCs

47.1 Introduction

Vehicles greatly facilitate people's lives. But with more and more vehicles, environmental pollution and energy crisis happen seriously. At present, the world's major automobile manufacturers are developing the energy efficient vehicles. Among them, the pure electric vehicle is an important research direction. In China's "twelfth 5-year plan", pure electric vehicle is also being studied as one of the three key research fields of new energy vehicles.

S. Xu (✉)
School of Electrical Engineering and Automation,
Tianjin University, Tianjin, China
e-mail: hpuxsj@sina.com

S. Xu
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_47,

Batteries are the only energy source in the traditional pure electric vehicles (PEV). However, in spite of providing a significantly high energy density potential, commercially available battery systems present some drawbacks, such as low cycle-life, long recharging time and low power density. Besides, rapidly changing current values may damage the chemical structure of batteries [1]. The ultracapacitors (UCs) can be changed and recharged by a large amount of current in a significantly short time (0.3 s–15 min in a single charging or recharging process), long cycle-life (about one hundred thousand times), high charging and recharging efficient (98%), high regeneration energy efficiency (up to 70% in a traditional brake), working in a wide temperature range (−40 to 75°C), etc. [2]. The characteristics of the batteries and the UCs are complementary, so they are combined as a compound power supply in this paper. Compound power supply can be charged and recharged, reduce the charging and recharging current of batteries. It also can extend the cycle-life of batteries and improve the operating efficiency of the energy storage system. The literatures [3–5] have studied these problems.

ADVISOR is a set of model, data, and script files for use with MATLAB/Simulink. All the models and codes are open to uses and can be modified freely. The original ADVISOR can not simulate PEV with compound power supply, but can simulate it after redeveloping. The literature [4] elaborates the redeveloping method. Energy management strategy (EMS) has a crucial impact on the performances of the PEV. Rule-based EMS and fuzzy-logic-based are commonly used. The literatures [6, 7] proposed that UCs only were charged by the regenerated energy, without recharging the UCs by batteries. In this paper, the UCs are recharged by the batteries in order to improve the dynamic performance while the state of charge (SOC) of the UCs is below the desired value. This paper redevelops ADVISOR, and establishes the rule-based EMS according to the characteristics of the batteries and the UCs.

47.2 Working Modes of Compound Power Supply and the Design of EMS

Energy allocation and flow direction are different on account of different working modes. Analyzing the working modes of compound power supply entirely is the base to design the reasonable EMS.

47.2.1 Working Modes Analysis of Compound Power Supply

(1) UCs and batteries provide the requested power simultaneously while vehicles is accelerated or drive at a high speed and need a big power.
(2) Batteries provide the requested power while that is small.

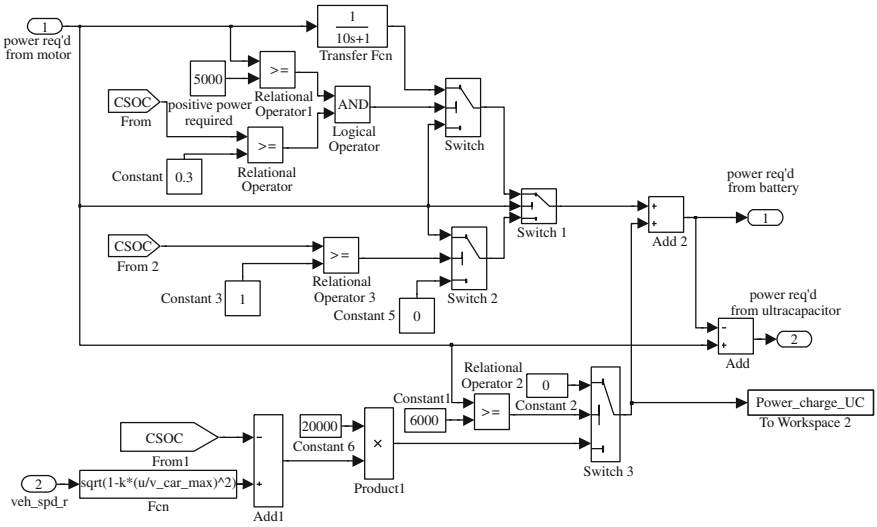


Fig. 47.1 EMS block diagram

(3) The UCs is charged by the regenerated energy firstly while decelerating or braking, then the batteries are charged after the UCs is full.

47.2.2 EMS Design of Compound Power Supply

The high charge–discharge efficiency and high power density should be considered entirely while designing the EMS, which can play the "load shifting" effect of the UCs. The big requested power is controlled by the filter function so that the UCs are allocated the instantaneous load in order to reduce the big current impact for the batteries and balance the output power of the batteries. The EMS is depicted as below:

1. The high-frequency load is distributed to the UCs, the low-frequency load is distributed to the batteries while the requested power is positive and larger than the defined value.
2. If the SOC of the UCs is less than 1, UCs are charged firstly while regenerating. Then, the batteries are charged while the SOC of the UCs reaches 1.
3. The batteries only provide the requested power while that is less than the defined value. In this paper, the defined value is 5 kW.
4. The batteries recharge the UCs while the SOC of the UCs is less than the desired value and the requested power is less than the defined value. Power_charge_UC is the charging power for the batteries in Fig. 47.1.

In order to ensure the vehicle dynamic performance, the UCs are recharged while their SOC is less than the ideal value. Judge as follows: where the variable *soc_uc* is the ideal SOC value of the UCs that is obtained by calculating, and CSOC is the practical SOC value of the UC.

1. While CSOC $\geq$ soc_uc, the UCs can provide energy for motor to drive the vehicle.
2. While CSOC $<$ soc_uc, the UCs need to be recharged for future accelerating. soc_uc is relevant to the vehicle velocity, and the formula is as follows [8]: $\text{soc\_uc} = \frac{V_{\text{cap}}}{V_{\text{cap}}^{\max}} = \sqrt{1 - k\left(\frac{v_{\text{car}}}{v_{\text{car}}^{\max}}\right)^2}$. Where V_{cap} is the actual voltage of the UCs and $V_{\text{cap}}^{\max}$ is the maximum voltage of the UCs. v_{car} is the actual velocity of the vehicle and $v_{\text{car}}^{\max}$ is the maximum velocity of the vehicle. The factor k is the energy utilization coefficient of the UCs, the value is 0.75 in this paper. The EMS based on MATLAB/Simulink of the compound power supply is show in Fig. 47.1.

If the filter time constant is too small in the EMS, the batteries will be allocated a big instantaneous load. If it is too large, the UCs will discharge very quickly. Comprehensively considering and according to the literatures [5], this paper takes 10 s. The charging power from batteries to UCs is defined by a constant multiplying the difference of soc_uc and CSOC. If the constant is too large, the UCs will be prior to surplus charge. If it is too small, the UCs will prior to be short of charge, which the UCs can not play their role fully. In this paper, the constant is 20,000. Whether charging the UCs is controlled by another constant. If the battery output power is larger than the constant, the batteries will not charge the UCs. On the contrary, the batteries will charge the UCs. If the constant is too large, the UCs will be prior to surplus charge. On the contrary, the UCs will be prior to be short of charge. In this paper, it is 6,000 W.

47.3 Simulation Results and Analysis

The vehicle simulation model is established for use with MATLAB/Simulink and the developed EMS is embedded, shown as Fig. 47.2. The induction motor from the company Westinghouse is adopted, which working voltage is 320 V. The battery pack is consist of 25 pieces of 12V25AH lead-acid batteries in series. Generally, if the voltage of the batteries and the UCs is close, the DC/DC efficiency will be high [5]. So, in this paper, the UC bank is consist of 120 pieces of Maxwell PC2500 UCs in series, and the maximum voltage is 300 V. The UC parameters is show as the Table 47.1 [8], and the vehicle parameters is depicted as the Table 47.2. In this paper, the EMS can realize the functionality of the DC/DC, the requested power is distributed to the UCs and batteries by the developed EMS.

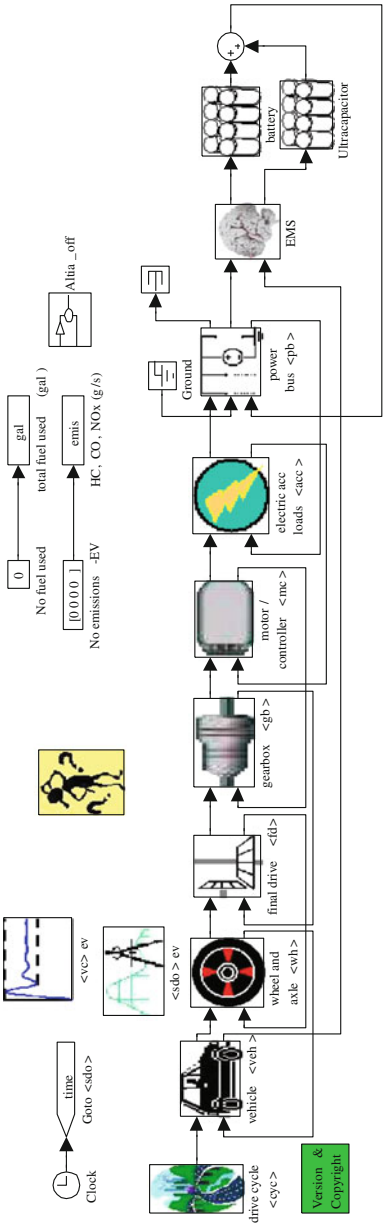


Fig. 47.2 PEV vehicle simulation model

Table 47.1 PC2500 ultracapacitor parameters

Ultracapacitor parameters	Value
Capacitance	2,700 F
Series (DC)	1.0 m Ohm
Voltage	2.5 V
Rated current	635 A
Mass	725 g
Volume	0.6 L
Temperature	−40 to 70°C

Table 47.2 Vehicle parameters

Parameter	Value
Max motor power (kW)	75
Aerodynamic drag coefficient	0.335
Frontal area (m^2)	2.0
Vehicle mass (kg)	1229
Rolling resistance coefficient	0.009
Initial SOC of Bat	1
Initial SOC of UC	1

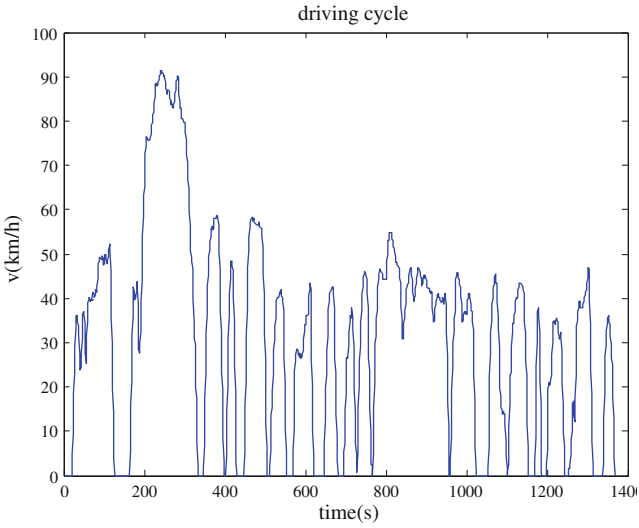


Fig. 47.3 UDDS driving cycle

The simulation is done in ADVISOR, and the urban driving cycle is adopted. The simulation results are shown in (Figs. 47.3–47.5).

It can be seen from Fig. 47.4 that the UCs peak power is larger than that of the batteries, which reduces the peak load of the batteries. It also can be seen that the regenerated power is absorbed by the batteries entirely. Energy recovery efficiency

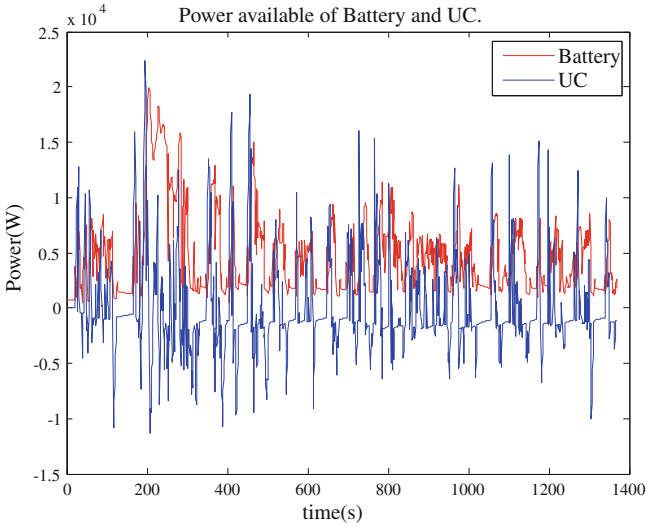


Fig. 47.4 Output power of the batteries and the UCs

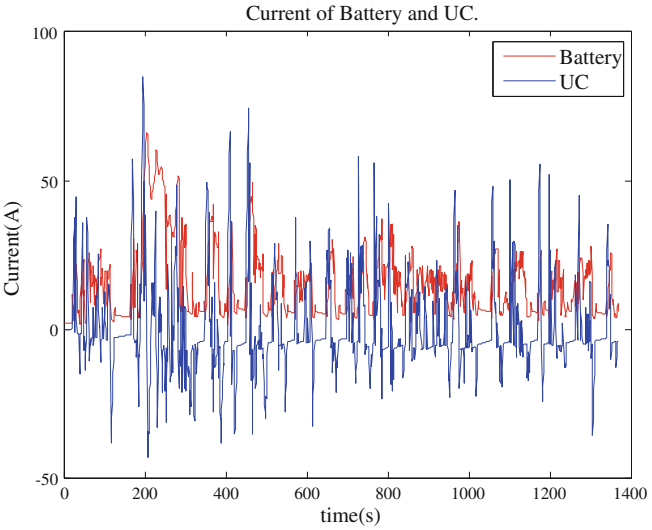


Fig. 47.5 Current of the batteries and the UCs

is higher than batteries only because the internal resistance of the UCs is less than that of the batteries. It is shown in Fig. 47.5 that the maximum current of the batteries is much less than that of the UCs.

While the batteries is the only energy storage system of the PEV, the output power of the batteries track the requested power instantaneously. The output power is provided by the batteries, so any EMS is not needed. The simulation is

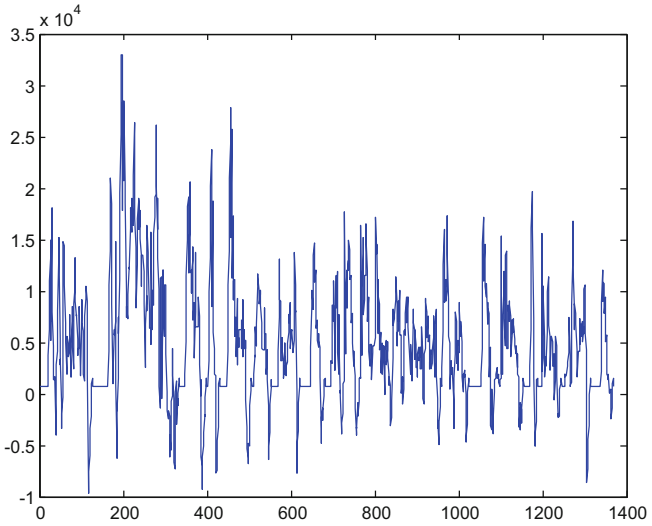


Fig. 47.6 Output power of the batteries

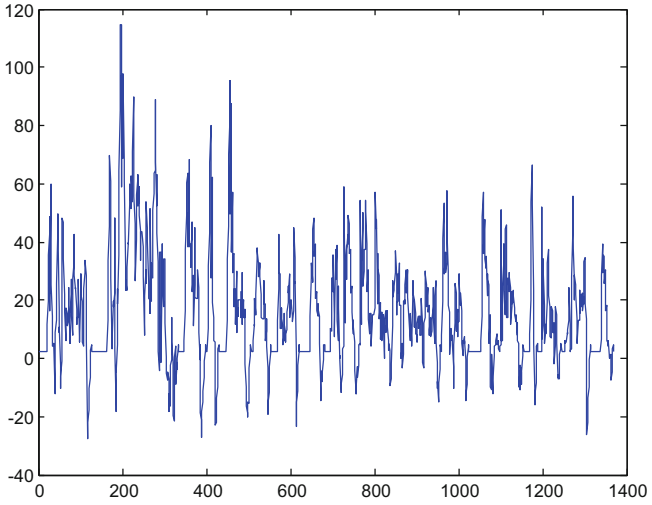


Fig. 47.7 Current of the batteries

done again in the UDDS driving cycle at the same vehicle parameters, and the output power and the current of the batteries are show as Figs. 47.6 and 47.7. The dynamic performance of the vehicle is compared in Table 47.3 with the batteries and with compound power supply. Compared with Figs. 47.5 and 47.7, the battery current of the compound power supply is much larger than that while

Table 47.3 Dynamic performance comparison of between PEV with battery power and compound power supply

Parameters	Battery	Compound power supply	Improvement (%)
0–100 km/h (s)	12	6.9	42.5
Max. accel. (m/s^2)	5	5	0
Max. speed (km/h)	157.6	157.6	0
Distance in 5 s (m)	53.4	61.7	15.5
Energy use (kWh/100 km)	18.86	17.39	7.8

the batteries are as the only energy storage system. According Symons premises [9], the battery cycle-life of the compound power supply is extended.

It is shown in Table 47.3 that the compound power supply has the greatest impact on hundred kph compared with the batteries powering only. but it is no effect for maximum velocity and maximum acceleration. In the UDDS driving cycle, compared with batteries only, energy consumption is reduced 7.8% with compound power supply.

47.4 Conclusions

The simulation results show that the UCs play a "load shifting" effect, and reduce the peak power and current. The compound power supply play the advantages of the batteries and UCs and improve the PEV energy efficiency. Compared with batteries only, in the UDDS driving cycle, compound power supply reduces energy consumption 7.8%. The dynamic performance is improved greatly, among them, hundred kph acceleration time is reduced 42.5%. Moreover, the UCs help extend the battery cycle-life, making up the shortcomings of the battery, such as the low power density and poor low temperature performance.

In order to improve energy recovery efficiency, the literature [10] designed a fuzzy controller to allocate the braking force. In addition, the battery cycle-life is extended because of UCs, the literatures [9, 11] studied the problem. This chapter does not study these problem, but will study later, which helps obtain better overall performance EMS.

References

1. Erdinc O, Vural B, Uzunoglu M (2009) A wavelet-fuzzy logic based energy management strategy for a fuel cell/battery/ultra-capacitor hybrid vehicular power system. J Power Sour 194(1):369–380
2. Dong Y-h, Yan C-l, Zhang J-w et al (2008) Simulation research for power and economy performance of parallel hybrid electric bus equipped with supercapacitor and battery. J Syst Simul 20(23):6529–6534

3. Xi Z, Mi CC, Masrur A et al (2008) Wavelet-transform-based power management of hybrid vehicles with multiple on-board energy sour including fuel cell, battery and ultracapacitor. J Power Sources 185(2):1533–1543
4. Yu Y-b (2004) Study on the parametric design and control strategy of hybrid power system for hybrid electric vehicle. Jilin University, Changchun
5. Li F-w (2006) Study on the parametric design and control strategy of hybrid power system. Jilin University, Changchun
6. Yang T (2010) Investigation and simulation of combined energy storage system on electric vehicle. Chang'an University, Xi'An
7. Shi Q-s (2009) Key technologies research on energy management problems of pure electric vehicles. Shandong University, Jinan
8. Walker AG, Simpson GR (2002) Lifecycle costs of ultracapacitors in electric vehicle applications [C]. Power electronics specialists conference 2002, pp 1015–1020
9. Drouilhet S, Johnson BL (1997) A battery life prediction method for hybrid power applications (preprint) [C]. The 35th AIAA aerospace sciences meeting and exhibit, Reno, Nevada, pp 1–14
10. Zhang C-l, Zhang Y-j, Yan M-d et al (2011) Fuzzy control modeling and simulation of regenerative braking system for pure electric vehicle with dual-source energy storage system. J Syst Simul 23(2):233–238
11. Wei X-z, Dai H-f, Sun Z-c (2007) Selection design of the assistant power batteries for fcevs. Chin J Power Sources 131(10):772–776

Chapter 48
Analysis of the Shortest Path of GPS Vehicle Navigation System Based on Genetic Algorithm

De En, Huanghe Wei, Jinxian Yang, Nana Wei, Xiaoguang Chen and Yangyang Liu

Abstract The shortest path is one of the key problems in the GPS vehicle navigation system. Genetic algorithm is one of the useful methods in solving the shortest path problem. First, the genetic algorithm is analyzed and improved. Its population is optimized before being initialized aiming at reducing the needless paths and nodes. A combination of roulette and elite protecting method is used to protect the population's best individual in each generation. The improved genetic algorithm can advance the convergence rate and the quality of solution in searching the shortest path. The functional demands of GPS vehicle navigation system can be met by using this method.

Keywords GPS · Shortest path · Genetic algorithm · Population optimization · Genetic manipulation

48.1 Introduction

In the 1980s, America, Europe and other developed countries began to develop the digital road map which used in the field of traffic management for the purpose of navigation and query. In the 1990s electronic navigation equipment began to be

D. En (✉) · H. Wei · J. Yang · N. Wei · X. Chen · Y. Liu
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: ende@hpu.edu.cn

H. Wei
e-mail: weihuanghe@yeah.net

J. Yang
e-mail: yangjinxian@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_48,

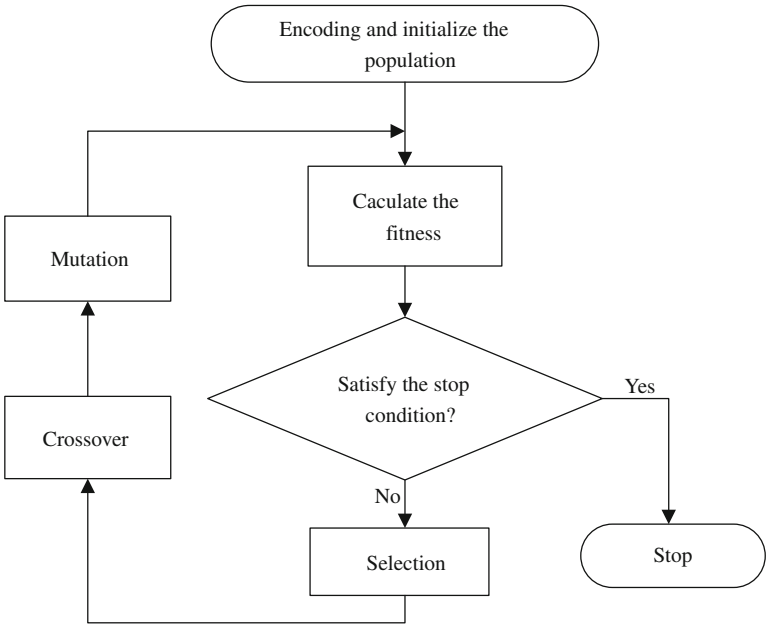


Fig. 48.1 Genetic algorithm's basic process

used in the cars, the users in the city can easily and fast to reach the destination. Selecting and providing the shortest path are the basic functions for the current GPS vehicle navigation system. With the rapid economic development, the advent of the information age, people pay more attention to time and efficiency. The shortest path problem becomes the object of study and discussion [1].

The genetic algorithm is analyzed and improved. Its population is optimized before being initialized, aiming at excluding the one which is obviously not necessary to be added to the process of algorithm and a combination of roulette and the elite protecting method is used, which can improve search efficiency and the quality of solution of the shortest path of GPS vehicle navigation system.

48.2 The Basic Structure of Genetic Algorithm

Genetic algorithm is the search algorithm which is a process of generate-and-test [2]. Its basic process is shown in Fig. 48.1.

As the Fig. 48.1 shows, the genetic algorithm is a population-base operation, which takes all the individuals in the group as its objects. Selection, crossover and mutation are the three major operating operators of the genetic algorithm, which constitute the so-called genetic manipulation and give the genetic algorithm the features that other traditional methods do not have.

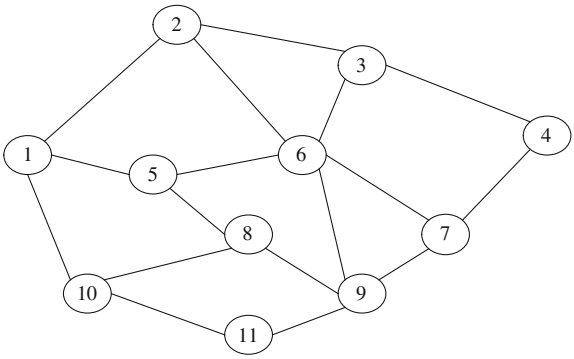


Fig. 48.2 A model of urban road network

48.3 The Design of Genetic Algorithm and its Improvement

The basic ideas of genetic algorithm are as follows: first of all, solution space in which there are problems needed to be solved should be transferred into the solution space of genetic algorithm, namely, choosing a code of a question. Secondly, initial population is given, and individual fitness is calculated [3]. In order to easily explain the problem, a model of urban road network is assumed as shown in Fig. 48.2.

48.3.1 Genetic Algorithm Encoding and Initial Population Generation

Using the symbol coding method can directly and concretely express the practical significance of the path. Initial population is a potential solution set, so the quality of the initial population is directly related to the quality of the final solution, the traditional way to generate the initial population is random, so an improvement is carried out at the beginning of the initialization operation, the nodes and branches which need not to be computed are cut down. For instance, the path from node 1 to 6 as shown in Fig. 48.2, there are several individuals, such as (1, 2, 6), (1, 5, 6), (1, 2, 3, 6) (1, 2, 3, 4, 7, 6), it is obvious that the last path and node 7 are not necessary to participate in operations. Just as the path from Beijing to Zhengzhou, there is no need from Beijing to Shanghai then comes to Zhengzhou, obviously, the last path is remote. According to the prior knowledge accumulation, the known knowledge is used to reduce the needless paths and nodes, and the speed of optimal path searching will be improved.

48.3.2 The Fitness Function

In order to guide the forward direction of the genetic algorithm, a fitness function is set and defined as:

$$f_i = \frac{1}{\sum_{j=1}^{l-1} G(j,j+1)} \tag{48.1}$$

where: f_i—the number i chromosome's fitness; l—the length of bit string; $G(j, j + 1)$—the distance from node j to $j + 1$.

48.3.3 Selection

The selecting operation is that choose the excellent individuals from the old generation. The commonly used method is based on the proportion of fitness. The individuals are selected by the means of roulette based on selective probability [4].

The function is used to calculate the selective probability:

$$P_i = \frac{f_i}{\sum_{j=1}^{n} f_j} \tag{48.2}$$

where: P_i—number i individual's selective probability; f_i—the number i chromosome's fitness; f_j—the number j chromosome's fitness; n—the total number of individuals.

But the roulette selection has some flaws. When an individual's fitness value is very large, even far more than any other individual fitness, it will be selected by the roulette with great probability. Opposite, when an individual's fitness value is too low, it will be discarded with a high probability, however this individual may contain some excellent genes, then, the population diversity would be destroyed. In order to protect the population's best individual in each generation, the best individual is directly singled out rather than being crossed and mutated, which is so-called elite protection. Therefore, the combination of elite protection and roulette can be a method to improve the solution quality.

48.3.4 Crossover

Crossover is a specific pair of chromosomes, they mutual exchange the genes. After crossover operation, its new individual's fitness can be even improved [5]. A concrete example is taken to illustrate the crossover operation, a pair of parent individuals A and B is assumed and the target from the start node 1 to node 7 is set as follows:

Parent individual A (1, 2, 6, 7)
Parent individual B (1, 5, 6, 9, 7)

It can be seen in the number 6 cross position can get a new offspring by crossover individuals, respectively, as follows:

Offspring individual A_1 (1, 2, 6, 9, 7)
Offspring individual A_2 (1, 5, 6, 7)

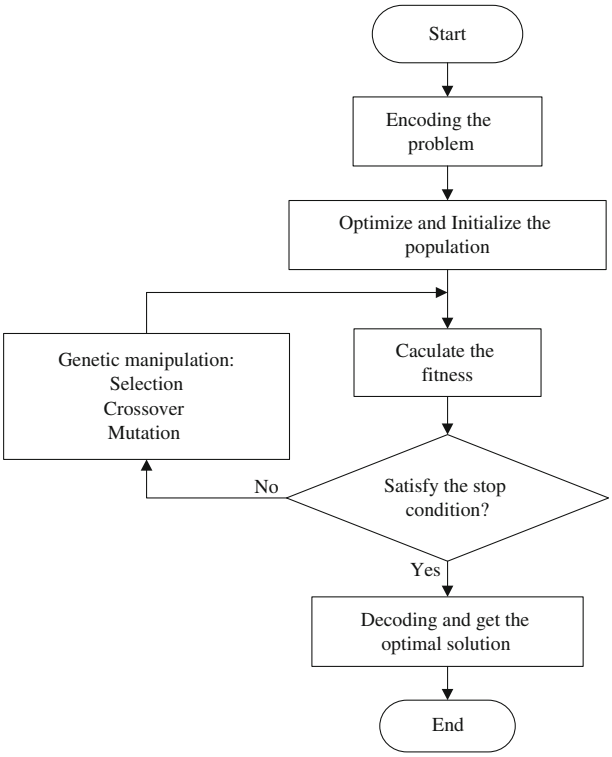


Fig. 48.3 The overall program flow chat of genetic algorithm

48.3.5 Mutation

In order to increase population's diversity and prevent getting a local optimal solution and the phenomenon of premature convergence, the mutation operators need to be added into the algorithm, but should avoid the phenomenon of short circuit and loop [6].

48.4 The Overall Design of Solving the Shortest Path by Using the Genetic Algorithm

The convergence rate and the quality of solution in searching the shortest path can be improved through using the optimized population and genetic manipulation. The selections of crossover probability and mutation probability have a great impact on the genetic algorithm, but here they need not to be repeated again [7, 8]. Finally, the overall programming of genetic algorithm is shown in Fig. 48.3.

48.5 Conclusion

The basic methods and strategies to solve the shortest path problem are discussed, a study of the algorithm implementation process is developed, the population is optimized before being initialized and the genetic manipulation is improved. It has a certain reference value in solving the problems of optimal solution's quality and convergence rate. People are now studying the theory of genetic algorithm and it will be widely used in different areas in the future.

Acknowledgments This paper is supported by the followings: (1) National Natural Science Foundation of China (41074090) (2) Key Technologies R&D Program of Henan Province (022102210360) (3) China Postdoctoral Science Foundation (2005038468) (4) Ph.D. Programs Foundation of Henan Polytechnic University (B2008-10).

References

1. Wu L (2010) The exploring of genetic algorithm for the shortest path based on urban road. University of Science, Nanjing
2. Chen G, Wang X, Zhang Z (1996) Genetic algorithm and its application. The People's Posts and Telecommunications Press, Beijing
3. Xu Q, Chen R, Guan Y (2003) The shortest path analysis based on genetic algorithms. J East China Geol Inst 2:168–172
4. He P, Pan J, Xue Q (2009) Dynamic route guidance for drivers based on genetic algorithms. Mod Electron Technol 15:205–207
5. Li Q, Xie S, Tong X (2006) A self-adaptive nenetic algorithm for the shortest path planning of vehicles and its comparison with Dijkstra and A*algorithms. J Univ Sci Technol Beijing 11:1082–1085
6. Xu Q, Ke X (2008) Genetic algorithm analysis for shortest path. Comput Eng Des 6: 1507–1509
7. Li Q, Zhang W, Yin Y (2006) An improved genetic algorithm for optimal path planning. Inf Control 4:444–447
8. Zhu S, Lin J, Guo X (2004) Application of improved genetic algorithm to the search of optimum route for intelligent vehicle navigation system. J Soochow Univ (Eng Sci Ed) 5: 99–102

Chapter 49
Radical Extraction for Handwritten Chinese Character Recognition by Using Radical Cascade Classifier

Enzhi Ni, Minjun Jiang and Changle Zhou

Abstract Radical extraction is the key step of radical-based Chinese character recognition. In this paper, we present a method of radical extraction by using radical cascade classifier, which detects radicals within characters. Haar-like features are applied in the cascade classifier, as they can adsorb radical distortion and guarantee the high speed of radical detection. Two methods of radical detection are proposed according to the characteristics of Chinese characters. The experiments are conducted on HITPU database and the results show the methods are efficient.

Keywords Handwritten Chinese character recognition · Radical extraction · Cascade classifier · AdaBoost

49.1 Introduction

Chinese characters are composed of radicals. Accordingly they can be recognized in terms of their radicals and the relative positions of these radicals. This kind of approach of Chinese character recognition is to be called radical-based approach.

E. Ni (✉) · M. Jiang · C. Zhou
Mind, Art and Computation Lab, Fujian Key Lab of the Brain-like Intelligent Systems, School of Information Science and Technology, Xiamen University, Xiamen 361005, China
e-mail: nienzhi@gmail.com

M. Jiang
e-mail: anna_xmu@sina.com

C. Zhou
e-mail: dozero@xmu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_49,

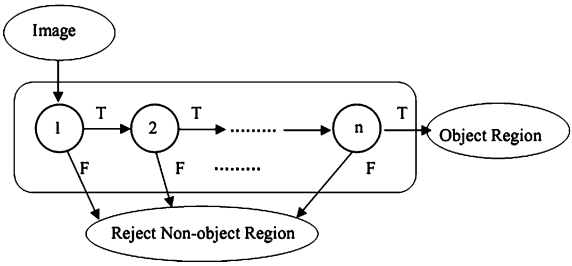


Fig. 49.1 The structure of cascade classifier

Its advantage is to simplify a very complicated classification task, which classifies around 20,902 Chinese characters [1], into a relatively simple one, which classifies only 560 radicals. Radical extraction is the key step of radical-based approach. There are several researches on radical extraction.

Wang and Fan [2] proposed a recursive hierarchical scheme for radical extraction of Chinese character. The proposed scheme includes three modules, which are character pattern detection module, straight cut line detection module, and stroke clustering module. Chung and Ip [3] proposed a radical segmentation algorithm based on D-snakes. D-snakes are used to locate the gaps between radicals, so that characters can be segmented into radicals according to these gaps. Shi et al. [4] used active shape models to extract radicals from Chinese characters. A set of examples of radicals is first represented by landmark points, then radicals are modeled as active shapes using kernel PCA, and finally unseen radicals are matched to the reference models using a genetic algorithm to search for the optimal shape parameters.

In this paper, we propose a method of radical extraction by using radical cascade classifier. Although it has been applied in many fields, cascade classifier is for the first time used in radical extraction. Feature selection for radical cascade classifier is discussed in this paper. Two methods of radical detection are presented according to the characteristics of Chinese characters.

49.2 Cascade Classifier and Adaboost Algorithm

Viola and Jones [5] proposed cascade classifier. In the beginning, it was used in rapid face detection. For its excellent performance, it has been applied to varieties of tasks. The idea of the cascade classifier is that a lot of simple and weak classifiers can make up complex and strong classifiers. This method allows non-object regions of an image to be quickly discarded and spends more computation on promising object-like regions. The strong classifiers then form a degenerated tree, which is a cascade classifier. Figure 49.1 shows the structure of a cascade classifier.

In Fig. 49.1, each stage in the cascade is a strong classifier. The stages are constructed by training classifiers using AdaBoost algorithm and then adjusting the threshold to minimize false negatives. In general, a lower threshold yields higher detection rates and higher false positive rates. Note that the default AdaBoost threshold is designed to yield a low error rate on the training data, therefore the default AdaBoost need to be modified for training cascade classifier [5].

The idea of AdaBoost [6] is to use the weak classifiers to form a highly accurate classifier (strong classifier) by learning weak classifiers repeatedly on different distributions over the training examples. Provided the error rate of these weak classifiers are less than 1/2 (with respect to the distribution on which it was trained), the error rate of the strong classifier decreases exponentially. The AdaBoost has some extensions such as real AdaBoost, LogitBoost and gentle AdaBoost [7]. They have very similar scheme. Gentle AdaBoost is used in the paper. And as mentioned above, the gentle AdaBoost needs to be modified for training cascade classifier. Following is a modified gentle AdaBoost algorithm:

Given maximum false detection rate $f_{\max}$, minimum detection rate $d_{\min}$ and example $(x_1, y_1), \ldots, (x_N, y_N)$ where $y_i = 0, 1$ for negative and positive examples, respectively.

Let f be false detection rate, d be detection rate, t be threshold of classifier.

1. Start with weights $w_i = 1/N$, $i = 1, 2, \ldots, N$, $m = 0, f = 1$, $H(x) = 0$.
2. If $f > f_{\max}$, $m = m + 1$ and repeat:
   (a) Train a weak classifier $h_m(x)$ by weighted least-squares of y_i to x_i with weights w_i, where $h_m(x) = P_w(y = 1|x) - P_w(y = 0|x)$.
   (b) Update $w_i \leftarrow w_i \exp(-y_i h_m(x_i))$ and renormalize
   (c) The current strong classifier is $sign[H(x)] = sign\left[\sum_{k=1}^{m} h_k(x) - t\right]$ where t is updated to ensure $d > d_{\min}$.
   (d) Update f.
3. Output the strong classifier $sign[H(x)] = sign\left[\sum_{k=1}^{m} h_k(x) - t\right]$.

This algorithm trains a strong classifier in example $(x_1, y_1), \ldots, (x_N, y_N)$, where x is the feature vector of an example, y is the label of an example. Given maximum false detection rate $f_{\max}$, minimum detection rate $d_{\min}$, the detection rate of the strong classifier has to be greater than $d_{\min}$ after each training. Train the strong classifier until the false detection rate is less than $f_{\max}$. In Step 2, the weak classifier $h_m(x) = P_w(y = 1|x) - P_w(y = 0|x)$, where $P_w(y|x)$ is the weighted class probability. We can approximately calculate $P_w(y = 1|x)$ and $P_w(y = 0|x)$ from the distribution of a training set. $h_m(x)$ has already been determined by the training set, the weights of training samples and the form of $h_m(x)$. Therefore, the threshold of the strong classifier t has to be adjusted in order to ensure current detection rate of the strong classifier is greater than $d_{\min}$. In each cycle of Step 2, the weights of examples are updated to $w_i \exp(-y_i h_m(x_i))$. By updating w_i, training weak classifiers will focus on those examples which still can not be classified correctly.

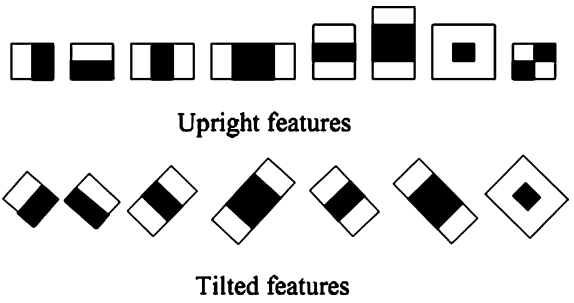


Fig. 49.2 Haar-like features

49.3 Feature Selection

A variety of features have been used to recognize handwritten Chinese characters, but it is uncertain whether they are suitable for training radical cascade classifier. We argue that a feature is suitable unless it satisfies the following three requirements:

1. The feature has to be high dimensional.

As mentioned above, the cascade classifier is formed by many strong classifiers, which are made up of lots of weak classifiers. A weak classifier corresponds to at least one dimension of the feature vector. Consequently, only high dimensional features are able to provide enough weak classifiers.

2. The feature has to be scalable or approximately scalable.

For radicals in handwritten Chinese characters are not in fixed sizes, cascade classifier needs to detect in the same images of different sizes. However, image scaling is time-consuming. Instead of scaling the image, we can scale the detection window of the cascade classifier so that the cascade classifier can rapidly detect radicals in different sizes. For this reason, the feature has to be scalable.

Definition 1 Let $F = (f_1, f_2, \ldots, f_n)$ be a feature vector of image I. After I scaled, the feature became F'. If for arbitrary scaling, there exists $F' = \lambda F$, where λ denotes arbitrary positive number and is directly proportional to the scale factor, the feature is scalable.

3. Since handwritten characters have much distortion, the feature must have the ability to reduce the influence of the distortion.

Haar-like features were originally used in face detection [5] and gave a good performance. These features are made up of several rectangles (see Fig. 49.2). The value of the features is the difference of the weighted sums of the pixels between black and white rectangular regions. The weights are in reciprocal proportion to the areas of rectangles. In Fig. 49.2 there are 12 haar-like features which are commonly used.

Fig. 49.3 The same haar-like features of three different radical "kou"

Haar-like features are high dimensional. For example, as to a 20×20 image, the number of its haar-like features is more than 120,000. By using integral image, any rectangular sum can be computed in four array references so that haar-like features can be computed very rapidly. Meanwhile, haar-like features are obviously scalable. Moreover, haar-like features have an advantage in radical detection. As we know, there is much distortion on handwritten Chinese characters. We usually blur the features to reducing the influence of the distortion. Haar-like features can absorb the distortion by expanding their rectangles. The expansion has the same effect with blurring (see Fig. 49.3). Haar-like features satisfy three requirements we mentioned above. And so we use haar-like features in this paper.

49.4 Radical Detection

Radicals in handwritten Chinese characters have the following characteristics: the positions of the same radicals in different characters can be different (see Fig. 49.4a–c); the same radicals can exist in different positions in a character (see Fig. 49.4d); even the sizes of the same radicals can be different when the radicals are in different positions. In addition, even the aspect ratios of the same radicals might change greatly (see Fig. 49.4a, b, e).

Because of these characteristics, radical detection has to be carried out in the whole images of Chinese characters by using the radical templates of different sizes and aspect ratios. The templates using haar-like features are approximately scalable and easy to transform to different sizes and aspect ratios. However, after the aspect ratios of templates being changed, the tilted haar-like features are not rectangular any more and their tilt angles are also changed. As a result, their values can not be computed with the original integral graph. Computing integral graph with every tilt angles needs much time, so it is not suitable for rapid radical detection. There are two solutions for this problem. One is to change the aspect ratio of character images instead of the aspect ratios of radical templates; the other is to avoid using the tilted haar-like features. Obviously, the first one will reduce efficiency and the second one will lose many efficient weak classifiers. Two diagrams of the two solutions are showed in Fig. 49.5a, b respectively. The ranges of the aspect ratios and sizes of different radical categories vary widely. We set the ranges according to training samples.

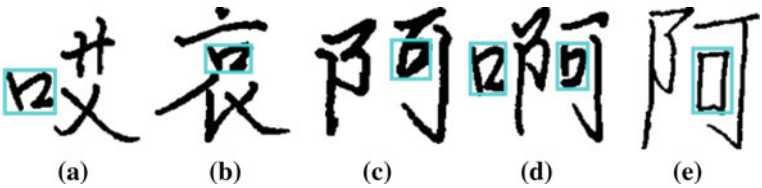


Fig. 49.4 Radical "kou" in different characters

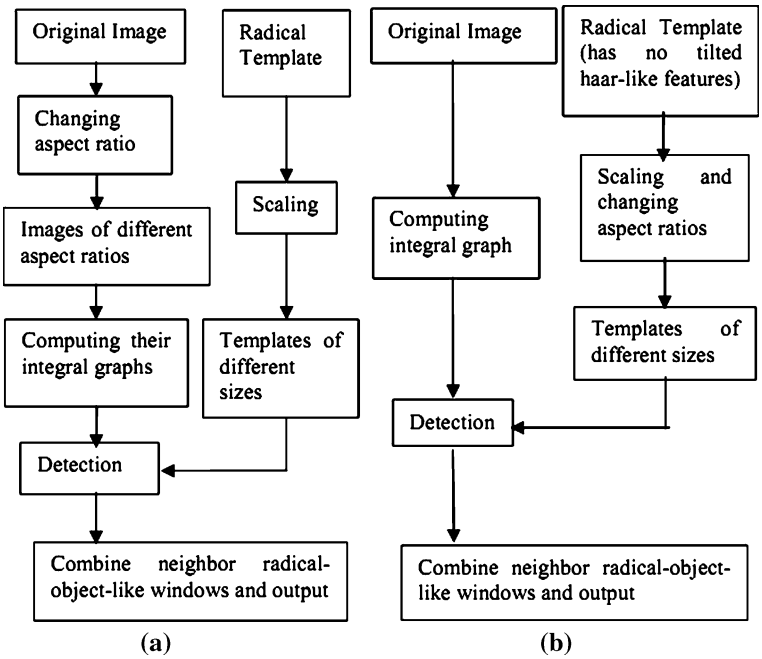


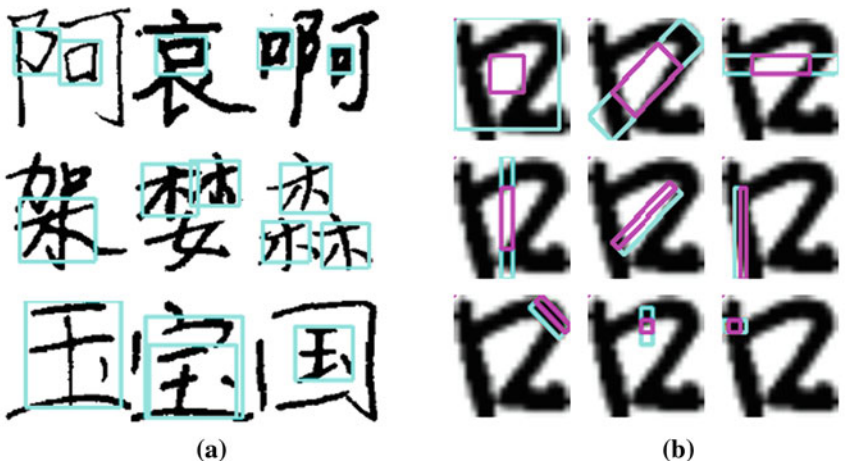
Fig. 49.5 **a** Method 1: detection in the same images of different aspect ratios. **b** Method 2: detection with the same templates of different aspect ratios

49.5 Experiments and Results

Experiments are performed on Intel Pentium 4 CPU (3.00 GHz) equipped with 1 GB RAM. The source code is written in C++ with OpenCV. All experimental samples come from HITPU, a database collected by Harbin Institute of Technology and Hong Kong Polytechnic University, which comprises a collection of 751,000 loosely constrained handwritten Chinese characters, consisting of 3,755 categories written by 200 different writers. All training radical samples are manually extracted from HITPU, including six radical categories. These radical categories could have much distortion in handwritten characters. The training set for

Table 49.1 Experimental results of Method 1/Method 2

Radical	Detection rate (%)	Average detection time (s) per character	Average number of object-radical-like windows per character	
			Samples containing object radical	Samples not containing object radical
口	91.70/94.50	0.144/0.076	2.02/2.41	1.49/1.81
Surrounding 口	98.51/98.86	0.100/0.041	1.44/1.44	0.88/0.81
木	95.35/94.03	0.113/0.055	1.69/1.71	0.66/0.70
王	94.73/97.99	0.039/0.017	2.26/2.34	1.06/1.02
玉	96.23/94.34	0.067/0.030	0.96/1.06	0.09/0.31
足	89.19/97.10	0.047/0.019	2.39/1.88	1.61/0.71



(a) (b)

Fig. 49.6 **a** Some results of radical detection. **b** Some haar-like features used in the cascade classifier of radical "kou"

each radical consists of 1,000 positive samples (500 for radical "yu" 玉) which are radical images and 6,000 negative samples which are character images not containing this radical. The test set for each radical consists of 1,000 character images which contain this radical (53 character images for radical "yu" 玉) and 5,000 character images which do not. The test samples cover a large number of character categories from different writers. Our proposed method has achieved detection rates of 94.29% (method 1) and 96.14% (method 2) and average detection time per character of 0.085 s (method 1) and 0.040 s (method 2). The detailed experimental results are showed in Table 49.1. In the table, the object-radical-like windows denote detected radicals, which can be correct object radicals or not. The results show method 2 is more rapid than method 1, meanwhile its performance is slightly

better than method 1. Some results of radical detection are showed in Fig. 49.6a and some haar-like features used in the cascade classifier of radical "kou" are showed in Fig. 49.6b. From the experimental results, we concluded that our methods are efficient for radical extraction of Chinese characters.

49.6 Conclusions

In this paper, we propose a method of radical detection for handwritten Chinese character recognition by using radical cascade classifier. Experiments conducted on HITPU database show this method has a good performance. Since collecting training radical samples and training cascade classifiers are very time-consuming, there are currently only six radicals in our experiments. Nevertheless, the methods of training cascade classifiers of all radicals are same. In our further work, we will continue training the classifiers of other radicals and implement a radical-based method of Chinese character recognition by using these classifiers.

Acknowledgments Thank Prof. Shi for providing HITPU database, which is very helpful to this research. The work described in this chapter was supported by National Nature Science Foundation of China under grant No. 60975076 and grant No. 60903129.

References

1. Zhang DS (2007) The goal of the Chinese character component standard and the disassembling method. Chin Character Res, 2:229–233
2. Wang AB, Fan KC (2001) Optical recognition of handwritten Chinese characters by hierarchical radical matching method. Pattern Recognit 34:15–35
3. Chung FL, Ip WWS (2001) Complex character decomposition using deformable model. IEEE Trans Syst Man Cybern Part C Appl Rev 31:126–132
4. Shi DM, Ng GS, Damper RI, Gunn SR (2005) Radical recognition of handwritten Chinese characters using GA-based kernel active shape modelling. IEE Proc Vis Image Signal Process 152:634–638
5. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of IEEE conference computer vision and pattern recognition, pp 511–518
6. Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55:119–139
7. Friedman J, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. Ann Stat 28:337–407

Chapter 50
State Filter for Descriptor Systems with Packet Losses

Lifang Liu and Shuli Sun

Abstract This paper is concerned with the design of the optimal linear filter for a networked descriptor system with multiple packet losses. A descriptor system is converted to two reduced-order subsystems. Based on projection theory, the optimal linear filters for the two reduced-order subsystems with packet losses are presented in the linear minimum variance sense, respectively. Further, the linear filter for the original descriptor system is obtained. The proposed filter only depends on the data arrival rate but do not depend on the knowledge whether the sensor data are received at a particular time. A simulation example shows its effectiveness.

Keywords Packet losses · Reduced-order state filter · White noise filter · Descriptor system

50.1 Introduction

The descriptor system is official mentioned in the complex networked systematic engineering due to the wide applications in robotics, economic systems and complex chemical process, which is more common to describe the actual systems than conventional systems. In networked control systems and sensor networks, the

S. Sun (✉)
School of Electronics Engineering, Harbin 150080, China
e-mail: sunsl@hlju.edu.cn

L. Liu
School of Mathematics Science, Heilongjiang University, Harbin 150080, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_50,

packet losses usually occur in data transmission because of the limit of communication and the disturbance of outside environment [1].

In recent years, the estimation problem for systems with packet losses has attracted lots of attention and some novel results are obtained. The optimal estimators including filter, predictor and smoother are investigated in the linear minimum variance sense via an innovation analysis approach in Ref. [2]. However, only some components of the system state are sometimes useful for some practical applications, the full and reduced-order optimal filters for systems with multiple packet losses are designed in Refs. [3, 4] has considered the filtering problem for systems with finite consecutive multiple packet losses via transform the original system into one with the random delayed measurements and moving average (MV) colored measurement noise. In Refs. [5] and [6], the information fusion Kalman filter and smoother with colored noises for descriptor system are given, but they don't consider the case of missing data. The optimal fusion estimation problem for descriptor system with uncertain disturbance is developed in Ref. [7] where both additional uncertain disturbance and multiplicative uncertain observation are taken into account. However, the optimal filtering problem for the descriptor system with packet losses is seldom studied.

In this paper, we consider the optimal linear state filter for a descriptor system with packet losses, where the phenomena of packet losses are described by a Bernoulli-distributed random variable. The descriptor system with packet losses is converted into two reduced-order subsystems by non-singular transformations. The state and white noise filters of the subsystems are designed based on projection theory. Further, the filter for original descriptor system with packet losses is obtained. The proposed filter only depends on the data arrival probability but do not need to know whether the sensor measurements are received at a particular time.

50.2 Problem Formulation

Consider a linear discrete stochastic descriptor system with packet losses

$$Mx(t+1) = \hat{\Phi}x(t) + \hat{\Gamma}w(t). \tag{50.1}$$

$$z(t) = \hat{H}x(t) + v(t), \tag{50.2}$$

$$y(t) = \xi(t)z(t) + (1-\xi(t))y(t-1), \tag{50.3}$$

where $x(t) \in R^n$ is the state, $z(t) \in R^m$ is the measured output, $y(t) \in R^m$ are the measurement received by the filter to be designed. $w(t) \in R^r$ and $v(t) \in R^m$ are the correlated white noise with zero means and variances Q_w, Q_v and cross-covariance S_0. $\xi(t)$ is the white Bernoulli-distributed random variable with probabilities $\text{Prob}\{\xi(t)=1\} = \alpha$ and $\text{Prob}\{\xi(t)=0\} = 1-\alpha$, $0 \le \alpha \le 1$ "Prob"

denotes probability, and is uncorrelated with other random variables. $\hat{\Phi}$, $\hat{\Gamma}$ and $\hat{H}$ are the time-invariant matrices with suitable dimensions.

Assumption 1. M is a singular matrix, rank $M = \mathrm{n}_1 < \mathrm{n}$.

Assumption 2. The initial state $x(0)$ is uncorrelated with $w(t)$, $v(t)$ and

$$\mathrm{E}[x(0)] = \mu_0,\ \mathrm{E}\left[(x(0)-\mu_0)(x(0)-\mu_0)^{\mathrm{T}}\right] = P_0, \tag{50.4}$$

where the symbol E denotes mathematical expectation and T denotes transpose.

Assumption 3. Systems (50.1)–(50.3) are regular, i.e., $\det(\lambda M - \hat{\Phi}) \neq 0, \forall \lambda \in C$, where C denotes complex number set.

We aim to find the linear minimum variance state filter $\hat{x}(t|t)$ for descriptor system with packet losses based on the received measurements $(y(1), y(2), \ldots, y(t))$.

In view of Assumptions 1–3, the descriptor systems (50.1)–(50.3) can be equivalently converted to two reduced-order subsystems. There exist non-singular matrixes U and Q, such that [8]

$$UMQ = \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix},\ U\hat{\Phi}Q = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}, U\hat{\Gamma} = \begin{bmatrix} \hat{\Gamma}_1 \\ \hat{\Gamma}_2 \end{bmatrix},\ \hat{H}Q = \begin{bmatrix} \hat{H}_1 & \hat{H}_2 \end{bmatrix},\ x(t) = Q\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \tag{50.5}$$

where $x_1(t) \in R^{n_1}$, $x_2(t) \in R^{n_2}$, $n_1 + n_2 = n$, I_{n_1} denotes an $n_1 \times n_1$ identity matrix.

So descriptor systems (50.1)–(50.3) can be written by

$$\begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \hat{\Gamma}_1 \\ \hat{\Gamma}_2 \end{bmatrix} w(t) \tag{50.6}$$

$$z(t) = \begin{bmatrix} \hat{H}_1 & \hat{H}_2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + v(t) \tag{50.7}$$

$$y(t) = \xi(t)z(t) + (1-\xi(t))y(t-1). \tag{50.8}$$

Assume Φ_{22} non-singular, the descriptor systems (50.6)–(50.8) are equivalent to two reduced-order subsystems

$$x_1(t+1) = \Phi x_1(t) + \Gamma w(t). \tag{50.9}$$

$$z(t) = Hx_1(t) + Cw(t) + v(t). \tag{50.10}$$

$$y(t) = \xi(t)z(t) + (1-\xi(t))y(t-1). \tag{50.11}$$

$$x_2(t) = Tx_1(t) + Dw(t). \tag{50.12}$$

where $\Phi = \Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21},\ = \hat{}_1 - \Phi_{12}\Phi_{22}^{-1}\hat{}_2,\ T = -\Phi_{22}^{-1}\Phi_{21},\ D = -\Phi_{22}^{-1}\hat{}_2,$ $H = \hat{H}_1 + \hat{H}_2 T,\ C = \hat{H}_2 D.$

50.3 Optimal Linear Filter

In this section, the optimal linear filter will be derived. It is easily known that the filter for the descriptor system (50.1)–(50.3) can be obtained based on the filters of the two subsystems (50.9)–(50.11) and (50.12) in the section above. So we will derive the related filters including the state filter and white noise filters of the two subsystems firstly.

For the first subsystems (50.9)–(50.11), from [2] and [9] we can directly obtain the optimal state linear filter $\hat{x}_1(t|t)$, the variance matrix $P_{x_1}(t|t)$, the state filtering gain matrix $K(t)$, the innovation sequence $\varepsilon(t)$, the innovation variance matrix $Q_\varepsilon(t)$, and the optimal white noise filters $\hat{w}(t|t)$, the white noise filtering gain matrix $M_W(t|t)$ and the variance matrices $P_w(t|t)$ and $P_v(t|t)$. Here the details are omitted due to limited space.

50.3.1 Optimal State Filter of the Second Subsystem

Based on the optimal state filter and white noise filter of the first subsystems (50.9)–(50.11), we can obtain the optimal state filter of the second subsystem.

Theorem 1 *For system (50.12), the optimal state filter is given by*

$$\hat{x}_2(t|t) = T\hat{x}_1(t|t) + D\hat{w}(t|t). \tag{50.13}$$

$$P_{x_2}(t|t) = TP_{x_1}(t|t)T^{\mathrm{T}} + DP_w(t|t)D^{\mathrm{T}} + TP_{x_1 w}(t|t)D^{\mathrm{T}} + DP_{x_1 w}^{\mathrm{T}}(t|t)T^{\mathrm{T}}. \tag{50.14}$$

where

$$P_{x_1 w}(t|t) = -[I_{n_1}\ 0]K(t)Q_\varepsilon(t)M_W^{\mathrm{T}}(t|t)[I_r\ 0]^{\mathrm{T}}. \tag{50.15}$$

Proof Taking projection of the both sides of the (50.12) onto the linear space spanned by $(y(1),\ y(2), \ldots, y(t))$ yields (50.13). Subtracting (50.13) from (50.12), the state filtering error equation of the state $x_2(t)$ can be obtained as

$$\tilde{x}_2(t|t) = x_2(t) - \hat{x}_2(t|t) = T\tilde{x}_1(t|t) + D\tilde{w}(t|t), \tag{50.16}$$

where $\tilde{x}_1(t|t) = x_1(t) - \hat{x}_1(t|t)$ and $\tilde{w}(t|t) = w(t) - \hat{w}(t|t)$. The variance matrix (50.14) is obtained by computing $P_{x_2}(t|t) = \mathrm{E}[\tilde{x}_2(t|t)\tilde{x}_2^{\mathrm{T}}(t|t)]$. $P_{x_1 w}(t|t) = \mathrm{E}[\tilde{x}_1(t|t)\tilde{w}^{\mathrm{T}}(t|t)]$ defined by (50.15) denoting the filtering error cross-covariance

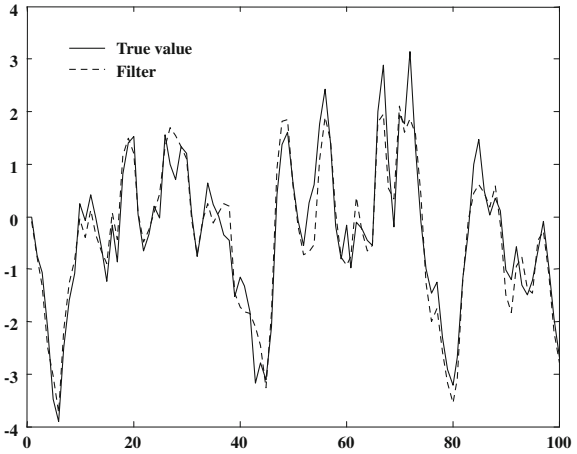


Fig. 50.1 Filter for the first state component

matrix of the state $x_1(t)$ and white noise $w(t)$, which can be easily derived by using Refs. [2] and [9].

By now, we have gained the optimal state filters of the two subsystems. In the following, we will give the filter of the original descriptor system with packet losses.

50.3.2 Optimal Linear Filter of the Original Descriptor System

According to $x(t) = Q\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$, the optimal linear filter of the original descriptor system with packet losses can be given by the following Theorem.

Theorem 2 *The optimal linear filter of descriptor systems (50.1)–(50.3) is computed by*

$$\hat{x}(t|t) = Q\begin{bmatrix} \hat{x}_1(t|t) \\ \hat{x}_2(t|t) \end{bmatrix}. \tag{50.17}$$

$$P_x(t|t) = Q\begin{bmatrix} P_{x_1}(t|t) & P_{x_1x_2}(t|t) \\ P^{\mathrm{T}}_{x_1x_2}(t|t) & P_{x_2}(t|t) \end{bmatrix} Q^{\mathrm{T}},\ P_{x_1x_2}(t|t) = P_{x_1}(t|t)T^{\mathrm{T}} + P_{x_1w}(t|t)D^{\mathrm{T}}. \tag{50.18}$$

Proof Based on the projection theory, we have (50.17) from the defined $x(t) = Q\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. Using (50.5) and (50.17), the filtering error equation is computed by

$$\tilde{x}(t|t) = Q\begin{bmatrix} \tilde{x}_1(t|t) \\ \tilde{x}_2(t|t) \end{bmatrix}. \tag{50.19}$$

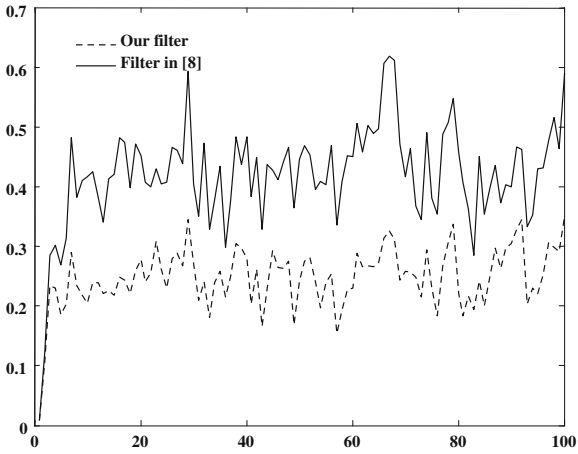


Fig. 50.2 Comparison of MSE of our filter and [8]

So the state filtering error variance (50.18) can be obtained by computing $P_x(t|t) = \mathrm{E}\left[\tilde{x}(t|t)\tilde{x}^{\mathrm{T}}(t|t)\right]$ form (50.19), where the filtering error cross-covariance matrix $P_{x_1x_2}(t|t)$ of the state $x_1(t)$ and $x_2(t)$ is defined by $P_{x_1x_2}(t|t) = \mathrm{E}[\tilde{x}_1(t|t)\tilde{x}_2^{\mathrm{T}}(t|t)]$ which comes from (50.16).

50.4 Simulation Example

Consider the stochastic descriptor systems (50.1)–(50.3) with packet losses, where

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ \hat{\Phi} = \begin{bmatrix} 0.5 & 0 & -0.1 & 0 \\ -2 & 0 & 0.2 & 0 \\ 2.55 & -0.5 & -0.93 & 0 \\ -2.4 & -1 & -0.76 & 2 \end{bmatrix}, \ \hat{\Gamma} = \begin{bmatrix} -0.1 & -0.4 \\ 0.5 & 2.2 \\ 0.4 & -1.1 \\ 1.6 & 2.01 \end{bmatrix},$$

$$\hat{H} = \begin{bmatrix} -0.9 & 1 & 0.7 & 0 \\ -0.5 & 0 & 2 & 1 \end{bmatrix}, \ U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The measurement noise $v(t) = cw(t) + \zeta(t)$ is correlated with the Gaussian input noise $w(t)$, c which is the correlated coefficient and $\zeta(t)$ is a two-dimension Gaussian white noise with zero mean and variance matrix Q_ζ. Our aim is to find the optimal linear state filter $\hat{x}(t|t)$. In the simulation, we take $Q_w = 4I_2$, $Q_\zeta = 0.01I_2$, $c = 0.9$, $\lambda = 0.3$, $\alpha = 0.8$. The initial values $x(0) = [0 \quad 0 \quad 0 \quad 0]^{\mathrm{T}}$, $P_0 = 0.1\, I_4$.

The filter of the descriptor system with packet losses is shown in Figs. 50.1 and 50.2 show the comparison of the MSE (mean square error) of 200-time Monte Carlo test for our filter and the filter by filtering algorithm in Ref. [8]. From Fig. 50.2, we can see that the accuracy of our filter is better. The filtering algorithm in [8] has lost the optimality when there are packet losses.

50.5 Conclusion

In this paper, the optimal linear state filtering problem for a descriptor system with packet losses has been investigated. The descriptor system is equivalently converted into two reduced-order subsystems. Based on projection theory, the optimal state linear filters for two reduced-order subsystems are obtained in the linear minimum variance sense. Further, the optimal filter for the original descriptor system with packet losses is obtained.

Acknowledgments This work was supported by Natural Science Foundation of China under Grant NSFC-60874062, Key Project of Chinese Ministry of Education under Grant 209038, Program for New Century Excellent Talents in University under Grant NCET-10-0133, 1154-NCET-01 and Program for High-qualified Talents under Grant Hdtd2010-03.

References

1. Sinopoli B, Schenato L, Franceschetti M, Poolla K, Jordan M, Sastry S (2004) Kalman filtering with intermittent observations. IEEE Trans autom control 49:1453–1464
2. Sun, SL, Xie LH, Xiao WD, Soh YC (2008) Optimal linear estimation for systems with multiple packet dropouts. Automatica 44:1333–1342
3. Sun L, Xie LH, Xiao WD (2008) Optimal full-order and reduced-order estimators for discrete-time systems with multiple packet dropouts. IEEE Trans Signal Process 56:4031–4038
4. Sun SL, Xie LH, Xiao WD, Xiao N (2008) Optimal filtering for systems with multiple packet dropouts. IEEE Trans Circuits Syst-II: express briefs, vol 55. pp 695–699
5. Tao GL, Liu, WQ, Sun CF (2009) Information fusion Kalman filter with colored noises for descriptor systems. J Daqing Pet Inst 33:93–97, Daqing Petroleum Institue, Diqing, China
6. Tao GL, Liu WQ, Hou JY (2010) Information fusion Kalman smoother for descriptor system with ARMA colored observation noises. Sys Sci Math 30:205–216
7. Qu DM (2010) The fusion estimation for descriptor system with uncertain disturbance. Heilongjiang University Master Degree Theses, Harbin
8. Deng ZL (2001) Kalman filtering and Wiener filtering modern time series analysis method. Harbin Institute of Technology Press, Harbin
9. Ma J, Liu LF, Sun SL (2011) White noise filters for systems with multiple packet dropouts. In: The 30th Chinese control conference, Yantai, China

Chapter 51
Real-Time Property Analysis and Design of Dual-Channel Simultaneous Data Acquisition and Processing System

Sheng Tang, Xiaohui Li and Ya Liu

Abstract Real-time property is one of the most important properties of the data acquisition and processing system. The working principle and designing ways of the data acquisition and processing system was briefly introduced and the factors affecting the real-time property of the system was emphatically analyzed in this paper. In order to design and implement a real-time dual-channel simultaneous data acquisition and processing system, the techniques of dual CPU, HPI (Host Port Interface) and ping-pong buffer strategy were adopted, all of which were important guarantees of the real-time property. Some conclusions obtained in this paper can be of significance to the similar engineering.

Keywords Data acquisition and processing · Real-time property · Dual CPU · HPI · Ping-pong buffering

This program is founded by the National Nature Science Fund 61001076.

S. Tang (✉) · X. Li · Y. Liu
National Time Service Center, Chinese Academy of Sciences,
Xi'an 710600, China
e-mail: tangsheng@ntsc.ac.cn

S. Tang
Graduate University of Chinese Academy of Sciences,
Beijing 100190, China

S. Tang · X. Li · Y. Liu
Key Laboratory of Time and Frequency Primary Standard,
National Time Service Center, Chinese Academy
of Sciences, Xi'an 710600, China

S. Tang
School of Information and Technology, Northwest University,
Xi'an 710127, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_51,

51.1 Introduction

Data acquisition and processing integrates the analog-to-digital conversion by sensor or A/D sampling circuit and the signal processing by CPU after data is gathered to obtain desired information.

There are two conventional approaches to the technique of data acquisition and processing: (1) MCU is adopted as the main CPU to accomplish data acquisition, processing and transmission, with the strength of relatively simple circuit structure but low precision, and (2) PC is supplemented with special data acquisition card to develop the system, which has high precision but large system size and poor portability.

Data acquisition and processing plays an increasingly important role in the area of modern industry, and the performance requirements like the system precision and size vary according to different application areas. With the development of modern industry toward high performance, efficiency, networking and intelligentization, the demand of data acquisition and processing system with time constraint grows, such as measurement of electric parameters, acquisition and analysis of radar signal, field control of complex industrial process, etc. And all these application require real-time property to some extent besides the performance specifications of precision and size [1–3].

The real-time property allows the data acquisition and processing system to respond within the limited time with high speed and transfer the result to the destination in time when the sampling data being generated, otherwise false result or system malfunction may be caused.

This paper analyses the mechanism of data acquisition, transmission and processing in the data acquisition and processing system, and determines the main factors affecting the real-time property of the system, then a dual-channel data acquisition and processing system is realized by using the techniques of dual CPU architecture, HPI and ping-pong buffering strategy.

51.2 Analysis on Real-Time Factors Influencing the System

51.2.1 Task Conflict Between Data Acquisition and Processing

Data acquisition and processing are carried out simultaneously as parallel tasks for the data acquisition and processing system, and the time of data processing should be less than the time of data acquisition.

The operating sequence chart is shown in Fig. 51.1. Data acquisition and processing are carried out simultaneously in a periodic way. The system with real-time property should meet the condition as follows:

$$T_s > T_h + T_t + T_p \tag{51.1}$$

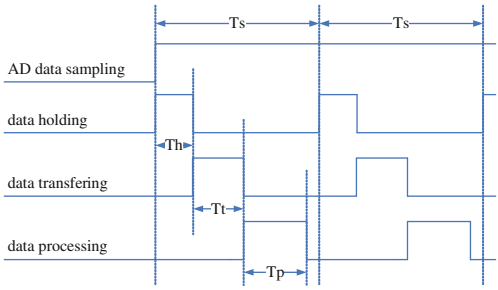


Fig. 51.1 Operating sequence chart of data acquisition and processing system

In the conventional system designing methods, data acquisition and processing are implemented by the same CPU, and real-time property can be guaranteed if the data flow is relatively small and the data processing algorithm is simple. However, inaccuracy in the time of data acquisition and processing will yield and then real-time property cannot be achieved if the processing capacity of CPU is limited for a large data quantity.

51.2.2 Conflict Between Storing and Reading in Data Transmission

Sampled data are usually computed after they have been accumulated to a certain amount in the process of data acquisition. Therefore, memory space is needed for data accumulation, storage and computation. The real-time property of the system will be affected when the data storing and data reading conflict for the lack of memory space and improper memory allocation or management mechanism.

51.3 Real-Time Performance Design of the System

Specification and technical requirements are as follows in our data acquisition and processing system:

(1) Two sine or cosine analog input channels;
(2) The frequency of the signal is 100 Hz and the range of input voltage is in the voltage range of is ±2.5 V;
(3) Long-duration signal and large data quantity.

Simultaneous acquisition with high-precision of the above signals is required to work at the sample rate of 10 kHz, and the digital signal processing algorithm (correlation algorithm) is adopted for calculating phase difference between the dual signal mentioned above. The calculating result should be outputted every 1 s, so that the system design will satisfy the performance requirements of precision, portability and real-time property.

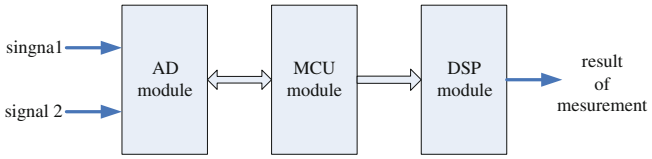


Fig. 51.2 Hardware structure block diagram

51.3.1 Design of Hardware Structure Based on Dual CPU

The designing method of PC supplemented with data acquisition card can be neglected due to the poor portability, while the method based on standard MCU can also be ignored because it can not perform multi-task such as data acquisition and processing from different channels simultaneously. Therefore DSP is a good choice as the core CPU of the system for its superiority in operation precision and speed over MCU.

However, complex system task scheduling will be caused with only one DSP performing the tasks such as data acquisition, analysis and computation, system control and real-time display. Thus, measurement results of the output cannot satisfy the real-time property.

As a result, a dual CPU structure of a DSP integrating with a MCU is adopted in our design to balance between the performance and cost. The two CPUs are working with clear responsibilities respectively in order to improve the working efficiency and guarantee the real-time property. MCU realize the function of data sampling control, liquid crystal display, keyboard control, serial data output, etc. At the same time, DSP mainly completes various real-time tasks of analysis and computation.

Structure of data acquisition and processing system based on dual CPU is presented in Fig. 51.2.

Taking account of the large amount of data to be processed and strict real-time demand of the output results, TMS320C6713 is selected as the core CPU to implement the real-time analysis and computation. As a high-performance floating-point DSP of 32-bit data word, TMS320C6713 delivers up to 1800 million floating-point operations per second (MFLOPS) operating at 300 MHz, and with dual floating-point multiplier up to 600 million multiply-accumulate operations per second (MMACS) [4].

MSP430F149 chip is used as the MCU in the design. MSP430F149 produced by TI is a mixed signal microcontroller, with a lot of on-chip external components. The maximum code conversion efficiency of MSP430F149 can be achieved by 16 registers and constant generators in its CPU. Its maximum working frequency is 8 MHz, and command cycle is only 150 ns. All these are out of range of 51 series single chip computer. The strong data processing ability can reach 2MIPS that is very suitable for embedded system which request very high processing ability.

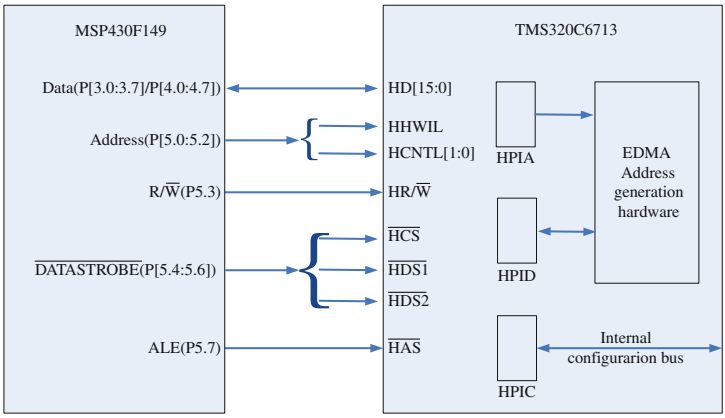


Fig. 51.3 Design of dual CPU interface

60 KB FLASH ROM and 2 KB RAM provide a great convenience for operation processing [5].

Considering that the system requires simultaneous acquisition of dual-channel data, the ADS8364 is adopted as the AD converter. The ADS8364 includes six, 16-bit, 250 kHz ADCs (Analog-to-Digital converters) with 6 fully differential input channels grouped into two pairs for high-speed simultaneous signal acquisition, and offers a flexible high-speed parallel interface with a direct address mode, a cycle and a FIFO mode. The output data for each channel is available as a 16-bit word [6].

51.3.2 Design of Dual CPU Interface

The structure of dual CPU is a good choice in the design of data real-time acquisition and processing system, because it is helpful to coordinate the tasks of data acquisition, transmission and processing. The data transmission mode between the dual CPU is various, such as serial port, USB, dual-port RAM, etc. But most of modes mentioned above have the shortcoming of occupying many hardware and software resources. In our design, HPI is adopted in data transmission between DSP and MCU.

The HPI is a 16 bit parallel port through which a host processor can directly access the CPU memory space [7]. In our design, MCU (MSP430F149) is the host processor, while DSP (TMS320C6713) is the slave processor. The interface between the host processor and the slave processor is shown as the Fig. 51.3, of which signals are presented in Table 51.1.

Table 51.1 HPI signals description

Signal name	Signal count	Host connection	Signal function
HD[15–0]	16	Data bus	Data transfer
HHWIL	1	Address or control lines	Half-word identification input
HCNTL [1–0]	2	Address or control lines	HPI access type control
HR/$\overline{W}$	1	Read/write strobe, address line, or multiplexed address/data	Read/write select
$\overline{HCS}$	1	Address or control lines	Data strobe inputs
$\overline{HDS[1\text{-}2]}$	2	Read strobe and write strobe or data strobe	Data strobe inputs
$\overline{HAS}$	1	Address latch enable, address strobe, or unused (tied high)	Differentiation between address and data values on multiplexed address/data host

The MCU functions as a master of the interface, which increases ease of access. The MCU and DSP can exchange information via internal or external memory, so that the task of data processing in DSP cannot be interrupted. Connection with the DSP memory space is provided through the direct memory access enhanced direct memory access (EDMA) controller of DSP. Both the MCU and the DSP can access the HPI control register (HPIC). The MCU can access the HPI address register (HPIA), HPI data register (HPID) and HPIC by using the external data and interface control signals.

In our design, MSP430F149 sends the data acquired by ADS8364 directly into the SDRAM of DSP. The whole procedure has no influence on the normal task of DSP, which is the important guarantee for real-time property of the system.

51.3.3 Design of Ping-Pong Buffering

In order to coordinate the tempo of data processing and transmission, a strategy of establishing send-and-receive buffer in the extended memory (i.e. SDRAM) is adopted in our design, which can be called as ping-pong buffering strategy or double buffering strategy.

Ping-pong buffering strategy includes two pairs of buffer: ping buffer and pong buffer. Data in one pair of buffer are being transmitted while data in another buffer pair are being processed. Each pair of buffer is comprised of input buffer and output buffer, thus we have 4 buffers in all, which can be defined as InBuffA, OutBuffA, InBuffB, OutBuffB in this paper. Data transmitting is handled by EDMA in ping buffer while data processing is manipulated by the CPU of DSP in Pong buffer in our design. Once completed, CPU and EDMA switch ping and pong buffer pairs. Data transmitting and data processing is always being executed in parallel, as shown in Fig. 51.4.

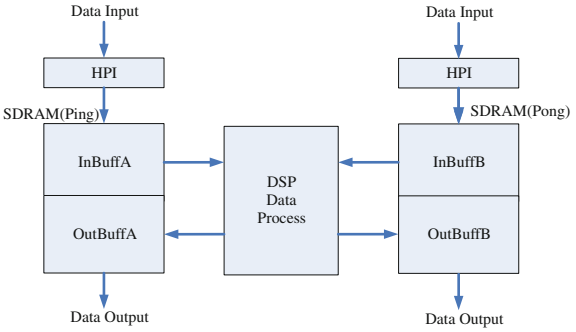


Fig. 51.4 Ping-pong buffering principle

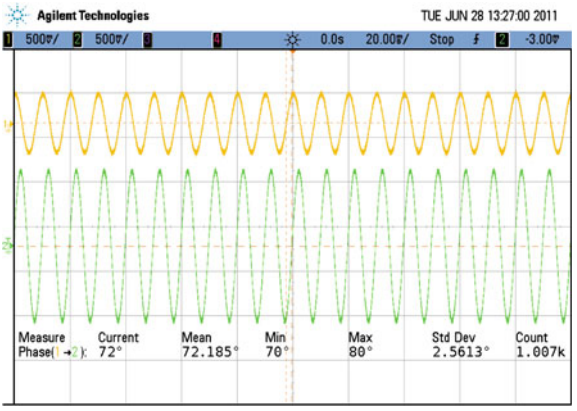


Fig. 51.5 Measurement result by Agilent DS5054A

51.4 System Test and Result Analysis

The measurement result of the dual signal by the oscillograph (Agilent DS5054A) is shown as Fig. 51.5, and the phase difference of the two signals is about 72.18°. The wave data of the two signals temporarily stored in the ping buffer of DSP at a certain time from the system presented in this paper is shown in Fig. 51.6 (intercepted from CCS3.0, with the voltage as abscissa and the sampling number as ordinate), and the real-time measurement result of the phase difference is presented in Fig. 51.7 (intercepted from Stable 32).

Processing of the sampled data is completed by DSP, and the correlation algorithm is adopted in phase difference calculating. Compared with the traditional phase difference measurement methods (e.g. counting method), correlation algorithm is complicated, but with high precision and good noise suppression ability to guarantee the measurement precision of the system [8]. According to the specifications, sampling of the dual-channel signal is taken every 0.1 ms uninterruptedly, and complicated data processing algorithm (i.e. phase difference calculation) is implemented every 1 s. The results of experiment showed that the proposed

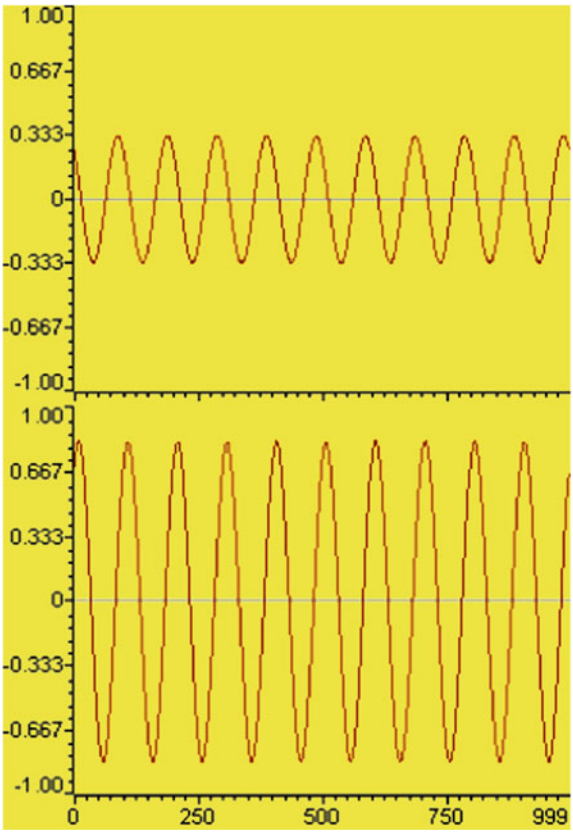


Fig. 51.6 Real-time data (dual signal to be measured) received by DSP

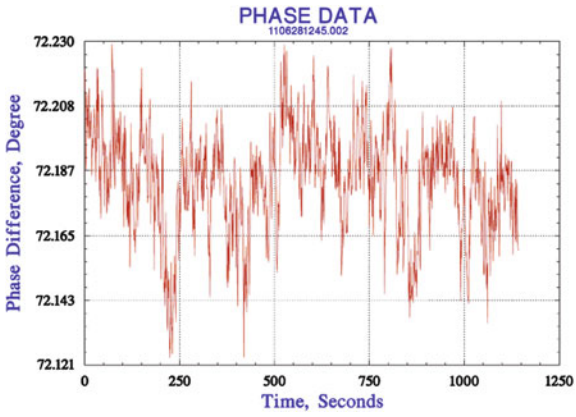


Fig. 51.7 Real-time data (phase difference of dual signal) output by DSP

system can output the measurement result every 1 s to meet the real-time requirement by using the method based on the strategies of dual CPU structure, HPI and ping-pong buffering.

51.5 Conclusion

Real-time property is one of the most important performance requirements of data acquisition and processing system. The primary task of the design is to determine the various factors influencing the real-time property, such as selection and architecture of CPU, task scheduling of data acquisition and processing, control of store and read in data transmission, etc. The real-time property of the data acquisition and processing system in our design is guaranteed by using the techniques of dual CPU architecture, HPI, Ping-Pong buffering, which is of reference significance to practical engineering design.

References

1. Zhao F, Zhou Y (2005) Real time optimization analysis of phased array radar system simulation. J Syst Simul 17(8):2001–2003
2. Wang W, Han S (2010) Song: research on real-time data acquisition system of vessel nosise multisensor. Ship Sci Technol 32(6):60–63
3. Manivannan M, Kumaresan N (2011) Design of on-line interactive data acquisition and control system for embedded real time applications. In: IEEE emerging trends in electrical and computer technology (ICETECT), 23–24 Mar 2011
4. TMS320C6713B (2006) Floating-piont Digital Signal Processor (Literature Number: SPRS294B). Texas instruments
5. MSP430x14x (2004) Mixed Signal Microcontroller (Literature Number: SLAS272F). Texas instruments
6. Analog-to-Digital Converters ADS8364 (Literature Number: SBAS219C) (2006) Texas Instruments
7. TMS320C6000 DSP Host Port Interface Reference Guide (Literature Number: SPRU578C). Texas Instruments (2006)
8. Ren H, Yin W, Hu F (2010) Contrast and analysis of three measuring methods for phase difference based on LabView. Sci Technol Eng 10(1):263–268

Part IV
Electrician Theory and New Technology

Chapter 52
Design of a Dual-Band Microstrip Antenna for WLAN

Yating Gan, Suling Wang and Jie Yu

Abstract A dual-band microstrip antenna with an E-shaped slot is proposed in this paper. Numerical simulation is performed for the characteristics of the antenna with the software HFSS 10.0. The simulated results show that the antenna operates at 2.40–2.50 GHz and 5.07–5.43 GHz, and the bandwidth achieves 100 and 360 MHz, respectively, which can cover ISM dual-band completely. In addition, the antenna has good radiation performance and simple configuration, which is suitable for wireless communication system.

Keywords Microstrip patch antenna · Dual-band · WLAN

52.1 Introduction

In recent years, with the development of the wireless local area networks (WLAN) technology, it will be a trend to deploy a multi-specification and multi-mode wireless network, which can be compatible with IEEE 802.11a, IEEE 802.11b and IEEE 802.11g. Therefore, the WLAN antenna with multi-band, wide bandwidth and small size will be required.

Y. Gan (✉) · S. Wang · J. Yu
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, Henan, China
e-mail: gan323@163.com

S. Wang
e-mail: tjslwang@163.com

J. Yu
e-mail: 13210575@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_52,

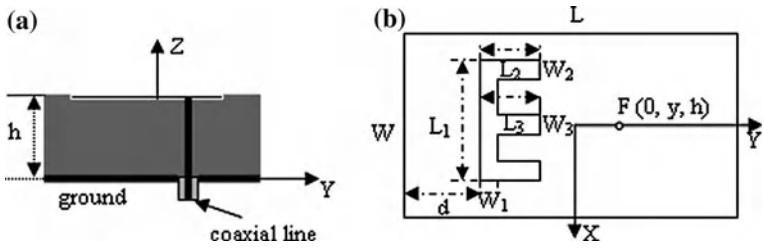


Fig. 52.1 Layout of the proposed dual-frequency microstrip patch antenna. **a** Cross-sectional view. **b** Top view

Microstrip antennas have been widely applied and researched in many fields thanks to its low profile, easy to conformal with other objects which load them, simplicity in manufacture, easy to perform circular polarization, dual-band, dual polarization and other advantages. Dual-band microstrip antennas are usually implemented with multi-patch, multi-mode in single patch and single patch loaded. A mass of development in dual-band microstrip antennas for WLAN have been presented by scholars at home and abroad. Such as the T-shaped element antenna [1], the F-shaped element antenna loaded with slots [2], the microstrip patch antenna loaded with an H-shaped slot [3], the triangular microstrip patch antenna loaded with slots [4] and so on. However, the above methods also have some shortcomings. For instance, it is complicated to fabricate them even if the fabrication is not a question, and bandwidth, gain and other performance of microstrip antennas may be not very satisfied.

A microstrip patch antenna loaded with an E-shaped slot is proposed in this paper. The overall dimension of this antenna is 30 × 48 × 6 mm. The simulation results show that it has a good double resonant characteristic, and its center resonant frequencies of are 2.45 and 5.25 GHz, respectively. The bandwidth covers ISM dual-band completely when VSWR is equal to 2. And it has good radiation performance in its wavelength band.

52.2 Configuration of the Antenna and Theoretical Analysis

The configuration of the microstrip antenna proposed is shown in Fig. 52.1, where (a) is the cross-sectional view and (b) is the top view.

Same as the common patch antennas, double-face circuit board covered with copper is adopted in this design. The microstrip on the upper surface is the radiation patch. A foam material, Rohacell IG-51, with ε_r of 1.06 and a thickness of 6 mm, is adopted as base-chip, and an E-shaped slot is loaded on the radiation patch. The size and location of the slot and the microstrip patch decide the radiation performance. The E-shaped slot is symmetrical about the *Y*-axis, and the length and width of the slot are L_1, L_2, L_3 and W_1, W_2, W_3 as depicted in

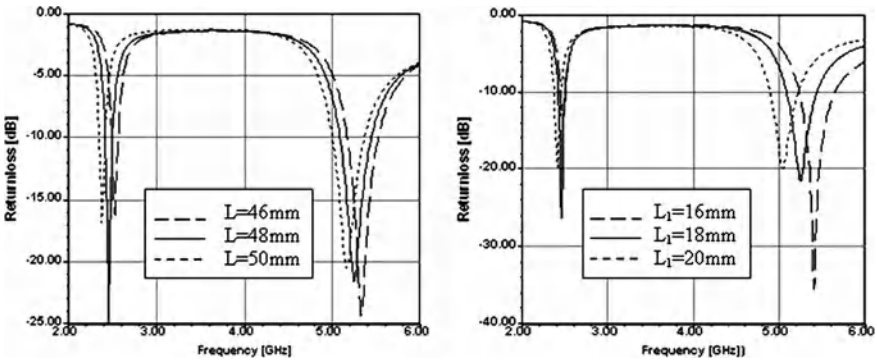


Fig. 52. 2 Frequency characteristic of return loss varying with the parameters L and L_1

Fig. 52.1b. The distance between the bottom margin of the slot and the patch boundary is d. The antenna is excited using coaxial line with 50 Ω as shown in Fig. 52.1a. The feed point is on the *Y*-axis, and its coordinate is F (0, y, h).

Dual-frequency operations are behaved by etching a narrow E-shaped slot dividing the current distribution of the rectangular microstrip patch antenna. The current path between the feed point F and the radial edges of the rectangular patch gets stretched out because of the slot. According to the microstrip theory, the resonant frequency of the patch, the lower frequency of the dual-band microstrip patch antenna, is reduced. At the same time, the E-shaped slot blocks the current fed to the microstrip patch by the coaxial line, so there will be an imaginary radiation patch in areas surrounded by the slot, and the higher resonant frequency of the proposed antenna is produced accordingly.

The parameters of the E-shaped slot: the coordinate position, the length of the slot, the width of the slot and so on, all of which will influence the far-field direction pattern. When the parameters change, the far-field direction pattern will change. Therefore, the double frequency property of the dual-band microstrip patch antenna can be adjusted through changing the size of the patch and the size and location of the E-shaped slot.

52.3 Simulation Research of the Antenna Performance

According to the design, a series of parameter simulation are carried on for further research and analysis of the antenna performance. The effects of parameters L, L_1, *d*, W_1 on the antenna performance are discussed subsequently.

The curve diagrams of frequency characteristic of return loss varying with the parameters L and L_1 are shown in Fig. 52.2. From the figure, we can see that the two resonant frequencies of the antenna both decrease with the increase of L, but

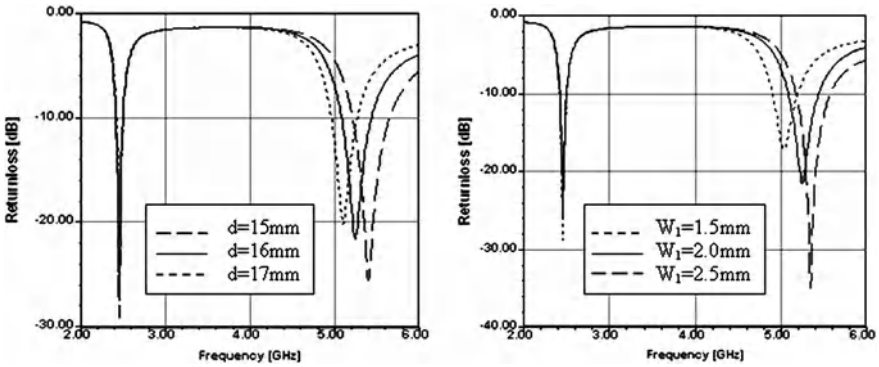


Fig. 52.3 Frequency characteristic of return loss varying with the parameters d and W_1

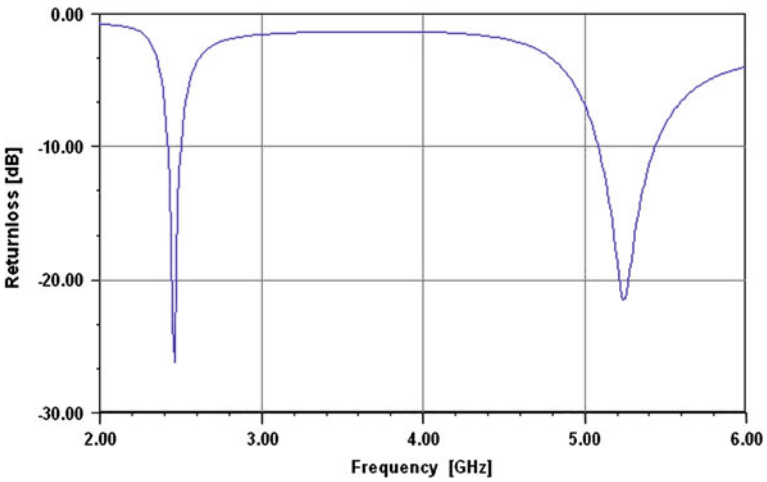


Fig. 52.4 Return loss of the dual-band microstrip patch antenna

the lower resonant frequency varies obviously. Therefore, the change of the L parameter mainly influences the lower frequency. Similarly, the two frequencies all decrease with the increase of L_1. The difference is that the higher frequency changes obviously. That is to say, the L_1 parameter has greater effect on the higher frequency than the lower one.

The curve diagrams of frequency characteristic of return loss varying with the parameters *d* and W_1 are shown in Fig. 52.3. We can find that the higher frequency decreases with increase of the distance between the bottom margin of the slot and the patch boundary d. However, the variation of d has less effect on the lower frequency. In contrast, the higher frequency increases with increase of W_1. But the variation of W1 has less effect on the lower frequency too.

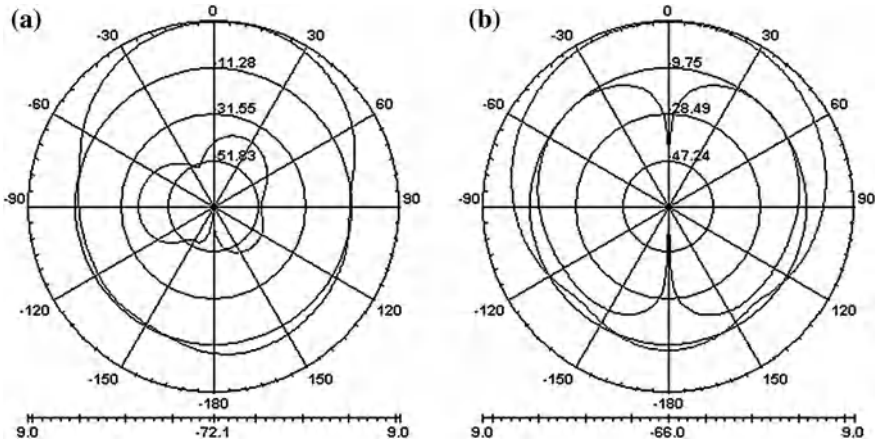


Fig. 52.5 Radiation patterns of the dual-band microstrip patch antenna at $f = 2.4$ GHz. **a** E-plane. **b** H-plane

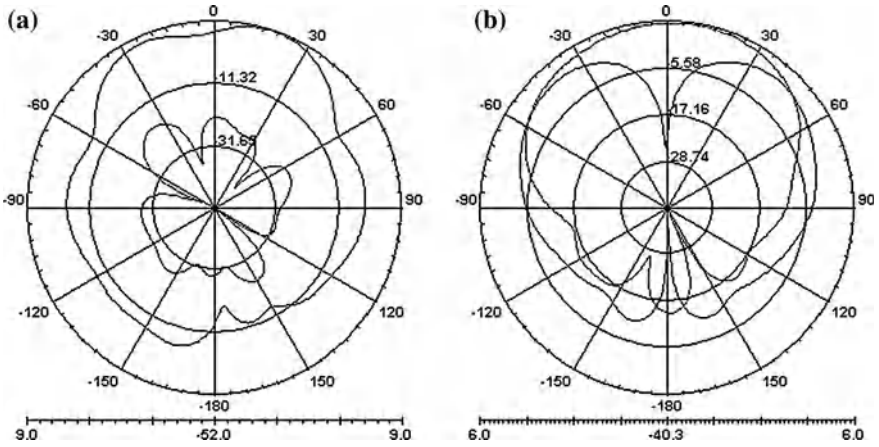


Fig. 52.6 Radiation patterns of the dual-band microstrip patch antenna at $f = 5.2$ GHz. **a** E-plane. **b** H-plane

With the variation of the parameters, the location of the feed point F should be changed to achieve impedance matching. When the distance between the feed point F and the origin coordinates is 2–3 mm, a good impedance match can be achieved.

52.4 Design

A large number of simulation and optimization are carried on in this paper. The size of the dual-band microstrip patch antenna is determined as: W = 30 mm, L = 48 mm, L_1 = 18 mm, L_2 = 9 mm, L_3 = 8 mm, $W_1 = W_2 = W_3$ = 2 mm, d = 16 mm, F (0, *y*, *h*) = (0, 3,6 mm).

The S parameters of the proposed antenna are showed in Fig. 52.4. For the 2.4 GHz ISM band (2.4000–2.4835 GHz), the patch antenna has a sufficient bandwidth of approximately. 100 MHz. For the 5.2 GHz ISM band (5.15–5.35 GHz), the center resonance frequency of the antenna is 5.25 GHz, and the band width achieves 360 MHz.

The E-plane and H-plane radiation patterns of the two operation frequency are shown in Figs. 52.5 and 52.6, respectively. The dual-band antenna has a good cross-polarization property, for which, the E-plane is better than the H-plane obviously.

52.5 Conclusion

In this paper, a dual-band microstrip antenna with an E-shaped slot is proposed. This microstrip antenna is compact, cost-effective and simple to manufacture, and has a sufficient bandwidth and a good cross-polarization property for all relevant WLAN standards, IEEE 802.11a, IEEE 802.11b and IEEE 802.11g. The proposed antenna might have a good application prospect for WLAN in nowadays or future.

References

1. Su CM, Chen MS (2005) Shorted T-shaped monopole antenna for 2.4/5.2 GHz WLAN Operation. Microw Opt Technol Lett 41(03):202–203
2. Wang YS, Lee MC, Chung SJ (2007) Two PIFA-related miniaturized dual-band antenna. IEEE Trans Antennas Propag 55(03):805–811
3. Kim SM, Son JM, Kim HJ, et al (2005) Design and implementation of dual-band square patch antenna for wireless LAN of 2.4 GHz and 5.7 GHz. Microw Conf Proc APMC 5:4–7
4. Augustin G, Shynu SV, Mohanan P et al (2006) Compact dual-band antenna for wireless access point. Electron Lett 42(9):502–503

Chapter 53
Research on Electromagnetic Compatibility of LED Display of Mine Safety Monitoring System

Zhigang Liu

Abstract Based on the introduction of the constitution and working principle of LED in mine monitoring system display screen, based on switching power supply, the panel itself, from the oscillator, signals line layout analysis of three aspects of the monitoring system of LED display by electromagnetic interference, combined with the specific test case, separately from the suppression of the radiated and transmitted to harass two to carry out rectification, finally through the signal line, power line and ground wire layout solve the antenna vertical polarization direction on radiation the unqualified.

Keywords Electromagnetic compatibility · Communication interface · Mine safety monitoring system · Surge protection circuit

53.1 Introduction

Point's giving out light emitting diode (LED) monitor is a mineral to apply a kind of most monitors with the monitor. No matter take out check or mineral to use the monitor examination of entrusting of manufacturer in NCTC in, the author discovers because LED the monitor interference is over-standard, but cause the product become available in the market period extension or be ordered by the

Z. Liu (✉)
Wuhan University of Technology, Wuhan 430070, China
e-mail: liuzhigang@swfu.edu.cn; xwqandfcc@163.com

Z. Liu
School of Computer and Information, Southwest Forestry University, Kunming 650224, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_53,

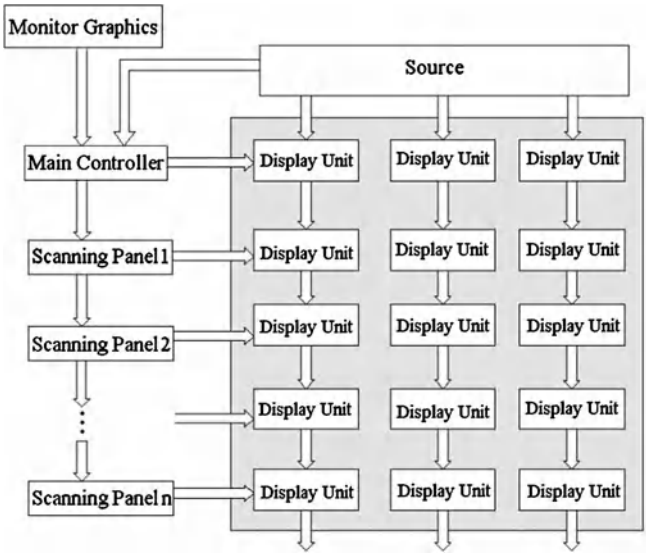


Fig. 53.1 Structure of LED display screen

government to take back have been already sold of product, the responsibility becomes whole change, use a monitor for business enterprise and mineral applied the customer all bring many troubles. For this, author exclusively to LED electromagnetism and permitted sex to carry on more thorough research, the achievement help obtained a lot of products pass an examination and hope to contain certain instruction function to the manufacturer follow-up product quality control.

53.2 The Constitution and Working Principle of LED Display of Mine Safety Monitoring System

The mineral is used monitor LED monitor mainly from lord controller, scan plank and show to control unit, the power and LED a monitor body several parts constitute. Its basic working principle is as Fig. 53.1 shows.

The load controller shows card to obtain a hold the each color bright number of degrees of each pixel from the calculator according to, then assign to some scanning planks, each scanning plank is responsible for control LED monitor up of some line (row), but each (row) LED of top show that the signal then goes way to pass profession with the string of each one shows that controlling unit class is allied to deliver, each one shows to control unit to directly face to a LED monitor body.

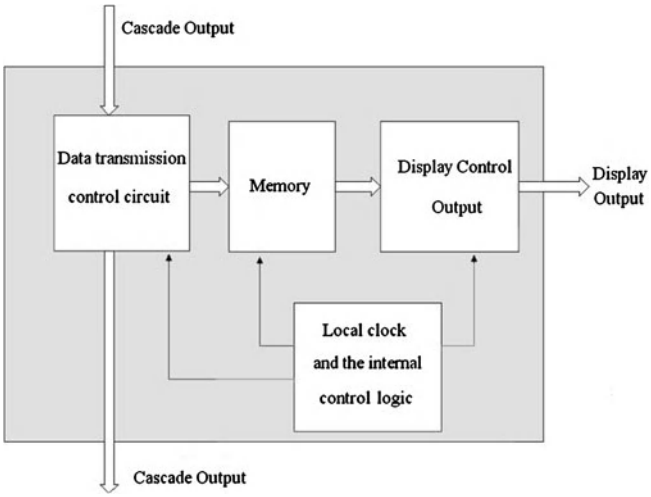


Fig. 53.2 Function diagram of scanning board

Scan plank to have accept up under the Qi of function, it on the other hand receives the lord knothole control signal, on the other hand, it belong to the data of this class and deliver to that own each one shows to control unit, at the same time still not belong to the data of this class get down a class unite of the scanning plank deliver, such as Fig. 53.2 show.

53.3 The Analysis on the Electromagnetic Interference on LED Display of Mine Safety Monitoring System

The mineral is mainly used the electromagnetism interference of the monitor of monitor LED to result in as follows from a few factors:

(1) The electromagnetism interference of switch the power creation. The mineral is used monitor LED monitor great majority all what to use is switch the power, however lead a moment in the switch tube of switch the power, the transformer entry-level coil will produce to flow out to flow very greatly, and appear a higher wave to flow out peak electric voltage in the entry-level coil both ends; Break to open a moment in the switch, because of the leaking of entry-level coil Ci, cause a part of energies did not deliver to two coils from a coil, hoard at leaking part of energies in the feeling will and the pole electric capacity, electric resistance formation of switch take the Shuai of having the peak to reduce to flap to concuss and fold to add the pass power failure that takes care of in the switch to press up, the formation closes peak electric voltage. This interference signal will conduct importation exportation to carry, the formation conducts interference.

Secondly, Gao repeatedly transformer entry-level coil, switch tube and filter Gao that the electric capacity constitutes repeatedly the switch electric current wreath road may also produce bigger space radiation, the formation radiates interference. As for this kind of interference have already described in detail in this chapter Sect. 1, no longer launch concrete treatise here.

(2) Monitor of crystal flap to result in of interference. Is crystal to flap mainly is in order to show that the system provides a basis frequency, it is a kind of interference signal Gao Ci Xie's wave of period pulse signal producing. This kind of interference signal will match to hold through various ways Ou inside of in the cable, then with bad mold or total mold formation radiation or conduct interference. Along with the development of science and technology, monitor's breaking velocity will be more and more high, crystal the basis frequency flapping also immediately become Gao, from crystal flap causable Gao Ci Xie wave the interference is stronger.

(3) Because signal line inside the monitor and power cable arrange more airtight, the interference signal will be also holding mutual Ou of of inside track cable to match, finally radiate space or go out through a power cable conduction and form interference.

53.4 Rectification Measures on the Electromagnetic Interference on LED Display of Mine Safety Monitoring System

53.4.1 Interference Analysis and Repress Measure on Conduction

The mineral is influenced this test a little bit greatly with the power of the monitor of monitor LED, the quality of the power function directly relates to whether this index sign is qualified. Sometimes also exist the power to do electromagnetism alone and permit experiment is qualified, once packing the whole machine, because of other parts in the whole machine at a certain repeatedly order to have stronger interference of signal, filtering of the power an unit can not completely filter in addition to the interference's signal, thus cause to conduct a test result of over-standard. For the power port harass electric voltage of over-standard, main should from as follows a few paths solve.

First, expel the interference of the power factor, under the situation that the condition allows can take out the power alone, link a sum to settle a pure Zu load to carry on experimenting. If the detection was over-standard to repeatedly order to have no, explained that the harassment that the frequency orders came from a lord to control plank at first. At this time, should put point in the lord to control knothole to filter wave up, lord's controlling a main interference source in plank is crystal to flap, should to crystal flap to carry on to goodly filter wave and connect ground; Secondly, it is crystal to flap is also radiation to shoot to test item an

Table 53.1 Corresponding relationship between load impedance, source impedance and filter structure

Source impedance	Circuit structure	Load impedance
High	C, π, multi-level π	High
High	Γ, multi-level Γ	Low
Low	Inverse Γ, multi-level, inverse Γ	High
Low	L, multi-level T	Low

over-standard main factor. The check lord controls plank in crystal flap and the signal line connect ground, the power to connect whether the ground is good, at promise this what time under the all intact circumstance, if conducting to test is still unqualified, explain interference signal to is really very strong, can carry to add the whole filter or *X* and *Y* electric capacity in the importation of the power at this time, strengthen filtering of the power a function. Remarkable BE, when the filter chooses should pay attention to inserting of the filter different frequency to exhaust a circumstance, also want according to the height of the source resistance and the load resistance, choose filtering of accommodation an electric circuit, it to should relate to see Table 53.1.

When the filter installs, ensure the filter connects a ground of good, just can promise that the filter has to the bad mold and the total mold interference signal like this of repress a function be not let up.

If need to install *X* and *Y* electric capacity, should notice, the *X* and *Y* electric capacity fuse should be possibly short, in order to prevent weaken filtering of *X*, *Y* electric capacity a function because fuse inductance and electric capacity lead greatly. The *X* and *Y* electric capacity permits a value to also want to combine to unqualifiedly and repeatedly select to choose a suitable value. Under the general circumstance, the electric capacity is worth more big, its closing frequency will be more low.

In fine, regardless adopt what kind of measure, have to order knot to put together with the unqualified frequency of product measured, just can effectively of cancelation interference.

53.4.2 Suppression Measures on Radiated Emission

(1) Filter: such products as digital pulse signal the existence of radiation emission is generally stronger than that can be connected next to the crystal filter bypass capacitors and the crystal to ensure good grounding, grounding resistance as small as possible. If conditions permit, you can use through the spread spectrum crystal clock to ensure the circuit does not affect the condition and the oscillation frequency in the frequency range in a smaller deviation occurs, a single frequency energy is dispersed, so the overall radiation is reduced. This spread spectrum technology is actually mine our research section supervisor in

the switching power supply electromagnetic compatibility, using the frequency jitter is the technology in the mining monitor different components of the application. Because the presence of common mode currents cause radiation emission is the main factor is too large, so you can also display the power cord and internal signals between the various display units online using a ferrite bead high frequency common mode interference currents filtering. Of course, the choice to combine ferrite bead insertion loss versus frequency curve to choose the right size, the effect will be good.

(2) Shield: for the display has been formed, the shield is to prevent radiation emission is an important measure, general mine supervisor monitors the front panel is displayed by the array of LED lights, so on the front panel shielded shielding effect is the whole key to good and bad. Recommend the use of sheet metal box made of metal mesh screen front panel, LED lights that between rows, columns and columns using electrical conductivity between the metal mesh better, so will the overall radiation emission energy have a certain attenuation. You can also the cabinet front panel and control panel to add a metal shield between the grid, the effect would be better. Board inside the box between each scan signal line to make use of shielded cable, and ensure its energy and metal shield box equipotential.

(3) Grounding: to make a good shielding box, make sure the box shielded enclosure and the earth equipotential. Metal box seams should overlap to ensure good resistance to overlap as small as possible. Especially the back cover of the box to connect closely with the cabinet as possible to prevent the seam from the hole caused by secondary emission. Second, the main control board in the ground should be as independent as possible of the crystal to avoid the interference caused by common ground. Main control board and scan board ground should be as wide as possible, it is best to ground into a mesh-like fabric, this will reduce the signal return area, effectively suppress radiated emissions.

53.5 Rectification Cases

Case 1: Figure 53.3 is a mine with a LED display monitor in the transmission line test failed N spectrum. The products in the 16.25 and 21.25 MHz Department, more than the average limit, the distribution of the interference signal was comb, which is the digital pulse signal of high-order harmonic interference typical symptoms. View screen-circuit structure, found in the main control board and the signal line of the crystal, the crystal and not be filtered, resulting in excessive conduction. Different manufacturers for a piece of the main control board to pass the tests, shown in Fig. 53.4.

Case 2: A mine supervisor LED display in the radiation tests, the antenna vertical polarization in the direction of excessive, as shown in Fig. 53.5.

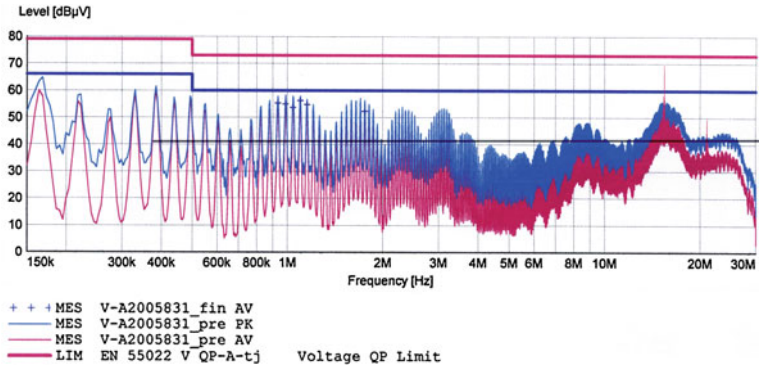


Fig. 53.3 Spectrum diagram of N-line disqualification of conduction disturbance

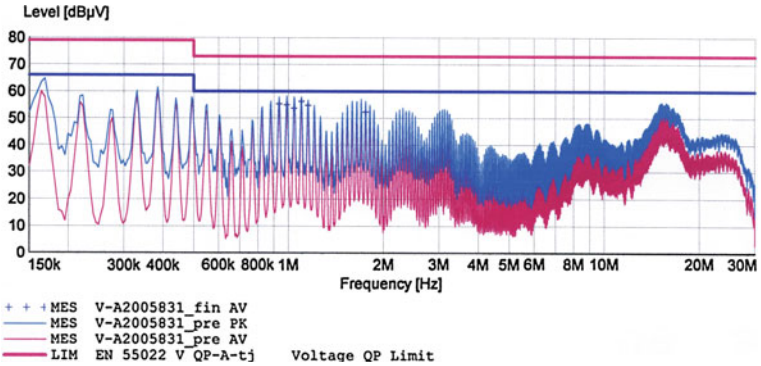


Fig. 53.4 Spectrum diagram of N-line qualification of conduction disturbance

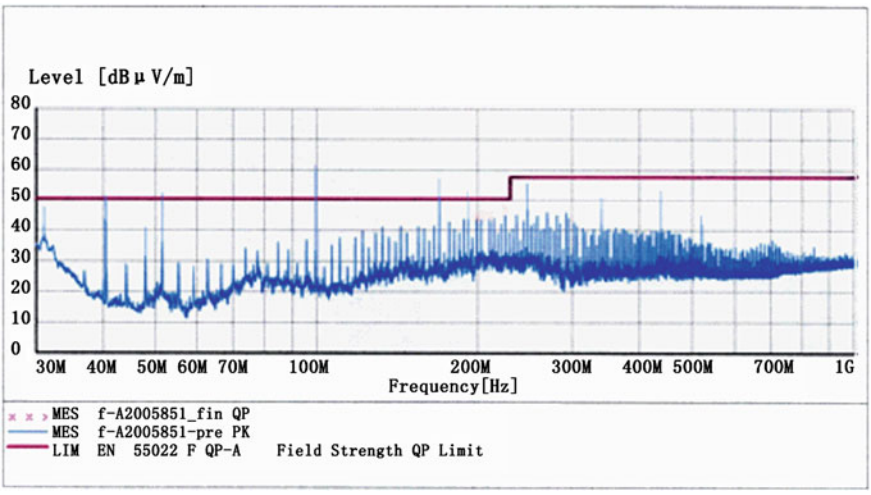


Fig. 53.5 Spectrum diagram of disqualification of radiation disturbance (antenna vertical polarization direction)

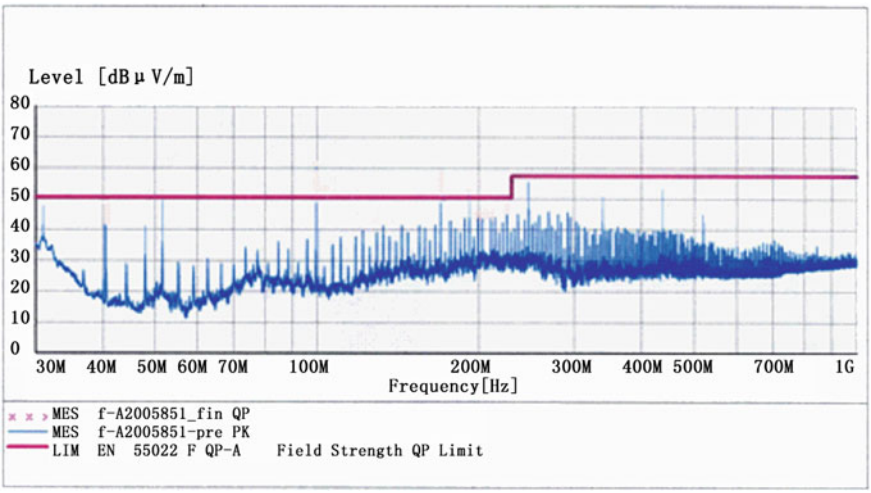


Fig. 53.6 Spectrum diagram of qualification of radiation disturbance (antenna vertical polarization direction)

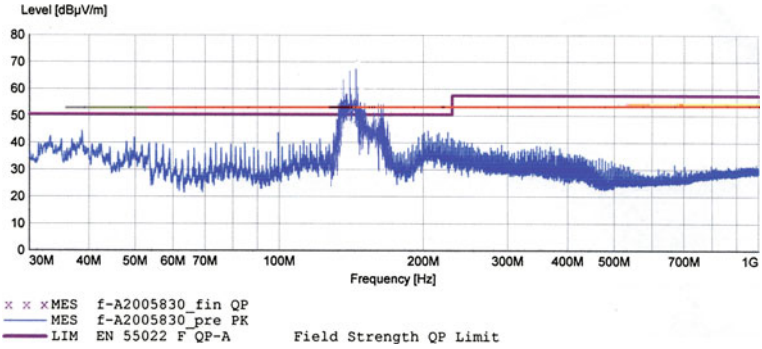


Fig. 53.7 Spectrum diagram of disqualification of radiation disturbance (antenna vertical polarization direction)

Corrective measures, in the main control board and scan board and the signal line between the scanning signal line between the board plus a suitable ferrite bead filter, and display between the front panel and control board strengthened with metal mesh screen. After rectification, the test results are satisfactory, as shown in Fig. 53.6.

Case 3: A mine supervisor LED display in the radiation tests, the antenna vertical polarization in the direction of excessive, as shown in Fig. 53.7.

The product is more than the 150–170 MHz band limit, the curve is evident in the background between 100 and 200 MHz higher, slightly shaking the power cord, interfering signals change more obvious, so the inference from the power line interference signal is radiated out. Corrective measures, the display's power cord

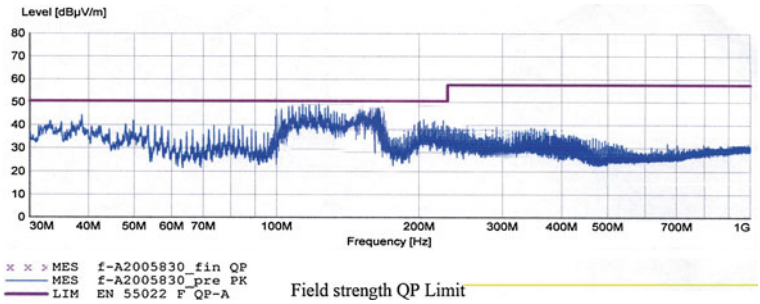


Fig. 53.8 Spectrum diagram of qualification of radiation disturbance (antenna vertical polarization direction)

Fig. 53.9 Winding displacement of ground wire in LED

into live magnetic power cord. After rectification, to pass the tests, shown in Fig. 53.8.

These are the rectification of electromagnetic compatibility LED panel encountered most frequently exceeded the failure of the three cases, there is a fault in the LED display is made of air electrostatic discharge immunity test test, LED garbled or does not show. Electrostatic air discharge test is shot through with the monitor shell hit the mine voids or gaps, to test static medium of mine through the air monitor internal circuit. Mine the gap generally monitor the installation of the LED housing, the keyboard, the upper and lower shell machine shell connection.

Air discharge test is generally encountered a problem and the cable design. As the LED in the shell and lower shell of the circuit board is generally connected by cable to the air discharge test than when the phenomenon mainly electrostatic charge through the cable crosstalk to the LED, you can connect via cable design avoided.

53.6 Conclusion

In conclusion, adding between the various signal GND and the power cord can not only effectively solve such problems, but can also can reduce the effect of radiation which is also useful indications. Theory is best able to do when each signal line has

a ground wire, but it is not possible, hence we must aim for atleast a single ground wire between two signal linest as shown in Fig. 53.9.

Acknowledgments This work was supported by the National Natural Science Foundation of China (NO. 50804038 and70971104). In the process of doing experiments,I have got some help from others. Especially, I would love to thank Engineer Wuhan (ChinaNational Computer Products Quality Supervising Test Center) for t some data and experiments done in their labs. I would also like to thank Xiaozhi Wang for photoprocessing.

Chapter 54
Lateral Air-Gap Control of a Novel Detent-Force-Based Magnetic Suspension System

Xiaozhuo Xu, Xudong Wang, Haichao Feng and Jikai Si

Abstract A novel detent-force-based magnetic suspension system (DMSS) is introduced. To guarantee system's stability and safety of running, the lateral air-gap must be controlled to prevent the collisions between the mover and track. The lateral air-gap control system with differential control strategy of double-electromagnets is studied. To achieve no static error control of steady-state position, closed-loop state-space model of double-electromagnets is established. The fuzzy PID algorithm is used to achieve the control process based on fuzzy control theory. A fuzzy PID controller is designed to have the ability of online self-tuning PID parameters. Simulation result shows that the DMSS with the fuzzy-PID algorithm has good dynamic and anti-jamming performance.

Keywords Magnetic suspension system · Detent-force · Lateral air-gap control · Fuzzy PID

Xiaozhuo XU: doctoral candidate, research interest: special motor theory and control.

X. Xu (✉) · X. Wang · H. Feng · J. Si
College of Electrical Engineering and Automation,
Henan Polytechnic University, Shiji road NO. 2001,
Jiaozuo 454003, Henan, China
e-mail: xxz@hpu.edu.cn

X. Wang
e-mail: wangxd@hpu.edu.cn

H. Feng
e-mail: fhc@hpu.edu.cn

J. Si
e-mail: sijikai@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_54,

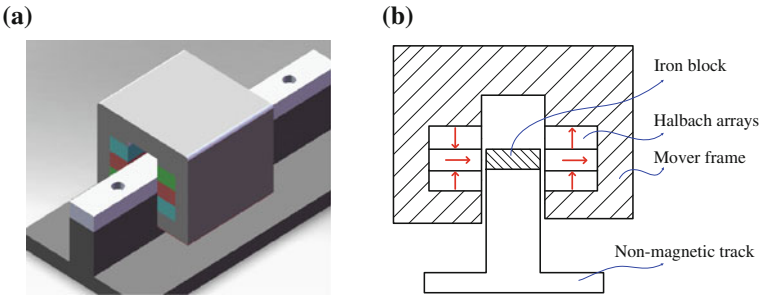


Fig. 54.1 Schematics of the DMSS. **a** Three-dimensional solid figure. **b** Cross-sectional profile

54.1 Introduction

The principle of minimum reluctance has been used to achieve one or more degrees of freedom passive suspension. It brings levitation applications with simple structures and low costs, and has been widely used in bearingless motors, low-speed maglev train CMS03A, and GA maglev vehicles suspension system [1–3]. New detent-force-based magnetic suspension system (DMSS) works with the principle of axial minimum reluctance, having the advantage of large suspension force, energy saving, and low cost, and have a wide range of applications in damping devices and magnetic levitation train [4]. The lateral air-gap of DMSS is extremely important to ensure stable operation of the entire levitation system. In the suspension process, the possible lateral offset caused by random external disturbance forces in DMSS may lead to car-body instability in lateral direction, so this must be considered and controlled. In this paper, differential control of double-electromagnet lateral air-gap control system is studied. The mathematical model of lateral air-gap control system is established, and fuzzy-PID control strategy is discussed to achieve the fast and robust control of DMSS. The PID parameters on-line self-adaptive is also studied.

54.2 Detent-Force-Based Magnetic Suspension System

Figure 54.1 shows the basic structure of new detent-force-based magnetic suspension system. Two parallel Halbach arrays are placed inside of U-shape frame of DMSS. The permanent magnets magnetized in the direction of the arrow shown in Fig. 54.1. The composite track located in the middle of two Halbach arrays is composed of iron block and non-magnetic track frame.

The main parameters of DMSS are listed in Table 54.1. Finite element model of DMSS are built by the FEA method. The axial length of the model is about 100 mm. Figure 54.2 shows the magnetic field distribution. Figures 54.3 and 54.4 show

Table 54.1 Main structural parameters of DMSS

Items	Value	Unit
Iron-core thickness	10	mm
Iron-core width	16	mm
Middle PM thickness	10	mm
Upper and lower PM thickness	10	mm
PM width	15	mm

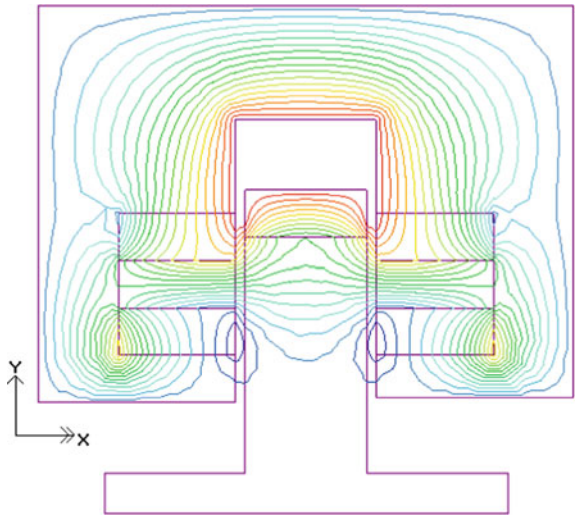


Fig. 54.2 Magnetic field distribution

the characteristics of vertical suspension force and lateral force, respectively. It can be seen that the maximum suspension force almost do not change when the mover has the lateral offset, but the lateral force increases linearly by the lateral offset increase. In other words, the suspension force and lateral force is decoupled and it is helpful to design the control system only considering the lateral force.

54.3 Modeling of Lateral Air-Gap Control System

Electromagnet differential control method is adopted in modeling of lateral air-gap control system [5, 6]. As the lateral attractive force is influenced by mover position, system structure and leakage magnetic field and so on, it is difficult to establish the mathematical model of lateral attractive force and here it is treated as the interference factor. Structure diagram of lateral double-electromagnet single point control system is shown in Fig. 54.5.

In the figure, $c_L(t)$and $c_R(t)$are the left and right air-gap between electromagnet and stator track. $u_{cL}(t)$ and $u_{cR}(t)$ are control voltages. $i_L(t)$ and $i_R(t)$ are the currents of left and right electromagnet. $F_{gL}(t)$ and $F_{gR}(t)$ are electromagnetic

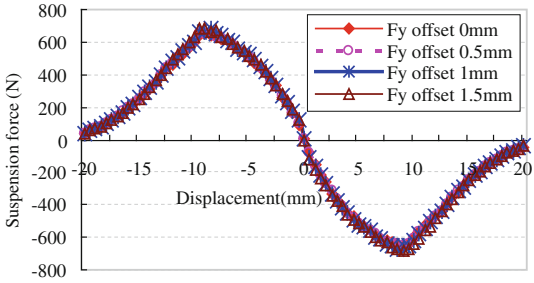


Fig. 54.3 Suspension force characteristics

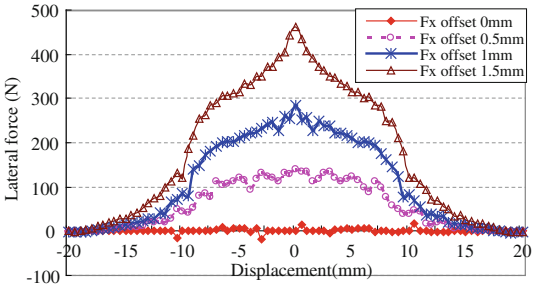


Fig. 54.4 Lateral force with different lateral offset

forces. F_d is external interference force in lateral direction. The mover is assumed as a rigid object and the lateral deformation of the track is ignored.

When the left and right gap equal to $c_{0,}$ the right and left electromagnet current are same and lateral attractive force have the same values but opposite direction. The mover will deviate from the lateral equilibrium position when the interference force is applied to the mover at time t. Air-gap sensors detect the air-gap offset and the controller will transform these displacement signals into control signals which then is changed into current Δi through the power amplifier. The electromagnet current increases by Δi in the side of gap increased, while the electromagnet current decreases by Δi in the side of gap decreased. Thus, a lateral back force is produced to bring the mover to the equilibrium position. In this stage, the system is double-electromagnet control mode. And when the lateral interference force is bigger than a certain value, leading to the current of one electromagnet is zero, then only another electromagnet works, the system becomes into a single electromagnet control mode.

Lateral air-gap control system is mainly working in the dual-electromagnet mode, so the dual-electromagnet differential model is applied to design controller. Suppose the system is in lateral equilibrium position, while, $c_L(t) = c_R(t) = c_0$, $i_L(t) = i_R(t) = i_0, u_{cL}(t) = u_{cR}(t) = u_0$, the vacuum permeability μ_0, quality m, magnet pole area A, number of turns N, resistance R. Taking the left electromagnet for example, the system model is established in the balance after linearization.

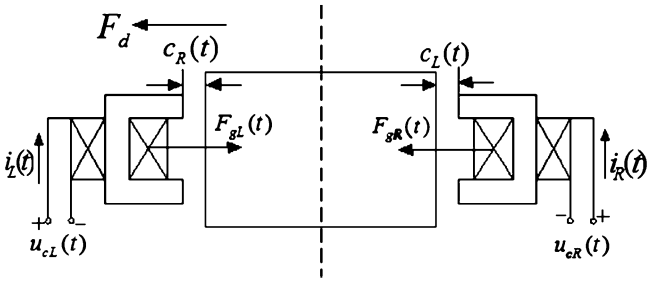


Fig. 54.5 Schematics of the lateral air-gap control system

$$m\Delta\ddot{c} = 2k_c\Delta c - 2k_i\Delta i + \Delta F_d \tag{54.1}$$

$$\Delta u = R\Delta i + L_0\Delta\dot{i} - k_i\Delta\dot{c} \tag{54.2}$$

$$\text{where}, k_c = F_c(i_0, c_0) = \frac{\mu_0 AN^2 i_0^2}{2c_0^3}, k_i = F_i(i_0, c_0) = \frac{\mu_0 AN^2 i_0}{2c_0^2}, L_0 = \frac{\mu_0 AN^2}{2c_0}.$$

Selecting $x = [\Delta c \quad \Delta\dot{c}]^T$as state variables and choosing current as a control input, we get state-space model express as.

$$\begin{aligned} \begin{bmatrix} \Delta\dot{c} \\ \Delta\ddot{c} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 2\dfrac{k_c}{m} & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta c \\ \Delta\dot{c} \end{bmatrix} + \begin{bmatrix} 0 \\ -2\dfrac{k_i}{m} \end{bmatrix} \cdot \Delta i \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{bmatrix} \Delta c \\ \Delta\dot{c} \end{bmatrix} \end{aligned} \tag{54.3}$$

In order to achieve no static error control of steady-state position, here we introduce a new state variable $\int \Delta c dt$, and use control law $\Delta i = k_I \int \Delta c dt + k_P \Delta c + k_D \Delta\dot{c}$, then obtain the following closed-loop state-space model.

$$\begin{aligned} \begin{bmatrix} \Delta c \\ \Delta\dot{c} \\ \Delta\ddot{c} \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -k_I\dfrac{2k_i}{m} & \dfrac{2k_c}{m} - k_P\dfrac{2k_i}{m} & -k_D\dfrac{2k_i}{m} \end{bmatrix} \cdot \begin{bmatrix} \int \Delta c dt \\ \Delta c \\ \Delta\dot{c} \end{bmatrix} \\ y &= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} \int \Delta c dt \\ \Delta c \\ \Delta\dot{c} \end{bmatrix} \end{aligned} \tag{54.4}$$

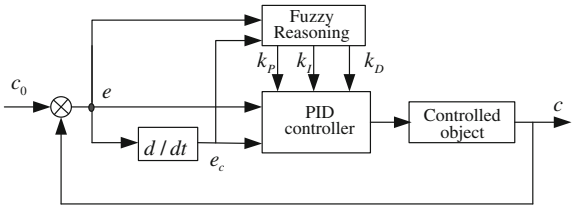


Fig. 54.6 Structure of fuzzy adaptive PID control system

54.4 Fuzzy-PID Controller Design

Figure 54.6 shows the typical structure of fuzzy adaptive PID control system [7]. It is composed of a conventional PID controller and a fuzzy controller. Choosing error e and error change rate e_c as the input, the three PID parameters k_P, k_I, k_D are online real-time correcting through the fuzzy control rules to meet the PID parameters self-tuning requirements at different times.

Finding the fuzzy relationship between the PID parameters and the error e and error rate e_c is the key of fuzzy control system design. It is necessary to reduce the overshoot, increase system responsiveness, while improving system stability. Therefore, we reference to the previous practical experience, establish the initial fuzzy rule table shown in Table 54.2.

Known that, when the system is stable, the two sides have the same air-gap 2 mm. Here, the basic domain of deviation e is taken as $[-0.5, 0.5]$ mm, the basic domain of e_c is $[-0.1, 0.1]$ m/s, and the fuzzy quantification domain of e and e_c are defined as $(-5, 5)$ and divided into seven fuzzy sets expressed as {NB, NM, NS, ZE, PS, PM, PB}, denoted by {negative big, negative middle, negative small, zero, positive small, positive middle, positive big}. Similarly, fuzzy sets of output PID parameters k_P, k_I and k_D, are also expressed as {NB, NM, NS, ZE, PS, PM, PB}, the corresponding scale factors are selected based on the basic theory and experience. In the paper, Mamdani fuzzy reasoning algorithms is adopted to achieve the fuzzy reasoning. And gravity method is used to do fuzzy solution.

54.5 Control Simulation

Simulation parameters: electromagnet pole area A = 450 mm^2, turn of winding $N = 250$, resistance $R = 0.7\ \Omega$, mover mass $m = 60$ kg. Quiescent bias current is 7 A, equilibrium air-gap is 2 mm.

In view of the above system parameters, on the equilibrium position, we obtain the following values, $k_c = 108240$, $k_i = 30.9251$, $L_0 = 0.0088$. Using the pole assignment method, conventional PID parameters are obtained: $k_D = 3492.3$, $k_P = 1932600$, $k_I = 548860000$. Due to robustness of PID control and flexibility of fuzzy controllers, it requires less precision of the initial values to PID parameters, so the conventional PID control parameters k_P, k_I, k_D are selected as the initial values.

Table 54.2 Fuzzy control rule table of k_P

e	e_c						
	NB	NM	NS	ZE	PS	PM	PB
NB	PB	PB	PM	PM	PS	ZE	ZE
NM	PB	PB	PM	PS	PS	ZE	NS
NS	PM	PM	PM	PS	ZE	NS	NS
ZE	PM	PM	PS	ZE	NS	NM	NM
PS	PS	PS	ZE	NS	NS	NM	NM
PM	PS	ZE	NS	NM	NM	NM	NB
PB	ZE	ZE	NM	NM	NM	NB	NB

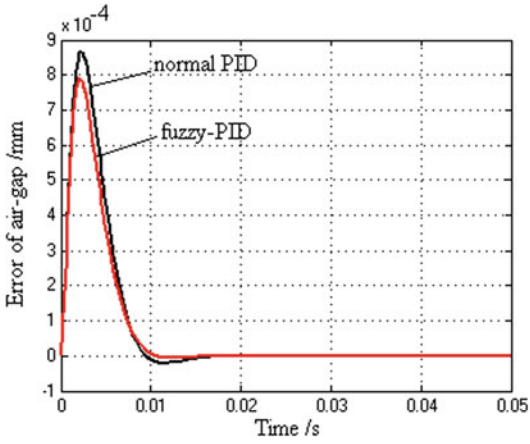


Fig. 54.7 Comparison of lateral air-gap control

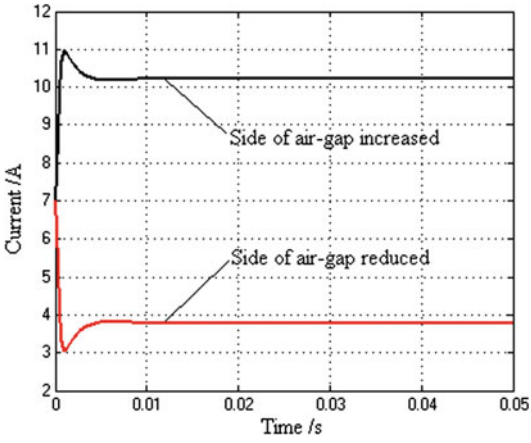


Fig. 54.8 Current curve of fuzzy-PID control

At 0 s, 200 N static interference force is applied to the lateral air-gap control system. Figure 54.7 is the comparison of lateral air-gap control effect with conventional PID and fuzzy PID. Figure 54.8 shows the current curve of fuzzy-PID control.

Conventional PID and fuzzy-PID controllers are designed with integral air-gap feedback to eliminate static error. From the simulation results can be seen, the conventional PID and fuzzy-PID control system can achieve lateral stability control, but fuzzy-PID controller has smaller overshoot, faster response, higher steady-state precision, and stronger robustness.

54.6 Conclusion

The detent-force-based Magnetic Suspension System is a novel suspesion system. The lateral air-gap control of DMSS is the key influence factor to the stable operation. Model of the lateral air-gap control system is established by adopting differential control strategy of double-electromagnets. A fuzzy-PID controller is studied based on fuzzy control theory. The parameters of fuzzy-PID are discussed and comfirmed. Simulation analysis has proved that the fuzzy-PID controller has good dynamic performance and robustness, the proposed suspension control system provides an important theoretical basis for future hybrid suspension device in the actual application.

Acknowledgments. The work was possible due to the support given by the National Science Foundation of China (61074095), Research Fund for the Doctoral Program of Higher Education (NO. 20104116120001), Innovation Project for University Prominent Research Talents of Henan Province (2005126) and Natural Science Program of Henan province(2011A470003), especially the Innovation Talents Foundation of Henan Province (104200510021) to the authors.

References

1. Liao QX, Wang XL, Deng ZQ (2008) 2D analysis of passive suspension characteristics of bearingless slice motors. Electric Mach Control 12(2):117–121
2. Liu SK, Ni HY, Zhang KK (2007) Analysis on guidance ability for mid and low speed Maglev train. Electric Drive Locomot 2:36–38
3. Li YG, Chang WS, Yan YZ (2006) Analysis and comparison of new Maglev transport technology in USA. Electric Drive Locomot 3:6–9
4. Xu XZ, Qin XF, Wang XD (2011) Characteristics analysis of a novel detent-force-based magnetic suspension system. Micromotors, vol 44, No 1. Xi'an Micromotor Research Institute, Xi'an, China, pp 58–60
5. Hao AM, She LH, Chang WS (2008) Adaptive controller design of guidance system of EMS high speed Maglev train. Control Eng China 15(2):116–119
6. Ahmad AK, Saad Z, Osman MK, Isa IS, Sadimin S, Abdullah SS (2010) Control of magnetic levitation system using fuzzy logic control. In: 2nd international conference on computational intelligence, modelling and simulation. IEEE Computer Society, Hoes Lane United States, pp 51–56
7. Yang X, Gao Y, Yang B, Du JG (2009) Research on the control of the magnetic suspension stifness in Gantry Machining Center. Modular Mach Tool Autom Manuf Tech 1(5):64–67

Chapter 55
Performance Test and Analysis of Phase-Locked Loop

Guoyong Wang, Wenli Wang and Xiaohui Li

Abstract Methods that can improve the frequency stability of phase-locked loop are proposed. In order to analyze the presented methods, a Phase-Locked Loop is developed and an experiment system is constructed to test the performance of the phase-locked loop. Experiments are carried out to testify the validity of the given methods. After analyzing the experiment results, the following conclusion can be drawn that adjusting the value of frequency division coefficients, increasing shielding measure of VCO circuit and changing the time constant of filter can obviously enhance the performance of phase-locked loop.

Keywords Phase-locked loop · Frequency stability · Frequency division coefficients · Shielding measure

55.1 Introduction

Phase-locked loop (PLL) is a closed-loop automatic control system which is able to track the input signal phase, with excellent features of carrier tracking, modulation tracking, low-threshold, etc. Currently, it has been widely used in

G. Wang (✉) · W. Wang · X. Li
National Time Service Center, Chinese Academy of Sciences,
Xi'an 710600, China
e-mail: wangguoyong321@163.com

G. Wang
Graduate University of Chinese Academy of Sciences,
Beijing 100190, China

G. Wang · W. Wang · X. Li
Key Laboratory of Time and Frequency Primary Standard,
National Time Service Center, Chinese Academy of Sciences,
Xi'an 710600, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_55,

modulation and demodulation, frequency synthesis, carrier tracking, TV color subcarrier extraction, synchronous detection of analog and digital signals, bit synchronization in data transmission and many other fields [1, 2]. Applicable scope from TV, radio in daily life to satellite orbit velocity, deep space exploration and others in space field [3], PLL has played an irreplaceable role. The performance of PLL can not only affect people's daily life, but also has a profound impact to aerospace and other high-tech fields.

The frequency stability is particularly important in many indicators of the PLL [4]. This paper studies the methods to improve the frequency stability of the PLL, and builds an experimental system to test effect of PLL performance in the case of no reference signal, increasing the shielding measure and adjusting the frequency division coefficients. Through analysis of the data, the proposed frequency stability improvement methods were fully validated.

55.2 PLL Principle

55.2.1 PLL Basic Principle

PLL is an error control system about phase negative feedback. The operating principle of the PLL is shown as Fig. 55.1. It consists of three basic functional blocks: a voltage-controlled oscillator (VCO), a Phase detector (PD), a loop filter (LF).

The phase detector is a phase comparison device that is used to detect the phase difference $\theta_e(t)$ between the input signal (reference signal) and output signal (feedback signal), and output error signal $U_d(t)$.

The loop filter should not only filter out the high frequency component of error signal, output DC signal, show low pass properties, but also play a decisive role in loop parameters adjustment.

VCO is a voltage-frequency conversion device, its oscillation frequency changes when the input control voltage changes linearly.

If the phase error $\theta_e(t)$ was not zero initially, the PD would develop a nonzero output signal $U_d(t)$. After some delay, the loop filter would also produce a finite signal $U_c(t)$. This would cause the VCO to change its operating frequency in such a way that the phase error finally vanishes [5].

55.2.2 PLL Circuit Design

The overall design block diagram of PLL the experiment used is shown in Fig. 55.2. The PLL uses 10 MHz signal as an input signal, and adopts a first-order active PI filter as a loop filter, meanwhile chooses Oven Controlled Crystal Oscillator (OCXO) for VCO and selects the phase frequency detector (PFD) as a

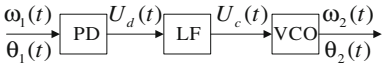


Fig. 55.1 Block diagram of PLL principle

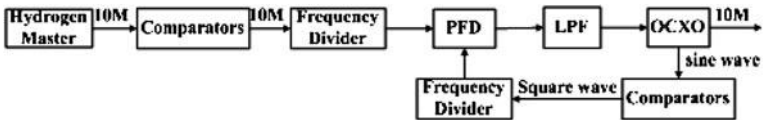


Fig. 55.2 Block diagram of PLL overall design

Phase detector. After the sine input signal comes from Hydrogen clock through high-speed comparator, a square wave signal is produced, then do frequency division to the square wave signal. After frequency division to the input and output signals, two same frequency signals do Phase Discrimination in the PD, then they pass first-order active PI filter. Later on the DC voltage is output to control the OCXO, make the OCXO lock to external reference signal.

55.2.3 Methods to Improve the Performance of Phase-Locked Loop

(1) Component Selection

PLL usually uses Phase detector, VCO, frequency divider, capacitor resistors and other components. For Phase detector selection, the phase Discrimination characteristics must be first considered, that is selecting matching Phase detector according to the characteristics of signal for phase discrimination; Second, the slope of the phase discrimination characteristics, the output voltage amplitude, working frequency, impedance of input and output, parasitic component output, threshold and other factors should also be paid attention to. Considering the above factors, in addition phase frequency detector (PFD) has the characteristic of capture range of theoretically infinite, in fact, it is corresponding to the actual frequency range generated from VCO, the PFD is used in this paper as the PLL's PD [5, 6].

For the VCO selection, frequency stability, control sensitivity, control characteristics of linearity, the wide linear region, phase noise, jitter, voltage frequency stability, load frequency stability, power consumption and other indicators must be considered. Based on the above factors, VCO of PLL in this paper adopts the OCXO which second level frequency stability is 2.0E−12.

There are many types of capacitors and resistors. To keep stable performance of loop filter, the PLL adopts low temperature coefficient metal film resistors and monolithic capacitors as loop filter components; meanwhile, in order to

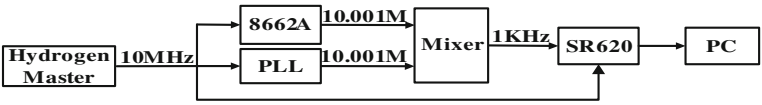


Fig. 55.3 Block diagram of test system

save volume, the chip fixed capacitors and resistors are chose except loop filter and power.

(2) Adjustment of the Loop Filter Parameters
In the PLL, once some type of the PD and VCO are selected, their indicator parameters have been fixed and owned less adjustable space, so only the loop filter parameters can be wide-range adjusted. After select filter type and order numbers, estimate first the filter parameters with consideration of damping factor, phase margin, synchronous bandwidth and other factors, and then do fine-tuning. Because first-order active PI filter which is an active lead-lag filter has two independent time constant, the natural frequency and damping factor can be independently chose and loop gain can meet a good tracking required size. Therefore PLL in this paper adopts a first-order active PI filter as the loop filter [5].

(3) Effect of Adjustment of Frequency Division Coefficient
Frequency division coefficients will impact the entire loop gain, so its adjustment is also essential. The PLL in this paper adopts phase discrimination with same frequency and increases the frequency division module to meet the optimal phase discrimination range of PDF and loop gain adjustment [7].

55.3 PLL Performance Test System

In order to study ways of improving the frequency stability of PLL, this paper makes full use of self-developed PLL, builds a performance test systems, which structure is shown in Fig. 55.3. PLL frequency stability measurement system consists of 10 MHz signal source generated by the hydrogen maser, PLL (no shielding measure, frequency division coefficients is 1,000), SR620 counter, 8662A signal source, mixer, computer, stable 32 data acquisition software.

The test system uses the principle of the beat frequency method for testing [4, 8], 10 MHz output signal generated by the hydrogen maser in National Time Service Center is respectively access to 8662A, PLL, SR620 as their reference signal, then adjust the 8662A to generate 10.001 MHz signal, later on access the output signal of PLL and 8662A to mixer and access the 1 kHz beat frequency signal to SR620, now can make use of computers for data collection and processing.

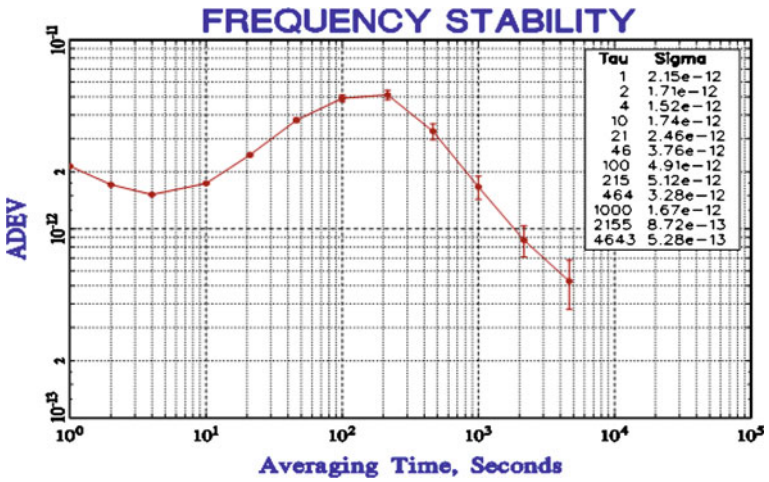


Fig. 55.4 Frequency Stability with reference signal but no shielding measure when N = 1,000

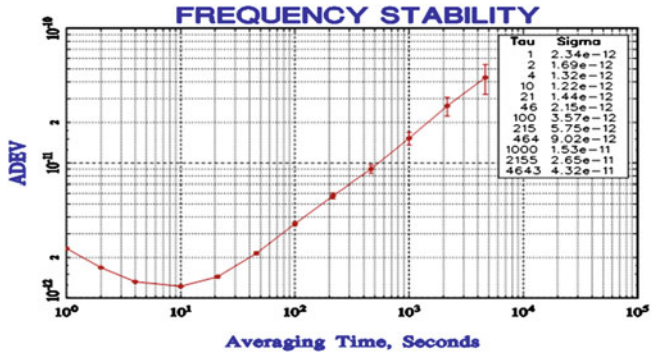


Fig. 55.5 Frequency Stability without reference signal and shielding measure when N = 1,000

55.4 PLL Test Result Analysis

(1) Test of effect of external reference on PLL performance.
Connection as Fig. 55.3, gets the measured PLL frequency stability with external reference signal, as is shown in Fig. 55.4; when removing the PLL external reference signal, gets the PLL frequency stability results shown in Fig. 55.5.

(2) Analysis of effect of increase shielding measure on the PLL performance.
Connection as Fig. 55.3, increase the shielding to VCO in the PLL circuit, measure the frequency offset and stability, get the result as shown in Figs. 55.6 and 55.7.

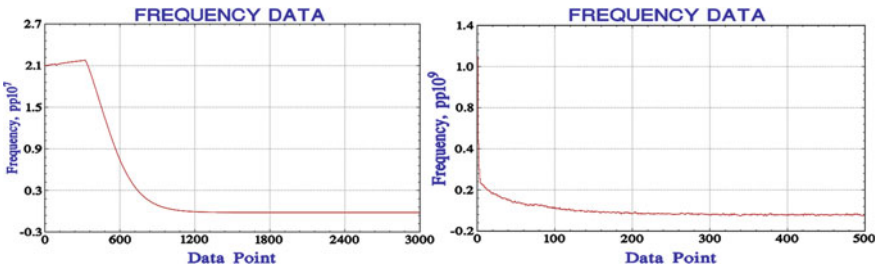


Fig. 55.6 Frequency offset with reference signal and shielding measure when N = 1,000 and N = 10

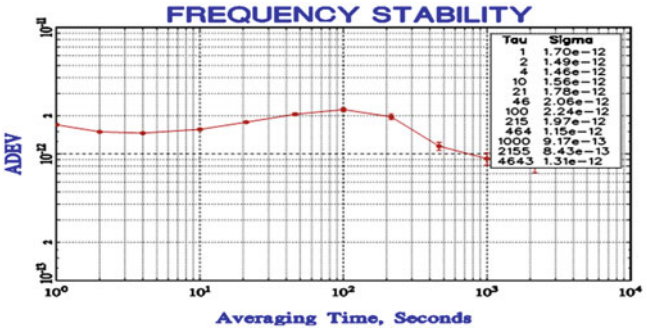


Fig. 55.7 Frequency stability with reference signal and shielding measure when N = 1000

(3) Analysis of effect of adjusting Frequency division coefficients on PLL performance.
Connection as Fig. 55.3, increase the shielding to VCO in the PLL circuit, decrease frequency division coefficients N of PLL frequency divider from 1,000 to 10, measure the frequency offset and stability, get the result as shown in Figs. 55.6 and 55.8.

Compare PLL frequency stability with reference signal or not by Figs. 55.4 and 55.5. From Fig. 55.4, it can be seen that the short-term frequency stability of PLL output signal is better when having external reference signal, and Allan deviation achieved 2.15E−12 at 1 s which is close to the used crystal oscillator. Meanwhile long-term frequency stability of PLL is significantly improved with respect to the situation of without external reference. For an example, when T = 4643S, Allan deviation is 5.28E−13 with external reference signal while Allan deviation is 4.32E−11 without external reference signal, that is to say that long-term frequency

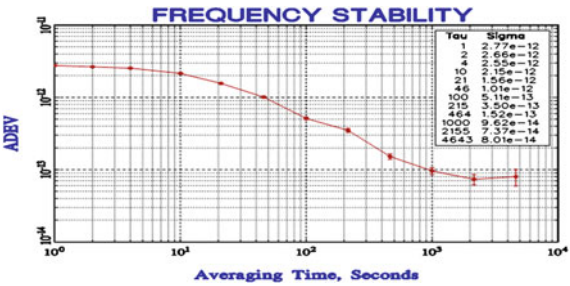


Fig. 55.8 Frequency stability with reference signal and shielding measure when N = 10

stability of PLL increases nearly two orders of magnitude. In fact, if there is no external reference, the PLL output signal is the output signal of crystal oscillator with the addition of noise, largely reflecting the frequency stability in the case of crystal oscillator with the addition of noise.

Compare PLL frequency stability with shielding measure or not in VCO circuit by Figs. 55.4 and 55.7. By compare, PLL frequency stability has a obvious improvement when increasing shielding measure. It can be included that increasing shielding measure to crystal oscillator can improve frequency stability to some extent.

Compare PLL frequency stability when adjusting Frequency division coefficients by Figs. 55.7 and 55.8. It can be seen, the PLL short-term frequency stability when N = 1,000 is slightly better than N = 10, but the PLL long-term frequency stability is poorer than N = 10. From Figs. 55.7 and 55.8, it is clearly seen that the frequency stability curve has an upward trend when N = 1,000, 100 s later it declines rapidly, while the frequency stability curve keeps a downward trend for a long time when N = 10.

Compare PLL frequency offset when adjusting frequency division coefficients by Fig. 55.6, it can be drawn that the capture time Tp is about 1,800 S when N = 1,000, while Tp is about 180 S when N = 10. It can be concluded that decreasing frequency division coefficients can greatly decrease the capture time.

When frequency division coefficients N becomes smaller, phase demodulation frequency become correspondingly larger, that is the adjusting number of VCO's input voltage per unit time will increase so that PLL is easy to fast lock. Therefore, capture time is much shorter when N = 10 than N = 1,000, but when N decreases, frequent adjustment of VCO input voltage will increase the instability of crystal oscillator, and result in poorer short-term frequency stability when N = 10 than N = 1,000.

55.5 Conclusion

Based on the above analysis we can see: increasing shielding measure to crystal oscillator can improve frequency stability to some extent; After a larger frequency division coefficients is used, the PLL output signal has a high short-term frequency

stability and capture time is long; After a smaller frequency division is used, the PLL output signal has a high long-term frequency stability and fast lock. Actually, PLL makes full use of the respective advantages of reference signal and crystal oscillator to meet its short-term frequency stability and long-term frequency stability expectations. All the methods that adjusting frequency division coefficients, increasing shielding measure and adjusting the time constant of loop filter in practice can improve the performance of PLL.

Acknowledgments The authors would like to thank Cuimin Zhao, Danni Wang, Huijun Zhang, Zhixiong Zhao, Sheng Tang, YA Liu, Longxia Xu, Shaohua Shi, Yanrong Xue, Feng Zhu, Zhenzhong He and Yuwei Li for their help. This work was supported by the National Natural Science Foundation of China (Grant Nos.11033004 and 61001076.)

References

1. Zhang JS, Zheng JY, Wan XP (1994) The phase lock technique. Xi Dian Press (1994) 1–11, 129–180
2. Yang M, Jin X, Li S et al (2001) The Application of ISP technology in digital phase-locked loop. Chin J Sci Instrum 22:312–315
3. Greenhall CA, Kirk A, Stevens GL: A multichannel dual-mixer stability analyzer: progress report*.In: 33rd annual precise time and time interval (PTTI) systems and applications meeting, Nov 2001
4. Yan D (2007) The noise bandwidth and dead time in the measurement of short-term frequency stability. J Astronaut Metrol Meas 27:17–20
5. Best RE writing (2007) Li Y translation: Phase-locked loops: design, simulation, and applications (5th edn). Tsinghua University Press (2007) 1–90
6. Zhang Y, Yang Y (2004) Application of digital PLL circuit in synchronous control of low frequency. J Astronaut Metrol Meas 25:896-905
7. Wang C, Jin X, Jin Z et al (2009) Effect of digital frequency dividers on spurs and PM noise in PLL. J Astronaut 30:2334–2338
8. Sui J, Chen Z, Gao W et al (2010) A high precision frequency stability meter based on frequency offset locked loop beat method. J Astronaut Metrol Meas 30:22–24

Chapter 56
Design of Quaternary Logic Circuits Based on Multiple-Valued Current Mode

Haixia Wu, Shunan Zhong, Qilong Cai, Qianbin Xia and Yueyang Chen

Abstract In order to improve the performance of arithmetic VLSI system, a novel structure of quaternary logic gates is proposed based on multiple-valued current mode (MVCM) by using dynamic source-coupled logic (SCL). Its key components, the comparator and the output generator, are both based on differential-pair circuit (DPC), and the latter is constructed by using structure of DPC trees. The pre-charge evaluates logic style, makes steady current flow cut off, thereby greatly saving the power dissipation. The combination of multiple-valued source-coupled logic and differential-pair circuit makes its power lower and its structure more compact. The performance is evaluated by HSPICE simulation with 0.18 μm CMOS technology. The power dissipation, transistor numbers and delay are superior to corresponding binary CMOS implementation. Multiple-valued logic is the potential solution for the high performance arithmetic VLSI system in the future.

Keywords Multiple-valued logic · Multiple-valued current mode · SCL circuit

56.1 Introduction

It is now an era for the rapidly developing integrated circuits. However, more transistors means more power dissipation and more heat produced [1, 2]. It is known that one of the key circuit techniques for reducing power dissipation is to use multiple-valued current mode circuitry [3, 4]. MVCM integrated circuits have two potential advantages: one is to reduce the wiring complexity. A wire in such

H. Wu (✉) · S. Zhong · Q. Cai · Q. Xia · Y. Chen
School of Information and Electronics,
Beijing Institute of Technology, Beijing, China
e-mail: wuhaixia@bit.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_56,

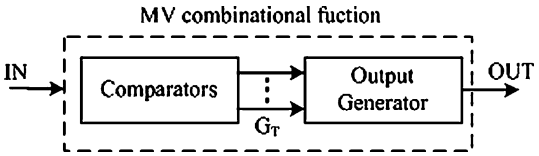


Fig. 56.1 Diagram of MVCM combinational function

circuit can carry more than 1-bit data, so the number of wires can be effectively reduced. Another one is to reduce the number of active devices in arithmetic and logic unit. The frequently used linear sum operation can be performed simply by wiring with no active devices [5–7]. In fact, several arithmetic circuits combined with MVCM logic circuits have been designed and fabricated, resulting in better performance compared with the corresponding binary implementation [8, 9].

As a typical logic unit of an arithmetic circuit, quaternary logic AND, OR and XOR circuits have been designed in this paper. The combination of multiple-valued source-coupled logic, dynamic logic and differential-pair circuit is taken for reducing the power dissipation together with device and interconnection counts while keeping a high-speed switching capability.

56.2 Basic Components of SCL MVCM Circuit

The dynamic multiple-valued source-coupled logic-based (MVSCL) circuit has an attractive feature that its input voltage swing is small enough while maintaining a high current-driving capability. A DPC-based voltage comparator can operate with small signal voltage swing, so that low-power operation will be realized while maintaining high-speed switching. A pre-charge evaluates logic style can make steady current flow cut off, which results in great reduction of the power dissipation. Figure 56.1 shows basic structure of MV combinational circuits, which consist of comparators and output generator [7, 10].

56.2.1 Comparator

Figure 56.2 shows the circuit diagram of a comparator. The current input is converted into voltage signal by the I–V convertor, which then goes as a differential-pair circuit input and is compared with the threshold voltage V_T. The output signal G_X and its inverted signal $G_{X'}$ are generated corresponding to the input I_X and the relationship can be described as (56.1) [9].

$$(G_X, C_{X'}) = (0, 1) \quad \text{if} \quad V_X > V_T \tag{56.1}$$

$$(G_X, C_{X'}) = (1, 0) \quad \text{if} \quad V_X < V_T$$

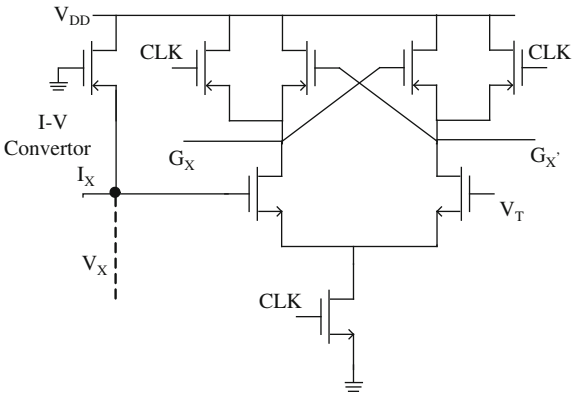


Fig. 56.2 Comparator circuit

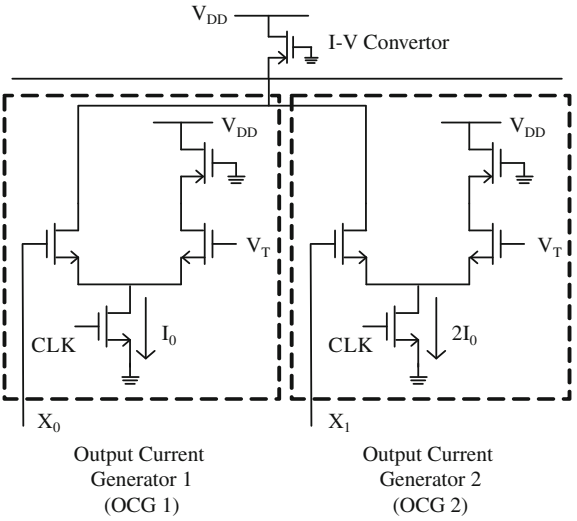


Fig. 56.3 Circuit diagram of the output generator and four-valued encoding

Table 56.1 Four-valued encoding

Binary	Four valued
0 0	0
0 1	1
1 0	2
1 1	3

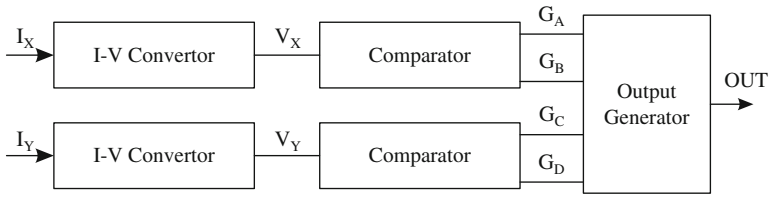


Fig. 56.4 Implementation diagram of quaternary logic circuit

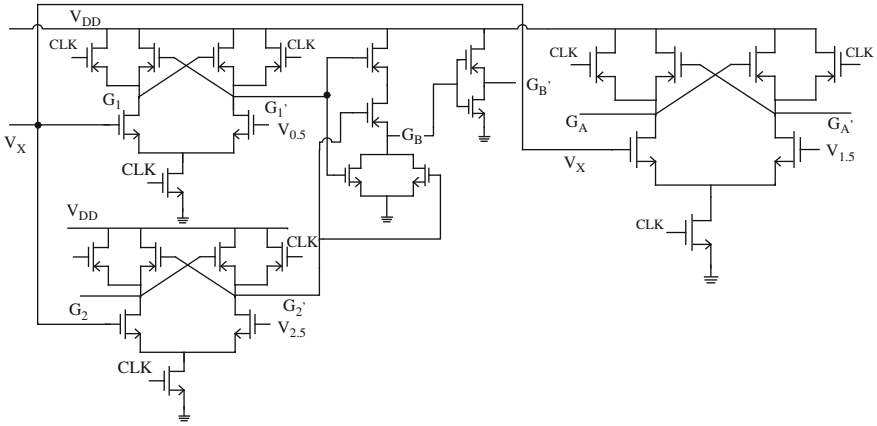


Fig. 56.5 Current diagram of the four-valued comparator

56.2.2 Output Generator

Figure 56.3 shows the diagram of an output current generator (OCG) which consists of two SCL circuit blocks [10]. In each block, the branch is charged whether close or not due to the control signal and the two parts work independently. The current flows generated by the current sources can directly sum into a four-valued output by wiring so that the two branches can combine into four output results 0, I_0, $2I_0$, $3I_0$. Table 56.1 shows the relationship between the output of the OCG block and the 2-bit binary input signal.

56.3 Design of Quaternary Logic Circuits

Figure 56.4 shows the block diagram of implementation for the quaternary logic circuit. The four-valued current input I_X and I_Y is converted into voltage signal V_X and V_Y, and then put into the next circuit. The comparator compares the voltage

input with the threshold voltages and leads to corresponding binary differential-pair output signals G_A, G_B, G_C, and G_D, which will later be used to control the output generator.

56.3.1 Four-Valued Comparator

The four-valued comparator is implemented by three basic comparators proposed in Fig. 56.5. As a result, three threshold voltages, $V_{0.5}$, $V_{1.5}$, and $V_{2.5}$, are defined to divide the voltages into four levels so that the four-valued logic can be transferred into binary form. Figure 56.4 shows the diagram of the four-valued comparator. The function is described in Eqs. (56.2)–(56.3):

$$G_A = \begin{cases} 0 & \text{In} > 1.5 \\ 1 & \text{In} < 1.5 \end{cases} \tag{56.2}$$

$$G_B = \begin{cases} 0 & \text{In} < 0.5 \text{ or In} > 2.5 \\ 1 & 0.5 < \text{In} < 2.5 \end{cases} \tag{56.3}$$

56.3.2 The Design of Output Current Generator

The use of tree structure makes it quite easy to design an output circuit for a specific function. Figure 56.6 shows the tree structure of two quaternary variables. The 16 branches, from 0 to 15, present the different combination of the control signals, G_A, G_B, G_C, G_D, which are generated by the comparator. Only one of the branches can be ON in one time.

Figure 56.6 also shows the truth table, K-map, and the tree structure of quaternary AND logic. The coding rules between the control signals and the branch label in the tree structure are shown in the K-map. In comparison with the truth table and K-map, it can be found how to design the OCG block, for example, for AND logic, blanch 5, 6, 9, 10 should be chosen for I_0 block and blanch 10, 11, 14, 15 for $2I_0$ block. Put the two blocks together by wiring and then comes the final OCG circuit for AND logic. In this way, logic OR and XOR can also be implemented.

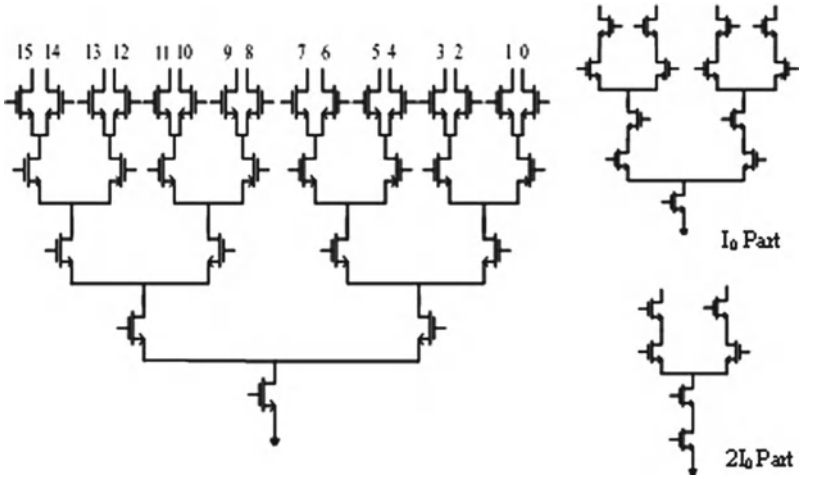
Figure 56.7 shows the truth tables and OCG circuits for four-valued OR and XOR logic.

(a)

AND	X				
		0	1	2	3
Y	0	0	0	0	0
	1	0	1	0	1
	2	0	0	2	2
	3	0	1	2	3

K-map		G_D			
		G_C'		G_C	
G_B	G_A'	0	1	3	2
		4	5	7	6
	G_A	12	13	15	14
		8	9	11	10

(b)



(c)

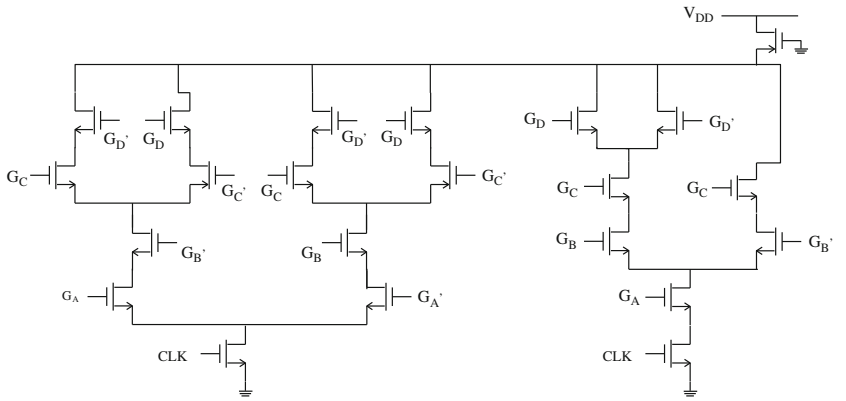
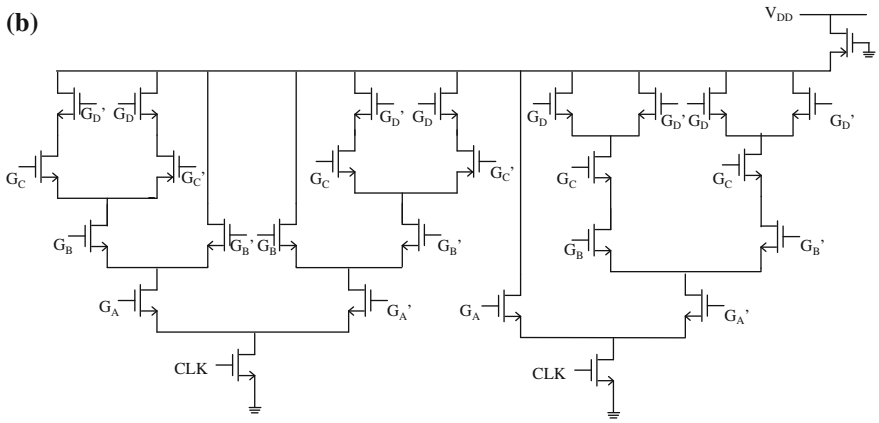


Fig. 56.6 OCG design of AND logic. **a** Truth table of AND logic and K-map. **b** Tree structure of two quaternary variables. **c** AND output generator

(a)

OR	X				
		0	1	2	3
Y	0	0	1	2	3
	1	1	1	3	3
	2	2	3	2	3
	3	3	3	3	3

XOR	X				
		0	1	2	3
Y	0	0	1	2	3
	1	1	0	3	2
	2	2	3	0	1
	3	3	2	1	0



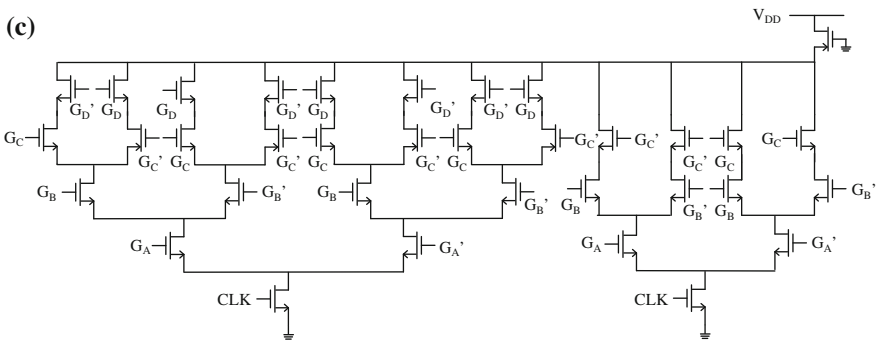


Fig. 56.7 OCG design of OR and XOR logic. **a** Truth table of OR and XOR logic. **b** OR output generator. **c** XOR output generator

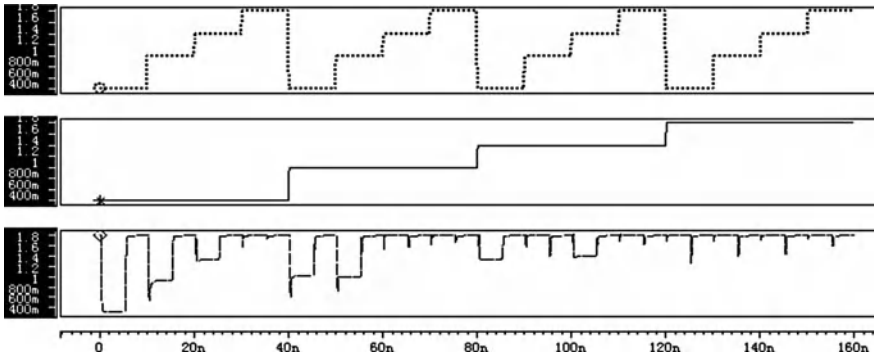


Fig. 56.8 Simulate results of AND logic

Table 56.2 Current value, voltage value and MVL

Current (uA)	Current logic	Voltage (V)	Voltage logic
0	0	1.8	3
50	1	1.4	2
70	2	1	1
90	3	0.4	0

Table 56.3 Threshold logic value and threshold voltage

Threshold logic (T)	Threshold voltage (V)
0.5	0.6
1.5	1.2
2.5	1.6

Table 56.4 Performance of AND, OR and XOR

	AND	OR	XOR
Power@100 MHz (uW)	234	297	268
Number of Trs.	76	81	89
Delay (ps)	89	143	134

56.4 Evaluation

The proposed quaternary logic units are evaluated by HSPICE simulation in a 0.18 μm CMOS technology and the supply voltage of 1.8 V. Figure 56.8 shows the parts of simulating result of AND. Input signal X and Y are complete voltages in four-valued logic. Table 56.2 shows the current value and voltage value corresponding to multiple-valued logic in proposed implementation. Table 56.3 shows relationship between the threshold logic value and the threshold voltage. Table 56.4 summarizes the performance of the proposed logic units. The power dissipation of AND, OR, XOR circuits are 234, 297 and 268 μW at 100 MHz, the transistor counts are 76, 81, and 89, the delay are 89, 143, and 134 ps, respectively.

56.5 Conclusion

In this paper, the four-valued logic units based on MVCM are proposed for high performance arithmetic and logic unit in VLSI system. The performance of the logic units is better in comparison with the CMOS implementation. The combination of multiple-valued source-coupled logic, dynamic logic and differential-pair circuit seems to be an approach for low-power, more compact and high-speed arithmetic VLSI system.

As a future research work, it is also important to save the static power dissipation and to immune noise. If these problems are successfully solved, multiple-valued logic will become the reality solution for high performance arithmetic VLSI system. We may well see a multiple-valued chip used in such a VLSI system.

References

1. Iwai H (1999) CMOS technology-year 2010 and beyond. IEEE Solid-State circuits SC-34, 3, 357/366
2. Geer D (2005) Chip makers turn to multi-core processors. Computer 38(5):11–13
3. Zhirnov VV (2005) Emerging research logic devices. IEEE Circuits Devices Mag 37–46
4. Venkatraman V, Burleson W (2008) An energy-efficient multi-bit quaternary current-mode signaling for on-chip interconnects. In: Proceedings of the custom integrated circuits conference, pp 301–304
5. Matsuura T, Shirahama H (2009) Timing variation aware multiple-valued current-mode circuit for a low-power pipelined system. In: 39th international symposium on multiple-valued logic. ISMVL 2009, pp 60–65
6. Natsui M, Arimitsu T, Hanyu (2010) Low-energy pipelined multiple-valued current-mode circuit with 8-level static current-source control. In: Proceedings of the international symposium on multiple-valued logic, ISMVL 2010, pp 235–240

7. Shirahama H, Mochizuki A, Hanyu T, Nakajima M, Arimoto K (2007) Design of a processing element based on quaternary differential logic for a multi-core simd processor. In: Proceedings of 37th IEEE international symposium on multiple-valued logic, May 2007
8. Jerraya A et al (2005) Multiprocessor systems-on-chips. Computer 38(7):36–40
9. Geer D (2005) Chip makers turn to multicore processors. Computer 38(5):11–13
10. Ike T, Hanyu T, Kameyama M (2002) Fully source-coupled logic based multiple-valued VLSI. In: Proceedings of 32nd IEEE international symposium on multiple-valued logic, 270/275, Boston, Massachusetts, U S A, May 2002

Chapter 57
Uniform Control of Single-Phase Two-Stage Grid-Connected Generation and Active Power Filter

Zhaoyin Yang, Guohai Liu, Zhiling Liao and Zhaoling Chen

Abstract In this paper a single-phase active power filter based on a two-stage photovoltaic grid-connected system is presented, which can not only inject PV power into gird but also can always act as an active power filter to compensate the load harmonics current and reactive power. Control method of a double loop of the inverter DC voltage and current control of grid is used. When local nonlinear load is changed, DC voltage can be kept stable. Simulation results of both conditions that harmonics exist or do not exist show that the designed inverter is stable and practicable. Single-phase harmonic detection method does not need coordinate transformation and the algorithm is simple. The feasibility of control strategy and system topology is verified by the simulation and experimental results.

Keywords Photovoltaic grid-connected · Harmonic detection · Uniform control · Nonlinear load

Z. Yang (✉) · G. Liu · Z. Liao · Z. Chen
College of Electrical and Information Engineering,
Jiangsu University, Zhenjiang 212013, Jiangsu, China
e-mail: yzhaoy@126.com

G. Liu
e-mail: ghliu@ujs.edu.cn

Z. Liao
e-mail: liaozhiling@ujs.edu.cn

Z. Chen
e-mail: czl908@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_57,

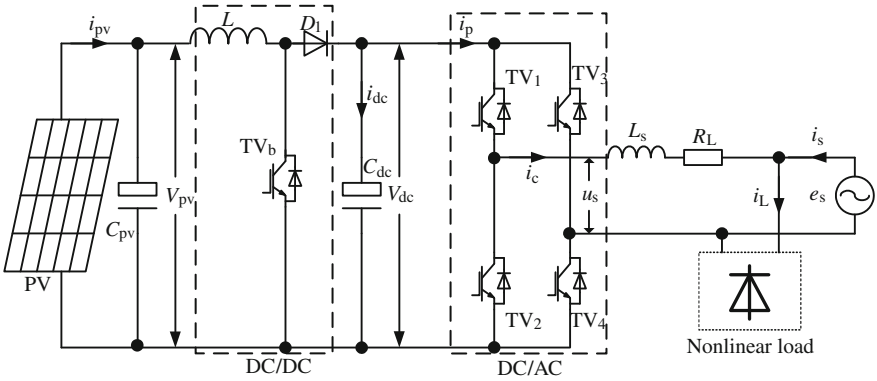


Fig. 57.1 Main circuit topology structure

57.1 Introduction

Solar energy is one of the cleanest and the most promising renewable energy in the world recently. With the extensive use of distributed generation systems, energy generated feedback to the grid can greatly ease the current tense situation in the domestic energy demand [1]. In the ends of grid, especially the remote parts away from the power grid, there is reactive and harmonic current which greatly affects the quality of power supplying other users on a transmission line and impact the voltage stability of utility. In order to solve the aforesaid power quality problems, active power filter (APF) systems have been rapidly expanded in last two decades. Many APF topologies and control algorithm for single-phase system are reported.

In the presented work, the function of APF is added in the existing inverter of PV system by making suitable modification/improvement in the control methodology. Hence, this proposed system does not need any additional hardware/power-circuit for enabling existing inverter to act as an APF also. In this paper, a unified control method is proposed to transmit active power, compensate reactive power and eliminate harmonics.

57.2 The Principle of Uniform Control

57.2.1 System Model of the Inverter

The system being modeled is shown in Fig. 57.1. It consists of a PV array that feeds power into the grid, DC/DC, DC/AC, filter inductance L_s, and local nonlinear load [2].

To obtain the high-quality power flow [3], the output current of the inverter is often controlled to be a sinusoidal current in phase with the grid voltage, however,

the harmonics and the reactive current produced by the local nonlinear load will impact the grid [4]. The PV system will perform dual functions, acting as a photovoltaic generation on the daytime and as an APF all day as long as there is local nonlinear load.

In the PV generation, the main problem is how to improve efficiency and the stability of the whole system. The DC voltage V_{pv} is controlled to track the maximum power point of the PV array [5]. The whole design for DC/AC converter is composed of two control loops with the voltage loop outside the inner current loop, which function is to shape the grid current into a sinusoidal in phase with the grid voltage. The two functions work in two transformations, and realize respectively.

57.2.2 Single-Phase Harmonics Detection Method

Based on extension of the *p–q* theorem, which essence is to separate fundamental current from harmonic current through the linear transformation [6]. In order to detect reactive and harmonic current accurately and quickly, an average current filter is used instead of conventional LPF, which has good real-time character. It is simply realized, and with high-precision detects the harmonic current.

Suppose that single-phase instantaneous grid voltage as follows:

$$e_s = \sqrt{2}U_s \sin \omega t \tag{57.1}$$

Nonlinear load current can be divided into three parts as follows:

$$i_L(t) = i_p(t) + i_q(t) + i_h(t) \tag{57.2}$$

Here i_p(t), i_q(t) and i_h(t) are active power component, reactive power component of fundamental current and harmonic current, respectively.

Using Fourier series, Eq. (57.2) can be written as

$$i_L = I_p \cos \omega t + I_q \sin \omega t + \sum_{n=3}^{\infty} I_n \cos(n\omega t + \varphi_n) \tag{57.3}$$

When Eq. (57.3) is multiplied by 2 cos ωt the result is

$$\begin{aligned} 2i_L \cos \omega t &= I_p + I_p \cos 2\omega t + I_q \sin 2\omega t + \sum_{n=3}^{\infty} 2I_n \cos(n\omega t + \varphi_n) \cos \omega t \\ &= I_p + I_p \cos 2\omega t + I_q \sin 2\omega t + \sum_{n=3}^{\infty} I_n \{\cos[(n+1)\omega t + \varphi_n] \\ &\quad + \cos[(n-1)\omega t + \varphi_n]\} \end{aligned} \tag{57.4}$$

The transformed current involves DC components and AC components. Through integration in a period, the integration of AC components is zero, but DC

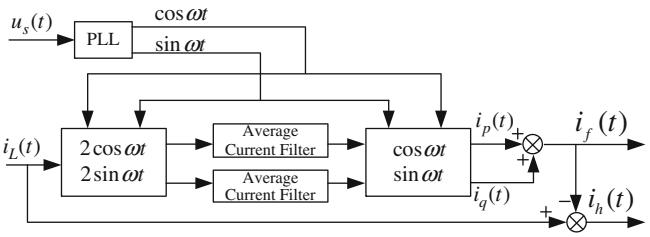


Fig. 57.2 Block diagram of single-phase harmonics detection method

component is not zero. So we can obtain the DC components, which is the current average principle. Theoretically, the DC components I_p can be separated using an average current filter instead of conventional LPF. The instantaneous active power components can be obtained as

$$i_p(t) = I_p \cos \omega t \tag{57.5}$$

As the same, Eq. (57.3) can be multiplied by 2 sin ωt the result is

$$\begin{aligned} 2i_L \sin \omega t &= I_q + I_p \sin 2\omega t - I_q \cos 2\omega t + \sum_{n=3}^{\infty} 2I_n \cos(n\omega t + \varphi_n) \sin \omega t \\ &= I_q + I_p \sin 2\omega t - I_q \cos 2\omega t + \sum_{n=3}^{\infty} I_n \{\sin[(n+1)\omega t + \varphi_n] \\ &\quad - \sin[(n-1)\omega t + \varphi_n]\} \end{aligned} \tag{57.6}$$

The instantaneous reactive power components can be obtained as

$$i_q(t) = I_q \sin \omega t \tag{57.7}$$

Therefore, the harmonic components are derived as

$$i_h(t) = i_L(t) - i_p(t) - i_q(t) \tag{57.8}$$

The block diagram of this algorithm is shown in Fig. 57.2. The reactive power detection channel is switched off when harmonics and reactive power need to be compensated at the same time.

57.2.3 Design of Controller

In this paper, control method of a double loop of the inverter DC voltage and current control of grid is used. Output voltage of PV array is regulated in voltage loop to generate given amplitude of the reference current. The tracking control of

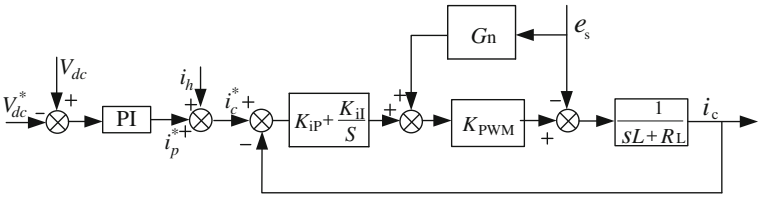


Fig. 57.3 Current and voltage trails principle control block diagram

grid-connected current and compensation current perform in current loop to ensure rapidity and less error in current tracking.

The grid voltage $E(s)$ is the disturbance input for the system, which cannot be ignored in the experiment. In this case, the harmonic current i_h is seen as a disturbance signal for the grid current control loop. Current and voltage trails principle control block diagram is shown in Fig. 57.3.

PI regulator is used in current loop, and at the same time grid voltage feed-forward compensation operated. After modified, the output current transfer function of the grid-connected inverter can be written as

$$I_c(s) = \frac{\frac{1}{sL+R_L}(K_{PWM}G_n - 1)}{1 + (K_iP + \frac{K_{iL}}{s})K_{PWM}\frac{1}{sL+R_L}}E(s) \tag{57.9}$$

As known in the equation above, by selection of $G_n = 1/K_{PWM}$, the impact of grid distortion or disturbance to output current can be eliminated.

57.2.4 Main Circuit Working State Analysis

Over a half period, operation of the inverter can be mainly divided into three modes, and they are described as follows:

Mode 1: In the initial mode, all the active switches are turned off as shown in Fig. 57.4a. Energy stored in the filter inductor starts to discharge and the DC-link capacitor store energy from PV array. Grid current is transferred to the DC-link capacitor through the body diodes to maintain the voltage level. DC-link voltage is set to the peak value of grid. Then operation of the inverter will shift to mode 2.

Mode 2: Switches Q3 and Q4 are turned on as shown in Fig. 57.4b and inverter current nonlinearly ramps up and inductance and capacitor store energy.

Mode 3: Switches Q1 and Q4 are turned on as shown in Fig. 57.4c which forces DC-link voltage to reduce continuously. PV array and DC-link capacitor will supply energy to inverter to transmit the active power and compensation current to the grid. Once the voltage drops below the peak value of the grid, all switches will be turned off again.

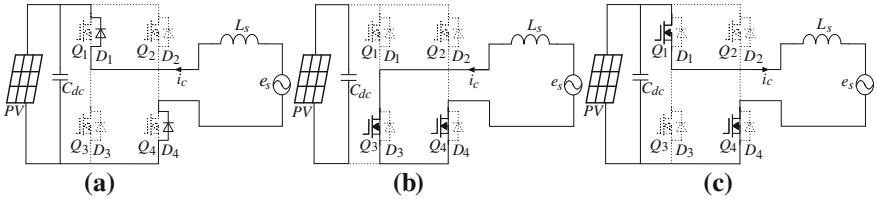


Fig. 57.4 Equivalent circuit of the inverter in mode 1–3

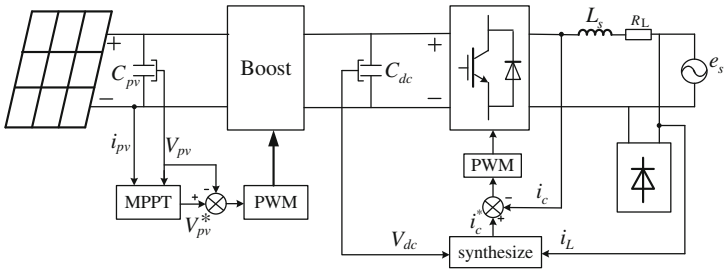


Fig. 57.5 Structure of uniform control

57.3 Uniform Control of Grid-Connected Generation and Active Power Filter

Structure of uniform control is shown in Fig. 57.5. In the daytime, system inject PV power into grid, at the same time, part of the active current is used to maintain DC bus voltage a constant. System can not only inject active power into grid, but also inject reactive power and suppress harmonic. The PV system maintains the voltage level by transmitting the active power to the grid. When there is no irradiance, the system will maintain the voltage level by absorbing power from the grid for acting as an APF.

To ensure the normal operation of the system, the accuracy and real-time of the command current is the key factor. In order to realize PV grid-connected and active power filter, the active power from PV generator can be added to the active power detection channel as shown in Fig. 57.6.

Active DC component of PV can be obtained by active power amplitude gained by PI regulator in Fig. 57.6 multiplying with grid voltage synchronizing signal.

Instruction signal $i_c^*(t)$ of the whole system can be obtained in Fig. 57.6.

$$\begin{aligned} i_c^*(t) &= i_L(t) - i_f(t) = i_L(t) - [i_p(t) + i_q(t) - i_{pv}^*(t)] \\ &= i_L(t) - i_p(t) - i_q(t) + i_{pv}^*(t) = i_h(t) + i_{pv}^*(t) \end{aligned} \tag{57.10}$$

As the results shown above, the preceding section is the harmonic components in p–q theory, and the later section is the active power from PV generation. The

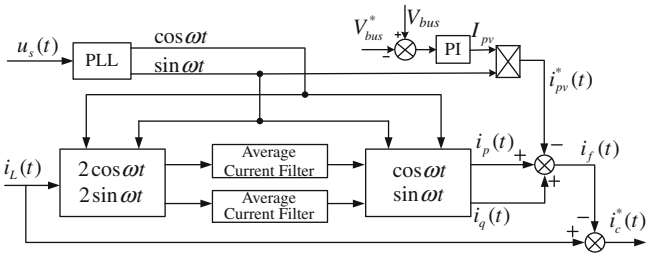


Fig. 57.6 Instruction signal synthesis

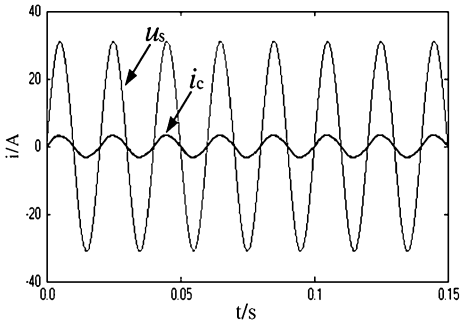


Fig. 57.7 Output current of the inverter and grid voltage

inverter inject the command current calculated into grid can compensate harmonic current and perform grid-connected.

57.4 Simulation and Experiment Results

In order to verify the correctness of the design of the system, a simulation model is built in Matlab/Sinmulink, in which the simulation parameters are as follow: inverter output filter inductance is 10 mH; DC side capacitor is 2,200 μF. Packaged PV array module [7] parameters are as follow: max power is 600 W, current at Pmax is 8.15 A, voltage at Pmax is 73.6 V, short-circuit current is 8.6 A, open-circuit voltage is 90.4 V.

Figure 57.7 shows the simulation waveform of grid voltage us and out current i_c without local nonlinear load, which can be seen that the current at the side of grid-connected are in the same phase with grid voltage.

Figure 57.8 shows the simulation waveform of grid current is and out current i_c with local nonlinear load at t = 0.08 s. With the compensation of harmonic, the grid current is shaped to be a sinusoidal current.

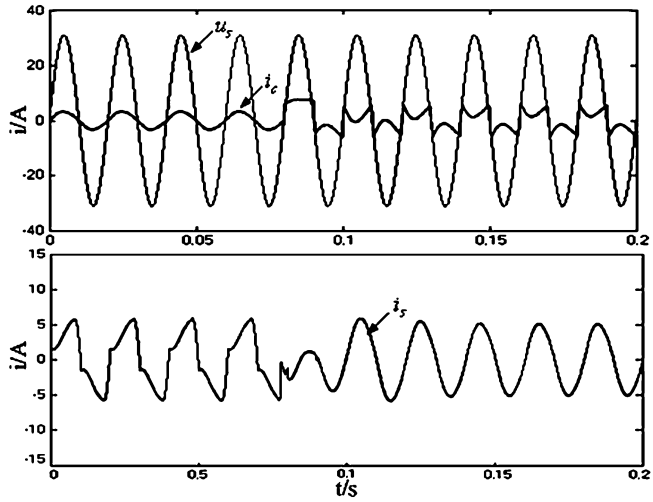


Fig. 57.8 Grid current and output current of inverter

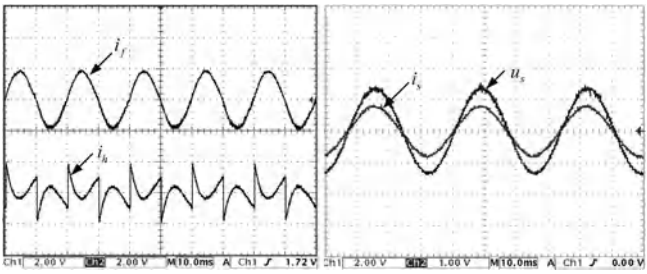


Fig. 57.9 Experimental results

To validate the efficiency of grid-connected PV inverter with the function of compensating the load harmonics and reactive power, a 1,000 W experimental model is built. Figure 57.9a shows experimental waveform of the fundamental components i_f and harmonic current i_h calculated by average current harmonic detection method. Figure 57.9b is the waveform of grid voltage us and grid current i_s with compensation of the load harmonics and reactive power.

The results of the experiment show that the improved single-phase harmonic detection method can detect harmonic and reactive power current accurately and quickly. Command current conducted by instruction signal synthesis control the output of the inverter, which can greatly compensate harmonic and reactive power current in grid current.

57.5 Conclusion

Considering the advantages and defects of photovoltaic and the APF, a two-stage V-source inverter including boost converter and H-bridge converter is proposed, which is simple and flexible to inject PV power into grid. Improved single-phase harmonic detection method does not coordinate transformation and algorithm is simple, which can detect harmonic and reactive current accurately and quickly. Results of simulation and experiment certify the correctness and feasibility of control strategy and system topology.

Acknowledgments A project funded by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (10KJB470003, 11KJB470005), the priority academic program development of Jiangsu higher education institutions (PAPD), and the major project subject to tender of the key research base of the college philosophy social science in Jiangsu province (2010-2-9).

References

1. Dong MI, Luo AN (2006) Design and control strategies of inverters for a grid-connected photovo-ltaic power system. Autom Electric Power Syst 30(20):97–102
2. de Souza KCA, Goncalves OH, Martins DC (2006) Study and optimization of two dc–dc power structures used in a grid-connected photovoltaic system. In: IEEE power electronics specialists conference, PESC06, pp 3247–3251
3. Barbosa PG, Braga HAC, Rodrigues MCB, Teixeira EC (2006) Boost current multilevel inverter and its application on single-phase grid-connected photovoltaic systems. IEEE Trans Power Electron 21:1116–1124
4. Kwon JM, Nam KH, Kwon BH (2006) Photovoltaic power conditioning system with line connection. IEEE Trans Ind Electron 53:1048–1054
5. Femia N, Petrone G, Spagnuolo G, Vitelli M (2005) Optimization of perturb and observe maximum power point tracking method. IEEE Trans Power Electron 20:963–973
6. Akagi H, Watanabe EH, Aredes M (2007) Instantaneous power theory and applications to power conditioning. Wiley, Hoboken
7. Roman E, Alonso R, Ibanez P, Elorduizapatarietxe S, Goitia D (2006) Intelligent PV module for grid-connected PV systems. IEEE Trans Ind Electron 53:1066–1073

Chapter 58
Novel Current Control Method for Active Power Filter

Zhaoling Chen, Mingyang Wei, Guohai Liu and Zhiling Liao

Abstract A novel hysteresis current control strategy based on voltage space vector is proposed for shunt active power filter. The differential relationship between reference voltage and current is then calculated discretely to determine the location region of reference voltage vector. Then the appropriate voltage space vector at each instant is determined by using 4-level hysteresis comparators with narrow hysteresis band working in α–β frame. The proposed method can efficiently improve the utilization of voltage and the current tracking performance, and decrease the switching frequency. Finally, simulation results show that the proposed method is valid and feasible.

Keywords Active power filter · Voltage space vector · Hysteresis current control · Harmonic

58.1 Introduction

In the past few years, the problem of the deterioration of power quality resulting from the proliferation of nonlinear loads in power systems has attracted much attention. Harmonic distortion is known to be source of several problems, such as

Z. Chen (✉) · M. Wei · G. Liu · Z. Liao
College of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, People's Republic of China
e-mail: czl908@126.com

M. Wei
e-mail: wmy06666@163.com

G. Liu
e-mail: ghliu@ujs.edu.cn

Z. Liao
e-mail: liaozhiling@ujs.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_58,

increased power losses, excessive heating in rotating machinery, significant interference with communication circuits, voltage flicker, audible noise, incorrect operation of sensitive loads, and so on [1, 2]. To solve the harmonic pollution, the application of active power filter (APF) has been considered as the best solution for its excellent dynamic performance [3–5].

The circuit structure of the typical shunt active power filter is shown in Fig. 58.1. The 3-phase load current i_L including its harmonic current to be compensated is measured and its harmonic component is denoted as the reference current i_c^*. The APF output current i_c should track the reference current i_c^*. If their difference $\Delta i = i_c^* - i_c$ is larger than a certain tolerance, the current tracking control of the APF will be activated.

This paper proposes a new vector-based hysteresis current controller to reduce the APF switching frequency. The differential relationship between reference voltage and current is then calculated discretely to determine the location region of reference voltage vector. Then, the appropriate voltage space vector at each instant is determined by using 4-level hysteresis comparators with narrow hysteresis band working in α–β frame. The proposed method could efficiently improve the utilization of voltage and the current tracking performance. Finally, simulation results show that the proposed method is valid and feasible.

58.2 Principle of Vector-Based Method

The 3-phase APF equivalent circuit based on voltage resource is typically shown in Fig. 58.2, in this figure, actual switches are replaced by ideal switches.

The system equation in phase co-ordinates is

$$\begin{cases} u_a = Ri_{ca} + L\mathrm{d}i_{ca}/\mathrm{d}t + e_a \\ u_b = Ri_{cb} + L\mathrm{d}i_{cb}/\mathrm{d}t + e_b \\ u_c = Ri_{cc} + L\mathrm{d}i_{cc}/\mathrm{d}t + e_c \end{cases} \tag{58.1}$$

In order to eliminate the interference between phases, the space vector is introduced into which also called static α–β frame of axes. When the α axis and a axis are in the same direction, the relation of the two frame of axes is

$$\begin{bmatrix} f_\alpha \\ f_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} \tag{58.2}$$

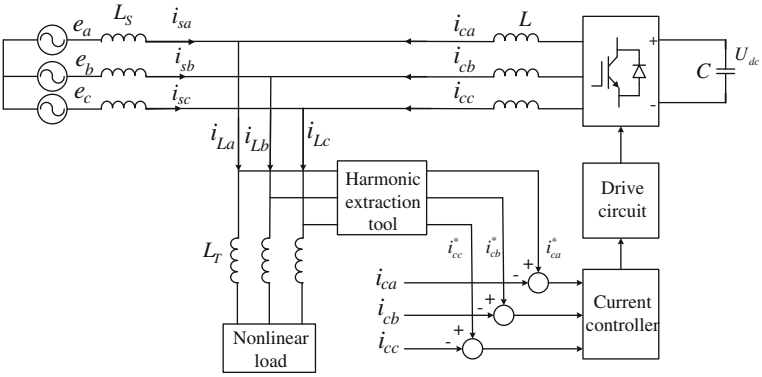


Fig. 58.1 Schematic diagram of SAPF

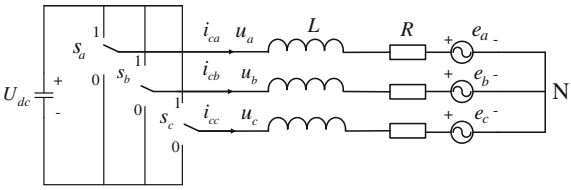


Fig. 58.2 Equivalent circuit diagram of APF

After this transformation, the 3-phase voltage or current in Fig. 58.2 can be expressed by a voltage space vector $\boldsymbol{u}$ or a current space vector $\boldsymbol{i}$. Equation 58.1 can be expressed as a space vector equation

$$\boldsymbol{u} = \mathrm{R}\boldsymbol{i} + \mathrm{L}\frac{\mathrm{d}\boldsymbol{i}}{\mathrm{d}t} + \boldsymbol{e} \tag{58.3}$$

where $\boldsymbol{u}, \boldsymbol{i}, \boldsymbol{e}$ are all space vector. For a certain VSI switch state S = (s_a, s_b, s_c), the corresponding space vector of voltage $\boldsymbol{u}_{abc}$ can be expressed as a complex number on the α–β plane.

$$U_k = \begin{cases} 2U_{dc}e^{j(k-1)\pi/3}/3 & (k = 1 \sim 6) \\ 0 & (k = 0, 7) \end{cases} \tag{58.4}$$

If $\boldsymbol{i}^*$ is the APF reference current space vector, the corresponding reference voltage space vector $\boldsymbol{u}^*$ will be

$$\boldsymbol{u}^* = \mathrm{R}\boldsymbol{i}^* + \mathrm{L}\frac{\mathrm{d}\boldsymbol{i}^*}{\mathrm{d}t} + \boldsymbol{e} \tag{58.5}$$

Defining the APF current tracking error as $\Delta\boldsymbol{i} = \boldsymbol{i}^* - \boldsymbol{i}$ and assuming R = 0, from Eqs. 58.3 and 58.5 we have

$$L\frac{d\Delta i}{dt} = \boldsymbol{u}^* - U_k \quad (k = 0, 1, \ldots, 7) \tag{58.6}$$

From the Eq. 58.6, we can infer that the rate of change of the current error vector $\Delta \boldsymbol{i}$ is decided by the deviation between reference voltage vector $\boldsymbol{u}^*$ and output voltage vector of APF. If we want to control the current error vector $\Delta \boldsymbol{i}$, the only things we should do is to control the change rate of the current error vector $d\Delta \boldsymbol{i}/dt$. In order to do that, we should select a suitable output voltage vector of APF $U_k(k = 0, 1,\ldots, 7)$, for the reference voltage vector $\boldsymbol{u}^*$ is a specified value.

58.3 Proposed Vector-Based Method

58.3.1 Principle of Current Control

The principle of current control of this method is shown in Fig. 58.3, compared instruction current i_c^* with feedback current i_c, error current Δi_a, Δi_b, Δi_c are got. After abc/$\alpha\beta$ transformation, Δi_a, Δi_b, Δi_c are projected on α–β coordinates, Δi_α, Δi_β are got, and then through two group of 4-level comparators, they are transformed into the corresponding comparative value S_α, S_β. By means of voltage space vector select logic circuit and judging which region the reference voltage vector $\boldsymbol{u}^*$ is in, the system gets a suitable voltage space vector U_k.

58.3.2 Calculation of $\mathbf{u}^*$

The output voltage vectors of APF $U_k(k = 0, 1,\ldots, 7)$, which are corresponding to the eight switch modes, divide the vector space into six triangle regions as I ~ VI. The reference voltage vector $\boldsymbol{u}^*$ is always in one of the 6 triangle regions, as shown in Fig. 58.4.

The reference voltage $\boldsymbol{u}^*$ can be theoretical calculated through the di^*/dt, however, due to the rapid change of the harmonic current, calculation of di^*/dt as so is very difficult. For this reason, this paper uses an approximation method of di^*/dt in order to quickly get reference voltage $\boldsymbol{u}^*$. In the sampling time point k, the three phase currents of a, b, and c are calculated discretely.

$$\frac{di_x^*}{dt} \approx \frac{i_x^*(k) - i_x^*(k-1)}{T} \tag{58.7}$$

where x represents a, b, and c, respectively. In the current sampling time point k, $i_x^*(k)$ are replaced by the reference compensation current signals i_a^*, i_b^*, i_c^*, $i_x^*(k-1)$ are replaced by the compensation currents i_a, i_b, i_c generated by APF at the same time. If system impedances are ignored, the output reference voltages of APF are got as follows:

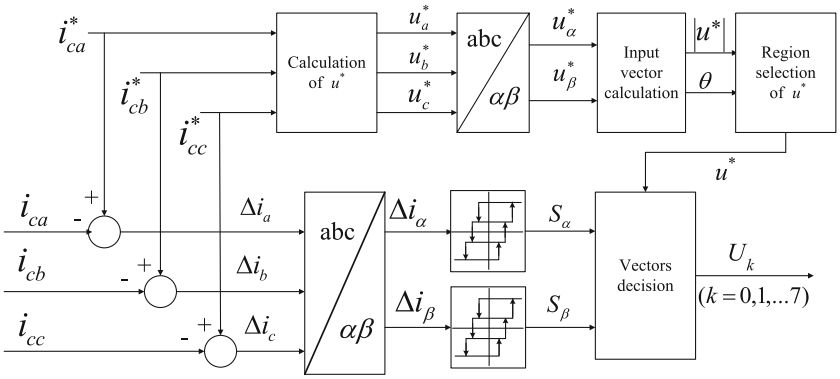
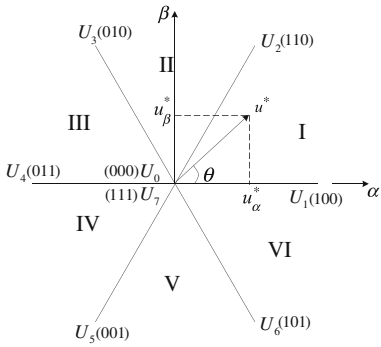


Fig. 58.3 Principle of current control

Fig. 58.4 Region division of $\boldsymbol{u}^*$



$$\begin{cases} u_a^*(k) = e_a + L\frac{i_a^*(k) - i_a(k)}{T} \\ u_b^*(k) = e_b + L\frac{i_b^*(k) - i_b(k)}{T} \\ u_c^*(k) = e_c + L\frac{i_c^*(k) - i_c(k)}{T} \end{cases} \tag{58.8}$$

After coordinate transformation, $u_a^*(k)$, $u_b^*(k)$, $u_c^*(k)$ are projected to α–β frame and transformed into components $\boldsymbol{u}_\alpha^*$, $\boldsymbol{u}_\beta^*$.

$$\begin{bmatrix} u_\alpha^* \\ u_\beta^* \end{bmatrix} = 2/3 \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix} \begin{bmatrix} u_a^*(k) \\ u_b^*(k) \\ u_c^*(k) \end{bmatrix} \tag{58.9}$$

So the region of reference voltage vector $\boldsymbol{u}^*$ can be determined according to the values $\boldsymbol{u}_\alpha^*$, $\boldsymbol{u}_\beta^*$.

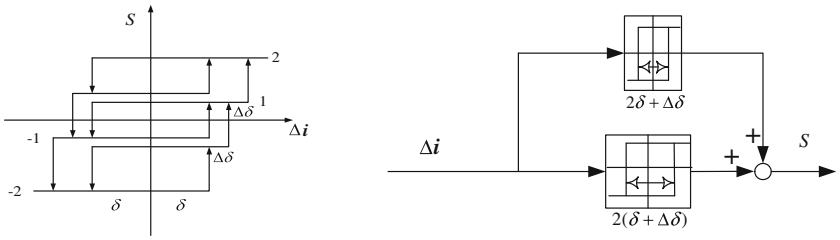


Fig. 58.5 Practical implementation of 4-level α–β hysteresis comparator

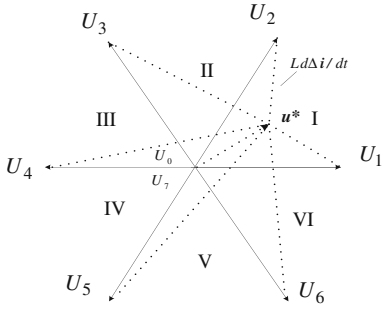


Fig. 58.6 Distribution of Ld$\Delta \boldsymbol{i}$/dt,$\boldsymbol{u}^*$

58.3.3 Optimal Voltage Space Vector Choice for APF

First of all, the error current Δi_a, Δi_b, Δi_c are transformed to the component of $\Delta \boldsymbol{i}_\alpha$, $\Delta \boldsymbol{i}_\beta$ on the α–β coordinates.Then, the error current vectors $\Delta \boldsymbol{i}_\alpha$, $\Delta \boldsymbol{i}_\beta$ pass through two groups of 4-level comparator and change into the corresponding output value S_α, S_β, whose state is one of -2, -1, 1, 2 as shown in Fig. 58.5.

In order to describe the operation of the proposed current controller scheme, suppose that reference voltage vector $\boldsymbol{u}^*$ is located in Region I, as shown in Fig. 58.6; thus whose adjacent voltage vectors U_6, U_1, U_2 and U_0 are selected to be applied. Under this condition, if $\Delta \boldsymbol{i}_\alpha$ and $\Delta \boldsymbol{i}_\beta$ touches the highest and lowest hysteresis bands respectively, thus $S_\alpha = 2$ and $S_\beta = -2$. This implies that the voltage space vector U_1 must be applied to make the corresponding d$\Delta \boldsymbol{i}$/dt have the maximum value in opposite direction of the current error vector to increase the α-component and simultaneously decrease the β-component of the current vector. Similarly, if $S_\alpha = -1$, $S_\beta = -1$, U_0 or U_7 must be applied to make the corresponding d$\Delta \boldsymbol{i}$/dt have the minimum value in opposite direction of the current error vector to decrease α-component and β-component of the current vector and simultaneously reduce switching frequency, etc. Using similar control strategy for other regions, the switching table of U_k given in Table 58.1 is extracted.

Table 58.1 Switching table of U_k

$s_\alpha s_\beta$ / u^*	-2, -2	-2, -1	-2, 1	-2 2	-1, -2	-1, -1	-1, 1	-1, 2	1, -2	1, -1	1, 1	1, 2	2, -2	2, -1	2, 1	2, 2
I	U_5	U_5	U_3	U_3	U_5	U_0,U_7	U_6	U_3	U_1	U_4	U_6	U_2	U_1	U_1	U_1	U_2
II	U_5	U_5	U_3	U_3	U_5	U_0,U_7	U_2	U_3	U_6	U_0,U_7	U_6	U_2	U_6	U_1	U_2	U_2
III	U_5	U_6	U_4	U_3	U_6	U_3	U_2	U_3	U_6	U_0,U_7	U_2	U_2	U_6	U_6	U_1	U_2
IV	U_5	U_5	U_4	U_3	U_5	U_1	U_3	U_4	U_1	U_1	U_0,U_7	U_2	U_1	U_1	U_1	U_2
V	U_5	U_5	U_4	U_3	U_5	U_1	U_0,U_7	U_3	U_1	U_5	U_0,U_7	U_2	U_1	U_1	U_2	U_2
VI	U_5	U_5	U_4	U_3	U_5	U_5	U_0,U_7	U_4	U_6	U_5	U_4	U_3	U_6	U_6	U_2	U_2

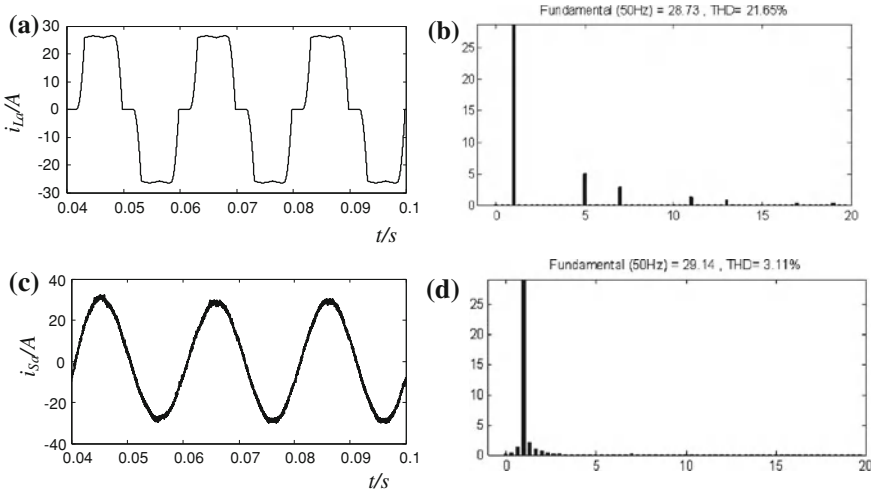


Fig. 58.7 **a** Waveform of load current i_{La}. **b** Spectrum and THD of load current i_{La}. **c** Waveform of current i_{Sa} after compensation. **d** Spectrum and THD of current i_{Sa} after compensation

58.4 Simulation Results and Analysis

The power system module (simpower systems) in MATLAB/Simulink environment is used to test the validity and feasibility of the proposed control scheme in this paper. Main circuit structure as shown in Fig. 58.1, and the harmonic current detection is i_p–i_q method based on instantaneous reactive power theory. Simulation parameters as follows: the voltage of source is 220 V/50 Hz, and the power system impedance was ignored. Harmonic source is three-phase bridge uncontrollable rectifier circuit with RL load, $L_d = 44$mH, $R_d = 10.8\ \Omega$, DC voltage $U_{dc} = 800$ V, inductance $L_T = 1.8$ mH, $L = 3$ mH.

The results of the voltage space vector control method are shown in Fig. 58.7, the Fig. 58.7a is the load current curve, the Fig. 58.7c is source current curve after compensat. The Fig. 58.7b and d show that the current total harmonic distortion

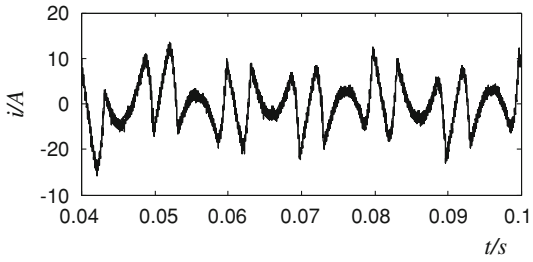


Fig. 58.8 i_{ca}^{*} and i_{ca}

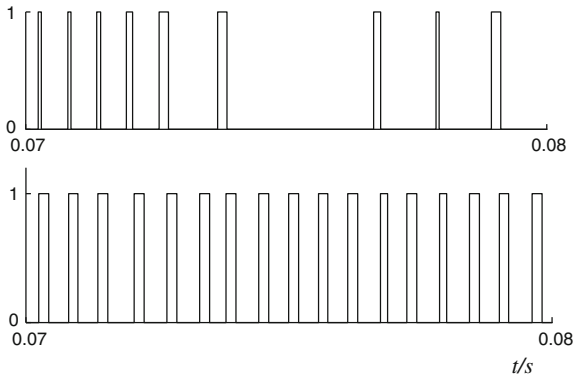


Fig. 58.9 The switching number of two methods

value (THD) from the original 21.65% is reduced to 3.11%, the Fig. 58.8 is the waveform of the compensation current i_{ca} and the reference current i_{ca}^{*}, it is observed that the compensation current i_{ca} can instantaneously track the reference current i_{ca}^{*}. The Fig. 58.9 shows the comparison result between proposed method's switching number (top) and hysteresis control method's switching number (bottom), it is obviously that proposed method's switching frequency and switching loss is lower than hysteresis control method's.

58.5 Conclusion

An improved vector-based hysteresis current controller is proposed in this paper. According to the result of u*, at each instant, two 4-level comparators working in α-β stationary frame determine the suitable voltage space vector. The proposed method could efficiently improve the utilization of voltage and the current tracking performance, and decreases the switching frequency. Finally, simulation results show that the proposed method is valid and feasible.

Acknowledgments A project funded by the Natural Science Foundation of the Jiangsu Higher Education institutions of China (11KJB470005), the priority academic program development of Jiangsu higher education institutions (PAPD).

References

1. Singh B, Al-Haddad K, Chandra A (1999) Review of active filters for power quality improvenment. IEEE Trans Ind Electron 46(5):960–971
2. El-Habrouk M, Darwish MK, Mehta P (2000) Active power filters—A review. Proc IEE-Elect Power Appl 147(5):403–413
3. Peng FZ (2001) Harmonic sources and filtering approaches. IEEE Ind Appl Mag 4(7):18–25
4. Jou HL, Wu JC, Chang YJ et al (2005) A novel active power filter for harmonic suppression. IEEE Trans Power Deliv 20(2):1507–1513
5. Mohseni M, Islam S (2009) A novel current controller for three-phase voltage-source inverters[C]. In: 35th annual conference of IEEE. Porto, Portugal 11:76–81

Chapter 59
Investigation of Busbar-Structure for High Power Converter

Yifeng Zhu

Abstract In high power converter design, low-inductance busbar connecting DC capacitors and power devices is main concern to improve the quality of the whole power electronics system. This paper analyzes and designs a busbar structure in detail for prototype of the high-speed railway traction converter. The method of partial element equivalent circuit and the Q3D software were used to extract the stray inductance. A hollow-shaped structure is proposed with lower inductance than the initial U-shaped structure. The simulation results are given to illustrate the effectiveness.

Keywords Stray inductance · Busbar-structure · Converter · Hollow-shape

59.1 Introduction

In high power converters, the stray inductance existing in the commutation loop should be as low as possible considering safe and high-efficiency operation of switch devices and converters. In fact, the stray inductance distributed in the DC bus usually accounts for the main part of the whole loop inductance [1]. The oversize stray inductance distributing on the busbar, which links DC capacitor

Y. Zhu (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zyfny@hpu.edu.cn

Y. Zhu
Key Laboratory of Power Electronics and Electric Drive,
Institute of Electrical Engineering Chinese Academy of Science,
Beijing 100190, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_59,

bank to converter power devices, may bring excessive transient voltage overshoots, which can result in high voltage stress, increased heating/power losses and exceeding of the device safe operating area [2]. Therefore, it is important to make out a low-inductance busbar structure in early design stage.

There are many literatures on the study of DC busbar to decrease the stray inductance in different aspects. Skibinski [3] compared various busbar shapes in terms of internal and external impedance to show the superior performance of a parallel plate structure. Dimino [4] discusses the application of a laminar busbar structure. A laminated-structure bus bar provides as low effective inductance as possible for a converter. The structure is to laminate a thin piece of dielectric material between the DC positive and negative copper plate [5]. This construction allows for the largest mutual inductance mitigation directly along the power distribution path. The closer these plates are laminated together, the more the mutual cancellation that will be realized. Reference [6] proposed to achieve the optimal busbar design by genetic algorithms.

In this paper, we will focus our study on the optimal design of planar busbar used in high power converter. A physical structure for a prototype of high-speed railway traction converter is constructed with lower inductance.

59.2 The Distribution and Extraction of Stray Inductance

The schematic diagram of the high-speed railway traction converter is shown in Fig. 59.1. The converter including four parts: two single-phase two-level PWM rectifiers in parallel connection, LC filter circuit to cancel the 100 Hz ripple wave in DC-bus voltage, DC-link capacitors and three-phase two-level PWM inverter. Figure 59.1 also gives the distribution of the parasitic inductance in the converter considering the interconnecting busbar, DC-link capacitor tank and IGBT modules as well, where L_{pi} and L_{ni} (i = 1,2,…,7) are stray inductance of positive bus elements and negative bus elements. L_{md} and L_{sc} are stray inductance of a IGBT module and a interconnecting screw, respectively. Among these parasitic parameters the inductance of the busbar is most important, because different from other stray parameters such as L_{md} and L_{sc} the parameter L_n and L_p are changeable.

A physical structure for the prototype converter with symmetrical planar busbar is constructed by proper electrical, thermal and mechanical design. The converter takes forced air cooling. The structure dimension of the busbar like 'U'-shape are 602 × 467 × 2 mm (length × width × thick-ness) with some grooves and holes, which can be seen from Fig. 59.2. The material of the busbar is copper and the dielectric plate with 2 mm thickness takes epoxy glass as material, which has a relative permittivity of approximately 4. Another advantage of epoxy glass is its high dielectric strength of approximately $E_{max} = 20$ kV/mm.

The model of the busbar is difficult to construct due to the complex structure with holes, notches and cross section discontinuity. Fortunately, the influence of the holes, notches and cross section discontinuity in large size conductor can be

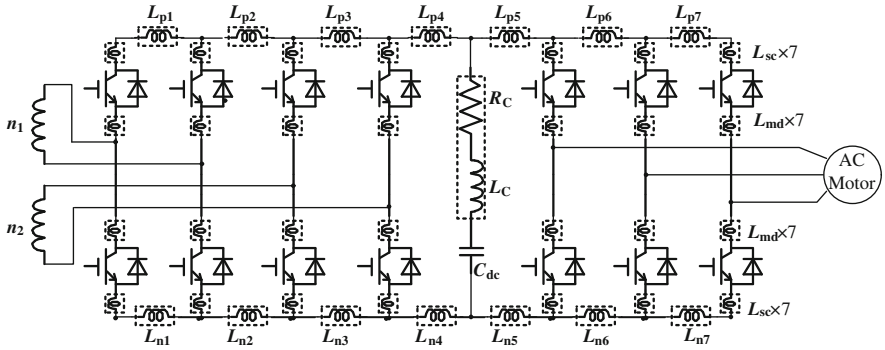


Fig. 59.1 The circuit of the traction converter and the distribution of stray inductance

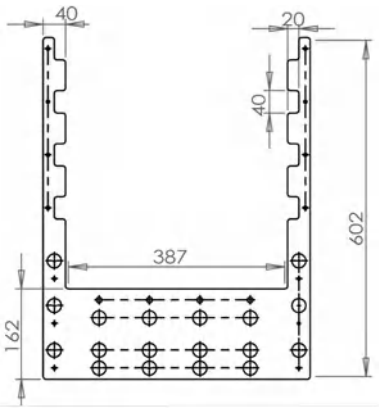


Fig. 59.2 The structure dimensions of the initial busbar used in the converter

neglected when the size of them is not very large [6]. The method of partial element equivalent circuit and the Q3D software were used to extract the stray parameters. In the software the stray inductance are the compositive results of self inductance and mutual inductance of busbar element, which are right emphasis considered to suppress the over-voltage within the power devices turn-off. The inductance of this busbar consists of an internal inductance L_i resulted from flux linkages within busbar plates and an external inductance L_e caused by the flux flowing through the space area surrounding the busbar plates and determined by the geometry of the two current carrying conductors.

At the same time, finite element software is used to simulate the current distribution in the busbar and extract the stray inductance. The maximum current density of the busbar occurs near the spot connecting to switching devices, which is approximately equal to 29.6 kA/m^2. The stray inductance extracted from the busbar is shown as Table 59.1.

Table 59.1 The stray inductance of the initial U-shape busbar

Self inductance of positive busbar (nH)	Mutual inductance (nH)	Self inductance of negative busbar (nH)
873.1	−746.6	890.2
$L_{eip} = 873.1-746.6 = 126.5$ nH		$L_{ein} = 890.2-746.6 = 143.6$ nH

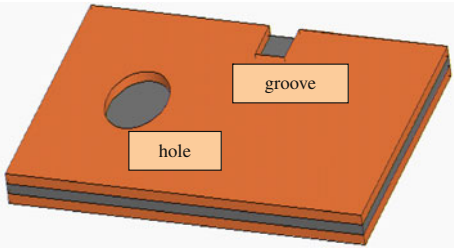


Fig. 59.3 The diagram of a busbar with a hole and a groove

59.3 The Influence of Holes and Grooves to Stray Inductance

In order to optimize the structure of the power busbar to minimize the stray inductance the influence of different structure factors such as width, length, distance, hole size, groove size could be studied firstly. Reference [1] gives mathematical formula to computer the stray inductance of regular plate busbar. However, these formulas don't adapt to the busbar with complex structure, especially with holes or grooves as shown in Fig. 59.3. The dimensions of the busbar in Fig. 59.3 are 60 × 120 × 2 mm.

The influencing laws of different structure parameters of a busbar to stray inductance are simulated. The simulated results show that: (1) the length of the copper plate is proportional to the inductance; (2) the wider the copper plate, stray inductance lower; (3) the closer the copper plates, stray inductance lower. Besides, the influencing laws of holes and grooves to stray inductance are shown in Fig. 59.4 and 59.5. Figure 59.4 implies that the bigger the radius of holes, the higher the stray inductance, because the holes in the copper plate decrease the mutual inductance. However, the inductance increases little when the hole-size is much smaller than the size of the busbar. Figure 59.5 shows that the bigger the depth of grooves, the higher the stray inductance, because the grooves in the copper plate decrease the mutual inductance.

Some optimization methods may be obtained on the basis of the above analysis. The busbar plate should be as wide and short as possible, and the copper plates should be very close to each other. Moreover, the holes and grooves should be designed small to minimize the stray inductance of a busbar. Again, new shape of a busbar may bring new current circuits, which may result in smaller stray inductance. As a conclusion, the voltage overshoots on the IGBTs resulted from the stray inductance will be largely decreased and the converter may operate more safely.

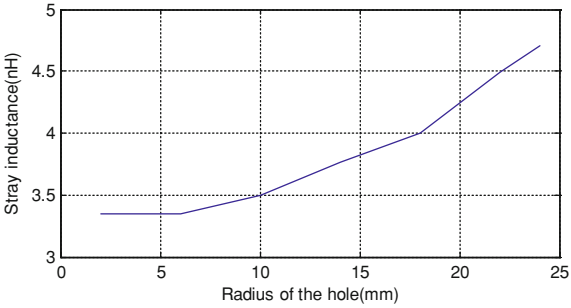


Fig. 59.4 The influence of holes on the stray inductance

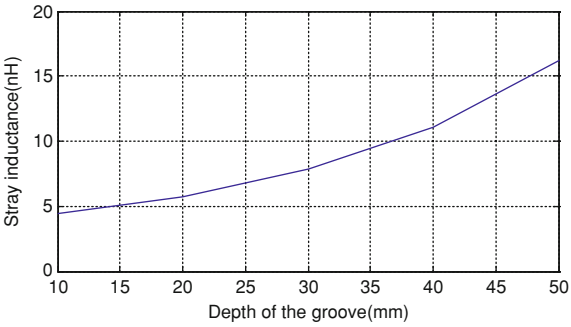


Fig. 59.5 The influence of grooves on the stray inductance

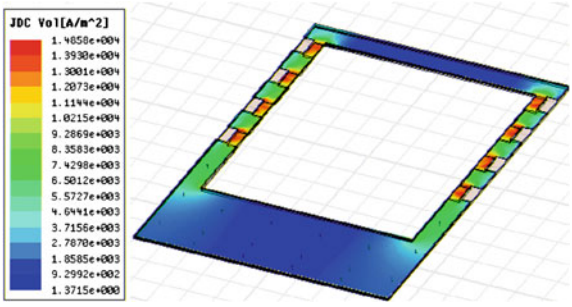


Fig. 59.6 The proposed hollow-shape busbar and the simulated current density

59.4 Optimization and Simulation

The initial U-shape structure of the busbar for the prototype of high-speed traction converter is like a large-size groove made in the busbar. According to the foregoing analysis, the large-size groove may increase the stray inductance much. Therefore, a hollow-shape structure for the busbar is proposed to decrease the parasitic parameter, which cancels the large-size groove. The novel structure is to add a small-size plate in the top of the busbar with the dimensions 467 × 40 × 2 mm as given in Fig. 59.6.

Figure 59.6 shows the simulated result of the hollow-shape busbar. The maximum current density exits near the position connecting to IGBTs and is

Table 59.2 The comparison of simulated stray inductance of two kinds of busbar

Regular	U shape	Hollow shape
Stray inductance (nH)	144	34
Copper area (cm^2)	1109	1295
Ratio of stry inductance versus copper area (nH/cm^2)	1298	263

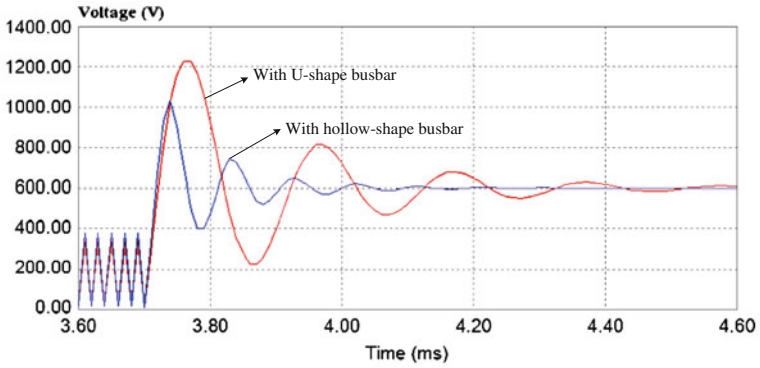


Fig. 59.7 IGBT V_{ce} waveforms with U-shape busbar and hollow-shape busbar respectively

approximately equal to 14.9 kA/m^2, which is much smaller than 29.6 kA/m^2 of the initial U-shape structure. Lower current density may produce lower thermal temperature, which will improve the reliability of the IGBTs. Furthermore, the stray inductance of the new structure busbar is also simulated as shown in Table 59.2.

It can be drawn from Table 59.2 that the stay inductance of the hollow-shape buabar is equal to 34 nH, which is lower than 144 nH of the U-shape busbar. At the same time the ratio of stray inductance vs. copper area can be taken into account. The ratio of the hollow-shape busbar is much lower than that of the U-shape busbar.

Figure 59.7 shows the voltage stress on IGBT module during the switching process with U-shape busbar connection and proposed hollow-shape busbar connection respectively. The voltage spike is higher than 1,200 V when U-shape busbar is adopted, while it reduces sharply when hollow-shape busbar is utilized. The reason is that a higher stray inductance leads to a lower oscillation frequency, which amplifies the voltage spike and oscillation time. Lower transient overshoots are helpful to guarantee devices in the safe operating area.

59.5 Conclusions

The paper analyzes the stray inductance of the busbar and presents a method to optimize the busbar structure for high power converter interconnection to improve the quality of the whole power electronics system. A hollow-shape busbar is better

than a 'U'-shape busbar considering low-stray inductance and over-voltages minimization. The analysis method and conclusion could be a great help for busbar designer to develop power converter interconnection.

References

1. Caponet C, Profumo F, De Doncker RW, Tenconi A (2002) Low stray inductance bus bar design and construction for good EMC performance in power electronic circuit. IEEE Trans on Power Electron 17(2):225–231
2. Zhiqiang W, Guozhu C (2009) Study on planar busbar regarding stray inductance minimization and oscillation suppression for high power converter, sustainable power generation and supply. SUPERGEN '09. International conference, pp 1–7
3. Wen H, Wen X, Liu J, Guo X, Zhao F (2007) A low-inductance high-frequency film capacitor for electric vehicles. In: International conference on electrical machines and systems. ICEMS, pp 2046–2050
4. Pasterczyk RJ, Martin C, Guichon J-M, Schanen J-L (2005) Planar busbar optimization regarding current sharing and stray inductance minimization, 2005 European conference on power electronics and applications, 9 pp
5. Beukes HJ, Enslin JHR, Spee R (1997) Busbar design considerations for high power IGBT converters. Power electronics specialists conference. PESC '97, Record., 28th Annual IEEE, vol 2, pp 847–853
6. Li F, Sun X, Huang L, Jiang J (2009) Main loop equivalent circuit of H-bridge power unit for EMI analysis. In: 4th IEEE conference on industrial electronics and applications. ICIEA, pp 2945–2948

Chapter 60
Digital Monitoring and Control System Based on Ethernet for Twin-Arc High Speed Submerged Arc Welding

Qi Li, Xue Jun Li, Kuan Fang He, Ke Wang and Zong Qun Deng

Abstract The Ethernet communications was applied to dynamic monitoring and control of high-speed twin-arc submerged arc welding process, with its characteristics of high transmission speed, large capacity and reliability, a digital monitoring and control system based on Ethernet was designed. The sequential control of twin arc-starting and -ending was achieved, and the formation of weld seam was improved with the guarantee of stability and order of welding process, which achieved the technology of twin-arc high speed arc welding.

Keywords Twin-arc high speed arc welding · Ethernet · Digital monitoring

60.1 Introduction

Submerged arc welding is a high efficiency welding method that the arc is burning under the flux layer, which is one of the most widely applied technologies in welding manufacture. With characteristics of high efficiency, deep penetration and

Q. Li (✉) · X. J. Li · K. F. He (✉) · K. Wang · Z. Q. Deng
Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China
e-mail: richie.lsq@gmail.com

K. F. He (✉)
e-mail: hkf791113@163.com

X. J. Li
e-mail: hnkjdxlxj@163.com

K. Wang
e-mail: wangkekeer@126.com

Z. Q. Deng
e-mail: dengzongqun@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_60,

high degree of mechanization, it is particularly applicable to the welding of heavy plate and long weld seam. In recent years, twin-arc and multi-arc submerged arc welding has become the research hotspot both at home and abroad. In dynamic monitoring of twin-arc submerged arc welding, feedback of arc signals were often used to conduct the monitoring and control of the welding process [1]. Because of the complexity of welding process, arc signals were seriously influenced by the electromagnetic interference, how to collect arc signals precisely, quickly and steadily is the key technology of research [2]. Currently, in the study of monitoring system of submerged arc welding, system communication were mostly realized by field bus, RS-485, RS-232 or PCI interface, limited by the speed of data transmission and environment, real-time capability and reliability of the system can not be guaranteed [3, 4]. In recent years, with the development of embedded Ethernet technology, which has characteristics of high transfer rate, anti-interference ability, large capacity, simple structure and low cost, has been received a considerable application in the field of data acquisition and real-time monitoring [5]. A digital monitoring and control system based on Ethernet for twin-arc high speed submerged arc welding was built in this paper, which provided a hardware and software platform of high real-time capability, high reliability and low cost for the further analysis of arc signals and welding quality monitoring and control.

60.2 System Hardware Structure

The digital monitoring system mainly consist of computer, data acquisition and monitoring module based on TCP/IP, control box of twin-arc submerged arc welding and sensing devices. The disposition of each part in the twin-arc high speed submerged arc welding system is shown in Fig. 60.1. The current sensor CS1 and CS2 and voltage sensor VS1 and VS2 are used to detect current and voltage of the master and slave arc. The data acquisition and monitoring module based on TCP/IP is core of the system, which is responsible for sensors signal processing, communication with host computer and execution of control instructions. The control box of twin-arc submerged arc welding is used to control the master and slave power source, wire feeding device and welding vehicle according the instructions from monitoring module. The host computer is used to set welding parameters and control instructions, and to display the current and voltage collected from the master and slave arc.

60.2.1 Data Acquisition and Monitoring Module Based on TCP/IP

The data acquisition and monitoring module based on TCP/IP consists of ARM controller, A/D, D/A, Ethernet interface and digital IO, its structure is shown in Fig. 60.2. Ethernet was used to realize intercommunication, compared with data

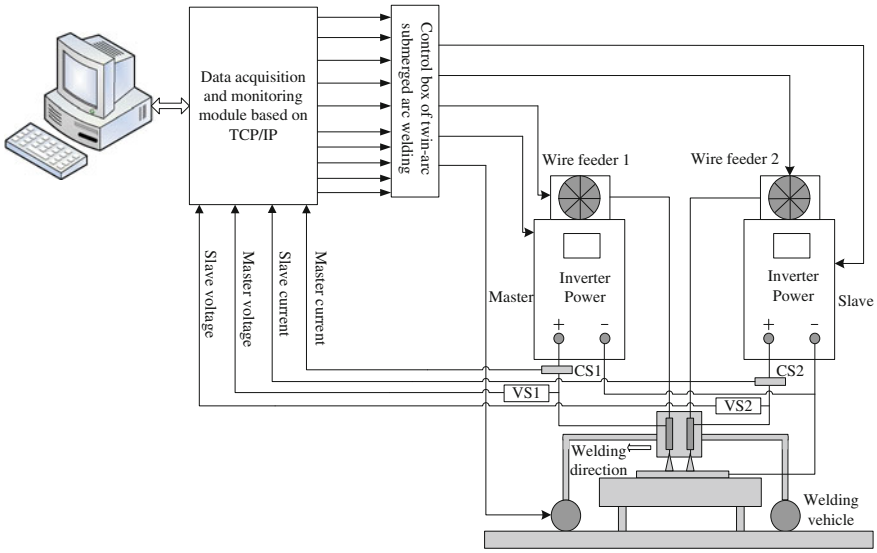


Fig. 60.1 Twin arcs high speed submerged arc welding system

transmission methods of USB, RS-232 and PCI, the transmission mode based on Ethernet TCP/IP protocol have characteristics of high transfer rate, large transmission capacity, good anti-interference ability, long transmission distance and easy installation. The welding process is complicated and changeable, and the arc signals were seriously influenced by the electromagnetic interference, as a result the monitoring system is required to have good real-time capability and reliability. The Ethernet communications was applied to the monitoring and control of high-speed submerged arc welding, which can effectively overcome these difficulties, improved the real-time capability and reliability of the system and ensured the reliable and stable of welding process.

60.2.2 Interface of Each Module

Interface circuits mainly consist of Ethernet interface, analog input and switching value output circuits. The Ethernet interface connects the host computer and monitoring module, which ensured the rapid and steady communications of the system. According to the functions and requirements of twin-arc submerged arc welding process monitoring, in combination with the data acquisition and monitoring module based on TCP/IP, the Ethernet interface, welding current and voltage sampling circuits were designed in this paper.

Ethernet interface. The Ethernet interface is composed of RF45 connector and Ethernet controller DM9000AEP. The DM9000AEP is a fully integrated and

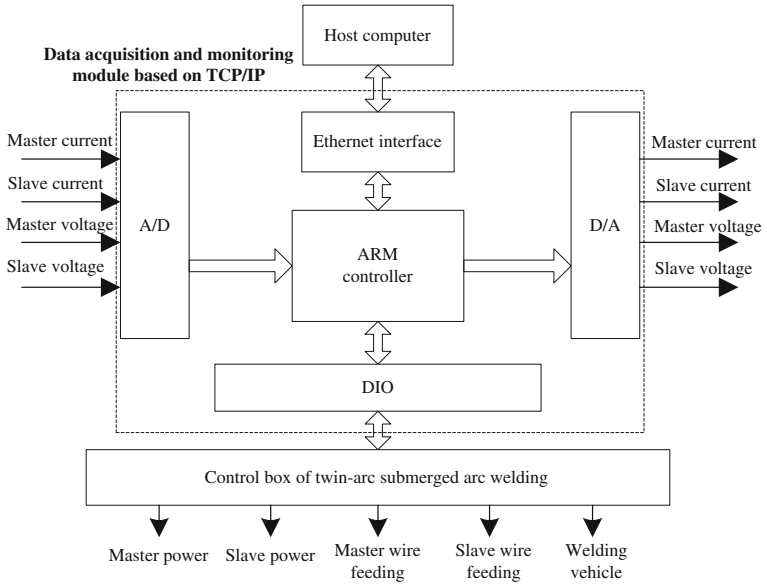


Fig. 60.2 Data acquisition and monitoring module based on TCP/IP

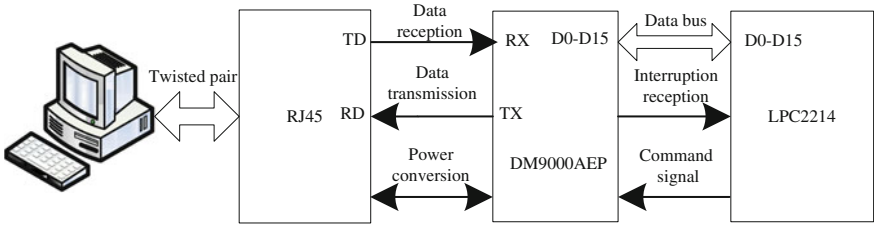


Fig. 60.3 Hardware schematic diagram of the Ethernet interface

cost-effective low pin count single chip Fast Ethernet controller with a general processor interface. It conform to the Ethernet standard, supports IEEE802.3x flow control for full-duplex mode and semi-duplex CSMA/CD mode; Integrated 10/100 M transceiver with AUTO-MDIX; Integrated 16 K Byte SRAM; Supports 8/16Bit data bus, interrupt request and I/O base address selection and the operation of the internal register is simple. The Ethernet interface uses a RJ45 connector integrated with network transformer, and is connected to the host computer by twisted pair wiring. Hardware schematic diagram of the Ethernet interface shows in Fig. 60.3.

Sampling circuit of welding current. The current and voltage changed greatly during the welding process, which placed higher demands on reliability and stability of the feedback circuits. Hall current sensor was used in welding current

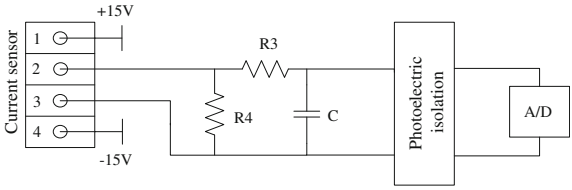


Fig. 60.4 The current sampling circuit

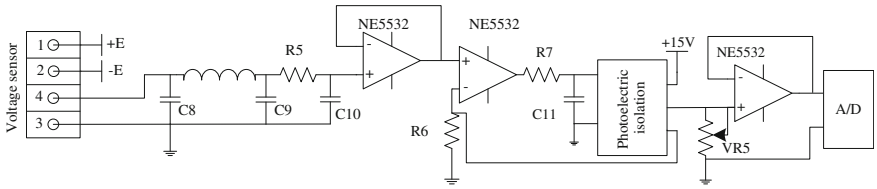


Fig. 60.5 The welding voltage sampling circuit

sampling, which has a range of 1500A, ±12 V power supply is required, and can be used to measure AC and DC current. Technologies of precision constant current, linear temperature compensation, high-speed digital acquisition and rapid processing of synchronous data were adopted in the sensor, effective values were transforming into standard output signals, the sensor have characteristics of strong overload capability, high port isolation, metal shell and strong anti-interference performance. The current sampling circuit is shown in Fig. 60.4.

Sampling circuit of welding voltage. Because the main circuits working under the heavy current, but monitoring system working under weak current, signal insulation were employed to cut off the electrical contact but preserve the control signals in the feedback of welding voltage. Electrical isolation circuits with linear opt coupler TLP521-2 at the core were used in the welding voltage sampling, which is shown in Fig. 60.5, the collected signals were filtered by LC circuit and isolated by TLP521-2, feedback signals were obtained and inputted into the A/D module.

60.3 Software Composition of System

The running of high-speed twin-arc submerged arc welding devices is complicated and tedious, reasonable software flow and structure is crucial for the effective control of welding process. Software of the monitoring system is designed in modularization, which contains communication procedure, welding process monitoring procedure and welding sequential control procedure.

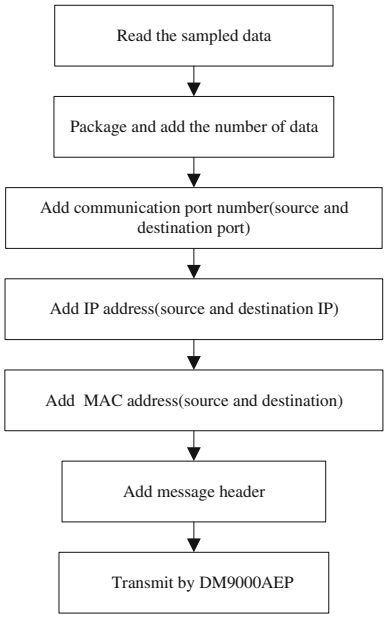


Fig. 60.6 Communication flow chart

60.3.1 Communication Procedure

The high speed communication based on Ethernet is the priority to ensure the sampled data can be transmitted to host computer rapidly, Ethernet controller DM9000AEP is used to realize data communication between the monitoring module and the host computer.

Because the data packets will be received by host computer, the collected welding current and voltage signals should be packed following the TCP/IP protocol standard, and then transmitted to the host computer by DM9000AEP. According to the requirements of digital monitoring system for high speed submerged arc welding, communication program based on TCP/IP protocol was designed in this paper. The flow chart is shown in Fig. 60.6.

60.3.2 Monitoring Procedure of Welding Process

The welding process monitoring procedure is used to complete the acquisition and real-time display of welding parameters, the A/D transformation subprogram is called to convert the collected current and voltage signals into digital, then packed by the Ethernet controller DM9000AEP and transmitted to the computer, receive and display program were called to realize dynamic monitoring of the welding process. The flow chart is shown in Fig. 60.7.

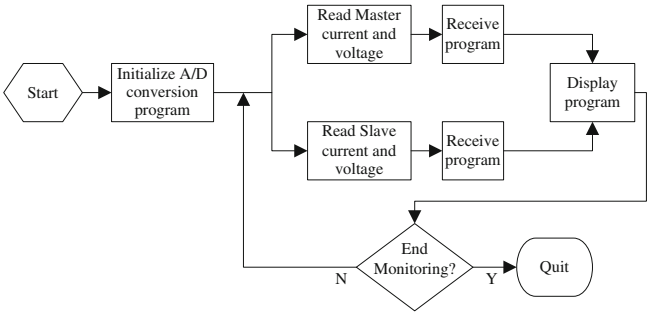
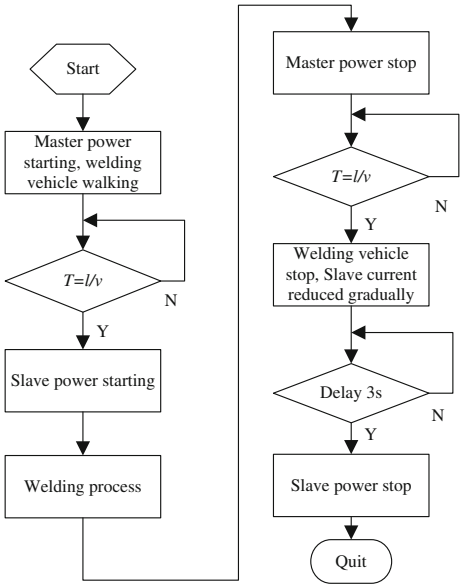


Fig. 60.7 The welding process monitoring procedure

Fig. 60.8 The flow chart of welding sequential control



60.3.3 Welding Sequential Control Procedure

In the control of twin-arc submerged arc welding, to obtain satisfactory weld seam quality, sequential control of the master and slave arc-starting and -ending is the utmost importance. In ordinary twin-arc submerged arc welding process, because of the space between two wires, there were two sections of the same length with the wire spacing at the head and trail of the weld seam, which were actually only suffered from the thermal effect of a single arc, as a result of this, collapse phenomenon would be caused at the head and trail of the weld seam.

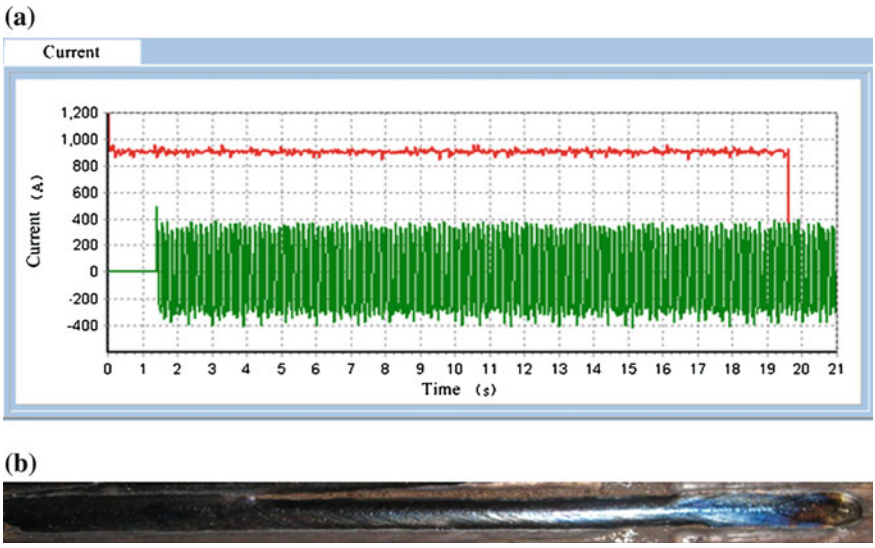


Fig. 60.9 Waveform of welding current and weld seam appearance

Table 60.1 Criterions and effects of twin-wire SAW

Current (A)		Voltage (V)		Welding speed (m/min)	Wire space (mm)	Welding conditions
Front wire	Back wire	Front wire	Back wire			
910	770	42	43	1.4	30	Arc-starting and -ending well, welding process stable and welding seams performing well

To guarantee the uniformity of weld width and reduce the collapse level, starting and ending of the two welding power sources could be controlled in a certain procedure. At the starting of welding process, start the master power and welding vehicle first, when the back wire moved to the initial start position of the front wire, start the slave power. At the ending of welding process, stop the master power and wire feeding when the front wire moved to the preinstalled destination, the welding vehicle continues walking, when the back wire moved to the preinstalled destination, stop the welding vehicle and reduce current and voltage of the back wire gradually, the program delay 3 s, use the arc back burning to fulfill the crater, then cut off the slave power supply and stop wire feeding. By the sequential control of the master and slave power supply and welding vehicle, narrow weld seam at the head and trail of the whole weld seam can be prevented, and the collapse level at the ending of weld seam also be reduced, a better weld seam quality was obtained. The flow chart of welding sequential control is shown in Fig. 60.8.

60.4 Test Experiment

The monitoring system designed in this paper was applied to high-speed twin-arc submerged arc welding to do test experiment. Test conditions: MZ1250 + MZE1000 submerged arc welding inverter power, low-carbon steel plate, thickness of plate is 20 mm, front wire Φ5 mm and back wire Φ4.8 mm, wire grade H08A, welding flux HJ431 and surfacing method was adopted. In these test conditions, chose the arc-starting and -ending program and set the wire spacing when monitoring system started. The criterions and parameters shows in Table 60.1, relevant current waveform and weld seam appearance shows in Fig. 60.9. It can be seen from the figure that there were no short circuit and arc interruption in the welding process and the process was stable, arc-starting and -ending well and perfect forming of weld seam (Table 60.1).

It can be seen from the test result that, the experiment of twin-arc submerged arc welding based on Ethernet monitoring system, which can real-time display electrical parameters in welding process, with well performance of stabilization, the collapse level was improved and the weld seam performing well.

60.5 Conclusions

The Ethernet communications which has characteristics of high speed, large capacity and high reliability was applied to the dynamic monitoring and control of high-speed twin-arc submerged arc welding process in this paper, hardware and software design of the monitoring system were introduced in detail. The hardware part concludes the principle of system components, data acquisition circuits and control interfaces. In the software part, communication procedure based on TCP/IP protocol was designed, and sequential control of twin arc-starting and -ending was achieved, which guaranteed the stability and order of welding process. The test results indicated that the system was achieved the technology of twin-arc high speed arc welding, and the collapse phenomenon at the head and trail of the weld seam was effectually avoided, the weld seam performance was improved.

Acknowledgments Project supported by Hunan Provincial Natural Science Foundation of China (11JJ2027), National Natural Science Foundation of China (51005073), Scientific Research Fund of Hunan Provincial Education Department (10C0682), Project of Human Provincial Research Scheme (2011GK3052), CEEUSRO special plan of Hunan province(2010XK6066), Industrial Cultivation Program of Scientific and Technological Achievements in Higher Educational Institutions of Hunan Province (10CY008), also from Aid program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province, are gratefully acknowledged.

References

1. He KF, Huang SS, Li XJ (2011) China Mech Eng 22(2):235–239
2. Xue HT, Li JY, Zhang XN (2003) Transactions of the China Welding Institution. vol 24–1, pp 19–22
3. Zhang HB, Huang SS, Zhou YQ (2004) Electric welding machine. vol 34–11, pp 25–27
4. He KF, Huang SS, Sun DY (2008) A expert system for submerged arc welding. J South China Univ Technol (Nat Sci Ed) 36-8:79–82
5. Chai Y, Wang YT, Chen H (2004) Computer automated measurement and control. 12:1188–1191

Chapter 61
Characteristic Testing and Experiment to Alternating Current Square Wave Submerged Arc Welding Inverter

Kuan Fang He, Qi Li, Si Wen Xiao and Yi Dao Yu

Abstract Based on the second inverting main circuit topology, a high-power alternating current square wave submerged arc welding power source is developed by combining limiting current and inverting parallel connection technology. The proposed inverting power source was tested and done technologic experiment. The results indicated that the novel power source can be used in thick wire or the multi-wires highly effective submerged arc automatically welding well because it possesses characteristics as follows: the novel power source were good external characteristics, excellent dynamic performance, quick speed of alternating polarity and good stabilisation of welding process, the shape of welding joint is smoothly and artistic.

Keywords Second inverting main circuit · Alternating current square wave · Power source · Testing and experiment

K. F. He (✉) · Q. Li · S. W. Xiao · Y. D. Yu
Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China
e-mail: hkf791113@163.com

Q. Li
e-mail: richie.lsq@gmail.com

S. W. Xiao
e-mail: xswxsw888@sohu.com

Y. D. Yu
e-mail: yydhnxt@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_61,

61.1 Introduction

In the submerged arc welding production process, the high-power alternating current square wave power source is mainly made by thyristor, which is relatively reliable and more technically mature. But the equipment is bulky, heavy, high energy consumption, low efficiency and the poor arc stability because of slow zero-crossing rate. The dynamic and static characteristic of the high-power power source are not ideal because of its structure. The emergence of large capacity insulated gate bipolar transistor (IGBT) improve this situation, the using of IGBT inverter technology is the best way to solve these problems with the comprehensive benefits of faster switching speed, the control current capability, low conduction losses and the small required driving power. Therefore the IGBT is particularly suitable for high alternating current square wave inverting circuit with the following characteristics. First, it can improve the speed of current polarity change output of the welding power source and the stability of alternating current arc burning. Second, it is easy to improve electrical performance and control precision and intelligent control by digital control, Third, the IGBT switching saturation voltage is low, which helps to reduce the tube power loss [1, 2]. However, the capacity of IGBT modules appeared on the domestic market is limited within a certain range of the voltage of 1,200 V and the maximum current of 600A. In addition, the single high frequency transformer currently can only achieve power output of 25–30 kW as the constraints capacity of magnetic material in domestic production, which limits applications of single high-power IGBT inverter. Therefore, the high-power digital controlled arc welding inverter is one of the key technologies of high-speed submerged arc welding equipment [3, 4]. Reference [5] proposed an alternating current square wave submerged arc welding power source based on parallel inverter to achieve high current output of 1,000 A. In this paper, the performance and technology are tested and further researched to provide technical basis for the popularisation and application of high-power alternating current square wave submerged arc welding power source.

61.2 Principle of High-Power Alternating Current Square Wave Arc Welding Inverter

Figure 61.1 shows the schematic diagram of IGBT alternating current square wave arc welding inverter, which is made up of the IGBT direct current arc welding inverter and the second full-bridge inverting circuit. The inverter can output the required characteristics of submerged arc welding direct and alternating current square wave electricity. Alternating current square wave is transformed by the second full-bridge inverting circuit. The principle is as follows: full-bridge IGBT four arm is composed by VT_1, VT_4 and VT_2, VT_3 which has access to the arc load diagonal. When VT_2, VT_3 are inspired, VT_1, VT_4 are turned off, positive current

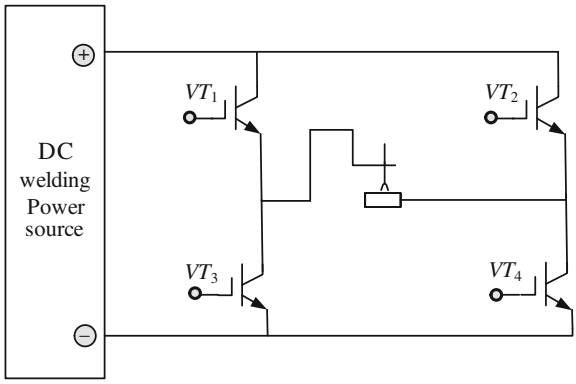


Fig. 61.1 IGBT alternating current inverter technology

polarity is offered to arc through the inverter. When VT_2, VT_3 are turned off, VT_1, VT_4 are inspired, negative current polarity is offered to arc through the inverter. Variable frequency, polarity and adjustable duty cycle of alternating current square wave current waveform can be obtained by controlling the time and their relative proportion of the two groups of IGBT. No-load output voltage and external characteristic of alternating current square wave arc welding inverter are achieved by regulating and controlling of direct current arc welding inverter.

Alternating square wave outputting current is achieved by using full-bridge inverter technology and advanced power electronic switch IGBT module, which first increases frequency of 50–20 kHz alternating current through inverting transformation, and go through a fast rectifier to get direct current, at last get alternating square wave current by transformation of the second inverter. The weight, size and materials of welding power source are greatly reduced contrasting to the traditional one efficiency. Submerged arc welding process is achieved by constant current feedback characteristics, the variable speed adjusting system constituted by the arc negative feedback and automatic wire feeding control circuits, running mechanism and flux protection.

The main circuit topology is composition of two bridge inverter by limiting their current in order to increase the output capacity and solve the safe and reliable problem of parallel operation of inverter. The current distribution, integrated circuits and programming of control systems are used to coordinated control and precise adjustment for welding parameters, which is to create conditions for matching the welding parameters and improving weld quality. The good, high stability of arc and welding performance are gained by optimising output reactor of main circuit and time constant of controlling loop and improving the dynamic characteristics and consistency of dynamic parameters of inverter. In design, the rational distribution of the protection circuit parameters of IGBT is important to improve its efficiency and meet the reliability and safety requirements of high-power inverter.

Figure 61.2 is the limiting parallel block diagram of double bridge arc welding inverter. In the figure, the rated output capacity is gotten by the two parallel

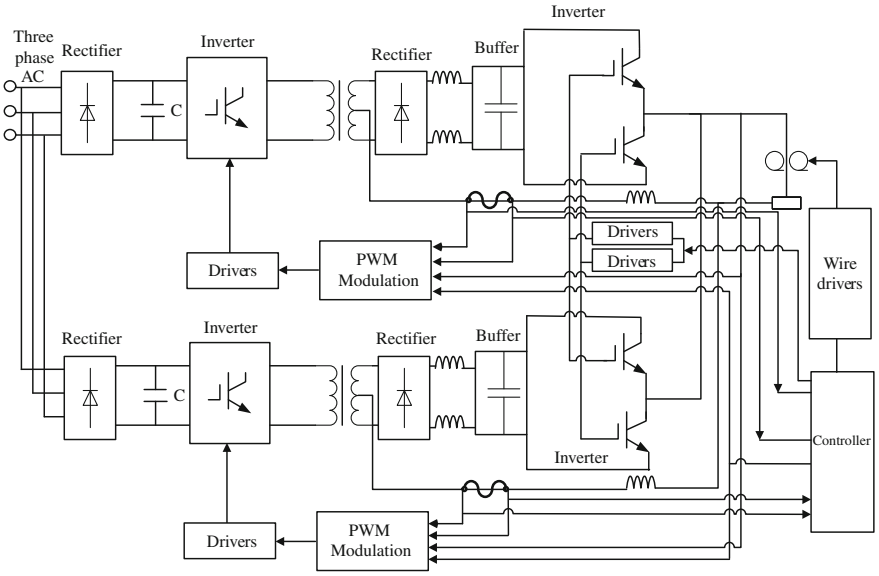


Fig. 61.2 Parallel connection principle of the inverter

full-bridge inverters. The output current of two basic inverters reaches equilibrium by limiting control to negative feedback, respectively.

61.3 Measurement and Analysis of Submerged Arc Welding Inverter

61.3.1 Testing Platform

The 1,000 A arc welding power source testing platform are composed of TDS2012B digital oscilloscope of TEKTRONIX Company, waveform analysis software of latest WaveStar^{TM} Software for Oscilloscopes V1.2 Version and other components.

61.3.2 The External Characteristics of the Submerged Arc Welding Inverter

Figure 61.3 shows the measured voltage and current values in the given output currents of 54, 320, 540, 800 and 1,000 A. In the figure, no-load voltage of the developed submerged arc welding inverter is 80 V, constant current characteristic

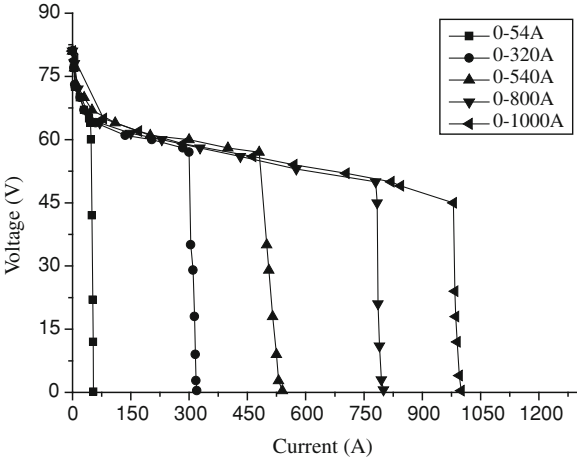


Fig. 61.3 The measured constant current external characteristics

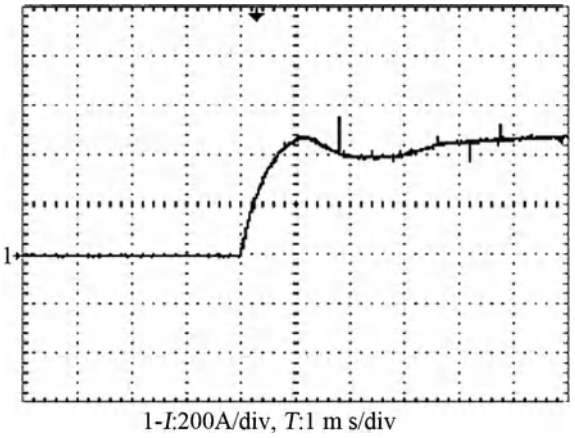


Fig. 61.4 The current response of 0–430 A

is good at different output currents, which shows constant current characteristics of the submerged arc welding inverter is ideal.

61.3.3 The Dynamic Characteristics of the Submerged Arc Welding Inverter

The curve of Fig. 61.4 is current response when current changing from 0 to 430 A under the simulated load of 0.1 Ω. In Fig. 61.4, the current changing from 0 to 430 A just needs time of 2 ms. The Fig. 61.5 is current response when current changing from 400 to 0 A under the simulated load of 0.1 Ω. In Fig. 61.6, the current changing from 400 to 0 A just needs a time of 3 ms.

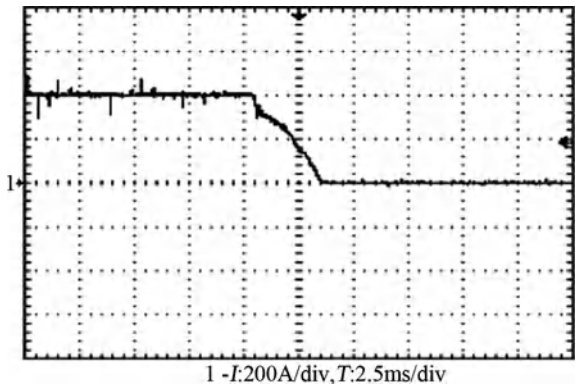


Fig. 61.5 The current response of 400–0 A

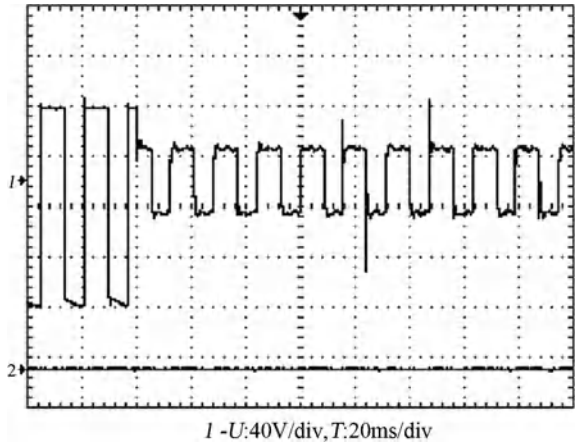


Fig. 61.6 The voltage response of no-load to load

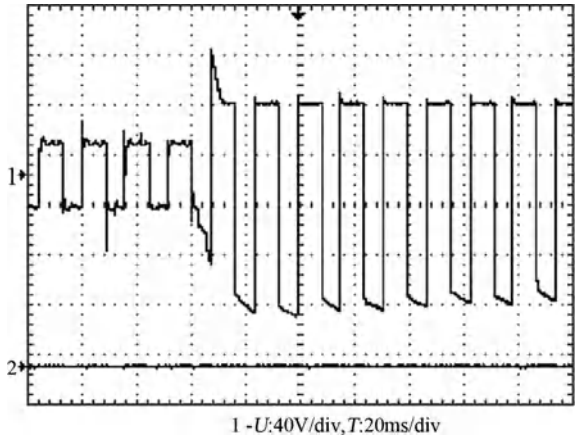


Fig. 61.7 The voltage response of load to no load

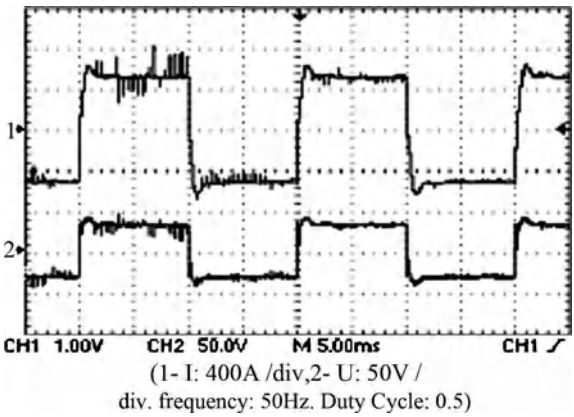


Fig. 61.8 The current and voltage waveform

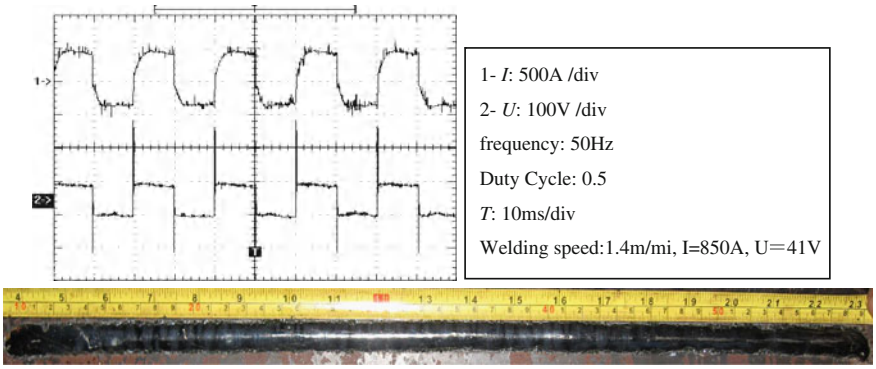


Fig. 61.9 Waveform of current and voltage and welding seam

Figure 61.6 is the response voltage curve of no-load to load. Figure 61.7 is the response voltage curve of load to no-load. It can be seen from the figures, the outputting voltage change very rapidly, over time is very short, and the welding inverter response to changing of external load is very fast. The testing result concludes that alternating current square wave submerged arc welding inverter also has better dynamic performance.

Figure 61.8 is output current and voltage waveforms of the second square wave welding inverter. It can be seen from Fig. 61.8 the polarity switching of alternating current inverter is high speed to the zero-crossing with stable waveform.

61.4 Experiments

Experiments have been done by the developed alternating current square wave submerged arc welding machine of MZE1000. The material of work piece is low carbon steel with slab thickness of 20 mm, the welding wire trademark is H08 A

and its diameter is 4.0 mm, welding flux is HJ431. Figure 61.9 shows experimental results in the welding current of 850 A and arc voltage of 41 V. Seen from the measured current and voltage waveform in Fig. 61.9, the welding process is stability and high speed to the zero-crossing. Seen from the photographs of weld seam in Fig. 61.9, the appearances are smooth and ripple fine.

61.5 Conclusions

The proposed power source was tested and done technologic experiment. The results show it possesses characteristics as follows: the novel power source are good external characteristics, excellent dynamic performance, quick speed of alternating polarity and good stabilization of welding process, the shape of welding joint is smooth and artistic. The novel power source can be highly effective used in thick wire and the multi-wires submerged arc automatically welding.

Acknowledgments Project supported by Hunan Provincial Natural Science Foundation of China (11JJ2027), National Natural Science Foundation of China (51005073), Scientific Research Fund of Hunan Provincial Education Department (10C0682), Project of Hunan Provincial Research Scheme (2011GK3052) also from Aid program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province, are gratefully acknowledged.

References

1. Li HQ (2006) Developing status of SAW at home and abroad. Electr Weld Mach 36:1–6
2. Wang ZM, Huang SS (2006) Inverter for soft-switching double-wire pulse metal active gas welding. J South China Univ Technol Natl Sci Ed 34:31–34
3. Li YB, Huang SS, Wang ZM (2003) Theoretic analysis of static characteristics of novel soft-switching submerged arc welding inverter. J South China Univ Technol Natl Sci Ed 31:39–42
4. Du GP, Huang SS (2002) Study on peak-current mode in high power soft-switching inverter. Power Electron 36:19–24
5. He KF, Huang SS (2008) Research on high-power submerged arc welding AC square-wave inverting power source. J South China University of Technol Natl Sci Ed 36:79–82

Chapter 62
Study of Current Controller Design and Performance Based on PMA RSM Model Decoupling

Yu Guo, Taihua Wang and Yong-le Ai

Abstract The permanent magnet assisted reluctance synchronous machine (PMA RSM) is a controlled object of multi-variable, nonlinearity and strong coupling. The fact that *d*-axis current can be controlled by *d*-axis voltage directly, same as the *q*-axis current, can not be realized even if the field current and torque current have been decoupled in the synchronous coordinates. A decoupling method for the electrical model is presented, which means the speed voltage and mutual inductance voltage are subtracted in the machine model proposed. Therefore the control of d- and *q*-axis components is changed to single-loop independent control system. Then a current controller is designed, which is applied to the current control of PMA RSM. The result from simulation shows that it is of high performance of dynamic and robustness.

Keywords Motor model decoupling · Current controllor · Matlab simulation

62.1 Introduction

As the core of the power PM control system, the PMA RSM has a wide range of applications in power locomotive traction, submarines and aerospace for small volume, light weight, rotor no fever, control system simpler than the asynchronous motor [1–5].The rotor field-oriented vector control mode is a wide range control mode of the PMA RSM [6]. The basic principle is through the coordinate

Y. Guo (✉) · T. Wang · Y. Ai
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: guoyu@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_62,

transformation, and in the rotor field-oriented synchronization of the magnetic field on an electric motor current and torque current decoupling control, have the same operation characteristic as the DC motor. But in general, it is unable to control d-axis voltage through the control d-axis current, and to control q-axis voltage through the control q-axis current directly. In order to simplify control, many literatures simplify the model of the motor the further, such as ignore the effect of the resistance. But when the motor speed range is very wide, resistance, the parameters such as reactance produce larger error to d- and q-axis current control, thus affecting the control accuracy and the dynamic response speed.

Aiming at this issue, based on the theory of the rotor field-oriented vector control [7], the author puts forward some method of the motor ontology decouple, designs a kind of current controller, and through the Matlab/Simulink verified some performance of the motor under the action of the current controller.

62.2 Motor Ontology Model Decoupling

The equations for the electrical model of the PMA RSM are coupled: there is a q-axis term in the d-axis equation and vice versa, which is given by:

$$v_d = r_s i_d + L_d' \frac{\mathrm{d}i_d}{\mathrm{d}t} + M_d' \frac{\mathrm{d}i_q}{\mathrm{d}t} - \omega_{re}(\lambda_q - \lambda_{\mathrm{pm}}) \tag{62.1}$$

$$v_q = r_s i_q + L_q' \frac{\mathrm{d}i_q}{\mathrm{d}t} + M_q' \frac{\mathrm{d}i_d}{\mathrm{d}t} + \omega_{\mathrm{re}} \lambda_d \tag{62.2}$$

in the equations:

L_d'	stator self-inductance;
ω_{re}	rotational speed;
λ_d, λ_q	d- and q-axis flux linkage;
L_q'	rotor self-induction;
v_d, v_q	d- and q- voltage;
i_d, i_q	d- and q- current;
M_d'	mutual inductance;

The block diagram that represents Eqs. (62.1) and (62.2) is shown in Fig. 62.1.

Note that this block diagram is drawn in Matlab's Simulink. Inputs and output are indicated by numbered capsules. Unfortunately the Greek alphabet is not available in Simulink and therefore the electrical rotor speed is written as Wre, the d-axis mutual inductance as Md, the q-axis flux linkage as Q_flux etc. Also note that the instantaneous inductance L_d' is written as L_d in the simulation diagram and $L_d \neq L_d$, where L_d is the linearized inductance.

From the block diagram in Fig. 62.1 it is evident that to control the d-axis and q-axis currents would be difficult, since they are coupled. Moreover, cross

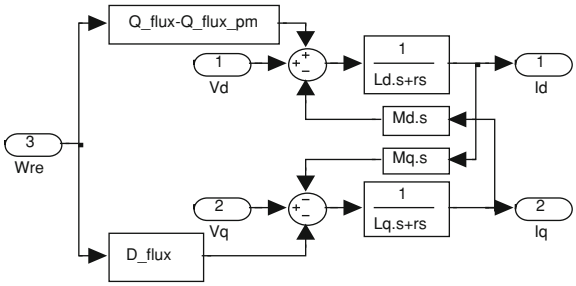


Fig. 62.1 Coupled model of the PMA RSM

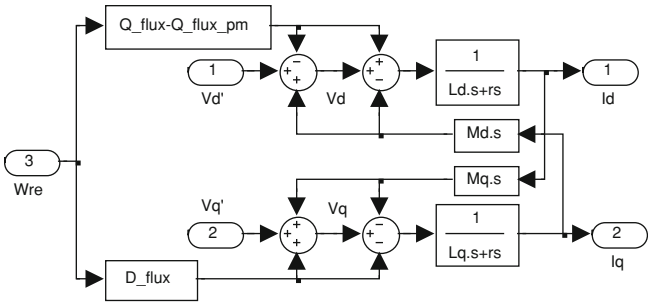


Fig. 62.2 Decoupling procedure for the electrical model

coupling voltage is relevant to other factors, such as the speed. Variable frequency speed regulation, cross coupling potential changes with ω_{re}, the existence of the coupling potential directly affect speed control performance of the speed regulating system. Therefore, a decoupling procedure is suggested to simplify the design of the current controllers. A decoupling procedure is suggested in Fig. 62.2. The speed voltages and mutual inductance terms are subtracted by the controller where they are added in the machine and vice versa.

This procedure results in decoupled differential equations represented by the block diagram shown in Fig. 62.3.

It is important to make the distinction between the real machine voltages, V_d and V_q, and the applied control voltages, V'_d and V'_q, in the current controllers. The relationship is given below:

$$V_d = V'_d + M_d \frac{\mathrm{d}i_q}{\mathrm{d}t} - \omega_{\mathrm{re}}(\lambda_q - \lambda_{\mathrm{pm}}) \tag{62.3}$$

$$V_q = V'_q + M_q \frac{\mathrm{d}i_d}{\mathrm{d}t} + \omega_{\mathrm{re}}\lambda_d \tag{62.4}$$

Profile usually refers to the flatness of the rolled material and whether partial waves, buckle folding, lateral bending and so on, appear on the transverse places

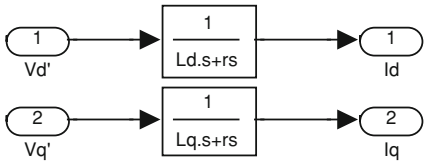


Fig. 62.3 Decoupled electrical model

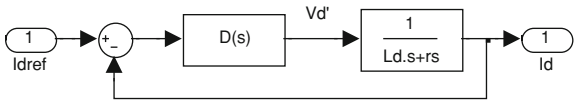


Fig. 62.4 *d*-axis current control system

of the foil material. The profile's quality depends on whether the extension beyond board band in the direction of width is equal.

62.3 Design the Current Controllers

Consider the block diagram in Fig. 62.4 which describes a closed loop control system for the *d*-axis part of the decoupled electrical machine equations.

Because scale controller has strong anti-jamming capability, transition time is short, this chapter presents the design of proportion, the controller are calculated respectively the *d*- and *q*-axis proportion coefficient of the current loop.

The controller gain K_d, should be chosen so that loop gain at the Nyquist frequency is −10 dB:

$$20\log_{10}\left(K_d\cdot\left|\frac{1-e^{-\frac{r_sT}{L'_d}}}{r_s\left(-1-e^{-\frac{r_sT}{L'_d}}\right)}\right|\right)=-10\Rightarrow K_d=10^{-10/20}\cdot\left|\frac{r_s\left(-1-e^{-\frac{r_sT}{L'_d}}\right)}{1-e^{-\frac{r_sT}{L'_d}}}\right|$$
$$=10^{-10/20}\cdot r_s\cdot\frac{1+e^{-\frac{r_sT}{L'_d}}}{1-e^{-\frac{r_sT}{L'_d}}} \tag{62.5}$$

Using a series expansion for the exponential terms, the computational intensity of the formula given by Eq. (62.5) can be reduced, an approximation for Eq. (62.5) can be derived:

$$K_d = 10^{-10/20} \cdot \left| \frac{r_s\left(-1 - \mathrm{e}^{-\frac{r_s T}{L'_d}}\right)}{1 - \mathrm{e}^{-\frac{r_s T}{L'_d}}} \right| = 10^{-10/20} \cdot r_s \cdot \frac{1 + \mathrm{e}^{-\frac{r_s T}{L'_d}}}{1 - \mathrm{e}^{-\frac{r_s T}{L'_d}}}$$
$$= 10^{-10/20} \cdot r_s \cdot \frac{1 + \left(1 - \frac{r_s T}{L'_d} + \left(-\frac{r_s T}{L'_d}\right)^2 + \cdots\right)}{1 - \left(1 - \frac{r_s T}{L'_d} + \left(-\frac{r_s T}{L'_d}\right)^2 + \cdots\right)} \tag{62.6}$$

The derivation of Eq. (62.6) assumes that the sampling time is sufficiently small so that any terms with T^2 or greater powers are negligible, then Eq. (62.6) can be evaluated:

$$K_d \approx 10^{-10/20} \cdot r_s \cdot \frac{\left(2 - \frac{r_s T}{L'_d}\right)}{\frac{r_s T}{L'_d}} = 10^{-10/20} \cdot \frac{L'_d}{T} \cdot \left(2 - \frac{r_s T}{L'_d}\right) \tag{62.7}$$

Generally, $r_s T / L'_d$ is small enough, it can be ignored, then Eq. (62.7) can be simplified:

$$K_d \approx 10^{-10/20} \cdot \frac{L'_d}{T} \cdot 2 \tag{62.8}$$

If the sampling frequency is 5,000 Hz, then Eq. (62.8) can be evaluated:

$$K_d = 10^{-10/20} \cdot 5000 \cdot L'_d \cdot 2 = 3162.278 \cdot L'_d \tag{62.9}$$

Because d-axis value scope of self-induction is for $1 \sim 8$ mH, so in any sample moment, d-axis current controller gain must be calculated. Similarly, the q-axis current controller gain is given by:

$$K_q = 3162.278 \cdot L'_q \tag{62.10}$$

Here, the q-axis self-induction is typically 0.75 mH. Evaluating Eq. (62.10) for $L'_q = 0.75\,\mathrm{mH}$, the q-axis current controller gain is given by:

$$K_q = 2.372 \tag{62.11}$$

The complete block diagram for the current controllers is shown in Fig. 62.5.

Let the controller be a constant gain: $D(s) = 10$. The closed loop transfer function is then $H_{cl}(s) = \frac{10}{L_d s + r_s + 10}$. Suitable parameters are $r_s = 15\,\mathrm{m}\Omega$ and $L'_d = 3.16\,\mathrm{mH}$ The closed loop Bode plot and step response is shown in Fig. 62.6. From the closed loop Bode plot it can be seen that the bandwidth of the closed loop system is 5,000 rad/s. Furthermore, the settling time for a step input to the closed loop is $\tau_s = 0.8\,\mathrm{ms}$. Therefore the controller and plant in a closed loop can give the required performance needed for current control of the PMA RSM.

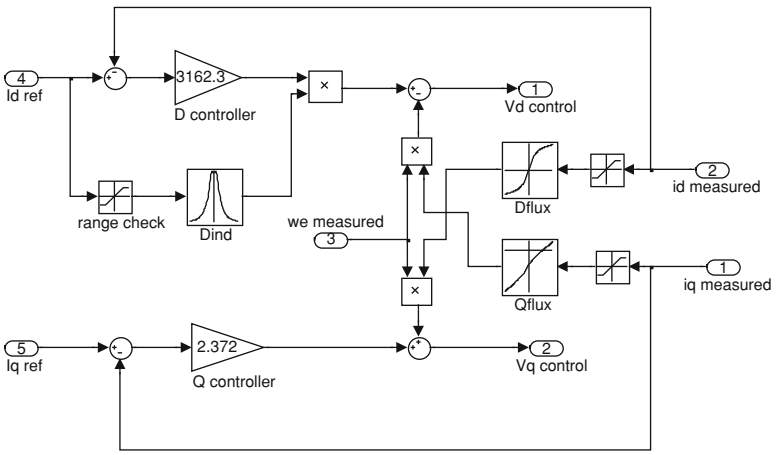


Fig. 62.5 Current controllers with decoupling

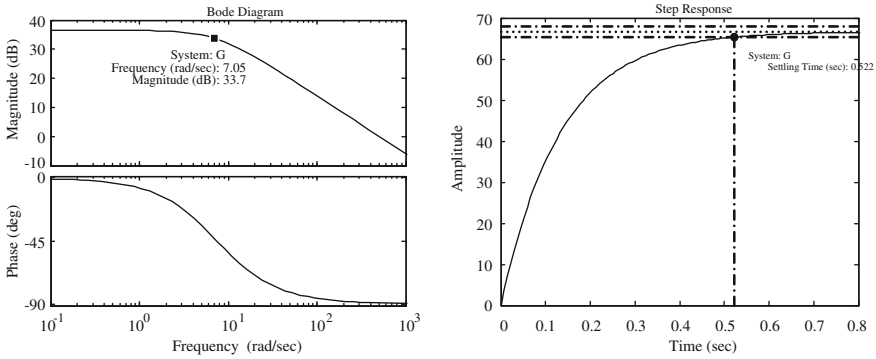


Fig. 62.6 Bode plot (*top*) and step response (*bottom*) for closed loop controller

62.4 Simulation of the Current Controllers

In order to test the performance of the current controller, it is necessary to build current control system model, which consist of Electrical model subsystems and control subsystem, with current references and output blocks are shown in Fig. 62. 7.

The measured speed is set to zero: this can be viewed as a locked rotor test. Note that the current output of the Electrical model is followed by a ZOH element. This represents the analog-to-digital conversion. As a first test, the q-axis current reference is set to zero and the d-axis current reference to 100 A at $t = 1$ s. The second test is to set the d-axis reference current to zero and the q-axis current reference to 100 A at $t = 1$ s. The simulation results are shown in Fig. 62.8.

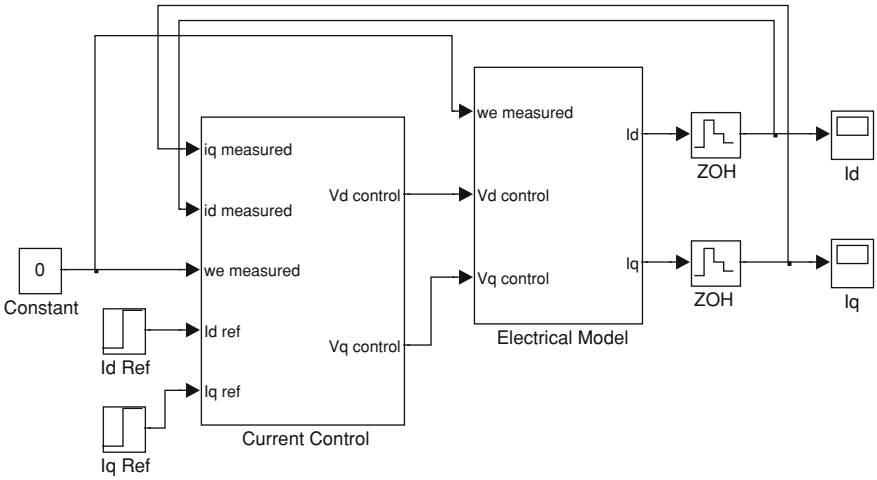


Fig. 62.7 Current controller test setup

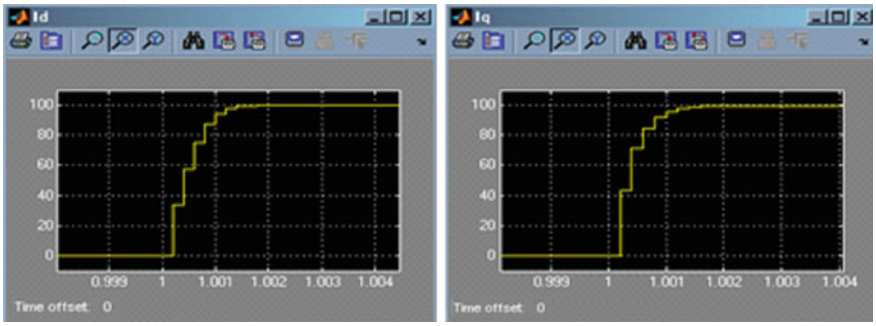


Fig. 62.8 *d*-axis current response (*left*) and *q*-axis current response (*right*)

From Fig. 62.8 it can be seen that both currents respond to their reference commands within 2ms, which shows the response of the current controller designed in this chapter is faster. *d*-axis current and *q*-axis current is decoupling, independent of each other.

62.5 Conclusion

According to the characteristics of PMA RSM model, this chapter puts forward the motor ontology decoupling method, and designs the stable current controller. In the design of the current loop, The proportion controller is used, and simulated experiment to the step response of d- and *q*-axis current given. The simulation

results show that the decoupling measures of motor ontology presented in this chapter is viable, and the current controller has better robustness and dynamic response performance.

References

1. Lipo TA (1991) Synchronous reluctance machines-A viable alternative for AC drives. Electr Mach Power Sys 19(6):659–671
2. Matsuo T, Lipo TA (1994) Rotor design optimization of synchronous reluctance machine. IEEE Trans Energy Convers (2):359–365
3. Kamper MJ, Mackay AT (1995) Optimum control of the reluctance synchronous machine with cageless flux barrier rotor. SAIEE Trans 86(2):49–56
4. Germishuizen J, Van der Merwe FS, Van der Westhuizen K, Kamper MJ (2000) Performance comparison of reluctance synchronous and induction traction drives for electrical multiple units. IEEE IAS Conference (Rome), Oct 2000
5. Erashima M (1997) Novel motors and controllers for high-performance electric vehicle with four in wheel motors. IEEE Trans Iad Appl 44(1):28–38
6. Guglielmi P, Pastorelli M, Pellegrino G, Vagati A (2004) Position-sensorless control of permanent-magnet-assisted synchronous reluctance motor. IEEE Trans Ind Appl 40(2): 615–622
7. Wildi T (2002) Electrical machines, drives, and power systems. Prentice Hall, Englewood Cliffs

Chapter 63
Harmonic Analysis Based on Independent Component Method

Ai Yongle and Zhang Haiyang

Abstract Harmonic analysis plays an important role in power quality research due to the harmful impact to power system and electric apparatus. It is the demand to accurate analysis of harmonics in case to filter. This paper presents a blind source separation algorithm, which combines independent component analysis (ICA) and least square method. During the harmonic analysis, ICA is used to determine harmonic frequencies first, then the least square method is used to determine amplitudes and phases of harmonic components. By MATLAB simulation, the result is shown that harmonics can be separated effectively with a high precision by using the method proposed under the condition with low real-time demand.

Keywords Harmonic analysis · Independent component analysis · Least square method · Matlab simulation

63.1 Introduction

With the development of power electronic technology, problems relevant to harmonics are becoming more and more serious due to growing applications of non-linear load devices such as inverters, converters, switching power supply,

A. Yongle (✉) · Z. Haiyang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: aiyongle@hpu.edu.cn

Z. Haiyang
e-mail: haiyang5256@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_63,

which seriously affect the safe operation of power system and electromagnetic environment surrounding the lines. Therefore, harmonic analysis becomes one of the key issues to improve the system's power quality [1].

There are a variety of harmonic analysis methods [2, 3], which are of their own characteristics. Such as, the method of instantaneous reactive power theory has the properties of fast dynamic response speed, small time delay and strong real-time, but its topology structure is very complicated; Fourier transform and its improved harmonic method which are high precision and multi-functional have a large amount of calculation and exist spectrum leak and fencing effects; The method of neural network can apply neurons' self-learning algorithm to the harmonic detection, but its convergence will be influenced by the size of learning rate.

Independent component analysis (ICA) developed in 1990s is one kind of blind source separation (BSS) methods and it is applied to many fields. This paper models the harmonic load identification problem as a blind source separation task and to solve it using a method combining FastICA and Least Square method together. First, the proposed method uses FastICA to determine harmonic orders by constructing proper number of virtue observed channels; Then the Least Square method is used to determine amplitudes and phases of harmonics containing in the mixing signals. The result is shown that the method presented in this paper to solve the harmonics is fast speed and high precision.

63.2 Principle of ICA

ICA is a new BSS technology [4–8], which is defined that multiple independent components are decomposed from observed signals by optimization algorithm to obtain estimation of the source signals according to principle of statistical independence.

Assume that we observe M linear mixtures $x_1, x_2,\ldots,x_M$ of N independent sources:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{iN}s_N, \text{ for all } i. \tag{63.1}$$

Now rewrite the above equation by using vector–matrix notation. Denote by x the random vector whose elements are the mixtures $x_1, x_2,\ldots,x_M$, and likewise by s the random vector with elements $s_1, s_2,\ldots,s_N$. Denote by A the matrix with elements a_{ij}. Therefore, the above mixing model (63.1) can be written as:

$$x = As. \tag{63.2}$$

When the mixing signal is collected by K times, the model will be written as:

$$X = AS = \sum_{j=1}^{n} a_j s_j. \tag{63.3}$$

In Eq. (63.3), X, A, S are a M × K, M × N, N × K matrix, respectively. The goal is to find a linear transform matrix W so that $Y = WX = WAS$ approach S as much as possible without knowledge of the mixing matrix A and original sources S. The following assumptions are adopted for ICA: (1) The components s_i are statistically independent; (2) At most one of the components s_i is Gaussian distributed; (3) The number of observed mixtures M is no less than that of independent components N; (4) For simplicity, any noise terms are omitted.

In this paper, a FastICA algorithm which is based on a fixed-point iteration scheme for finding a maximum of the non-Gaussian is used.

However, before applying FastICA algorithm on the data, it is usually very useful to do some preprocessing. The most basic and necessary preprocessing is to center X, i.e. subtract its mean vector $m = E\{X\}$ so as to make X a zero-mean variable. After centering, another useful preprocessing strategy is to whiten the observed variables. The X is transformed linearly so that a new Z which is white, i.e. its components are uncorrelated and their variances equal unity can be obtained.

The basic form of the FastICA algorithm is as follows:

Let $i = 1$;

Choose an initial weight vector $w_i(0)$(It satisfies$\|w_i(0)\| = 1$), and let $k = 1$;

Let $w_i(k) = E\left\{z\left(w_i(k-1)^T z\right)^3\right\} - 3w_i(k-1)$;

Normalize $w_i(k)$, $w_i(k) = w_i(k)/\|w_i(k)\|$;

If not converged, let $k = k + 1$ and go back to step (3);

If $i < N$, let $i = i + 1$ and go back to step (2).

The resulting separated matrix is $W = [w(1), w(2), \ldots, w(M)]^T$, then the mixing matrix A can be obtained by computing the inverse of W.

63.3 Application in Harmonic Detection Based on ICA

63.3.1 Construction of Virtue Observed Channels

The electric signals which can be directly observed is a cyclical non-sinusoidal signals, and denote it by x_1. When x_1 is decomposed by Fourier transformation, fundamental and a series of harmonic components can be obtained.

$$\begin{aligned} x_1 &= a_{11}s_1 + a_{12}s_2 + \cdots + a_{1N}s_N \\ &= a_{11}\sin(\omega t + \varphi_1) + a_{12}\sin(2\omega t + \varphi_2) + \cdots + a_{1N}\sin(N\omega t + \varphi_N) \end{aligned} . \quad (63.4)$$

During the process of harmonic analysis, there is only one actual observed channel, while it is demanded that the number of observed channels M is no less than that of independent components N for ICA model. Therefore, it is needed that at least another (M-1) channels should be constructed, then it can be constructed by the following (M-1) expansion equations

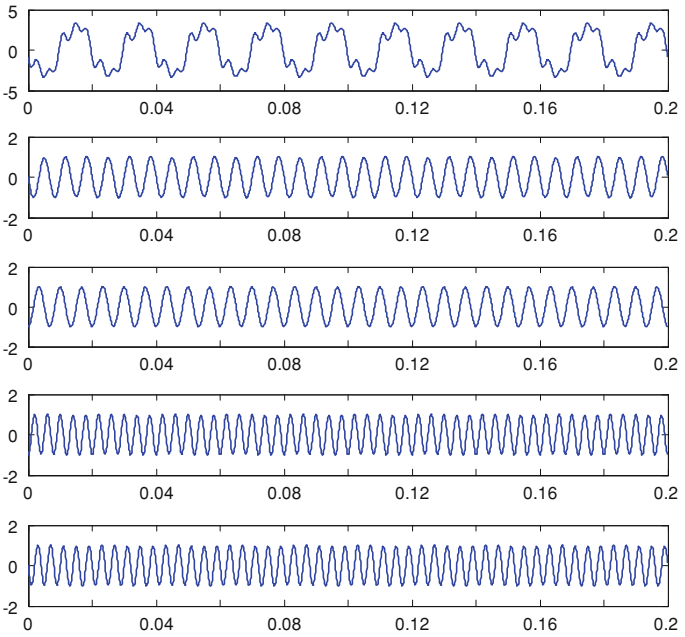


Fig. 63.1 Five observed signals

$$\begin{cases} x_2 = a_{21}s_1 + a_{22}s_2 + \cdots + a_{2N}s_N = a_{21}\sin(\omega t + \varphi_1) + \cdots + a_{2N}\sin(N\omega t + \varphi_N) \\ x_3 = a_{31}s_1 + a_{32}s_2 + \cdots + a_{3N}s_N = a_{31}\sin(\omega t + \varphi_1) + \cdots + a_{3N}\sin(N\omega t + \varphi_N) \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \\ x_M = a_{M1}s_1 + a_{M2}s_2 + \cdots + a_{MN}s_N = a_{M1}\sin(\omega t + \varphi_1) + \cdots + a_{MN}\sin(N\omega t + \varphi_N) \end{cases} \tag{63.5}$$

63.3.2 Determination of Harmonic Components

First, a period of timed-sampling data is obtained from the signals x_1, x_2, …, x_M, i.e. size (x_1) = size (x_2) = … = size (x_M) = (1×K). Then a matrix $X = [x_1, x_2, \ldots, x_M]^T$, i.e. size (X) = (M×K), is made up of. Next, the separate matrix W and mixing matrix A are obtained by using the FastICA algorithm.

However, there exists two ambiguities in the ICA model in Eq. (63.3). (1) We can not determine the variances of the independent components; (2) We can not determine the order of the independent components. As a result, amplitudes of the harmonic components need to be determined by the least square method after all the harmonic frequencies are obtained.

63.4 Simulation and Results Analysis

63.4.1 Simulation

One observed signal consisting of fundamental, 3rd and 5th harmonics and 4 other virtual signals are simulated by MATLAB.

$$\begin{aligned} x_1 &= 3\sin(\omega t) + 0.6\sin(3\omega t + \pi/3) + 0.8\sin(5\omega t + \pi/5) \\ x_2 &= \sin(3\omega t) \\ x_3 &= \cos(3\omega t) \\ x_4 &= \sin(5\omega t) \\ x_5 &= \cos(5\omega t) \end{aligned} \quad . \tag{63.6}$$

Ten cycles of time-sampling date are gathered during the simulation, and the five signals waves are shown in Fig. 63.1.

It is assumed that the data is preprocessed by centering and whitening. First, the harmonic frequencies are determined by using FastICA algorithm, and the resulting separated matrix W and mixing matrix A are as matrix (63.7) and (63.8), respectively.

$$W = \begin{bmatrix} -0.9998 & 0.0004 & 0.0007 & 0.0008 & 0.0006 \\ -0.0089 & -1.0000 & -0.0000 & 0.0000 & -0.0000 \\ 0.0068 & 0.0000 & -1.0000 & 0.0000 & 0.0000 \\ 0.0131 & 0.0000 & -0.0000 & -1.0000 & 0.0000 \\ 0.0076 & -0.0000 & 0.0000 & -0.0000 & -1.0000 \end{bmatrix} \tag{63.7}$$

$$A = \begin{bmatrix} -1.0002 & -0.0004 & -0.0007 & -0.0008 & -0.0006 \\ 0.0089 & -1.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.0068 & -0.0000 & -1.0000 & -0.0000 & -0.0000 \\ -0.0131 & -0.0000 & -0.0000 & -1.0000 & -0.0000 \\ -0.0076 & -0.0000 & -0.0000 & -0.0000 & -1.0000 \end{bmatrix} \tag{63.8}$$

Then the least square method is used to determine the amplitudes and phases of harmonic components, and the resulting waves of separated signals are shown in Fig. 63.2.

Error analysis of harmonic detection is shown in Table 63.1.

From Table 63.1, it is shown that the observed signals can be separated with a high precision by using the proposed algorithm.

63.4.2 Error Analysis

For testing the effectiveness of the combined algorithm, it is used to process data obtained from actual scenes. The original time-sampling date wave is shown in Fig. 63.3.

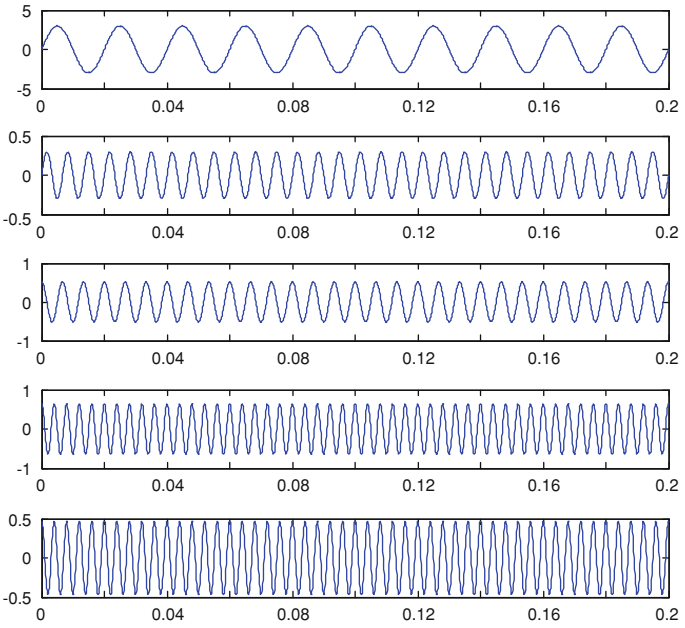
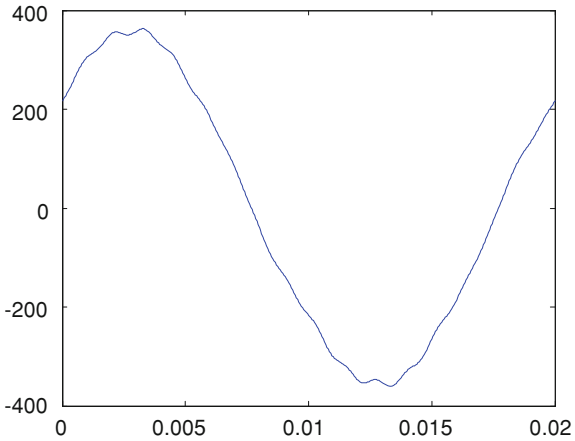


Fig. 63.2 Final separated signals

Table 63.1 Error analysis of harmonic detection

Harmonic orders	Fundamental		Third		Fifth	
	$\sin(\omega t)$	$\cos(\omega t)$	$\sin(3\omega t)$	$\cos(3\omega t)$	$\sin(5\omega t)$	$\cos(5\omega t)$
Amplitudes	3.0000	0.0000	0.3	0.5196	0.6472	0.4702
Amplitude errors	0		0.0022%		0.00344%	
Phases	0		$\pi/3$		$\pi/5$	
Phase errors	0		0		0	

Fig. 63.3 Original data



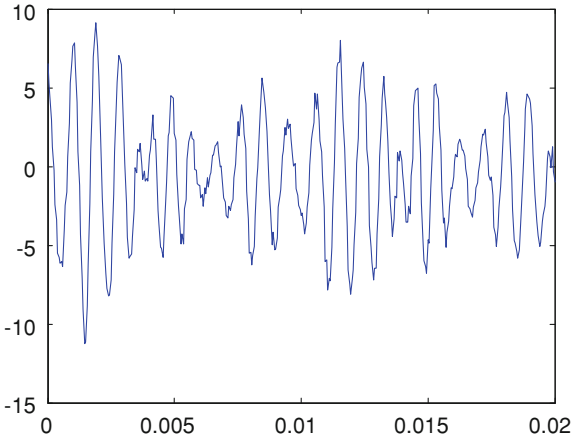


Fig. 63.4 Errors between real observed signals and separated signals

Similarly, errors are analyzed for the algorithm, and the errors between real observed signals and separated signals are shown in Fig. 63.4.

From Fig. 63.4, it is seen that the errors are quite small, and it is suggested that the actual signals can also be separated with a high precision by using the proposed algorithm.

63.5 Conclusion

This paper models the problems of harmonic detection as blind source separation task, and presents the famous FastICA algorithm to carry on harmonic analysis. However, there is only one observed signal that can be used, virtual signal channels needs to be constructed for ICA analysis. Besides, there exists two ambiguities in the ICA model. Therefore, the least square method is used to determine amplitudes and phases of harmonic components. The result is shown that fundamental and harmonics can be separated with a high precision and fast speed by using the algorithm presented in the paper.

References

1. Xiao X (2004) Power quality analysis and control. China Electric Power Press, Beijing
2. Li H, Yang S (2004) Summarization of the harmonic measurement method in electric power system. Mod Electr Power 21(4):39–44
3. Li S, Zhu Y, Zhou Y (2004) The over view of detecting methods for harmonic in power system. High Volt Eng 30(3):39–41
4. Aapo H, Juha K, Erkki O (2007) Independent component analysis. Electronic Industry Press, Beijing, p 6

5. Comon P (1994) Independent component analysis, a new concept. Signal Process 36(3):287–314
6. Aapo H, Earhunen O (2000) Independent component analysis: algorithms and applications. Neural Netw 13(4–5):411–430
7. Zhang J (2007) Independent component analysis for harmonic detection. J Electr Power Sys Autom 19(1):74–78
8. Ji Y, Sun Y, Li J (2009) Harmonic detection based on independent component analysis. Power Sys Prot Control 37(8):14–18

Chapter 64
Health Diagnosis Strategy for Coal Mine Underground Cable Line Based on Neural Network and Expert System

Caixia Gao, Yunping Wen, Fuzhong Wang and Chunbo Wang

Abstract In order to evaluate the operation condition of coal mine underground cable line, and to realize the advanced repair for coal mine underground cable line. Through the analysis of health Condition characteristics of the coal mine underground cable line, health diagnosis system of the combination of the neural network and expert system is designed for coal mine underground cable line. The whole structure design of health diagnosis system for coal mine underground cable line is introduced, Establishments of knowledge database, inference machine and interpreter are represented. The simulation results show that the health diagnosis system can realize high real time and high diagnosis accuracy.

Keywords Coal mine underground cable line · Health diagnosis · Neural network · The expert system

64.1 Introduction

Coal mine underground cable line for the underground coal production equipment to transport the task of electric power, the safe, and reliable operation of the coal mine safety in production is the most fundamental and most important guarantee.

C. Gao (✉) · Y. Wen · F. Wang
Henan Polytechnic University, Jiao zuo 454000, Henan, China
e-mail: gcx@hpu.edu.cn

Y. Wen
e-mail: wenyp@hpu.edu.cn

C. Wang
Wuzhi Electric Power Bureau, Jiao zuo 454950, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_64,

Therefore, the immediate access to reflect the circuit of cable and circuit breaker electrical components health condition characteristic parameters to evaluate the current health condition, predict the future development trends of degradation and the trend of the poor state of equipment timely release early warning message, and state effectively failure mode and the analysis of the reason, this to advance equipment maintenance, prevent faults, ensure that coal mine underground cable line the security and reliability of operation to have the important meaning. At present, for coal mine underground cable line health diagnosis, still stay in artificial test method, not only the workload is too big, but also easy to go wrong. In this paper, the neural network and the expert system, the construction of the combination of the neural network expert system (ANNES) the combination of coal mine underground cable line health diagnosis system [1]. The system to avoid the artificial collect and sort out the hard labor, real-time data, and avoids the person's knowledge and lack of experience and human feelings, the accuracy of the evaluation system is high.

64.2 Health Condition Characteristics of the Coal Mine Underground Cable Line

Coal mine underground cable line composed of underground cable and 6 kV Distribution Units.

64.2.1 Coal Mine Underground Cable Line Health Condition Characteristic

Coal mine underground cables in tunnels, because in the long operation, environment moisture, and corrosion of harmful material, makes its insulation to age, causes its cable to reduce, cable for insulation joint contact undesirable, serious when will occur leakage, break the fault. Cable connector contacts undesirable electrical characteristics for cable joint fever.

64.2.2 6 kV Power Distribution Units Health Condition Characteristic

6 kV power distribution units main equipment is vacuum circuit breaker, vacuum circuit breaker as the head of the vacuum between touch insulation and arcing medium, mainly by load flow arcing equipment and operation system of two parts. The main health type and warning signs are as follows [2]:

(1) vacuum arcing chamber leak or chronic got a flat

Vacuum arcing chamber leak or chronic leak will cause the vacuum arcing chamber reduced. If vacuum arcing chamber is reduced to a certain value the vacuum degree, makes its open circuit ability and reduces the levels of insulation, serious which will lead to circuit breaker failure from breaking, have great power supply accident. The electrical characteristics of lower vacuum.

(2) vacuum circuit breaker's operation structure action abnormalities

The vacuum breaker mechanical operation organization in the operation process, as the running time increases, break-brake electromagnet coil will appear controling circuit break, short circuit, and iron core sports card acerbity, points off the abnormal rate mechanism motion loose parts, or fall off, dc motor aging anomalies, their unusual characteristics are: The points off the current waveform coil abnormal; Vacuum circuit breaker points, closing speed, abnormal temperature control the abnormal.

64.3 Health Diagnosis System Design for Coal Mine Underground Cable Line Based on Neural Network and Expert System

64.3.1 The Whole Structure Design of System

For coal mine underground cable line health condition characteristic parameters of complexity and diversity and fuzzing features, this article will neural network and the expert system, the construction of the fusion based on the neural network expert system of coal mine underground cable line health diagnosis system. This can make full use of the neural network and the strength of the expert system, will simply pure logic reasoning to expert system to complete, and fault mode matching through the neural network is to finish, parallel computation, can improve the fault diagnosis system efficiency, and also make the diagnosis of building the system easier. Coal mine underground cable line health diagnosis expert system (the ANNES) as shown in Fig. 64.1 shows, using a serial connection method, the characteristics of experts in the field for the reliance of the relatively small, system set up concise [3]. The system consists of knowledge storage system, learning system and inference machine, interpreter and man–machine interface components [4].

64.3.2 The Establishment of Knowledge Database

Knowledge which ANNES used is divided into knowledge structure (including system operation parameters and process parameters), health condition and health diagnosis knowledge sample knowledge three knowledge system [5]. For coal

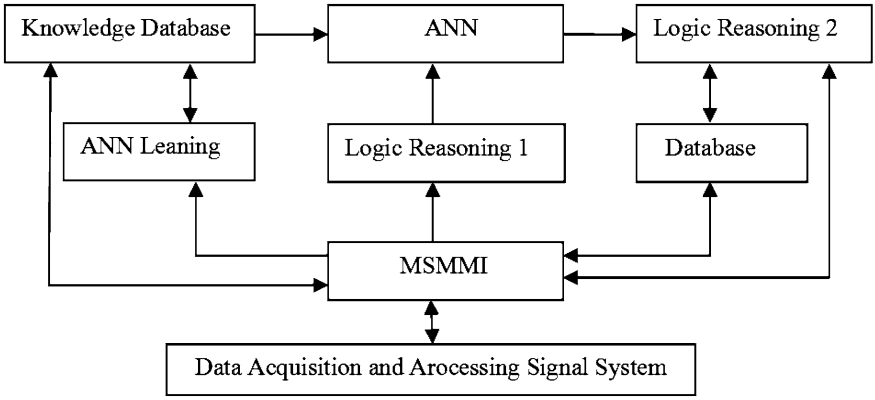


Fig. 64.1 Serial neural network health diagnosis expert system structure diagram

Table 64.1 Cable joint health condition sample knowledge

Type code	Health condition types	Abnormal condition characteristic samples
L1	Cable connector contact undesirable	Cable joint temperature of 80° above
L2	Cable joint moderate contact undesirable	Cable joint temperature of 70–80° above
L3	Cable joint mild contact undesirable	Cable joint temperature of 60–70° above

Table 64.2 Circuit breaker health condition characteristic signal

Characteristic signal code	Characteristic signal name	Normal range
S1	Points off coil current RMS (A)	1.8–2.2
S2	Points off coil current time (ms)	58–66
S3	Points off both ends coil voltage (V)	200–220
S4	Move the contactor average speed (m/s)	4.6–4.8
S5	(°C temperature control box)	40

mine underground cable line health diagnosis system, the structure of knowledge mainly includes: cable insulation resistance, cable joint temperature, the vacuum arcing chamber, coil current waveform, closing points closing speed, control points, the parameters such as temperature; Cable health sample knowledge as is shown in Table 64.1, circuit breaker health condition characteristic signal as shown in Table 64. 2, circuit breaker health knowledge such as sample Table 64.3 shows; Health diagnosis knowledge is through the neural network learning system using the method of learning to be.

Table 64.3 Circuit breaker health sample knowledge

Type code	Health Condition types	Abnormal condition characteristic samples				
		S1	S2	S3	S4	S5
D1	No	Normal	Normal	Normal	Normal	Normal
D2	The iron core is beginning a card closing acerbity	Big	Long	Normal	Low	Normal
D3	Operators have card of color	Big	Long	Normal	Normal	Normal
D4	Auxiliary switch action contact undesirable	Big	Long	Normal	Normal	High
D5	Iron core, winding aging	Big	Long	Normal	Low	High
D6	Low power operation	Small	Short	Low	Small	Low

64.3.3 The Establishment of Learning System

In the study, the system ANNES mainly refers to the neural network system with knowledge engineers build the health condition of neural network training samples, and take the health diagnosis knowledge to neural network weights is stored in the form of neural network of each neuron, namely the learning system establishment is a neural network model.

Because RBF networks in approximation ability, classification ability and learning speed and so on, are better than the BP network. Therefore, this paper using RBF neural network learning algorithm of coal mine underground cable line health diagnosis [6].

64.3.4 The Establishment of Inference Machine

In health diagnosis system, reasoning is according to the input cable and the state of the circuit breaker feature information get health condition types, evaluation at present to health, according to the development trend forecast fault type process. In ANNES, inference machine mainly by logic 1 and neural network of two parts, the reasoning process as follows: input sampling data; Logic 1 sampled data processing, get the neural network's input; Neural network to get input parallel computing network output. Because neural network of reasoning is the parallel competition mechanism to complete to health state reasoning, it is a kind of numerical calculation process, so fast, and avoid the reasoning of expert system matching conflict.

The establishment of the 2.5 interpreter

In ANNES, through the logical reasoning in the diagnosis of 2 interpret the results. In logic reasoning 2, according to different diagnosis object summarizes the diagnosed the object logical inference rules of the neural network, the diagnosis, treatment with output in the health Condition in the form of statements output, in healthy state statements are given in the prediction of the health, serious degree and the treatment method. That is to say, this paper is through the logic reasoning 2 and health Condition report for the diagnosis of neural network to interpret the results.

Table 64.4 The training sample

Health condition types	Input node (characteristic signal)					Output node (health condition types)		
	S1	S2	S3	S4	S5	O1	O2	O3
D1	2.21	59	225	4.8	10	0	0	0
D2	2.25	68	222	4.46	15	0	0	1
D3	2.26	69	219	4.8	10	0	1	0
D4	2.27	68	218	4.70	28	0	1	1
D5	2.28	65	220	4.38	45	1	0	0
D6	1.6	40	160	2.5	12	1	0	1

64.4 The Early Warning Model Realize of Coal Mine Underground Cable Line Based on Neural Network and Expert System

In this paper the design based on neural network expert system of coal mine underground cable line health diagnosis system by cable health diagnosis module, breaker health diagnosis module two diagnosis module, this system the parallel combination, which can improve the diagnostic speed and the accuracy, and can also diagnosed DuoZhong health. Here in circuit breaker health diagnosis module as an example, this paper based on neural network expert system of coal mine underground cable line early warning model of the working process.

64.4.1 The Determination of Input and Output Vector

In a circuit breaker state monitor, winning five group states, such as Table 64.4 sample data shown (Table 64.4 list only 1 group), have the normal state, operational power supply is too low, close brake iron core start a card acerbity, operating mechanism has a card color, core, and winding aging, auxiliary switch action contact undesirable the six different state data. Due to the circuit breaker is according to the points off coil current RMS, points off coil current time, points off on both ends of the coil voltage, move the contactor average speed, control the temperature rise this five characteristic signal to determine the value of the circuit breaker health state, therefore, based on the five characteristics signal as the input vector. Output vector using O1, O2, O3 as the output node, the binary code said of the process of tie breaker closed six health condition types: normal data D1 (000), closing at the start of the iron core is card acerbity D2 (001), operating mechanism has a card color D3 (010), auxiliary switch contact undesirable action D4 (011), iron core, winding aging D5 (100), low power operation D6 (101).

In order to improve the neural network learning speed, and that data in the RBF neural network can handle the range [RBF neural network can handle range

Table 64.5 ANNES diagnosis result

Input node (characteristic signal)					Diagnosis result			Health
S1	S2	S3	S4	S5	O1	O2	O3	
2	62	219	4.6	12	0	0	0	D1
2.26	69	221	4.45	15	0	0	1	D2
2.25	68	219	4.9	10	0	1	0	D3
2.28	67	221	4.8	29	0	1	1	D4
2.27	66	220	4.3	45	1	0	0	D5
1.6	40	160	2.5	12	1	0	1	D6

for (1, 1)], this article on between input neural network is normalized data, the input vector transformation sample to [0, 1], between concrete realization method for:

$$y = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{64.1}$$

Type (64.1), $y_{\min}$ $y_{\max}$, y_i, respectively, the same features input mode said the amount of data and the first I a minimum, maximum; Y—for the normalized the value after the input and output (network mode).

Network input/output vector, after the confirmation of the MATLAB newrbe create RBF networks, and the sample data for 5 groups as the training sample RBF network for training.

64.4.2 Health Condition Diagnosis Result Analysis

The six groups circuit breaker health Condition of the validation sample build ANNES system, and the simulation results such as shown in Table 64.5. By the diagnosis of Table 64.5 results that the RBF network based on the expert system can accurately identify and the training sample with a similar health law health condition, and high accuracy.

64.5 Conclusion

In view of the present situation of repairing difficultly of coal mine underground cable line and mostly repairing after the fault, fault prediction method for coal mine underground cable line is proposed using the neural network and expert system, and health diagnosis system is designed based on neural network and expert system. The system has the following features:

(1) It can realize healthy condition diagnosis rapidly of signals obtained from online monitoring device. The new confirmed health condition will be added

to the knowledge base through some operating procedures, solves the problem of the "bottleneck" of the expert system on a certain extent;

(2) Neural network based-reasoning is the parallel competition mechanism, fault reasoning is completed through the neural network parallel computing of input vector, it is a kind of numerical calculation process, so fast, and avoids the reasoning of expert system matching conflict.

(3) The health diagnosis system based on neural network and expert system has high real time and high diagnosis accuracy, and can predict multiple health condition type at the same time.

References

1. Liu Z, Zhong W, Deng Y et al (2010) Electric railway substation diagnosis with model-based method. Proc CSEE 30(34):36–41 (in Chinese)
2. Wang X, Rong M, Wu Y et al (2007) Method of quick fault diagnosis and new knowledge obtainment for high voltage circuit breaker expert system. Proc CSEE 27(3):95–99 (in Chinese)
3. Xie S (2001) The study of traction substation fault analysis based on expert system. Southwest Jiaotong University, Chengdu (in Chinese)
4. He Z, Zheng W (2008) Deformation prediction of deep foundation p it based on BP neural net work. J South China Univ Technol Nat Sci Ed 36(10):92–96
5. Hu S, Li Y (2009) Prediction of coalmine accidents based on neural net works. J Saf Sci Technol 5(4):1062–1109
6. Lei Y, Yang Z (2010) Application of neural network to fault diagnosis of hoist. J Saf Sci Technol 38(2):67–72

Chapter 65
The Out-of-Step Protection Strategy for the Vertical-Moving Permanent Magnet Linear Synchronous Motor

Qiaolian Wang, Caixia Gao, Fuzhong Wang and Chunbo Wang

Abstract In order to prevent the safety accidents causing by the out-of-step of the vertical-moving Permanent Magnet Linear Synchronous Motor (PMLSM), the out-of-step protection scheme of the combination of the mechanical braking and dynamic braking has been established according to the common features of various out-of-step of PMLSM, on this basis, the mathematical model of dynamic braking of the vertical-moving PMLSM is established, the dynamic braking characteristics of the motor are obtained, and at the same time, the relationship equation between the braking force and the reactance and resistance of armature circuit is established, and the dynamic braking circuit of the vertical-moving PMLSM is designed. The experiment shows that the established out-of-step scheme of PMLSM is practicable, as well as the established mathematical model and theory, can effect prevent the safety accidents causing by the out-of-step of the vertical-moving PMLSM, and thus guaranteed the motor's safety.

Keywords Permanent magnet linear synchronous motor · Out-of-step protection · The dynamic braking · The relationship equation of dynamic braking

Q. Wang (✉)
Guangzhou Institute of Railway Technology,
Guang zhou 510430, Guangdong, China
e-mail: gcx81@126.com

C. Gao · F. Wang
Henan Polytechnic University,
Jiao zuo 454000, Henan, China
e-mail: gcx@ hpu.edu.cn

C. Wang
Jiaozuo BRAKE Limited Company,
Jiao zuo 454000, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_65,

65.1 Introduction

The vertical-moving Permanent Magnet Linear Synchronous Motor inevitably encounters various kinds of turbulences in operation, such as the change of Air-gap magnetic field resulted by thrust disturbance and ends effect, and the fluctuations of power supply voltage leaded by line contact undesirable, along with mechanical wear and load change [1, 2], etc. All of these disturbances would be able to lead the out-of-step of the vertical-moving PMLSM. If the motor is out-of-step in operation, there is no reliable safeguard procedures, the mover and the cage are dropped down quickly by gravity, which causes severe accidents. To take effective and accurate detection means to prevent the safety accidents causing by the out-of-step of the vertical-moving PMLSM have always been the problem that people want to resolve. The out-of-step prevention strategies of the vertical-moving PMLSM are established according to the electrical characteristics after the out-of-step of the vertical-moving PMLSM.

65.2 The Out-of-Step Protection Scheme of the Permanent Magnet Linear Synchronous Motor

The kinds of out-of-step of the PMLSM are as follows [1, 2]: asynchronism under-excitation out-of-step, with excitation out-of-step and power-off out-of-step and so on. Here are their common characteristics: when the PMLSM is out-of-step, its thrust would reduce greatly or the average value is down to zero, the running speed of the motor would fall down with the criterion of free fallers acceleration downward motion, and motor's power angle would be beyond its stable working area. Therefore,on the basis of the decline extent of motor's thrust and its change of running speed, along with the size of the motor power angle and power supply voltage, it is possible to estimate whether the motor is out-of-step or not. When out-of-step happened, we can take protection measures to prevent safety accidents' taking place [3]. the axiom is shown as the Fig. 65.1: when the controller verdicts that the motor's thrust is less than Fm, the downward running speed is higher than Vm, and the motor voltage is near to zero or the motor's power angle is bigger than θ m, it will start mechanical braking and dynamic braking at the same time. Mechanical braking is realized by using hydraulic device. When the motor's thrust, running speed or the motor's power voltage is accord with motor out-of-step conditions, the controller will cut the hydraulic equipment's power supply. The hydraulic equipment make the linear motor brake through a mechanical spring. The dynamic braking is realized by out-of-step power characteristics of the vertical-moving PMLSM. And the dynamic braking mode is mainly introduced.

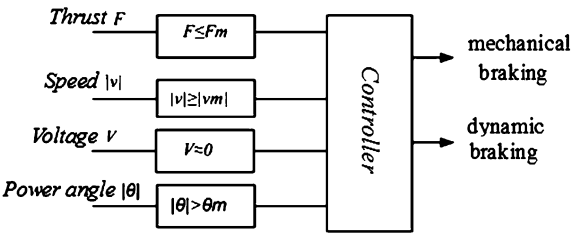


Fig. 65.1 The out-of-step protection block diagram of PMLSM

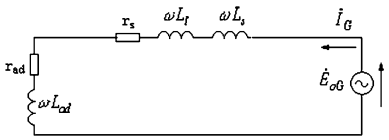


Fig. 65.2 Equivalent circuit of dynamic braking

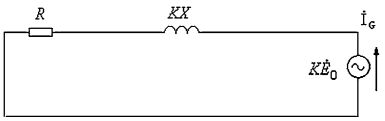


Fig. 65.3 Equivalent circuit of dynamic braking (the fundamental frequency is normal frequency)

65.3 The Dynamic Braking Protection Strategy for the Vertical-Moving Permanent Magnet Linear Synchronous Motor

65.3.1 The Dynamic Braking Mathematical Model of the PMLSM

In the operation process of the vertical-moving PMLSM, when met with disturbance, the out-of- step occurs. Then the mover's permanent magnet together with the load of cage conductor will fall down because of the strength of gravity. This produces a traveling magnetic field between the mover and the stator which moves along with the mover. This magnetic field cuts stator winding thus producing three-phase induction electromotive force on it. If we short circuit the three-phase armature winding or the added impedance form a closed loop circuit at this time, the permanent magnet linear synchronous motor will be in a moving state as a loaded generator, and meanwhile produced induction current in the closed three-phase armature winding of the stators, the magnetic field produced by the current and the traveling magnetic field caused by the mover permanent magnet will cause interaction and produces a braking force which is opposite to the falling direction of the mover, when the upward resultant external force balances the downward resultant external force, the mover and the cage will be lowered at a pretty low constant speed, so we can limit the falling speed of the mover to a certain low level, and limit the falling speed of hoist to the safe range. The corresponding

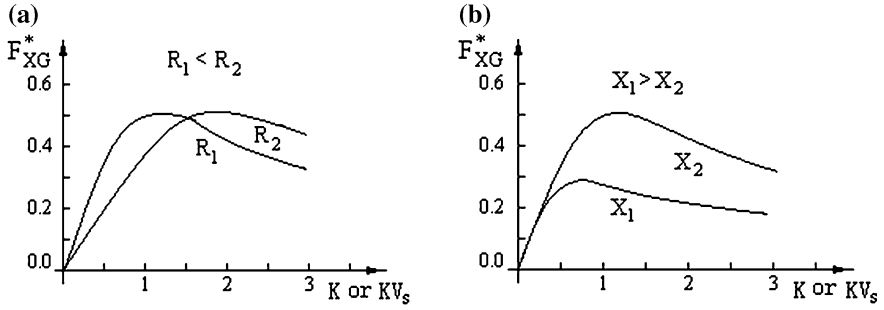


Fig. 65.4 The characteristic of dynamic braking

equivalent circuit If we short circuit the three-phase armature winding or the added impedance form a closed loop circuit is like what shows in Fig. 65.2. In Fig. 65.2, Ll and Ls respectively represent leakage inductance and armature inductance of the permanent magnet linear synchronous motor. Rad and Lad are the outside series resistance and inductor, ÌG represents the current of the dynamic braking, È0G represents excitation electric potential of the dynamic braking. If we suppose that R $= r_s + r_{\text{ad}}, X = \omega(L_l + L_s + L_{\text{ad}}), \text{K} = \frac{f}{f_e}$, then the equivalent circuit shows in Fig. 65.2 can be simplified to what shows in Fig. 65.3. f is the generating braking frequency of the generator, f_e is the nominal frequency of the generator.

Suppose Us = 0, we can get equation of braking force when the dynamic braking happens as Eq. (65.1).

$$\text{F}_{xG} = -\frac{3KE_0^2R}{v_s(R^2 + K^2X^2)} \tag{65.1}$$

In Eq. (65.3) "−" represents that the braking force and the thrust are in the opposite direction. If we calculate the differential coefficient of Eq. (65.4), we can get the frequency (or speed) when the max braking force appears.

$$\frac{\text{dF}_{\text{xG}}}{\text{d}K} = 0 \tag{65.2}$$

From Eq. (65.2), we can get the calculation method of kmax the proportion between max braking rate f and f_e.

$$K_{\text{max}} = \frac{R}{X} \tag{65.3}$$

When combining Eqs. (65.3) and (65.5) together, we can get the equation of max braking force when the dynamic braking happens as Eq. (65.4).

$$F_{G\text{max}} = -\frac{3E_0}{2v_sX} \tag{65.4}$$

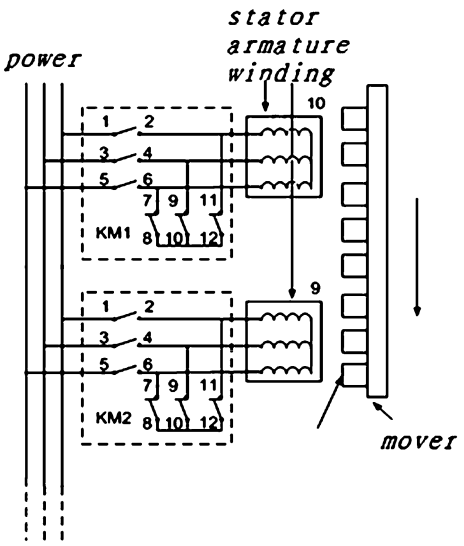


Fig. 65.5 The circuit of power supply of PMLSM

In order to analyze easily, suppose both R and X are Fig. 65.5 that we can get arbitrary value, and at the same time, when we take nominal frequency, parameters in Eq. (65.6) are all reference value, when combine Eqs. (65.7), (65.4) together, we can get the Eqs. (65.5), (65.6) as dynamic braking force that represents by normalized value.

$$F_{\mathrm{xG}}^{*} = -\frac{KR^{*}}{R^{*2} + K^{2}X^{*2}} \tag{65.5}$$

$$F_{G}\max = -\frac{1}{2}\frac{1}{X^{*}} \tag{65.6}$$

Equation 65.5 is the characteristic of the vertical-moving PMLSM's dynamic braking. From the Eq. (65.5), we can know that the characteristic of Vertical run permanent magnet linear synchronous motor's dynamic braking varies with R*and X*. From Eq. (65.6), we can know that the max braking force has nothing to do with the armature circuit resistance. From Eq. (65.5) and Fig. 65.4a, we can know the position of the max braking force has something to do with the return circuit resistance. From Eq. (65.6) and Fig. 65.4b, we can know that the max braking force and return circuit reactance concern into inverse ratio.

65.3.2 The Design of Dynamic Braking Circuit

Energy consumption braking is to cut the three-phase stator armature winding of permanent magnet linear synchronous motor from the power supply, and then short circuit the three-phase armature winding or make the added impedance form a closed

loop circuit. Aim at the segmental hoisting system drived by the permanent magnet linear synchronous motor, in order to realize the dynamic braking, we have designed the power supply line of permanent magnet linear synchronous motor segmented as shown in Fig. 65.5. The basic principle is: according to the position of mover detected by the position sensor, we supply power to the stator's armature winding which the mover will coupling in advanced, for example in the process of down weight, if the current the armature winding stator of mover coupling is NO.10 and NO.9, then the mover will coupling with NO.8 stator the armature, so we should supply power to NO.8 stator the armature. Before the NO.8 the armature winding stator is filled with current, to cut the power supply of NO.11 stator the armature and short circuit the windings. When the coupling between NO.9 and NO.8 stator armature winding and mover is finished, first, we should cut the power supply of NO.10 stator armature winding, and short circuit the windings, and then give power supply to NO.7 stator armature winding. The power supply of armature winding of the stator is finished by the move closed contacts of contactor NO.1 to NO.10, but the dynamic braking is completed by move open contacts of NO.1–10 contactor. When the controller judges that the motor thrust, operation speed, motor power angle accords with motor out-of-step conditions, the power supply of 1–10 co contactors are cut off automatically, or when the total power supply is lost, move closed contacts of 1–10 contactor opened automatically, move closed contacts close automatically, enter into the process of dynamic braking automatically.

On the other hand, in order to prevent problems caused by controller that may cause the out-of-step protection failed, we designed the automatic closing of unworked stator armature winding in the design of power supply line (as shown in Fig. 65.5). No matter what fault has caused the irregular down movement, the automatically close loop of stator armature winding can always realize dynamic braking [4, 5].

65.4 The Out-of-Step Protection Experimental Analysis Based on Dynamic Braking

Permanent magnet linear synchronous motors (PMLSM) load 600 kg (not including mover' weight 270 kg) by the speed of 0.312 m/s lifting object. When it gets the height of 2.5 m, we suddenly cut off the electricity. Then by using the velocity sensor and weighing sensor, send the change data on speed and ascension load when PMLSM get out-of-step and the protected measures put to the whole process to the computer, then work it out.

Figures 65.6 and 65.7 are the motor speed and ascension load curves in the PMLSM out-of-step protection experimental. According to Fig. 65.6 we can see, after power off, the rate of linear electric motors' mover and lift car change immediately from rised speed to descend speed -0.13 m/s, and the motor is in state of power generating. Then after about 100 ms (time of protection device is put into), contactor close three-phase stator armature winding of PMLSM, dynamic

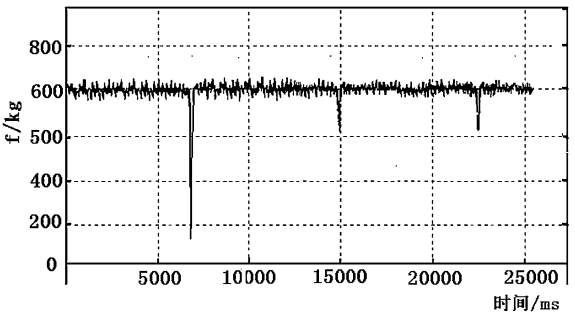


Fig. 65.6 The velocity of motor in desynchronism in operation without power of PMLSM

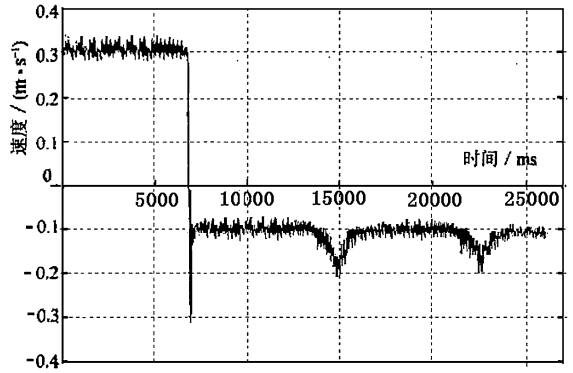


Fig. 65.7 The various hoisting load of motor desynchronism operation without power of PMLSM

braking put into operation, at the same time linear electric motors' descend speed reduce to about −0.01 m/s. In Fig. 65.6 the speed fluctuation is caused by the speed change which caused by the change of electromagnetic force when 0mover going through stator armature winding segmented joints. After dynamic braking, PMLSM descend by the speed of 0.1 m/s, biggest wave as 0.2 m/s, which guarantee the security of PMLSM hoisting system.

Figure 65.7 is the curve of ascension load after PMLSM's power off, which work on the weighing sensor. We can see that when the PMLSM power off, the tension disappear suddenly, the tension stress on the weighing sensor proximate zero. When descend speed reduce to −0.01 m/s, the tension stress on weighing sensor equals to loading weight and changes with the motor' descend speed.

65.5 Conclusion

(1) According to the common characteristics of the vertical-moving PMLSM's under asynchronism under-excitation out-of-step, with excitation out-of-step and power-off out-of-step, established the out-of-step protection scheme of

PMLSM which is input with motor thrust, motor speed, the motor power supply voltage, power angle, and output with the mechanical braking and dynamic braking. The experiment shows that this protection strategy can operate quickly after the out-of-step of the vertical-moving PMLSM, so that it can prevent the accident brought by the out-of-step and at the same time ensures the safe and steady operation of PMLSM driven lifting system.

(2) The mathematical model of dynamic braking of the vertical-moving PMLSM is established, the dynamic braking characteristics of the motor are obtained, and at the same time, the relationship equation between the braking force and the reactance and resistance of armature circuit is established, and a theoretical basis for the design of dynamic braking control circuit of vertical-moving PMLSM is provided.

(3) The dynamic braking circuit is designed, which can ensure the dynamic braking operation no matter the controlling equipment is normal or not, and a reference for the design of the motor power supply circuit and the protection circuit.

References

1. Wang F, Yuan S, Jing P (2010) The out of step preventing strategy of permanent magnet linear synchronous motor moving for vertical movement. J China Coal Soc 35(4):696–700
2. Jiao L, Yuan S (2000) Effect of constant current source supply on the electromagnetic force of permanent magnet linear synchronous motor for vertical movement. J China Coal Soc 25(4):420–422
3. Jiao L, Yuan S (2002) Study on operating characteristics of permanent magnet linear synchronous motor for vertical movement. Proc CSEE 22(4):37–40
4. Wang F, Wang X, Jiao L et al (2001) Research on control strategy of electro—magnetic force angle and the maximum of thrust of permanent magnet linear synchronous motor for vertical movement. J China Coal Soc 26(3):307–312
5. Wang F (2010) Strategies of out-of-step fault symptom acquisition and prevention control for permanent magnet linear motor. Dissertation, China University of Mining and Technology, Beijing

Part V
Computer Network and Security

Chapter 66
Secure Sensitive Data Transmission in Wireless Sensor Network Based on Compressive Sensing

Jiping Xiong, Lifeng Xuan and Tao Huang

Abstract A novel sensitive data secure transmission method in wireless sensor network based on compressive sensing technology is presented. In the proposed method, the sensor nodes encrypt sensitive data using only linear operations such as simple addition and multiply instead of traditional symmetric encryption algorithm, and base station can decrypt sensitive data even when the wireless channel is lossy. Theory analysis and simulation result demonstrate the effectiveness of our method in practical deployment.

Keywords Compressive sensing · Wireless sensor network · Secure transmission · Lossy channel

66.1 Introduction

Single-hop Wireless Sensor Network (WSN) [1] is consists of massive tiny sensor nodes distributing in target region and powerful data collecting centers (or called base station, BS). The sensor nodes send the measured data collected by sensor units to BS using wireless channels. WSN is widely used not only in public

J. Xiong (✉) · L. Xuan · T. Huang
College of Mathematics, Physics and Information Engineering,
Zhejiang Normal University, Jinhua 321004, Zhejiang, China
e-mail: xjping@zjnu.cn

L. Xuan
e-mail: xuan.lifeng@zjnu.cn

T. Huang
e-mail: tao.huang@zjnu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_66,

domains but also in security-required domains. In practice, data collected by sensor nodes are usually sensitive and require to be kept secret. Symmetric Encryption technique is one of the most commonly used methods at present [2], however, it is complex in the sense of computing. Since the operation systems installed in sensors are always simple with low processing ability and limited battery energy, the lifetime of wireless sensor network will decrease rapidly if such encryption technique was adopted.

Information hiding technology is another reliable method to guarantee data secret in telecommunication. Xiao et al. [3] first introduced information hiding technology to wireless sensor network. They embedded sensitive information to normal information using LSB algorithm which can get rid of being detected by a malicious client. However, this approach has two drawbacks: (1) sensitive information is still needs to be encrypted by using symmetric encryption method before applying LSB algorithm; (2) This approach will fail to retrieve any sensitive information when data losing occurs during transition. Thus, it is not quite practical to apply that kind of information hiding technology into a WSN.

Compressive sensing (CS) is a newly developed technology in signal processing field [4–9]. It can reconstruct the origin signal (which is sparse signals or sparse in certain transform-domain) accurately with less measurement. The theory of compressive sensing was first introduced by Candès et al. [4] and Donoho [5] and has already made a remarkable effect and is already being used in massive areas [10–12], such as radar, cognitive radio network, MRI and image et al. Authors can get more information through [11]. There are some compressive sensing related researches on large-scale distributed data aggregation in multi-hop WSN which mainly concentrates on data compressive measuring for multi-nodes to realize data aggregation [14]. But there is less CS research in single-hop WSN that focus on secure data transmission.

The paper is organized as follows. In Sect. 66.2, we give out the basic background of compressive sensing theory which is used throughout our paper. Prototype implementation details will be given out in Sect. 66.3. In Sect. 66.4, simulation is conducted. Finally, we draw some conclusions and sketch future work.

66.2 Theory Of Compressive Sensing

The theory of compressive sensing was first introduced by Candès [4] and Donoho [5]. If an unknown signal $x \in R^n$ with length N is k-sparse or k-sparse under the transform of a certain orthonormal basis ψ, such as wavelet basis, Curvelet basis, discrete cosine basis or Fourier basis et al., then x can be expressed as the combination of k non-zero coefficients of $\{\psi_i\}_{i=1}^{k}$, such as:

$$x = \sum_{i=1}^{k} \alpha_i \psi_i = \psi \alpha \tag{66.1}$$

where $\alpha = \psi^T x$ is the transform coefficient with only k non-zero numbers.

If we project signal x to an incoherent measurement basis $\Phi : M \times N, M \ll N$ with the sparsifying basis Ψ, then we can get measurement set $y : M \times 1$. Compressive sensing proves that we can accurately reconstruct the origin signal x from the small number of measurement through solving the below formula

$$\min_x \left\| \Psi^T x \right\|_1, \quad s.t. \quad y = \Phi x \tag{66.2}$$

In CS theory, the measurement matrix Φ needs to satisfy the famous restricted isometry property (RIP) condition:

$$(1 - \delta_K)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 - \delta_K)\|x\|_2^2. \tag{66.3}$$

where $0 < \delta_K < 1$.

Many matrixes have already been found that satisfy the RIP condition, Gaussian matrices is one of them which has been proved to be simple and efficient. In this paper, we also select Gaussian matrices as the encoding and decoding matrices.

Candès and Tao solve the signal recovery problem in the lossy channel using sparse reconstruction technology in compressive sensing [7]. Suppose that we get corrupted measurement $y = Af + e$, where A is a m by n matrix (here, we assume throughout this paper that $m > n$), e is an arbitrary and unknown signal of errors and $f \in R^n$, then we can exactly recover the input signal f from y. This theory motivates our novel design of securing sensitive data transmission in WSN.

66.3 Propose Method

In typical WSN, each tiny sensor nodes connects with base station (BS) through wireless channel. We assume that each sensor nodes has multi-sensing modules which can collect both sensitive and regular data. Before deployment of WSN, each sensor nodes shares a same secret key with BS.

66.3.1 Topological Structure of WSN

Figure 66. 1 shows the network structure of this paper which is a typical single-hop WSN. The deployed sensor nodes (N1, N2 and N3) are with multi-sensing modules. The BS has a high processing ability and abundant resource. The

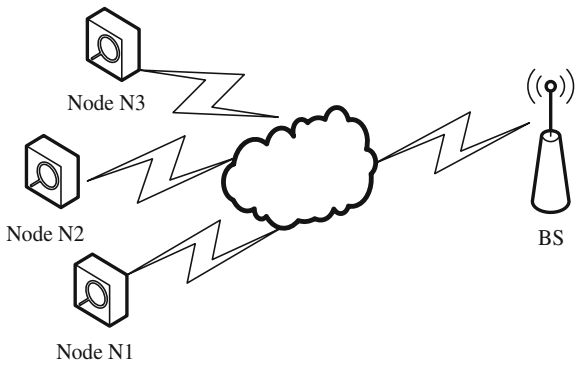


Fig. 66.1 The topological structure graph of the WSN

communication channel between nodes and BS is wireless with packet drop probability. Before deploying the network, the secret key k_i shared by BS and each node is stored in the flash memory area of each sensor node in case of data missing caused by power off.

Figure 66.2 describes the general encoding and decoding steps of our proposed method.

66.3.2 Encoding Phase

At the initial stage of communication, sensor node Ni generates and stores a full rank Gauss random matrix A, where $A \in \mathrm{R}^{m \times n}$ with $m > n$, using the secret Ki shared with BS as the seed. After initial stage, sensor node follows below steps to do the encoding.

Step1: Sensor node Ni collects sensitive data S ($\in$R) and regular data r ($\in$R) and keeps putting them into individual buffers. When sensitive data buffer is full or has enough data, node Ni forms sensitive vector S_n with length n from the buffer and generates measurement vector $y_m = A^{m \times n} \times S_n$.

Step 2: Sensor node Ni constructs sparse signal r_m which contains zeros and regular data from regular data buffer. In r_m, the number of zero data should be larger than the number of regular data. More formally, the number of regular data in r_m is limited by

$$||r_m||_{l_0} := |\{i : r_i \neq 0\}| \leq \rho \cdot m \tag{66.4}$$

where $\rho > 0$ is a small trade-off positive number which can affect the successful decoding probability.

Step 3: Sensor node Ni generates transmission vector fm as follow,

$$f_m = y_m + r_m = A^{m \times n} \times S_n + r_m \tag{66.5}$$

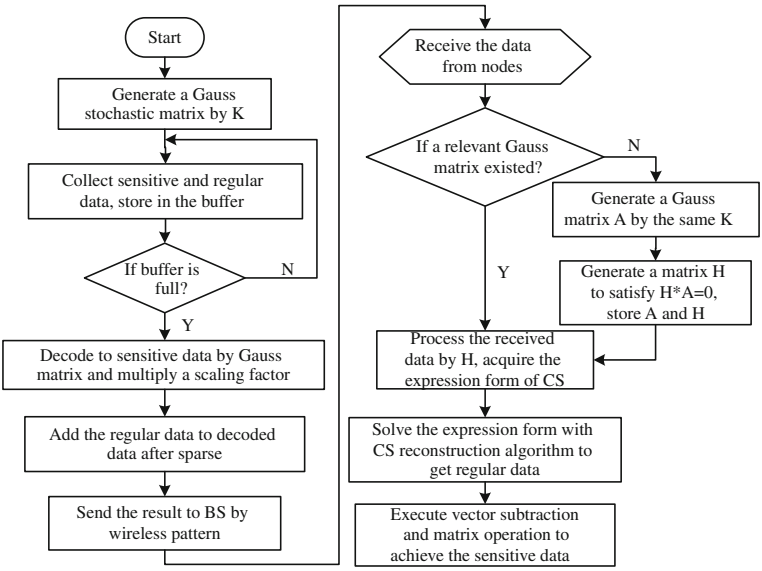


Fig. 66.2 The general encoding and decoding flowchart

Then Ni transmits the vector fm to BS through WSN's wireless channel. As we can see from those steps, the sensitive data is embedded into regular data only using addition and multiply operations which are very simple and makes our proposal appealing in practical deployment.

66.3.3 Decoding Phase

After receiving fm from sensor node Ni, BS does the following steps.

Step 1: BS checks whether the matrix $A_{m\times n}$ shared with Ni does exist. If the matrix does not exist, BS generates and stores it using the same shared key Ki.

Step 2: BS calculates and stores sensing matrix $H_{k\times m}$ $(k < m)$ that satisfies the following,

$$\mathrm{H}_{k\times m} \times \mathrm{A}_{m\times n} = 0 \tag{66.6}$$

where $H_{k\times m}$ can be constructed using the kernels in the range of matrix A in R^m [7].

Step 3: BS applies $H_{k\times m}$ to received data vector fm and get a new vector y' $(\in \mathrm{R}^k)$ as follows,

$$y' = \mathrm{H}_{k\times m} \times \mathrm{f}_m \tag{66.7}$$

Since $\mathrm{f}_m = y_m + r_m$, y' can be expressed as follows,

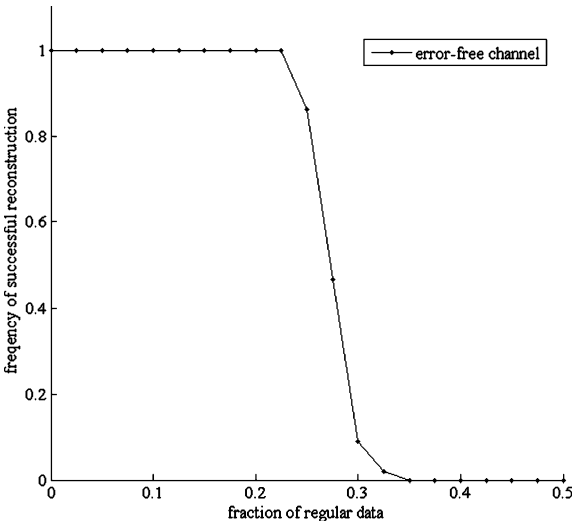


Fig. 66.3 Successfully reconstruction frequency under error-free channel for $m{:}n = 512{:}128$ and $k = 300$

$$\begin{aligned} y' &= H_{k\times m} \times (y_m + r_m) \\ &= H_{k\times m} \times (A_{m\times n} \times S_n) + H_{k\times m} \times r_m \\ &= 0 \times S_n + H_{k\times m} \times r_m \\ &= H_{k\times m} \times r_m \end{aligned} \tag{66.8}$$

Now, as for BS, y' is a known signal, r_m is an unknown sparse signal and $H_{k\times m}$ satisfies the Restricted Isometry Property (RIP) [7]. BS can recovery r_m exactly using any current CS reconstruction methods, such as Orthogonal Matching Pursuit (OMP).

Step 4: After obtaining r_m, BS can get sensitive data S_n as follows,

$$S_n = A^{-1}_{m\times n}(y' - r_m) \tag{66.9}$$

66.4 Simulation Results

We evaluate our proposed method of successful reconstruction probability under two scenarios: one with error-free channel, where the wireless network between sensor nodes and BS is ideal, and one with lossy channel, where packets losing or erasing will happen during the data transmission. Reconstruction was deemed successful if the MSE of the reconstructed sensitive vector was on the order of machine precision (10^{-12} in our simulation). We iterate each simulation for 100 times and compute the frequency of successful reconstruction.

Figure 66.3 illustrates the empirical probabilities of successful reconstruction under error-free channel. Horizontal axis is the parameter ρ in Eq. 66.1 which

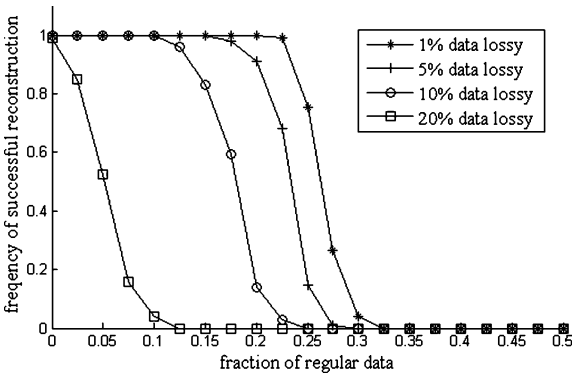


Fig. 66.4 Successfully reconstruction frequency under lossy channel for $m{:}n = 512{:}128$ and $k = 300$

controls the ratio of regular data number in sparse signal rm. From the experiments, the sensitive data can always be successfully reconstructed as ρ is below 20%.

Figure 66.4 gives out the simulation result of our proposal under lossy channel with varying random data drop probability. We can find that sensitive data can still be successfully reconstructed in lossy channel with 10% data drop probability if the sparse ratio ρ is below 10%. It is one of our major advantages that our method can tolerant lossy channel without the usage of any error correcting codes. In practice, we can adjust sparse ratio ρ manually or adaptively to meet different wireless environment.

66.5 Conclusions

The scope of this paper has been to evaluate the introduction of compressive sensing tools in order to securely transmit data on wireless sensor network. Our proposed method is a new alterative approach for securing data transmission. In our method, only simple linear operations on sensor nodes and robust on lossy channel make our proposal very appealing in practical deployment. But more investigations need to be done on the combination of WSN and CS. Our future research work include detail formally security analyzing, energy consumption analyzing and more simulation and practical deployment experiments.

Acknowledgments The project is sponsored by the Projects of Department of Education of Zhejiang Province, No. Y200805325.

References

1. Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. Computer Networks 52:2292–2330
2. Syed TA, Vijay S, Ashay D, Diethelm O (2010) Secure key loss recovery for network broadcast in single-hop wireless sensor networks. Ad Hoc Netw 8:668–679
3. Xiao X, Sun X, Yang L, Chen M (2007) Secure data transmission of wireless sensor network based on information hiding. Fourth annual international conference on mobile and ubiquitous systems. Aug 2007 pp. 1–6
4. Candès E, Romberg J, Tao T (2006) Robust uncertainty principles: EXACT signal reconstruction from highly incomplete frequency information. IEEE Trans Inform Theory 509(2):489–509
5. Donoho DL (2006) Compressed sensing. IEEE Trans Inform Theory 52(4):1289–1306
6. Candès E (2006) Compressive sampling. In: Proceedings of international congress of mathematicians. european mathematical society publishing house, pp 1433–1452
7. Candès E, Tao T (2005) Decoding by linear programming. IEEE Trans. Inform Theory 51:4203–4215
8. Donoho DL (2006) Compressed sensing. IEEE Trans Inform Theory 52(4):1289–1306
9. Candès E, Tao T (2006) Near optimal signal recovery from random projections: universal encoding strategies. IEEE Trans Inform Theory 52:5406–5425
10. Yang Y, Au OC, Fang L, Wen X, Tang W (2009) Perceptual compressive sensing for image signals. ICME pp 89–92
11. Wang Y, Pandharipande A, Polo Y Lamelas, Leus G (2009) Distributed compressive wide-band spectrum sensing. In: Proceedings of the information theory and applications (ITA 09), San Diego (CA), IEEE, February 2009
12. Kang LW, Lu CS (2009) Distributed compressive video sensing. in Proc. of 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing, Taipei, Taiwan, April 2009
13. Luo C, Wu F, Sun J, Chen CW (2009) Compressive data gathering for large-scale wireless sensor network. MobiCom, Beijing
14. Mohamed Yacoab MY (2010) An adaptive traffic aware data aggregation technique for wireless sensor networks. Am J Sci Res (10) pp 64–77

Chapter 67
Study on Promotion of E-Commerce Websites Based on SEO

Xiaoyan Liu and Jian Wang

Abstract Search engine plays a vital role in improving website traffic. Through Search Engine Optimization (SEO) techniques, website can be friendlier to search engine. This paper minutely analyzes PageRank (PR) algorithm and puts forward strategy based on SEO from the aspects of keywords, website development and page links. Keywords optimization is focused on keywords selection, density and distribution. The point of website optimization is title, Meta tag, domain name etc. Optimization in page links is concerned with inbound links and outbound links.

Keywords SEO · Search engine optimization · Website optimization · Keywords optimization

67.1 Overview of SEO

Search Engine Optimization (SEO) has been a popular network marketing mode during recent years. Its primary purpose is to increase the exposure of specific keywords, so as to increase the visibility of the website and then aggrandize the marketable chance. SEO is a technology that is used to optimize the webpage to improve its search engine rankings, thus improve visits of the websites and promote marketing capacity or publicity capacity by understanding some technology,

X. Liu (✉)
School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan, China
e-mail: xyanliu@hpu.edu.cn

J. Wang
School of Safety Science and Engineering, Henan Polytechnic University, Jiaozuo, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_67,

such as how each search engine captures internet webpage, how to index and how to determine the rankings of search results for some specific keywords and so on. By using reasonable means indexed easily by search engines, it makes all the basic elements of the website be appropriate for the search principles of search engines and more user-friendly. Search engine can help users find what they want and reach where they want to go by using SEO technique [1]. For this reason, website builders rank their brain for the key to the problem how to make their own websites be included by major search engines, and then get a higher ranking.

67.2 Strategy of E-Commerce Websites Promotion Based on SEO

When users search information with keywords by using a search engine, then the search engine seeks from database. If search engine finds the websites in line with users' requirements, it then uses a special algorithm. Usually basing on the matching extent of the keywords in websites, placement/frequency, quality of link and so on, figure out correlation and ranking level of each webpage, then return these webpage links back to users in sequence according to the level of correlation. Search engine contains four parts, including web spider, index module, information retrieval and user interfaces. Web spider feature can determine the pages the search engine first crawls, only crawls or do not crawl. Link analysis is the heart of algorithms. PageRank algorithm is a well-known link analysis algorithm [2]. The algorithm formula is as follow.

$$\mathrm{PR}(I)=(1-D)+D\sum J[\mathrm{PR}(J)/N(J)]. \tag{67.1}$$

I and J mean webpage. PR function shows the importance of webpage and ranges from 0 to 10. Webpage is more important, the PR value is higher. N function shows the number of webpage export link. J is a webpage link to i. D is damping coefficient and its value usually is 0.85. Apparently, authoritative web pages are closely related with the keyword. Web export link, internal link structure and site topology also have impact on the ranking.

Aiming at the principle of search engine optimization, this paper puts forward strategy of e-commerce websites promotion based on SEO and analyzes the problems of websites optimization from following aspects.

67.2.1 Keywords Optimization

Keyword is a mathematical factor of search engine algorithm, which determines the correlation between webpage with specific search. The most relevant pages with a keyword search that search engine algorithm consider will be ranked out. Keyword optimization is the core of search engine optimization [3].

1. Keywords selection
   (1) Use organization name and service name to name. For example, Kodak's online printing service. (2) Do not have too many keywords. It is needs to meet the requirements of search tools. Try to avoid using popular keywords. The number is usually around 5. Not only single words, multi-word phrase is more useful than a single word. Because it is difficult for search engine to search related results using a single word and such conditions will produce too many results. Searcher is not going to browse hundreds of results pages. (3) Take seasonal changes into account. For example, sites of clothing types, should include keywords such as spring clothing or summer clothing. (4) Take dialects, slang and spelling errors into account. (5) Describe the location of products and services. (6) Use keyword analysis tools. Overture search engine provides free keywords analysis tool. Keyword Generator tool can test keyword frequency. Enter a keyword, the tool lists the high frequency phrases associated with that keyword or phrase can be used as reference. The function of Word Tracker software is to view the frequency of your keywords used in other pages, and the number of people use these keywords when searching on the major search engines in the past 24 h.
2. Keywords density
   The proportion of keywords and characters of this page is called keywords density. This is also an important factor of search engine optimization strategy. Usually, there are many words in a webpage. Search engine counts the number of the words in this page, then count the importance of each word by using its own algorithm. Those recurring words or phrases are considered to be more important. Therefore, keywords density is best not to exceed 3% of the number of text or less. Search in the search engine to optimize the keywords, then count the keywords density in the top websites can be done. A lot of specific statistics keyword density tool can be used. Remember not to blindly pile, try to be natural and meet the basic grammar rules. Do not continuously use a keyword in the same line two or more times. Do not use cheat techniques such as keyword stuffing cheat which unreasonably repeat keywords by the method that the search engine pays close attention to the keywords that emerge in text of page, title and HTML Meta tags. This technique is easy to be detected by search engines and punished accordingly [4].
3. Keywords distribution
   The principle of keywords distribution is to place keywords at the right place. First, count the number of needed keywords according to keywords density. Then, consider how to place keywords in a prominent position on the page, including top, left, title text within the first 200 words. Keywords are usually placed in the Title tag, MATA label, title (Headings), the hyperlink text, URL text and text in the top of the page. (1) Title tags are the most important tags in the page. It is part of the page to see first, and it is better to be put in the front. So it is very important to add some keywords in the title. The number should not exceed 20 words. (2) Meta tags contain some hidden information, in which the content of the attribute description can be a bits of sentences to reflect the

theme of website. So add main keywords in the description to a great extent. Keywords in attribute keyword should be separated by commas, and generally should be controlled within five words. (3) Headings. Keywords appear in headings for improving the site's ranking is highly desirable. (4) Hyperlink text. Linked webpage is certainly associated with the content of the text. So it is important to place keywords in hyperlink text. When other sites link to your sites, make use of keywords as link text as much as possible. This will help improve your sites' importance. (5) URL text. URL is a uniform resource locator. Keywords appearing directly in URL will have a huge impact on search engine ranking. This kind of keywords is named as URL text. For search engine optimization, easy domain name is not the most important, and the most important is whether the domain name contains the optimal keywords. (6) The top text of page, bottom text and each beginning of content are keywords or navigation bar. So it is better to contain keywords in such places [5].

67.2.2 Website Optimization

1. Title optimization
   Web page title is a highly condensed summary. Generally speaking, title of the home is the official name of the site, title of the page of web site is the title of the article and title of home column is the title of column. When searching keywords of website, google gives a higher weight to keywords appears in their title tag. So, it needs to be ensured that the site's title tag contains the most important keywords. In other words, focus on the most important keywords to determine the content of the page title. The title of the page should not be too long, usually best to contain 35–40 characters. In practice, the title should not be too short or too long. Too short title is hard to express webpage information completely. Too long title is not conducive to user identification, and increases the difficulty of identifying the core words for search engine. The title of page should summarize the core of page. Search results of search engine are usually title of page, summary information of page and links. To cause user's attention, it is important to have a highly summary of webpage's title. For example, the title of the home page of Dell China website is Dell China-computer, notebook PC, desktop computer, printer, workstation, server, memory, electronics, accessory and so on. Dell's home page, not only covers the most important company information, but also the company's main products. These are the core keywords. When searchers use keywords such as notebook PC, desktop computer to search by google, Dell's pages are ranked in the top few positions [6].
2. Meta tag optimization
   Meta is an element mark, used between <head> and </head>. Meta tag has a lot of functions. At present, almost all of the search engine robots classify

webpage by searching meta value automatically to determine the quality of web content. The most important of all are description and keywords. Place body-related keywords in description with not too many words and do not exceed the maximum number of search engine summary. Keywords prompt search engine that the contents of this website start around these terms. Choose the most webpage-related core words as keywords and let them emerge in description. Another consideration of keywords is keyword density. Ensure key words and key phrases emerge in entire webpage for several times. Remember, "too far", 6–10% keyword density is the best.

3. Static website design
   Compared with the dynamic site, search engine spiders like static websites.
4. Page capacity rationalization
   Do not let the page is too large. Reasonable page size can increase the speed of page display and the friendly degree of search engine program.
5. Web directory and content
   The more web directory, richer content, the easier to attract concerns of search engine "spiders". This is why big sites are more easily indexed by search engines than small sites.
6. Domain name optimization
   No matter how long the domain name is, it will not affect the friendliness of search engine on domain name. It is better for domain name to have a certain meaning and significance. From the view of site optimization, containing keywords is the best way to register domain names. The length of time the domain name exists has certain impact on search engine optimization. Search engine gives different weights to different suffixes of domain names. Currently, search engine gives higher weights to .gov, .edu, .org, .com than general domain name. Relatively, .cn domain name has low weights.

67.2.3 Optimization in Pages Link

When search engines determine the ranking of a site, it should not only analyze the content and structure of the webpage, but also analyze the links of the website. The search engine spider is also tracing and penetrating deeply along the pages link.

1. Optimization in inbound links
   Inbound links refers to the links from external sites to e-commerce websites. Generally, it is thought that the following external links will help improve site ranking in search engines: (1) The external site is a static page; (2) There is a snapshot of the external page in the mainstream search engines, and it must be record more than 45 days; (3) The value of PR is larger than 3, and the larger the better; (4) The correlation between the external pages is large; (5) A website with few outbound links, no hidden links, no Flash links or Rich Media links; (6) A website with a frequent update. When the spider captures the update

pages, it will catch our pages; (7) External links are Blogroll. The most popular methods are publishing article or comments with links of your own websites in some forum, blog or space with high popularity, publishing original articles in some authoritative website, or publishing original news in web portal of news; (8) Submitting the website to some professional directory website, such as DMOZ, yahoo, ODP and so on. These directories are manual classified catalogs, with which website obtain some important and high-quality external links.

2. Optimization in outbound links
   Outbound links means the links that is added in e-commerce website and is concerned with the keywords of the site. This will not only benefit enriching the site content, but also help advance the search engine ranking. Of course, it differs to inbound. The number of outbound links should be controlled, although the theme is dependent with the e-commerce website and its outbound links. Because of from the "page level" perspective, the more the outbound links, the less the benefits of other pages in this website. The number of outbound links of a page should be controlled within 15, and the number in home page should be controlled with 10, and extra links should be arranged in subpage.

67.3 Conclusion

Search engine optimization is a long process and runs through all stages of e-commerce website from planning to promotion. It needs continuing research to ensure its effect. This paper puts forward promotion strategy of e-commerce websites based on search engine optimization to improve the website's friendliness of search engine and facilitate users to browse and access.

References

1. Fan Y (2010) Study on search engine optimization. Comput Appl Softw (1)
2. Liao W (2010) Research of website design optimization based on SEO Sci Technol Inf (11)
3. Li L, Wang X, Xue W (2010) Research of the website information edit strategies of SMEs based on SEO. J Mod Inf (12)
4. Yang Y, Su G, Deng J (2010) Campus network search engine ranking method based on user behavior analysis. Comput Eng (24)
5. Jiang J, Liu T, Liu Y (2010) Strategies for optimizing websites based on search engine optimization technologies and template engine technologies. Inf Stud Theory Appl (5)
6. Lin Y (2009) Simple analysis on application of search engine optimization technology. Softw Guide (11)

Chapter 68
Research of P2P Technology and Its Application

Xingmin Xu

Abstract Peer-to-Peer (P2P) technology is an active research area of next-generation networks, which guide the network computer shift from centralized mode to distributed style. The core of network applications spreads from a central server to the edge of the terminal equipment. P2P technology has been widely used in file-sharing, distributed computing and streaming media technology. This paper describes the P2P technology's definition, features and key technologies, focused on the status of the application of P2P technology and the problems of P2P technology in the final outlook the development of P2P technology's prospect.

Keywords P2P technology · File-sharing technology · Distributed computing · Streaming media technology

68.1 Introduction

P2P is the short form of Peer-to-Peer and also known as "point to point". P2P technology is a new network technology in recent years. It is an exchange pattern of network information relative to client/server model (C/S), which depends on the network participants' computing power and bandwidth instead of gathered in the dependent on a few servers. After the 1990s, with the improvements in computer

X. Xu (✉)
Computer and Communication Engineering Department,
WeiFang University, WeiFang 261061, China
e-mail: sdxuxingmin@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_68,

performance, due to the growth of network bandwidth and the growing user number demand, P2P technology has begun to be applied. P2P network technology has changed the way that people use the network, and also provides a new way of thinking for the future development of the network.

68.1.1 P2P Definition

There is no other definition for the unity of the P2P technology. Two of the most common of the P2P technology are defined as follows:

Definition one: P2P networking is an application that runs on a personal computer and shares files with other users across the Internet. P2P networks work by connecting individual computers together to share files instead of having to go through a central server.

Definition two: P2P is a type of Internet network allowing a group of computer users with the same networking program to connect with each other for the purpose of directly accessing files from one another's hard drivers.

68.1.2 Characteristics of P2P Technology

The major characteristics of P2P technology are the following:

1. Decentralization: The network's resources and services are distributed across all nodes. The achievement of information's transmission and services are directly between the nodes. It can go without the involvement of intermediate links and servers, avoiding the possible bottleneck. The P2P's basic characteristics of decentralization bring its advantages of scalability, robustness and other areas.
2. Scalability: In the P2P networks, with the increasing number of users, not only the demand for services increases, but also the overall system resources and service capacity expansion are synchronized. So it is always able to meet the needs of users more easily. Because the whole network system is all distributed, there is no bottleneck. In theory, we can consider its scalability almost infinite.
3. Robustness: The architecture of P2P has inherently advantages of being resistant to attack and high fault tolerance. Because the services are distributed between the various nodes, some node or the network destruction has small effect on other parts of the network.
4. High performance/price ratio: The excellent performance is an important reason for the widespread concern of P2P. It can be effectively useful for a large number of spread ordinary node in the internet with the use of P2P architecture. Because the use of the idle computing power or storage space can achieve high performance computing and mass storage purposes. In addition, the use of a

large number of idle network resources provides lower cost higher computing and storage capacity.
5. Privacy Protection: In the P2P network, the transmission of information scatters in between the nodes without having to go through a centralized part of the user's private information leaked by eavesdropping and the possibility of greatly reduced. In P2P, all participants can provide relay functions, thereby greatly improving the flexibility and reliability of anonymous communications, to provide users with better privacy protection.
6. Load balancing: Since each node even if the server is the client, P2P network environment reduces the traditional C/S structure server computing power, storage capacity requirements, because the resources are distributed across multiple nodes, to achieve better the entire network load balancing.

68.1.3 P2P Network of Key Technologies

P2P, as a new Internet application technology, is mainly related to the following key technology areas:

1. Efficient search strategy. As the P2P system resources are stored in each node instead of C/S mode and B/S mode, it will focus on the content stored on the server needed to provide a node for each node from the other resources needed to find access to the resources search mechanism. Resource search mechanism not only ensures that users can quickly find the information but also maximizes the bandwidth utilization of the current situation from the current study, there are three main file search algorithms: centralized indexing algorithm, algorithms, and distributed flood news find and message routing algorithm.
2. Communication protocol standard. Considering the current Internet, and exchange of data protocols such as XML, SOAP, UDDI and others, P2P software is a complete need to consider [4].
3. System's security mechanisms. Security is the basic need of all computer systems, because P2P networks are featured by the group, openness and anonymity of users in the network, it brings a lot of new security problems which can not be ignored. Current P2P networks for common security issues include: routing attacks, the target node overload attack, the system attacks, separated attacks, access attacks, exploits firewall attacks has been inconsistent behavior. How to address these security issues is the focus of P2P research.
4. P2P network monitoring technology. With the face of large-scale emergence of P2P virtual network, how effective inspection and evaluation of performance of P2P networks have become increasingly important. Due to lack of centralized monitoring, most P2P systems cannot provide current health information network, providing network adjustment and optimization is so difficult in the application by enterprises of all ages, which cannot truly achieve earning targets. Therefore, how to use the P2P network resources, cost-effectively

distributed to the others to provide centralized information network monitoring service is the key issue of P2P networks.

68.2 P2P Technology Applications and Research Status

68.2.1 The Main Application Areas of P2P Technology

The main driving force of Internet is the resource sharing. The file exchange needs the rise of P2P technology, which is not only the first P2P application, but also one of the most successful applications. With people's understanding and thinking of P2P technology, it promotes file exchange as well as the depth of search, distributed computing, instant messaging and personal work such as new applications.

1. File sharing and exchange: the use of P2P sharing of information resources can make better use of network bandwidth to improve the system efficiency of data communication, such as Napster [1] and Gnutella [2].
2. Distributed computing: the point is that the decomposition of a given number of nodes in the P2P network to collaborate computing done using the accumulated amount of ability to perform super-computing work (such as space exploration, molecular biology computing, extraterrestrial life search, etc.).
3. Search engine: the use of P2P technology is to develop a search tool in a P2P network nodes directly, in real-time search, and can reach the traditional centralized search engines, unparalleled depth of information search will be more effective and strong real-time [5]. Applications examples are InfraSearch, Pointera and other search soft wares.
4. Live video: P2P video live mainly in the application layer multicast technology, relying on other data between nodes to achieve the multicast forwarding function, some of the traditional video broadcast software are to solve the problem of limited bandwidth and load, the overall greatly improved quality of service. For example, Tsinghua University developed Gridmedia broadcast system, Huazhong University of Science and Technology R&D video broadcast system, AnySee, has been very popular on the Internet, as well as PPLive.
5. Storage services: The deployment of the Internet has given a number of storage servers, the storage of objects to be stored in different distributed servers, provide data storage services to ensure data reliability, availability, security and access efficiency. Such as Tsinghua University, the storage service system Granary, University of California, Berkeley OceanStore.
6. Collaborative work: P2P collaboration systems use a computer on the network node on the same computing platform to run together to complete a task or the sharing of information resources. Using P2P technology, you can create a secure enterprise collaboration platform, interactive supply and demand

information to help users maintain service distribution channels, etc. Typical applications are Groove, domestic KDT and so on.

68.2.2 Major Technology Company Research P2P Applications

1. Microsoft Corporation Pastry team are responsible for P2P computing technology research and development. Currently Microsoft has released a software package based on Pastry SimPastry/VisPastry.
2. August 2000, Intel announced the formation of P2P working group, formally launched the P2P research. After the establishment of working groups, the working groups actively cooperate with application developers to develop P2P applications platform. Intel released the 2002. Net-based architecture of the Accelerator Kit (P2P Acceleration Kit) and P2P Security API package, which makes Microsoft's. NET developers quickly build secure Web applications P2P.
3. Sun's Java technology has the background to carry out the JXTA project. JXTA is a Java-based open-source P2P platform, any individual or organization can join the project. JXTA has released a JXTA-based instant messaging software package. JXTA defines a set of core services: authentication, resource discovery and management. In terms of security, JXTA joined the encryption software package, which allows the use of the encrypted data packet encryption to ensure that the privacy of information can be authentication and integrity. The top of the JXTA core, also includes content management, information search have paid their management, including a variety of other optional JXTA services. In the core services and optional services, the user can develop a variety of P2P applications on the JXTA platform.

68.3 P2P Technology Problems and Challenges

Although P2P technology has been widely used and studied, the current application of the present situation, P2P technology itself, still has many problems to be overcome. P2P technology problems and challenges are reflected in the following aspects:

1. Protocol standards. Currently, the major problem is the lack of the development standard, which is difficult to set, also becoming one of the main difficulties in P2P development. P2P developers currently use different methods of developing P2P applications, the lack of uniform standards, has resulted in a lot of duplication, and product interoperability, scalability and platform independence problems.
2. Network security issues. Security issues in P2P networks is more serious, which will directly affect the P2P large-scale commercial, so security issues must be proven through a variety of security technologies to prevent malicious attacks,

virus and other issues to address in the user trust, user authorization, data accurate delivery and other issues.
3. Bandwidth problem. Since the constant expansion of P2P networks and the increased number of users to download files, P2P applications, which swallow network bandwidth become insurmountable obstacles. For network operators, a simple block P2P traffic is not a wise move, how to effectively manage the enormous P2P application traffic, how to use P2P technology has been a major issue for the development of value-added services.
4. Copyright issues. P2P copyright issue has been the development of an uncertain factor in the prevalence of P2P sharing network, intellectual property protection. P2P sharing software piracy boom has accelerated the dissemination of information, increased intellectual property protection more difficult. Napster offers free MP3 files because of exchange by the U.S. Recording Industry Association (RTAA) to court [3].

68.4 P2P Technology Development Prospects

P2P with its unique advantages, and the traditional model of technology, provides another effective network solutions that enable people to expect in the "open, free 'Internet vision into reality, which shows great potentialities for development. Currently, file sharing, P2P distributed computing and P2P streaming media technology into the field in the application of P2P technology has become a trend. This year, Microsoft, SUN, IBM, Intel and other large companies have begun to put into research in this area. Although there is still a bandwidth of P2P, copyright, network security and other issues, but its impact on the future of the network is beyond doubt. I believe that with the gradual deepening research of P2P technology, those factors restricting the development of P2P will eventually be gradually resolved, P2P will play a greater role on the Internet.

References

1. Peng S, Kong T, Wu J (2006) The development of P2P technology. Mod comput 2:35–38
2. Liu H (2008) Classification of P2P networks and key technical analysis. Micro-comput inf 3-3:112–114
3. Han T, Geng Y (2009) Development and application of P2P technology. Comput Inf Technol 3:78–81
4. Meng F, Yang C (2008) Application of web service in P2P network. Comput Eng Design 7:1688–1689
5. Liu J (2011) Study of the AVL-index tree range query based on P2P networks. Microelectron Comput 2:11–14

Chapter 69
Study on Safety On-Line Monitoring and Warning Systems of Tailings Reservoir

Jianxiong Zhang, Xinhui Wu and Guoqiang Zhao

Abstract With the development of select mining industry in all kinds of mineral resources in our country, it has become increasingly a serious problem that safety in production and environment of various tailings reservoir which go with select mine refinery, by means of the modem technology to make the on-line monitoring and warning for tailings reservoir security is a practicable method for ensuring safety in tailings reservoir. Presented design principle of the system, the structure of tailings reservoir on-line monitoring and warning systems, contents of monitoring, information communication style and software system. Focused on the monitoring methods in system construction, such as dam deformation monitoring, interior displacement surveying of dam body, water level monitoring of reservoir areas, automatic monitoring for saturation line and so on, these can be a basis for on-line monitoring and warning system.

Keywords Tailings reservoir · Monitoring · Warning

J. Zhang (✉)
School of Energy Science and Engineering,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: zhangjx3s@126.com

J. Zhang · X. Wu · G. Zhao
School of Surveying and Land Information Engineering,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: toyowxh@gmail.com

G. Zhao
e-mail: zgqabcd@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_69,

69.1 Introduction

There are abundant category mineral resources in our country, after 60 years of exploitation and development since 1949, metallic and non-metal mines have more than 100,000 that have the enormous select mine refineries which go with the mine production, however, a very great deal of the select mine refineries has tailings reservoir which serve as the important facilities of mine production, it well done or not will has direct bearing upon the mine production efficiency and the safeties of lives and property of the people.

According to the recent material, there are more than 12,000 tailings reservoir in our country, for production units takes an excessive interest in economic benefit in the past, safety production and management consciousness to tailings reservoir is thin, yet parts of tailings reservoir have no formal design and approval procedures, embankment technology is simple, plus failing in construction supervision and safety production situation gives some cause for the concern. Especially the "10·28" tailings reservoir dam-break serious accident happened in XiangFen Xinta Mining Company in the Year of 2008, 277 people were killed, 4 people were missing and 33 injured and the direct economic loss hit 9619.2 million yuan, this serious accident has caused wide public concern about tailings reservoir security. For security of tailings reservoir, besides further improve the safety monitoring of tailings reservoir, we should build safety monitoring system, forming more reliable, efficient, fast safety monitoring mechanisms and management decision information channel. Recent years, with the development of the computer science and micro-electronic technology, the popularization and application of tailings reservoir safety production and automatical monitoring and warning technology offers possibility. At the same time, "Notify on Safety Production 'Eleventh Five-Year Plan' Which State Council General Office Issued" in the 2006 has explicitly demanded that tailings reservoir which design reservoir contain over 10 million or design main dam is more than 60 m high must establish and improve security monitoring system.

69.2 Systems Design of' Principle

According to the demand of on-line monitoring and warning system of tailings reservoir, system design should include the following principle:

(1) Functions of the system including stability collect, display, storage, data communication, management, auto detection and warning of system.
(2) System has remote-control function, it can make use of the network to have telecontrol monitoring for the host computer through serial ports, implement all functions of the data acquisition software, and adopted to access right of historical data for data acquisition software.

(3) System can make compilation of observation data at stated intervals time that provides materials of design, construction and management for later work.
(4) System can perform analyses for observation materials anytime, undertake technical appraisement for operating condition of dam body, and sum up of the experience in the operation, which provide data for improving running way, establishing safeguards and evaluating production conditions.
(5) System can calculate the flood height, safety height of reservoir area through the real-time acquisition data and make warning; can make warning for the state of reservoir area water level through data of reservoir area water level and flood capacity; can real-time analyse the safe state of dam body at the moment through geological data in designing of tailings reservoir and in combination with the scene of data acquisition; can draw automatically a saturation line of dam body through the scene of data acquisition and give relevant data; and can analyse the settlement and horizontal displacement of dam body and predicate the development of deformation through the analysis result.
(6) System can integrate the historical data and real-time data of seepage, water level, deformation and give some related analysis in according to with the relevant standard provisions.

69.3 The Mainly General Structures and Functions of the System

69.3.1 The Mainly General Structures of Tailings Reservoir Security in the On-Line Monitoring and Warning

System consists of two major parts: on-the-spot automatic monitoring warning and analysis issue, on-the-spot automatic monitoring warning consists of each measurement subsystem systems, data processing subsystem systems, data communications subsystem systems and monitoring alarm subsystem systems; analysis issue consists of data analysis issue and information sharing system. Composition schematic diagram of on-line monitoring systems of tailings reservoir is shown in Fig. 69.1.

69.3.2 The Mainly Functions of the System

System mainly includes on-line monitoring of tailings reservoir and analysis and issue emergency handling functionality. Concrete compositions of the modules show as Fig. 69.2.

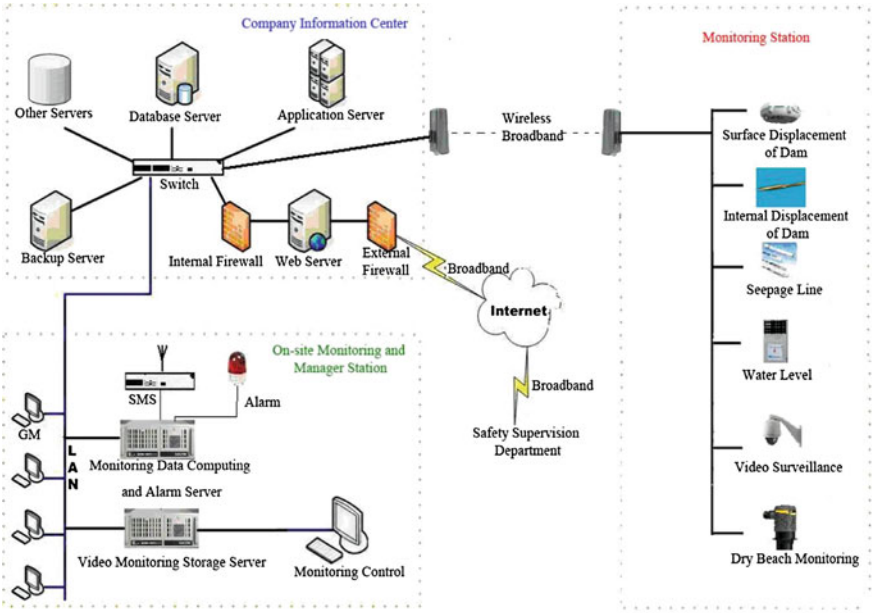


Fig. 69.1 Composition schematic diagram of on-line monitoring systems of tailings reservoir

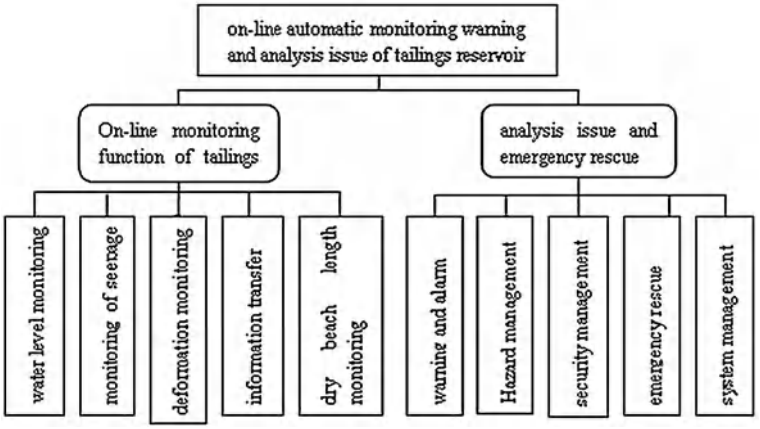


Fig. 69.2 The system function module diagram of on-line monitoring systems of tailings reservoir

69.3.2.1 On-Line Monitoring Function of Tailings Reservoir

The function consists of on-line monitoring modules which includes water level monitoring of reservoir areas, monitoring for saturation line, dam deformation monitoring, safe superelevation monitoring and dry beach length monitoring.

Data of real-time monitor and other data provide from the alarm and information transmission to analysis and issue emergency handling functionality.

69.3.2.2 Safe Alarming of Tailings Reservoir and Emergency Handing Joint Function

Control system has automatic warning previously function, when the monitoring parameters evolves to dangerous state, system will issue forewarning information; when the monitoring parameters is more than preset warning number, system sill issue alert information. So it can provide against accidents and check the hidden peril of the accidents in the bud.

69.4 The Construction of System

69.4.1 Monitor Contents of System

The main reasons that causes tailings reservoir disaster are dam body break, dam body seepage and earthquake, are next. Therefore, while constructing the safety monitoring system of tailings reservoir, fixing the displacement of a dam body, capability of flood control (through monitor dry beach length and reservoir water level to indirectly gain) and saturation line mainly monitor contents.

69.4.1.1 Deformation Monitoring of Tailings Reservoir

Dam deformation monitoring is mainly set in the GPS observation stations in the dambody. GPS receiver antenna receives GPS signals in real-time and sends signals to the control center through a certain transmission. After the real-time three-dimensional deformation analysis system software computes the signals, we obtain real-time three-dimensional coordinates of each point which reflect trace changes in the displacement. A more economic and facilitate the implementation of the current “one machine multi-antenna” GPS deformation monitoring mode structure is shown in Fig. 69.3.

69.4.1.2 Dam Internal Displacement Monitoring

When make dam internal displacement monitoring, people usually adopt the method that the inclinometer monitoring buried in the drilling within the dam to make internal displacement monitoring. The displacement sensor is put in the inclinometer monitoring, so we can collect the real-time deformation of inclinometer.

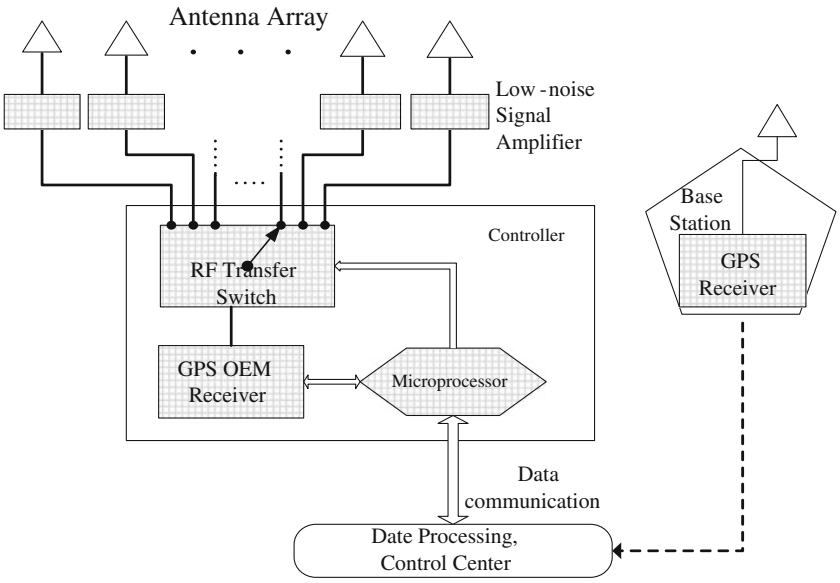


Fig. 69.3 "Antenna and a machine" GPS deformation monitoring system structure diagram

Fig. 69.4 Inclinometer

According to the monitoring data obtained, we can obtain dam deformation data by analysis and calculate.

Inclinometer is mainly used for long-term automatic measurement of horizontal displacement and angle of the dam. Through the drilling, we can put inclinometer probe buried the ground through the connecting rod means. When generating deformation, the platform is tilted. The inclinometer probe in pace with the platform, the signal cable is connected into the ground test measurement instrument. According to the size of angle values, we can accurately measure the horizontal displacement ΔX, ΔY to forecast. With automated data collection devices, Inclinometer can automatically and continuously monitor. Multiple sensors can be installed to obtain deflection profiles along the tube axis inclinometer (Fig. 69.4).

69.4.1.3 Tailings Water Level Monitoring

Tailing water level is the free surface elevation for a benchmark zero. Water level Monitoring generally is set up automatic monitoring instrument on the

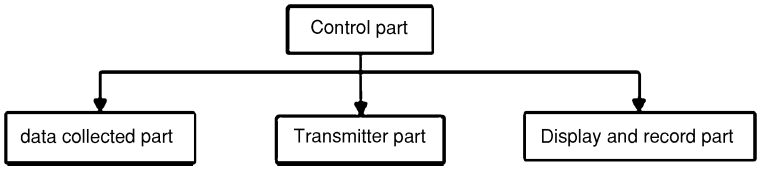


Fig. 69.5 Video monitoring module

bank drainage structures. The measured signals are accurately and real-time passed to the indoor receiver handle for water level. It should be noted that, drainage structures are generally located in the tailings repository, tailings water clarified is around drainage structures. Therefore we need to choose flexibly the construction plan for the actual situation of specific tailings before the monitor system is layout. In terms of the selection of monitoring equipment on water level, because ultrasonic level meter owns the advantages such as easy-to-directional emission, the direction, strength and easy control, and does not require direct contact with the object being measured, the millimeter accuracy in the precision level measurement etc., it becomes the ideal equipment in tailings water level monitoring.

69.4.1.4 Automatic Monitoring of Seepage Line

The location of seepage line is one of the most important parameters to analyze the stability of tailings dam. So it is also an important feature for determining whether or not dam is safety. Automatic monitoring of seepage line mainly puts the osmotic pressure sensors layout directly in the observation hole of seepage line, collects automatically the data regularly, and transfers these data to monitoring center for display and storage by routing.

69.4.1.5 Video Monitoring

Real-time video monitoring is a supplementary and auxiliary means in safety. Intelligent digital monitoring system is a new generation of monitoring system to combine with video monitoring, real-time monitoring, multi-split screen, multi-split screen display, PTZ camera control. Video monitor module is shown in Fig. 69.5. By the establishment of surrounding security monitoring and warning systems, video surveillance can effectively protect key areas and staff safety, prevent maximum from various invasion to improve the reflecting speed of handling with various emergencies and provide a good working environment to security personnel for the safety of the entire enterprise.

69.4.1.6 Dry Beach Length Monitoring

There are many monitoring methods of dry beach length, based on photogrammetry principle, in which bolt is placed in a fixed position and points to dry beach, and waterline is automatically identified with the help of the graphical comparison and optical imaging, then dry beach length is automatically measured by use of three-D triangulation method with three-D coordinates of key points or also artificial points in the poor weather conditions, whose accuracy will be better than 5 m.

69.4.2 Communication of System

The communication part of monitoring and warning system includes mainly two parts: (1) Communication design from monitoring equipments (the sensor for example) to monitoring station; (2) Communication design from monitoring station to on-site monitoring management station. Monitoring station is set up in observation room on the dam topper (If no room, equipment location room will be built lonely near by observation room). Wired Cable is used to data transmission from monitoring equipments (the sensor for example) to monitoring station, and radio or wire transmission can be selected to that from monitoring station to on-site monitoring management station based on the distance and visibility between dam and on-site monitoring management station. When the data and video are transported into on-site monitoring management station, they are also transported to customs controlled computers by internal broadband network, contemporarily, are mapped and published onto Internet by internal broadband router (firewall) port.

69.4.3 Software Systems

The overall safety of tailings is determined by the securities of the dam, leakage and flood regulation, and online monitoring system of tailings is related to the automatic monitoring of the three parts. The dam security includes dry beach length and secured elevation difference. In monitoring system, each factor affected tailings is monitored online, monitoring data of each module is analyzed professionally and calculated quantitatively, then the real-time dynamic process and safety of tailings are exported by software user interface of the system, the danger or detrimental factors occurred in running process of tailings and the related practical countermeasures (expert system database) are summarized and proposed by deep analysis to monitor data, and the convenient and reliable methods are provided for daily operation and management of tailings.

The system sets up warnings of grade III, and gets monitoring data by 365 days and 24 h of continuous monitoring to significant danger. When monitoring data above the warning value, the alarm information is transferred automatically to various terminal of company leaders and related personnel; and automatically associated contingency plan, detailed graphic information and analysis reports can be looked up, deal with and cancelled anytime and anywhere on internet.

69.4.4 The Key Technical Indicators of Generally Monitoring System

(1) The accuracy of dam seepage line more than 10 mm.
(2) Dam displacement monitoring accuracy of horizontal displacement better than 5 mm and the accuracy of Vertical displacement better than 8 mm.
(3) The accuracy of Reservoir water level monitoring is better than 10 mm.
(4) Video loop recording more than 20 days.

69.5 Conclusion

We analyse the on-line monitoring and warning system on composition, function and other aspects, this system is essentially a management system, it includes information and data collection, processing, conclusions, take measures, information feedback and other major aspects. Its fundamental purpose is to improve efficiency and keep lives and property safety. With the science development various high-tech were used in the system especially the information technology was used in the system successfully. It reduced the labor intensity of traditional monitoring. And it also improved the accuracy of monitoring information and automated analysis. In our country, we will promote the monitoring of tailings warning systems, it can improve tailings management level of safety. It enhanced the warning capacity of enterprise, community, government about disaster.

References

1. Li X (2010) Disposal and experience for tailing pond emergency. Nonferrous Metals (Mining Section) 62(6):39–41
2. Li Q, Xie X, Chen S et al (2008) Key points of tailings monitoring. Labour Protect (12):97–99
3. Mei G, Wang Y (2010) Statistic analysis and countermeasure study on tailings pond accidents in china 6(3):211–213
4. Lin Z (2007) Study on GPS deformation monitoring system of mine tailings dam. Hohai university, Jiangsu

5. Tian W, Xue J (2006) Safety technical and management for the tailing pond. Coal Industry Publishing House, Beijing
6. Yin W, Li L (2009) Utilization and management for tailing. Metallurgy Industry Press. Beijing
7. Liu X, Li M, Zhang S (2003) Study of the expert system for dam safety monitoring. J Os South Inst Metall (5):1–5

Chapter 70
Research on Database Security Technology

Fuguo Li

Abstract With the rapid development of computer technology, database technology has been applied to a wide range of fields, thus the security of it has also become more and more important. The paper discusses the concept and safety system database security, the main risks facing database system, commonly used technology of database security, data security transmission protocol and the tendency of commonly used database security technology research.

Keywords Database security · Security system · Security technology · Security agreement

70.1 Introduction

With the rapid development of computer technology, database has been widely applied to all fields, and the problems following data security has also become more and more prominent. As information aggregation, database system is the core parts of the computer information system, thus its security is very important. We have been paying more and more attention to such problems concerning the security issues of large amounts of data existing in all kinds of application in the system's database and ways to prevent tampering and stealing sensitive data.

F. Li (✉)
WanFang College of Science and Technology,
Henan Polytechnic University, Jiaozuo, China
e-mail: lfg@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_70,

70.2 Database Security Concepts and Safety System

Database security means to protect the database from excessive using, stealing, changing or data destroying. Database security involves many aspects, for instance SQL Server database security including their own security mechanism, external network environment, the operator's technical level, etc. Therefore, database security can be divided into three levels, namely network system security, operating system security and database management system security [1]. The three levels built into database security system are in a close relationship with data security.

70.2.1 Network System Security

System safety is the first safety barrier of database. The main threats facing current network system are Trojans, network deceive, invasion and viruses, etc.

70.2.2 Operating System Security

Operating system security problems mainly come from network used in the operating system. For example at present there are mainly three defects in the operating system Windows 2003 Server including: the defect of the operating system itself; The security features of operating system, namely the relevant security strategy configurations; The threat of virus.

70.2.3 The Security of Database Management System

Based on the fact that different database management systems are used, we will apply different security Settings. For SQL Server 2000 which can be classified by database administrators, and different database users have different access privileges; We will use view information isolation to prevent users basic operation of tables; At the same time database backup operation should be prepared regularly to avoid date loss.

70.3 The Main Risks Facing Database System

There are many problems existing in the application of database system, which would result in all kinds of security problems. The security risks facing database system can be divided into operating system risk, management risk and the risk from users.

70.3.1 Operating System Risk

The security of database system relies on the safe working environment provided by operating system and hardware devices, if the operating system allows users' direct access to files in the database, then even the most reliable safety measures would not guarantee the security of the system [2].

70.3.2 The Management Risk

The management risk mainly refers to users' security awareness, the attention that paid to the information network security and relevant measures of safety management.

70.3.3 The Risk from Users

It is mainly displayed in the users' accounts and operation rights to specific database.

70.4 The Commonly Used Database Security Technology

70.4.1 Access Management Technology

Access management technology refers to a process which prevents unauthorized use and access to database through running program, including user authentication technology and access control technology [3]. User authentication technology includes the user's identity verification and user identity; Access control model can be divided into DAC, MAC and RBAC.

70.4.2 Security Management Technology

Safety management deals with the management system that is applied in the authority distribution of database management. And it can be divided into centralized control and distributed control. The centralized control of single control system provides security maintenance to the whole system, which ensures our more convenient safety management; Distributed control is the use of the available management program in controlling different parts of the database which ensures the security maintenance of the system [4].

70.4.3 Database Encryption

Database encryption is an effective means to prevent data leak, in this way the safety of the user information can guaranteed, thus reducing the damage caused by backup medium stolen or lost [5]. Ways of database encryption mainly include: the library encryption, machinery encryption, hardware encryption, etc.

70.4.4 Audit Tracking and Attack Detection

When the auditing is in operation, it will automatically record in the audit log all the operation records. Attack detection system is the detection of internal and external attack [6]. After analysis the weaknesses of the system, it represents the reasons that causes the current situation and then tracks down that convinces related.

70.4.5 Information Flow Control

Information flow control grades all the elements of the mechanism and categories all the components. Information flow control is responsible for the checking of the information flow, so that the information of the high level of protection objects will not be transmitted to low levels, which could avoid the perception of secret information by some malicious users from a lower level of protection.

70.4.6 Reasoning Control

Inference control is the acquiring of unauthorized data or information indirectly. The goal is to prevent users' deeds of this kind.

70.4.7 Data Backup and Recovery

It is very important to have data backup which contribute a great deal to the recovery of database in case of virus or accidents.

70.5 The Commonly Used Agreement of Data Security Transmission

70.5.1 SSL Protocol

Secure Socket Layer (SSL) protocol has become the standard to identify network websites and web visitors' identity, as well as the communication standard for browser and the web server in globalize encryption.

70.5.2 IPSec Agreement: IPSec (Internet in Security)

IETF is the safety standards defined by IETF framework, aiming to provide end-to-end encryption and certification services of public and private network. It specifies the various optional network security services, and organizations can choose these services according to their own security strategy, which is comprehensive in theory on constructing IPSec security solutions [7]. In this way data can be sent to improve confidentiality, integrity and reliability. The agreement provides three basic elements, namely "validation head" and "package" and "the Internet security load key management agreement" for the protection of the network communication [8].

70.5.3 HTTPS Protocol

Hypertext Transfer Protocol over Secure Socket Layer (HTTPS) is developed by the Netscape and placed inside their browsers, using for data compression and operation decompression [9]. After all these it would go back in the network and return the results and uploaded.

70.6 Research Trends of Database Security Technology

With the development of computer technology and the expansion of database technology application, database security technology will have the following trends.

70.6.1 Research on the Security Model

Research on the security mode includes the actual old application security model and new security model.

70.6.2 Hidden Problem

Hidden problem is ways of thoroughly detection and elimination through information flow control, inference control and other safety control.

70.6.3 Database Audit Trail

The size of the fine audit takes time and space. How to improve the audit system or whether we can use the high efficiency of the automation tools and DBMS audit integration.

70.6.4 Database Technique and Other Related Technique Combine

The influence of database technology and other combination of related technical on database security should combine.

70.6.5 Research on Multilevel Security Database

The semantic research on multilevel security and the improvement of database and DBMS multilevel security protection system are further perfect.

70.6.6 Database Use with the Kit of Other Security Products

The vulnerability of database system can be easily taken advantage of, as a result the database security must be supported by the intrusion detection system, firewall and other safety products.

70.6.7 Security Research on Application System and Database

Research on of the application system and connection of the program itself is also part of the generalized database system safety research.

70.6.8 Research on Database Encryption Technology

Research on database encryption technology and application will be the focus of research and extension in finance, business database and other important departments in the future [10].

70.7 Conclusion

Database security, network security, operating system security and database management system security are closely linked to each other. Our analysis must base on the application environment. We should also work out unified security management strategy to ensure the safety of database system. More importantly,

we must strengthen internal safety control and safety training for administrators. Safety is a long-term problem which still needs further security maintenance and management.

References

1. Li S, Dai J, Zhang L (2009) The model of security enhancement for database based on rule self-study. Comput Appl Softw 26(1):110–111
2. Bai K (2010) Analysis and process of application system database security issues. Electr Des Eng 18(7):91–93
3. Chen J (2007) Research of security management for DDBS. Control Manag 23(15):154–156
4. Hu S, Ye X, Peng Y, Xie F (2009) Design and development of a defense-in-depth model for database security. J Comput Res Dev 46(z2):99–104
5. Xiu M, Feng X, Sun Z, Ruan H (2007) Network database security and application. Comput Knowl Technol 3(17):1272–1273
6. Zhang R, Liu C (2007) Research on web database security mode. J Beijing Union Univ Nat Sci 21(4):25–27
7. Yang T (2008) Study of database security and maintenance. Comput Knowl Technol 4(31):776–777
8. Zhou S (2010) Research on the network database and application. Comput Knowl Technol 6(5):1038–1040
9. Gu Z, Du G (2007) Research on database security of SQL server application. Comput Eng Des 28(15):3717–3719
10. Lin J, Huang T, Yang C (2008) Realization of a database security auditing forensic model. Control Autom 24(27):132–134

Chapter 71
Petri Net-Based Component of the Reconfigurable

Ping Zhang and Bin Tao

Abstract Software reuse is one of the realistic ways to solve "software crisis". Software architecture (SA) and component-based software development technology (CBSD) start to play an important role in software development. In the component-based software development, construction of components and division method directly affects the reusability and work performance of component. This paper proposes a new component generation method which is able to merge or optimize the components based on Petri net components, reduce the difficulty of component reuse, speed up the software development cycle and improve software performance.

Keywords Petri nets · Components · Construction · Components division

71.1 Introduction

With the rapid development of information technology, how to further enhance the depth and professional level of software development has become an important issue faced by the nation to realize information construction strategy [1–4]. In the software field, there are basically three kinds of development model: process-oriented, object-oriented and component-based. Component is a special object,

P. Zhang (✉) · B. Tao
Computer Science and Engineering Department, Henan Engineering College, Zhengzhou 451191, China
e-mail: zpings@sina.com

B. Tao
e-mail: tb3190@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_71,

owning independence function of software modules that follows certain technical standards, such as COM, CORBA and EJB, which makes it easier to implement modular software systems and has wide application in reusable software area. Component-based software system development method is currently a hot research in software engineering field. The current study mainly focuses on the static build of software system functions, which is to process the new aggregation and inclusion of the components of original system when the system is putting forward optimized or expanded functions, but it is not very easy to achieve. Therefore, the stratified design of existing components in source code level can repartition or combine the new components based on hierarchical relationship between calls easily and quickly. Components of the software system is a discrete systems, which could be used to study how to dynamically build the software system components by establishing common Petri net model of component-based software system [5–8].

71.2 Petri Net Modeling

Petri net is a mathematical representation of discrete parallel system. Petri nets is invented by Karl A. Petrie in the 1960s, suitable for describing asynchronous, concurrent computer system model [9–12]. Petri net has both rigorous mathematical formulation, but also intuitive graphical expression, suitable for discrete system modeling and analysis. The basic structure of Petri nets composed by the library, the composition change and the arc, describing the static structure of the system; on the basis of this site in the base to increase the library logo, library capacity, the weight of arcs and change the rules occur, so that they can describe the dynamic behavior. For component-based software system that studies in the paper, library arcs weight capacity and the base make a good description of this site to the structure and dynamic behavior of systems, so it apply the basic Petri nets to a modeling tool [9, 10].

One of the main function Petri net workflow is to process definition which is premised on the need to build models of software systems, that is, transforming the software system model into Petri net model for numerical analysis [13–18]. There are many method involved in models of software systems. Generally there will be four basic entities: activities, roles (a resource for participation), information (activity data needed) and function (related constraints with activity). In order to make the entities and basic structure of the software system model transformed into the corresponding Petri net, it has to be in accordance with the following principles.

Activities of the corresponding Petri net consists of two changes and a library of the composition, the number of tokens in the library marked the ability of activities.

The roles of the corresponding Petri nets make up of one or more libraries, each of the libraries represent a feature of the role in it, and the number of token in library stand for the ability of roles.

The corresponding information with Petri nets posed by a library, and the number of tokens in it represents the ability of information.

Function of the corresponding Petri net, according to the different constraints, the structure is changed. With the time constraints, posed by a library; conditions or time constraints, posed by the two libraries.

The Petri nets that corresponding with parallel architecture posed by three libraries. The number of tokens in library represents the current state of the structure.

The Petri nets that corresponding with serial structure posed by the two libraries. The number of tokens in library represents the current state of the structure.

The transformed Petri net model needs to be simplified that in accordance to the enterprise business process model in order to better analyze the characteristics of Petri net model. The basic principle is to simplify the Petri net to maintain the same basic features of the case or transfer multiple abstraction libraries into a library or transfer, and to ensure that the simplified Petri net model is reached, bounded and active, so that Petri net is safe and feasible with these characteristics, and it can be based on the realization of business processes [19–24]. Petri nets can be simplified according to different performance requirements, or to merge multiple components optimized to generate a larger particle size or performance of better components.

71.3 Building and Zoning of Components

The component is redefined here as: the applied package that has the function of restratification. Component should provide the external interface of each functional layer addresses and parameters; the less functional layer of components, the smaller the particle size, and the minimum is 1; user can build different components in demand of construction; there must be a main function layer component calls other sub-layer functional layer, respectively, if it has multiple function layers.

All the sub-function layer of a component may be a separate component, which is called C-type components that is no longer sub-component unit. For a multi-component, it is called Z-component that can call all the other components of C-component. C–C, CA and AA can be combined into a new component and generated the external interfaces and components described functions according to the original components [25–30].

A complete software system includes multiple steps, one or more steps can be implemented as a component. The components are connected by front and rear components work together according to the order, in that way it is able to complete the software system. Components generally have two work modes: synchronous and asynchronous. Synchronization refers to the component tasks completed before returning control to the previous component; asynchronous refers to the

components in the processing returns control to the beginning, and the first component to a new treatment can begin to work. Asynchronous serial structure can achieve the parallel effect, greatly improve efficiency. In the use of components of building systems, the system's state transition model would be differing based on the work of different components.

Here, the Ti ($i = 1, 2, 3, \ldots, n$) represents a synchronous component, Yi ($i = 1, 2, 3, \ldots, n$) represents an asynchronous component and Hi ($i = 1, 2, 3, \ldots, n$) represents a component of the time consuming complete the task. The merging rules of component are as follows.

The closely connected multiple simultaneous components are synthesized to a new synchronization components, its performance does not improve, which is still time consuming for the sum of the component parts.

The bounded connected component T1 and a synchronized asynchronous component Y2 is synthesized to a new component, and its performance does not improve dealing with a mandate. The performance increased dramatically coping with N consecutive processing tasks. When $H1 < H2$, the new component time-consuming is $H1 + N \times H2$; When $H1 > H2$, the new component time-consuming is $N \times H1$.

For the synchronized component T3 includes multiple inputs, it still plays an synchronization role with the input components, and its performance does not improve when they merged into a new component, processing one or more consecutive mandate.

For the asynchronous components Y4 includes multiple inputs, it still plays a synchronization role with the input components, and its performance does not improve, when they are merged into a new component processing an ordered mandate. The performance increased dramatically dealing with N consecutive processing tasks. When the maximum time-consuming Hx of input component is less than H4 time-consuming, the time-consuming of synthesized components is $Hx + N \times H4$; when $Hx > H4$, the new component time-consuming is $N \times Hx$.

For the components T5 that contains multiple outputs, it is a serial output operation when the output components are both synchronized components, and its work performance does not improve, combining them into a new component and processing one or more consecutive mandate.

For the components T7 that contains multiple outputs components it is a serial output operation, when the output components are asynchronous components, and its work performance does not improve, combining them into a new component and processing one consecutive mandate. The performance increased dramatically handling with N consecutive processing tasks. When the maximum time-consuming Hx of input component is less than T7 time-consuming, the time-consuming of synthesized components is $N \times H7$; When $Hx > H7$, the new component time-consuming is $H7 + N \times Hx$.

According to the above rules, for different needs, Petri net model can be based on rational division components. This classification can greatly improve the flexibility and the adaptation component of the overall efficiency of the system.

71.4 Conclusion

This paper presents a new understanding of component, on this basis, and a new building and zoning method of component is proposed based on it. The method can be divided into components dynamically to improve the overall performance mode on the basis of Petri net model; it can be quickly generated into composite components to improve reusability in accordance with the given generation component way, speed up the software development cycle and improve software performance.

References

1. Xiang JL, Yang J, Mei H (2004) ABC-Tool—an architecture-based component composition. J Comput Res Develop 41(6):956–964 (in Chinese with English abstract)
2. Mayer A, McGough S, Furmento N, Lee W, Newhouse S, Darlington J (2003) ICENI dataflow and workflow: composition and scheduling in space and time. In: UK e-Science All Hands Meeting, pp 627–634. http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/132.pdf
3. Mei H, Chen F, Feng YD, Yang J (2003) ABC: an architecture based, component oriented approach to software development. J Softw 14(4):721–732 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/14/721.htm
4. Scheben U (2005) Hierarchical composition of industrial components. Sci Comput Program 56(1–2):117–139
5. Gössler G, Sifakis J (2005) Composition for component-based modeling. Sci Comput Program 55(1–3):161–183
6. Sadaoui S (2003) Composition of structured process specifications. Electron Notes Theoret Comput Sci 82(5):1–12. http://www.elsevier.nl/locate/entcs/volume82.html
7. Kim J, Spraragen M, Gil Y (2004) An intelligent assistant for interactive workflow composition. In: Vanderdonckt J, Nunes NJ, Rich C (eds) Proceedings of the international conference on intelligent user interfaces (IUI-2004). ACM Press, New York, pp 125–131
8. Alda S, Cremers AB (2005) Towards composition management for component-based peer-to-peer architectures. Electron Notes Theor Comput Sci 114:47–64
9. Regehr J (2002) Scheduling tasks with mixed preemption relations for robustness to timing faults. In: Baruah S (ed) Proceedings of the twenty-third IEEE real-time systems symposium. IEEE Press, Austin, pp 315–326
10. Tokuda L, Batory D (2001) Evolving object-oriented designs with refactorings. J Autom Softw Eng (8):26–33
11. Caballero R, Demurjian S (2002) Towards the formalization of a reusability framework for refactoring. In: Gacek C (ed) Proceedings of the seventh international conference on software reuse. Lecture notes in computer science, vol 2319. Springer, Berlin, pp 293–308
12. Russell N, ter Hofstede AHM, Edmond D, van der Aalst WMP (2004) Workflow data patterns. QUT Technical report, FIT-TR-2004-01. Queensland University of Technology, Brisbane. http://is.tm.tue.nl/staff/wvdaalst/publications/p220.pdf
13. Hollingsworth D (1995) The workflow reference model, Document No. WfMC-TC-1003. Workflow management coalition. http://www.wfmc.org/standards/docs/tc003v11.pdf
14. Fan YS (2001) Foundation of workflow management technique. Tsinghua University Press, Beijing, pp 63–66 (in Chinese)
15. Jacobson I, Booch G, Rumbaugh J (1999) The unified software development process [M]. Addison-Wesley, Boston

16. Subrahmanyam A, Cedric B, John D (2001) Professional Java Server Programming. J2EE 113 Edition [M]. Wrox Press, USA pp 694–696
17. Xiang JL, Yang J, Mei H (2004) ABC-Tool—an architecture-based component composition. J Comput Res Develop 41(6):956–964 (in Chinese with English abstract)
18. Mayer A, McGough S, Furmento N, Lee W, Newhouse S, Darlington J (2003) ICENI dataflow and workflow: composition and scheduling in space and time. In: UK e-Science All Hands Meeting, pp 627–634. http://www.nesc.ac.uk/events/ahm2003/AHMCD/pdf/132.pdf
19. Mei H, Chen F, Feng YD, Yang J (2003) ABC: An architecture based, component oriented approach to software development. J Softw 14(4):721–732. http://www.jos.org.cn/1000-9825/14/721.htm (in Chinese with English abstract)
20. Scheben U (2005) Hierarchical composition of industrial components. Sci Comput Program 56(1–2):117–139
21. Gössler G, Sifakis J (2005) Composition for component-based modeling. Sci Comput Program 55(1–3):161–183
22. Sadaoui S (2003) Composition of structured process specifications. Electron Notes Theor Comput Sci 82(5):1–12. http://www.elsevier.nl/locate/entcs/volume82.html
23. Kim J, Spraragen M, Gil Y (2004) An intelligent assistant for interactive workflow composition. In: Vanderdonckt J, Nunes NJ, Rich C (eds) Proceedings of the international conference on intelligent user interfaces (IUI-2004). ACM Press, New York, pp 125–131
24. Alda S, Cremers AB (2005) Towards composition management for component-based peer-to-peer architectures. Electron Notes Theoret Comput Sci 114:47–64
25. Regehr J (2002) Scheduling tasks with mixed preemption relations for robustness to timing faults. In: Baruah S (ed) Proceedings of the twenty-third IEEE real-time systems symposium. IEEE Press, Austin, pp 315–326
26. Tokuda L, Batory D (2001) Evolving object-oriented designs with refactorings. J Autom Softw Eng 8: 89–120
27. Caballero R, Demurjian S (2002) Towards the formalization of a reusability framework for refactoring. In: Gacek C (ed) Proceedings of the seventh international conference on software reuse. Lecture notes in computer science, vol 2319, Springer, Berlin, pp 293–308
28. Russell N, ter Hofstede AHM, Edmond D, van der Aalst WMP (2004) Workflow data patterns. QUT Technical report, FIT-TR-2004-01. Queensland University of Technology, Brisbane. http://is.tm.tue.nl/staff/wvdaalst/publications/p220.pdf
29. Hollingsworth D (1995) The workflow reference model, Document No. WfMC-TC-1003. Workflow Management Coalition, http://www.wfmc.org/standards/docs/tc003v11.pdf
30. Fan YS (2001) Foundation of workflow management technique. Tsinghua University Press, Beijing, pp 63–66 (in Chinese)

Chapter 72
Digital Watermark Encryption Algorithm Based on Arnold and DCT Transform

Wei-yuan Han, Yan Yang and Hui-lai Zhi

Abstract Digital watermarking is one of the effective methods which can protect the copyright of multimedia data. This thesis presents a digital watermarking algorithm based on Arnold and DCT Transform. First, Arnold transform is used to shuffle the watermark image, and then the algorithm makes 8×8 DCT on the carrier image. In order to improve the robustness of watermark, this thesis adopted the multi-embedded strategy. The DC coefficients in the block of the same location of each color components are embedded the same robust watermark which is transformed by Arnold. Experiments show that the improved algorithm is an effective approach which can resist the attacks and has good invisibility.

Keywords Digital watermarking · Arnold transform · DCT transform

W. Han (✉)
Department of Information Engineering, Jiyuan Vocational and Technical College, JiYuan 454650, Henan, China
e-mail: hwy9202@126.com

Y. Yang
Department of Art Design, Jiyuan Vocational and Technical College, JiYuan 454650, Henan, China
e-mail: yy6463@163.com

H. Zhi
School of Computer Science, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zhihuilai@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_72,

72.1 Introduction

Using the Internet for multimedia information transmission process reveals more and more problems with development of network and multimedia technology. Digital watermarking technology (Digital Watermarking) is a new information security technology proposed in this case [1]. This article proposed an algorithm based on Arnold and DCT transform. In order to increase the difficulty of being deciphered, we use an improved Arnold transform algorithm to encrypt binary image watermark algorithm. Because DCT transform in watermarking technology is more mature, small amount of calculation, and international popular compression and encoding standards-compliant, etc. [2]. This article considers that embed watermark in the DCT domain, the energy of image is mainly focused on DC coefficients and AC coefficients in low frequency after DCT transforming, we can obtain good robustness by digital watermark be embedded into these components. The algorithm has better robustness on conventional image processing and resists attacks effectively through experimental comparison.

72.2 Color Space Selection

The research of digital watermarking technology for color image, the key problem is how to choose the color space to embed watermark because the original carrier image is color images. The color of color image are most described by RGB, but RGB and other similar color model cannot be adapted to practical to the sensitivity of people very well. Because the HSI color description is more natural and intuitive, more suitable for human visual properties, and in color and brightness of separation can deal with brightness in the images of the weight. So this article chooses HSI color space to embed the watermark in order to achieve both good robustness and make the color images with good visual effect.

72.3 Arnold Transform of Watermark

We need to finish the pretreatment before watermark embedding, this plays a key role in the robustness of the system, in the digital watermarking system security aspects of the role is also very prominent, it increases the difficulty of the attacker in the guessing attack. This article uses Arnold transform because of its briefness, convenient and periodic characteristic.

Images can be seen as a planar region of the dual function: $Z = F(x, y), (x, y) \in R$, region R is a rectangle usually, for any point(x, y) in R, $F(x, y)$ Representative gray value of the image. The dual function $Z = F(x, y)$is corresponding to a two-dimensional discrete lattice after the image is digitized. Elements in the matrix

Fig. 72.1 Original image watermark

represent image information. We can get the new matrix by doing some kind of transformation, in order to realize image scrambling. The image of the pixel position will rearrange after Arnold transformation assuming digital image pixel coordinates to be $(x, y) \in \{0, 1, 2, \ldots, M-1\}, M$ expresses the matrix order of digital image [3], Arnold transform defined as formula (72.1):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod \mathrm{M} \tag{72.1}$$

Image pixel (x', y') is the new pixel coordinates of image after transformation, mod M express modulo of x' and y', the purpose is to ensure that x' and y' in the range of $\{0, 1, 2, \ldots, M-1\}$. Arnold transformation is actually a point location of mobile; the pixel position of original image will rearrange watermark scrambling through Arnold transformation, and getting messy images, as in Figs. 72.1 and 72.2. Figure 72.1 is the original image watermark, Fig. 72.2 is the Arnold scrambling transform image, image space changed, but the pixel value and image size did not change.

72.4 Watermarking Embedding Position Analysis and Selection

Selection of watermark embedding positions is very important to ensure the system's invisibility and robustness. In order to make full use of the human eye to the frequency of the masking effect, we need process the image from spatial domain to transform domain.

72.4.1 DCT

2 d DCT transform the way to use the pixels of its operation in image processing. Because the whole realization of the image 2 d DCT transformation calculation is relatively large, in this article, the image region DCT transform. When we realize DCT transform, $N = 8$, the image is divided into 8×8 block, the main reason is the efficiency increases not much but the complexity increases a lot when N is greater than eight. Through the discrete cosine transform, the left corner of Images is the low frequency coefficient, coefficient value is larger.

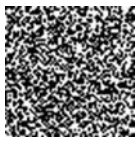


Fig. 72.2 Scrambling watermark image

72.4.2 Watermarking Embedding Coefficient Selection

In the DCT domain digital watermarking algorithm, in order to ensure the watermark has good imperceptibility and robustness, the embedded watermark DCT coefficient can still be better preserved through conventional image processing or attack, this method can guarantee the robustness of watermarking [4]. At the same time, we can ensure embedding strength of watermarking information based on visibility. This algorithm selects the robust frequency coefficients as being embedded watermark coefficient.

72.5 Digital Watermarking Realization Process Based on Arnold and DCT Technology

72.5.1 Pretreatment of Watermark

Set the original image is M * N 24 bits RGB image K, the embedded watermark $\frac{M}{8} \times \frac{N}{8}$ is a binary watermarking WR. When the watermark is about to be embedded, we need preprocess the watermark image first, realize the value sequence transformation:

(1) Do M times Arnold scrambling transformation on two value watermark image WR, in order to remove the watermark image's pixel space correlation, and enhance watermarking algorithm is robust and security. Here M is watermark image encryption key [5].
(2) Convert WR to {0,1} sequence, each element is defined as WR_k $\left(k = 1, 2, hellip;, \frac{M \times N}{8 \times 8}\right)$.
(3) "0" transforms into "−1" of *WR,* realize $WR_k \in \{-1,1\}$, so can remove the influence of rounding error.

72.5.2 Watermark Embedding Algorithm

(1) Conversion of color space
First of all, transform the original color images from RGB mode to K HSI mode, the algorithm uses the column conversion, conversion formulas such as (72.2):

Fig. 72.3 Original carrier image

Fig. 72.4 Embedded watermark image

$$I = \frac{1}{\sqrt{3}}(R+G+B); S = 1 - \frac{\sqrt{3}}{I}\min(R,G,B); H = \begin{cases} \theta, & \text{if } G \geq B \\ 2\pi - \theta, & \text{if } G < B \end{cases} \tag{72.2}$$

$$\theta = \cos^{-1}\left[\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right].$$

(2) Extracting luminance component I and realize block
Extracting luminance component I of image, the original image is divided into 8×8 image block that does not cover each other, get a block matrix by $\frac{M}{8} \times \frac{N}{8}$ block composition, and each block can be expressed into f_k^I(x, y)(x, y = 1, 2,…8, k = 1, 2,… $\frac{M \times N}{8 \times 8}$), each block f_k^I (x, y) corresponding to the watermark information WR_k of robust watermarking WR.

(3) DCT transform and coefficient quantization
In order to get the DCT coefficient matrix F_k^I (u,v), we do DCT transform each block f_k^I(x,y),and (u,v = 1,2,…8, k = 1,2,… $\frac{M \times N}{8 \times 8}$). Then we do Quantization processing. Treatment such as formula (72.3), among them F_k^I(u,v) is the coefficient that needs to quantify, β as the quantification factor.

$$B_K^I(u,v) = \left\lfloor \frac{F_K^I(u,v)}{\beta * Q(u,v)} + 0.5 \right\rfloor \ (u,v = 1,2\ldots 8) \tag{72.3}$$

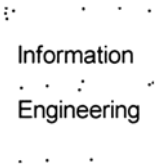


Fig. 72.5 Copyright watermark extracted

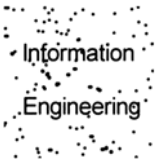


Fig. 72.6 The extracted watermark after 1×3 median filtering

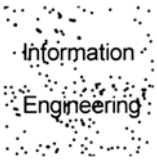


Fig. 72.7 The extracted watermark after 3×3 median filtering

(4) Robust watermarking embedding
Select the DC coefficient C_k^I (72.1) of embedding robust watermark information in each block, using plus embedded watermarking embedding way, the method of embedding such as formula (72.4) shown below. $c_k^I(1)$ is the DC coefficients of the original image. α is the embedding strength of watermark, its value will directly affect the effectiveness of the algorithm. After repeated test, and $\alpha = 0.79$ is more appropriate, we can combine the invisibility and robustness [6, 7].

$$C_k^{I'}(1) = C_k^I(1) + \alpha W R_k \tag{72.4}$$

72.6 Experimental Results and Performance Test

72.6.1 Performance Evaluation

We need an objective evaluation method to make quantitative evaluation of the algorithm. Generally, we use peak signal to noise ratio (PSNR) to evaluate image visual quality, PSNR is larger, and the original carrier image's damage is smaller. We use the normalized mean square error NC to evaluate the extracted watermark of the original watermark similarity, NC is larger, and both more similar. Peak signal-to-noise ratio PSNR are as defined in equation as formula (72.5):

Table 72.1 The extracted watermark effect after median filtering

Median filter	1×3	3×3	5×5
NC	0.981	0.963	0.913

Table 72.2 The extracted watermark after JPEG compression of different QF

QF	90	80	70
NC	0.973	0.956	0.922

$$\mathrm{PSNR} = 10\log_{10}\frac{w_{\max}'^2}{\frac{1}{M}\sum_{i=1}^{M}\left(w_i - w_i'\right)^2} \tag{72.5}$$

w_i' expresses the extracted watermark signal, w_i expressed the initial watermark signal, M expressed the length of the watermarking signal. The PSNR of the image is presented by formula (72.6).

$$NC = \frac{\sum_{i=1}^{M} w_i * w_i'}{\sqrt{\sum_{i=1}^{M} w_i^2} * \sqrt{\sum_{i=1}^{M} w_i'^2}} \tag{72.6}$$

72.6.2 Invisibility Test

In order to validate the efficiency of the method, algorithm was simulated based on MATLAB7.0. According to this algorithm, embedding robust watermark in the original image, and robust embedding strength $\alpha = 0.79$, Fig. 72.3 is the original carrier image, Fig. 72.4 is the embedded watermark image. Through testing, we did not find that the quality of the watermark image obviously drop, and did not find differences on the vision, confirmed the watermark image's Invisibility. Figure 72.5 is the embedded watermark image from the extracted watermark image.

72.6.3 Attack Resistance Test

(1) Median filtering
Figures 72.6 and 72.7 are the extracted watermark image superposition of different noise. After median filtering, the extracted watermark effect as shown in Table 72.1. Experiments show that the copyright watermark information has good ability of resistance to the median filtering.

Fig. 72.8 QF = 90

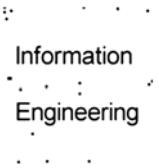


Fig. 72.9 QF = 80

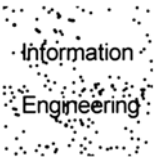


Fig. 72.10 QF = 70

(2) JPEG compression
Use quality factor 90, 80, 70, 60 on the picture to JPEG compression. Compression effect mainly use the QF (Quality Factor) to express. As shown in Figs. 72.8, 72.9 and 72.10 the quality factor QF is larger, the image compression ratio is smaller, the extracted watermark image effect is better. The watermark extraction effect through the different quality factor for JPEG compression is shown in Table 72.2.
Through the experiment, we found that after JPEG compression, can extract the watermark better, and reflect the watermarking robustness.

72.7 Conclusion

This article presents an algorithm based on Arnold scrambling and DCT transform digital watermarking. This article chooses 24 RGB model of color image into the original image, select the HIS color space to embed the watermark. In order to increase the decoding difficulty, the improved Arnold transform is used to encrypt watermark image. The watermark is embedded into the DCT domain in low frequency based on the original carrier image DCT transform. Through the contrast test shows that, this algorithm has good invisibility and resistance to attack, can effectively resist attacks such as filtering, compression, with strong robustness.

References

1. Liu S (2003) Trends and some problems of digital watermarking techniques based on information hiding. Comput Eng Appl 12:158–160
2. Gao JG, Fowler JE, Younan NH (2001) An image adaptive watermark based on redundant wavelet transform [C]. Processing of the IEEE international conference on image processing. Metaichmio press, Thessaloniki
3. Ding W (2001) Digital image scrambling technology based on Arnold transformation. J Comput Aided Design Comput Graph 13:338–341
4. Wang B (2003) Digital watermarking technology. Xian University of Electronic Science and Technology Press, Xian pp 4–7
5. Li J (2007) Digital image watermarking technology against geometric attacks [M]. Intellectual Property Press, pp 2–6
6. Tirkel A, Rankin G, van Schyndel R (1993) Electronic watermark. In: Proceedings of DICTA, pp 666–672
7. Yi K (2001) Digital watermarking techniques: an introductory review. J Image Graph 6(2): 111–117

Chapter 73
Research on Snort Intrusion Detection System and Key Match Algorithm

Guo-zheng Zhou and Jun-ya Li

Abstract Traditional encryption and firewall technology cannot fully meet the needs of information security, intrusion detection technology as a necessary means of security, network security plays in its unique role. Snort as a typical lightweight network intrusion detection system (NIDS) is a free open-source projects, design principles and implementation of Snort study of the characteristics can serve as the development of commercial intrusion detection system the cornerstone of a strong academic significance and higher commercial value. The architecture and workflow of Snort was analyzed and key match algorithm (BM algorithm) was studied. The research result has theoretical and practical significance on improvement and optimization of Snort and other intrusion detection systems.

Keywords Snort · Intrusion detection · Pattern match · BM algorithm

Computer and network technology to bring convenience to people, but also brings security issues that cannot be ignored. Computer viruses, hacker attacks, an unprecedented development and dissemination of information quickly, so that the status of network security is increasingly grim [1, 2]. IDS (Intrusion Detection, ID) technology as an effective complement to traditional security mechanisms, network security technology has become the main research directions. Snort is a misuse-based detection of network intrusion detection system, using C language, with open-source, cross-platform, lightweight features [3–5]. It is not only capable of protocol analysis, content retrieval, content matching, and can be used to detect buffer overflow,

G. Zhou (✉) · J. Li
Office of Academic Affairs, Jiyuan Vocational and Technical College,
Jiyuan 454650, Henan, China
e-mail: zgz.zhou@163.com

J. Li
e-mail: junya006@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_73,

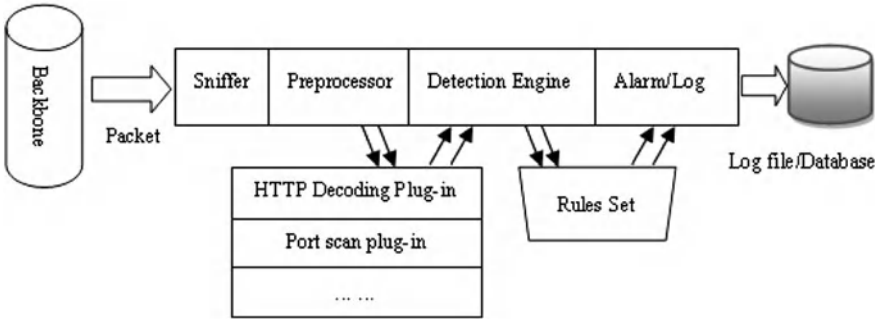


Fig. 73.1 Snort architecture

concealed port scans, CGI attacks, OS fingerprinting, a large number of attacks and probes, are widely used in the industry known as the "Swiss Army Knife."

73.1 Snort Architecture

Figure 73.1 shows the Snort architecture consists of four basic modules, namely, packet sniffer, preprocessor, the detection engine, the alarm output module [6].

(1) A packet sniffer. From the card using libpcap to capture network packets, and packet decode engine through the fill to the link layer protocol packet structure in order to high-level protocol decoding.
(2) Pre-processor. With the appropriate plug-ins (such as HTTP decoding plug-in) to analyze the data packets, from these data that the "conduct"—the original performance of the application layer packet. After pre-processing data packets transmitted after the detection engine.
(3) Detection engine. Snort is the core module. Based on different characteristics of rules and information packet to detect and plug-ins through the detection engine to provide additional detection.
(4) Alarm output module. If the detection engine of a rule is matched, it will trigger an alarm, the alarm information to a selected output means, such as syslog, MySQL and so on.

73.2 Snort Workflow

Snort entry main () function defined in the file snort.c. main () function for a simple command-line parameters and then calls a function to determine SnortMain () [7, 8]. SnortMain () constitutes the main body of the whole system, describes the whole system workflow.

(1) Command-line argument parsing, and other initialization routines. Snort contains a number of important global variables and data structures. Which the PV (defined in the file snort.h) and the Packet (defined in the file decode.h in) two structures of the most massive and important. PV system configuration includes a variety of information and identification field. Function ParseCmdLine () for parsing command-line parameters will enter the results of PV. Function fpInitDetectionEngine () initializes the detection engine, to specify the speed rule matching module's default configuration parameters, including the default mode search algorithm (BM algorithm) and so on. Function InitDecoderFlags () initialization protocol parser responsible for parsing the agreement specify alarm information generated in the process of the type of error.
(2) Call the function OpenPcap () initialize pcap and network interface devices.
(3) Initialization of the plug-in. Call the function InitOutputPlugins (), InitPreprocessors () and InitPlugins () respectively output plug-in, pre-plug-in, testing plug-in initialized.
(4) Read the rules, create a rule list. Function CreateDefaultRules () to establish the framework of the rules list, the function ParseRulesFile () reads the configuration file snort.conf file *. rules and rules for filling the rules list, input and output plug-ins.
(5) Fast rule matching optimization. Call the function fpCreateFastPacketDetection () read all the rules in the rule list list head (RuleTreeNode type) and the rule options node (OptTreeNode type), built on its new layer of data structure, so for fast rule matching purposes.
(6) Call the function InterfaceThread () into the packet processing loop. The function is very simple, the main call pcap library pcap_loop (), and set the packet handler callback ProcessPacket ().Interface function ProcessPacket () for each packet is assigned a Packet structure that records the levels of all agreements and other information required test data, the process of decomposition of the agreement, including the structure of the various fields in different stage is properly filled. And then determine if the system in intrusion detection mode, the pre-processing module is invoked Preprocess ().Interface function Preprocess () through pre-processing plug-in list PreprocessList to achieve pre-processing, detection module is called Detect (). fpEvalPacket () is the main interface function of rapid detection rules, which Detect () is called. The function level based on the current data packet protocol type (TCP, UDP or ICMP), respectively, call fpEvalHeaderTcp (), fpEvalHeaderUdp (), fpEvalHeaderIcmp () for fast matching rules, only if the current packet is an IP packet, call fpEvalHeaderIp (). The fpEvalHeaderXXX () function in different subset of the rules match the rules of the implementation of specific tasks. After the success of matches into the alarm module, order through the output plug-ins list, trigger an alarm or log an event. The entire packet processing shown in Fig. 73.2.

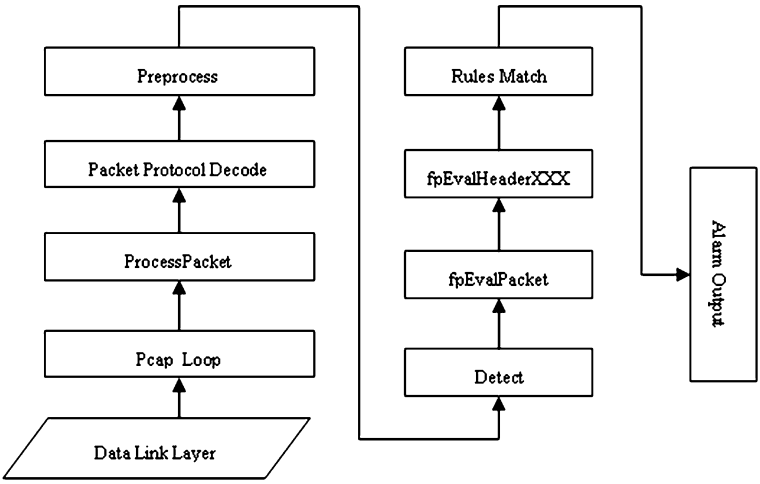


Fig. 73.2 Packet processing

73.3 BM Algorithm

Snort uses rule-based pattern matching to detect attacks. Pattern matching refers to the target text string to find a substring, so that the pattern string with a known pattern is equal [9, 10]. According to experimental statistics: Snort detected during the pattern matching functions is called when the time is needed to detect the total time of 25.2%. Therefore, the efficiency of pattern matching algorithm of Snort is directly related to operating efficiency.

Boyer-Moore (BM) algorithm is a kind of sub-linear time string matching algorithm. The algorithm is widely used in IDS, Snort can use this algorithm for pattern matching, the code contained in the file mstring.c.

BM algorithm's basic idea is: starting from the right side of string pattern matching, more often than from the left to get more inspiration. Therefore, the algorithm first goal of the mode string pattern and string of string left-justified, then the right-most characters from the pattern started to the left compared with the corresponding character in the string. When characters in a pattern corresponding with the character string b mismatch at the same time use two kinds of heuristic rules, rules that the bad character and good suffix rules to determine the distance to the right, which grew up under the rules of moving step, which was adopted.

(1) Bad character rules:

i. If b is not included in the pattern string pattern, then the string from the b, length strlen (pattern) of the substring and pattern matching cannot be successful, skip directly to the region.

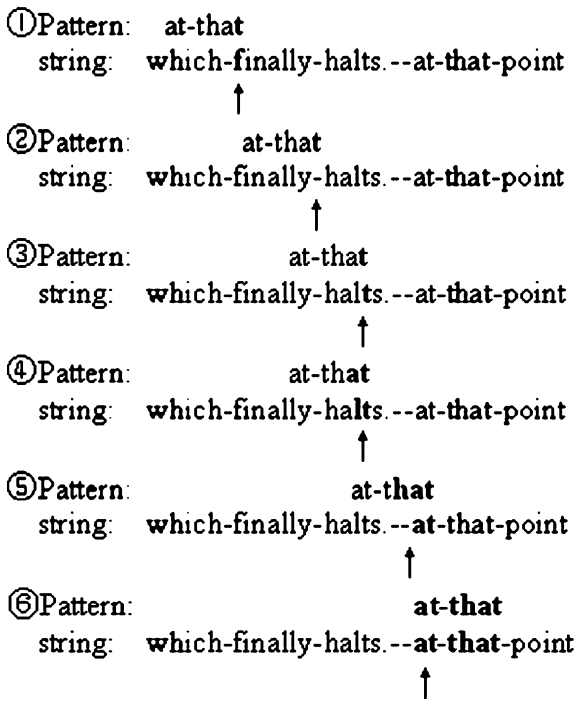


Fig. 73.3 BM algorithm matches the sample

ii. If b is included in the pattern, and in a left, then the pattern of the b and b in the string alignment (left more than one if a b, choose the right b); if the right of b in a, then the pattern shifts one character.

(2) Good suffix rule:
In a and b mismatch at the same time, there are a length l substring s match, according to the following two cases to discuss:

i. If the pattern is the same in the presence of other sub-string and s t, and t s the first character and the previous character (i.e, a) is not the same, then the pattern shifted to the right so that t s alignment with the string (if more than one t, choose the right t).
ii. If the pattern does not exist in this sub-string t, and there is a longest prefix pattern t 'is equal to s, a suffix, the pattern shifted to the right to t' and s, corresponding suffix alignment.

The following brief description of an example of the main steps BM algorithm. Pattern in Fig. 73.3, said pattern string "at-that", string the target text, namely in the string to find the pattern.

(1) The pattern with the string of left-justified, right-most characters from the beginning of pattern and string comparison. Because the "f" does not appear in the pattern, by the bad character rule i, the pointer right skip strlen (pattern) = 7 characters.

(2) Every time right after starting from the rightmost pattern comparison. "-" and "t" range, but this character "-" appears in the pattern, by the bad character rule ii, the pattern pointer in the right "-" and in the string "-" alignment.
(3) The right-most two characters match, the pointer left one character.
(4) "L" and "a" range, and l is not included in the pattern, by the bad character rule i, the pointer right seven characters.
(5) The first two characters match, move the pointer left "-" and the pattern in the "h" mismatch. According to the bad character rule ii, the pointer should be shifted to the right four characters, so that pattern in the "-" and in the string "-" alignment; According to the good suffix rule i, the pointer should be right seven characters, so that pattern in another sub-string "at" with the string in the "at" alignment. Should choose which one? Choose to move further grow that! BM algorithm is the key.
(6) Each character corresponds to equal matches.

In the file mstring.c, we can see the definition of BM algorithm Snort-related functions where the function make_skip () and make_shift () mode string of bad character rule and apply a good suffix rule pretreatment, resulting in bad character and good suffix table. Function mSearch (), mSearchCI () and mSearchREG () are the standard BM algorithm, the BM algorithm is not case sensitive and supports wildcards of BM algorithm.

BM algorithm best-case time complexity is O (n/m), worst-case time complexity is O (m * n), in which n = strlen (string), m = strlen (pattern).

73.4 Conclusion

Network intrusion detection system is an important defense in depth system component, and is seen as second line of defense behind the firewall. Snort intrusion detection system using a rules-based pattern matching to detect attacks, the BM algorithm is considered to be the fastest character matching algorithm. However, with increasing speed of computer networks, pattern matching speed must also be a corresponding increase in research faster and more efficient algorithms is necessary. And intrusion detection based on pattern matching cannot meet the increasingly diversified and new intrusion detection, intrusion detection system, thus changing the existing structures, and neural networks, genetic algorithms, data mining and other methods combined, the future research directions.

References

1. Caswell B, Beale J, Foster JC, Posluns J (2004) Snort 2.0 intrusion detection. National Defence Industry Press, Beijing, pp 26–29
2. Lan J, Wang Y (2008) Snort research and improved BM algorithm. Comput Eng Design 29(9):5–6

3. Song J (2004) Network intrusion detection. National Defence Industry Press, Beijing, pp 116–117
4. Boyer RS, Strother J Moore A (1997) Fast string searching algorithm. Commun ACM 20(10):762–772
5. Li H (2005) Based on Snort system-specific string matching algorithm. Harbin Institute of Technology, Harbin
6. Yang W, Liu X (2006) An improved BM pattern matching algorithms. Comput Appl (2):156–162
7. Caswell B, Beale J, Foster JC, Poslum J (2001) Snort2.0 instrusion detection. Syngress Pubishing Inc., Boston, pp 118–122
8. Lou Z, Zhang T (2007) Snort network intrusion detection and its application. Softw Guide (13):156–161
9. Hu J, Zuo M (2007) Snort intrusion detection rules based on matching study. Comput Secur 2:073–078
10. Wu F (2010) Snort intrusion detection system with improved BM algorithm [D], vol 10. Southwest Jiaotong University, Chengdu, pp 042–046

Chapter 74
Research on DoS Attacks and Resist Method Based on 4-way Handshake in 802.11i

Jun-ya Li and Yan Yang

Abstract As the transmission medium is open, the use of wireless LAN communication requires higher performance and better security. By analyzing the latest generation of wireless LAN security protocols in 802.11i, four-way handshake agreement was introduced, two against the DoS attacks. The first four-way handshake based on fake news of a DoS attack was initiated, proposing added authentication solution. Robust security on the second associated information was proposed by modifying the Robust Security Network Information Element (RSN IE) elements in the important position and launching DoS attacks on the proposed four-way handshake process to improve it. Through analysis, the proposed method can effectively prevent DoS attacks from occurring, to a greater extent to improve the effectiveness of wireless data transmission.

Keywords 802.11i · 4-way handshake · Message integrity code · Robust security network information element

Wireless LAN with its fast, cheap, ease, and many other advantages of networking in enterprises and public hot spots within the areas of application have made great strides. However, due to the openness of its transmission medium, it also determines that the wireless LAN security is facing a major problem [1]. To enhance security, the IEEE 802.11 committee has developed the 802.11i protocol [2].

J. Li (✉)
Office of Academic Affairs, Jiyuan Vocational and Technical College, Jiyuan 454650, Henan, China
e-mail: junya006@163.com

Y. Yang
Department and Art Design, Jiyuan Vocational and Technical College, Jiyuan 454650, Henan, China
e-mail: wyzzxx@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_74,

This paper analyzes the 4-way handshake in 802.11i protocol, two against the denial of service attacks, and the corresponding improved method.

74.1 IEEE 802.11i Summary

The IEEE in June 2004 approved the WLAN security standard 802.11i. The IEEE 802.11i robust security network proposed the joint Robust Security Network Association (RSNA) concept, RSNA provides Temporal Key Integrity Protocol (TKIP) and Counter-Mode/CBC-MAC Protocol (CCMP), which are two encryption protocols, including TKIP to ensure forward compatibility and remain in the 802.11i encryption protocol in the interim. Still based on RC4 encryption algorithm, you can upgrade the existing firmware and device drivers methods to improve WLAN security. The CCMP is based on Advanced Encryption Standard (AES) encryption technology, next-generation Advanced Encryption Standar. The greatly increased level of WLAN security is to achieve the mandatory requirements of RSN. 802.11i uses 802.1X [3] mechanism to achieve user authentication and key distribution. 802.1X port-based access control mechanism. The 802.IX work is an Extensible Authentication Protocol (EAP) [4] protocol, which is responsible for authentication transmission define message format and transmission rules [5]. Mechanism through 802.1X EAP authentication protocol supports the user after the client STA (Station) and the authentication server will also have a Master Session Key (MSK). MSK will be passed to the authentication server by the client and the authenticator after a four-way handshake process of negotiation of reciprocal temporary key Pairwise Transient Key (PTK) to encrypt data transmitted between the two [6].

The RSNA establishment process consists of the client STA (Station), certified by the access point Access Point (AP) Authentication Server (AS) with three entities. The RSNA establishment procedure shown in Fig. 74.1, SPA and AA on behalf of STA and the AP's MAC address, SNounce and ANounce represent the STA and the AP generates a random number, SN on behalf of serial number, msg1, 2,3,4 representing different message types. MIC is the message integrity check code; messages 2,3,4 in the MIC check is carried out based on PTK.

The RSNA building process can be divided into the following six stages [7].

The discovery phase of the network and security capabilities.

This stage by the news (1) to the message (3) the composition, AP periodically broadcasts Beacon frame by or sent to the STA's Probe Request frames (two frames contains a description of the ability of AP RSN IE security field) response Probe Response frame to inform the STA of its security capabilities. STA may be listening passively through the initiative to send Beacon frames or Probe Request frame to get the AP's security capabilities.

802.11 authentication and association phases.

This stage by the news (4) to the message (7) the composition, STA from the available list, selects an AP and try to establish certification and associate with.

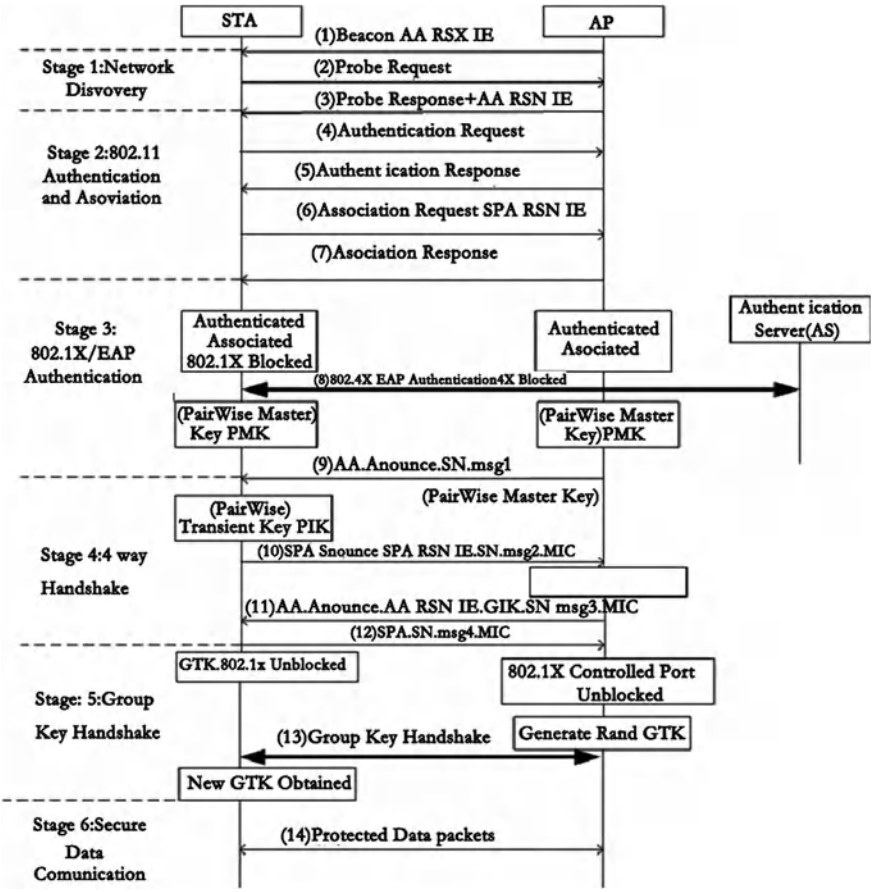


Fig. 74.1 RSNA building process

The default to select the strongest signal strength to connect. STA should be specified in the Association Request frame and its security capabilities. After completion of this phase STA and AP were certified and associated with status. But 802.1X port is still closed and not allowed to send any packets.

EAP/802.1X certification stage.

After completion of this phase STA and AP will share the key between the two so that the master key of the Pairwise Master key (PMK) in this process is the figure of the message (8), because the focus of this paper is to analyze the four handshake protocol. This process introduces omitted details, please see [2].

4-way handshake phase.

This stage by the news (9) to the message (12) composition, STA and AP, confirm each other using four-way handshake that it has received PMK, PMK

generated by using the message encryption, integrity checks and other security measures in the key. After the four-way handshake, 802.1X port is open, STA and AP can exchange data between themselves. Therefore, the 4-way handshake protocol in the 802.11i standard plays a vital role.

Group key handshake phase.

This phase diagram in the news (13), taking into account the multicast problem, 802.11i defines a set of keys at this stage to complete the Group Transient Key (GTK) handshake and distribution.

Data security communication phase.

At this stage is the message in Fig. 74.1 (14). STA and AP use of PTK or GTK encrypted transmission of data between each other.

Through the above six stages, STA and AP achieve certification for each other and establish a data transmission to protect the security association.

74.2 Four-way Handshake Protocol Analysis and DoS Attacks

If the STA and AP establish a shared PMK, AP will launch a four-way handshake process. The purpose is to confirm the STA side held by the PMK and they are the same, while the STA and AP were used to generate the PMK after PTK for both data encryption and message authentication [8]. The four-way handshake process the message format is as follows:

Message 1: AP-> STA
[AA, ANonce, SN, msg1]
Message 2: STA-> AP
[SPA, SNonce, SPA RSN IE, SN, msg2, MIC]
Message 3: AP-> STA
[AA, ANonce, AA RSN IE, SN +1, msg3, MIC]
Message 4: STA-> AP
[SPA, SNonce, SN +1, msg4, MIC]

The AP sends messages to STA 1 start a handshake process four times. This message includes PTK used to produce the random number ANonce, receive the news after 1 get STA ANonce, use the sharing of PMK and generate its own, with the SNonce PRF function, and then calculate the PTK to AP sending messages 2. Contains SNonce and STA RSN IE, and with the new calculated news for MIC PTK 2 authentication. AP receives news after using PMK, 2, ANonce and get SNonce generation, and of news PTK PTK 2 for MIC check. If calibration failure will discard the news, the correct AP will continue to check whether the STA of RSN IE before the frame and Association agreement, if not RSN IE consistent AP will stop with STA connection. If agreed, messages are sent to STA 3. Message 3 contains requirements into STA PTK and AP's RSN IE, which receives the news

Element ID 1 octet	Length 1 octet	Group Cipher Suite 4 octets	Group Cipher Suite 4 octets	Pairwise Cipher Suite list 4-m octets	AKM Suite Count 2 octets	AKM Suite list 4-n octets	RSN Capabilities 2 octets	PMKID Count 2 octets	PMKID List 16-s octets

Fig. 74.2 RSN IE format

after 3 STA judgment of whether ANonce of first with news of the same ANonce 1. If different, it directly discards the news, if the same AP continues to check whether the news RSN IE and received by the Probe Beacon and frame RSN IE consistent, if not consistent with the AP it will suspend STA connection, if consistent STA will continue this message for MIC check. If calibration failure it will discard the news and the right of the AP of sending a message to 4. Has PTK load. AP news in received four and calibration is correct with PTK. So far 4-way handshakes process finished.

After a four-way handshake process of analysis and found that there are two possible DoS attacks. The first is due to message 1 transmitted in clear text, an attacker can easily be forged. An attacker can send message 2 after message 3 before sending bogus AP STA to send fake messages to 1, then the STA will be based on the forged Message 1 to recalculate a PTK′. As the generated PTK′ used ANonce and SNonce calculated with PTK is different from before, it is clear that PTK′ to the AP after receiving a message generated by PTK 2 is inconsistent. This STA will receive the message 3 MIC validation failure due to drop, creating a DoS attack.

In addition to the proposed 802.11i standard STA storage PTK, the need to also store a Temporary PTK (TPTK). STA only 3 in the right when the message receives PTK updates. This inconsistency can be avoided due to PTK launching DoS attacks. But the same program has a lot of deficiencies. Because the attacker can forge an infinite number of messages sent to STA, the STA receives a message every 1 to produce a SNonce, and calculate the TPTK, and then assign the corresponding memory space. STA 3 receives the correct message for each TPTK and will try to verify the MIC, which is likely to result in the CPU and memory exhaustion STA, the formation of new DoS attacks (Fig. 74.2).

Another four times in the DoS attacks are shaking hands agreement using RSN IE verification mechanism of the flaw of the launched. RSN IE format as below.

AP in Beacon Probe and send 5 frames, send STA (Re) Association that will include their frames RSN IE. STA and AP consultation with good safety group executive authentication and key management, using good password consultation agreement between each other groups to encrypt data transmission. In order to verify the authenticity of the requirements, IE RSN STA in four times in the news of the process to shake hands with 2 contains Association (Re) that the same frame RSN IE. Requirements in four times the AP news in the process to shake hands with Beacon and contains 3 5 in the same frame Probe RSN IE. After received the news in 2, AP will be checked by bit of RSN IE and whether (Re) Association in

the same frame that RSN IE. After received the news in 3, STA will also check the RSN IE and Beacon and Probe whether the same RSN IE 5 frames. If RSN IE not equal, the AP field and STA will lift the each other between authentication and will record a security mistakes. This confirmation mechanism can stop the attacker by forging RSN IE deceive STA and AP make them the weak security mechanism, but this mechanism also exist DoS attack holes.

The attacker can through the protected reliably against detective Beacon frames, modify one of the not important, not a is the meaning of these important a revision will not affect the legitimacy of the frame and the authentication protocol group selection. For example, keep a replay count a, these a RSN ability in field is not important. The attacker through the radio modified not important a fake Beacon frame to change in the AP STA storage of RSN IE. Because of the change is not important a so STA and AP still can use between the effective safety groups for the authentication and key management. However, because RSN IE not through four times to authentication handshake agreement can never succeed. If the use is active scan STA way, not passive way found the wireless network, the attacker can be sent through modified RSN IE forge Probe to influence the process, 5 shake hands formed DoS attack.

74.3 The Improved Method of Four-way Handshaking Protocols

For the first message a forgery initiated by DoS attacks, we propose to add 1 in the message authentication of the AP, allowing the attacker to not be forged and to prevent DoS attack purposes. This proposed solution based on the fact that in the 802.11i 4-way handshake upon completion of the third stage before the start of the process, AP and STA have a shared between the key PMK. AP can send a message 1, using the PMK as a MIC key, the message in order to achieve a certification. The 4-way handshake process ensures smooth progress.

In order to reduce the news to the format of the modified 1, we use the news of the serial number 1 encrypt SN, such as a MIC when STA receives news after 1, which will be encrypted SN after decryption and unencrypted SN were compared, agreement, receive news, and that this news 1 from legal AP.

Because in the original news 1 no MIC, so STA received news after the 1, will not update the Replay class Counter: (Replay), news 1 count field in the MIC fields, received the news if the MIC STA validation success will update its Replay, so that even if the attacker class Counter: won the SN plaintext/ciphertext pairs, due to Replay class Counter: cannot pass verification field, and cannot use the SN Ming ciphertext get to Replay.

For the second RSN IE by manipulating the important bits and launch DoS attacks, we proposed four-way handshake protocol for authentication on the RSN IE bit by bit from the verification, to ignore one of the important bits and only the

RSN IE in the authentication and key management group selector for verification. Because in the RSN IE, only the authentication and key management group selector for the handshake after the process is necessary. AP and the STA after the completion of certification and be able to share a secret key encryption protocol group consultation, if the attacker does not change the authentication and key management group of selectors, the RSN IE is authenticated. If an attacker to change the authentication and key management group of selectors, in the beginning when the connection can be detected, and STA will cause the connection fails immediately open a new connection process.

74.4 Conclusion

In this paper, the four-way handshake in 802.11i protocol analysis is discussed, pointing out the existence of security vulnerabilities, based on false information for a DoS attack launched by the MIC field for membership, to prevent the fake attackers. By modifying the RSN IE in the least bit of DoS attacks launched by four-way handshake protocol for change RSN IE authentication strategies to prevent attacks. The above improvements to the four-way handshake protocol can effectively prevent DoS attacks against them to protect the wireless LAN security.

References

1. Zhou J, Liu D, Yu YP (2008) IEEE802.11 Wireless LAN technology and security analysis. Netw Security Technol Appl (2):15–17
2. IEEE Std 80211i. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification Amendment 6: Medium Access Control (MAC) security enhancements, pp. 521–527
3. IEEE Std 802. Ix (2001) IEEE standards for local and metropolitan area network: Port Based Network Access Control, pp 97–102
4. Blunk L, Vollbrecht J (1998) PPP Extensible Authentication Protocol (EAP). RFC2284, Issue3, pp 55–59
5. Nai L (2004) Wireless LAN (WLAN) eleven principles, techniques and applications [M]. Xidian University Press, Xi'an
6. Edney J, Arbaugh WA (2006) Wireless LAN security practices—WPA and 802.11i [M]. In: Zhou Zheng M (ed) People's Posts and Telecommunications Press, Beijing
7. He C, Mitchell JC (2005) Security analysis and improvements for IEEE 802.11i [C]. Proceedings of the twelfth annual network and distributed system security symposium (NDSS'05) pp 90–110
8. He C, Mitchell JC (2004) Analysis of the 802.111 4-way handshake. In: Proceedings of the ACM workshop on wireless security (WiSe '04) [C], Philadelphia, USA, pp 43–50

Chapter 75
Implementation of SNMP-Based Network Management System's MIB for Dual-Stack

Lishen Yang and Qiansheng Meng

Abstract IPV4 to IPV6 transition is an inevitable trend that has been widely recognized; however, the transitions of IPV4 to IPV6 is a long process, and double-stack technology will play an important role in this process. During this period, the dual-stack network management is very important. In the platform of SNMP management, through access to MIB objects the network management purposes is achieved. Therefore, the MIB's access is the key to the network management. MIB was analyzed in this paper, and access to MIB simple variable with JAVA code is given.

Keywords Dual stack · Network management · IPV4 · IPV6 · SNMP

75.1 Introduction

With the development of IP networks, IPv4 cannot meet the needs of the expanding and emerging network applications, such as address depletion, the rapid expansion of the routing table, lack of security, poor mobility and so on. The emergence of IPv6 [1] solved the above problems. However, IPv4's design has successfully created a global Internet with structure of different hardware, different operating systems, interconnection of different network systems and to accommodate the scale of the network to grow exponentially over the past decade. The wide-ranging implications of IPv4 network will lead to a long period of coexistence between the two.

The transition from IPv4 to IPv6 is a gradual evolution process. During this period, there are many complex issues related to IPv6 transition. Among them, the

L. Yang · Q. Meng (✉)
School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: mengqiansheng@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_75,

network management is one aspect greatly affected by the transition. Interim network management research focuses on how to achieve a smooth transition to IPv6 networks under the existing IPv4 network environment, and the original IPv4 applications have not been altered in this case to achieve free use under the IPv6 environment and with IPv6 applications communicate with each other.

75.2 SNMP

Simple Network Management Protocol (SNMP) [1] is the most commonly used network management protocol. It was raised by the Internet Engineering Task Force research group, in order to solve the management issues on the Internet routers, in 1988. The predecessor of SNMP is Simple Gateway Monitoring Protocol (SGMP), it become the SNMP by great changes. From its release it has been widely used and supported because of the effect of SNMP in network management. At present, SNMP has become the de facto industry standard, and most network management systems and platforms are based on SNMP.

With SNMP development, SNMPv1 [2], SNMPv2C and SNMPv3 evolved. SNMPv1 had two major disadvantages: firstly, management is imperfect and inefficient; secondly, the lack of effective security mechanisms. Then IETF released SNMPv2 in 1993. SNMPv2 by extending the data types, the type of protocol operations and other methods not only enhanced management capabilities, but also put forward solutions in terms of security. On the basis of previous version, SNMPv3 added safety factor to ensure that SNMP itself increased security and remote configuration capability.

In fact, SNMP refers to a series of network management standard collection, which is formed by a series of agreements and norms, including the agreement itself, the definition of data structures and some related concepts. SNMP network management framework consists of three parts: Structure of Management Information (SMI), Management Information Base (MIB) and SNMP [3]. SMI is used to define the rules that accessible object by Network Management Protocol, the data types used by MIB and the network resources of naming and said. MIB is a collection of the structure of managed objects maintained by the equipment. Each object has an object mark with particular object type, the object mark only marks the object of an object value. SNMP is the communication protocol between the management process and the agent process.

75.3 IPV6 AND IPV4

75.3.1 IPV6 Appearance

IPv6 is to solve current Internet problems and birth. Existing IPv4 network lurks two crises: the address depletion and rapid expansion of the routing table. Inheriting the advantages of IPv4, IPv6 have drastic changes and features to expand according to

the experience of IPv4 running over the years, and its capacity is more powerful than IPv4, and is more efficient. IPv6 is the one to cause least controversy compared to other technical concepts after the emergence the Internet development process.People have a consensus: IPv6 to replace IPv4 is an inevitable trend.

75.3.2 Compared With IPv4, IPv6 has the Following Advantages

75.3.2.1 IPv6 has a Larger Address Space

IPv6 was originally designed mainly to solve the shortage of IPv4 addresses, IPv4 address-specified length is 32 bits, that is, 232-1 addresses; and that of the IPv6 address is 128 bits in length, that addresses 2128-1, almost provides address unlimitedly. According to conservative estimates of the actual allocation, IPv6 can be allocated more than 1,000 addresses per square meter on the whole planet.

75.3.2.2 IPv6 Uses Smaller Routing Tables

IPv6 address allocation follows aggregation principle, which makes the routers a subnet by an entry in the routing table, reduces the length of the routing table in the router and raises the rate of router transmit packet.

75.3.2.3 IPv6 Adds the Support of Enhanced Multicast and the Support of Flow

Flow Control that may contribute to the rapid processing of real-time services in the source and the destination and conducive to the business terminal with low performance support voice, video and other IPv6 applications and provides Quality of Service (QoS) control with a good networking platform.

75.3.2.4 IPv6 Plug and Play Ease of Management Capacity

IPv6 added to the support of Auto Configuration. This is the DHCP of improvement and expansion, making the network (especially LAN) management more convenient and faster.

75.3.2.5 IPv6 has a Higher Security

The user of the IPv6 network can encrypt the data of network layer and verify the IP packet, which greatly enhances network security.

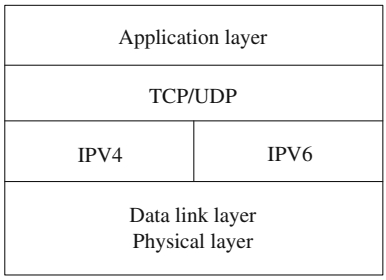


Fig. 75.1 Dual-stack TCP/IP architecture

75.4 Dual Stack

IPv6 has the same function as IPV4 in the network layer. The two protocols are based on the same platform. For computer network TCP/IP architecture, if a network node (host or router) supports the IPv6 and is compatible with the IPv4, then the node can not only communicate with the host of support IPv4, but also can communicate with the host of support IPv6. We call the node as "IPv4/IPv6 nodes" that is the dual-stack node. Dual-stack technology means a single node to support both IPv4 and IPv6 stack, which can be communicated with two types of node. Dual-stack node TCP/IP architecture is shown Fig. 75.1. Dual-stack hosts communicate with IPv6 hosts to use IPv6 addresses, while communicate with IPv4 hosts is to use IPv4 addresses. Dual-stack hosts can query the DNS to know what kind of destination the host addresses. If the DNS returns an IPv4 address, the source uses IPv4 stack in the help of 4 over 6. But when the DNS returns an IPv6 address, the source host uses IPv6 stack. IPv6 node to access IPv4 node, at first application for a temporary IPv4 addresses from dual-stack server, while gets the IPv6 address of the gateway router's tunnel endpoint from dual-stack server. Then IPv6 nodes form a 4 over 6 IP packet, the 4 over 6 packet passed through the IPv6 network to gateway router, gateway router removes the IPv6 header and IPv4 packets will be sent to the IPv4 node through IPv4 network. Gateway router is used to remember the correspondence between IPv4 temporary address and IPv6 source address, so that the IP packet from the IPv4 node forwards to the IPv6 node.

Dual-Stack technology is an effective IPv4 to IPv6 transition technique. Nodes in the network support both IPv4 and IPv6 stack; the source node selects different protocol stacks according to the destination node, and network equipment selects different protocol stack processing and forwards the packet according to the type of packet; there is no need to purchase IPv6 routers, thus saving hardware investment. Among the core IPv6 router use special local link to overcome many shortcomings of the tunnel. Dual-stack mechanism is the most direct way in which IPv6 nodes are compatible with IPv4 nodes, interoperability and easy to understand. In the early stages of IPv6 network construction, due to relatively ample number of IPv4 addresses, the implementation of this program was feasible. In terms of hosts and routers, a successful migration to IPv6 is to support dual-stack operation approach; although the allocation mechanism may be different,

dual stack node address can also be configured with two versions. The main benefit of Dual-stack host is that they minimize the NAT devices of network. Because IPv6 and IPv4 are similar network layer protocol in functionality, based on the same physical platform, and the transport layer protocols (TCP and UDP) of load on it are basically same, the dual stack node communicates both with the IPv4 node and the IPv6 node.

75.5 MIBs Implementation

In the system management module, a description of the network resource information is extremely important. MIB as the network management of managed resources is expressed in the form of objects, each object is one aspect of the managed resource properties, and a collection of the same aspects of objects form a MIB module. Each MIB object or variable records the network status, traffic data, device configuration information, the number of errors occurred and the current contents of the internal data structures and other information. Network manager achieves five network management functions through accessing the MIB.

75.5.1 MIB Definition

MIB objects are defined using ASN.1 language. ASN.1 is a language which describes the structure and content of the structured object. SMI provides object information coding, Basic Encoding Rules (BER) and thus unity encoding the object information, a description of the data has nothing to do with the system and manufacturers. MIB object definition is specified by SMI, it is used to define types and values. With ASN.1 it is described as follows [4]:

OBJECTNAME OBJECT-TYPE
DESCRIPTION:(description)
SYNTAX:(syntax)
MAX-ACCESS/ACCESS:(access)
STATUS:(status)
:: = {(Parent)number}
OBJECTNAME is the name of managed object, and unique;
OBJECT-TYPE is keyword of node object;
SYNTAX is type keyword of managed object. (syntax) is the object type;

MAX-ACCESS/ACCESS is access method keyword of managed object. MAX-ACCESS clause is similar to the ACCESS clause. The ACCESS clause uses the keyword of access method in SNMPv1, while MAX-ACCESS [5] clause uses the keyword of access method in SNMPv2. (Access) is the access method of a managed object.

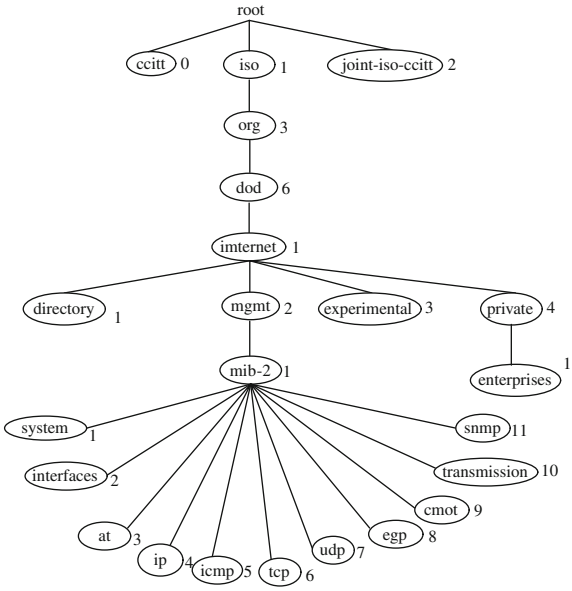


Fig. 75.2 MIB tree structure

STATUS is status keywords of managed object, (status) is the status of a managed object.

DESCRIPTION is description keywords of managed object, (description) is the description of a managed object. Parent represents the parent node in the MIB tree, and number represents child nodes.

75.5.2 MIB Tree

Object identifier (OID) represents each MIB object uniquely. It is a set of numbers and other number integers in the middle with a dot-separated; it represents the node's exact location in the ASN.1 tree. Each MIB object is a node with a label, each node is with two parts: numbers and text descriptions [4]. Text description describes the node with a label. A node with a label can have other nodes with label in its subtree; if not, it is the leaf node, also known as objects. Each object is represented by the sequence of the label on the path from the root to the object node uniquely (Fig. 75.2).

75.5.3 MIBs Required to Achieve

For dual-stack, MIB is needed to support separated management of IPv6 and IPv4 defined in 1998 (RFC2452/2454/2465) or after the definition of the unified

management MIB (RFC2851, RFC3291). For SNMP, the MIB is needed to support the three versions SNMPv3, SNMPv2 and SNMPv1.

75.5.4 MIB Access Implementation

75.5.4.1 SNMP Operations

SNMP provides five types of PDU (protocol data unit); the SNMP message is used to interact between the management process and the agent process. The primitives of SNMP network management operations are:

Get-request: from the agent process get one or more parameter values.

Get-next-request: from the agent process get the next parameter value followed by the current parameter values. It is used to traverse a set of parameters of the object in network device MIB.

Set-request: set the proxy process, one or more parameter values.

Get-response: Return one or more parameter values. This operation is sent from the agent process, which is response operation of the front three operations.

Trap: Proxy process unsolicited messages, reporting the occurrence of an abnormal situation to the manage the process.

Previous three kinds of operation are management process sent to the agent process, the others are the agent process sent to the management process. In the SNMP operation, generally, the agent process uses port 161 accepted message of Get-request, Get-next-request or Set-request, and the management process uses port 162 to accept Trap message.

75.5.4.2 Program Implementation Access to MIBs

The following describes a program which accesses the variable ipAdEntAddr in MIBs. The program uses the java language, developed using the open source ObjectSNMP components. Readers can map out the IPV6 address table and other tables and variable operation corresponding to the MIB according to ObjectSNMP development component description.

```
import com.zhtelecom.common.snmp.SNMPAPI;
import com.zhtelecom.common.snmp.SNMPFactory;
import com.zhtelecom.common.snmp.SNMPNodeParam;
import com.zhtelecom.common.snmp.test.ipAddrTable;
public static void main(String args[])

  {
  try{

SNMPFactory.init(SNMPFactory.Mode_Server_Local, null);
```

```
snmpapi = SNMPFactory.getSNMPAPI();
nodeParam = new SNMPNodeParam();
nodeParam.nodeIP = ''122.206.79.205";
nodeParam.readCommunity = ''public";
java.util.List list = snmpapi.getAllTableData(ipAddrTable.class,nodeParam);
System.out.println("…… Network address number:" + list.size());
  for (int i = 0; i < list.size(); i ++)
{ipAddrTable tcpAddress = (ipAddrTable) list.get(i);
System.out.println(i + " IP address:" +tcpAddress.getIpAdEntAddr());}
  } catch (Exception ex)
  {ex.printStackTrace();}
  }
```

75.6 Conclusion

Management MIBs are key aspects of network management. In order to better play network management system good access to MIBs are required.This achieved the variable access of the dual-stack based on SNMP MIBs, and discussed the IPV4 and IPV6 dual-stack issues.

Acknowledgments This paper is supported by Henan Polytechnic University Postgraduate Dissertations Foundation. The author would like to thank the Editor for helpful comments that greatly improved paper.

References

1. Deering S, Hinden R (1998) Internet protocol, Version6 (IPv6) Specification. RFC 2460
2. Case J, Fedor M, Schoffstall M, Davin J (1990) Simple network management protocol (SNMP). RFC 1157
3. Pan B, Zhan X (2008) Solution for IPv6 Network Management Based on SNMP
4. Luo Y (2010) On the MIB database accessibility based on SNMP
5. IETF RFC2578 (1999) Structure of management information Version 2 (SMIV2)[S]

Chapter 76
Network Security Monitoring-Oriented Application-Level Protocol Identification Technology

Dong-xia Wang and Yan-ru Feng

Abstract Along with the development of the Internet and strengthening of network protocol, it becomes more and more difficult to monitor network. And some sensitive information have been leaked outside the application environment. Application protocol identification plays an important role in network security monitoring. With accurate application protocol identification, it can improve the accuracy and robustness of the Network Security Monitoring (NSM). A new method of application protocol identification based on classification and characteristic matching is proposed in this paper. According to the experiment, the method can enhance the accuracy of identification. In this paper, we outline the traditional methods of protocol identification, and then propose a layer protocol identification method based on classification and characteristics analysis, which is applied to NSM. The method classifies the identified data stream by coarseness to match the different characteristics, achieving the purpose of identifying more intrusion. Experimental results show that the method greatly improves the efficiency of protocol recognition.

Keywords Security monitoring · Protocol identification · Characteristic match

With the rapid development of network, the exchange of the information becomes more and more convenient, but it brings more serious network securities at the same time. More and more users put forward higher requirements about the

D. Wang (✉) · Y. Feng
Information Engineering, Jiyuan Vocational and Technical College,
Jiyuan 459000, Henan, China
e-mail: wdxbest@163.com

Y. Feng
e-mail: fyr404@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_76,

information security. People pay more attention to the information security. The traditional firewalls, intrusion detection and many other technologies, which just can protect the network, do not receive the accusations from outside. The internal network security from inside becomes a serious problem that needs to be solved. In this situation, then comes the Network Security Monitoring (NSM). NSM through the application layer of network traffic monitoring and analysis of data, identifies a variety of application layer data, and takes a different security policy by the variety of different data. It has no effect on openness and flexibility of the network. At the same time, it ensures the safety of the sensitive data that the network owns inside.

During the implement of the NSM, Only the first goal of the network traffic is used to confirm the application layer protocol, then we can make an analysis of the data flow. We can get the loophole of the agreement and the contents in the agreement. From all above, the application layer protocol identification is the foundation of the NSM. At the same time, it is the Prerequisite to make Security Monitoring come true. For the important position of protocol identifier, In this paper, the traditional protocol identification method based on the analysis, combination of safety monitoring for identifying the characteristics of high efficiency requirements is presented. We then propose a classification based on feature recognition and agreement and recognition method.

76.1 The Traditional Application-Layer Protocol Identification

76.1.1 Port-Based Identification

Port-based identification is the use of various application-layer protocols IANA [1] in the registered port numbers to identify the protocol. For example, if a TCP stream uses port 25 or 110, it will be marked as the e-mail traffic and if the method is simple, it consumes less system resources. Some using a fixed port for the protocol recognition rate is relatively high. However, using a random port for the application layer communication protocol is powerless. Overall, the algorithm error rate is higher than the correct rate, due to its simple calculation, less the required information, the time complexity of all algorithms in the application layer protocol identification lowest.

76.1.2 Depend on it Load Identification

Identification is mainly based on the load [2] in the data stream by examining the TCP header of each packet of data after the load. Pre-load data and analysis should be matched layer protocol characteristics. If the match is successful, the agreement

will mark the corresponding. So by the theory, as long as there is enough workload analysis protocol and the actual message to be interactive features of the agreement, such algorithm can identify all of the agreements. Such recognition accuracy rate is relatively high, but its identification is based on the entire data stream to match the application layer based on the data load, and its space complexity is the entire application layer protocol identification algorithm in the highest.

76.2 Classification and Feature Matching Based on the Application Layer Protocol Identification

Because a wide range of application layer protocols are proposed, and as the Internet continues to evolve, new applications continue to join layer protocol. The traditional recognition technology has been difficult to meet the needs of practical application. In order to more quickly identify the existing application layer protocols, this paper presents a feature-based classification and analysis of application layer protocol identification program.

76.2.1 Identification Scheme

Protocol identification system includes classification and identification of two major components of device, as shown in Fig. 76.1. Data flow through the classifier are classified with coarseness into recognizer, and the recognizer for different flow characteristics matches the different protocols, enabling application-layer protocol identification. Specifically, the classification is based on some basic features of the data flow and port mapping table; identifier in the data stream classification and the combination of protocol signatures match the query. The different types of data streams to different protocols by feature recognition. Feature selection protocol is identified by the treatment protocol. Detailed analysis to identify the interaction is different from any other agreement in the field, as a feature of the agreement.

76.2.2 Data Flow Classification

Classifier function is aim to achieve network data stream diversion, because the focus is on safety monitoring network to monitor contents of the file, so we just need to identify the protocol for file transfer, through dozens of application layer protocol analysis. This paper can be used for files to transfer application layer protocol, which is divided into three categories, as shown in Table 76.1.

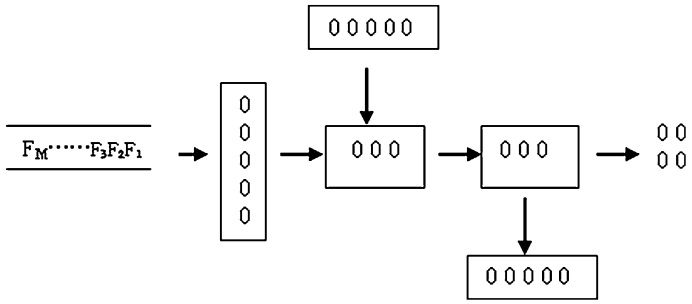


Fig. 76.1 Protocol identification flow chart

To consider the classification problem as a data stream, from the received data to the stream in which the first paragraph of the agreement by carrying the process all over this time, the classifiers which extract the relevant information to be processed can abstract the data stream. The abstract data flow will determine the ownership of a category.

Basic elements

M = {m0, m1, m2 : Collection of data stream type,m0 for the classic type, m1, said P2P class, m2 for other classes.

SrcIP = {srcip1, srcip2,… srcipn} indicates that the source IP set
DstIP = {dstip1, dstip2,… dsipn} for the destination IP set
SrcPort = {srcipn | $1 \leq$ srcportn ≤ 65535} indicates that the source port set
DstPort = {dstipn | $1 \leq$ dstportn ≤ 65535} for the destination port set.
P = {SrcIP × DstIP × λ SrcPort × DstPort} indicates a collection of packets to
Pi = {srcipi, dstipi, srcporti, dstporti} the IP address of the host to use its srcporti srcipi port the packet is passed to the IP address of Pi dspipi of dstporti port.
PT = {25, 110, 21, 80…} is a collection of classic protocol port
F = {fsrc (srcipn, srcportn), fdst (dstipn, dstportn)} is a collection of functions, fsrc (srcipn, srcportn) that the source IP address and source port are srcipn and srcportn the number of packets, fdst (dstipn, dstportn) indicates that the destination IP address and destination port, respectively, dstipn and dstportn the number of packets.

Data stream S is a finite length of the whole number of bytes of binary data, which consists of a finite length of a certain form of data packets P.

These are some basic variables and functions definitions about the classification processes discussed below.

For most of the P2P hosts, when use the P2P to download, it is host to many resources required to download the data, that the use of a port with multiple host ports for communication and data transfer. Therefore, there will be a single data stream {dsipn, dstportn} corresponding to multiple {srcipn, srcportn} phenomena [3], and for most of the protocols that are concerned, there is no such feature.

Table 76.1 Protocol and application layer protocol mapping table

Protocol type	Application layer protocol
Traditional classic	http, ftp, smtp, pop3, etc.
P2P class	Gnutella, eDonkey, DirectConnect, BT, etc.
Other	MSN, ICQ, webhard etc.

So for the data stream and a feature, the basic flow can be determined that the P2P data stream, so the following decision rule.

Rule 1. When the data stream S in the packet Pi, fdst (dstipi, dstporti) and fsrc (srcipi, srcporti) the difference between two values is greater than a specified threshold, the data stream is a P2P data stream.

Verified by experiments, when the threshold is greater than 10, using Rule 1 for classification can reach 90% accuracy.

For the classification of classical class, although there is a lot of application layer is protocols that have not been registered with the IANA port, many protocols use a fixed port. For this part of the agreement, this can still use certain techniques to identify the communications port by the following decision rule:

Rule 2. Data stream S is a classic data flow which refers to the data stream S is not a P2P data stream, but the packet Pi, there srcporti $\in$ PT or dstporti $\in$ PT.

Two by the rule that, through the port corresponding to the matching table can be achieved on a fixed port protocol data stream diversion, and this agreement will also be attributed to this type of agreement for the traditional classical category, and HTTP, FTP and other protocols in the subsequent feature matching in a more fine-grained identification.

Rule 3. Data stream S is the other type of data flow, data flow when the flow is not a P2P data stream are not classic, then S to other data streams.

We can see from the above three decision rules, only under the premise of the first transport layer protocol analysis, the data stream for a coarse-grained classification, so as to reduce the workload of the subsequent feature recognition.

76.2.3 Feature-Based Recognition

Feature recognition implementation, mainly in the extraction of features to be identified after the agreement, these characteristics attributed to the corresponding feature class. Therefore, when the data flow is separated, the feature-based recognition can be based on different data streams to different feature classes match, the traditional feature matching and classification through the feature matching, the process depicted in Fig. 76.2. The figure shows that, after a category and then match the characteristics of the protocol, due to narrowing the range of feature matching, so can increase the speed of feature recognition. String in the feature

selection, this paper in the literature [4] based on improved, making identification more accurate.

This paper adds to the Gnutella protocol identification, although the Gnutella protocol specification fully use HTTP for file transfer, but the flow through the Gnutella protocol analysis and monitoring, it was found to be testing for the Gnutella data stream flow, then:

(1) In the TCP header immediately after the load data in the first string is 'GNUTELLA', 'GET' or 'HTTP';
(2) If (1) as described in the string is not 'GNUTELLA', then load the data in the back will certainly be the following string appears:

User-Agent: <Name>
UserAgent: <Name>
Server: <Name>
<Name> For the character set in which:

{LimeWire, BearShare, Gnucleus, MorpheusOS, XoloX, MorpheusPE, gtkgnutella, Acquistion, Mutella-0.4.1, Qtella, AquaLime, NapShare, Comeback, Go, PHEX, SwapNut, Mutella-0.4.0, Shareaza, Mutella-0.3.9b, Morpheus, FreeWire, Openext, Mutella-0.3.3, Phex} a member. So long as the character set to join the library as a P2P protocol features characteristics of a class that can identify the Gnutella protocol.

76.3 Experimental Design and Analysis

76.3.1 Experimental Design

This section of the program through the experiment to verify the performance. Experimental data set on the campus network for the data collector for network traffic data to monitor the total number of packets to 11.0G. Experiments to identify the eight kinds of application layer protocol commonly used in China for example, from the accuracy and efficiency in the implementation of the program on the assessment, in order to compare the performance of the program, this port-based identification method and the literature [4] methods were compared.

76.3.2 Experimental Analysis

Figure 76.3 shows the three methods commonly used protocol for eight correct rate comparison chart, from the figure we can clearly see, the port identification for BT, eDonkey and Gnutella P2P protocol identify three kinds of almost zero, that

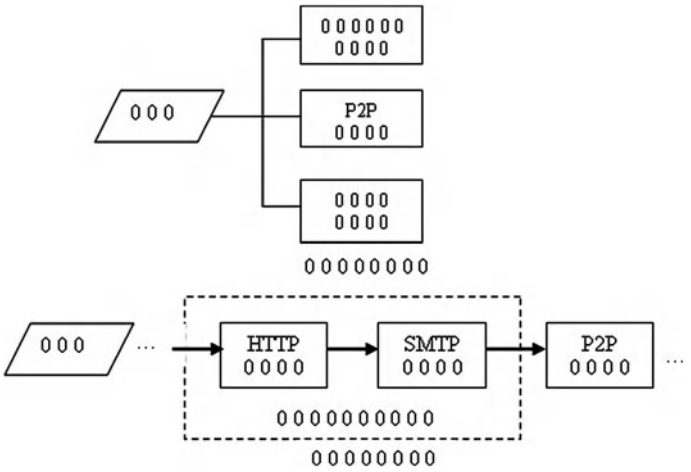


Fig. 76.2 Classification feature matching with the traditional feature matching process

the port identification method P2P protocol for the identification has been impossible, features string matching methods in the identification of the HTTP traffic is higher than that proposed in this paper on the program.

According to analysis and comparison, string matching characteristics of this program with the identification number of the HTTP traffic difference, and this program on the Gnutella protocol identification were similar, which can be inferred [4] in the string matching the characteristics of flow due to the Gnutella protocol for HTTP traffic, but this paper will be by classification, classified as P2P like Gnutella, and then combine the signatures on the Gnutella protocol to identify, so the identification of the protocol is superior to the literature on the characteristics of four string matching.

Three kinds of identification method of comparing the efficiency shown in Table 76.2. Characteristics of string matching and classification first proposed in this paper further feature matching method in space and time consumption were higher than port-based identification method, and the proposed method but also on the time consuming method than feature string. The experimental environment are carried out in the local environment, and the number of agreements and protocols in the load flow characteristics of the position of the string are fixed, so regardless of classification is to further direct feature matching or feature matching, they consume time will not be too large. However, with the agreement features a string of ever-increasing, this identification method proposed in this paper's advantages in terms of time consumption will be increasing.

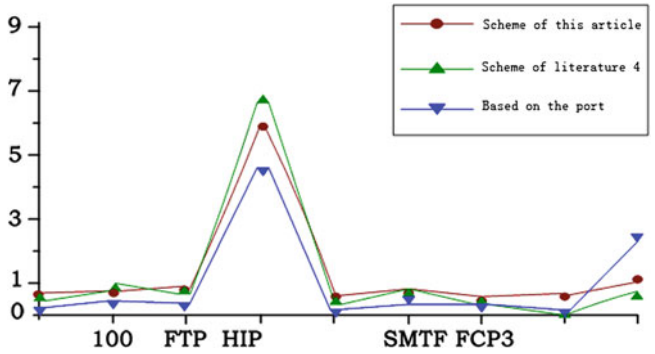


Fig. 76.3 Comparison of three methods of accuracy

Table 76.2 The efficiency comparison of three kinds of identification methods

Time Method	Time consuming
Port identification method	350 min
String matching characteristics	397 min
This program	387 min

76.4 Conclusion

Based on the advantages and disadvantages of the traditional application layer protocol analysis, combined with the data stream classification thinking of a new type of protocol identification method. In this method, a data stream for identifying coarse-grained classification, and then take a different for different identification categories of feature matching. After experimental tests proved that the proposed method on the recognition accuracy is higher than the port-based identification method, which will have higher efficiency in the implementation of string matching based on feature recognition method. However, the identification method is how the inadequacies of the existing cover more application layer protocol, accurate extraction of the characteristics of the application layer protocol, application layer protocol to further improve recognition accuracy, which is the focus of this next step to solve the problem.

References

1. IANA http://www.iana.org/assignments/port-numbers
2. Thomas K, Andre B, Michalis F, Claffy Kc (2004) Transport layer identification of P2P traffic. In: IMC'04. ACM Press, New York, pp 121–134
3. Liang C, Jian G, Xuan X (2009) Feature-based application layer protocol identification string. Comput Eng Appl 42(24):16–19
4. Kang HJ, Kim MS, Hong JW-K (2005) A method on multimedia service traffic monitoring and analysis. In: DSOM 2003. Heidelgerg, Germany, pp 93–105

Chapter 77
The Application of Chinese Word Segmentation in Malicious Code Detection

Xiaoguang Yue, Guangzhi Di, Tonglin Zhao, Yuyan Yang, Ziqiang Zhao and Yuanyuan Dong

Abstract At present, malicious code becomes more and more harmful in common people's daily life. In order to solve this problem, the relative methods and theories of Chinese word segmentation and malicious code detection are illustrated in this paper. A brief discussion on the new method of malicious code detection based on Chinese word segmentation has been made basically on a summary description.

Keywords Chinese word segmentation · Maximum matching · Malicious code detection

X. Yue (✉) · G. Di · T. Zhao · Y. Yang
School of Computer and Information, Southwest Forestry University, Kunming 650224, China
e-mail: xgyue@foxmail.com

G. Di
e-mail: swfcdgz@gmail.com

T. Zhao
e-mail: zhtl0563@sina.com

Y. Yang
e-mail: kathy-kathy@km169.net

Z. Zhao
Network Center, Shanxi University of Traditional Chinese Medicine, Taiyuan 030024, China
e-mail: ziqiang42@163.com

Z. Zhao
Complex System and Computational Intelligence Laboratory, Taiyuan University of Science and Technology, Taiyuan 030024, China

Y. Dong
School of Information Science and Engineering, Guilin University of Technology, Guilin 541004, China
e-mail: 303656966@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_77,

77.1 Introduction

Malicious code is a program that violates users' wishes and breaks into computer systems, and takes malicious action against infected machines. Malicious Code involves five main aspects: virus, Trojan, worm, ActiveX control and attack script. And they are not only a dangerous risk to computer systems, but also cause economic loss [1]. And Trojan horses, worms or any other similar programs that may damage computers and operation systems. For example, thousands of computers were infected by CIH in 1998; red code infects computers running Microsoft windows serial systems and their internet information server (IIS) software and then makes infected machines scan the internet for more victims; xiongmaoshaoxiang, called nimaya, whboy or jinzhubaoxi, is a worm circulated that tricked computer users by changing files' images into pandas in 2007. Malicious code exists mainly in website, USB disk, CD and software, and is spread via the virus-infected computers and systems.

77.2 The Algorithm of Chinese Word Segmentation

The purpose of Chinese word segmentation (CWS) is that it gives the segmentation of words results of Chinese characters [2]. There are many methods of Chinese word segmentation: rule-based CWS, statistics-based CWS, understanding-based CWS, multi-features-based CWS and so on. The most common methods are forward maximum matching (FMM) algorithm and reverse maximum matching (RMM) method. FMM and RMM belong to rule-based CWS. FMM is an efficient algorithm that can be used in many research areas. The principle of forward maximum matching is that compares with the word dictionary and segments Chinese word, then finds words and unknown words. There are string ABC and dictionary Dic, $A \in Dic$, $AB \in Dic$, $C \in Dic$, $ABC \notin Dic$, then ABC will be segment into AB/C (/ is a separator). There is string DEF, $D \notin Dic$, $DE \notin Dic$, $F \notin Dic$, $DEF \notin Dic$, so DEF is an unknown word [3]. For example, maximum matching method is shown in Fig. 77.1.

Word segmentation specification, ambiguous solution and unknown words recognition are the three common problems of Chinese information processing. Although there are many Chinese word segmentation specifications for information processing, Chinese word segmentation material still can not unite in information processing for variety causes. Ambiguous solution is the second question for Chinese information processing. For example, 结合成 can segment into 结合/成 or 结/合成. Unknown words are frequently found on person name, location name, organization name, league name, specialized vocabulary and abbreviation, like "路易斯.费利佩.马德拉.卡埃罗.菲戈 (Madeira Caeiro Figo Luis Filipe)", 碘-, 日本宫城县, 中科院计算机所 and so on.

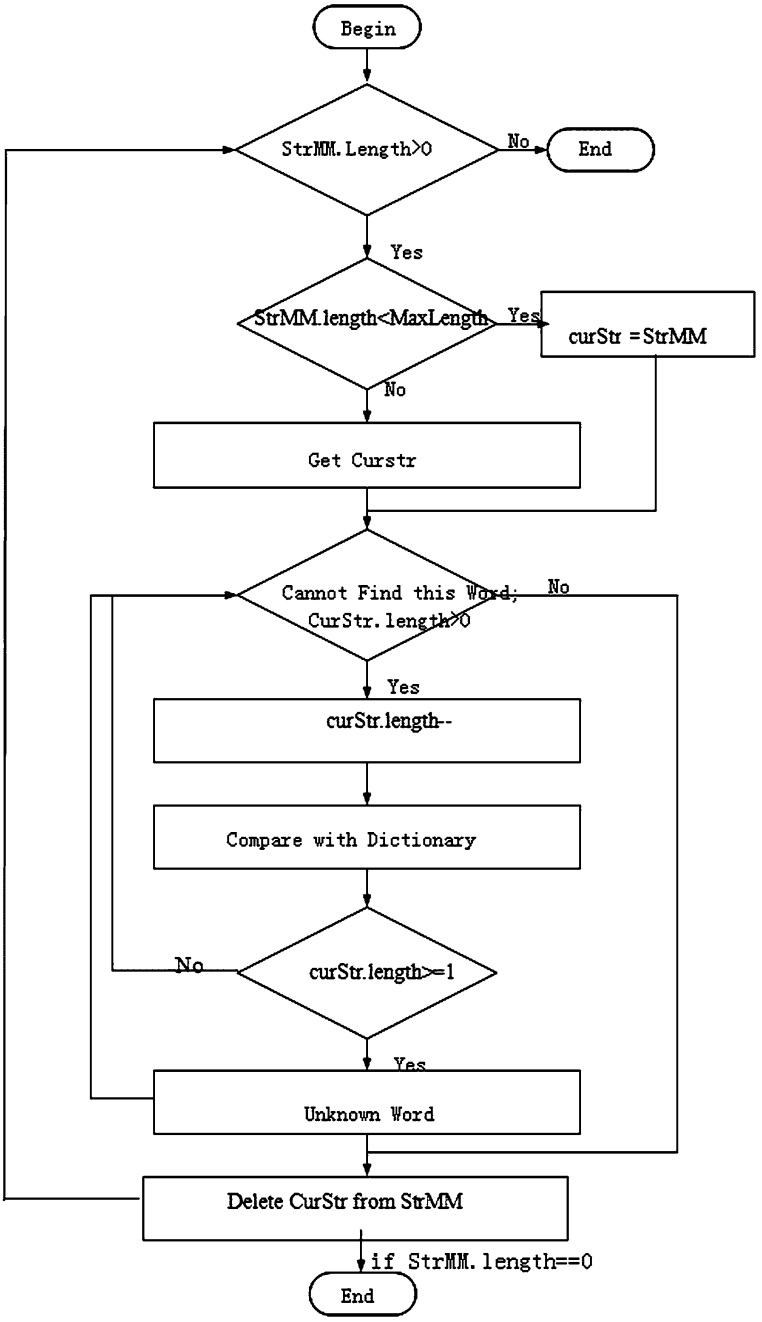


Fig. 77.1 Maximum matching method

77.3 Malicious Code Detection

The mineral is mainly used the electromagnetism interference of the monitor of monitor LED to result in as follows from a few factors:

(1) The electromagnetism interference of switch the power creation. The mineral is used in LED monitors, what great majority uses is the power switch , however lead a moment in the switch tube of the power switch, the transformer entry-level coil will produce great flow, and a higher wave appears to flow out peak electric voltage in the entry-level coil at both ends; Break to open a moment in the switch, because of the leaking of entry-level coil Ci, causes a part of energies did not deliver to two coils from a coil, hoard at leaking part of energies in the feeling will and the pole electric capacity, electric resistance formation of switch take the Shuai of having the peak to reduce to flap to concuss and fold to add the pass power failure that takes care of in the switch to press up, the formation closes peak electric voltage. This interference signal will conduct importation exportation to carry, the formation conducts interference. Second, Gao repeatedly transformer entry-level coil, switch tube and filter Gao that the electric capacity constitutes repeatedly the switch electric current wreath road may also produce bigger space radiation, the formation radiates interference. As for this kind of interference have already described in detail in this chapter Sect. 77.1, no longer launch a concrete treatise here.
(2) Monitor of crystal flap to result in of interference. Is crystal to flap mainly is in order to show that the system provides a basis frequency, it is a kind of interference signal Gao Ci Xie's wave of period pulse signal producing. This kind of interference signal will match to hold through various way Ou inside of in the cable, then with bad mold or total mold formation radiation or conduct interference. Along with the development of science and technology, monitor's breaking velocity will be more and more high, crystal the basis frequency flapping also immediately become Gao, from crystal flap causable Gao Ci Xie wave the interference is stronger.
(3) Because signal line inside the monitor and power cable are more airtight, the interference signal will be also hold mutual Ou of inside track cable to match, finally radiate space or go out through a power cable conduction and form interference.

77.4 The Application of Chinese Word Segmentation in Malicious Code Detection

Boyer–Moore algorithm is used in many detection systems. Qing Ye, Xiaoping Wu and Jin Cheng have pointed out some definitions of Boyer–Moore algorithm [4].

Definition 1 T is a major string for matching and its length is n.

Definition 2 P is a pattern string for matching and its length is m.

Definition 3 Pattern matching is the process of finding P in T

Based on these three basic definitions and hypothesis, features dictionary *FDic* have been built based on malicious code features. *FDic* keeps records of malicious codes. There is a pattern string $P(1)$ for test. Define reverse matching substring of $P(1)$ as $\mathrm{Pr}(1,i)$ $(i = 1, 2, 3, \ldots)$. Define forward matching substring of $P(1)$ as $Pf(1,j)$ $(j = 1, 2, 3, \ldots)$. Define the length function as $\mathrm{Len}(X)$ $(X = P(1), P(2), P(3), \ldots)$. Define *fr* as a reverse matching cumulative variable of malicious features $(\mathrm{int}\, fr = 0)$. Define *ff* as a forward matching cumulative variable of malicious features $(\mathrm{int}\, ff = 0)$. Define the feature threshold of codes as *Th*.

First, reverse maximum matching is for $P(1)$.

Step 1: $P(1)$ $(\mathrm{Len}(P(1)) = m)$ is segment from the last few characters (this substring is $\mathrm{Pr}(1,1)$, and $Len(\mathrm{Pr}(1,1)) = l$);
Step 2: Compare $\mathrm{Pr}(1,1)$ with *FDic*;
Step 3: If $\mathrm{Pr}(1,1)$ can be found in *FDic*, $\mathrm{Pr}(1,1)$is a full malicious feature; $fr = fr + 1$; then $\mathrm{Pr}(1,1)$ is deleted from $P(1)$ $(\mathrm{Len}(P(1)) = m - l)$; return new $P(1)$ $(\mathrm{Pr}(1,2))$ to step 1.
Step 4: If $\mathrm{Pr}(1,1)$ cannot be found in *FDic* $(l > 1)$, $\mathrm{Len}(\mathrm{Pr}(1,1)) - -$; return new $\mathrm{Pr}(1,1)$ to step 2; if $l = 1$, $\mathrm{Pr}(1,1)$ is an unknown feature that should be recorded in database;
Step 5: If all substring of $P(1)$ $(\mathrm{Pr}(1,1), \mathrm{Pr}(1,2), \mathrm{Pr}(1,3)\ldots)$, have been fully detected, count the result of *fr*.

Second, forward maximum matching is for $P(1)$.

Step 1: $P(1)$ $(\mathrm{Len}(P(1)) = m)$ is segment from the first few characters (this substring is $Pf(1,1)$, and $(\mathrm{Len}(Pf(1,1)) = l)$;
Step 2: Compare $Pf(1,1)$ with *FDic*;
Step 3: If $Pf(1,1)$ can be found in *FDic*, $Pf(1,1)$ is a full malicious feature; $ff = ff + 1$; then $Pf(1,1)$ is deleted from $P(1)$ $(Len(P(1)) = m - l)$; return new $P(1)$ $(Pf(1,2))$ to Step 1.
Step 4: If $Pf(1,1)$ cannot be found in *FDic* $(l > 1)$, $Len(Pf(1,1)) - -$; return new $Pf(1,1)$ to Step 2; if $l = 1$, $Pf(1,1)$ is an unknown feature that should be recorded in database;
Step 5: If all substring of $P(1)$ $(Pf(1,1), Pf(1,2), Pf(1,3))\ldots)$ have been fully detected, count the number of *ff*.

Finally, Compare the before two matching results. If $ff \geq Th$ or $fr \geq Th$, $P(1)$ is a malicious code. If there are ambiguous substrings in two ways, system administrators and experts should be notified. And ambiguous substrings will be analyze and solve by system administrators and experts.

77.5 Conclusion

There are many articles about malicious code detection [5, 6]. Many ways are very useful. Malicious code detection is good for society, especially for computer and system. And malicious code detection methods become more advanced and complicated, the purpose of all of them is to detect malicious code in computer programs and systems in a better way. This article briefly describes the basic theories of Chinese word segmentation and malicious code detection. Maximum matching method is shown in figure one. Chinese word segmentation is a part of artificial intelligence. So, the application of Chinese word segmentation in malicious code detection can be regarded them as a combination of Intelligence malicious code detection. At last, with analyzing and studying the method of Boyer–Moore algorithm, the article gives a new method of malicious code detection based on Chinese word segmentation.

Acknowledgments This work is supported by National Natural Science Foundation of China for Youths (61003053), Science and Technology Development Project of Shanxi Province Colleges and Universities (20101105) and Natural Science Foundation of Guangxi Autonomous Region (0991254). Special thanks go to Prof. Fanjin Mai and Dr. Zhigang Liu and Dr. Yeqing Zhao and Mr. Fei Xiong for their invaluable comments and suggestions.

References

1. Wang C, Wang F, Zhang Y et al (2007) Malicious code modeling and analysis in weighted scale-free networks. Wuhan Univ J Natur Sci 12:51–54
2. Mai F, Yue X, Zhao Z (2010) Intelligence search engine model based on similarity computation. In: Proceedings of the 2010 international conference on future computer and communication, Wuhan, pp 106–108
3. Zhai F, He F, Zuo F (2006) Chinese word segmentation based on dictionary and statistics. Mini-Micro Syst 27:1766–1771 (In Chinese)
4. Ye Q, Wu X, Cheng J (2010) Detection of malicious codes by using feature matching of rules optimization and sort. J Naval Univ Eng 22:103–106 (In Chinese)
5. Assaleh TA, Cercone N, Keselj V et al (2003) Detection of new malicious code using N-grams signatures. International conference on intelligent information systems (IIS). Springer, Poland, pp 193–196
6. Zhang Z, Guo F (2010) Malicious code detection by taint analysis. 2010 international conference on multimedia communications, pp 206–209

Chapter 78
Application of Windows Tilt Correction Technology in Literature Digitization Conversion

Wenliang Zheng

Abstract Due to the document scan tilt easy to have noise often, the stand of the tilt correction will directly affect the quality of the digital document. This paper puts forward the tilt correction method based on Windows, and gives the results. This algorithm is used to test the layout of this chapter the size is A4, with a resolution of 300 dpi scan input, layout images include Chinese text, English text, images and tables, calculate on a Pentium 4 clocked at 3.0G implemented on a microcomputer. The results show that the results of this basic algorithm close to the actual layout of the tilt angle.

Keywords Document · Tilt correction · Image recognition

78.1 Introduction

Along with the information and the development of computer technology and popularization, human society has entered into the information age. The information has two categories, that is, paper publications and electronic publications. When transfer the paper document into electronic documentation process, because of the noise produced by the document scanning tilt, after scanning the quality of the image will directly affect the quality of the digital document. The existing tilt correction methods are projection method, the nearest neighbor clustering method, Fourier transform method and so on.

W. Zheng (✉)
Library of ShenYang Ligong University, Shenyang 110000,
People's Republic of China
e-mail: zhengwe3049_cn@sina.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_78,

78.1.1 The Method of Projection

Kindly assure that the Contact Volume Editor is given the name and email address of the contact author for your paper. The Contact Volume Editor uses these details to compile a list for our production department at SPS in India. Once the files have been worked upon, SPS sends a copy of the final pdf of each paper to its contact author. The contact author is asked to check through the final pdf to make sure that no errors have crept in during the transfer or preparation of the files. This should not be seen as an opportunity to update or copyedit the papers, which is not possible due to time constraints. Only errors introduced during the preparation of the files will be corrected.

78.1.2 The Nearest Neighbor Clustering

The Nearest neighbor clustering method is based on the following facts: the general direction of the document layout is the direction of text lines, text, arranged in order. Characters on the same line with each other by the close, can use their distance and spatial relationship between the angle of position detection. From this we can first find the connected domain document image analysis which connected domain on one line, each connected domain with its center point or the midpoint of the bottom of said center line of these points using the least square method or the straight line Hough transform to extract, the inclination of the straight line is the request [1, 2].

78.1.3 Method Based on Fourier Transform

The method based on Fourier transform detection dip Angle, will all pixel layout for Fourier transform layout, corresponding to make Fourier space fixing the most dense direction. Because of a large amount of calculation, this method has little practical use.

78.2 Windows-Based Layout Tilt Detection Algorithm

In order to detect tilt Angle of the layout quickly and accurately and reduces computation. This paper uses a method which is based on the layout tilt detection of the windows transformation. First, this algorithm will deal the details of the text and pictures in the windows with changing resolution processing, and then fit a straight line for its edge, so the tilt angle of layout will be detected.

78.2.1 Algorithm Definition

Use 10-point type for the name(s) of the author(s) and 9-point type for the address(es) and the abstract. For the main text, please use 10-point type and single-line spacing. We recommend the use of Computer Modern Roman or Times. Italic type may be used to emphasize words in running text. Bold type and underlining should be avoided.

Definition 1 I is the layout image, the size of the image is H × W.

Definition 2 We suppose there is image J in the layout, and $J \subseteq I$, then J is defined as a sub graph of the layout.

By the Definition 1, 2 shows that if $\varphi(I)$ like tilt angle of I, then $\varphi(J) = \varphi(I)$;

Definition 3 If the sub graph of the layout J correspond

$$J(x, y) = \{(x, y)|x \in [w_1, w_2], y \in [h_1, h_2], (h_2 - h_1) \geq nh\} \tag{78.1}$$

punctuate a displayed equation in the same way as ordinary text but with a then the image J is the selected window.

In the equation: $w1 = W/8$, is the left edge of the window image;

$w2 = W/4$, is the right edge of the window image;
$h1 = H/6$, is the upper border of the window image;
$h2 = H/3$, is the nether border of the window image;
h is the line height;
n is the number of lines which is contained in the windows.

Definition 4 We suppose the image S(x,y) rotate θ based on the lower left corner, and get image S(u,v), then

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{78.2}$$

According to above definition, steps of the layout window the based on tilt detection algorithm can be summarized as follows:

(1) Choose the appropriate window J in the layout image I;
(2) Use the rotation transformation, while project in the window J and judge the projection results;
(3) Use variable-precision processing for the window;
(4) Extract the corresponding feature points, and fit a straight line for the points, determine the tilt angle of the layout.

78.2.2 Implementation of the Algorithm

Selection of Window. Paragraph text in space of a whole page is usually typeset by the same direction, and the leading is fixed. So the projection of the paragraph text in the layout image takes on periodic. There will be a lot of noise in the layout edge when they are scanned. Because of that, we should consider to remove all the layout edges and ensure the windows are big enough when choosing windows, that is, meet Definition 3.

If the line number of the text is less than appointed n except images and tables in the window, the window is reselected. The layout usually contain text, tables, graphs or images, so the window content can be detected by the projection of horizontal and vertical when the window is selected.

Rotation of window. The numbers If the content of the selected window is different, the projection is different too. If text line is contained in the window, and the angle of inclination of the window is less than θ_0, we can see the recurrent pulses from horizontal projection directly which are provided with certain width and peak value. If straight line is contained, the peak value is more than in text line and the breadth is more narrow (Fig. 78.1). In the figure, the origin of coordinates is at left bottom of the "window picture", the ordinate of horizontal projection and abscissa of vertical projection is, respectively, the height and width of the selected window. The unit is pixel points and the unit in the ordinate of horizontal projection and abscissa of vertical projection is all Accumulative total black pixel points.

The selected window is rotated according the Definition 4 in this algorithm to obtain the best recurrent pulse in contrast. The rotation is implemented based on the far left of the window as baseline and the as the increment. When it is rotated once, we can ensure the existence of text line according the estimate of projection value in recurrent pulses. When the pulsing peak value of is steady, we can have the approximate degree of tilt angle. If a narrow high-peak-pulse exists, we can confirm that straight lines exist [3–5].

Then turn to the step "estimate the tilt angle" with the point of the straight line as characteristic point and fit the straight lines to get the tilt angle.

The process of variable-precision. Since the tile angle is irrelevant with the layout content, the image of the window is processed in horizontal direction based on the Eq. (78.3) to reduce the precision of texts or images in the algorithm.

$$f(x,y) = 1, \sum_{i=\triangle d}^{k-\triangle d} \sum_{j=\triangle d}^{k-\triangle d} f(i \pm \Delta d, j \pm \Delta d) \geq \frac{d^2}{2}, \Delta d = \frac{d}{2} \tag{78.3}$$

In Eq. (78.3): d represents the word space;
R represents the pixel number of the window in horizontal direction;
C represents the pixel number of the window in vertical direction.

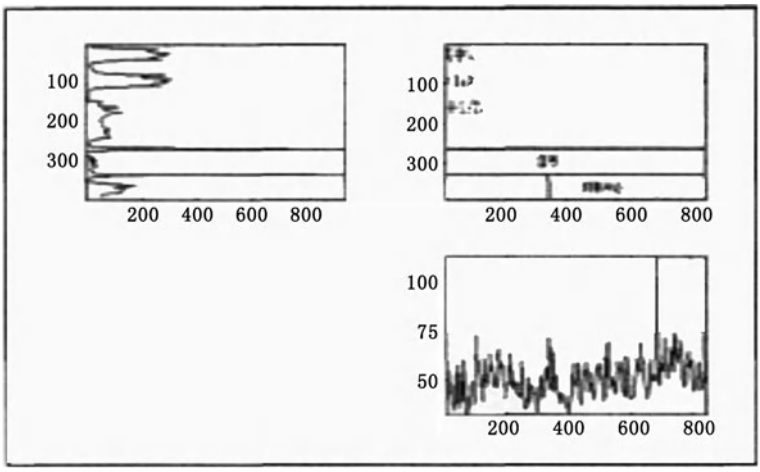


Fig. 78.1 Windows contain text and lines

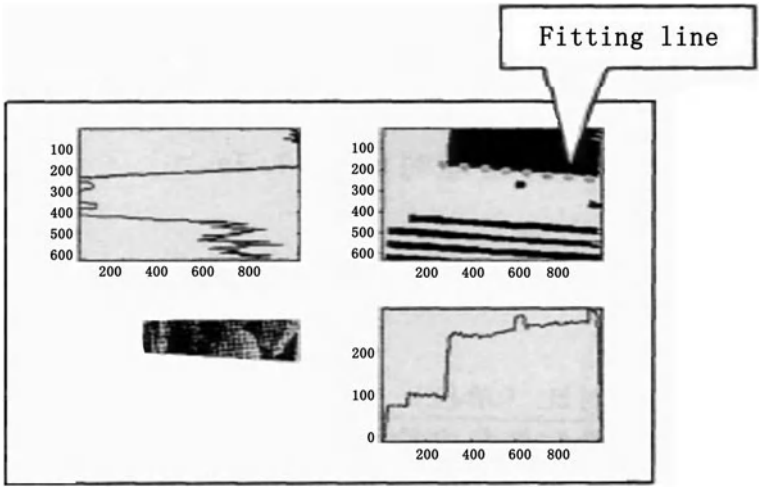


Fig. 78.2 Result of windows with images

The Eq. (78.3) indicates that if the sum of the black pixel points exceeds the half sum of any $d \times d$ area pixel points, then we replace this area pixel points with black pixel points.

Determine the angle of layout. Start from the window edge, top-down selection the feature points that the point of change the value of adjacent pixels, and composition of a collection of feature points with adjacent feature points. If feature points set points long enough, fitting a straight line with the method of the minimum squares, find the angle with the horizontal line, it is the layout angle; Otherwise, start from the windows along and from the left to right to find feature

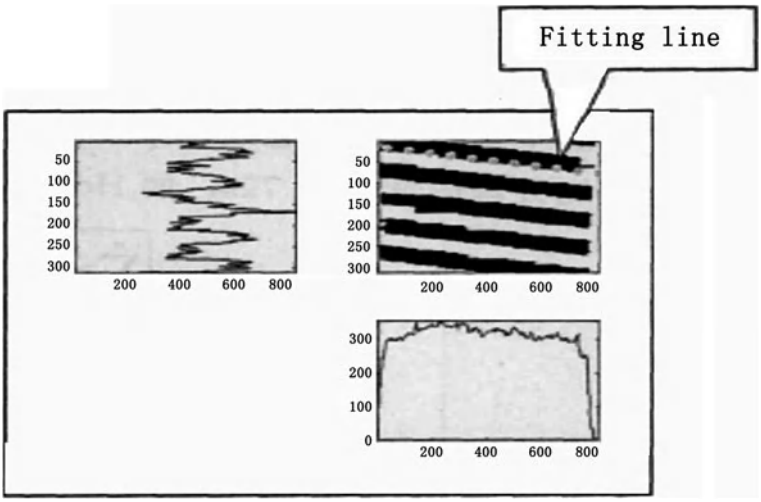


Fig. 78.3 Result of windows with texts

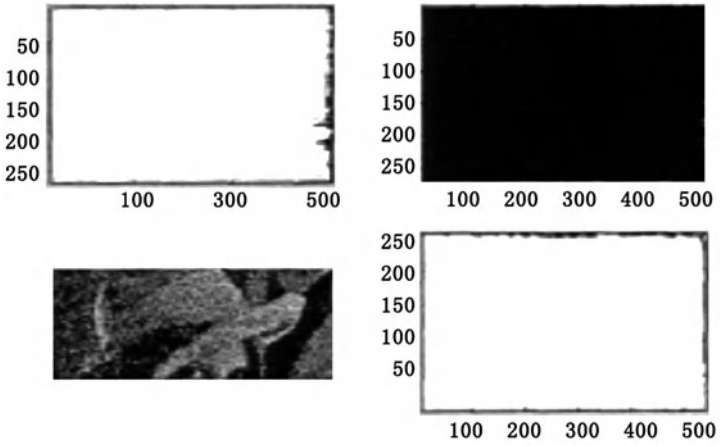


Fig. 78.4 Without the image side of the window

points form a set of feature points, then fitting a straight line with the method of the minimum squares, find the angle with the vertical line, it is the layout angle (For Figs. 78.2 and 78.3).

If all of the selected window is the image that there is no image borders (For Fig. 78.4), then select the original window left border to right border of the same size window, repeat the above variable-precision processing, find the layout tilt angle Otherwise, reselect the window.

Table 78.1 Experimental results

The actual layout image	Layout tilt angle results
−3.50	−3.53
−2.50	−2.51
−1.50	−1.49
−0.50	−0.47
0.50	0.51
1.50	1.48
2.50	2.51
3.50	3.44
Average	0.02

78.3 Results Statistics

This algorithm is used to test the layout of this chapter size is A4, with a resolution of 300 dpi scan input, layout images include Chinese text, English text, images and tables, calculate on a Pentium 4 clocked at 3.0G implemented on a micro-computer. The results shown in Table 78.1, it shows that the results of this basic algorithm close to the actual layout of the tilt angle.

References

1. Lv Y, Chen J, Lu X (2006) Content-based document image tilt correction. Comput Simul 23(12):192–196
2. Zhong H, Sun S, Liu Q (2008) Layout analysis and reconstruction of Chinese Shenyang Architectural University. Natur Sci 24(2):333–336
3. Liu J (2008) Document image understanding research sections. PhD thesis. Chongqing University, Chongqing
4. Chen S, Zhao H, Kong M et al (2007) 2D-LPP: a two-dimensional extension of locality preserving, projections. Neurocomputing 70(4–6):912–921
5. Tsui PPC, Basir OA (2006) Wavelet basis selection and feature extraction for shift invariant, ultrasound foreign body classification. Ultrasonics 45(1–4):1–14

Part VI
Image Processing

Chapter 79
Improved Method for Image Segmentation Based on Cellular Neural Network

Man Guo and Dongqing Feng

Abstract In this paper, we proposed an improved method for gray image segmentation. The proposed algorithm first searches a threshold automatically that distinguishes the object and background. Then divides the gray image into homogeneous regions by cellular neural network. After some simulations based on the proposed algorithm, we make a comparison among cellular neural network segmentation, traditional single-threshold-based segmentation and Otsu algorithm. The simulation results indicate that cellular neural network segmentation processes a better segmentation effect.

Keywords Image segmentation · Cellular neural network · Otsu algorithm

79.1 Introduction

The goal of still color segmentation is to divide the image into homogeneous regions. There are several ways to dine homogeneity of a region that are based on a particular objective in the segmentation process. For instance, it may be measured by color, gray levels, texture, motion, depth of layers, etc. Overlaps among regions are not permitted, thus, each pixel belongs only to a single region. Two neighboring regions should be merged if the new combined region is homogeneous. Consequently, each region is anticipated to be as large as possible under its certain characterization. Then, the total number of regions is reduced [1].

M. Guo (✉) · D. Feng
Department of Electrical Engineering, Zhengzhou University,
Kexue Road, Zhengzhou, Henan, People's Republic of China
e-mail: gmest@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_79,

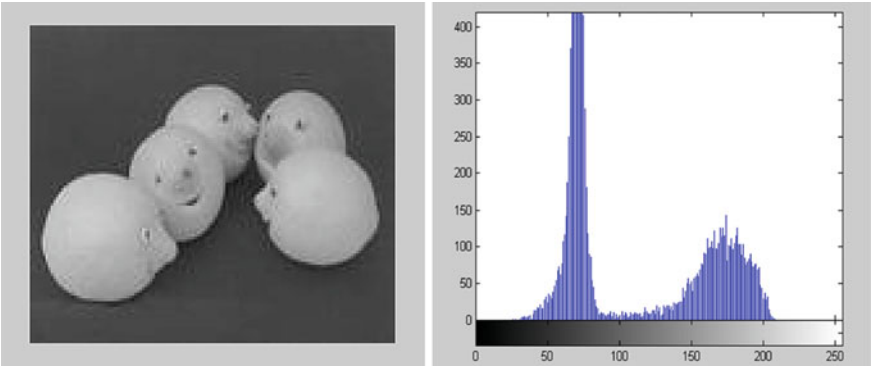


Fig. 79.1 Image and its histogram

Most of the histogram-based algorithms deal with gray level images, which are represented as one-dimensional (1D) histogram. The range of intensities is assumed to be constant. The histogram is considered as being a probability density function of a Gaussian and the segmentation problem is reformulated as parameter estimation followed by pixel classification [2].

Histogram threshold on the other hand specifies the gray distribution of the image. As shown in Fig. 79.1, 100 maybe is a good segmentation threshold. However, most of the histogram-based methods were found to fit specific images. These methods achieved reasonable performance when the input is characterized without noise and with small number of regions. Moreover, the number of potential segmentation classes in the image is usually assumed to be known beforehand.

It is well known that Hopfield neural network (HNN) requires fully connected and grows exponentially with the size of the array. Thus it is very difficult to implement, even in the modest array sizes [3]. A novel class of information processing system called cellular neural cutwork (CNN) was proposed by Chua and Yang in 1988, which came from the Hopfield neural network and cellular automata as an effective combination of both characteristics [4, 5]. Moreover, the CNN has two prominent features: real-time digital signal processing capability and local connection; On one hand, the characteristic of real-time signal processing has been extensively exploited in various applications such as parallel signal processing, image edge detection, connected component detection and various morphology operations (dilation, erosion and hole filling, etc.). On the other hand, the characteristic of local connection makes it applicable to VLSI implementation and allows operating at a very high speed in real time.

With deep submicron technology (0.25–0.33 um), an array of 100 $\times$ 100 large analog processors array can be implemented on a single chip, whose theoretical computation speed can be at least a thousand times faster than the current digital processor [6]. Some smaller operational test chips have also been designed. As a result of this rapid development, the CNNs have been widely studied for practical

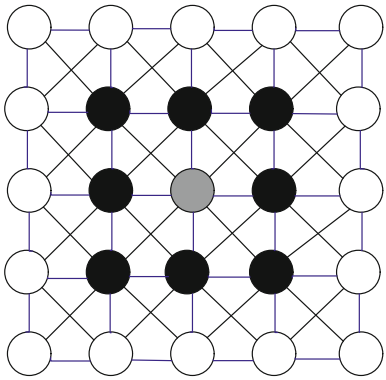


Fig. 79.2 The structure of cellular neural network

applications in image and video signal processing, robotic and biological visions and higher brain functions [7–9].

79.2 CNN Threshold

The basic unit of cellular neural network is known as cell. Each cell is only connected to its neighboring cells, according to this rule, many models can be constructed in theory. CNN proposed by Chua is called Standard CNN: it consists of m rows by n columns cell array arranged in the form of single-layer network.

As shown in Fig. 79.2, a 3*3 grid network with a gray cell in the middle and 8 black cells around is a typical structure of CNN. The state of the cell in the middle only has relationship with its rounding cells. Compared with neural network, CNN is computation saving.

The human eyes can be recognized as the best image processing 'system'. As can be seen in Fig. 79.2, the way cells distribute is similar with digital image pixels' arrangement. From this point, that cellular neural network simulates human eyes in image processing makes sense. CNN image processing is to choose network with the same scale of image to be processed, each pixel in the image corresponds to a cell in the network that is why CNN processes the capability of parallel computation.

In following, the model of a two-dimensional CNN is briefly described which is composed of basic processing units called cells. The dynamics of each cell can be described by the following state equations [1, 2]:

Every cell can be described as a dynamic system with its input detonated by u_{ij}, state by x_{ij} and output y_{ij}. In order to facilitate computation, the continuous equations are converted to discrete form. Any cell's state equation and output equations has the following form [4, 5]:

$$x_{ij}(n+1) = \sum_{C(\mathrm{k},l)\subset N(i,j)} A(i,j;k,l)y_{kl}(n) + \sum_{C(k,l)\subset N(i,j)} B(i,j;k,l)u_{kl}(n) + z^* \quad (79.1)$$

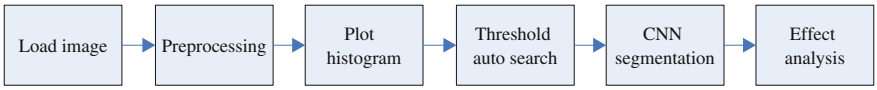


Fig. 79.3 The process of image segmentation based on CNN

$$y_{ij}(n+1) = \frac{1}{2} * \left(\left|x_{ij}(n)+1\right| - \left|x_{ij}(n)-1\right|\right) \tag{79.2}$$

$1 \le i \le M, 1 \le j \le N, x_{ij}, y_{ij}, u_{ij}, z_{ij}$ represents the state, the output, the input and threshold, respectively. The initial condition $\left|u_{ij}\right| \le 1$ and static input $\left|u_{ij}\right| \le 1$. A(i, j;k, l), B(i, j;k, l) denote the connection templates from cell C(k, l) to cell C(i, j); Z_{ij} represents the bias of (i, j)th cell in the grid. From Eq. (79.1), (79.2), it follows that the state and the output of each cell are affected by the inputs and outputs of its neighboring cells.

79.2.1 Standard CNN Threshold

To apply CNN properly to image processing, the design of a cellular neural network template is an important problem, and has received wide attention. Many algorithms have been used for templates design, including a variety of optimization methods, such as GA (genetic algorithms), particle swarm optimization and simulated annealing. The basic step is to design a sample of some of the binary image to find its corresponding desired output image and find optimal parameters to minimize the objective function in a certain search space. Then apply the trained parameters for the gray-scale images. Csaba Rekeczky and other scholars in Hungarian Academy of Science summed up a lot of templates for specific use and built up a Temlib (template library). Here we use one of templates [10] for threshold.

$$A = \begin{bmatrix} 0,0,0 \\ 0,2,0 \\ 0,0,0 \end{bmatrix}, \quad B = \begin{bmatrix} 0,0,0 \\ 0,0,0 \\ 0,0,0 \end{bmatrix}, \quad Z = -z^* \tag{79.3}$$

$$-1 < z^* < 1 \tag{79.4}$$

Considering a gray-scale image I with threshold z^*, if $U(t) = 0$, $X(0) = I$ and the cells on the boundary of image with a state of zero, then the gray value of the pixels in I is 'black' when $P_{ij} > z^*$, else the gray value is 'white'.

79.2.2 CNN Algorithm

The histogram plots the number of pixels in the image (vertical axis) with a particular brightness value (horizontal axis) as shown in Fig. 79.3. In the field of Computer

Fig. 79.4 Original image and segmentation based on Atsu of Lena

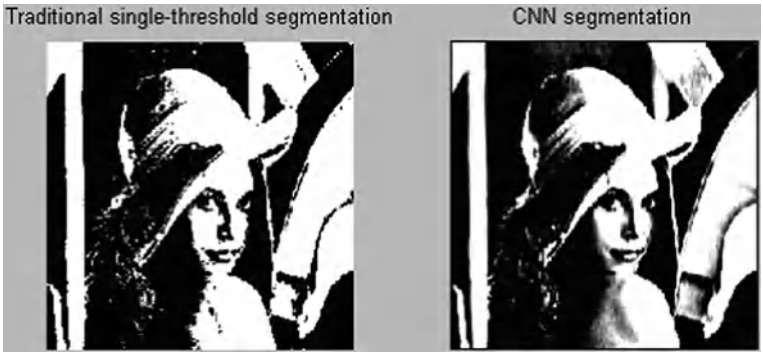


Fig. 79.5 Traditional single-threshold segmentation and CNN segmentation of Lena

Vision, image histograms can be useful tools for threshold. Because the information contained in the graph is a representation of pixel distribution as a function of tonal variation, image histograms can be analyzed for peaks and/or valleys which can then be used to determine a threshold value. This threshold value can then be used for edge detection, image segmentation and co-occurrence matrices.

In many situations, background and target is not so definitely divided. It is not easy for us to figure out the gray of object and background in image histogram, so we would better propose a searching method of the two peaks. Here we should construct a relationship between gray values $v_{\max}$ and the frequency they occur. There is no doubt that the pixel with maximum frequency corresponds to one of peaks. Considering the fact that threshold must located in between the interval $[0,v_{\max}]$, we just need to find the point which has a lower frequency of occurrence than the points following it.

Using the searching method discussed above, we can determine the threshold value explicitly. Having discussed the method threshold searching, we'll give the algorithm of segmentation based on CNN.

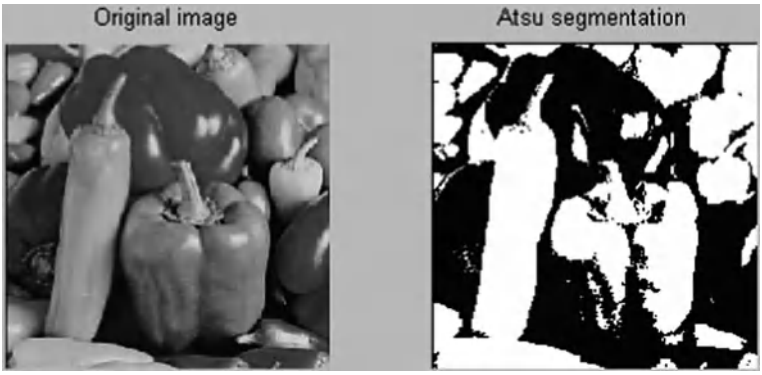


Fig. 79.6 Original image and segmentation based on Atsu of vegetables

Step 1: Image preprocessing, including denoising and image enhancement and so on;
Step 2: Use the discussed method to search for T-CNN i.e. z^*;
Step 3: Normalize the image gray matrix

$$p_{ij} = \frac{2 \times I_{ij} - \max I - \min I}{\max I - \min I} \tag{79.5}$$

The gray value of ith row, *j*th column pixel denoted by I_{ij}, while the normalized gray value denoted by p_{ij}.

Step 4: Use CNN algorithm for image segmentation. To be specific, the algorithm flows as the block diagram Fig. 79.4 shows below.

79.3 Simulations

After discussing the T-CNN algorithm, in this section, an example is presented to illustrate the effectiveness of the proposed method. Figs. 79.5, 79.6, 79.7 show the process of segmentation and make comparison of results of segmentation between Atsu algorithm and CNN algorithm, single-threshold segmentation and CNN algorithm. Here we take 'lena' and 'vegetables' for examples, use CNN algorithm for segmentation.

Image normalizing is to make every cell satisfy the premise of CNN, $|u_{ij}| \leq 1$.The image contrast does not change but the gray value of pixels distribute in a relative small range. After comparison of the three methods we can find that traditional single-threshold segmentation is simple to realize but calls for obvious difference between backgrounds and objects. For a common image, if the threshold is not carefully chosen, the segmentation results can be extremely bad. Using Atsu algorithm we can get a good segmentation result but it costs a relatively long time.

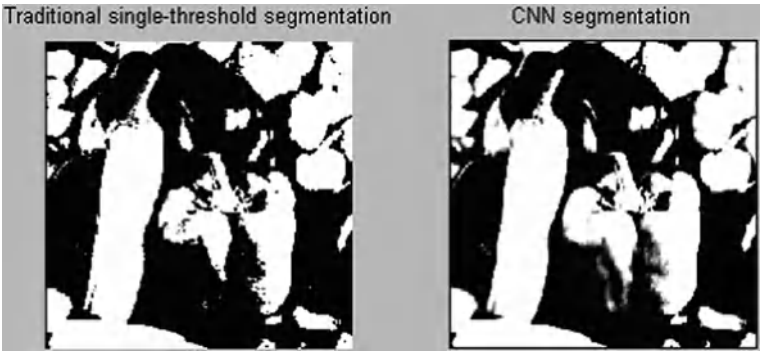


Fig. 79.7 Traditional single-threshold segmentation and CNN segmentation of vegetables

Compare the three processed images, we can find that the results of single-threshold-based segmentation and Atsu segmentation are has the problem of edge blur, while CNN segmentation edges are clear.

79.4 Conclusions

In this paper we proposed an improved method. First, do some processing including image denoising and image enhancement. Second, search the threshold automatically using the algorithm we discussed. Third, segment the image based on CNN. Large numbers of experiments show that the algorithm has high detection rate. What distinguishes the traditional method is that CNN does not need to set the threshold, but construct the relationship between image pixels and cells by setting the appropriate template. If templates are properly chosen, good segmentation results can be achieved and the improved method is computation saving. CNN are bound to be great prospect in terms of image processing and computation vision.

Acknowledgments This work is supported in part by the National Natural Science Foundation of China under grant no. 60974005.

References

1. Navon E, Miller O, Averbuch A (2004) Color image segmentation based on adaptive local thresholds. School of Computer Science, Tel-Aviv University
2. Haralick RM, Shapiro LG (1985) Survey: image segmentation techniques. Comput Vis Graph Image Process 29:100–132
3. Li H, Liao X, Li C, Huang H, Li C (2011) Edge detection of noise images based on cellular neural networks. State Key Laboratory of Power Transmission Equipment: System Security and New Technology, College of Computer Science, Chongqing University

4. Chua LO, Yang L (1988) Cellular neural networks: theory. IEEE Trans Circ Syst 35(10):1257–1272
5. Chua LO, Yang L (1988) Cellular neural networks: applications. IEEE Trans Circ Syst 35(10):1273–1290
6. Zarandy A (1999) The art of CNN template design. Int J Circ Theory Appl 27(1):5–23
7. Harrer H, Nossek JA, Roska T, Chua LO (1994) A current-mode DTCNN universal chip. In: Proceedings of IEEE international symposium on circuits and systems, pp 135–138
8. Cruz JM, Chua LO, Roska T (1994) A fast, complex and efficient test implementation of the CNN universal machine. In: Proceedings of third IEEE international workshop on cellular neural networks and their application. Rome, pp 61–66
9. Espejo S, Dominguez-Castro R, Rodriguez-Vázquez A, Carmona R (1995) CNNUC2 user's guide, Centro Nacional de Microelectrónica, Seville
10. Chua LO, Roska T (2002) Cellular neural networks and visual computing. Cambridge Press, London, pp 52–56

Chapter 80
Denoising and Recognition for Road Signs Based on Markov Random Fields

Ailan Yang, Weifeng Liu and Chenglin Wen

Abstract Since the road signs (RS) are susceptible to noise and difficult to recognize, the paper proposed a method to remove noise of RS and recognize it, where the parameters were estimated by the EM algorithm. The algorithm consists of two steps: first, the RS were modeled and restored by the Markov random field (MRF); Second, recognizing restored RS was through calculating the invariant moments. Simulation experiment compares the median filtering method and Gaussian smooth method. The results show that the proposed algorithm is better than other two methods in restoration and easily to recognize the RS.

Keywords Road signs · Markov random field · EM algorithm · Moment invariants

This work was supported in part by the NSFC (91016020, 60934009, 61175030), the ZJNSF (Y1101218, Y1080422), and the Hangzhou Dianzi University (KYS065609051).

A. Yang · W. Liu (✉) · C. Wen
School of Automation, Hangzhou Dianzi University,
Hangzhou 310018, Zhejiang, China
e-mail: Liuwf@hdu.edu.cn

A. Yang
e-mail: yangailanhao@163.com

C. Wen
e-mail: wencl@hdu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_80,

80.1 Introduction

The importance of road signs (RS) recognition shows in the following two aspects: First, it can help the driver to understand road information and guide their driving. The RS can also reduce the driver's fatigue, and thus decrease traffic accidents caused by the fatigue driving. Second, RS recognition is a key issue in the guiding the intelligent vehicle's driving.

The RS recognition methods include the color [1] and the shape [2, 3] based methods. The color-based method uses the color threshold to segment and identify the RS. Nevertheless, the color pixel value is often affected by light conditions and the weather. The shape-based method identifies the RS just from the shape angle, however, two RS shapes may have similar appearance (such as go straight or not and turn left, banned left, and so on). So the shape-based method does not recognize the RS correctly. The RS recognition with these drawbacks, the paper considers the correlation of the pixels, gray, the particularity of RS, and the shape information through Markov random field (MRF) modeling and shape moment invariants matching.

The article uses moment invariants method such as the image of the translation, rotation, gray, scale, and so on, which are not sensitive to distortion and noise matching the shape of road signs image and recognizing the road signs. This method is widely used in pattern recognition, image classification, target recognition, and scene analysis.

80.2 Image Modeling Using Markov Random Fields

Image modeling is divided into two parts: one is to search objective optimization function. The other is to get the solution of the objective optimization function. We use MRF to build the optimization function. Markov random field [4, 5, 6, 7] can describe the correlation. It has the following positivity and Markovianity.

$$P(f) > 0, \forall f \in F \text{ (Positivity)} \tag{80.1}$$

$$P(f_i|f_{S-\{i\}}) = P(f_i|f_{N_i}) \text{ (Markovianity)} \tag{80.2}$$

where $S-\{i\}$ is the set difference, $f_{S-\{i\}}$ denotes label set at the sites in $S - \{i\}$, f_{N_i} stands for label set at the sites neighboring i.

According to the equivalence [8] Markov random field and Gibbs field, the paper uses Gibbs field to describe the RS image. The Gibbs distribution is defined by

$$P(f) = Z^{-1}e^{-U(f)} \tag{80.3}$$

where $Z = \sum_{f \in F} e^{-U(f)}$ is normalizing constant. $U(f)$ is energy function. $U(f) = \sum_{c \in C} V_c(f)$ is a sum of clique potentials $V_c(f)$ over all possible cliques C. Obviously, Gaussian distribution is a special condition of the Gibbs distribution. The paper uses second order Markov random field and recognizes the RS image as a mixture distribution model. So, the RS image is described below

$$f(d_i|\theta) = \pi_1 f_1(d_i|\theta_1) + \cdots \pi_j f_j(d_i|\theta_j) + \cdots \pi_m f_m(d_i|\theta_m) \tag{80.4}$$

where d_i is the observed data, that is, the noise image in this paper. π is the weight. When it is Gaussian mixture model $\theta = \{\mu, \sigma\}$, μ is the mean of the Gaussian distribution (here, recognizes color as a Gaussian distribution), σ is standard deviation.

Objective optimization function uses Maximum a posteriori (MAP) rule [9] of the MRF. The combination of the MRF and the MAP rule which is implied to the image processing were first proposed by Geman and Geman. Prior distribution considers the relevancy of the pixels of the image and is described by Gibbs field [10].

When noise follows Gaussian distribution, the likelihood distribution is

$$p(d|f) = \frac{1}{\prod_{i=1}^{m} \sqrt{2\pi\sigma^2}} e^{-U(d|f)} \tag{80.5}$$

where $U(d|f) = \sum_{i=1}^{m} (f_i - d_i)/2\sigma^2$ is the likelihood energy. With the prior Gibbs distribution, the posterior probability is given by

$$P(f|d) = \propto e^{-U(f|d)} = e^{-[U(d|f)+U(f)]} \tag{80.6}$$

where $U(f|d) = U(d|f) + U(f) = \sum_{i=1}^{m} (f_i - d_i)/2\sigma^2 + \sum_{i=1}^{m} (f_i - f_{i-1})^2$ is the posterior energy. The MAP rule minimizes the posterior energy function

$$\widehat{f} = \arg\min_{f} U(f|d) \tag{80.7}$$

where d, f are the gray pixel values for the noise image and restored image, respectively; $\widehat{f}$ is the estimation of observing image corresponding standard image.

80.3 Unsupervised Learning Process of Expectation–Maximization (EM) Algorithm

In actual transport environment, collected images have noises and the number of components in the collected image is unknown. We combine the EM algorithm, the MRF-MAP theory and the technology of component management [11] to get

the removed noise image. Combining the equations (80.4), (80.7) and (80.9) can get equations (80.10) and (80.11). Solving this problem is getting the solution to the model (80.7). About the second problem, EM algorithm optimizes objective optimization function through choosing the number of components according the experience, not the real number of components in the RS image. In order to get the correct number of the components, the paper combines the EM algorithm and the technology of component management. We list the algorithm as follows.

(a) *Initialization*. In Gaussian mixture model, the means are initialized by uniform in the data plane. Variances adopts unit matrix. Weights are with equal values.

$$\mu_m = \frac{m \times 255}{cnum},\ \sigma_1 = \cdots = \sigma_m = 1, \pi_1 = \cdots \pi_m = 1/cnum \tag{80.8}$$

where *cnum* is the number of the distributions in the observing RS image.

(b) *E step*. Calculating the expectation of likelihood function of Gaussian mixture distribution.

$$Q(t) = \sum_{i=1}^{n} [e_{i1} \log P_1(d_i|\theta_1) + \cdots + e_{im} \log P_m(d_i|\theta_m)] \tag{80.9}$$

(c) *M step*. Generally, searching max value through numerical optimization method cannot get a specific analytical formula. While under Gaussian mixture condition, we can get the analytical formula in M step.

$$e_{ij} = \frac{\pi_j N(d_i, \mu_j, \sigma_j^2)}{\pi_1 N(d_i, \mu_1, \sigma_1^2) + \cdots \pi_m N(d_i, \mu_m, \sigma_m^2)} \tag{80.10}$$

$$\mu_j = \frac{\sum_{i=1}^{n} d_i e_{ij}}{\sum_{i=1}^{n} e_{ij}}, \sigma_j = \frac{\sum_{i=1}^{n} (d_i - \mu_j)^2 e_{ij}}{\sum_{i=1}^{n} e_{ij}}, \pi_j = \frac{\sum_{i=1}^{n} e_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m} e_{ij}} \tag{80.11}$$

where e_{ij} is labeling variable, i is the pixel number, j is the distribution number. d_i is the observed value. μ_j is the mean value of the jth distribution. σ_j is the variance of the jth distribution. π_j is the weight of jth distribution.

(d) Combining and deleting components. The rule is

$$\|\mu_1 - \mu_2\| < \lambda_\mu, \|\sigma_1 - \sigma_2\| < \lambda_\sigma \tag{80.12}$$

$$\pi_j < \lambda_\omega \tag{80.13}$$

where λ_μ, λ_σ are the threshold of the mean and variance.

(e) Return to (b), until all parameters are convergent.

80.4 Recognizing Road Signs through Moment Invariants Method

Collected images are always influenced by noise. For example, rotation, size change, shift occur in an image. Therefore, the paper adopts the moment invariants method [12, 13, 14] which is not sensitive to the noise. In this section, use the shape information of the RS images.

The probability density function of continuous random function of two dimensions $f(x, y)$ with $(p + q)$order moment is defined as follows,

$$m_{pq} = \int_{-\infty}^{+\infty} x^p y^q f(x, y) dxdy, p, \ q = 0, 1, 2, \ldots \tag{80.14}$$

Regulation moment of discrete function $f(m, n)$ with two dimensions $(M \times N)$ is defined as follows,

$$m_{pq} = \sum_{m=1}^{M} \sum_{n=1}^{N} m^p n^q f(m, n) \tag{80.15}$$

Complete set of n order regulation moment include all $\{m_{pq}\}$ which satisfy $p + q \leq n$. Thus it has $\frac{1}{2}(n + 1)(n + 2)$ elements. Monomial product $x^p y^q$ is the basement function defined by the corresponding moment. According to Uniqueness Theorem, it is known that if $f(x, y)$ is a bounded function, and it is piecewise and continuous. That is, if limited area of the x, y plane exist nonzero value, then all moments exist moment sequence $\{m_{pq}\}$ and $f(x, y)$ can be unique confirmed.

Like equation (80.14), the central moment of $f(x, y)$ is defined as follows,

$$\bar{\mu}_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q dxdy \tag{80.16}$$

Where $(\bar{x}, \bar{y})$ is barycentric coordinate of the image; x, y is the coordinate of the image; $f(x, y)$ is the value of the pixel.

For the RS images, the paper uses two and three order central moments to make up the moment set Hu moment invariant. It is invariant to translation, rotation and size scales. It can get the RS shape information through Hu moment. The different shapes correspond the different value of Hu moment. For example, the sixth and eighth standard images in the Fig. 80.1 have different shape information. That is, they have different moment values. From Table 80.1, it can be seen the sixth and eighth standard images have different moment values. So, we can get the RS image shape information through moment invariants method.

The error between restored image and standard images adopts following equation,

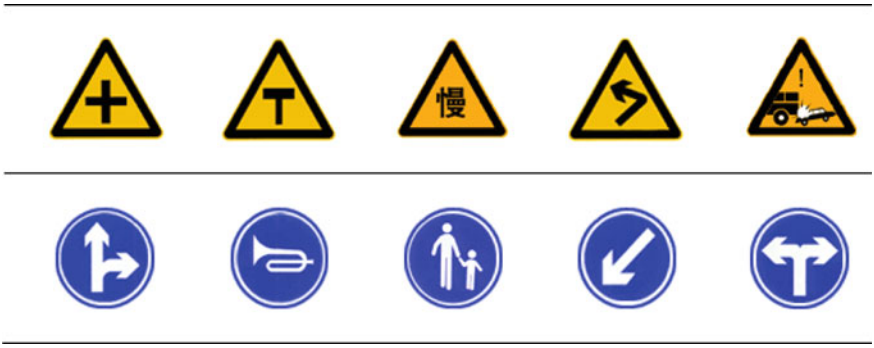


Fig. 80.1 Select 10 standard RS images randomly. The sequence is labelled as numbers 1,2,...,10 from top-left to lower-right

Table 80.1 The moment invariant values of the standard RS images 6 and 8

Moment image	1	2	3	4	5	6	7
The 6th RS	15.9	178.2	28.2	35.3	1110.6	−0.02	−121.1
The 8th RS	16.2	183.3	28.6	37.3	1216.6	0.08	50.0

$$Er = \sum_{i=1}^{7} |r_i - s_i| \tag{80.17}$$

r_i, s_i is the ith moment invariant value of restored image and standard image, respectively.

80.5 Simulation Experiment

In this section, we choose 10 standard RS images randomly. The sixth RS was added to the Gaussian noise and had some rotation. Removing the noise of the sixth RS uses the MFR and the EM algorithm, respectively. Finally, we calculate the invariant moment value of the RS image which was removed the noise and the standard RS images.

It can be seen that the proposed algorithm has a clear image in Fig. 80.2. However, Gaussian smoothing method and median filtering method are fuzzy. This shows that the proposed method is easier to identify the RS than Gaussian smoothing method and the Median filtering method (Figs. 80.3, 80.4). From Fig. 80.3, the proposed method can better fit the real distribution.

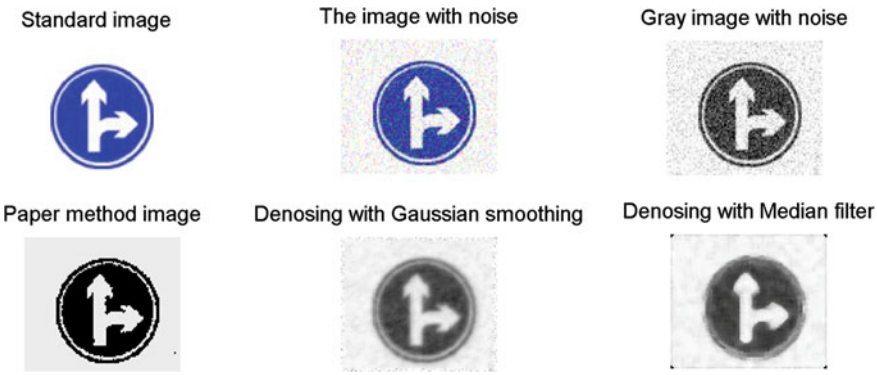


Fig. 80.2 The RS images restored through three methods

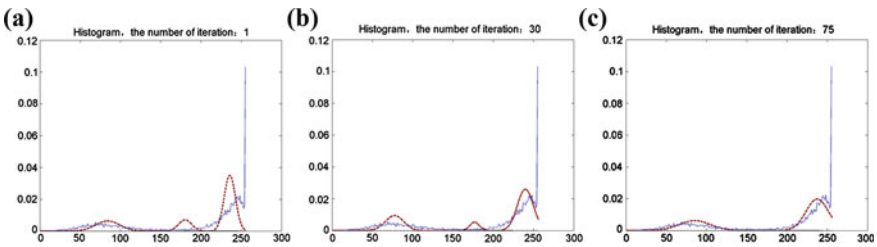


Fig. 80.3 The approximate figure of iterating curve. **a**, **b**, **c** give the unsupervised learning process of the EM algorithm in difference iteration step. Where blue line represents the curve of gray image with the noise; red line represents learning gray curve of the EM algorithm; black line represents the individual distribution

In Fig. 80.4, the errors of the first five standard images are larger than the latter five. The main reason is that the restoration image is different from the first five standard images both in the shape and the color information, but only the shape information is different in the last five ones. So the error of the moment invariant is much closed to that in the first and third images. The proposed method is more accurate than the other two methods in the step of recognizing the RS.

Table 80.2 lists the error of moment invariant between the images removed noise by the proposed method, Gaussian smoothing method and Median filtering method, respectively. We use the three methods to all identify the RS. And the accumulated sum of the error with the standard image is 0.375, 0.958 and 0.735 using Eq. 80.17. Obviously, the value of the proposed method is the smallest. It is twice smaller than the other two methods and can better identify the RS.

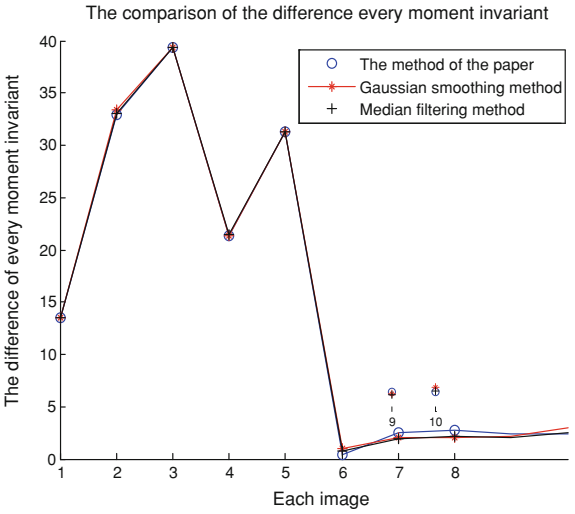


Fig. 80.4 The error of moment invariants. Horizontal label represents ten standard images; longitudinal label represents the error of moment invariants. The blue line represents the error of moment invariant between the unknown image and the standard image; The red line represents the error by the Gaussian smoothing method; The black line represents the error by the median filtering method

Table 80.2 The error of the moment invariants

Standard image		1	2	3	4	5
Image after removing noise	Paper	13.477	32.913	39.318	21.373	31.248
	Gaussian	13.462	33.414	39.303	21.358	31.233
	Median	13.534	33.079	39.375	21.431	31.305
Standard image		6	7	8	9	10
Image after removing noise	Paper	0.375	2.513	2.748	2.455	2.363
	Gaussian	0.958	2.093	2.082	2.198	2.972
	Median	0.735	1.931	2.175	2.039	2.529

80.6 Conclusion

RS image recognition by using Markov random fields is discussed in this paper. The proposed algorithm consists of two major parts: Building objective optimization function and finding the solution of the objective optimization function. The paper can get objective optimization function by adopting MAP rule, and get the solution of the objective optimization function by adopting unsupervised learning of EM algorithm. In real image, the number of components in the RS is unknown. We assume that the number of components in the RS is not beyond a certain

empirical value. After obtaining the RS image restored, we adopt the method of moment invariants to recognize the RS. Simulation results show that the parameters such as mean value and variance in the paper have a better estimation than Gaussian smoothing method and Median filtering method. Nevertheless, in some cases, this paper estimates the number of components is not accurate enough, so that the effect of RS image restored is not ideal, and future research may focus on this aspect.

References

1. Zhang QZ, Kamata SI (2008) Automatic road sign detection method based on color Barycenters hexagon model. In: The 19th international conference on pattern recognition, 8–11 Dec, Tampa, Florida, U S A, pp 1–4
2. Broggi A, Cerri P, Medici P, Porta PP, Ghisio G (2007) Real time road signs recognition. In: The 2007 IEEE intelligent vehicles symposium, Istanbul, Turkey, pp 981–986
3. Miura J, Kanda T, Shirai Y (2000) An active vision system for real-time traffic sign recognition. In: The 2000 IEEE intelligent transportation systems proceedings, Dearbom, U S A, pp 52–57
4. Li SZ (2009) Markov random field modeling in image analysis. Springer, Berlin, pp 199–208
5. Lu MJ, Wang RS (2000) The markov random fields in computer vision. J Electronics Inf Technol 22(6):1028–1037. (in Chinese)
6. Chen WK, Zhang LZ, Zhang QZ, Zeng T, Cui G (2007) Segmentation of moving object based on markov random field. Microcomput Appl 7(28):673–677. (in Chinese)
7. Wu Q, Zhao XB, Zhou HQ (2006) Detection of vane defects based on markov random field model. Comput Appl 9(6):187–212. (in Chinese)
8. Hammersley JM, Clifford P (1971) Markov field on finite graphs and lattices. Unpublished
9. Geman S, Geman D (1984) Stochastic relaxation Gibbs distribution and Bayesian restoration of image. IEEE Trans Pattern Anal Machine Intell 6(6):721–741
10. Wang KQ, Bai XB (2006) Classification of wood surface texture based on Gauss-MRF model. J Forestry Res Northeast Forestry Univ Harbin China 17(1):57–61
11. Liu WF, Han CZ, Shi Y (2009) Unsupervised learning for finite mixture models via modified gibbs sampling. J Xi'an Jiaotong Univ 2(43):15–19
12. Wang HX, Xu XF, Xu SY (2004) An invariant moment approach to recognize objects. Opt Precis Eng Z2(12):1–6. (in Chinese)
13. Xia YQ, Liu ZD, Yang JY (2005) Application of invariant moment approach in region matching. J Comput Aided Design Comput Graph 10(17):2152–2156. (in Chinese)
14. Qi SS (2008) Studies on recognition of cell images and initial application based on invariance moments. Inner Mongolia Agriculture University, Huhhot, China, pp 10–15. (in Chinese)

Chapter 81
Soccer Robots' Color Logos Recognition Based on HSI Model and Eigenvalues

Lei Yao, Deng Xiaolu and Wang Yufeng

Abstract In robot soccer games, target recognition is a very important technology. In order to improve the accuracy of recognition, based on the analysis of the color library, HSI color model is used to replace RGB color model and some eigenvalues are used as the key factors to recognize robots' color logos. Experiment shows that the location of color logo can be recognized accurately. The theory has been successfully applied in Climber Robot Soccer Team.

Keywords Robot recognition · Color library · HSI model · Eigenvalues

81.1 Introduction

Target recognition is to extract the target from an image, and replace the non-targets with the background. The traditional method is to determine the target based on RGB model. However, in a real environment, light changes lead to the dramatic changes in RGB values, which often results in misjudge even the target cannot be recognized. So we choose HSI model to present colors. Because the colors used in logos are blue, green, red, and yellow, the differences among them are great. We propose some eigenvalues that can describe the differences. By these two methods, the accuracy of recognition improves a lot.

L. Yao (✉) · D. Xiaolu · W. Yufeng
Wuhan University of Science and Technology, Wuhan, China
e-mail: 846985325@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_81,

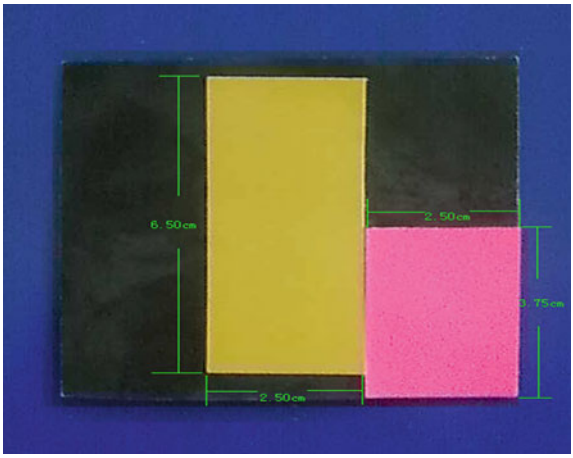


Fig. 81.1 The *blue* team

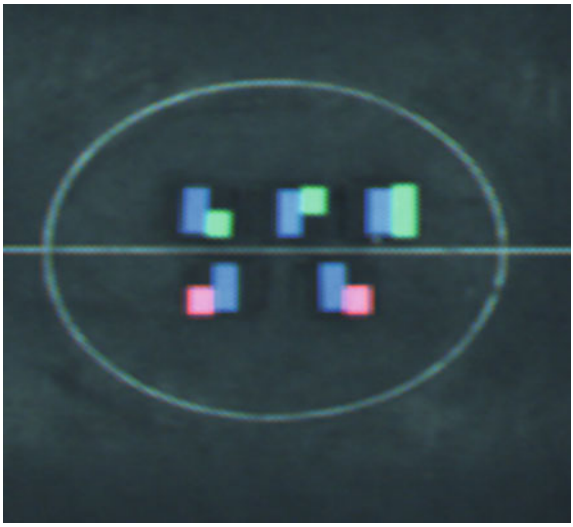


Fig. 81.2 The format of color logo

81.2 Analysis of Color Logos

Color logos for robots have the following characteristics:

First, the kinds of colors are few. There are four colors altogether. They are blue, green, red, and yellow. The differences among them are great, which is good for recognition. The background color is black which is good for extraction.

Second, the combination of these colors is the key to identify different robots. See it in Fig. 81.1. One robot is identified by three colors, the main color, the first secondary color, and the second secondary color. The main color logo is 6.5 cm long and 2.5 cm wide. The two secondary colors are both 3.75 cm long and 2.5 cm wide (Fig. 81.2).

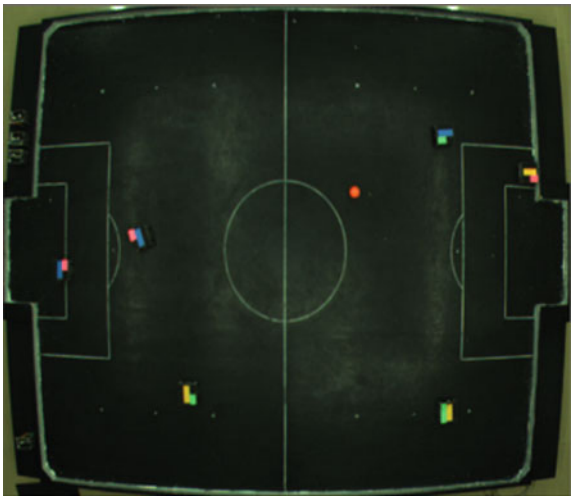


Fig. 81.3 The image captured by image card. There are three *blue* robots

81.3 Color Library

Soccer robot vision system should meet the following requirements:

(1) Identify the position and direction angle of each robot's correctly.
(2) Good real-time ability, a high number of frames' process.
(3) Stable operation
(4) Color library should be complete but not too large.

Establishing color library of the targets is one of the main tasks of a soccer robot vision system [1]. The quality of color library determines the accuracy of recognition. Too small a color library leads to the result of segmentation instability, even makes the targets lost because the segmentation is too small. To the contrary, too large a color library will introduce a lot of noise, even color mixing phenomenon. Additionally, a good color library should have a good maintainability, easy to adjust (Fig. 81.3).

81.4 Color Model

The defects of RGB color model

Because the lookup table does not support fuzzy search, while the same color logo's color information changes in a very wide range [2]. It often takes a long time to repeat sampling to generate a complete color library. It is very difficult make all the information covered by the color library. As a result, identification results often have empty zone. What is more, this lookup-table has a poor anti-interference

Fig. 81.4 The result of recognition. The target color is *blue*

ability to changes of environment. Once the environment changes, like the change of light, it needs to sample again. This cannot adapt to the rapid adjustment.

2. HSI color model

HSI color model uses hue, saturation and intensity to represent colors [3]. It is similarly to the way objects are observed by human. In HSI space, H describes the property of color. S describes the degree of dilution by while I reflect the intensity of light.

The equations of changing RGB model to HSI model are as follows:

$$H = \begin{Bmatrix} \theta & B \le G \\ 360 - \theta & B > G \end{Bmatrix} \ni = \cos\left\{ \frac{[(R-G)+(R-B)]/2}{\sqrt{[(R-G)+(R-B)(G-B)]}} \right\} \tag{81.1}$$

$$S = 1 - 3/(R+G+B) * [\min(R, G, B)] \tag{81.2}$$

$$I = 1/3(R+G+B) \tag{81.3}$$

The model's character is that same color's distribution in this space is continuous, that is to say the target's colors are continuous in HSI space and in different regions of color space [4]. To create a color library in HSI space, we only need to record the coverage of H, S, I of the subject color, which means that we should record the maximum and minimum values of H, S, I. In HSI model, it is easy to establish color library because we just need to place the color logos at the several most special areas of the venue, such as gate area, the circle, and corner. Sampling in these areas can determine the range of the three components. This kind of color library is easy to adjust and maintain. Because what most image capture cards provide is RGB color model, we need to generate conversion table in the memory. When recognizing each pixel, the RGB information is transformed into HSI information to determine whether it is the points to be searched (Fig. 81.4).

Table 81.1 The conditions of identifying colors

Color	FT1	FT2	FT3
Blue	≥ -2	<0	>0
Green	>0	<0	≤ 0
Red	≤ -3	>0	<0
Yellow	≥ 0	>0	<0

Table 81.2 The results of using HSI model and eigenvalues

Time (min)	Found	Lost	Accuracy (%)
0	11	0	100
5	11	0	100
10	10	1	91
15	9	2	82
30	9	2	82

81.5 Eigenvalues

H component is used to describe the hue of image. It is the main basis to identify different colors [5–7]. H component is multiplied by $\pi/180$ to convert to 0°–360° to correspond to circle angle. Followed the same plane is red, yellow, green, cyan, blue, magenta. Saturation also has a very important effect on color, so we have to consider its effect, but we can use it without processing. However, these are not enough. Because the color of background and the line have influence on the identification, we need to apart them from the robots. What is more, the light changes during the time, the same color may have different shades. So we propose several features that can improve the effect of identification.

$$\mathrm{TI} = (32/4) * (I/32 + 2) \tag{81.4}$$

$$\mathrm{FT1} = (2 * G - B - R)/4 \tag{81.5}$$

$$\mathrm{FT2} = (2 * R - B - G)/4 \tag{81.6}$$

$$\mathrm{FT3} = (2 * B - R - G)/4 \tag{81.7}$$

In Eq. (81.4), 4 presents the total grads of intensity, 2 is the average grads. The Eq. (81.4) is used to reduce the affect the background and white line have on the robot. According to our experiment, it is often greater than 16.5. Other three equations are used to improve the identification of robots. Because the colors which are used are quite different, the values calculated are in different regions, see in Table 81.1. What is more, the other three eigenvalues can also be used to reduce the influence of black background and white line.

81.6 Test Code

```
bool CLiveViewDlg::determineColor(int no,int m, int n, RGBTRIPLE *P)
    {
    int r,g,b,data;
      int FT1,FT2,FT3;
      r=((P+m+(479-n)*640)->rgbtRed)/8;
      g=((P+m+(479-n)*640)->rgbtGreen)/8;
      b=((P+m+(479-n)*640)->rgbtBlue)/8;
      if((r&0x1f)==0) r=32;
      if((g&0x1f)==0) g=32;
      if((b&0x1f)==0) b=32;
      FT1 = (2*g-r-b)/4;
      FT2 = (2*r-g-b)/4;
      FT3 = (2*b-r-g)/4;
      if(((GetIValue(RGB(r,g,b))/32+2)*(32/4)>=16.5)&&(abs(FT1)+abs(FT2)+a
bs(FT3))>=5)
if(GetHValue(RGB(r,g,b))<=colorScale[no].HMax*1.1&&GetHValue(RGB(r,g,b))>=
colorScale[no].HMin*0.9)
if(GetSValue(RGB(r,g,b))<=colorScale[no].SMax*1.1&&GetSValue(RGB(r,g,b))>=
colorScale[no].SMin*0.9)
if(GetIValue(RGB(r,g,b))<=colorScale[no].IMax*1.2&&GetIValue(RGB(r,g,b))>=
colorScale[no].IMin*0.8)
                    switch(no)
                    {
                    case 0: if(FT1>=-2&&FT2<0&&FT3>0)
                       return true;
                            else if(FT1>=0&&FT2>0&&FT3<0)
                       return true;
                                                  break;
                    case 1: if(FT1>0&&FT2<0&&FT3<=0)
                                          return true;
                            break;
                    case 2: if(FT1<=-4&&FT2>0&&FT3<0)
                                          return true;
                                                 break;
```

```
            case 3: if(FT1>=-2&&FT2<0&&FT3>0)
                                    return true;
                                else if(FT1>=0&&FT2>0&&FT3<0)
                                return true;
                                break;
            case 4: if(FT1<0&&FT2>0&&FT3<=-5)
                 return true;
            break;
            default : return false;
            }
            return false;
        }
```

81.7 Algorithm of Identification

The algorithm of identification are as follows:

Get the image's RGB values
Change RGB model to HSI model
Calculate the eigenvalues
Identify the color logos by eigenvalues

81.8 Result

In order to test the accuracy of recognition, we use Climber robots to do the experiment. This method has been successfully used in our Climber robots. Runtime Environment is: 2.6 GHZ of CPU, the memory is 1 G, and the camera is basler1394 digital camera (Tables 81.2, 81.3).

81.9 Summary

By using HSI color model, the color library is more accurate and reliable. It can cover most of values of each color logo. The use of eigenvalues makes recognition more precise. It can also reduce the affect of background and white line. The experiment shows it is a good method and it has got good results in practical applications.

Table 81.3 The results of using RGB model

Time (min)	Found	Lost	Percentage (%)
0	11	0	100
5	9	2	82
10	7	4	64
15	6	5	55
30	5	6	45

References

1. Haibo L, Jing S, Song G (2010) Digital image processing using visual C++. China Machine Press, Beijing, pp 524–544
2. Gonzalez RC, Woods RE (2009) Digital image processing, 2nd edn. Publishing House of Electronics Industry, Beijing
3. Xiaocan Z (1997) Remote sensing digital image processing. Zhejiang University Press, Zhejiang
4. Zheng Z, Yanping W, Guixiang X (2010) Digital image processing and machine vision. Posts and Telecom Press, Beijing
5. http://www.docin.com/p-15427511.html
6. http://www.docin.com/p-50186368.html
7. http://www.docin.com/p-32021447.html

Chapter 82
Application of Image Extraction Technology Based on Edge Detection Method

Yan Yang and Guo-zheng Zhou

Abstract Edge detection algorithm will directly affect the development of the performance of the system. In this article, the image edge detection algorithms are analyzed. We present a practical solution: detect the edge of the image again to deal with binary. A feasible solution is to detect the edge of the image again to deal with binary, the result is similar to the Canny operator edge effect, but the edge of the smoothness is a bit poor. A more effective method needs to be further studied.

Keywords Edge detection · Sobel operator · Kirsch operator

Goal of understanding the human visual system is divided into two steps: first, to separate the image edge and background; then, to perceive the image details, recognize the image contours. Computer vision is the process of imitating human vision [1]. Therefore in object detection edge, first, its contour points are roughly detected, then through the linking rules originally detected contour points are connected, and link the missing boundary point and remove the false boundary point. The edge of the image is the important characteristic of image, and it is the basis of computer vision [2], pattern recognition and so on. However, edge detection processing is a problem, because the image edge of actual scene is often a combination of types of edges and their fuzzy, and noise exist in the actual image signal. Noise and edge belong to the high frequency signal, so it is difficult to use frequency bands to choose.

Y. Yang (✉)
Department of Art Design, Jiyuan Vocational and Technical College,
Jiyuan 454650, Henan, China
e-mail: wyzzxx@163.com

G. Zhou
Office of Academic Affairs, Jiyuan Vocational and Technical College,
Jiyuan 454650, Henan, China
e-mail: zgz.zhou@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_82,

82.1 The Development of Digital Image Edge Detection

Digital image processing technology and the theory is an important realm of the computer application. “Figure” is the objects of transmitted light or reflected light distribution. “Image” is the impression or understanding of the visual system for figure receiving in the brain of the formation. Edge detection is the important content of image processing. The edge is the most basic features of the image. Edge is defined around the pixel gray with step change or roof changes of those pixels set. The image of the object is related to the boundary but it is different. The edge is a reflection of image discontinuity [3].

In 1959, edge detection of a word is put forward the earliest. Roberts begins the earliest system study based on edge detection [4]. Canny summarizes the theory and the practice achievement in 1986, and put forward “edge detection”. He proposed three principles: good detection results, good positioning and edge low-grade repetition response, and gives their mathematical expressions. It is a continuous utilization guidelines [5].

Through long-term research, Dider Demigny presents discreteness criterion of edge detection quality, he found that under the guidelines in this discrete role of the third person criterion threshold operation can be replaced, then get the discrete Canny standards under the optimal linear filter.

82.2 The Edge of Image and Its Characteristics

The edge of image the most significant part of the partial image brightness changes, and refers to a dramatic change in the gray level. On the edge of the adjacent two-dimensional image pixel gray value transform from one to another: and the gray variation curves show singular signal waveform: step signal or a roof-shaped pulse. The gray value changes significantly on both sides of step-edge; and the intermediate gray and the edge on both sides gray have obvious difference of roof-edge. Gray variation curves and its derivatives are often used in mathematics to describe the edge of change. The two edges are respectively for one order, two order derivative: the edge of a step-derivative point A to reach maximum, and the two order derivative cross at the point A and zero; one derivative cross at the ridge points B and zero point of roof-edge and at point B, the second derivative to reach extreme.

82.3 Several Main Edge Detection Methods

The first kind is testing the maximum gradient. Due to the edge generally occurs in the local where images’ gray value changes greater, so the derivative of the better operators will become a line of thought. Roberts operator, Prewitt operator, Sobel

operator, etc. are relatively simple and common examples. There is also a more intuitive approach to use some of the current pixel neighborhood of pixel values to fit a curved surface, then find the edge of the surface gradient at the current pixel.

The second is to test the zero-crossing point of the second derivative. This is because the edge of the gradient obtains the maximum value (positive or negative), we know that turning points of function of the two order derivative is 0.

Third, statistics-based method such as the use of hypothesis testing to detect the edge, Marimont analysis second order zero intersection to get probability of each pixel, and thus get the edge detection program.

Fourth, the wavelet multi-scale edge detection. In the nineties of the twentieth century, with the rapid development of the wavelet analysis, wavelet was used for edge detection of non-stationary signal. As a tool, wavelet in edge detection has the advantage. Mallat has done a lot in the work.

Of course, there are some other methods. For example, fuzzy mathematics method proposed many methods, such as using the edge flow to detect edge, edge detection based on integral transform methods, and tensor-based edge detection method [6].

82.4 Edge Detection Algorithm

Digital image is a 2D digital signal. In the practical application, the region mask template for convolution calculation is used. Although there are many kinds of operators, and new operators appear ceaselessly, but none algorithm can fit any common image. When solving specific problems, for the different characteristics of the image, even for the different characteristics of the region, we can put forward one or several suitable algorithms in order to achieve the purpose of identifying. Because of the particularity of recognition, we must consider the operator of noise sex, directional and computational complexity in the choice of edge detection algorithm.

Edge detection algorithm has a lot of methods, a gradient operator, the direction operator, Laplasse operator and Canny operator and so on.

82.4.1 Several Commonly Used Operators

(1) Sobel edge operator

Sobel marginal operator mask template is two 3×3 convolution kernels (shown in Table 82.1). Sobel marginal operator emphasizes the center pixel of the 4 edge pixels and its influence, and weaken the 4 diagonally adjacent pixels of its role [7]. It is directional, in the horizontal direction and the vertical direction is formed on the strongest edge.

Table 82.1 Sobel convolution kernel

1	2	1
0	0	0
−1	−2	−1
−1	0	0
−2	0	2
−1	0	1

(2) Kirsch edge operator

Kirsch edge operator is composed of a set of eight 3 × 3 convolution kernel composition (shown in Table 82.2). The angle between each of the two convolution kernel is 45°. For image T, module Wk (k = 1, 2,…, 8), the edge intensity at the point (x, y) is $E(x, y) = \max\{\mathrm{W}k \bullet \mathrm{T}\}$ $\{k = 1, 2, \ldots, 8\}$, the operator take care of image 360° in the 8 direction on edge extraction.

(3) Laplacian operator

Laplacian operator is non-directional operator with a convolution kernel can process the image. Typical convolution kernel is shown in Table 82.3.

(4) Robert operator

Robert operator is 2 × 2 operator, using partial differential operator to find the edge, calculating the first-order differential along direction of 45°. The gradient of the image is the sum of vector for the two 45-degree gradient. It does not contain smooth, and it cannot restrain noise, it responds best to the image of steep low noise.

(5) Canny operator

Canny operator is a first-order operator, the method is essentially using a quasi-Gaussian function for smooth operation, and then positioning derivative maximum value with the direction of first-order differential operators. It can be used to approximate the gradient of Gaussian function, in theory it is very close to the best edge operator with linear combination of four exponential functions. In actual application, its programming is more complex, and the operation is slower. It is the traditional first-order differential effect detected in the best step-edge one operator. It is the best one of the traditional first-order differential operators in order to detect step edge [8]. It is stronger than Prewitt operator, Sobel operator, Laplacian operator in denoising ability, but it is also easy to smooth out some edge information.

(6) LOG operator

(Laplacian of Gaussian) Marr-Hildreth first use the Gaussian function for smoothing original image, then using Laplacian operator operation without the direction, and then extracting zero-crossing point of the algorithm for edge

Table 82.2 Kirsch edge operator

5	5	5	−3	5	5	−3	−3	5	−3	−3	−3
−3	0	−3	−3	0	5	−3	0	5	−3	0	5
−3	−3	−3	−3	−3	−3	−3	−3	5	−3	5	5
−3	−3	−3	−3	−3	−3	5	−3	−3	5	5	−3
−3	0	−3	5	0	−3	5	0	−3	5	0	−3
5	5	5	5	5	−3	5	−3	−3	−3	−3	−3

Table 82.3 Laplacian operator

−1	−1	−1
−1	8	−1
−1	−1	−1

Fig. 82.1 Gray image

detection, its accuracy has improved significantly. Its advantage is filtering the noise, it may also smooth the original edges of the image, Gauss function of the variance directly affects the results of edge detection. δ is smaller for testing details, δ is larger for the detection of contours.

$G(x,y) = \frac{1}{2\pi\delta^2} e^{\frac{r^2}{2\delta^2}}$, the Gauss–Laplasse operator can be obtained according to the convolution.

$$G(x,y) = \frac{1}{\pi\delta^4}\left[1 - \frac{x^2+y^2}{2\delta^2}\right] e^{-\frac{x^2+y^2}{2\delta^2}}$$

82.4.2 Comparison and Analysis of Several Operators

License of the original image into a grayscale after treatment, shown in Fig. 82.1.

The Sobel operator processing image is almost cannot tell the border, as shown in Fig. 82.2. It is because the original image have the mottled sunlight and shadow, the highlight is not the edge of the license. As the shooting angle, license plate frame does not present a horizontal or vertical direction.

Fig. 82.2 Sobel operator processed image

Fig. 82.3 Kirsch operator processed image

Fig. 82.4 Laplacian operator

The image is processed after Kirsch edge operator as shown in Fig. 82.3, it is a little brighter than the image processed with convolution of the image, but we still distinguish the border.

Laplacian operator is a second derivative operator. As mentioned earlier, the second derivative of the step edge will be generated at the edge of a steep zero-crossing. Figure 82.4 shows that it is the best of several algorithms because the image processed by the Laplacian operator will have a clearer edge.

82.5 Summary

Edge detection in image processing and computer vision occupies a special position, it is one of the most important part in the underlying visual processing. The proposed multi-scale edge detection algorithm based on mathematical morphology has better noise immunity and real-time than traditional edge detection operator. But its denoising effect is not very ideal of Gauss noise image. A feasible solution is to detect the edge of the image again to deal with binary, the result is

similar to the Canny operator edge effect, but the edge of the smoothness is a bit poor. A more effective method needs to be further studied.

References

1. Poggio T, Voorhees H, Yuille A (1985) A regularized solution to edge detection, Technical Report MA, Rep AIM_833, MIT Artificial Intell Lab, May 1985
2. Julez B (1959) A method of coding tv signals based on edge detection. Bell Syst Tech 38(4):1001–1020 Compression, Video Television
3. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell PAMI-8(1):679
4. Demigny D, Lorca FG, Kessal L (1995) Evaluation of edge detectors performances with a discrete expression of canny's criteria. In: Proceedings of the international conference on image processing, Washington, MA, pp 169–172
5. Marimont DH, Rubner Y (1998) A probabilistic framework for edge detection and scale selection. Comput Vis 4–7:207–214 Sixth international conference
6. Zhao-yu P, Li-qun G (2009) An edge detection algorithm based on multi-phase processing. Nat Sci 5:25–29
7. Xi-ling J (2009) A new algorithm of edge detection based on two-value image. Sci-Tech Inf Dev Econ 4:123–124.1
8. Baldonado M, Chang C-CK, Gravano L, Paepcke A (1997) The stanford digital library metadata architecture. Int J Digit Libr 1:108–121

Chapter 83
Research of a Digital Image Watermarking Algorithm Based on DCT Transform

Liang Li and Na Gao

Abstract This paper proposes a digital watermarking algorithm based on Discrete Cosine Transform (DCT) transform. Arnold scrambling technology is used to binary watermarking information for pretreatment. Then, the binary watermarking information is embedded in the corresponding component of the DCT transform to implement watermarking embedding solutions. The algorithm is simple and easy to implement. The Matlab simulation results demonstrate that the algorithm has validity, reliability, resistance to attack, robustness and invisibility.

Keywords Digital watermarking · DCT · Arnold scrambling · MATLAB

83.1 Introduction

The network is widely used today, and copyright protection of digital products and safety certification issues more and more attention [1, 2]. Digital watermarking technology can add the authors and other copyright information to images, video and audio data products and attacks against unauthorized users [3]. The attack on the unauthorized users is robust to prove the creators of digital media copyright.

L. Li (✉) · N. Gao
School of Electrical Engineering and Automation,
Henan Polytechnic University, Henan, China
e-mail: liliang74@hpu.edu.cn

N. Gao
e-mail: gaona@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_83,

This paper presents a discrete cosine transform-based digital image watermarking algorithm, which will have the meaning of the gray watermark image (two-dimensional watermark) embedded in the protected image. The watermarking algorithm can resist various attacks, and has strong robustness.

83.2 Discrete Cosine Transform

Discrete Cosine Transform (DCT) is the cosine of a real number [4], and it is a kind of real number field transformation. Digital image X(m, n) is a matrix of M rows N columns. Two-dimensional DCT can be used for the image transformation from space domain to the DCT-domain, in order to weaken or remove the image correlation data.

Two-dimensional discrete cosine transform is defined as follows:

$$\mathrm{Y}(k,\, l) = \frac{2}{\sqrt{\mathrm{MN}}}\mathrm{c}(k)\mathrm{c}(l)\sum_{m=0}^{\mathrm{M}-1}\sum_{n=0}^{\mathrm{N}-1}\mathrm{X}(m,n)\cos\frac{(2m+1)k\pi}{2\mathrm{M}}\cos\frac{(2n+1)l\pi}{2\mathrm{N}} \quad (83.1)$$

where m, $k = 0, 1, \ldots, \mathrm{M}-1$; n, $l = 0, 1, \ldots, \mathrm{N}-1$.

In which

$$\mathrm{c}(k) = \begin{cases} 1/\sqrt{2} & k = 0 \\ 1 & k = 1,\ 2,\ \ldots,\ \mathrm{M}-1 \end{cases}$$

$$\mathrm{c}(l) = \begin{cases} 1/\sqrt{2} & l = 0 \\ 1 & l = 1,\ 2,\ \ldots,\ \mathrm{N}-1 \end{cases}$$

There are two DCT transforms for a given image: one is the directly DCT transform, then embedded watermark in a specific factor. Another method is to segment image into small 8 × 8 pieces with JPEG compression standard system, and transform each piece with DCT, respectively, then embed watermark information. For more good invisibility and robustness, we use the latter method.

83.3 Arnold Transform

Digital image scrambling technology is a commonly used digital image encryption. For example, If the cycle is A, we will get the same image as the original image after replacement 3 times with Arnold transform. If the original image has been replaced B times, then it will go back to the original image when continue to transform (A–B) times. Through the meaningful watermark preprocessing can enhance the security of the watermark, and improve the robustness of the watermark. So Arnold transformation is widely used in digital watermarking research [5].

Arnold transform Function is:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod N_{(x, y, x', y' \in \{0, 1, 2, \ldots, N-1\})} \tag{83.2}$$

X and y are pixel coordinates of watermark image before transformation. X' and y' are pixel coordinates after transformation. N is the order.

83.4 Digital Watermarking Embedding and Extraction

Watermarking Embedding

The paper use 256 × 256 gray image and 32 × 32 binary watermarking image. The embedding process is as follows:

(1) Arnold transform is used to binary watermarking image pretreatment.
(2) The original image is divided into 8 × 8 size image block (BLOCK(x, y)) which does not cover each other. Then BLOCK(u, v) is obtained to make DCT transform on each block from left to right and top to bottom.
(3) New spectrum coefficient can be obtained to conduct linear transform on one element (W(p, q)) of watermarking image and low-frequency coefficient (BLOCK($k1$, $k2$)) of image DCT transform. The formula is as follows:

$$\text{BLOCK}' = \text{BLOCK}(k1,\ k2) * (1 + a * \text{W}(\text{p}, \text{q})) \tag{83.3}$$

the coefficients($k1$, $k2$) can be regarded as the key, a is embedding strength which generally take 0.005–0.05;
(4) The image block which be embed in watermarking information is conducted IDCT transform.
(5) Combined image block, a image which will be embeded in binary watermarking information can be obtained.

Watermarking Extraction

(1) Every 8 × 8 size image block which contain the watermarking information is conducted DCT transform from left to right and top to bottom.
(2) According to the formula to calculate the value of a

$$a = \text{BLOCK2}(k1,\ k2)/\text{BLOCK1}(k1,\ k2) - 1 \tag{83.4}$$

BLOCK1($k1$, $k2$) is the DCT of original image block, BLOCK2($k1$, $k2$) is the DCT of image block which contain the watermarking image;
(3) According to the value of a, the binary watermarking information can be judged to determine its gray scale value which is 0 or 1;
(4) To repeat (2) and (3) step, until all binary watermarking information are judged;
(5) The proposed watermarking information are conducted inverse Arnold transform, which can recover the watermarking information.

Fig. 83.1 Embedded watermarking image

Fig. 83.2 Original watermarking

83.5 Simulation and Analysis of Results

The paper use Matlab for embedding and extracting algorithm simulation, and use Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Signal-to-Noise Ratio (SNR) as the image quality evaluation standard (Fig. 83.1).

The simulation take the scrambling number of times as 8, a (embedding strength) = 0.02 (Fig. 83.2).

JPEG Compression

JPEG compression is a most common method of digital image processing. Using some image processing tools it can be very convenient for the JPEG compression. Any watermarking system must be able to withstand a certain degree of lossy compression, and can extract watermarking from the compressed image.

From comparison of data in Fig. 83.3, when the compression factor Q is larger, there is less influence. Based on the above data, the algorithm in JPEG compression has good robustness.

Noise

Image in the transmission process will be attacked by noise, such as the Gauss noise. Most of noise obey normal Gauss noise with a mean of 0.

From the data in Fig. 83.4 can be obtained, at the same Gauss noise case, along with the Gauss noise variance increases, images are affected by noise more strongly. But in a certain intensity range, watermarking image and extracted watermarking can still be ensured that the recovery is quite good. It is shown that the proposed algorithm has good robustness to Gauss noise.

Filtering

Filtering is actually an image enhancement method. In the filtering process, the embedded watermarking information is inevitably affected as well. So good

Fig. 83.3 Compressed watermarking image in different values of Q. **a** 15, **b** 35, **c** 75

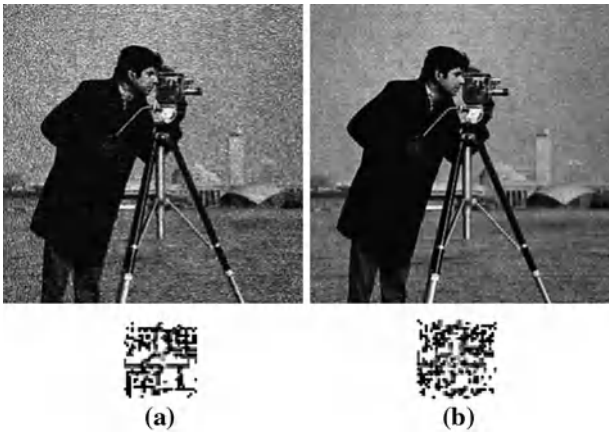


Fig. 83.4 Watermarking image in Gauss noise. **a** variance = 0.002, **b** variance = 0.001

immunity to image processing in a variety of filtering operation, is a watermarking algorithm must be considered. This paper uses the common median filter to attack the embedded watermarking.

From the data in Fig. 83.5 could be drawn, with median filtering template size increases, the influence of noise on image is increasingly strong. However, with the certain template size range, extraction recovery watermarking can still reflect the original watermarking information completely. It is indicated that this watermarking algorithm has good robustness to median filter.

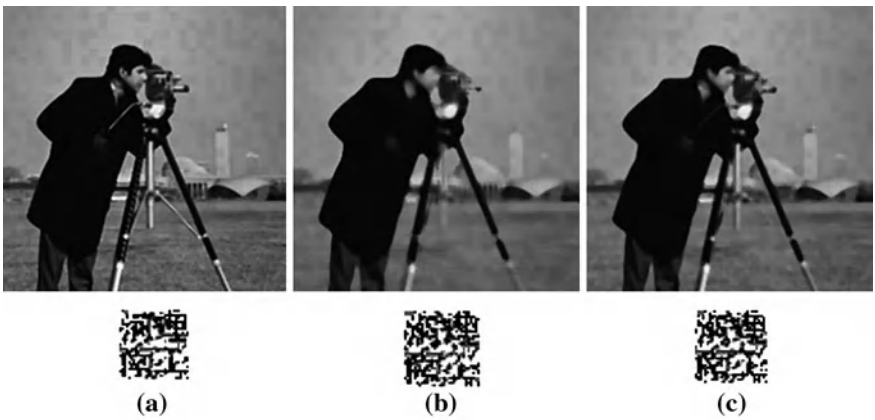


Fig. 83.5 Watermarking image with median filtering. **a** 3×3, **b** 5×5, **c** 7×7

83.6 Conclusion

This paper studies a DCT-based watermark algorithm, in which embedding and extracting Watermark in Low-frequency components, and it is simple and easy to implement. Simulation results show that the algorithm has good invisibility and can resist to compression, noise, filtering and other attacks.

References

1. Liu R, Wang J, Pan X, Deijun C (2010) Low complexity DCT-based distributed source coding with gray code for hyperspectral images. J Syst Eng Electron 21(6):927–933
2. You-kuang Z, Feng Y (2010) A digital watermarking algorithm based on chaos and SVD. Inf Technol Informatization 6:46–48
3. Jie Y, Min J (2010) Scheme of information hiding based on new anti-Arnold transformation and DCT coefficient blending. J Nanjing Inst Technol (Nat Sci Edition) 21(3):38–44
4. Wei E, Baosheng K (2008) The technology of digital image information hiding based on scrambling and amalgamation and its application. Comput Appl Softw 25(5):215–218
5. Xiaosu C, Zhiguang Z, Xiaofeng Y (2009) CDMA watermarking algorithm using WPT and DCT. J Huazhong Univ Sci Technol (Nat Sci Edition) 37(10):5–8

Chapter 84
The MATLAB Realization of Mandarin Digit String Speech Recognition

Peiling Zhang and Lingfei Cheng

Abstract This chapter has established a mandarin digit string speech recognition system which is based on the Mel Frequency Cepstrum Coefficient and the continuous hidden markov model. The system's preprocessing of speech signal, the characteristic parameter extraction, the training of the model, and the recognition algorithm is discussed in this chapter. Meanwhile the system has carried on the simulation and analysis using MATLAB pronunciation toolbox voice box. Some results of the recognition system are given. Several directions for improvement of the system are proposed in the end this chapter.

Keywords Mandarin digit string · Continuous hidden markov model (CHMM) · Speech recognition · Mel frequency cepstrum coefficient (MFCC)

84.1 Introduction

Mandarin digit string speech recognition refers to the pronunciation of a small vocabulary continuous speech recognition whose task is to identify from "0" to "9" in any combination of the ten digit strings [1, 2]. Now it is widely applied to phone voice dialing, home appliance control, industrial control, and other fields.

However, mandarin digit string speech recognition in Chinese has lower recognition ration than in English [3, 4]. The reason is as follows. First, Chinese has single syllable words, and English has more syllables. The experiment showed that

P. Zhang (✉) · L. Cheng
College of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: plzhang@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_84,

lesser the syllable, greater the degree of confusion between the speech. Second, pronunciation of continuous degree of mandarin digit string speech recognition in Chinese is higher than in English. Third, mandarin digit string speech recognition in Chinese has more coarticulation between digital phenomenons than in English.

So the technology of mandarin digit string speech recognition in English which is used in Chinese cannot get high recognition ration. The research value of mandarin digit string speech recognition in Chinese not only has wide application background, but also avails Chinese phonetic feature extraction and acoustic model research. At the same time, it has very significant enlightenment and reference to large vocabulary continuous speech recognition.

This chapter has established a mandarin digit speech recognition system, and has carried on the simulation and analysis using MATLAB pronunciation toolbox voice box to the system [5]. Experimental results show that the system achieves the experiment requirements.

84.2 System Design

Mandarin digit string speech recognition system's realization has two important processes. One is the train of sample modes, the others is the recognition of unknown pattern. In those processes, the extraction of voice's characteristic parameters and the recognition algorithm are two important technologies. The recognition system includes four parts: pretreatment, characteristic parameters extraction, training, and recognition.

84.2.1 Pretreatment

Pretreatment includes preemphasis and endpoint detection. Preemphasis is realized through a digital filter which is $H(z) = 1 - 0.97z^{-1}$ [6]. The endpoint detection detects the speech signal's starting point and end point, where mandarin digit string is divided into isolated numbers using manual methods in speech recognition system's training phase and using voice of short-time energy and the zero rate judge the digital border in speech recognition system's recognition phase.

84.2.2 Extraction Characteristic Parameters

Mel frequency cepstrum coefficients (MFCC) and linear predictive cepstral coding (LPCC) are common parameters [7, 8]. LPCC which is based on the pronunciation of the model simulate the channel characteristic of person from the point of linear forecast. It is synthetic parameters, but this parameter is not making full use of the

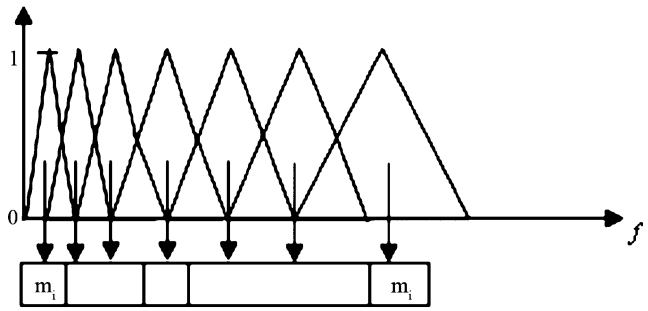


Fig. 84.1 Diagram of Mel scale division filter group

auditory characteristics. According to the results of the experiment to analyze hearing voice frequency, MFCC is based on the mechanism of the auditory. Compared with LPCC, its advantages are not dependent on the pole's speech production model, and in the noise environment shows stronger. A number of studies have suggested that MFCC is superior to LPCC from the recognition ratio and resistance to noise performance. So the system uses MFCC.

MFCC calculation methods are as follows:

1. Using a certain window long (this chapter takes 20 ms) and window move (for 10 ms) divide speech into one by one the overlapping frame. The following steps will signal for each frame ;
2. Each frame gets the high frequency pretreatment through digit filter which is $H(z) = 1 - 0.97\ z^{-1}$;
3. In order to reduce Jibbs effect, hamming windows function is used to each frame for adding window treatment. Its expression is as follows:

$$S_n^* = \left\{0.54 - 0.46\ \cos\left(\frac{2\pi(n-1)}{N-1}\right)\right\} S_n \tag{84.1}$$

Voice signal adding windows processing is done by 512 points discrete Fourier transform (DFT) whose expression is as follows:

$$X(k) = \sum_{K=0}^{K-1} x(n) \mathrm{e}^{-j2\pi k/K} \tag{84.2}$$

where k is 512. Then, signal's power spectrum is calculated which is as follows:

$$P(k) = |X(k)|^2 \tag{84.3}$$

The frequency mentioned above is transformed according to Eq. 84.4, and then power spectrum of each frame is done by Mel scale division filter group which is shown in Fig. 84.1.

$$\mathrm{Mel}(f) = 1127.0148 * \log_2\left(1 + \frac{f}{700}\right) \tag{84.4}$$

where from m_i to m_j is each frequency band's energy. This group of bandpass filter's center frequency shown in Fig. 84.1 is evenly arrayed by Mel scale, and each filter triangle's two bottom point of frequency are equal to two nearby filter center frequency, namely every two adjacent filter is mutual lapped, and sum of frequency response is 1.

Calculating each frequency band's power spectrum and getting m_j through the formula which is as follows:

$$m_j = \log_{10}\left[\sum_{k=0}^{(k/2)-1} \varphi_j(k)P(k)\right];\ 1 \le j \le Q \tag{84.5}$$

where Q denotes the number of filter and $\varphi_j(k)$ denotes the jth filter's weighted function.

m_j is done by DFT to get MFCC parameter.

$$\mathrm{c}_i = \sqrt{\frac{2}{Q}}\sum_{j=1}^{Q} m_j \cos\left(\frac{\pi i}{Q}(j - 0.5)\right) \tag{84.6}$$

where i denotes exponent number of Cepstrum coefficients ($i = 12$) and j denotes the number of filter ($j = 1{-}Q$, $Q = 24$). In order to reflect the dynamic characteristics and the energy effect of voice, MFCC's first order difference, first-order energy, and first-order difference energy are joined voice of characteristic vector which is mentioned above.

84.2.3 CHMM'S Train and Recognition Algorithm

84.2.3.1 CHMM Model

CHMM model is a statistical model whose model parameters are $\lambda(\pi, \mathrm{A}, \mathrm{B})$, where, π is the initial state distribution, A is the state-transition probability distribution, and B is the observation symbol probability distribution. This paper uses CHMM because of higher recognition ratio than DHMM. Its type is shown in Fig. 84.2.

Model parameters are stated as follows [3]:

1. N, the number of states in the model. The model labeled the individual states as $\{\mathrm{S}_1, \mathrm{S}_2, \ldots, \mathrm{S}_7\}$ and denote the state at time t as qt. Where $N = 7$.
2. The state-transition probability distribution $\mathrm{A} = \{a_{ij}\}$, where

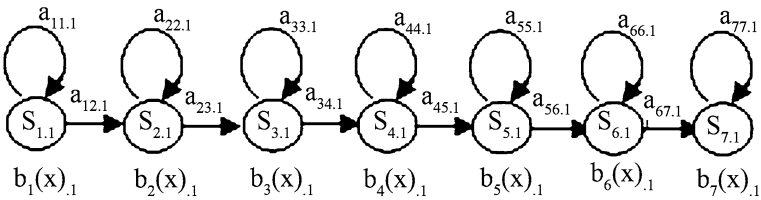


Fig. 84.2 A 7-state *left–right* continuous hidden markov model

$$a_{ij} = P\left[\mathrm{q}_{t+1} = \mathrm{S}_j / q_t = \mathrm{S}_i\right],\ 1 \le i, j \le N$$

From Fig. 84.2, $a_{ij} = 0, j < i, j > i + 1$.
The observation symbol probability distribution, $\mathrm{B} = \{b_j(k)\}$, in which

$$b_j(\mathrm{X}) = \sum_{k=1}^{M} c_{jk} N\left(X,\ \mu_{jk},\ U_{jk}\right),\ 1 \le j \le N$$

defines the symbol distribution in state j, $j = 1,2,\ldots,N$, where X is the observation vector being modeled, c_{jk} is the mixture coefficient for the kth mixture in state j and N is any log-concave or elliptically symmetric density [4]. M is Gaussian with mean vector μ_{jk} and covariance matrix U_{jk} for the kth mixture component in state j, where $M = 5$.

84.2.3.2 CHMM's Algorithm

The task of training includes two stages. First, using manual method mandarin digit string is best divided into isolated numbers. Then, the isolated numbers mentioned above from 0 to 9 will be come together as the training data. At last, parameter $\hat{\lambda}(\pi,\ \mathrm{A},\ \mathrm{B})$ is got through Baum–Welch algorithm, where, in order to avoid training process' underflow [5], the system uses multiple observation sequences reestimation formulas to calculate $\hat{\lambda}_i (i = 0, 1, \cdots 9)$.

In order to reduce recognition time, modified Viterbi algorithm is used which is described as follows [5]:

1. Initialization

$$\delta_i^v(0) = \begin{cases} \log a^v, & i = 1, v = 1 - \mathrm{V} \\ -\infty, & i \ne 1 \end{cases} \tag{84.7}$$

2. For $n = 1, 2, \cdots N$ in accordance with the recurrence formula to calculate which are shown as follows:

$$\delta_i^v(n) = \max\left\{\lambda_i^v(n),\ \eta_i^v(n)\right\} \tag{84.8}$$

$$\lambda_i^v(n) = \max\left\{\delta_{L_u}^v(n-1) + \log \mathrm{A}_{uv} + \log \mathrm{A}_{li}^v + \log p_i^v[\mathrm{y}_n]\right\} \tag{84.9}$$

$$\eta_i^v(n) = \max\left\{\delta_j^v(n-1) + \log \mathrm{A}_{ji} + \log p_i^v[\mathrm{y}_n]\right\} \tag{84.10}$$

If $\lambda_1^v(n) > \eta_1^v(n)$, then

$$\varphi_i^v(n) = \arg\max_u\left\{\delta_{L_u}^v(n-1) + \log \mathrm{A}_{uv} + \log \mathrm{A}_{li}^v + \log p_i^v[\mathrm{y}_n]\right\} \tag{84.11}$$

$$\phi_i^v(n) = L_{\phi_i^v(n)} \tag{84.12}$$

If $\lambda_1^v(n) < \eta_1^v(n)$, then

$$\varphi_i^v(n) = v \tag{84.13}$$

$$\varphi_1^v(n) = \arg\max_j\left\{\delta_j^v(n-1) + \log \mathrm{A}_{ji}^v + \log P_i^v[y_n]\right\} \tag{84.14}$$

Backtracking
Supposing $v_N = \arg\max_v \delta_{L_N}^v(N)$ is the last digital code of the best sequence and $x_N = S_{i_N}^{v_N}$, $i_N = L_{v_N}$ is the last state of the best sequence, then according to time sequence to backtrack which is described as follows:

$$\begin{cases} v_{n-1} = \varphi_{i_n}^{v_n}(n) \\ i_{n-1} = \varphi_{i_n}^{v_n}(n) \quad n = N, N-1, \ldots 1 \end{cases} \tag{84.15}$$

$v_1 \sim v_N$ is the single best word sequence, where a^v is the original state's probability, A_{ij}^v is the state-transition probability, and $P_i^v[y](v = 1 - V, i, j - 1 - L_v)$ is different state's vector output probability. $\delta_i^v(n)$ is the vth digital code's probability of its ith state. Variance n denotes moment. $\varphi_i^v(n)$ is the best digital code's number at $n-1$ moment. $\phi_i^v(n)$ is the best digital code's state at $n-1$ moment. $\lambda_1^v(n)$ and A_{uv} denote transition probability of between two digital. $\eta_1^v(n)$ and A_{ij}^v denote state-transition probability of one digital.

Given parameters a^v, A_{uv}, A_{ij}^v, $P_i^v[y]$ and Characteristic vector sequence $Y = \{\mathrm{y}_1, \cdots, \mathrm{y}_N\}$, the single best state sequence $X = \{x_1, \cdots, x_N\}$ can be gutted which is described as formula $\mathrm{x}_N = \mathrm{S}_{i_N}^{v_N}$, $i_N = L_{v_N}$.

84.3 Experimental Results

There are 50 male talkers' voices in data base including ten digits from 0 to 9. Each male talker speaks ten mandarin digit strings and length of each mandarin digit string is different. Where, voices of 30 talkers are used to train, others are used to test.

Table 84.1 Recognition results of speaker-dependent system

Model parameter	CHMM (7 states, 5 gauss) (%)	CHMM (7 states, 5 gauss) (%)
Acc	91.729	93.985
Corr	92.481	94.737
(E)	8.271	6.015

Table 84.2 Recognition results of speaker-independent system

Model parameter	CHMM (7 states, 5 gauss) (%)	CHMM (7 states, 5 gauss) (%)
Acc	78.212	84.637
Corr	85.475	91.899
(E)	21.788	15.363

The spectral analysis uses MFCC coefficients with the following characteristics: Sampling rate-16 kHz; Analysis window size-320 samples (20 ms); Analysis window shift-160 samples (10 ms); MFCC dimension-12; Cepstrum dimension-12;

Delta Cepstrum dimension-12; Cepstral window-hamming windows. The observation vector is a 26-component vector consisting of 12 cepstral coefficients, 12 delta Cepstrum coefficients, cepstral log energy, and delta log energy.

The type of CHMM is left-to-right models which have seven states and each state is mixture Gaussian probability density function, where the number of per-state's function is 5.

In this paper, a system of mandarin digit speech recognition which based on basic CHMM is accomplished using MATLAB. Types of error recognition include replace error (S), insertion error (I), and deleting error (D). S means one digit recognized the other digit. It means inserting a new digit to mandarin digit string. D means one digit is deleted from mandarin digit string. So accurate ratio, correct ratio, and error ratio are defined as follows (84.16), (84.17), (84.18).

$$\mathrm{Acc} = \frac{N - (\mathrm{S} + \mathrm{I} + \mathrm{D})}{N} \times 100\% \tag{84.16}$$

$$\mathrm{Corr} = \frac{N - (\mathrm{S} + \mathrm{D})}{N} \times 100\% \tag{84.17}$$

$$\mathrm{E} = \frac{\mathrm{S} + \mathrm{I} + \mathrm{D}}{N} \times 100\% \tag{84.18}$$

The system's results are shown in Tables 84.1, 84.2, where, the third column of Tables 84.1 and 84.2 are shown recognition results without replacing error of "8" is recognized as "2". From Tables 84.1 and 84.2, each index of recognition results of speaker-dependent system is obviously higher than that of speaker-independent system.

84.4 Discussions

Through the above experiments, conclusions can be gutted. One is that "8" is recognized as "2" is mainly replacing error which is to 46%. The other is insertion error which is the most serious among replace error, insertion error, and deleting error. The main reason is that using manual method to divide mandarin digit string is inaccurate. So there are some questions that are required to be researched, such as how to find better method to accurately divide mandarin digit string is important. How to find an effective model which can increase accurate ratio and correct ratio is important. In order to improve recognition ratio, how to find better parameters to distinguish between "8" and "2" is important.

Acknowledgments This work was supported by Henan province university science and technology innovation talent support plan (2011HASTIT013), Henan province university Fund of Teaching Reformation Project (2010GGJS-056), and Research Foundation of Youth Program of Henan Polytechnic University (Q2010-66).

References

1. Husheng L, Jia L, Runsheng L (2000) High performance digit mandarin speech recognition. J Tsinghua Univ 40(1):32–34
2. Husheng L, Jia L, Runsheng L (2001) High performance mandarin digit string speech recognition. Chin J Electron 29(5):595–599
3. Rabiner L, Juang BH (1999) Fundamentals of speech recognition, vol 6. The MIT Press, Cambridge
4. Levinson SE, Rabiner LR, Sondhi MM (1983) An introduction to the application of the theory of probabilistic functions of a markov process to automatic speech recognition. AT&T, Bell Syst Tech J 62(4):1035–1074
5. Hangjun Y, Junsheng C (1995) Speech signal digital processing. Electronic Industry Press, Beijing
6. Viterbi AJ (1980) Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. IEEE Trans Inf Theory 13:260–269
7. Abiner LR (1985) Recognition of isolated digits using hidden Markov models with continuous mixture densities. AT&T Tech J 64:1211–1234
8. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286

Chapter 85
A Fast Machine Vision for Automated Packaging

Yang Liu and Thanh Vinh Vo

Abstract This chapter introduces a fast, simple, cheap but accurate and reliable vision system designed for the automated packaging robot application, which is to replace manual product packaging. The objective is to develop a vision system and algorithm for such application and achieve pattern tracking in nearly real-time with a normal entry-level computer. This chapter describes in detail the implementation of the proposed vision system including image processing and pattern recognition. The application uses only a simple gray scale image of product as sample pattern to track the product in the scene at real-time speed, which has come near to replace human eyes. The proposed machine vision has been implemented and equipped on a real ABB robot arm with other related hardware/software to achieve automated packaging. The performance has been evaluated as a success in both timing and precision. Its stability and reliability are also proved under laboratory testing environment.

Keywords Vision system · Pattern recognition · Autonomous robot

Y. Liu (✉) · T. V. Vo
Department of Information Technology, Vaasa University of Applied Sciences, Vaasa, Finland
e-mail: yang.liu@puv.fi

Y. Liu
School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, China

T. V. Vo
e-mail: vthanhvinh@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_85,

85.1 Introduction

Automated packaging by using robot is a promising replacement of human labors. The increasing usage of autonomous packaging robots in industries demands related technological advancement such as the vision system. Vision system is a system that can see for specific purpose, and in particular in this application, it is a machine that is able to see and perceive position, pose and type of products to package. Faster, simpler, cheaper but more accurate and reliable vision system is vital to boost the development of autonomous packaging robot and improve its performance. This paper presents in detail the implementation of a fast vision system designed for the autonomous packaging robot application in order to fulfill the new demands.

85.2 Camera Model and Calibration

Based on the theory of [1, 2], a camera model can be summarized into one simple mathematic representation [3, 4]:

$$s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{85.1}$$

where

$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ is the coordinate of a point in the object coordinate system.

$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ is the rotation matrix.

$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ is camera intrinsic matrix.

$\begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$ is the translation vector.

$\begin{bmatrix} x \\ y \end{bmatrix}$ is the coordinate of a point in image plane.

From Eq. (85.1), the camera model is derived.

$$\rightarrow \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} * s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}^{-1} * \left(\begin{bmatrix} f_z & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} * s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} - \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \right) = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (85.2)$$

$$\rightarrow \text{A*s} - \text{B} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (85.3)$$

$$\text{With A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}^{-1} * \begin{bmatrix} f_z & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$\text{B} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}^{-1} * \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} * s - \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\rightarrow s = \frac{z + b_3}{a_3}; X = a_1 * \left(\frac{z + b_3}{a_3} \right) - b_1; Y = a_2 * \left(\frac{z + b_3}{a_3} \right) - b_2 \quad (85.4)$$

The fundamental mathematical model for vision calculation is provided in Eq. (85.4). The equation represents mere theory: given a calibrated camera (the rotation matrix, intrinsic matrix and translation vector are known), a detected object (its x and y coordinates in image plane are known) and the distance Z from the camera to the object, the X and Y coordinates of the object in the object coordinate system are derived.

From this model, all necessary steps are built up:

- Calibrating camera based on [5, 6].
- Detecting object.
- Measuring height distance from camera to object.

After finishing the above steps, the application should be able to calculate X, Y coordinate of the objects using Eq. (85.4).

In this application, EO-0413 M 1/3″ CMOS Monochrome USB Camera and TAMRON CCTV Lens are used as camera.

After calibrating, the result is as follows:
Intrinsic matrix

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 939.8787 & 0 & 291.6239 \\ 0 & 949.0118 & 204.8668 \\ 0 & 0 & 1 \end{bmatrix}$$

Extrinsic matrix

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} = \begin{bmatrix} 0.0738630 & 0.966409 & -0.2461678 & -51.76666 \\ -0.99509 & 0.055149 & -0.0820756 & 186.6437 \\ -0.065742 & 0.251023 & 0.965745 & 157.20998 \end{bmatrix}$$

Distortion parameters

$$\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} -0.453744 \\ -0.3256628 \\ -0.0071649 \\ 0.0236721 \\ 4.8843275 \end{bmatrix}$$

Apply those parameters into Eq. (85.2)

$$\begin{bmatrix} 0.0738630 & 0.966409 & -0.2461678 \\ -0.99509 & 0.055149 & -0.0820756 \\ -0.065742 & 0.251023 & 0.965745 \end{bmatrix}^{-1} *$$

$$\left(\begin{bmatrix} 939.8787 & 0 & 291.6239 \\ 0 & 949.0118 & 204.8668 \\ 0 & 0 & 1 \end{bmatrix}^{-1} * \begin{bmatrix} s*x \\ s*y \\ s \end{bmatrix} - \begin{bmatrix} -51.76666 \\ 186.6437 \\ 157.20998 \end{bmatrix} \right) = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 69.41944 & 917.13448 & 219.27853 \\ -935.273 & 52.33800 & -278.97862 \\ -61.78763 & 238.22094 & 33.22007 \end{bmatrix} * \begin{bmatrix} s*x \\ s*y \\ s \end{bmatrix} - \begin{bmatrix} -199.88765 \\ 0.27157 \\ 0.27157 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{85.5}$$

Equation (85.5) represents the camera model at a specific pose for the vision system.

85.3 Pattern Recognition

The pattern recognition is based on a fast scale- and rotation-invariant interest point and descriptor. The method is built on top of [7]. The aim of this method is to find the corresponding features between two images. To be more specific, this application will track in the scene for predefined product's existence and position. The primary challenge is speed which has to be near real-time in order to compete with human speed in recognition. Speeded up robust features (SURF) [8] is employed in the implementation which is several times faster than scale-invariant feature transform (SIFT) [9] and more robust against different image transformations than SIFT.

The aim of the approach is to extract distinctive invariant features from images that can be utilized to reliably match between different views of objects. The extracted features are invariant in terms of rotation and scale. Therefore, an object can be identified with high robustness regardless of changes in 3D viewpoint, noise or illuminations with a rapid speed (near real time performance).

At first, a sample image of the product under .png (portable network graphics) format is taken. Computer vision applications have been using this format because it supports gray scale images and does not require a patent license. Gray scale digital image has only one channel that carries the intensity information. Hence, the computer will run many times faster to process a gray scale image than a colorful image with more than one channel. Then, interest points in the image are extracted. The method to extract interest points is based on Hessian detector to find a list of points which are invariant in terms of rotation and scale.

Having a list of interest points, they will be compared against those interest points extracted from camera's frames. By comparison of interest points, it can be decided if they represent the same object points. In order to do the comparison, the interest region is divided into four smaller square regions. Then construct the vector for each interest region

$$v = \left(\sum d_x; \sum d_y; \sum |d_x|; \sum |d_y|\right) \tag{85.6}$$

where

d_x is Haar wavelet transform in horizontal direction.
d_y is Haar wavelet transform in vertical direction.

At final step, this vector is normalized. The interest point comparison is now done by comparing this vector, as shown in Fig. 85.1. After comparison, a list of matched features is achieved. For better accuracy, it will be voted for size and orientation to eliminate the matched features whose scale and rotation are not in harmony with the majority of them. Finally, a projection matrix (homography matrix) can be computed from those matched points to project the sample image into the frame images by using RANdomSAmple Consensus (RANDSAC) [10].

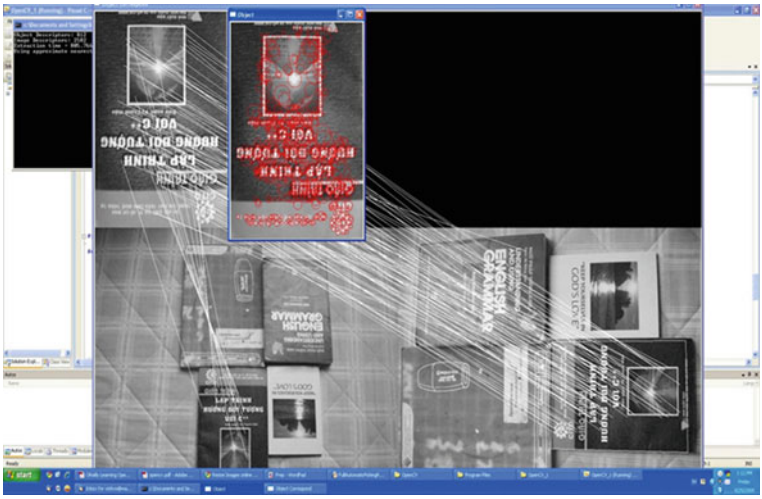


Fig. 85.1 Interest area extraction, comparison and matching

However, in case there are many products in the frame, the detection seems to reduce significantly the accuracy and efficiency. With one sample interest point, there are multiple matching points in the frame. RANDSAC, hence, fails to construct homography matrix. This redundancy is handled by dividing the frames into subregions. RANDSAC is applied for only matching points which belong to those subregions. In various experiments, with the particular size of boxes, the length and the width of the frame are both divided into two smaller sizes. This combination gives best performance in terms of speed and accuracy.

Obviously, the application does not intend to stop after tracking one product. Recursive algorithm is brought in use to track all or most of products inside the box. The list of matching points obtained from comparison step is filtered out, which means matching points from found products will be eliminated. Matching points are iterated and checked if they locate inside areas of founded product. Those that locate inside areas of founded product will be eliminated. After filtering matching points, the step of dividing subregions is repeated and RANDSAC again. As a result, a recursive loop is executed until no more products are found.

Last, this method, however, can sometimes give bogus detections. Hence, the validation method is added to verify the existence of product. There are two criteria used to validate the tracking result:

- The projected area must not be smaller than a certain minimum value.
- The projected are must be rectangular (because all images are rectangular).

After setting parameters for above two criteria, the detection is now trustful.

Fig. 85.2 Testing of the vision system applied on the autonomous packaging robot

Table 85.1 Precision testing results

x(mm)	y(mm)	X(mm)	Y(mm)	X (mm)	Y′(mm)	Deviation X(mm)	Deviation Y(mm)
25	30	284.91	29.0107	283.66	27.365	1.2575	1.64577
68	89	234.00	13.0976	232.68	14.12	1.3217	1.0224
155	143	189.23	98.4423	191.63	100.55	2.3977	2.10766
399	258	87.933	361.36	95.04	362.57	7.1064	1.21
620	467	139.58	651.509	136.44	647.22	3.1468	4.2894
					Average	3.04602	2.055046

x is the X-coordinate in the image(where product is tracked)
y is the Y-coordinate in the image (where product is tracked)
X is the X-coordinate of product in real world (achieved by using Eq. (85.4))
Y is the Y-coordinate of product in real world (achieved by using Eq. (85.4))
X′ is the X-coordinate of product in real world
Y′ is the Y-coordinate of product in real world
Deviation X = abs (X-X′)
Deviation Y = abs (Y-Y′)

85.4 Testing

The proposed fast vision system is applied on the autonomous packaging robot in laboratory environment as shown in Fig. 85.2 and explicitly tested with positive results. The real-time detection is demonstrated, in which the application is able to track multiple products at really fast speed. There is no significant lagging time observed during application operation. Usually, the vision system always has a queue of multiple products waiting robot to pick and place. As the result of fast tracking, robot will have constant tasks to perform. Therefore, there is no wasting

time in the system which means optimal speed can be achieved with faster robot movement. During the laboratory test, due to the constrained speed in robot for safety reasons, the robot has never been put to operate in maximum speed. The mostly used speed which is in manual operating mode is 250 mm/s. The pace of picking and placing one product is around 8–9 s/product at the manual operating mode. Moreover, some experiments to put robot to work at higher speed 600 mm/s are carried out. The result is positive with 4–5 s/product. This packing speed is really satisfactory if it can be applied it into real industrial environment.

The precision testing results are listed in Table 85.1.

The average error is around 2–3 mm, which is sufficient to make a draft calculation of X, Y and Z. Hence, the robot can navigate right above the product to take the precise distance measurement, pick up and drop down the product.

85.5 Conclusion

In this paper, the method of implementing a fast vision system for autonomous packaging robot is proposed and described in detail. It mainly consists of two important parts which are calibration for image detection and pattern recognition. Both of described methods are implemented and tested to produce good performance in terms of speed and accuracy.

Acknowledgments The author gratefully acknowledges the financial support of this research from China Hubei Provincial Science and Technology Department (Project No. 2010BFA007), and Vo Thanh Vinh for his valuable work in this project.

References

1. Bakstein H, Halir R (2000) Camera calibration with a simulated three dimensional calibration object. Czech Pattern Recognition Workshop 2000, Perslak, Czech Republic
2. Brown DC (1966) Decentering distortion of lenses. Photogramm Eng 32(3):444–462
3. Bradski G, Kaehler A (2008) Learning OpenCV: computer vision with the OpenCV library. O'Reilly Media, Sebastopol
4. OpenCV 2.0 C Reference. http://opencv.willowgarage.com/documentation
5. Zhang Z (2000) A flexible new technique for camera calibration. IEEE Trans Pattern Anal Mach Intell 22(11):1330–1334
6. Brown DC (1971) Close-range camera calibration. Photogramm Eng 37(8):855–866
7. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comp Vis 60(2):91–110
8. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) SURF: speeded up robust features. Comp Vis Image Underst 110(3):346–359
9. Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision, vol 2, p 1150–1157
10. Fischler MA, Bolles RC (1981) Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24:381–395

Part VII
Image Processing

Chapter 86
Activity Analysis in Video Scenes Using an Unsupervised Learning Approach

Peini Zhang and Zhichun Mu

Abstract In this paper, we introduce an unsupervised learning framework for automatically detecting motion patterns in the real video scenes containing multiple groups of heterogeneous objects. First we quantize the vectors of optical flow as the visual features from video scenes, and then these visual features are clustered using hierarchical dirichlet process to discover the key motion patterns in video scenes. We test our approach on two datasets, and experimental results show that this framework can effectively detect the motion patterns in complex scenes.

Keywords Activity analysis · Motion patterns · Unsupervised learning approach · Probability topic models

86.1 Introduction

Surveillance systems are widely applied in various facilities including traffic intersections, train stations, and other public facilities. Recently, with the installation of a large number of cameras, the video data of surveillance system is increased but not sufficiently exploited. In practice, most of the systems still need human expertise to determine whether there is abnormal situation in the video scene. In this paper, to improve the performance of surveillance we provide a framework which can automatically detect the key motion patterns in the video of surveillance using unsupervised clustering.

P. Zhang (✉) · Z. Mu
School of Automation, University of Science and Technology Beijing,
Beijing 100083, China
e-mail: zpeini@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_86,

86.1.1 Related Work

In a real video scene, the activities of multiple objects are affected by some factors including the scene structure (e.g. vehicles moving along the road avoiding obstacles, intersections), the scene status (e.g. traffic order is dictated by the lights), and the motion patterns (e.g. multiple objects move according to different directions), etc. Using the above information of scene structure and motion patterns, we can obtain the semantic regions for activity detection. With these key motion patterns in the regions, high level activity analysis and abnormal activity detection can be accomplished.

Various researches on activity analysis fall into two categories, that is, tracking-based approaches and non-tracking- based approaches. The tracking-based approaches [1–4] commonly use the trajectories of objects to model the scene. In [5], the motion patterns of trajectories are learned by the hierarchical self-organizing neural network, which is accompanied by high computational time cost. The spatio-temporal motion patterns are learned by using kernel density estimation [1]. The other related approaches [2, 3, 6] use multiple features including size, velocity, and location to clustering tracks. The reliance of all the above methods is based on the ability to tracking objects. However, in crowded scenes, tracking failures and errors often happened which lead to trajectory broken. Therefore, the tracking-based approaches are not fit to the crowded and complex video scenes.

Recent researches on activity analysis propose the non-tracking-based methods. The motions of multiple objects are modeled using a Dynamic Probabilistic Networks (DPNs) in [7]. Wang et al. [8] use text document analysis to clustering activities in complex scene. Yang et al. [9] use diffusion maps to detect motion patterns at different scales. Saleemi et al. [10] use GMM to discover motion patterns in static scene. Our proposed approach aims to overcome the drawbacks of object-based tracking. We propose an unsupervised learning framework to automatically detect the key motion patterns in the video scene. Instead of trajectories, we only need to obtain the low level motion features by computing optical flow between consecutive frames. And the video clips are represented by quantized visual bags of words which are optical flow direction histograms. The probability topic model, which captures the co-occurrence words, is used to learn from quantized visual data. We can automatically detect the key motion patterns in the video of surveillance.

Probability topic models have been applied to information classification and retrieval, object recognition [11–14], etc. Topic models mainly include probability latent semantic analysis, Latent Dirichlet Allocation, Hierarchical Dirichlet Process, etc. These models can capture word correlations in a set of textual documents with the set of multinomial distribution. The words co-occurring in same documents are clustered into the same topic. Therefore, our goal is to automatically detect the key motion patterns in the scene using the characteristics of this probability topic model.

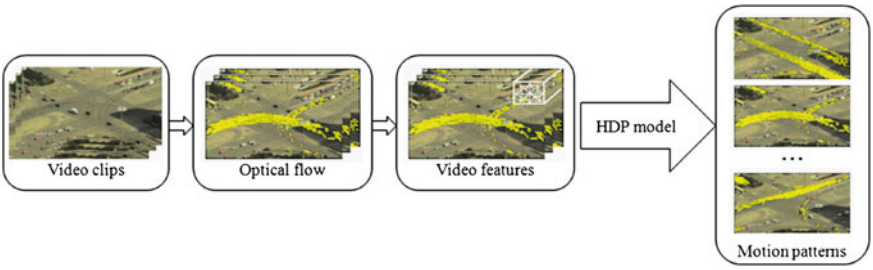


Fig. 86.1 The flowchart of our approach

86.2 Our Proposed Framework

In this section, we introduce our method and describe the process of discovering the key motion patterns in detail. Our method is outlined in Fig. 86.1.

86.2.1 Visual Features Quantization

Video scene contains complex activities of multiple objects. Local motions are used as low level features. Given a video, we first segment the video into clips, and divide a frame into N×N size cells. We compute optical flow between consecutive frames in each cell. Due to slight camera and variations in illumination, we set the threshold Th0. The optical flow is reliable when the magnitude of flow is greater than the threshold Th0. We quantize the motion of a moving pixel into one of eight directions according to the motion directions. Each moving pixel is represented by the visual word based on the location and the motion direction. Each clip is represented by the set of quantized words. If the size of each frame is $(N \times m) \times (N \times n)$, the size of codebook will be $m \times n \times 8$ ($m \times n$ is the size of cells, eight represents the directions of motion of pixels). A motion pattern is a distribution over quantized words. We model motion patterns as the quantized words co-occurring.

86.2.2 Hierarchical Dirichlet Process

Our main task is to divide the video data into a set of groups and capture the latent structure among these groups. To this end, we first segment the video into clips, and then quantize the local motion into quantized words (visual features) for each video clip, after finding the set of visual words with same characteristics in each clip, we can obtain a group of topics with the same meaning, at last, by clustering into the topics with same characteristics in clips, we may find the topics (if present)

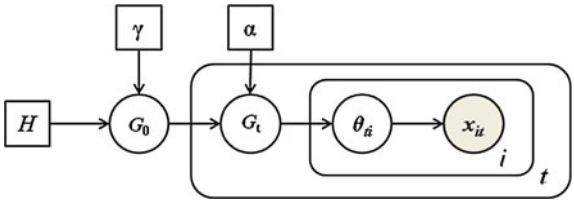


Fig. 86.2 HDP model

of the whole video, that is, we wish to find the sharing of clusters across clips. In this way, we will obtain the number of key motion patterns (topics).

The probability topic models such as LDA are generative models including words, topics and documents. In these models, words that co-occur in the same documents are clustered into the same topic. The topics can reflect the correlations among the words. However, LDA cannot automatically decide the number of topics. HDP is a non-parametric hierarchical Bayesian model and can automatically decide the number of topics [11]. This method models the data as groups and each group is characterized by a mixture model, in which a set of DPs is coupled with their own base measure and each data group is distributed according to a DP. The HDP model is shown in Fig. 86.2.

The HDP model is a distribution based on random probability measure. A global random measure G_0 is distributed as a Dirichlet Process with concentration parameter γ and based on probability measure H: $G_0 \sim (\gamma, H)$, where G_0 is a prior distribution over the whole corpus. The distribution can be expressed using the stick-breaking representation:

$$\begin{aligned} G_0 &= \sum_{k=1}^{\infty} \pi_k \delta_{\theta_k} \\ \theta_k &\sim H \\ \&pi_k^{'} &\sim \text{Beta}(1, \gamma) \\ \pi_k &= \pi_k^{'} \prod_{l=1}^{k-1} 1 - \pi_l^{'} \end{aligned} \tag{86.1}$$

A Dirichlet process generates a distribution G_0 with mixture of θ_k drawn from H, where the θ_k are parameters of multinomial distributions. H is defined as a distribution over multinomial distributions using a Dirichlet distribution:

$$\begin{aligned} \text{H} &= \text{Dir}(D_0) \\ \theta_k &\sim \text{Dir}(D_0) \end{aligned} \tag{86.2}$$

G_0 is a prior distribution over the whole corpus. For each document t, G_t is also distributed as a Dirichlet Process with concentration parameter α and base probability measure G_0: $G_t \sim (\alpha, G_0)$. G_t is a subset of G_0, and also used by the stick-breaking representation:

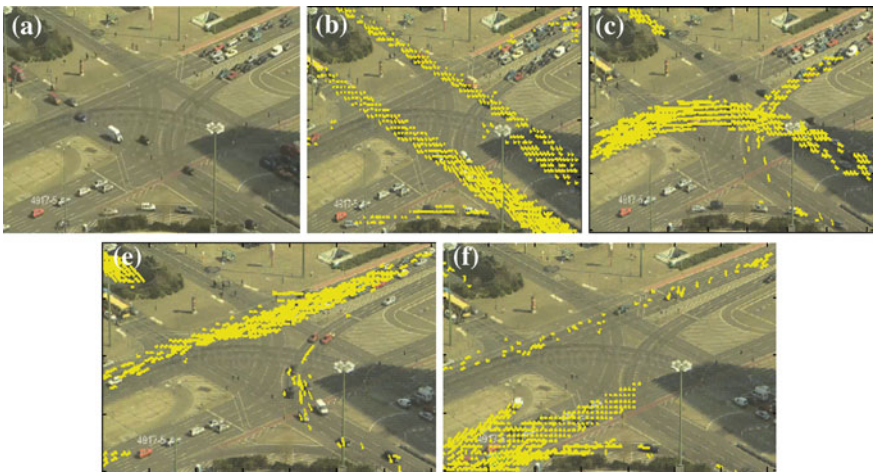


Fig. 86.3 Experimental results of the traffic intersection dataset. **a** Traffic intersection. **b–e** The motion patterns learned in the video clips

$$
\begin{aligned}
&\mathrm{G}_t = \sum_{k=1}^{\infty} \widehat{\pi}_k \delta_{\widehat{\theta}_k} \\
&\tau_k \sim \mathrm{Mult}(\pi_1, \pi_2, \ldots) \\
&\widehat{\theta}_k = \theta_{\tau_k} \\
&\pi_k^{'} \sim Beta(1, \alpha) \\
&\widehat{\pi}_k = \widehat{\pi}_k^{'} \prod_{l=1}^{k-1} 1 - \widehat{\pi}_l^{'}
\end{aligned}
\tag{86.3}
$$

We model activities as the quantized words co-occurring in the clip. G_0 is the global set of activities that shared by all clips. G_t is the subset of activities in each clip t. A motion flow pattern is multinomial distributions over the quantized words and activities are represented by motion flow patterns θ. In general, hypeparameters (γ and α) are priors on the distribution of activities, and D_0 is the parameters of the Dirichlet distribution. These parameters (γ, α and D_0) are set by the user.

In summary, we quantize visual features into words, and then automatically learn co-occurring activities using HDP based on visual words and clips.

86.3 Experimental Results and Discussion

We validated the proposed framework using two video datasets which are shot by single fixed camera (see Figs. 86.3a and 86.4a). Traffic Intersection Dataset: This consists of 24 video clips (each clip is 3 s, 25 fps, frame size is 720 × 400).

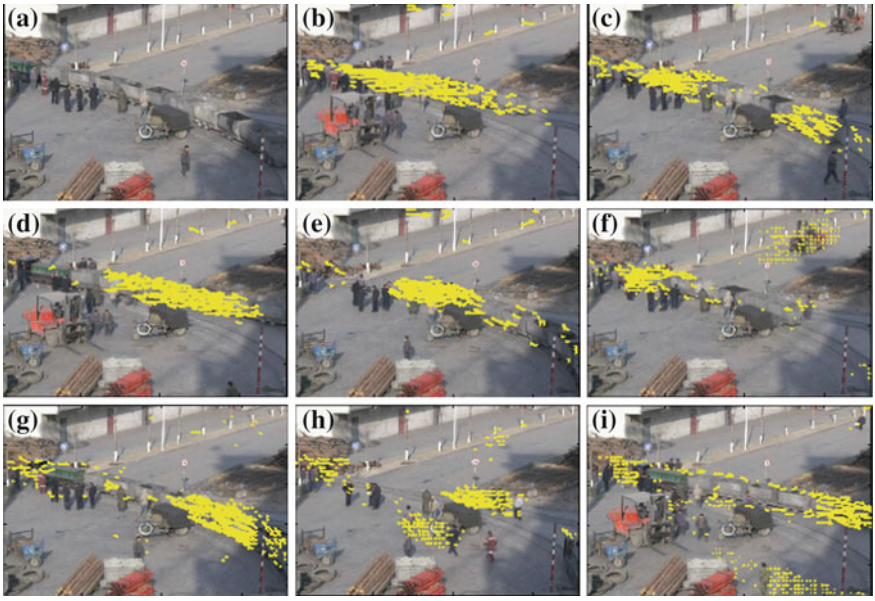


Fig. 86.4 Experimental results of the coal mine ground dataset. **a** Surveillance scene from mine surface plant of coal mine auxiliary shaft. **b–i** The motion patterns learned in the video clips

We compute optical flow between consecutive frames using [15], and quantize into eight directions. The cell size is 10×10 for the frame. Therefore, size of the visual features set (codebook) is 72 × 40 × 8. Coal Mine Ground Dataset: This consists of 110 video clips (each clip is 3 s, 29 fps, frame size is 720 × 480). The cell size is 20 × 20 for the frame. Therefore, size of the visual features set (codebook) is 36 × 24 × 8.

In the scene, the key motions are critical for activities analysis. We want to know there has the key motions in the scene. Our model can find the key motion patterns in the scene. In experimental video of traffic intersection, it contains four traffic flows in different directions which are regulated by traffic lights. In experimental video of coal mine ground, it contains two tram lanes, staff activities, etc. The scene is in a clutter of interactive activities due to random behaviors of staff. To create visual documents, a document should consist of the quantized words in the clip. Our model can automatically detect the key motion patterns in the scene (see Figs. 86.3, 86.4).

In the first experiment, we obtain fourmotion patterns (topics) using hyperparameters ($\alpha = 0.7$, $\gamma = 0.3$, $D_0 = 30$). The video describes that the different traffic flows are regulated by the traffic lights. The motion patterns are shown in Figs. 86.3b–e, (b) traffic in the east–west direction, and at the same time vehicles moving right-turn from south to east, (c) vehicles moving left-turn from east to south, and the same time vehicles stopping along north-east road, (d) vehicles moving along north–south road, and the same time vehicles moving left-turn from

north to east, (e) vehicles moving along north–south road, vehicles turning right from south to east, and some coming to the intersection and stopping along south-west road.

In the second experiment, we obtain eight motion patterns (topics) using hyperparameters ($\alpha = 2$, $\gamma = 1$, $D_0 = 60$). The topics are arranged from large to small according to the number of quantized words. The video describes that repetitive motions of tram, staff activities and forklift truck motions (e.g. tram moves along the lane from southeast to west until the tram completely moving out of the video screen, and the other tram moves along the lane from west to east until the tram completely moving out of the video screen after a period of time). These key motion patterns marked with yellow arrows discovered in the video are shown in Fig. 86.4b–i, where the key motion patterns shown are, (b) tram pulling out from west, (c) tram pulling out completely from southeast, (d) tram pulling out into east, (e) tram pulling out from southeast, (g) tram pulling out into west, (i) tram pulling out from west to east, (f) a forklift truck moving from northeast while tram moving and (h) staff activities.

Results of experiments show that our proposed method can clearly detect multiple meaningful motion patterns in the video containing numerous activities.

86.4 Conclusion

In this paper, we discuss the problem of activity analysis in real scenes and develop an unsupervised approach for motion pattern discovery based on non-tracking. After extracting the optical flow from frames and quantizing visual features into words, the system can automatically learn the co-occurring activities using hierarchical dirichlet process based on visual words and clips. We have tested our approach on two video datasets and obtained meaningful results, that is, it can provide the effective protection for modern management in surveillance system.

Acknowledgments This research was supported by the National Natural Science Foundation of China (60973064), and the Natural Science Foundation of Beijing, China (4102039).

References

1. Saleemi I, Shafique K, Shah M (2009) Probabilistic modeling of scene dynamics for applications in visual surveillance. PAMI 31:1472–1485
2. Wang X, Tieu K, Grimson E (2006) Learning semantic scene models by trajectory analysis. In: ECCV, pp 110–123
3. Basharat A, Gritai A, Shah M (2008) Learning object motion patterns for anomaly detection and improved object detection. In: CVPR, pp 1–8
4. Stauffer C, Grimson WEL (2000) Learning patterns of activity using real-time tracking. PAMI 22:747–757

5. Hu W, Xie D, Tan T et al (2003) A hierarchical self-organizing approach for learning the patterns of motion trajectories. Chinese J Comput 26:417–426
6. Makris D, Ellis T (2005) Learning semantic scene models from observing activity in visual surveillance. IEEE Trans Man Cybernet Part B 35:397–408
7. Xiang T, Gong S (2006) Beyond tracking: modelling activity and understanding behavior. J Comput Vis 67:21–51
8. Wang X, Ma X, Grimson E (2007) Unsupervised activity perception by hierarchical Bayesian models. In: CVPR, pp 1–8
9. Yang Y, Liu J, Shah M (2009) Video scene understanding using multi-scale analysis. In: ICCV, pp 1669–1676
10. Saleemi I, Hartung L, Shah M (2010) Scene understanding by statistical modeling of motion patterns. In: CVPR, pp 2069–2076
11. Teh Y, Jordan M, Beal M, Blei D (2007) Hierarchical dirichlet processes. J Amer Statist Assoc 101:1566–1581
12. Blei D, Ng A, Jordan M (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
13. Wei X, Croft WB (2006) LDA-based document models for ad-hoc retrieval. In: SIGIR, pp 178–185
14. Li F-F, Perona P (2005) A Bayesian hierarchical model for learning natural scene categories. In: CVPR, pp 524–53
15. Zach C, Pock T, Bischof H (2007) A duality based approach for real time tv-11 optical flow. In: DAGM, pp 214–223

Chapter 87
Modeling and Simulation of Ground Clutter for Radar Seeker

Shuanghong Liu and Xiaoyu Ma

Abstract In the way of overlooking, the influence of ground and sea clutter to active radar seeker should never be neglected. It will seriously affect the seeker to intercept and track the target, so it is necessary to analyze the influence of ground and sea clutter to active seeker, so as to improve its capacity of anticlutter. This paper introduced a basic method of studying power spectrum, which described how pulse Doppler radar seeker received ground and sea clutter, as well as the real-time simulation model with high accuracy based on the technique of data abstraction. It was found on analysis that the algorithm applied to closed-loop and real-time simulation experiment of active radar seeker and whole missile.

Keywords Pulse Doppler radar seeker · Ground and sea clutter · Real-time simulation · Abstraction

87.1 Introduction

In modern war, the hedgehopping of aircraft and cruise missile is a typical and efficient attacking means. Therefore, it is pressing for ground-to-air missile weapon system to obtain the ability of attacking the target at a minimum altitude. In the process of attacking the target, the missile seeker has to suffer clutter that

S. Liu (✉)
Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450005, China
e-mail: cloudyni@zzia.edu.cn

X. Ma
Henan Institute of Engineering, Zhengzhou 451191, China
e-mail: hnzzmxy@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_87,

changes constantly. The relative position between clutter and target spectra is in a dynamic change with the transformation of intercepted geometrical relationship and movement of maneuvering target. When the target maneuvers or chases, its spectrum may penetrate into main-lobe clutter, which will seriously affect the seeker to intercept and track the target [1]. Therefore, the simulation of clutter circumstances is as important as that of the target when evaluating the missile weapon system's ability of attacking low-altitude target among multiple clutters. Thus the clutter jamming must be modeled and reproduced by radio frequency. Taking active Doppler radar seeker as the study object, this paper introduced a basic method of studying power spectrum based on the radar equation, which described how pulse Doppler radar seeker received ground and sea clutter; proposed a new way to model the power spectrum of real-time clutter by memorizing some characteristic data in advance, so as to achieve the high-accuracy and real-time simulation [2]. It was found on analysis that the method satisfied the requirements for the closed-loop and real-time simulation experiment of active radar seeker and whole missile.

87.2 Calculation Principle of Ground and Sea Clutter Power Spectrum

Suppose that the radar seeker studied hereby belonged to mono-pulse Doppler system with high pulse repetition frequency and single range gate, the error of reflection coefficient under reflecting background was not taken into consideration, the traditional Doppler uniform mesh partition was adopted, the interval between Doppler equal frequency curves was less than half of the radar's frequency resolution, and the interval between equidistant curves was equal to the travel of light velocity in half of a recurrence interval, so it could ensure the accuracy and avoid too much redundant information as well [3].

The mesh partition and coordinate system definition for clutter calculation were shown as Fig. 87.1, in which Origin O of coordinates stood for the center of mass O of the missile, XZ plane referred to the one consisting of isa, the vector in horizontal direction and the altitude line of the missile, the included angle between the direction of missile velocity and axis X was defined as the dive angle ξ, the flying height was h, the flying velocity was V, the azimuth angle and the pitch angle of mesh A in the coordinates were φ and θ, respectively, and the included angle between mesh A and the vector direction of the missile velocity was β.

It was known from Ref. [4] that in a certain state, the power spectrum density of a frequency f_n was equal to the sum of the power spectrum densities brought about by all meshes that its Doppler loop cut across, i.e.:

$$\sigma(f_n) = \frac{P_t \lambda^2}{(4\pi)^3 L \Delta f} \sum_m \sum_n [G^2(\theta_m, \varphi_n) + G^2(\theta_m, \pi - \varphi_n)] \cdot D(R_m, f_n) \qquad (87.1)$$

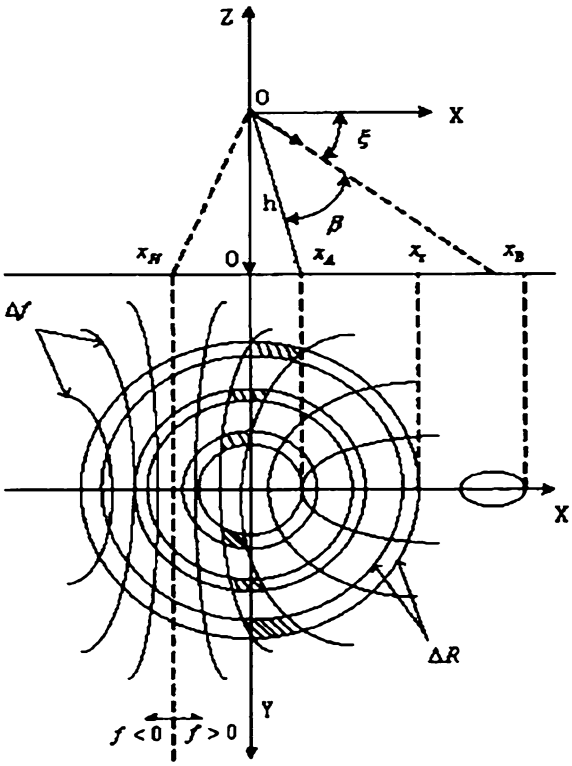


Fig. 87.1 Definition of Range-Doppler bins and coordinates

$$D(R_m,f_n) = \frac{1}{2}\int_A \frac{\sigma_0 dA}{R^4}; R_m \le R \le R_m + \Delta R. \tag{87.2}$$

In the formula, P_t was the average transmission power, λ was the wave length of radar, $G(\theta,\varphi)$ was the antenna gain, and L was the loss of the system; σ_0 was the backscattering coefficient of ground, $R_m = R_0 + m\Delta R, m = 0,\pm1,\pm2,\ldots$; $f_n = f_0 + nf_{PRF}, n = 0,\pm1,\pm2,\ldots; \Delta R = c/2f_{PRF}$; f_{PRF} was the pulse repetition frequency of radar, and c was light velocity. In order to calculate the integral expression $D(R_m,f_n)$ of the area of range-Doppler bins, $C(r,f_d)$ was defined as:

$$C(r,f_d) = \int_A \frac{\sigma_0 dA}{R^4}; \quad h \le R \le r; f \ge f_d. \tag{87.3}$$

Its physical meaning was the area intersected by inner Doppler curves and outer distant circles, then Formula (87.2) might be shown as:

$$\begin{aligned} D(R_m,f_n) = \frac{1}{2}[&c(R_m + \Delta R,f_n) - c(R_m,f_n) - c(R_m + \Delta R,f_n + \Delta f) \\ &+ c(R_m,f_n + \Delta f)]. \end{aligned} \tag{87.4}$$

In the formula, Δf was the bandwidth of frequency resolution.

It was obvious that the calculation of Formula (87.1) was transformed to the calculation of $C(r,f_d)$, i.e. different ranges r and Doppler frequency f_d, except for the correction of antenna pattern.

87.3 Integral Calculation of Clutter

In order to integrate Formula (87.2), the conversion relationship between rectangular plane coordinate system (X, Y, Z) and spherical coordinate system (R,θ,φ) might help:

$$\begin{cases} X = R\cos\theta\sin\varphi \\ Y = R\cos\theta\cos\varphi \\ Z = -h = -R\sin\theta \end{cases} \quad 0 \le \theta \le \pi/2, 0 \le \varphi \le 2\pi. \tag{87.5}$$

For the radar seeker, the flying height was very small compared to the radius of the Earth, so the ground could be regarded as a plane, then the integration of Eq. (87.5) could be represented by:

$$D(R_m,f_n) = \frac{\sigma_0}{2h}\int_A \sin\theta\cos\theta d\theta d\varphi. \tag{87.6}$$

For the clutter problem of radar seeker, the overlooking attack was the chief consideration, i.e. the flying velocity had a downward pitch angle ξ, then:

$$\beta = \arccos(f_n/f_{\max}); f_{\max} = 2v_d/\lambda, -\xi \le \beta \le \pi - \xi. \tag{87.7}$$

For the spherical coordinate system, the relationship was shown as below:

$$\cos\beta = \sin\xi\sin\theta + \cos\xi\cos\theta\cos\varphi. \tag{87.8}$$

According to the above equation, it was obtained:

$$\varphi = \arccos\left[\frac{\cos\beta - \sin\xi\sin\theta}{\cos\xi\cos\theta}\right]. \tag{87.9}$$

Substituted the result into Formula (87.2), the integration of $D(R_m,f_n)$ was transformed to one-dimensional integration of θ, which was given as $f(\theta)$; thus the closed solution could be obtained with the help of subsection integration.

$$L(\theta) = \frac{\sigma_0}{h^2}\Big\{\cos^2\theta\cdot\varphi - \cos\beta[(\sin\theta - \sin(\xi-\beta))(\sin(\xi+\beta) - \sin\theta)]^{1/2} - \sin\xi\sin^2\beta\cdot\varphi\Big\}. \tag{87.10}$$

in which, $\varphi = \arccos\left[\frac{\cos\beta - \sin\xi\sin\theta}{\cos\xi\cos\theta}\right]$.

Thereby, $C(r,f_d)$ was the result of $L(\theta_A) - L(\theta_B)$. For the whole range-Doppler space, there were other cases for the area A with the change of r and f_d; details referred to Ref. [5]. Here the general formula for calculation of $C(r,f_d)$ was given:

$$\begin{aligned} C(r,f_d) = {} & u(X_r - |X_A|)[L(\theta_{\max}) - L(\theta_{\min})] \\ & + u(-X_A)\{L(\theta_r) + u(X_r + X_A)L(\theta_A) - L(\theta_r)\} \end{aligned} \tag{87.11}$$

$$\text{In which, } \begin{array}{llll} X_r & = \sqrt{r^2 + h^2}; & \theta_r & = \arcsin(h/r) \\ X_A & = h\cot(\xi + \beta); & \theta_A & = \arctan(h/X_A) \\ X_B & = h\cot\theta_B; & \theta_B & = \max(\xi - \beta, 0) \end{array}$$

$$\theta_{\max} = \theta_A\,; \quad \theta_{\min} = \max(\theta_r, \theta_B)\,.$$

87.4 Real-time Simulation Modeling of Clutter

In order to solve the contradiction between real-time property and amount of calculation, the method of memorizing some characteristic data and then calculate power spectrum of real-time clutter was adopted. The way of modeling real-time clutter was elaborated as follows.

Formula (87.1) was analyzed. Suppose that the maximum flying speed of the missile was V and its corresponding Doppler frequency was $f_{\max}$, toovk the value that was less than or equal to half of the seeker's frequency resolution as the bandwidth Δf, the Doppler equal curves such as meshes were partitioned in the case of maximum Doppler frequency, then the curve of power spectrum density was obtained. With the same parameters, such as flight attitude and altitude, frame angle and reflecting background, except for the flying velocity, a corresponding curve of power spectrum density could be obtained by means of real-time abstraction of curves calculated at the maximum flying speed, which would not cause the loss of accuracy. The abstraction method was a mature technology, which was mentioned in many documents [6].

The influence of clutter to the seeker was studied in the case of overlooking attack, so only $\xi > 0$ was taken into consideration. In practical attacking process, the flying velocity direction of the missile could not be more than 45°, so only the case of $0 < \xi < 45°$was considered in this paper. The steps for modeling real-time clutter were as follows:

(1) According to the accuracy of angle measurement, memorized (87.12) at different pitch and drift angles of the antenna.

$$\frac{P_t\lambda^2}{(4\pi)^3 L\Delta f}G^2(\theta_m, \varphi_n). \tag{87.12}$$

(2) For certain missile and reflecting background, the value of the above formula would not change.
According to Formulas (87.4) and (87.6), calculated parameters φ and β corresponding to each mesh; and according to Formula (87.7), calculated (87.13) of the mesh.

$$D(R_m, f_n)\frac{\sigma_0}{h^2}. \tag{87.13}$$

Memorized the three parameters in advance.

(3) During the experiment, the parameters of each mesh were calculated in real time according to the input data, including missile attitude, maximum flying velocity of the missile, flight altitude, pitch, drift angles of the seeker's antenna, etc. With these parameters, the results of Formulas (87.12) and (87.13) were memorized previously and the backscattering coefficient σ_0 of the ground, the power spectrum was modeled in real time in the case of the maximum flying velocity.

(4) According to the maximum velocity, the power spectrum curve was obtained; and abstracted the power spectrum curve at corresponding velocity in real time. Considering the maximum probability, the quantity of meshes for equal Doppler-range curves was about 5×10^4, each mesh memorized 4 floating points, calculated the downward angle of visibility in 100 cases, and considered the symmetry of meshes to Axis X as well, then 30 MB zero-latency external storage space was required. During real-time calculation, it would cost the longest time to multiply/add the 5×10^4 floating points. With the help of digital signal processor, such as ADSP21060 chip, it would complete the multiply/addition operations of 4×10^7 floating points per minute, then it would only cost 1.25 ms to complete 5×10^4 operations. Considering other workload, such as the D-value operation between vectors of antenna and mesh directions, data abstraction, etc. the total calculation time would not exceed 50 ms, so it was very easy to achieve the purpose of real-time calculation and simulation in 100 ms.

Figure 87.2 showed the simulation result of ground clutter spectrum that the seeker confronted at a moment in attacking trajectory, from which we could see the power of main-lobe clutter was more than that of the target's echo signals. In attacking state, the frequency of the target's echo signal was out of the main-lobe clutter, so the influence of clutter to target detection might be neglected. However, in the case of chasing or target maneuvering, the target's spectrum would probably penetrate into the main-lobe clutter, so the seeker could not identify clutter and target from the frequency domain at this moment and the signal processor could not receive the target's observation data. If clutter was not identified in time and the target's echo frequency adopted the center frequency of Doppler filter that exceeded the detection limit in the group, it would probably result in the main-lobe clutter frequency was locked by velocity tracking loop, so the target would be lost and the tracking became a failure. This would seriously influence the track and acquisition of the seeker.

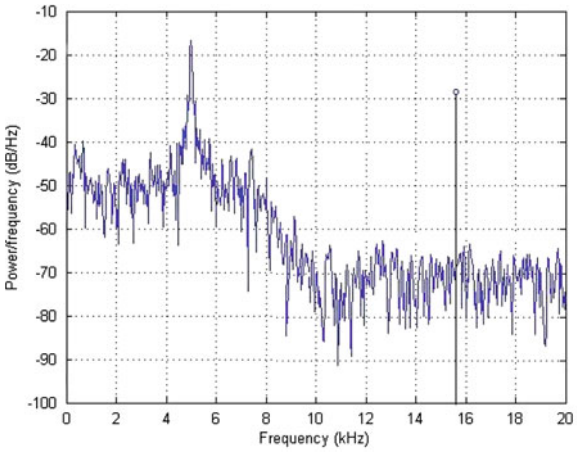


Fig. 87.2 Simulation result of ground clutter spectrum

87.5 Conclusion

It was a huge workload to calculate the integration of mesh areas by the traditional basic method, so the real-time calculation of power spectrum of ground and sea clutter could not be achieved. In that case, the power spectrum at current speed could be obtained by means of memorizing all mesh areas and their coordinate directions in advance and calculating the real-time power spectrum at the corresponding maximum flying velocity, then utilizing the method of data abstraction. It was possible to achieve real-time calculation for the method avoided the area integration, which was a huge workload. It could satisfy the requirements for the closed-loop and real-time simulation experiment of active radar seeker and whole missile.

References

1. Jao JK, William Gogginst BC (2001) efficient, closed-form computation of airborne pulse-doppler radar clutter [A]. IEEE Int Radar Conference, pp 17–22
2. Mao S (1990) Pulse Doppler radar (in Chinese) M. National Defence Industry Press, Beijing
3. Hong Mu (1996) Design of radar seeker for air-defence missile (in Chinese) M. Astronautic Publishing House, Beijing
4. William C (1994) Morchin airborne early warning radar M. National Defence Industry Press, Beijing
5. Li X (2006) Modeling and real-time simulation for clutter spectrum (in Chinese) . Comp Simul, (9):32–34
6. Orfanidis SJ (1996) Introduction to signal processing M. Prentice Hall, Englewood Cliffs, pp 644–727

Chapter 88
A Fast Fractal Image Compression Algorithm Using Improved Quadtree Partitioning Scheme

Hui Guo, Yunping Zheng and Jie He

Abstract By presenting an improved quadtree partitioning scheme and by modifying the composition method of the classical domain pool, we propose a fast fractal image compression algorithm. In our partitioning scheme, a square image is divided into four equal-sized square blocks, and then each block is tested to see if it meets some criterion of homogeneity. If a block meets the criterion, it is not divided any further. If it does not meet the criterion, it is subdivided again into four blocks, and the test criterion is applied to those blocks. This process is repeated iteratively until each block meets the criterion. Also, the proposed algorithm is validated and is compared with the classic fractal image compression algorithm. Experimental results show that the proposed algorithm can almost achieve the same reconstructed image quality as the classic algorithm, and can greatly reduce the encoding time and improve the compression ratio.

Keywords Fractal · Image compression · Quadtree · Human visual system

H. Guo (✉)
Department of Electronics Information Engineering, Wuzhou University,
Wuzhou 543002, China
e-mail: gxguohui@126.com

Y. Zheng
School of Computer Science and Engineering, South China University of Technology,
Guangzhou 510006, China
e-mail: zypdragon@163.com

J. He
Computer Science Department, Wuzhou University,
Wuzhou 543002, China
e-mail: hejie1213@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_88,

88.1 Introduction

Image compression is one of the most active research areas at present [1–3]. During the transmission of an image, the image compression technique can not only reduce the storage space, but can also shorten the transmission time. Currently, there are hundred kinds of image compression methods, but the compression quality, the compression ratio, the encoding and decoding time cannot meet the requirements of the current information age. Traditional compression algorithms generally have reached the compression limit and are not potential. However, the fractal image compression is a promising compression method [4–6]. When the compression ratios of 10000:1 is achieved under manual intervention conditions, decoding images still have good visual effect.

In this chapter, an improved quadtree partitioning scheme is presented after studying the advantages and disadvantages of some common partitioning schemes, such as the region segmentation [7], the quadtree segmentation [8] and the triangle segmentation [9]. In our partitioning scheme, a square image is divided into four equal-sized square blocks, and then each block is tested to see if it meets some criterion of homogeneity, i.e., the difference of the maximal and minimal pixel values among each block ranges from 10 to 99. In addition, we modify the composition method of the classical domain pool by taking a subsampling operation on the original image. Also, an improved quadtree-based fractal image compression algorithm is proposed and validated. Experimental results show that the proposed algorithm can almost achieve the same reconstructed image quality as the classic Jacquin's algorithm, and can greatly reduce the encoding time and improve the compression ratio. Therefore, the proposed algorithm in this chapter is an effective encoding method for fractal image compression.

88.2 Our Proposed Fractal Image Compression Algorithm

According to the improved partitioning scheme stated above, by redefining a new domain pool, we propose an improved quadtree-based fractal image compression algorithm in this section.

Step 1. For a given visual threshold Q, firstly divide the original image G with size $M \times M$ into four subimages each with size $M/2 \times M/2$, and then each subimage will be applied consistent standards to determine whether it satisfies.
Step 2. Set the maximum and the minimum sizes for range blocks.
Step 3. If a block meets the criterion, it is not divided any further. If it does not meet the criterion, it is subdivided again into four blocks, and the test criterion is applied to those blocks. This process is repeated iteratively until each block meets the criterion. Finally, the non-overlapping range blocks, which is denoted by a symbol R, with different sizes $2^i \times 2^i$ will be obtained.

Fig. 88.1 Original and restructured images using the classic algorithm and our proposed algorithm for 'Lena', 'Goldhill' and 'Barbara'. **a** Original. **b** Jacquin's. **c** Proposed. **d** Original. **e** Jacquin's. **f** Proposed. **g** Original. **h** Jacquin's. **i** Proposed

Step 4. Take a subsampling operation in the horizontal and vertical direction of the original image G, and we can get a subsampling image F with size $(M/2) \times (M/2)$.
Step 5. For each R, search a subblock $2^i \times 2^i$, i.e., a domain block (which is denoted by a symbol D), in the subsampling image F and make the square error between a new domain block D' and the range block R minimal, where D' s obtained by executing affine or space transformation in D.

Table 88.1 Total number of the range blocks for different thresholds

Image	*S*		
	$Q = 30$	$Q = 60$	$Q = 90$
Lena	3543	2671	2121
Goldhill	3785	3092	1869
Barbara	3923	3247	1977

Table 88.2 Performance comparison between jacquin's and our proposed algorithms

Image	*Parameters*	Jacquin's	Our proposed		
			$Q = 30$	$Q = 60$	$Q = 90$
Lena	*T*	196.58	62.45	49.09	35.82
	PSNR	31.28	31.03	30.72	29.87
	C	4.74	5.48	7.27	9.16
Goldhill	*T*	203.28	67.67	52.92	40.58
	PSNR	31.47	31.14	30.25	29.36
	C	4.74	5.13	6.28	10.39
Barbara	*T*	198.19	65.54	51.87	43.92
	PSNR	23.08	23.03	22.27	21.45
	C	4.74	4.95	5.98	9.82

Step 6. For each R, record five encoding parameters, i.e., the upper left corner coordinates (dx, dy), the affine number n, the gray contrast factor w and the gray shift factor g.

88.3 Experimental Results

In this section we implemented our proposed algorithm with the Matlab 7.0 software and made a comparison with the classic fractal image compression algorithm. In order that the proposed algorithm is comparable with the classic algorithm, the two algorithms both adopted our proposed method of constructing a domain pool. In the classic algorithm, the sizes of the range block and the domain block are 4×4 and 8×8, respectively. The horizontal or vertical movement step of the domain block is 4. Also, in our proposed algorithm, the minimum range block size is 4×4 and the maximum range block size is 8×8. According to the characteristics of HVS, the visual threshold Q is usually several orders of magnitude, i.e., Q ranges from 10 to 99.

By taking some standard grayscale images with size $256 \times 256 \times 8$, such as, Lena, Goldhill and Barbara for test objects (see Fig. 88.1a), the experimental results of our proposed algorithm are obtained in Table 88.1, where the S denotes the total number of the range blocks for different thresholds Q (30, 60 and 90). It can be seen from Table 88.1 that the bigger the threshold is, the less the total number of the range blocks.

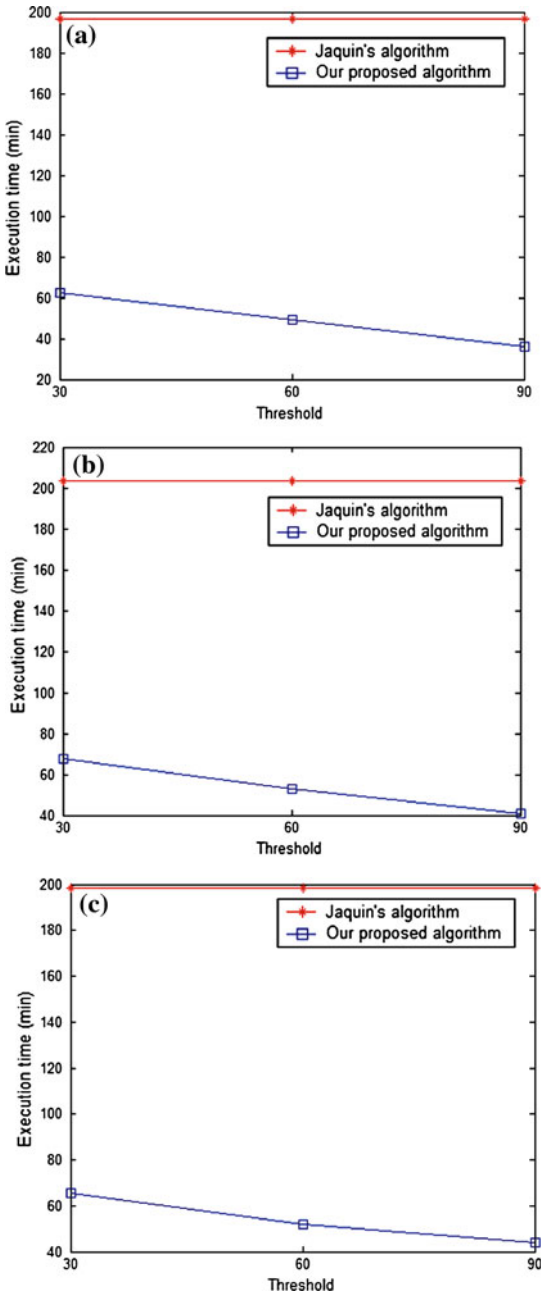


Fig. 88.2 Contrast of encoding time for the two algorithms. **a** Encoding time for 'Lena'. **b** Encoding time for 'Goldhill'. **c** Encoding time for 'Barbara'

Table 88.2 presents performance comparison of the two algorithms. It can be seen that selection of the value Q has some relations with the *PSNR* (take dB as unit), the code time T (take minute as unit), and the compression ratio C, where

PSNR in this table is specially defined as the signal-to-noise ratio of the eighth decoding. We can see from Table 88.1 that the smaller the valve Q is, the higher the decoding quality is, but the longer the code time. The compression ratio C is defined as $C = 256 \times 256 \times 8/(S \times (6 + 6 + 3 + 5 + 7))$. The top left-hand corner coordinate dx and dy of a domain block by the quantification are both six bits. The affine transformation matrix number n by the quantification is three bits. The gray contrast factor w and gray translation factor g by the quantification is five and seven bits, respectively.

Since the total number of the range blocks is fixed for a gray image when it is compressed by the Jacquin's algorithm, the compression ratio is also fixed according to the expression C. For example, in this experiment, since $S = 256/4 \times 256/4 = 4096$, we can deduce that $C = 256 \times 256 \times 8/(4096 \times (6 + 6 + 3 + 5 + 7)) = 4.74$ by using Jacquin's algorithm. However, the compression ratio ranges according to the threshold Q in our proposed algorithm. Different compression ratios can be obtained from Table 88.2 when Q equals 30, 60 and 90.

Figure 88.1 shows the original and restructured images using the classic algorithm and our proposed algorithm for 'Lena', 'Goldhill' and 'Barbara'.

Figure 88.1b, e and h are the three restructured images using the classic algorithm. Figure 88.1c, f and i are the three restructured images using our proposed algorithm under the condition $Q = 60$. It can be seen from Table 88.2 that the *PSNR*s of Fig. 88.1c, f, and i are 30.72, 30.25 and 22.27, respectively. The reason why the *PSNR* of Fig. 88.1 i is lower than the Fig. 88.1c and f is that Fig. 88.1i has a higher image complexity and its texture is relatively complex.

Figure 88.2 shows the comparison charts of the encoding time for Lena, Goldhill and Barbara by using the two algorithms.

In fact, the compression ratios of our proposed algorithm and Jacquin's classic fractal algorithm are obtained without using the entropy coding technology. If the entropy coding is used, the compression ratios could be further improved. From Fig. 88.2 and Table 88.2, it can be worked out that our proposed algorithm has not only higher compression ratio than the classic fractal algorithm, but also the encoding speed of ours is three to five times of Jacquin's whereas remaining a satisfactory decoding image quality. Thus, our proposed algorithm is an effective fractal image compression method.

88.4 Conclusions

In this chapter, by putting forward an improved quadtree partitioning scheme and by modifying the composition method of the classical domain pool, we propose a fast fractal image compression algorithm. The experimental results presented in this chapter show that the proposed algorithm can almost achieve the same reconstructed image quality as the classic Jacquin's algorithm, and simultaneously can greatly reduce the encoding time and improve the compression ratio.

Therefore, the proposed algorithm is an effective encoding method for fractal image compression.

Acknowledgments This work is supported by the Fundamental Research Funds for the Central Universities of China under Grant No. 2011ZM0074, the National Natural Science Foundation of China under Grant No. 60973085 and the Natural Science Foundation of Guangdong Province of China under Grant No. S2011040005815.

References

1. Tsai T, Lee Y, Lee Y (2010) Design and analysis of high-throughput lossless image compression engine using VLSI-oriented FELICS algorithm. IEEE Trans Very Large Scale Integr (VLSI) Syst 18:39–52
2. Park SY, Kim SD (2010) Efficient depth compression based on partial surface for 3-D object represented by layered depth image. IEEE Signal Process Lett 17:839–842
3. Lee C, Lee J (2009) Efficient compression for sampled color images. IEEE Trans Consumer Electron 55:2090–2096
4. Jacquin AE (1992) Image coding based on a fractal theory of iterated contractive image. IEEE Trans Image Process 1:18–28
5. Lai CM, Lam KM, Siu WC (2003) A fast fractal image coding based on kick-out and zero contrast conditions. IEEE Trans Image Process 12:1398–1403
6. Distasi R, Nappi M, Riccio D (2006) A range/domain approximation error-based approach for fractal image compression. IEEE Trans Image Process 15:89–97
7. Thomas L, Deravi F (1995) Region-based fractal image compression using heuristic search. IEEE Trans Image Process 4:823–838
8. Fischer Y (1995) Fractal image compression-theory and application. Springer, Berlin, pp 55–77
9. Davoine F, Antonini M, Chassery J (1996) Fractal image compression based on Delauney triangulation and vector quantization. IEEE Trans Image Process 5:338–346

Chapter 89
A New HVS-Based Fractal Image Compression Algorithm

Hui Guo, Yunping Zheng and Jie He

Abstract Inspired by the traditional quadtree fractal coding algorithm, we put forward a new partition principle of the quadtree according to the feature of HVS. Then we apply the feature of HVS to partition domain and range blocks. Also, we present a new HVS-based search method of finding out the best domain block for each range block in this chapter. Based on the new partition principle of the quadtree and the HVS-based search method, in this chapter, we proposed a new HVS-based fractal image compression algorithm. By comparing our HVS-based algorithm with the basic fractal image compression algorithm, the experimental results show that the encoding speed of our HVS-based algorithm is 8–27 times to that of the basic algorithm when the visual threshold is 30, 60 and 90, respectively, whereas maintaining a satisfactory reconstructed image quality. Therefore, our HVS-based algorithm is an effective image compression method.

Keywords Human visual system (HVS) · Fractal · Image compression · Quadtree · Search strategy

H. Guo (✉)
Department of Electronics Information Engineering, Wuzhou University, Wuzhou 543002, China
e-mail: gxguohui@126.com

Y. Zheng
School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China
e-mail: zypdragon@163.com

J. He
Computer Science Department, Wuzhou University, Wuzhou 543002, China
e-mail: hejie1213@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_89,

89.1 Introduction

Image compression is minimizing the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level, and it has been widely applied in many fields [1–3]. There are many methods for image compression for the time being. Based on the B-tree triangular coding (BTTC) approach, Distasi et al. [4] proposed the first spatial data structures (SDS) for representing the gray image. One advantage of BTTC over JPEG [5] is its shorter execution time. Later, based on the S-tree data structure and the Gouraud shading method, a new S-tree Coding (STC)-based approach for representing gray images was proposed [6]. Later, a novel hybrid gray image representation using spatial- and DCT-based approach (SDCT) was presented by Chung [7]. A novel approach pioneered by Michael Barnsley is to use the similarities on different scales throughout images to assist in compression. Fractal image compression enables an incredible amount of data to be stored in highly compressed data files [8]. The method is best suited for textures and natural images, relying on the fact that parts of an image often resemble other parts of the same image.

Some images perhaps contain pixels whose individual colors will not be perceived if they are below the spatial threshold of human visual system (HVS). This is a very important feature of HVS. If we can apply this feature to the fractal image compression, the compression will be raised. In this chapter, inspired by the traditional quadtree fractal coding algorithm, we put forward a new partition principle of the quadtree according to the above feature of HVS. In our partitioning scheme, a square image is divided into four equal-sized square blocks, and then each block is tested to see if it meets some criterion of homogeneity. We then apply the feature of HVS to partition of domain and range blocks. In addition, we also propose a new HVS-based search method of finding out the best domain block for each range block in this section. Experimental results show that the HVS-based algorithm can almost achieve the same reconstructed image quality as the classic Jacquin's algorithm, and can greatly reduce the encoding time and improve the compression ratio.

89.2 HVS-Based Search Method

The HVS-based search method that each range block uses when it finds out the best-matched domain block is that the search is carried out only among the domain blocks with the same size as the range block according to the feature of HVS. Therefore, in our proposed HVS-based algorithm in the following section there are no horizontal or vertical movement steps of the domain block. That is to say, our HVS-based algorithm need not choose a moving window. Therefore, the search room of our algorithm is much less than the classic algorithm, which is a very significant factor why our proposed algorithm is much faster than the classic algorithm.

89.3 Our HVS-Based Algorithm

Inspired by the traditional quadtree fractal coding algorithm, we put forward a new partition principle of the quadtree according to the feature of HVS. We then apply the feature of HVS to partition domain and range blocks. In addition, we also propose the new HVS-based search method of finding out the best domain block for each range block in this section.

Step 1. Given a visual threshold V, firstly divide the original image G with size $N \times N$ into four subimages each with size $N/2 \times N/2$, and then each subimage will be applied consistent standards to determine whether it satisfies.

Step 2. Set the maximum and the minimum sizes for range blocks.

Step 3. If a block meets the criterion according to V, it is not divided any further. If it does not meet the criterion, it is subdivided again into four blocks, and the test criterion is applied to those blocks. This process is repeated iteratively until each block meets the criterion.

Step 4. The non-overlapping range blocks, which is denoted by a symbol R, with different sizes $2^i \times 2^i$ will be obtained.

Step 5. Take a subsampling operation in the horizontal and vertical direction of the original image G, we can get a subsampling image F with size $(N/2) \times (N/2)$.

Step 6. Divide the subsampling image F into four subimages each with size $N/4 \times N/4$, and if each block cannot meet the criterion, execute Step 3.

Step 7. The non-overlapping domain blocks, which is denoted by a symbol D, with different sizes $2^i \times 2^i$ will be obtained.

Step 8. For each R, search a D with the same size as R in the subsampling image F and make the square error between the domain block D and the range block R is minimal. Then, record five encoding parameters, i.e., the upper left corner coordinates (dx, dy), the affine number n, the gray contrast factor w, and the gray shift factor g.

89.4 Experimental Results

By taking some standard grayscale images with size $256 \times 256 \times 8$ as test objects (see Fig. 89.1a), the experimental results of our HVS-based and the classic algorithms are obtained in Table 89.1, where the P (take dB as unit) denotes the signal-to-noise ratio of the eighth decoding, T (take minute as unit) denotes the encoding time, and C denotes the compression ratio. In the classic algorithm, the sizes of the range block and the domain block are 4×4 and 8×8, respectively. Also, in our HVS-based algorithm, the minimum range block size is 4×4 and the maximum range block size is 8×8. According to the characteristics of HVS, the visual threshold V usually ranges from 10 to 99.

The compression ratio C is defined as $C = 256 \times 256 \times 8/(S \times (6 + 6 + 3 + 5 + 7))$. The top left-hand corner coordinate dx and dy of a domain block by the quantification are both 6 bits. The affine transformation matrix number n by the quantification is 3 bits. The gray contrast factor w and gray translation factor g by

Fig. 89.1 Original and restructured images using the two algorithms. **a** Original images of 'Lena', 'Goldhill', and 'Peppers'. **b** Restructured images using the classic algorithm for 'Lena', 'Goldhill', and 'Peppers'. **c** Restructured images using HVS-based algorithm for 'Lena', 'Goldhill', and 'Peppers'

Table 89.1 Performance comparison between classic and HVS-based algorithms

Images	Parameters	Classic algorithm	HVS-based algorithm		
			$V = 30$	$V = 60$	$V = 90$
Lena	*T*	248.29	29.98	15.57	9.08
	P	32.98	32.15	31.76	30.83
	C	4.74	5.48	7.27	9.16
Goldhill	*T*	203.48	24.78	13.25	8.87
	P	32.85	31.78	30.23	30.02
	C	4.74	5.13	6.28	10.39
Barbara	*T*	225.86	28.05	14.32	9.12
	P	32.76	31.96	31.15	29.87
	C	4.74	5.76	7.25	9.18

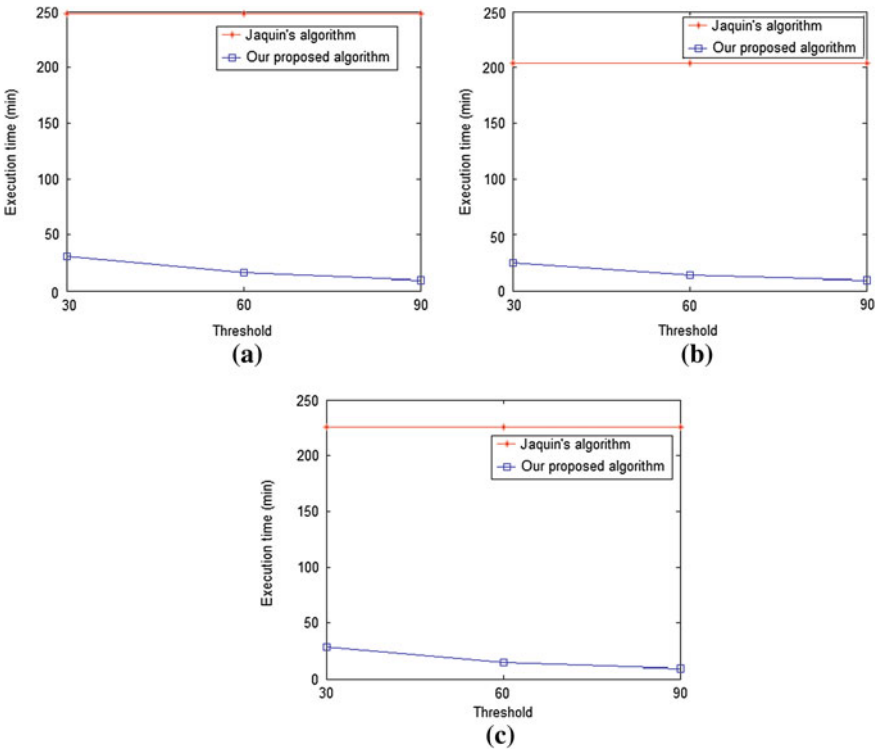


Fig. 89.2 Contrast of encoding time for the two algorithms. **a** Encoding time for 'Lena'. **b** Encoding time for 'Goldhill'. **c** Encoding time for 'Peppers'

the quantification is 5 and 7 bits, respectively. Since the total number of the range blocks is fixed for a gray image when it is compressed by the Jacquin's algorithm, the compression ratio is also fixed according to the expression C. For example, in this experiment, since $S = 256/4 \times 256/4 = 4096$, we can work out that $C = 4.74$ by using Jacquin's algorithm. However, the compression ratio ranges according to the threshold V in our HVS-based algorithm. Different compression ratios can be obtained from Table 89.1 when V equals 30, 60, and 90. Figure 89.1a–c presents three original and restructured images using the classic algorithm and the HVS-based algorithm, respectively. Figure 89.2 shows the comparison charts of the encoding time for Lena, Goldhill, and Peppers by using the two algorithms.

Figure 89.1a lists three original images for 'Lena', 'Goldhill', and 'Peppers'. Figure 89.1b shows three restructured images using the classic algorithm. Figure 89.1c shows three restructured images using our HVS-based algorithm under the condition $V = 60$. From Table 89.1, we know that the PSNRs of Fig. 89.1a are 31.76, 30.23, and 31.15, respectively. Under this situation, as a

matter of a fact, we cannot make a difference between the classic and HVS-based algorithms only according to the human visual system. From Table 89.1, it can be worked out that the encoding speed of our HVS-based algorithm is about 8–27 times of Jacquin's whereas remaining a satisfactory image quality. In addition, in the case of the compression ratio, our HVS-based algorithm is superior to the classic fractal algorithm.

As stated above, the experimental results in this section shows that our HVS-based algorithm can significantly deduce the encoding time whereas remaining a satisfactory image quality, and therefore it is an effective fast fractal image compression method.

89.5 Conclusions

In this chapter, we put forward a new partition principle of the quadtree according to the feature of HVS. Then we apply the feature of HVS to partition of domain and range blocks. Also, we present a new HVS-based search method of finding out the best domain block for each range block. Based on the new partition principle of the quadtree and the HVS-based search method, we proposed a new HVS-based fractal image compression algorithm. By comparing our HVS-based algorithm with the basic fractal image compression algorithm, the experimental results show that the encoding speed of our HVS-based algorithm is 8–27 times to that of the basic algorithm when the visual threshold is 30, 60, and 90, respectively, whereas maintaining a satisfactory reconstructed image quality. Therefore, our HVS-based algorithm is an effective image compression method.

Acknowledgments This work is supported by the Fundamental Research Funds for the Central Universities of China under Grant No. 2011ZM0074, the National Natural Science Foundation of China under Grant No. 60973085, and the Natural Science Foundation of Guangdong Province of China under Grant No. S2011040005815.

References

1. Thomas L, Deravi F (1995) Region-based fractal image compression using heuristic search. IEEE Trans Image Process 4:823–838
2. Park SY, Kim SD (2010) Efficient depth compression based on partial surface for 3-D object represented by layered depth image. IEEE Signal Process Lett 17:839–842
3. Distasi R, Nappi M, Riccio D (2006) A range/domain approximation error-based approach for fractal Image compression. IEEE Trans Image Process 15:89–97
4. Distasi R, Nappi M, Vitulano S (1997) Image compression by B-tree triangular coding. IEEE Trans Commun 45:1095–1100
5. Wallace GK (1991) The JPEG still picture compression standard. Commun ACM 34:29–44
6. Chun KL, Wu JG (2000) Improved image compression using S-tree and shading approach. IEEE Trans Commun 48:748–751

7. Chung KL, Liu YW, Yan WM (2006) A hybrid gray image representation using spatial- and DCT-based approach with application to moment computation. J Vis Commun Image Represent 17:1209–1226
8. Jacquin AE (1992) Image coding based on a fractal theory of iterated contractive image. IEEE Trans Image Process 1:18–28

Chapter 90
Study the Key Technologies of the Handheld Camera Panorama Mode

Xiaoliang Zhang and Lei Shi

Abstract For the characteristics of using ordinary digital camera or mobile phone to capture images, study of the key technologies of image registration and image fusion is conducted. Extracting and matching the feature points with the Harris corner detection algorithm, and filtering the feature points with the quick RANSAC method is proposed. Then the LM algorithm to optimize the transformation matrix, and the Multi-resolution fusion algorithm for image fusion to make the final output images achieve a satisfactory visual effect are used. Finally, the algorithm through experiment and splicing results are verified for satisfaction.

Keywords Panorama · Image · Registration · Image fusion · Harris corner · Detection

90.1 Introduction

Panorama stitching technique is based on the virtual reality technology. It can be a three-dimensional effect that uses ordinary images. The general picture and three-dimensional modeling cannot be achieved through this. Panorama's filming and production involves less investment and is simple to produce. A panoramic image

X. Zhang (✉)
Department of Computer and Information Engineering,
Henan Jiaozuo Teachers College, Jiaozuo 454000, Henan, China
e-mail: zxl970427@163.com

L. Shi
Department of Computer Engineering, Henan Polytechnic Institute,
Nanyang 473009, Henan, China
e-mail: shileiftp@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_90,

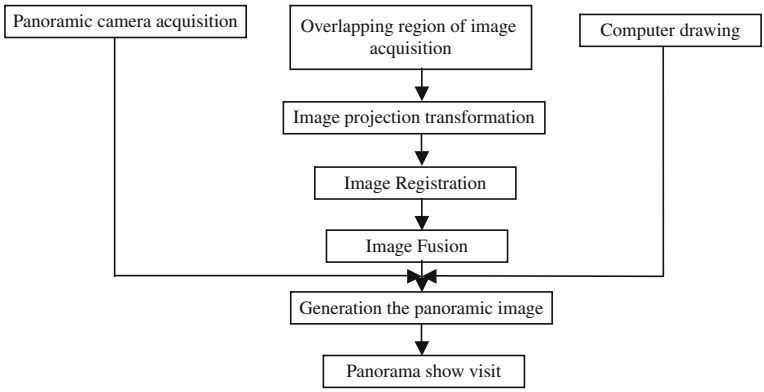


Fig. 90.1 Panorama generation process

from capture to release needs only a few minutes. it is a very small amount of data, the system requirements are low in a variety of ways, and can be watched on a variety of terminal devices.

90.2 Panorama Production Process

The basic idea of the panorama generation technology is through an ordinary camera or other handheld devices, sample the scene information with photo images. The panorama generation process summarized as follows Fig. 90.1.

90.3 Image Registration Technology

After projection transformation, the registration of all images in sequence to be projected onto a cylindrical coordinates, then these images can be spliced. In order to effectively complete the image splice, first, to conduct image registration, that is adjacent to accurately identify the two images overlap with the corresponding position, and then determine the transformation between two images, and finally splicing and integration. Currently, the typical image registration methods: a template matching method; based on image gray registration method; based on image feature registration method; phase correlation method and other registration methods [1, 2].

90.3.1 Image Registration Techniques Based on Corner Feature

According to computer vision theory, fixed-point shooting image with hand-held camera is used the projection transformation, for this transformation relationship

of the image registration more using feature point-based registration technique, as the hand-held camera is very flexible, so the selection of feature points is very important. In practice, the turning point of the image and profile of the end of the line are all corners. Corner features that with informative, able to adapt to changes in ambient light, easy to measure and expressed become the first features registration algorithm of choice.

The corner detection algorithms can be divided into three categories: based on edge detection, based on the template detection and based on the grayscale.

90.3.2 Harris Corner Detection Algorithms

In 1988, C. Harris and M. Stephens proposed a Harris corner detection algorithm. The algorithm is based on gray-scale image of the autocorrelation function. It is considered the neighborhood of pixel gray change, rather than the entire target edges. The formula for the corner detection following Eq. (90.1):

$$E(\Delta x, \Delta y) = \sum_{x,y} w(x,y)[I(x+\Delta x, y+\Delta y) - I(x,y)]^2 \tag{90.1}$$

where $w(x, y)$ is the window function, $[I(x + \Delta x, y + \Delta y) - I(x, y)]^2$, is the image gray value gradient. For each small amount of displacement $(\Delta x, \Delta y)$, by Eq. (90.1) can be approximated as Eq. (90.2):

$$E(\Delta x, \Delta y) \cong [\Delta x, \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \tag{90.2}$$

where $M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$.

Let λ_1, λ_2 are two eigenvalue of the matrix M, denotes the local curvature of the autocorrelation function, E can be approximated as a local cross-correlation function describing the shape of the point. The corner point response function can be written as Eq. (90.3):

$$R = \det(M) - k\mathrm{Trace}^2(M). \tag{90.3}$$

where $\det(M) = \lambda_1\lambda_2$, $\mathrm{Trace}(M) = \lambda_1 + \lambda_2$. When the determinant of the matrix M is large, it indicates that this is an edge or a corner; when a region of the main diagonal of the matrix M and large, it indicates that this is an edge. $k = 0.04 - 0.06$.

90.3.3 Feature Point Matching and Parameter Estimation

After a rough feature points match above, there are still a small amount of false feature points, if the least squares method to estimate the direct use of model

parameters, then this will bring relatively large registration error, so to use the RANSAC method to achieve precise matching and exclude the outer points.

To access points and the transformation matrix of the initial value, need to optimize the transformation matrix M through the set of interior points. Real parameter optimization is a nonlinear least squares estimation problem. In this paper, use the based on LM algorithm optimization, as follow Eq. (90.4):

$$\Delta x = -\left(J^T(x)J(x) + \mu I\right)^{-1} J^T(x)e(x). \tag{90.4}$$

Let the error function as follow Eq. (90.5):

$$E(x) = \frac{1}{2}\sum_{i=1}^{N} e_i^2(x). \tag{90.5}$$

where $ei(x)$ is the error, $E(x)$ is the error indicator function, $J(x)$ is a Jacobian matrix of $E(x)$, Eq. (90.4) $\mu > 0$ is a constant scale factor, I is the unit matrix. Since the LM algorithm uses the approximate second derivative information, it is much faster than the gradient descent method.

90.4 Image Fusion Technology

Used to splice the original image, due to shooting locations, the difference between light and vision, resulting in the overlapping area between images may vary greatly in the two images. If such images directly on the splicing, then the resulting splicing image in the overlapping area are blurring, distortions, and obvious seams phenomenon. So a smooth transition between images and between images into a seamless panorama splicing technique in another key technology, that is image fusion technology. Currently there are direct average fusion method, the weighted average fusion method, median filter fusion method, based on pyramid decomposition fusion method and multi-resolution spline technology fusion [3–5].

90.4.1 Image Fusion Based on Multi-Resolution

The multi-resolution fusion of specific process is as follows:

1. To build hierarchical pyramid structure of the source image, resulting in each image of the low-pass layer $G_0, G_1, \ldots, G_N$, as follow Eq. (90.6):

$$\mathrm{G}_l(x,y) = \sum \sum_{m,n=-2}^{2} w(m,n)G_{l-1}(2x+m, 2y+n) \tag{90.6}$$

where G_0 is the source image, $w(m,n) = w(m)w(n)$ is 5×5 window for the weighed function.

2. Use the images of the low-pass layer, decomposition of the image of the band-pass layer $L_0, L_1, \ldots, L_{N-1}$, as follow Eq. (90.7):

$$\mathrm{L}_l(x,y) = \mathrm{G}_l(x,y) - 4\sum_{m,n=-2}^{2} \mathrm{G}_l\left(\frac{2x+m}{2}, \frac{2y+n}{2}\right). \tag{90.7}$$

3. And then layer in each separate band-pass image fusion operation, the weighted average method is used here to achieve, is for the current layer of L_k, as follow Eq. (90.8):

$$L_{k\mathrm{out}}(x,y) = \frac{\sum_{i=0}^{i=N-1} \mathrm{L}_{kli}\left(x^{'}, y^{'}\right) w_i\left(x^{'}\right) w_i\left(y^{'}\right)}{\sum_{i=0}^{i=N-1} w_i(x^{'}) w_i(y^{'})}. \tag{90.8}$$

After this step, the output image corresponding to the band pass space is $L_{o\mathrm{out}}, L_{1\mathrm{out}}, \ldots, L_{N-1\mathrm{out}}$.

4. Combination of the band-pass layer, then get the final splicing image, as follow Eq. (90.9):

$$G_{\mathrm{out}} = \sum_{k=0}^{N} L_{k\mathrm{out}}. \tag{90.9}$$

It can be seen through experiments, multi-resolution fusion method is better fusion methods, makes up the image is clear and smooth.

90.5 Experiment Results

Using the methods that described above generate the panorama and analysis the results.

90.5.1 Experiment Process

1. Hand-held camera by the way, shots two images of the Jiaozuo Teachers College Library. Shown in Figs. 90.2 and 90.3.
2. Extract the feature points and remove the outer points. Shown in Figs. 90.4, 90.5, 90.6 and 90.7.
3. To get the panorama with the multi-resolution fusion algorithm. Shown in Fig. 90.8.

Fig. 90.2 Jiaozuo teachers college library a

Fig. 90.3 Jiaozuo teachers college library b

Fig. 90.4 Harris feature point extraction a

90.5.2 Experimental Results

More than two spliced images due to be shot from a distance, and laboratory equipment is limited, the images have a certain perspective distortion. Using the methods that study in this paper, the number of feature points extracted more reasonable, the false corner less, and the extraction time was significantly shorter. And the methods also have some anti-noise, fast match algorithm, and strong self-adaptability.

Fig. 90.5 After remove the outer point a

Fig. 90.6 Harris feature point extraction b

Fig. 90.7 After remove the outer point b

Fig. 90.8 Jiaozuo teachers college library panorama

Acknowledgments This subject stems from "Natural Science Basic Research Program" of the education department of Henan province (serial number: 2011C520014). The thanks would go to my beloved family for their loving considerations and great confidence in me all through these years.

References

1. Hua S, Zeng L, Ou Z (2006) Fast algorithm for cylindrical panoramic image mosaic [J]. J Data Acquis Process 21:434–438
2. Zhang J, Hu Z, Liu Z (2007) Image automatic mosaics based on contour phase correlation [J]. Frontiers Mech Eng China 2:228–234
3. Shum HY, Kang SB (2009) A review of image based rendering techniques [C]. IEEE SPIE visual communications and image processing, pp 2–13
4. Peleg S, Ben-Ezra M, Pritch Y (2008) Omni stereo: Panoramic stereo imaging [J]. IEEE Trans Pattern Anal Mach Intell 23, p 279
5. Chen H, Long A (2009) Standard hand-held camera by the uncertain structural panorama pictures [J]. China J Comput 32:328–335

Chapter 91
Image Quality Assessment Based on Complex Representation of Structure Information

Yong Wang, Xiaohui Zhao, Xiuling Mo and Yuqing Wang

Abstract Local variance and single pixel value are combined to describe image structure information using complex method in this paper in order to improve the consistency of objective image assessment result with that of subjective method. The effect of the detail structure information on the image quality was emphasized by this method accordingly. Singular value decomposition was performed on local variance distribution complex matrix. The angle between the singular value vectors of the reference image and the disturbed image was used to measure their structural similarity. Then the quality assessment process was achieved. Results from experiments show that the proposed method is better consistent with human visual system characteristics than MSE, PSNR, and SSIM.

Keywords Local variance · Complex matrix · Singular value decomposition · Image quality assessment · Angle

91.1 Introduction

It is necessary to evaluate image quality by objective method in order to obtain quantified results. Subjective evaluation method is a common way to perform the task, but the result is unreliable due to the influence of several personal factors.

Y. Wang (✉) · X. Zhao · X. Mo
College of Communication Engineering, Jilin University,
Changchun 130012, China
e-mail: wgyg82@yahoo.com.cn

Y. Wang
Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_91,

Furthermore, it is very difficult to combine the subjective evaluation method with the image processing systems or arithmetics. The index to assess the objective quality assessment method is the consistency of the result with that of the subjective human evaluation [1]. Typical full-reference methods, such as the mean-squared error (MSE) and peak signal to noise ratio (PSNR), are proved to be poor correlation with the human visual system (HVS) [2–4]. Due to the complexity of the HVS, the HVS based methods [5–7] do not appear to be very inaccurate and superior to the simple pixel-by-pixel assessment methods [3]. Structural Similarity (SSIM) method in paper [2] has been proved to be versatile and robust in many applications, though its performance is better than MSE and PSNR, its different sensitivity to different types of degradation results in the inconsistency with the assessment results obtained by the subjective judgment based on perceptive information [4, 8].

In this paper, the local variance distribution is taken into account, the image is encoded into a complex matrix for quality assessment and the structural information of an image which is sensitive to human eyes is emphasized. Singular value decomposition of the complex matrix was performed. Due to the unique properties of the singular value feature vector, it is used to represent the main structure information of an image. The quality of the distorted image is assessed by evaluating the structure similarity of two images which is performed by calculating the angle between the two singular value feature vectors of each two blocks.

91.2 Complex Representation of Structure Information

Based on the assumption that a great amount of the structural information of an image is coded in its local variance distribution, in paper [8], a new grayscale image quality assessment method was proposed. The index proposed in it has been proved to be very sensitive to the image distortion type of blur. We denote $\mathrm{Var}(I_{x,y})$ as the local variance of grayscale image I. Since local variance distribution is too sensitive for detail information losing to describe the image structure exactly [8], we use the combination of local variance and pixel value as:

$$J_{x,y} = \mathrm{Var}(I_{x,y}) + I_{x,y} \cdot i$$

where $J_{x,y}$ is used to compose the assessment matrix for assessing the quality of the distorted images. Singular value decomposition of each block is performed. The process can be described by Fig. 91.1.

The local variance can be computed within a local square window centered at the pixel (x, y) which moves pixel-by-pixel over the entire image I. $I_{x,y}$ represents an image block containing L pixels, and the pixel in it is denoted as η_p, then the local variance for a complex block is defined as:

$$\mathrm{Var}(I_{x,y}) = \frac{\sum_{p=1}^{L} \omega_p (\eta_p - \bar{I}x, y)^2}{\sum_{p=1}^{L} \omega_p} \tag{91.1}$$

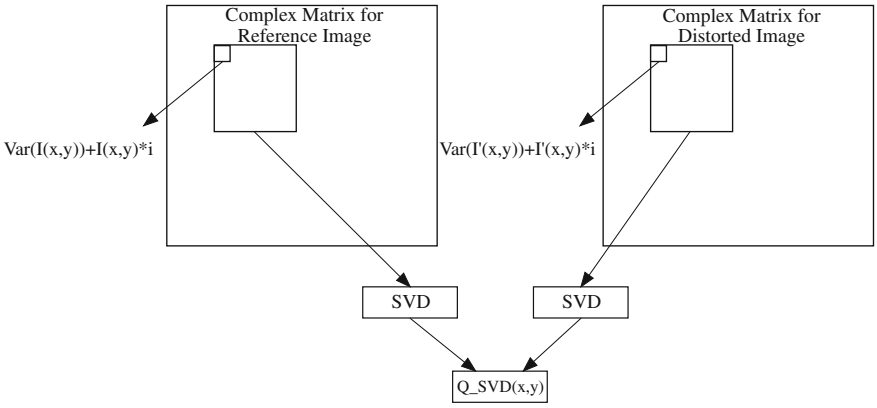


Fig. 91.1 Process of the proposed image quality assessment method

where

$$\bar{I}_{x,y} = \frac{\sum_{p=1}^{L} \omega_p \eta_p}{\sum_{p=1}^{L} \omega_p}$$

In the proposed method, $\mathrm{Var}(I_{x,y})$ and $\bar{I}_{x,y}$ are actually calculated using a weighted neighborhood η_p, and 11×11 circular-symmetric Gaussian weighting function $w = \{\omega_i | i = 1, 2, \ldots, N\}$ [2], with standard deviation of 1.5 samples, normalized to unit sum $\sum_{i=1}^{N} \omega_i = 1$ is used to perform the task.

Similarity of each two vectors can be measured by the angle between them [4]. In paper [3], a numerical measure for grayscale image was proposed based on the SVD of a grayscale image. We use angle instead of distance to measure the singular value feature vector similarity.

Let

$$x = (\sigma_1, \sigma_2, \ldots, \sigma_p, 0, \ldots, 0)^T, \quad \hat{x} = (\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_q, 0, \ldots, 0)^T$$

where x and $\hat{x}$ are the singular value feature vectors of the quaternion matrices corresponding to the two complex matrix J and J'. Then the angle between x and $\hat{x}$ can also be used to measure their similarity. We take the angle between x and $\hat{x}$ as the assessment index Q_SVD, Q_SVD of each block can be calculated as:

$$\mathrm{Q\_SVD}(x, y) = \arccos \frac{\sum_{i=1}^{k} (\sigma_i \times \hat{\sigma}_i)}{\sqrt{\sum_{i=1}^{k} (\sigma_i \times \sigma_i)} \sqrt{\sum_{i=1}^{k} (\hat{\sigma}_i \times \hat{\sigma}_i)}}$$

where $k = \min[p, q]$. The dynamic rang of $\mathrm{Q\_SVD}(x, y)$ is $[0, \pi/2]$, with the best value $\mathrm{Q\_SVD}(x, y) = 0$ achieved when there is no difference between the two blocks. Quality assessment result of the distorted image is:

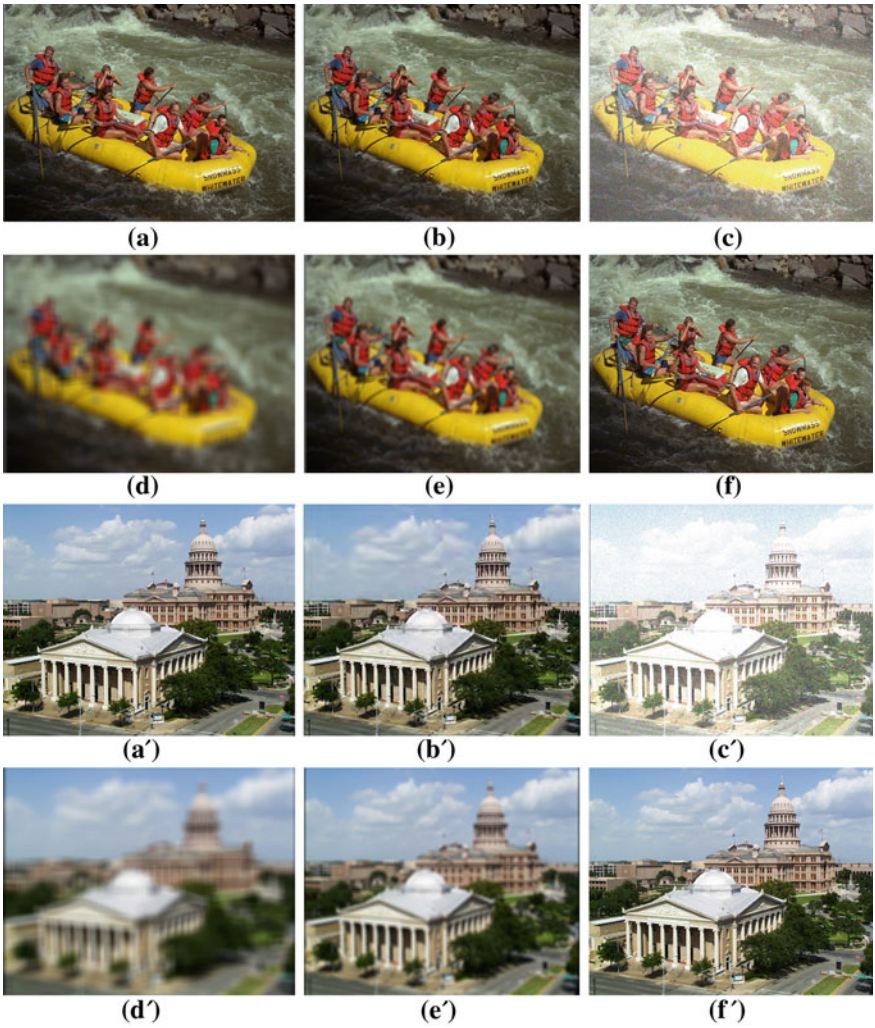


Fig. 91.2 **a** and **a′** Reference images. **b** and **b′** JPEG2000 compressed images. **c** and **c′** Images with additive noise. **d** and **d′** Blurred images. **e** and **e′** Blurred images. **f** and **f′** Images plus constant

$$\text{Q\_SVD} = \frac{1}{m}\left(\sum_{i=1}^{m} \text{Q\_SVD}(i)\right) \tag{91.2}$$

where m is the block number. The singular value feature vector is unique for the corresponding complex matrix. Since the matrix is the representation for some structural information of the image, the singular value feature vector can be applied to describe some of its important properties.

Table 91.1 Assessment results of proposed method and other methods for the luminance layer of the distorted images in Fig. 91.2

Figures	SSIM	MSE	PSNR(dB)	Q_SVD(rad)
91.2a	1	1	0	0
91.2b	0.7701	83	28.9382	0.0083
91.2c	0.4443	6475.7	10.0179	0.0114
91.2d	0.4081	516.8	20.9975	0.0206
91.2e	0.4993	318.2	23.1036	0.0185
91.2f	0.9955	35.9	32.5756	0.0004
91.2a′	1	1	0	0
91.2b′	0.8364	105.8	27.8857	0.0067
91.2c′	0.5245	5402	10.8053	0.0099
91.2d′	0.4412	1203.2	17.3273	0.0165
91.2e′	0.5552	681.6	19.7957	0.0138
91.2f′	0.9972	36	32.5734	0.0003

91.3 Experimental Results

Performance of Q_SVD proposed in this paper is investigated by some experiments. Figure 91.2a and a′ are the reference images. The distortion types for other distorted images are given below:

(1) The distorted images are generated by compressing the reference images (full color) using JPEG2000 at bit rates 0.41906 bits per pixel (bpp) for Fig. 91.2b and 0.38308 bits per pixel (bpp) for Fig. 91.2b′. Kakadu version 2.2 [9] is used to generate the JPEG2000 compressed image.
(2) White noise is added to the images, with 0 mean and $\sigma = 0.32$, Fig. 91.2c and c′.
(3) The reference images are blurred using a 21×21 averaging kernel, Fig. 91.2d and d′.
(4) The reference images are blurred using a 9×9 averaging kernel, Fig. 91.2e and e′.
(5) Reference images plus constant 6, Fig. 91.2f and f′.

Some other assessment methods such as SSIM [2], MSE and PSNR are used in this experiment for comparing. Since they can only assess the grayscale image, we obtain the assessment results by using these methods to assess the luminance layer. The assessment results are listed in Table 91.1.

In order to evaluate the behaviours of the proposed method, four types of assessment methods are evaluated by assessing the quality of the distorted images in Fig. 91.2. Loss of detail information in image results in rapid descending of image perceptive quality. Obviously, in Fig. 91.2, the quality of blurred images are worse than that of the other images, but as it can be easily seen from the assessment results in Table 91.1 that the Q_SVD method proposed in this paper assessed the quality of the disturbed images more rationally than what the other methods assessed. Performance of MSE and PSNR is the worst in the four

methods. Due to the low sensitivity to blur, SSIM assessed the blurred images such as Fig. 91.2e and e′ in accurately. The best quality of Fig. 91.2d and d′ are given by Q_SVD. For the blurred images its assessment results are worse than those of the other distorted images. Its assessment results for the distorted images in Fig. 91.2 with different types of distortion are better consistent with the HVS.

91.4 Conclusions

Human eyes are more sensitive to the detail information in an image due to some characteristics of HVS. Local variance distribution of an image can well represent its detail information. We encode some important structural information into a complex matrix in order to assess its quality by using local variance distribution. Singular value feature vector is taken as a tool to describe the information. Structural similarity of each two blocks can be measured accurately. From the experiments carried out in this paper we can see that the behaviors of the assessment method proposed in this paper are more consistent with the HVS than those of some of the state-of-the-art assessment methods.

References

1. Sheikh HR, Sabir MF, Bovik AC (2006) A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans Image Process 15(11):3441–3452
2. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
3. Shnayderman A, Gusev A, Eskicioglu AM (2006) An SVD-based grayscale image quality measure for local and global assessment. IEEE Trans Image Process 15(2):422–429
4. Wang Y, Liu W, Wang Y (2008) Color image quality assessment based on quaternion singular value decomposition. The 2008 international congress on image and signal processing, Sanya, Hainan, China, vol 3. p 433 (May 2008)
5. Westen SJP, Lagendijk RL, Biemond J (1995) Perceptual image quality based on a multiple channel HVS model. IEEE 1995 international conference on acoustics, speech, and signal processing, Detroit, MI, USA, pp 2351–2354
6. Fernandez-Maloigne C, Larabi M-C, Bringier B, Richard N (2005) Spatio-temporal characteristics of the human color perception for digital quality assessment. IEEE signals, circuits and systems ISSCS 2005 international symposium, IEEE, New York, pp 203–206
7. Barni M, Bartolini F, De Rosa A (2002) HVS modelling for quality evaluation of art images. IEEE 14th international conference on digital signal processing, Santorini, Greece, July 2002, pp 91–94
8. Aja-Fernandez S, Estepar RSJ, Alberola-Lopez C, Westin, C-F (2006) Image quality assessment based on local variance. IEEE 28th EMBS annual international conference, New York City, USA, pp 4815–4818
9. Taubman DS, Marcellin MW (2001) JPEG2000: image compression fundamentals, standards, and practice. Kluwer, Norwell

Chapter 92
Dust Monitoring System Based on Video Image Processing

Yi Chu, Guohai Liu, Congli Mei and Yuhan Ding

Abstract This paper describes the discharge of smoke during the process of smelting in the steel factory. A video dynamic monitoring method is proposed to monitor the dust emission and extract the dust by using the Otsu threshold segmentation method at any time. After the real-time calculating, it can make a judgment of the concentration and diffusion rate of the dust. The technology of adjusting and optimizing the power of the dust-fan by using the frequency-conversion can avoid the great energy waste compared with the traditional dust-fan which is running at full speed. The technology mentioned in this paper makes the factory save a lot of electricity.

Keywords Steel factory · Dust · Video monitoring · Frequency-conversion · Dust-fan · Energy saving

Y. Chu (✉) · G. Liu · C. Mei · Y. Ding
College of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China
e-mail: chuyi19871125_@126.com

G. Liu
e-mail: ghliu@ujs.edu.cn

C. Mei
e-mail: clmei@ujs.edu.cn

Y. Ding
e-mail: yhding@ujs.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_92,

92.1 Introduction

The electric arc furnace plays an important role in modern steel industry. The symbolic significance is the birth of large DC electric arc furnace and its systematization technology for reducing energy consumption, shortening the refining cycle, using a variety of resources, improving the quality, and reducing the environmental load. However, the steel industry has long been considered to be a factory with large dust emissions, much waste, and serious pollution. Indeed, the steel-making process of electric arc furnace is the main source of air pollution. To eliminate the damage to the environment of dust, the steel factory all equip the fans dust collector. Traditionally, the fun runs at full speed regardless in which run phase or the size of the dust which wastes a lot of energy [1].

Currently, widespread use of dust particle size test technology contains sieving method, microscopic method, sedimentation method, electric field induction region method and optical technology test method. Among these methods, sieving method is difficult to measure bond and reunite materials for the poor reproducibility result. Microscopic method can only check up less particles, sometimes cannot reflect the whole sample level with its large error. Sedimentation method takes a long time for its slow speed. Electric field induction region method must convert the conversion to get particles size information.

There are differences in effect between Computer vision and particle size testing technology due to image segmentation, pattern recognition methods and algorithms have their own characteristics in spite it made a lot of achievements in recent years. Indeed, this image analysis system is hard to operate and the interface is too complex. At present the main existing problems in the following respects [2]:

1. Domestic and foreign general function image instrument is stronger in function but because of the lack of and the combination of professional field, so general special application function is poor. It can get general image preprocessing or analysis but cannot make special application analysis in-depth.
2. Generally, it is limited to gray segmentation and false color segmentation and cannot undertake true color segmentation so the accuracy of the segmentation target extraction is poor.
3. It only can provide basic measuring data analysis and the materials need second processing which is hard to meet the customer needs.

In order to overcome these above problems, this paper proposes a static background based on the smoke monitoring system. This system contains several modules such as video image capture module, processing and basic parameter calculation module. The program and modules communication links is achieved by Visual C++. Firstly, take some time, a few adjacent frames, and then make the subtraction operation, threshold segmentation, binarization, morphological processing to the result. In the last, and calculate each of images of smoke and dust concentration and compare every data, from which the diffusion degree of dust can be judged. The basis is the adjustment of the rate for fan.

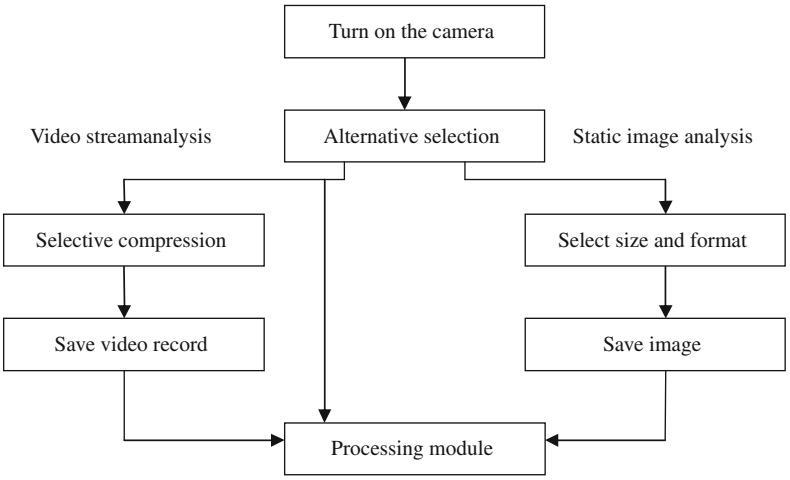


Fig. 92.1 Work flow chart of the system

92.2 Dynamic Monitoring Algorithm of Dust Based on the Static Background

92.2.1 Basic Principle of the System

The system uses a video monitoring method to the exhaust smoke with real-time calculation, which can be the adjustment of the rate for fan. The specific method is as follows: a cameras is connected to computers, control the camera, the video image module in the system processing and analysis the video image and the results is sent to the computer. The system will real-time alarm if the results is larger than the established threshold. The power of the fan is adjusted by the frequency converter, thus the energy saving goals is achieved.

92.2.2 Work Flow Chat of the System

The real-time camera shot the smoke and dust, then sent the data to computer synchronously. The system takes a different method of analysis between video and image in order to understand the movement of the smoke and some key parameters qualitatively and quantitatively. Work flow chat of the system is typically shown in Fig. 92.1.

The moving objects are obtained by filtering out the background in the video if we just want to know the movement of smoke and dust qualitatively. In the system, the program is programed by OpenCv function library for the process is very complex.

The input signal uses the image to calculate the rate of movement and concentration of the dust quantificationally. Moving objects is obtained by filtering out the background in the video with subtraction operation for the camera is fixed. Much sharper images is obtained from the smoke after the binarization transformation and morphological processing. The instant concentration of smoke can be computed according to the target image in the whole image of the area; compare the two data, the judgment that the smoke is spreading is obtained if the second result is larger than the first one. Of course it is the same in reverse.

92.2.3 Threshold Segmentation Method

The correct choice of threshold which will directly affect the segmentation effect is the most important step to set the target from a complex background extracted from the full [3]. The gray value is judged to be object if it is larger than the threshold, or it is background. The value is various in different environment. The Formula is as follows:

$$f_t(x,y) = \begin{cases} \text{background} & f(x,y) < T \\ \text{motion} & f(x,y) \geq T \end{cases} \tag{92.1}$$

The background point will be mistaken for the object if the threshold is too low; oppositely, the object point is mistaken for the background. In order to overcome the situation, this paper uses a classical threshold segmentation algorithm—Otsu threshold method, the basic principle as follows [4, 5]: the size of the image is $P \times Q$. The gray value in (x,y) is $f(x,y)$ which limits from 0 to $L-1$. The number of gray value i is n_i, then the probability is $P_i = n_i/(P \times Q)$, and $P_i \geq 0$, $\sum_{i=0}^{L-1} P_i = 1$. If the threshold is divided into two categories from gray level t :S_1(background) contains the pixel $i \leq t$ and S_2(object) contains the pixel $i \geq t$. The probabilities of S_1 and S_2 are $P_1 = \sum_{i=0}^{t} P_i$, $P_2 = \sum_{i=t+1}^{L-1} P_i$, $P_1 + P_2 = 1$. The mean values of the two categories are $w_1 = \sum_{i=t+1}^{L-1} iP_i/P_1$, $w_2 = \sum_{i=t+1}^{L-1} iP_i/P_2$. The mean value of the whole image is $w_0 = \sum_{i=0}^{L-1} iP_i$. The infra-class of the two categories is $\delta^2 = P_1[w_1 - w_0]^2 + P_2[w_2 - w_0]^2$ according to the pattern recognition theory. To achieve the best image segmentation result, the best classification results and the most suitable must be insured. The optimal threshold formula is $t^* = \arg\max(\delta^2)$.

92.2.4 Extraction of Object Dust

There will be a lot of little holes, isolated small points and broken contour lines in the object area in the difference image after binarization operation. This problem can be solved by morphological filtering method The basic idea is to define

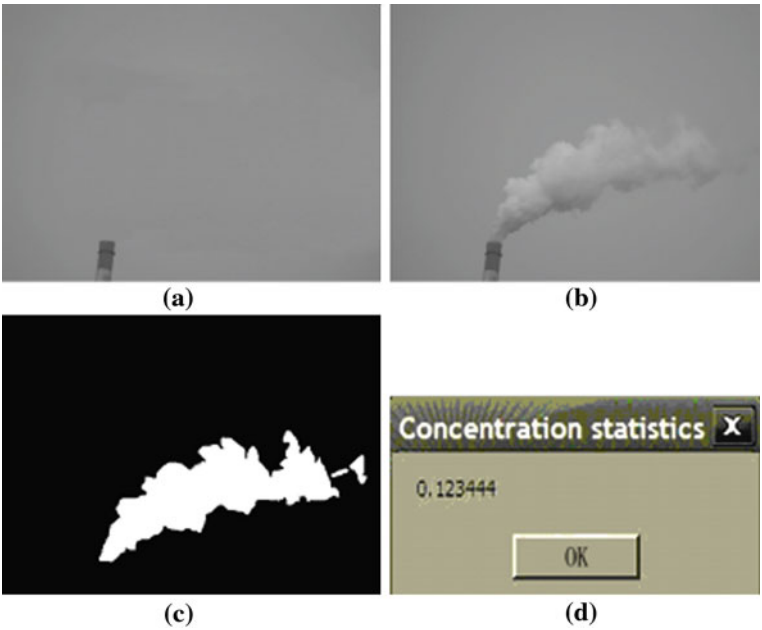


Fig. 92.2 Experiment result of time t_1. **a** Background image. **b** Image with dust. **c** Object after binarization. **d** Concentration

morphological elements, filtered out isolated noise points by corrosion template, use the same size of the expansion template filtering, process the fractured or empty part, then achieve the goal of image analysis and recognition. This means make two basic operations to the image: dilation and corrosion. The major purpose of dilation is to increase the number of pixels object and then fill the hole and the bridge crack. Corrosion is the inverse of dilation, its role is to eliminate image isolated small point and unnecessary small connection [6].

92.3 Design of Monitoring System for Dynamic Dust

92.3.1 Design Performance Indicators the System Should Have

(a) Dust emissions at any time, production environments and operation of electric arc furnace can be monitored in the monitoring center. The shooting modes of the camera can be controlled and it must have the function of photography and video in necessary.
(b) In the monitoring center, videos and images which are shot in the production environment can be compressed and saved according to the different forms of need.

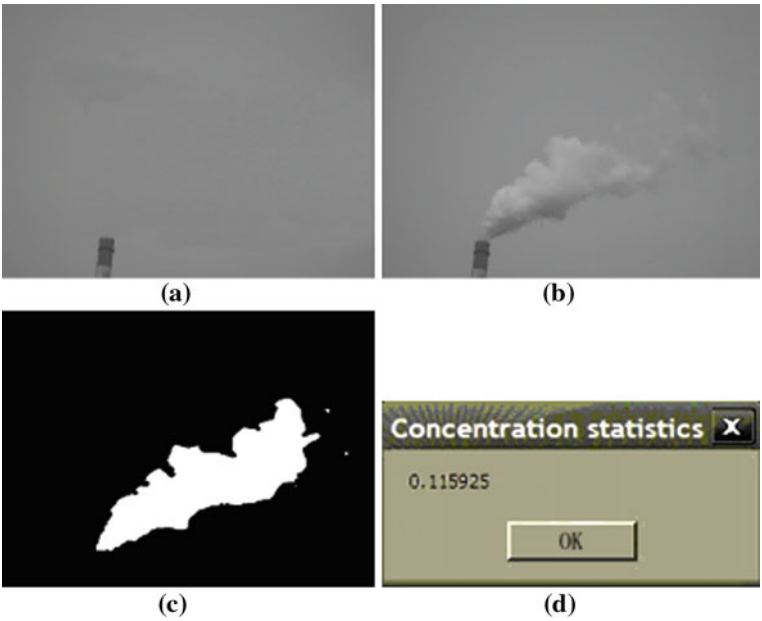


Fig. 92.3 Experiment result of time t_2. **a** Background image. **b** Image with dust. **c** Object after binarization. **d** Concentration

(c) Video and image processing system in the monitoring center can calculate the size of smoke concentration and compare with the empirical value.
(d) Computer in the monitoring center can record the data and save to the database, inquires at any time, print the record and statistics list every time.

92.3.2 Hardware Design

Computer configuration: Operating System of Windows XP Professional, Processor of Intel Pentium E6300, Mainboard of P5G41C-M, RAM of 2 GB DDR3 and Hard disk of 500G.

Camera parameters: Microvision industrial digital camera, Maximum resolution of 1024 × 768, capture rate of 60 f/s, focal distance of 4.8–46 mm, S/N ratio >52 dB, output mode of FireWire IEEE1394A.

92.3.3 Software Platform

In Windows operating system platform, Visual C++6.0 for integrated development environment, Call part function library of OpenCv and MFC and complete the video images process system based on MFC.

92.4 Experiment Results

The steel making cycle is 70—85 min: loading preparation 6–10%, melting 25–30%, oxygen blowing 30–35%, deoxidation period 15–20% and the steel slag rushing 6–8%. Based on the time interval above, the system shoots an image and analysis every 10 min. Do subtraction operation between the two images and the background, binarization and morphological processing. At the last, the concentration of dust is achieved, from which the situation of dusk spread of can be obtained quantificationally. Figures 92.2 and 92.3 are the results of time t1 and t2.

From the experimental results above, the concentration of dust at time t2 is lower than time t1, so the frequency converter should be adjusted to decrease the power of the fan, which can not only get the goal of adsorbing most of the smoke but also save much energy.

92.5 Conclusion

The system is mainly for the detection of the size and concentration of the dust which is exhausted in the steel factory. A video monitoring detection method which is different from the traditional technology is proposed. The original system is reconstructed overall due to its real-time and high efficiency. At the same time, the research in this paper provides a simple and accurate and fast image test methods to the ultra-fine particle detection in the chemical industry and the healthcare system. With the image processing technology and testing technology development, the application range of the particle detection images will be larger than now. Image particle detection will be toward the direction of the higher accuracy faster development. The application prospect of image particle detection is very wide. The development direction of measurement or statistical analysis to fine particle or dust based on the image is very glorious.

Acknowledgments This project was supported by the Special Technology Cultivation Project of Energy Saving and Emission Reduction of Social Development in Zhenjiang (SH2008005) and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

1. Yu XL (2007) Study on energy saving in power plant fan and its modification. Water Conserv Electr Power Mach 29(10):22–27
2. Sun ZB (2003) Application and research on the measurement of particle size using computer vision[D]. Jiangsu University, China
3. Liu JQ, Gao JM, Shen QM (2009) A background separation method of nonuniform image segmentation. IEEE industrial electronics and applications, ICIEA 2009. pp 3049–3053

4. Xu XY, Xu SZ, Jin LH et al (2011) Characteristic analysis of Otsu threshold and its applications. Pattern Recognit Lett 32(7):956–961
5. Zhang GM, Chen SP, Liao JN (2011) Otsu image segmentation algorithm based on morphology and wavelet transformation. Computer research and development (ICCRD), 2011 3rd international conference, pp 279–283
6. Francisco Estrada J, Allan DJ (2009) Benchmarking image segmentation algorithms. Int J Comput Vis 85(2):430–443

Chapter 93
Vector Cartoon Generating Method Based on Layer Representation

Ruxin Gao and Meihong Wang

Abstract This paper presents a vector cartoon generating system based on layer representation. First segment the image into regions and curves, and represent the whole image with an And-Or graph. The segmentation method is graph cut and the curves are the primal sketch in each region. Represent the regions and corresponding curves with layer representation method, this will have a layered vector graph.

Keywords Cartoon generating · Layer representation · 2.1D sketch · Primal sketch · Image segmentation

93.1 Introduction

Cartoon generating from image is an important and interesting task. Currently, some related researches [1] have been conducted. Most of the existing methods can only generate "rigid cartoon", whereby the users cannot edit or adjust any parts and they need to do many repeated tasks, which make it ineffective.

Tu et al. [2] think an image can be parsed into regional components and sketch components in paper. Region components interpret surface information and sketch components interpret contours, drapes, edges and some strong textures. Inspired by Tu's idea, this paper presents a vector cartoon generating system from image by

R. Gao (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, Henan, China
e-mail: gaoruxin@gmail.com

M. Wang
Construction Department, Henan Polytechnic University,
Jiaozuo, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_93,

interactive graph cut and primal sketch corresponding region and sketch components, respectively, and supposing region components and sketch components of objects without intersection and together represent full objects. The region components and sketch components are modeled together by 2.1D sketch, which includes layer information for objects, then the users can adjust the layers of objects and distort the shapes of objects.

93.2 Representation of Graph

93.2.1 Definition of Graph Structure

Given an image I with a number of objects, the cartoon coming from the image can be denoted as:

$$G = G_1 \bigcup G_2 \cdots \bigcup G_i \cdots \bigcup G_L \tag{93.1}$$

Here, G_i is a subgraph of a cartoon, and a basic group or mass composing full objects, which can be an object or a part of an object. Supposing there are $L(L \geq 1)$ subgraphs and each subgraph has region components and sketch components:

$$G_i = G_{\text{seg}_j} \bigcup G_{\text{sketch}_k} \tag{93.2}$$

G_{seg_j} is the region components of the ith subgraph, G_{sketch_k} is the sketch components of the ith subgraph. The whole cartoon region components set are:

$$G_{\text{seg}} = G_{\text{seg}_1} \bigcup G_{\text{seg}_2} \cdots \bigcup G_{\text{seg}_i} \cdots \bigcup G_{\text{seg}_{NN}} \tag{93.3}$$

Here, each region component G_{seg_i} can have a number $(N_i \geq 1)$ of regions, denoted as:

$$G_{\text{seg}_i} = R_1 \bigcup R_2 \cdots \bigcup R_{N_i} \tag{93.4}$$

There is no common region between subgraphs, that is $G_{\text{seg}_i} \bigcap G_{\text{seg}_j} = \Phi, \forall i \neq j$.

G_{sketch_j} is composed with a number of curves (supposing M, $M \geq 1$), denoted as:

$$G_{\text{sketch}} = G_{s1} \bigcup G_{s2} \cdots \bigcup G_{sj} \cdots \bigcup G_M \tag{93.5}$$

Each sketch components may have a number of curves (supposing K, $K \geq 1$), denoted as: $G_{sj} = C_1 \bigcup C_2 \cdots \bigcup C_j \cdots \bigcup G_M$, C_j is one of the curves of the subgraph.

In the graph representation, supposing each sketch component G_{sj} is in corresponding region R_j, and supposing all regions R_j and sketch components in it have the same layer information, and do not change the relation during reference.

93.2.2 Layered Representation of the Cartoon

Suppose the graph representation of the cartoon is denoted as $G = <V, E>$, detecting the 3-degree-vortex [3] using the method presented by Wu [4], the 2D graph representation is denoted as:

$$W_{2D} = (V_R, V_T) \tag{93.6}$$

The 2D representation W_{2D} includes two vortex sets: one for region, V_R and another for terminator, V_T [3]. The vertex sets for region are denoted as:

$$V_R = (R_1 \bigcup G_1, R_2 \bigcup G_2, \ldots, R_N \bigcup G_N) \tag{93.7}$$

Here, G_i is the curves set in region R_i.
The vertex set for terminator V_T is:

$$V_T = (T_1, T_2, \ldots, T_M) \tag{93.8}$$

The layered representation $W_{\text{layer}}$is defined as:

$$W_{\text{layer}} = (X, Y; Sf, Ct) \tag{93.9}$$

$$X = (x_{R_1}, x_{R_2}, \ldots, x_{R_N}) \tag{93.10}$$

X is the label set, used to interpret the layer of a region; Y is the variables set, used to descript the address variables assignment. The connected terminators which are attached to the same layer through address variables assignment and corresponding regions can be grouped into surfaces. Let Sf denote the surface set. Each surface can have a region or some merged regions which are cued with address variables assignment. Supposing recovered contours are denoted as Ct which are from paper [5]. The variables set are denoted as:

$$Y = (y_{a_1}, y_{a_2}, \ldots, y_{a_M}) \tag{93.11}$$

Suppose all vertexes are attached to a layer, and each layer has a label, and assuming there are K layers, then the layer set is denoted as: $l_i \in \{1, 2, \ldots, K\}$.

93.3 Cartoon Generating System

There are three tasks that need to be dealt with at the same time. The first is how to generate regions. If auto-segmentation method is used, then image may be cut into many fragmentary regions that cannot meet the requirements of cartoon, hence interactive graph cut method is used in this paper. The second is the sketch generation method in regions; there are many disorder and redundant line segments using class edge detection method such as Canny and primal sketch method used in this paper. This method can pursuit realistic contours from the original image. The third is to

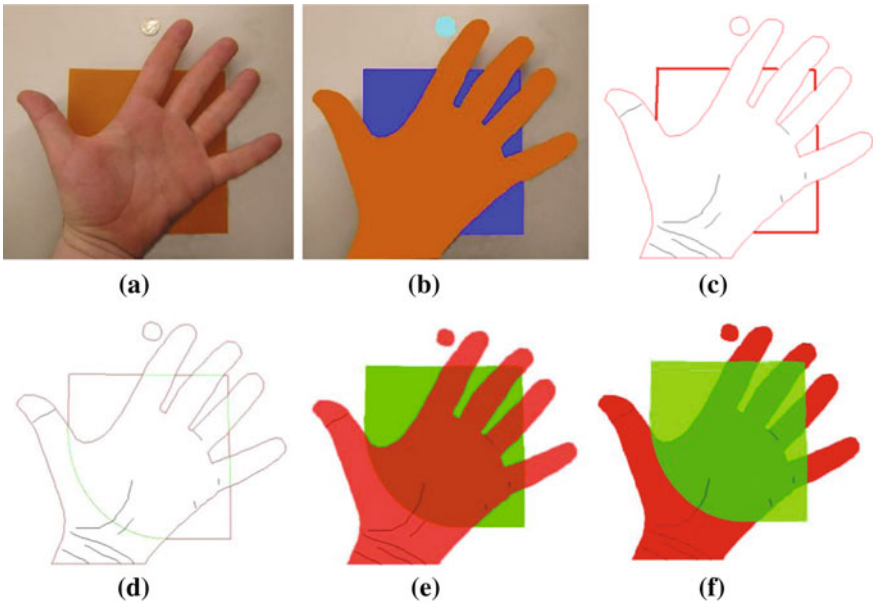


Fig. 93.1 Process of cartoon generating. **a** Input image. **b** Regions by interactive graph cut. **c** Primal sketch in regions (*black thin curves*) and terminator detecting results (*red broad curves*). **d** Recovered curves (*green thin curves*). **e**, **f** Two layered representation results by users editing

represent cartoon by a general method. There are no suitable methods to deal with the task,hence layered graph representation method [6] is used in this paper.

An example is shown in Fig. 93.1a, whereby the input image is segmented into several regions in Fig. 93.1b by interactive graph cut method. Then the sketch is computed in each region by primal sketch method, as shown in Fig. 93.1c, the black thin curves. The 3-degree-vertex is detected by using the method in paper [2] shown in Fig. 93.1c, the red broad curves, then contours recovering and layered representation are done in Fig. 93.1d, the green curves are the recovered curves, the layered results are shown in Fig. 93.1(e, f), the two results are edited by users by adjusting the layers of objects, when the hidden layer is as the foreground, the occluded curves is shown.

In order to make the cartoon more slinky and realistic, it is necessary to render the contours and sketches. The B-spline and snake [7] is used to smooth the curves and touch up the real boundaries. The cartoon generating diagram is shown in Fig. 93.2.

93.4 Experiments Results

There are different kinds of experiments shown in this paper; one is without occlusion relation of the objects and another is that there are several objects in the input image and one object occludes other objects, layers information is needed.

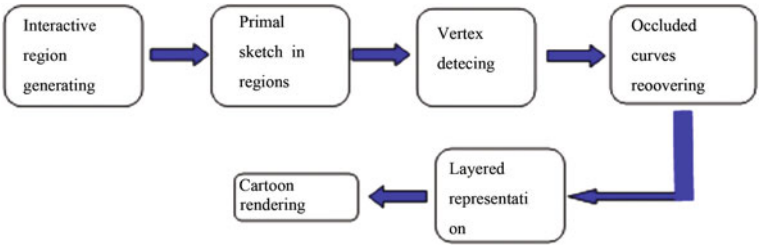


Fig. 93.2 Cartoon generating diagram

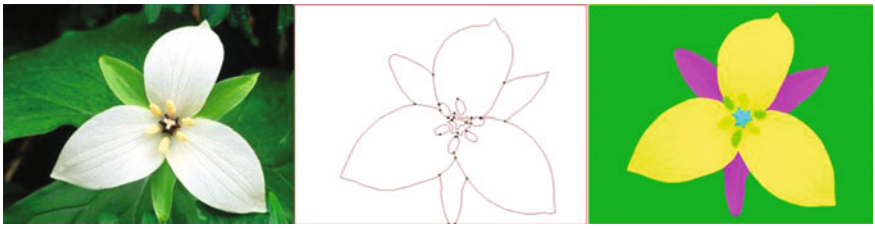


Fig. 93.3 Experimental result without occlusion relation

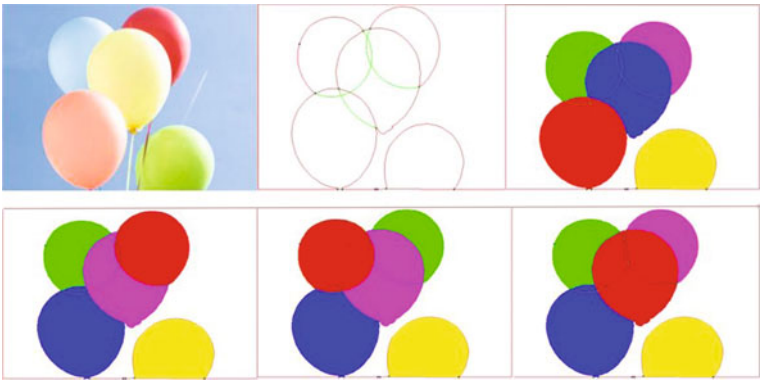


Fig. 93.4 Experimental result with several objects and occlusion relation. The occlusion relation can be edited by the user

An experimental result is shown in Fig. 93.3 as without occlusion relation from input nature image, all objects are visible and without occluded curves. The objects, the petals in original input image belong to the same object, a flower, which suits the common sense of people without occlusion relation.

Another experiment result is shown in Fig. 93.4, there are several balloons, each one is a single object, so it is suitable to build the layer relation for the objects in the example. The users can adjust the layer relation, that is the occlusion relation.

93.5 Conclusion and Future Work

Inspired by Tu's idea that objects in scene can be parsed into regions part and sketch parts, the paper presents a vector cartoon generating system, regions by interactive graph cut and sketches by primal sketch, then the whole sketch graph is represented by layered representation method is a living vector graph, and can be edited by users. The representation method is a general one and can be expanded into more applications.

Currently, the cartoon rendering is simple; background beautifying and color harmonizing of the whole cartoon is the future work.

Acknowledgments The work is supported by Doctors' Funds of Henan Polytechnic University (No: B2010-17).

References

1. Chen H, Zhu SC (2006) Composite templates for cloth modeling and sketching. In: IEEE international conference on computer vision, pp 943–950
2. Tu ZW, Zhu SC (2006) Parsing image into regions, curves and curve groups. Int J Comput Vis 69(2):223–249
3. Guo CE, Zhu SC, Wu YN (1997) Primal sketch: integrating texture and structure. Comput Vis Image Understanding 66(2):5–19
4. Wu TF, Xia GS, Zhu SC (2007) Computational boosting for computing hierarchical image structures. In: Proceeding of IEEE conference on computer vision and pattern recognition, June 2007, pp 1–8
5. Kimia BB, Frankel I, Popescu AM (2003) Euler spiral for shape completion. Int J Comput Vis 54(1–3):159–182
6. Gao RX, Wu TF, Sang N et al (2007) Bayesian inference for layered representation with mixed Markov random field. In: Proceedings of 6th international conference on energy minimization methods in CVPR, August 2007. Springer Lecture notes in computer science, vol 4679. Ezhou, China, pp 213–224
7. Lass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331

Chapter 94
Linguistic Description for Extracted Corner Features of the Regular-Shaped Rigid Objects

Lei Zhang, Yindan Zhou and Xingguo Zhang

Abstract In this paper, linguistic description for contour features of the regular-shaped rigid objects is studied for future syntactic pattern recognition. First, one-dimensional language is used to describe the contour corner features of every object candidate. Then some rules are made to distinguish between contour features and internal features. Last, but the most important, a modified non-recursive DFS traverse algorithm is put forward to extract the loop of the contour features for every connectivity component and to exclude the contour features which are not on the desired loop. Experimental results show the proposed algorithm can successfully describe the valid contour features in one-dimensional language, which make the preparation for future recognition.

Keywords One-dimensional language · Traverse algorithm · Corner feature

94.1 Introduction

Our future work is to recognize the regular-shaped objects through syntactic pattern recognition. When syntactic pattern recognition is applied to image processing and understanding, image features of the objects should be expressed in

L. Zhang (✉) · Y. Zhou · X. Zhang
School of the Mechanical Engineering, Nantong University, Nantong, China
e-mail: zhangleint@ntu.edu.cn

Y. Zhou
e-mail: zhou.yd@ntu.edu.cn

X. Zhang
e-mail: zhang.xg@ntu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_94,

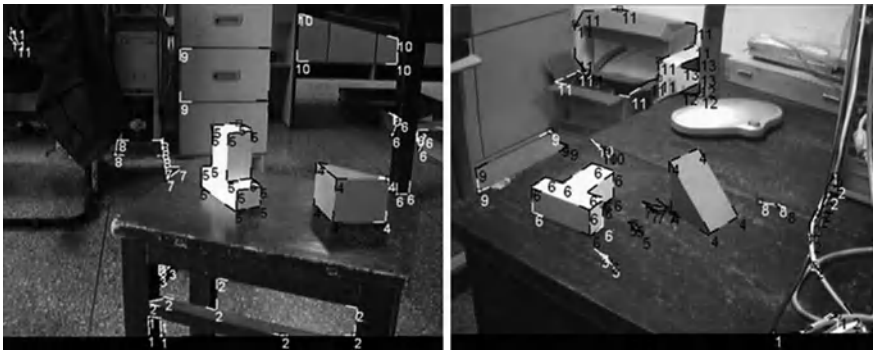


Fig. 94.1 Results of corner extraction in the previous study [2]

a linguistic mode before recognition. Structural features extraction from the regular-shaped objects were studied in the previous reports by the same authors [1, 2].

As before, the corner features are the specific structural features, and a corner in this includes not only the corner point, but also the edges of the corner angle as Fig. 94.1 shows [2]. In this paper, the linguistic description for the corner features of the regular-shaped objects will be studied.

Commonly, every terminal by which a natural or a computer language is presented is connected in one-dimensional space. Hence, the grammar on which the language is based is called "one-dimensional grammar". Such grammar can easily describe one-dimensional structure such as the image contour of an object. However, it has some difficulty in describing two or higher dimensional structure.

Two-dimensional grammar was studied for the past decades, including tree grammar, array grammar and graph grammar as well. But some inherent limitations hamper its application. First, two-dimensional grammar is a kind of elaborate description for the complex two-dimensional objects. Consequently, it is very sensitive to distortion and noise [3]. Second, compiling, translation and parsing theory that has been well developed and successfully used is almost all based on one-dimensional language. Although the translation and parsing theory for the tree grammar and graph grammar was studied by a few researchers [3–11], the successful application was rarely discovered. Thomas [3], truve [4] studied the ideal two-dimensional structural decomposition. Zhang [5, 6] and Zeng [8] focused on the attribute of the two-dimensional grammar, not on the parsing of the grammar and its application. Fang [7] applied the relational graph grammar to a simple application. However, his work was just limited to very ideal and simple line, rectangle and circle shape recognition. Recent reports [10–12] suggested that some new two-dimensional parsing theory was put forward and was studied in-depth. But when applied to practical case, simplification or modification in language presentation [10, 12] is required for less sensitivity to the noise and the variance in

structure. All above show little advantage of the two-dimensional language over one-dimensional language.

Another solution to the description for the two-dimensional structure is PDL language [13, 14]. This language is one-dimensional, yet uses some special terminals to represent some local two-dimensional structures. Unfortunately, PDL also suffers greatly from noise and distortion of the structure [13, 14]. After all, as a one-dimensional language, PDL cannot generally describe the two-dimensional structure [13, 14]. What is more, PDL often cannot be easily understood when corresponded to the structure of the objects.

In this paper, one-dimensional language is used to describe the extracted corner features in view of the following factors:

First, one-dimensional language has very mature compiling and parsing theory for recognition. Successful syntactic pattern recognition cases are almost all based on one-dimensional language.

Second, when combined with fuzzy grammar, one-dimensional language will be less sensitive to noise and distortion, which greatly improves its feasibility in practical application.

Third, it is much more convenient to implement one-dimensional language.

Fourth, in our case, one-dimensional language can basically represent the uniqueness of an object.

The following will be the detailed realization of the linguistic description for the regular-shaped objects.

94.2 Rules for Distinguishment Between Contour Features and Internal Features

In the previous study as Fig. 94.1 shows, the corner features were extracted and the relative graph that represented the connection of the corners was established. It can be found that the corners on the image projection contour (although all of the features are on the surface of an object, when projected on the image, some will be on the image contour and others will be inside the image contour. The former will be called “contour features” and the latter “internal features” in this paper.) of every object candidate play a vital role in representing the shape of the object candidate. Obviously, the contour features are connected in one-dimensional mode. Thereby, it is better to have the contour features to be described in one-dimensional language.

The first work we need to do is to distinguish between the contour features and the internal features. We find that a contour feature can not be directly confirmed only by the local structure of an individual corner feature. Hence, several rules are made according to the global structure and the relationship of the features:

(1) In the image coordinate system, the leftmost, the rightmost, the top and the bottom corners of every object candidate are the original contour features.

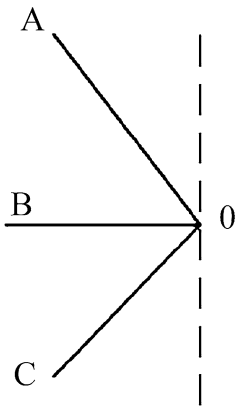


Fig. 94.2 An example of Rule3

(2) If a corner is a contour feature and the corner has only two straight edges, then the two edges are contour features.
(3) If a contour corner has three or more straight edges and all of the edges are distributed at one side of the image plane of the corner as Fig. 94.2 shows that OA, OB and OC are all on the left side of the dashed line, then the edges that are outside are contour features, ex. OA and OC in Fig. 94.2 and the edges that are inside are internal features, ex. OB in Fig. 94.2.
(4) If a corner connects a contour edge, then the corner is a contour feature.

The rules are executed with the help of depth-first search (DFS) algorithm for traveling every feature of every connectivity component. After all the contour features are confirmed or labeled, the left features will be viewed as internal features.

94.3 Traverse Algorithm For Contour Feature Loop

Although a graph that expresses the topological structure of the features is established and contour features are labeled, it does not mean the contour features could be immediately described in a string. A loop along the contour that passes through the related features should be specified and extracted. Here, the non-recursive DFS traverse algorithm is used again, but a new modified one is suggested.

94.3.1 Topological Structure of the Contour Features

According to our previous study as Fig. 94.1 shows, the contour features have the following three typical topological structures.

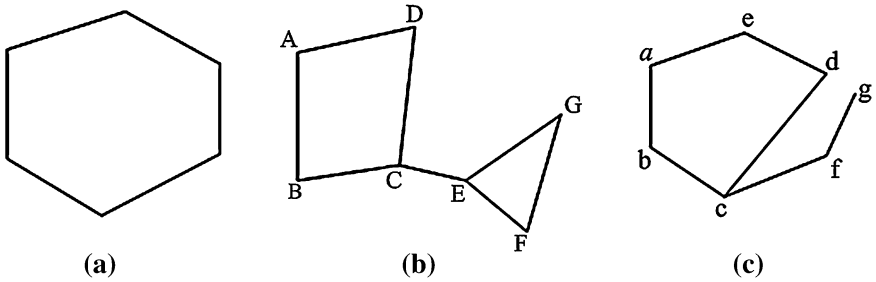


Fig. 94.3 Topology structure of the contour features

Figure 94.3a shows the normal topological structure. But there are some abnormal topological structures as Fig. 94.3b and Fig. 94.3c show. This is mainly caused by the noise and clutter background perturbation.

The algorithm to be designed should be able to specify the only loop for the given starting node (corner feature). For example in Fig. 94.3b, if "A" is a starting node, the only loop for it must be "ABCD". "EFG" loop should be another loop although "A" is connected to "EFG" via some paths. In Fig. 94.3c, "fg" can not be a part of "abcde" loop, and should be rejected as a kind of perturbation.

94.3.2 Modified DFS Traverse Algorithm for Extraction of the Contour Feature Loop

The algorithm for extraction of the contour feature loop is based on the classical non-recursive DFS traverse algorithm. But there are some special issues to be considered.

First, in our work, the traverse algorithm is not used to visit every node (corner feature), but used to visit the nodes that have been labeled "contour features". The internal features do not need to be visited.

Second, the traverse algorithm should have the function that stores the visited nodes in a single linked list that can be easily extracted for future language production.

Third, also a hard nut to crack, how to handle the structure of Fig. 94.3b and c. As for the classical DFS traverse algorithm, it will visit the nodes that are not supposed to be on the current loop. What's more, after those nodes are visited, it can still transfer to the current loop. For example, in Fig. 94.3b, for a starting point "A", when the algorithm visits along "B, C" and the current loop is "ABCD", the nodes "E", "F", "G" will be visited and then "D". Obviously, "E", "F", "G" should be eliminated. But the conventional traverse algorithm does not recognize the nodes "E", "F", "G".

It follows the solution:

Our algorithm is to extract the loop of the contour features. So when the traverse algorithm sets off from an edge of a starting node, it should return to another edge of the same node. If a path has reached its end and could not meet the aforementioned requirements, backtracking will happen. But during backtracking, the conventional traverse algorithm will simply transfer to another edge of the node where another path is possible.

In our improved algorithm, the visited node that has just been confirmed a node where the backtracking occurs should be labeled "an invalid node", like the node "g" in Fig. 94.3c. When the algorithm backtracks, the visited node that is the closest to the invalid node will be tried if there is any other edge that can be traveled. If no such edge, the node should also be labeled "an invalid node", like the node "f" in Fig. 94.3c, and backtracking continues. Otherwise the other edge of that node will be tried. To fulfill the performance, our modified traverse algorithm will use not only the visit label that has been used in the conventional traverse algorithm, but also a label that indicates if a node is on the desired loop. The latter is denoted by "ValidLoop Node" in later pseudocode of the proposed algorithm.

In our modified algorithm, a dead path means any path that passes through the nodes only once could not reach another edge of the starting node. For example, in Fig. 94.3b, for a starting node "A", when the path is along A-> B-> C-> E-> F-> G, it can no long travel further. Thus, E-> F-> G makes a dead path. Then backtracking occurs, and the node "E", "F", "G" will be labeled "an invalid node" until the path transfers to the other edge of the node "C". Then a complete loop A-> B-> C-> D is formed.

The pseudocode of the modified non-recursive DFS traverse algorithm is listed as below:

10 **Begin initialize:** establish a start node V_S, a start edge E_S, an end edge E_E, a temporary edge $E_{t,}$

a visit label of node: visited[],
a valid label of node: ValidLoopNode[];

20 **While** E_t is not empty **OR** the stack pointer has not pointed to the bottom of the stack
30 **do** E_t connects to a node called V_t,

E_E connects to a node called V_E;

40 **if** $V_t = V_E$ **then** a loop for V_S exists;**endwhile;**
50 **if** visited[V_t] = false **then**
60 **do** visited[V_t] ← true, ValidLoopNode[V_t] ← true;
70 V_t is pushed into a stack, the stack pointer move incrementally,

E_t ← the first edge information of the link in which

V_t stays;

80 **else** $E_t \leftarrow$ another edge information of the link in which E_t stays;
90 **while** E_t is empty **AND** the stack pointer has not pointed to the bottom of the stack;
100 **do** a node V_P is popped from the top of the stack,
look for unvisited adjacent node of V_P,
$E_t \leftarrow$ the related edge information;
110 **if** E_t is empty **then** ValidLoopNode[V_P] $\leftarrow$ false;
120 the stack pointer move decrementally;
130 **Endwhile;**
140 **if** E_t is empty **then** a loop for V_S does not exist;
endwhile;
150 **endwhile.**

Some explanation for the proposed algorithm is given below:

The algorithm must be based on the graph structure that is stored in adjacency list.

The content from line 30 to 80 is the main body of the algorithm, which is similar to the conventional non-recursive DFS traverse algorithm. It indicates that the travel will move forward as long as the current edge is not empty.

The conventional non-recursive DFS traverse algorithm pushes the other unvisited edges into the stack so as to directly visit one of those unvisited edges if forward travel can not continue. While our modified one pushes the currently visited node (NOT edge) into the stack so as to make preparation for judging whether the visited node is valid or not when backtracking occurs. Line 70 shows the key modification.

When backtracking occurs, the visited nodes will be popped out of the stack and another travel route will be tried. So a ValidLoopNode [V_P] that indicates whether a node is on the desired loop is added in our propose algorithm. It will change its sign on some condition on line 110.

94.4 Terminal Presentation For Contour Features

According to our previous study, the contour features can be categorized into three groups. A corner that has only two edges will be denoted by a. A corner that has three or more straight edges and all the edges of which are distributed at one side of the image plane of the corner will be denoted by b as Fig. 94.2 shows. A corner that also has three or more straight edges and all the edges of which are not distributed at one side of the image plane of the corner will be denoted by c.

Therefore, a, b and c will be the terminal notation for the corner features. The loop of the contour features needs a starting node. It is better to select a relatively unique feature as a starting node to avoid ambiguity and repeatability. We find that

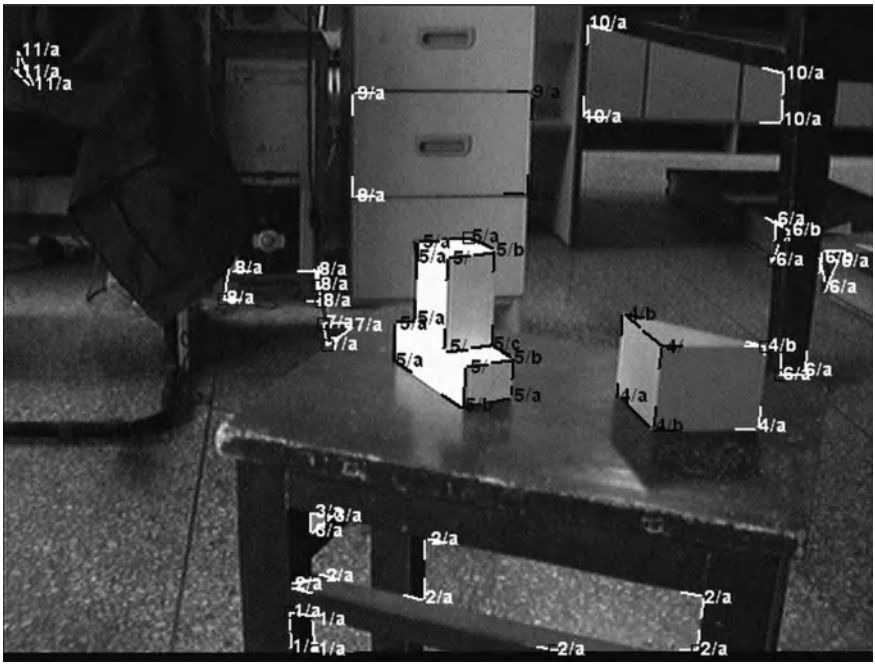


Fig. 94.4 Labeling results of the contour feature type of Fig. 94.1a

c type corner is less common than b type and b type corner is less common than a type, i.e. a type corner is the most common one. Thus, the priority sequence of choosing which type of the corner to be a starting node will be c, b and a.

94.5 Experimental Results

Based on the results of the previous study [2], the experiments are done to extract all possible loops of the contour features and present every loop in one-dimensional language. Figure 94.4 shows the labeling results of the type of the contour features. The numbers denote the label of every connectivity component of the corner features and the letters denote the related type of the corner features. A bar "/" is between the number and the letter. The labels of letters (the type of the corners) are not attached to the internal features to indicate the difference from the contour features.

Table 94.1 shows the representative terminal strings of every loop of Fig. 94.4. The direction of the loop is counter-clock wise.

For the connectivity component labeled 1, 2, 3, 7, 8, 9, 10, 11, the type of the starting node is a. There is more than one starting node for the related loop.

Table 94.1 Terminal strings of every connectivity component of Fig. 94.4a

Label of connectivity component	Terminal string 1	Terminal string 2	Terminal string 3
1	aaaa		
2	aaaaaaa		
3	aaa		
4	babba	babab	bbaba
5	cbaaaaaabab		
6	baa		
7	aaa		
8	aaaaa		
9	aaaa		
10	aaaa		
11	aaa		

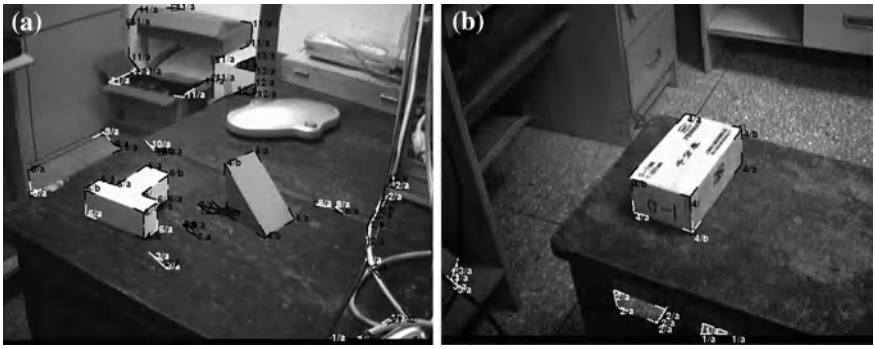


Fig. 94.5 Labeling results of the contour feature type of other images

It implies that there will be more than one string presentation for a loop. However, all of the strings turn out to be the same in linguistic mode. For the future syntactic pattern recognition, they will have the same recognition results. Therefore, only one string is presented in Table 94.1. The same handling will be done on any other similar cases to avoid repeatability.

For the connectivity component labeled 4, there are three strings, i.e. babaa, babab, bbaba. For the connectivity component labeled 5, there is only one string, i.e. cbaaaaaabab.

Figure 94.5 show the labeling results of the type of the contour features for other images. The related strings are also created. We can simply find some terminal strings from observation. For example, in Fig. 94.5a, the connectivity component labeled 6 has three strings: babaaabaaa, baaabaaaba, baaababaaa. Since they are different and the computer can not automatically know which one is the most appropriate, they will all be put into the syntactic recognizer in the future. Thereby, they should be preserved in the present experiments in this paper.

94.6 Conclusions

In this paper, linguistic description for regular-shaped objects is studied. This work is a continuous study after the feature extraction of the regular-shaped objects. One dimensional language is used to describe the contour corner features in view of simplicity and feasibility. We propose some rules for distinguishment between contour features and internal features. A modified non-recursive DFS traverse algorithm is put forward to extract the loop of the contour features for every connectivity component and to exclude the contour features which are not on the desired loop. Experimental results show the proposed algorithm can successfully describe the valid contour features in one-dimensional language. All these make preparation for future syntactic pattern recognition. However, since the previous feature extraction did not give the precise results, the linguistic description can not be precise too. Therefore, fuzzy syntactic recognition based on fuzzy grammar will be proposed to handle the imprecision and uncertainty in our future study. This will be reported in a future paper.

Acknowledgments Thanks for the support by two Applied Research Projects of Nantong City (K2010001 and K2010037).

References

1. Zhang L, Fu Q, Zhang X, Hua G, Zhu L (2010a) Structural feature extraction from regular-shaped rigid objects (1)—based on gradient direction constraint randomized hough transform. IEEE 3rd Int Sympos Comput Intell Des 2:209–214
2. Zhang L, Fu Q, Zhang X, Hua G, Zhu L (2010b) Structural feature extraction from regular-shaped rigid objects (2)—enhanced by prior knowledge. IEEE 3rd Int Sympos Comput Intell Des 2:215–220
3. Henderson TC (1983) Syntactic and structural methods. In: Faugeras OD (ed) Fundamentals in computer vision, pp 270–281
4. Truve S (1990) Image interpretation using multi-relational grammars. In: Proceedings of third international conference on computer vision, pp 146–155
5. Zhang S (1994a) Forest grammar (I). Scientia Sinica Mathemation 24(4):418–423 in Chinese language
6. Zhang S (1994b) Forest grammar (II). Scientia Sinica Mathemation 24(6):659–667 in Chinese language
7. Fang L, Xie L (1997) Relation graph grammar and its application. J Softw 8(2):87–92 in Chinese language
8. Xiaoqin Z, Kang Z, Kong J, Guang-Lei S (2005) Constructing VEGGIE: machine learning for context-sensitive graph grammars. IEEE symposium on visual languages and human-centric computing, pp 272–274
9. Han F, Zhu S-C (2005) Bottom-up/top-down image parsing by attribute graph grammar. IEEE Int Conf Comput Vis 2:1778–1785
10. Siskind JM, Sherman JJ, Pollak I, Harper MP, Bouman CA (2007) Spatial random tree grammars for modeling hierarchal structure in images with regions of arbitrary shape. IEEE Trans Pattern Anal Mach Intell 29(9):1504–1519

11. Kukluk JP, Holder LB, Cook DJ (2007) Inference of node replacement graph grammars. Intell Data Anal Archive 11(4):377–400
12. Han F, Zhu S-C (2009) Bottom-up/top-down image parsing with attribute grammar. IEEE Trans Pattern Anal Mach Intell 31(1):59–73
13. Miller WF, Shaw AC (Nov 1967) A picture calculus, emerging concepts in graphics. University of Illinois, Illinois (in press)
14. Shaw AC (1970) Parsing of graph-representable pictures. J Asso Comput Machin 17(3): 453–481

Chapter 95
Rigidity Recognition of the Cataract Lens Nuclear Based on Multiple Classifier Fusion

Aizhong Mi and Zhanqiang Huo

Abstract In phacoemulsification, controlling the ultrasonic frequency from the emulsification pinhead is essential, which depends on the accurate judgment of the rigidity of the lens nuclear. In this chapter, a multiple classifier fusion method was applied to the rigidity recognition of the cataract lens nuclear, which used different color components to train classifiers and fused the decisions of all the classifiers to judge the rigidity. The experimental results show that the method can achieve a high recognition rate, which meets the demand of the project.

Keywords Multiple classifier systems · Decision fusion · Cataract · Lens nuclear · Rigidity judgment

95.1 Introduction

Cataract patients in our country have a large population base and the number of the patients who need the treatment of phacoemulsification is great. The principle of phacoemulsification is: the emulsification pinhead at the end of the ultrasonic handle oscillates lengthways at the ultrasonic frequency. The turbid lens is emulsified by the cavity effect and machinery function and sucked out. Then, an intraocular lens is implanted, which enables the patient to see. Rigidity is a

A. Mi (✉) · Z. Huo
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454003, Henan, China
e-mail: miaizhong@163.com

Z. Huo
e-mail: hzq@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_95,

physiological characteristic and is closely related to the choice of clinical operation mode. In phacoemulsification, controlling the frequency of ultrasonic from the emulsification pinhead is essential. A low oscillating frequency will not reach the purpose of breaking the lens nuclear, while a high frequency will have an adverse impact on the retina.

The National Key Technological Project of the Eleventh Five Years Plan of China, "development of the efficient, intelligent and minimally invasive cataract treatment system" introduces pattern recognition and image processing technology into the phacoemulsification system, which can accurately recognize the rigidity of the lens nuclear by computer assistant technology and reduce the difficulty and learning period of the surgery.

Classification and recognition is one of the most important fundamental activities of human. Humans always do classification and recognition in daily life, social activities, production, research, study and work. A traditional pattern recognition system is always using a single classifier. However, in recent years, it had been observed in experiments that the samples incorrectly classified by distinct classifiers were generally not the same. It means that different classifiers can potentially offer complementary information about the object to be recognized. Driven by this discovery, Multiple Classifier System (MCS) attracts more and more concerns and has become one of the research hotspots of the pattern recognition field. Effective fusion of the complementary information is expected to considerably improve the performance of the pattern recognition system. By a large number of experiments and applications, scholars have proved that fusing classifiers not only can show better performance than the best single classifier used in isolation, but also can improve the efficiency and robustness of the pattern recognition system. At present, the viewpoint of doing pattern recognition based on MCS has been adopted by many applications areas, such as text analysis (character and handwritten recognition), fingerprint recognition, natural language understanding, medical diagnostics, voice recognition, text mining, remote sensing image recognition and analysis, earthquake prediction, automatically recognition of military objective and so on. Therefore, this chapter uses MCS to recognize the rigidity of the cataract lens nuclear in order to achieve higher precision.

The rest of the chapter is organized as follows: Sect. 95.2 describes the overall process of the cataract recognition system. In Sect. 95.3, the related approaches of the multiple system fusion are introduced. Experiments including the feature extraction, classifier design and experimental results are reported in Sect. 95.4 and conclusions are drawn in Sect. 95.5.

95.2 Overall Process of the Cataract Recognition System

The major task of the cataract recognition system is to judge the part touched by the emulsification pinhead is cataract or normal tissues. For the cataract, the relevant characteristics of the lens are extracted and the rigidity degree of the lens

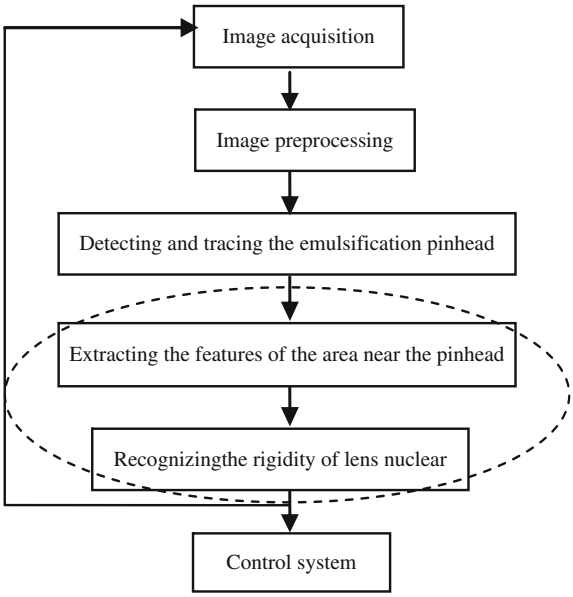


Fig. 95.1 Overall process of the cataract recognition system

nuclear is recognized. According to the different levels of the rigidity, different levels of signals are sent out to control the ultrasonic frequency. The key question in the process is to research the high precision recognition methods to ensure the accurate recognition of the rigidity of the lens nuclear. Figure 95.1 is the overall process of the cataract recognition system.

The system function mainly has the following two aspects:

(1) Detecting and tracing the emulsification pinhead

In the process of the surgery, the position of the emulsification pinhead changes continuously and has some deformation and sheltering in varying degrees under the influence of the light beam and the cataract lens nuclear. Therefore, an effective method to detect, locate and trace the pinhead timely is required.

(2) Recognizing the rigidity of the lens nuclear

The aim of detecting and locating the emulsification pinhead is to study the rigidity of the lens nuclear near the pinhead and sent out ultrasonic at appropriately frequency to crush the nuclear. In the process, the key problem is how to get the characteristics describing the rigidity of the lens nuclear and recognize the rigidity degree of the lens nuclear timely in the surgery.

The main work of this chapter is the design and development of the second system function (see the dash-dotted ellipse in Fig. 95.1).

95.3 Multiple Classifier Fusion

An essential issue in the research of MCS is how to combine all the classifiers' outputs to achieve the final decision. There are generally two types of multiple classifier combination: multiple classifier selection and multiple classifier fusion. Multiple classifier selection assumes that each classifier has expertise in some local regions of the feature space and attempts to find which classifier has the highest local accuracy in the vicinity of an unknown test sample. Then, this classifier is nominated to make the final decision of the system. For this scheme, a method of partitioning the feature space and estimating the performance of each classifier is required.

While in multiple classifier fusion, all classifiers are supposed to be equally "experienced" in the whole feature space and their outputs are fused in a certain way to achieve the final decision. According to the different output forms of the classifiers, multiple classifier fusion can be divided into three categories: when the output is a decision form, the fusion methods include majority voting, weighted majority voting, Naive Bayes, Behavior Knowledge Space (BKS) [1], etc. When the output is a rank form, the fusion methods include the Borda Count [2] and double-objective classifier fusion method [3], etc. When the output is a measurement form, the fusion methods include Max/Min/Mean/Product/Median, Decision Templates (DT) [4], Dempster-Shafer theory [5], Neural Networks fusion [6] and Support Vector Machine (SVM) fusion [7], etc.

The Decision templates (DT) method is a simple and effective multiple classifier fusion rule. The method has the following strong points [8]: roughly, the DT method consists in measuring the dissimilarities between previously learnt templates and a decision profile built from the individual classifier. Another strong point of this method is that it requires no strong assumptions compared to probability-based methods. Furthermore, it is less sensitive to the size of the learning set (overtraining) than other methods. The DT method is also intuitive, so understandable by experts, and is not time-consuming.

However, the DT method has two problems:

(1) The decision template used in the method is just the average of the decision profiles belong to a class and doesn't consider the performance difference of classifiers.
(2) It uses only the classification information of the decision template and doesn't make use of the information of the training samples fully.

Therefore, we had proposed a Self-adjusting weighted DT (SWDT) method in the literature [9]. The method measures the performance of classifiers by the confusion matrix, and self-adaptively assigns weights to classifiers based on their outputs. Reliable outputs are assigned to big weights so that the decision templates are more reliable. For a sample which is prone to be misclassified, besides the similarity between it and the decision templates, the information of the training samples around it are included to make a comprehensive decision. This chapter uses the SWDT method as the fusion rule.

95.4 Experiments

95.4.1 Feature Extraction

To recognize the rigidity of the lens nuclear need select and extract the image features of the objective firstly. In image processing, the common features include color, shape, texture, etc. Color is an important visual property of image and the most widely used visual feature in the image analysis and recognition because color is always very relevant to the objects or scenes in an image. In addition, compared with other visual features, color has less dependence on the size, direction and visual angle of an image and accordingly has stronger robustness. In the light of the specific situations of the lens nuclear recognition, color information is also the most important among the various features. Normal tissues and diseased tissues have different colors which can be directly distinguished by human eyes. Different rigidity of the lens nuclear also lead to the different colors of it.

The rigidity of the lens nuclear is mainly based on the rigidity grade standard for lens nuclear of Emery and Little, which judges the color of the lens nuclear to grade its rigidity according to the examination results of slit lamp.

According to the different research purposes and objects, color feature can be represented and described by different color models. The common color models can fall into two categories: one faces the hardware equipments, such as a color display or a printer. RGB is the common model of this category. The other faces the applications of the color processing. HIS is the common model belongs to this category. This chapter chooses the RGB model.

When we extracted the image features of the area near the emulsification pinhead, every picture was divided into $m*n$ pieces and the color value of a piece was the average of all the pixels in the piece. When the system was implemented, the parameters were selected as $m = 15$, $n = 5$. Thus, every picture was divided into 75 pieces and each piece had the approximate number of the pixels. Then, the average of R, G, B value of all the pixels in a piece was separately computed as a feature value of the piece. That is to say, each piece corresponds to a three-dimensional feature vector and each picture corresponds to a 225-dimensional feature vector.

95.4.2 Classifier Design

Classifier design is a key issue in the pattern recognition system. At present, there are many classification methods have been used in practice, such as the minimum distance method, the linear classifier, the Nearest Neighbor, Artificial Neural Networks and Support Vector Machine, etc. The Nearest Neighbor is a very intuitive classification method. Its basic idea is to find a sample from the training set which has the minimum distance to the objective needs to be recognized and

Table 95.1 The Recognition Rates (RR) of the test methods (in %)

Feature	Method	RR	Method	RR	Method	RR
R	3-NN	85.78	5-NN	85.78	8-NN	84.85
G	3-NN	92.12	5-NN	92.89	8-NN	92.58
B	3-NN	88.87	5-NN	89.03	8-NN	87.79
RGB	3-NN	91.50	5-NN	92.74	8-NN	91.50
	DT	93.05	DT	93.35	DT	93.05
	SWDT	94.82	SWDT	94.13	SWDT	94.67

take the sample's class as the recognition results. The method as a typical representative of non-parametric methods is simple and reliable. Therefore, this chapter uses *k*-Nearest-Neighbors method as the base classifier.

On the training set, a *k*-NN classifier was separately trained using R, G, B feature, i.e., the input of each classifier is a 75-dimensional feature vector. For comparison, a *k*-NN classifier was trained using all color features, i.e., the input of the classifier is a feature vector of 225-dimension. The system tests chose the DT method and the SWDT method as the fusion rules.

95.4.3 Experimental Results

The experimental images were captured from the surgery videos provided by Beijing Tongren Hospital. The image size has three types: 20(pix) ×20(pix), 30(pix) × 30(pix) and 40(pix) × 40(pix).

The size of training image set is:

649 = 95 (normal tissue) + 293(I level nuclear) + 184(II level nuclear) + 37(III level nuclear) +28(IV level nuclear) + 12(V level nuclear).

The size of test image set is:

647 = 91(normal tissue) + 296(I level nuclear) + 183(II level nuclear) + 33(III level nuclear) + 30(IV level nuclear) + 14(V level nuclear).

The software environment of the experiments is Window XP and Microsoft Visual C ++6.0. When the parameter $k = 3,5,8$, the experimental results are shown in Table 95.1.

It can be seen from Table 95.1 that the classifiers using the feature G have the highest recognition rate among the single classifiers, which are slightly higher than the classifiers using all the features. While the classifiers using the feature R, B have so much lower recognition rate. Fusing the decisions of three classifiers using single feature, the recognition rate of DT method is above all the single classifiers and lower than that of the SWDT, which shows the validity of the SWDT method again. When the parameter *k* takes different values, the results change accordingly. But the size-relationship is similar, that is to say, the conclusions have universality in a sense.

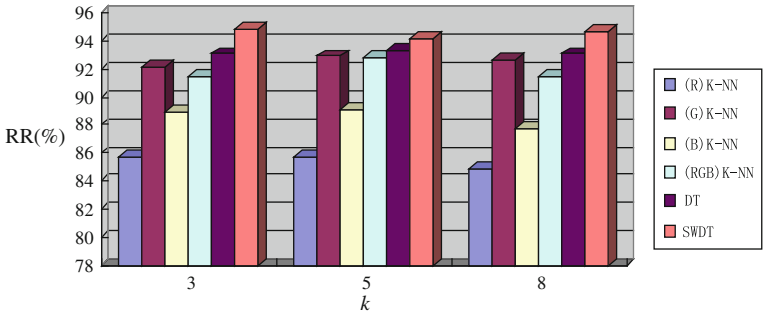


Fig. 95.2 The Recognition Rates (RR) comparison of the test methods

For comparing distinctly, the experimental results are shown in Fig. 95.2, where the conditions can be seen more clearly than that in Table 95.1.

95.5 Conclusions

This chapter applied multiple classifier fusion method to the rigidity recognition of the cataract lens nuclear, which used different color components to train classifiers and fused the decisions of all the classifiers by the SWDT method to judge the rigidity. The experimental results show that the SWDT method can achieve a higher recognition rate than the DT method and meet the basic recognition rate demand of the project, which shows the validity of the SWDT method again.

Acknowledgments This chapter is supported by the Basic and Leading Research Project of Henan Province (No. 092300410216), the Foundation Project for Youth Key Teacher of Henan Province (No. 2009GGJS-035) and the Doctoral Fund of Henan Polytechnic University (No. B2010-46).

References

1. Raudys Š, Roli F (2003) The behavior knowledge space fusion method: analysis of generalization error and strategies for performance improvement. Lecture Notes in Computer Science, vol 2709. Springer, Berlin, pp 55–64
2. Jie Z, Qiang J, Nagy G (2007) A comparative study of local matching approach for face recognition. IEEE Trans on Image Process 10:2617–2628
3. Ming L, Baozong Y, Zhenjiang M (2007) A double-objective rank level classifier fusion method. Acta Automatica Sinica 12:1276–1282
4. Kuncheva LI, Bezdek JC, Duin RPW (2001) Decision templates for multiple classifier fusion: an experimental comparison. Pattern Recogn 2:299–314
5. Denoeux T (2000) A neural network classifier based on Dempster-Shafer theory. IEEE Trans Syst Man Cybern Part A: Syst Humans 2:131–150

6. Tremblay C, Valin P (2002) Experiments on individual classifiers and on fusion of a set of classifiers. Proc 5th Int Conf Info Fusion 8–11:272–277
7. Chou CH, Lin CC, Liu YH et al (2006) A prototype classification method and its use in a hybrid solution for multi-class pattern recognition. Pattern Recogn 4:624–634
8. Valet L, Ramasso E, Teyssier S (2008) Quality evaluation of insulating parts by fusion of classifiers issued from tomographic images. Info Fusion 2:211–222
9. Aizhong M, Hongwei H, Xuefeng Z, Xuyan T (2009) A method of multiple classifier fusion with self-adjusting weights. Chin J Electr 11:2604–2609

Chapter 96
The Establishment of Sensor Virtual Laboratory Based on VR-Platform

Dahu Wang, Huan Liu and Hui He

Abstract This system comprehensively utilizes 3Ds max modeling tools and VR-platform which is 3D virtual reality interactive Platform. We develop a sensor virtual laboratory system which include teacher's interpretation, animation demo, independent operation, knowledge consolidating and test. The system represents 3d experimental apparatus, realizes the display of experimental instrument details, experiment process and the simulation of experimental phenomena by using VRP scripting language. Under the condition of collision detection, the system realizes the roam to the laboratory scene. The actual application indicates that the system provides a good solution for experimental teaching when some experimental conditions are restricted.

Keywords 3Ds max · VR-platform · Virtual reality · Sensor virtual laboratory

96.1 Preface

Experimental teaching as the important part of the teaching in universities plays an irreplaceable role to cultivate students' manipulative ability, ability to analyze the problem and ability to solve problems, the correct method of thinking and the

D. Wang (✉) · H. Liu · H. He
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: dahuwang@126.com

H. Liu
e-mail: liuhuan8802@126.com

H. He
e-mail: herunhaocili@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_96,

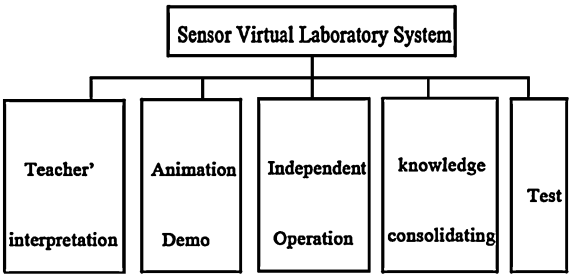


Fig. 96.1 System structure

rigorous work style. However, with the increasing number of college students, the traditional experimental teaching means are affected by teacher resources, venues, instruments and less equipment or teaching quantity is affected [1].

At present, the virtual experimental teaching systems mostly are demo systems, made by Java or Flash. These soft wares lack interactivity between people and system, so the experiment lacks authenticity and vitality [2]. To solve this problem, we must combine 3D visualization technology into the virtual experimental system. By running script language in client we make sure that the system outputs life-like 3D experiment scenes.

This sensor virtual laboratory system, as the basis of undergraduate experimental course of detection and automation technology, puts theoretical teaching and practical teaching together and solves a series of shortages during traditional teaching process due to the limits of objective conditions (instrument insufficiency, the risk of practical operation of experiment and a series of other difficulties). Virtual sensor laboratory realistically shows the dynamic characteristics of the experiment, its intelligent human–machine interactive way greatly enhances the learning interests of the students; its functions of independent experimental operation, knowledge summary and testing play an important role for students to deepen knowledge, understanding and grasping, and improve the student's practical ability, analyzing ability and the ability of solving problems [3].

96.2 System Analyses

In order to make this system more useful for students' learning, we divide it into five modules, and they are teacher's interpretation, animation demo, independent operation, knowledge consolidating and test, as shown in Fig. 96.1:

Teacher's interpretation and animation demo parts achieve the advance guidance for student's independent experimental operation by loading lectures video; Animation demo part, combining with the 3D models produced by 3Ds max, is produced by adjusting the key frames and call action library, and it has good dynamic characteristic. Independent operation namely intelligent human–machine interaction, let the students feel that they exist in the real laboratory personally, so increases students' interest to experimental courses. In the system, students can

design some simple experiments independently. Knowledge consolidate part is used to help students deepen understanding and mastering of theoretical knowledge. Testing system is a summary to this experiment course's technical operation and difficulty knowledge, and can also test students' insufficieny through this system, so that most students can self-study better.

96.3 The Key Technology of System Implementation

96.3.1 Entity Modeling Technology

(1) The drawing of models
This system uses polygon method of 3Ds max8 to model the CSY2001/B sensor system comprehensive test bench, temperature sensor experiments module and other equipments.
(2) The optimization of scene model
Under the condition of the experimental models, reducing the number of scene model's strips can not only make sure that the operation of virtual system is more fluent, but also can effectively reduce the requirements of the scene to computer memory and physical memory.
(3) The processing method to map models
Because of faces which are used too much in the scene, taking more memory resources, the models are displayed by maps in the process of the scene building, such as the upper part of the heating furnace of thermocouple experimental modules, there are too much broken surfaces as it is created by the method of Boolean, thus it costs a significant amount of memory resources.

We use Boolean operation to create entity model, and use the render technology of 3Ds max render images with.png or.jpg styles, then use Adobe Photoshop to delete the spare parts. The name of map cannot be named in Chinese, otherwise when we render the images using network rendering method in late stage, it will have problems. At the same time, the size of the texture also has strict requirements, we should ensure map dimensions are 2n, and make the map size as small as possible while guarantee map is clear enough.This can not only save computer memory, but also can optimize speed of the system.

96.3.2 Technology of 3D Animation

Technology of 3D animation is mainly applied in animation demo part, it is the important part of virtual laboratory, it can make the system be more vivid and can detailedly introduce the function of each experiment instrument, and the key content of the experiment.

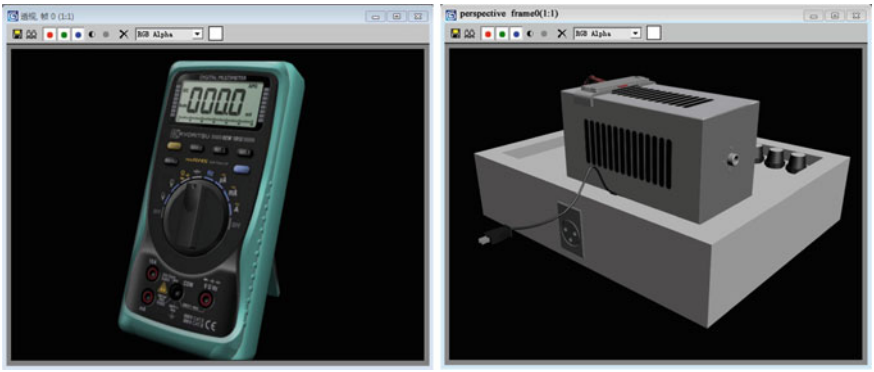


Fig. 96.2 Experimental equipment model for example

This technique mainly makes use of the animation function in 3Ds max8. After choosing good materials, playing good lighting and adjusting the key frames, we output images with.tag format, then we put the pictures sequences synthesis for film by using Adobe Premiere cs3, eventually payload to VRP.

96.3.3 The Establishment Technology of VR-Platform System

VR-Platform, a 3D interactive simulation platform developed by Zhong Shi Dian digital technology Company independently which is software of 3D virtual reality platform with fully independent intellectual property rights, this software has high applicability, simple operation, powerful, high visualization and its appearance will inject new vitality into the developing VR industry.

We use script language of VRP software to realize the browsing and interactive function of 3D scenes. It is both convenient and simple, and can realize a good effect.

96.4 System Constructions

96.4.1 Model Production

By using 3Ds max8, we model the modeling of CSY2001/B sensor system comprehensive test bench, temperature sensor experiments module and other equipments, partial models are shown in Fig. 96.2:

96.4.2 Model Baking and Importing

After finishing maps of models and decorating lamplight of the scene, and because the model area of laboratory equipment is lesser and some maps need to be tiled, we bake them using complete map way.

After being baked, its light and shade information on smooth surface of model will be stored in the file in the form of corresponding textures, after we import the scene into VR-platform, the whole scene will appear the characteristics of trenchant details, regular light and shade.

After rendering the images, click [*VRPlatform*] button, then in the open rollout, click the export button, after having confirmed, click append the VRP editor button, and we can successfully import 3D models into VRP editor.

96.4.3 Design of System Interface

Based on management system of client and server mode, the function module and functional interface are made firstly in drawing district of VR-platform; then we design images of users' interface in Photoshop and then cut figures.

The interfaces of the system are shown in Figs. 96.3 and 96.4:

96.4.4 The Realization of Interaction

96.4.4.1 The Adding Method of Background Music

Experimental teaching in university occupies more and more proportion, in order to add bit vigor for virtual laboratory system, so as to enhance the students' enthusiasm to experimental courses; we use VRP script language to add some soft background music. By using the following script files to realize this function:

Firstly, establish an initial function in system file:
Set the state value of object, music broadcast, 0
Then, add the following scripts in trigger function:
Switching the state values of object, music broadcast
compares the state values of the objects, music broadcast, 1
Play music, E:\zzz\ papers dubbing\ windancer.mp3, 0, 0, 1
Set the volume, 0, 5
Otherwise
Stop music, 0

Such, we can achieve the playing and closing of music through a button, and can also control the volume.

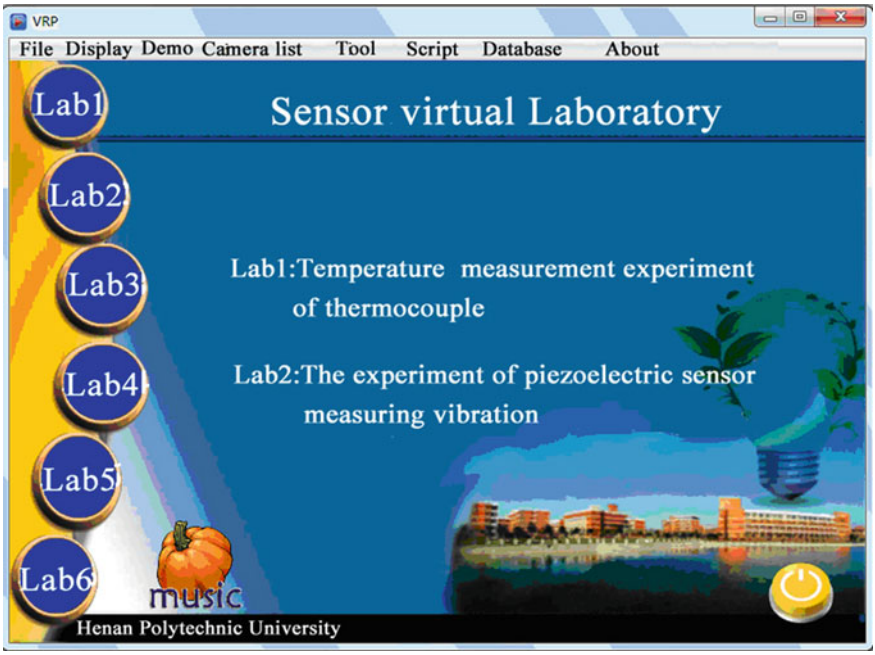


Fig. 96.3 Main interface

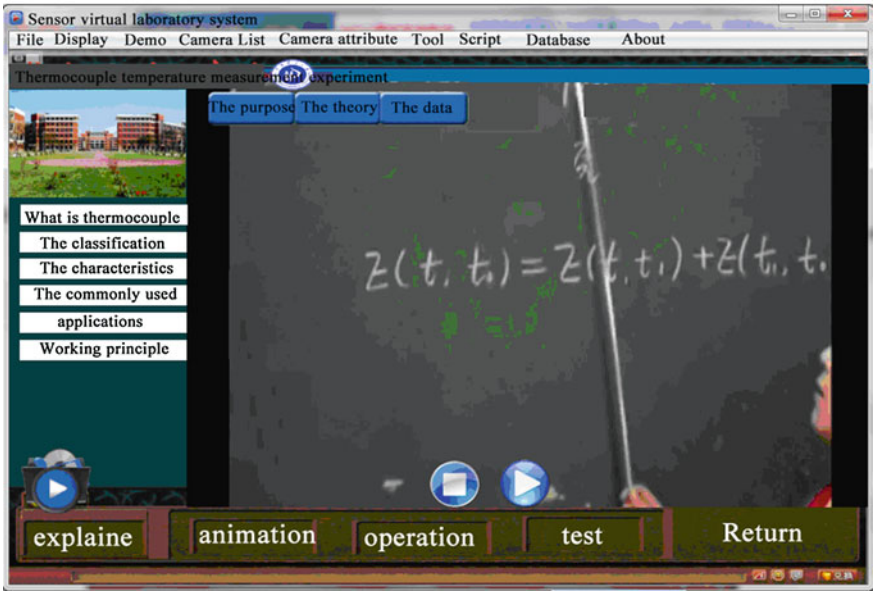


Fig. 96.4 Loading window of teacher's interpretation video

96.4.4.2 Load of Animation Video

In 3Ds Max files, we use the method of adjusting key frames and loading action library makes a small animation about the processing of experiment, after rendering, we import sequence frame images into Adobe Premiere CS3 to synthesize for .avi files.

We load.avi files which have been compounded into the sensor virtual laboratory system based on VRP. Firstly establish a 3D screen model in 3Ds max8, we can use cuboids instead, then we import it into VRP after rendering, and finally we realize video loading by using the following script commands:

Firstly, establish an initialization function in system file:
Set the state value of object, video broadcast, 0
Then, add the following scripts in trigger function:
Switching the state values of object, video broadcast
compares the state values of the object, video broadcast, 1
Load video, E:\zzz\ papers video\ thermocouple experiment..avi,0
Apply video to model, video broadcast, 1
Control to video broadcast, video broadcast, 1
Show hidden objects, 1, video broadcast, 1
Otherwise
Control to video broadcast, video broadcast,0
Show hidden objects, 1, video broadcast, 0

We can achieve the button control to the broadcasting and closing of the video, and can achieve the controlling to pause and broadcast operation by using the upper script language. Video loading window is shown in Fig. 96.4:

96.4.4.3 Production of Test System

Testing system is a simple system which achieves the control to static text and single marquee by using script languages, it can remember students' answers, and display test scores, and finally shows the correct answer. Its characteristic is every problem has only one chance to choose, after being chosen, the question becomes unavailable state, and record scores, after answering the last question, the system will automatically display 'displaying the correct answer' button, after clicking the button, the right answer can display, click 'start again' button can start again to answer. Test system as Fig. 96.5 shows:

96.4.4.4 The Design of Independently Operating System

Independently operating system is a student to 3D experimental models interactive system. Students can operate 3D equipment freely, and students in virtual laboratory like placing themselves in real laboratory, they can operate any of the

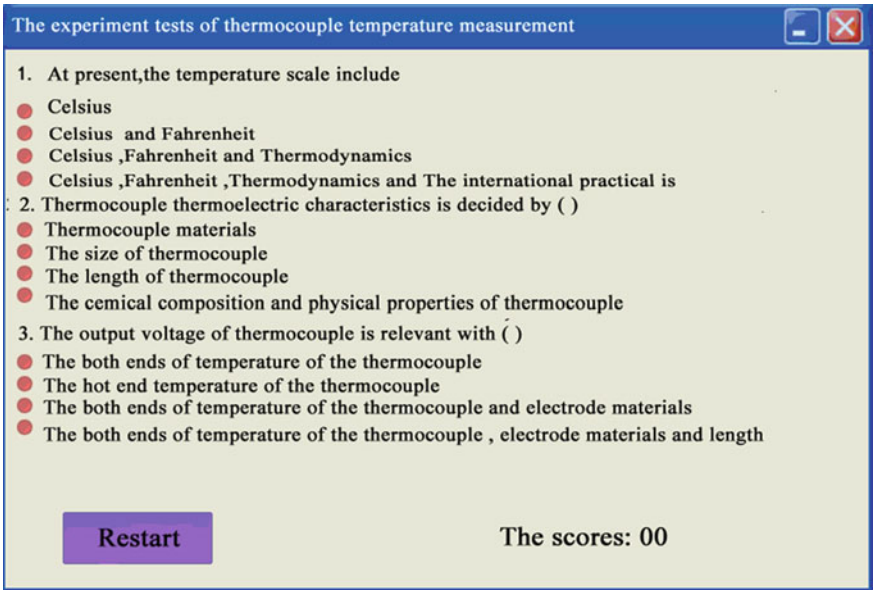


Fig. 96.5 Interface of test system for thermocouple thermometer experiment

experimental models independently, design experimental steps freely and can record their own process of experiment. So this system can facilitate students to analyze their operation, and it is useful to find their mistakes.

Each step of students' operation can immediately be observed, when the experiment is completed, students can keep their experimental results, and it is conducive to facilitate the completion of experiment report.

96.5 Conclusions

The establishment of the sensor virtual laboratory can improve the students' autonomous learning ability, and it can reduce teachers' burden, the more important factor is that it can solve problems of experimental teaching caused by restricting about experimental conditions, although virtual experiment just an auxiliary means for experimental teaching, and cannot completely replace the actual experimental operation, but virtual laboratory injects new vitality for experimental teaching, let experimental teaching extend to a broader space [1], and dynamic 3D scene can more stimulate the fun of student to experiment courses.

References

1. Qu Y (2010) Design and research for virtual laboratory based on the campus network. Mod Education Technol 5:23–27
2. Liang B, Xu D, Mo L (2003) The Research based on VRML for online 3D virtual laboratory. Comput Age 11:167–171
3. Zhang Y, Tan T, Zhao X, Guo Z (2010) The building design for virtual chemical experiment platform based on OSG. Comput Eng Des 31(12):87–91

Part VIII
System Analysis and Decision

Chapter 97
Modeling for Interference Coupling Path of Data Acquisition System of Microprocessor Protection Device

Wei Zhang and Jiaolong Zhang

Abstract The electromagnetic compatibility of the data acquisition system for microprocessor-based protection device was studied and the main work was modeling and simulation for interference coupling path under the electrical fast transient pulse interference signal. Firstly, the input and output data of the interference transmission channel was obtained through the standard electromagnetic compatibility test experiment. Secondly, the data was processed mainly by the identification and the vector matching, and then got the state-space model. Lastly, the simulation model was established by the equivalent circuit in Pspice, which was verified correct through the comparison between simulation and experimental results. The method to establish interference channel model and obtain conclusions have important practical significance to improve the electromagnetic compatibility of microprocessor-based protection.

Keywords Data acquisition system · Interference channel · Vector matching · State-space model

97.1 Introduction

Microprocessor-based protection devices are widely used in substations and power plants where electromagnetic environments are extremely harsh because of high intensity and high density. With the applications of the electronic technology,

W. Zhang (✉) · J. Zhang
School of Electrical Engineering and Automatization,
Henan Polytechnic University, 454000 Jiaozuo, China
e-mail: zwei1563@hpu.edu.cn

J. Zhang
e-mail: zjl@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_97,

computer technology, automation and communication technology applications into the power system in-depth, the microcomputer protection is developing in the direction of control, measurement and integration of intelligent communication, so that it is more sensitive and vulnerable to the electromagnetic interference. In recent years the growing concern is put on the electromagnetic compatibility of the protection device [1].

The hardware circuits of protection devices are generally composed of the power supply circuit, the AC signal acquisition circuit, the CPU and other peripheral circuits. Because the AC signal acquisition circuit is directly connected to the transformer of primary system the acquisition circuit is most vulnerable to transient interference. The access is filtered before analog–digital conversion, and the actual circuit impedance is relevant to the frequency, which is subject to parasitic effects. This kind of transient interference can make the protection device relay malfunction, resist the action and trigger accident [2], so the acquisition circuit must take effective measures to suppress transient interference. Suppressing electromagnetic interference mainly have two ways: one is to reduce the interference of high frequency electromagnetic energy emission source (i.e, radiated interference), and the other is to change the propagation of electromagnetic interference channel. Most of the existing measures use specific inhibition or interference circuit components, and there is no systematic approach, main reason is that the propagation characteristics of electromagnetic interference is not enough clear and comprehensive. Therefore, systematic study of the electromagnetic propagation characteristics has very important significance both interference suppression and interference prediction. In this chapter the data acquisition system that is most vulnerable to electrical fast transient interference pulse in the protection device [3, 4] was studied and the main work is to establish the interference channel model for simulation. The methods that the channel interference model is established and conclusions of the mathematical model can be used to guide the system electromagnetic sensitivity analysis and optimization of filter design.

97.2 Interference Channel of the Data Acquisition System

97.2.1 Composition of the Data Acquisition System

The data acquisition system is mainly composed of the form circuit of the voltage and current, analog filter and A/D conversion module. The voltage formation circuit includes the auxiliary current transformer and voltage converter, which transforms the secondary current and voltage output through the current transformer substation (TA) and voltage transformer (TV) into a weak signal (± 5 or ± 10 V) that a computer can recognize. The analog filtering unit maybe of active filter or passive filter. The A/D conversion unit has two kinds: one is VFC based on the relation of voltage and frequency convertor and the other is successive approximation type. As the VFC has the characters of strong

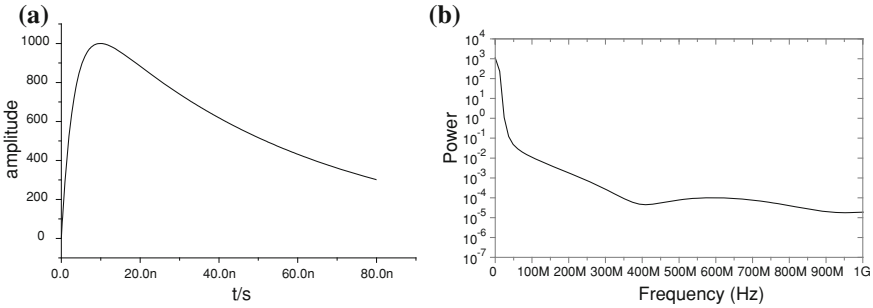


Fig. 97.1 Waveform and power spectrum of a single EFT pulse **a**. Waveform of a single EFT **b**. Power spectrum of a single EFT

anti-interference and simple interface with the CPU, it has been widely used in the field of computer protection in China. Because of the coupling capacitor between transformer primary and secondary it provides a channel for the interference signal and then the interference signals can directly enter into the device inner, which could make the device's operational amplifiers, AD p etc. work abnormally.

97.2.2 Characteristics of Electrical Fast Transient Disturbance

The IEC has developed a standard IEC 61000-4-4 (GB/T 17626) "electrical fast transient burst immunity test" [5] to simulate the impact of the electrical fast burst on the electrical and electronic equipment to test the product antiinterference.

A single pulse waveform is a double exponential decay pulse, which can be expressed as

$$u(t) = A_D\left(e^{-t/\tau_d} - e^{-t/\tau_r}\right)$$

where A_D for the constant of proportionality, τ_r and τ_d for the pulse rise time and duration of the relevant time, respectively. The fast transient burst test apparatus is designed according to the mathematical expression. The waveform of a single EFT is shown in Fig. 97.1a. Because the rising of a single pulse burst is 5 ns and the pulse width to 50 ns, it makes the burst harassment extremely rich in harmonic content and should be called broad-band disturbance signal [6]. Its power spectrum of the single pulse is shown in Fig. 97.1b.

97.3 Model for the Interference Channel

97.3.1 Parameter Identification

The studied interference channel mainly includes the transformer primary, conditioning circuit and AD converter circuit, which can be seen as a linear time-invariant system, and the relationship of input and output can be expressed as:

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{j=1}^{n} b_j u(k-j) + e(t)$$

where a_i and b_i for the system parameters, n for the order of the system and *e(t)* for zero-mean white noise.

The prediction model can be expressed as:

$$\hat{y}(k) = \sum_{i=1}^{n} \hat{a}_i y(k-i) + \sum_{j=1}^{n} \hat{b}_j u(k-j) \tag{97.1}$$

And it can be rewritten into matrix form

$$Y = XP \tag{97.2}$$

where P for the system parameters estimated matrix

$$P = \left[\hat{a}_1\, \hat{a}_2 \ldots \hat{a}_n\, \hat{b}_1\, \hat{b}_2 \ldots \hat{b}_n\right]^T$$

X for the data matrix consisting of data collection

$$X = \begin{bmatrix} y(n) & y(n-1) & \ldots & y(1) & u(n) & u(n-1) & \ldots & u(1) \\ y(n-1) & y(n) & \ldots & y(2) & u(n+1) & u(n) & \ldots & u(2) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ y(n+N-1) & y(n+N-2) & \ldots & Y(N) & u(n+N-1) & u(n+N-2) & \ldots & u(N) \end{bmatrix}$$

In order to avoid the inverse operation of matrix and iterative optimization conclusions parameters the steady-state Kalman filter method is adopted to identify parameters matrix P. Let X_i^T be the ith row vector of the matrix X, y_i be the ith component of Y.

The recursive algorithm is as follows:

$$\begin{aligned} P_{i+1} &= P_i + \frac{S_{i+1} x_{i+1}^T \left(y_{i+1} - x_{i+1}^T P_i\right)}{1 + x_{i+1} S_i x_{i+1}} \\ S_{i+1} &= S_i + \frac{S_{i+1} x_{i+1}^T x_{i+1} S_i}{1 + x_{i+1} S_i x_{i+1}^T}, \quad i = 0, 1, \ldots, N-1 \end{aligned} \tag{97.3}$$

The initial conditions are set: $P_0 = 0$, $S_0 = \alpha I$, α should be taken a real number greater than 10,000, I for a unit matrix of $(r + 1)c \times (r + 1)c$. So the parameter P can be got by the proposed method.

97.3.2 State-Space Model

The behavior of the nonlinear devices can be described in the time domain and expressed by the differential equation generally, but the state-space model of a system can analyze and process problem more conveniently and effectively. Through the link between the transfer function and the state-space model, we get the state-space model by using the vector matching method [7, 8].

The transfer function of interference channel can be obtained from the corresponding differential equation, it is expressed as:

$$G(s) = \frac{b_{n-1}s^{n-1} + \cdots + b_1 s + b_0}{S^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}$$

Convert it into the form of a pole–residue

$$G(s) = \sum_{i=1}^{n} \frac{k_i}{s - p_i} + d$$

The state-space model can be obtained by using Jordan-type [9].

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} A_r & 0 & 0 \\ 0 & A_c & 0 \\ 0 & 0 & A_c^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} B_r \\ B_c \\ B_c^* \end{bmatrix} u$$

$$y = \begin{bmatrix} C_r & C_c & C_c^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + Du$$

Since the complex poles have no physical meaning in the time domain and cannot correspond with the physical components, we need to transform by the transformation matrix T.

$$\bar{x} = Tx,\ T = \begin{bmatrix} I & 0 & 0 \\ 0 & I & I \\ 0 & jI & jI \end{bmatrix}$$

The final state-space model can be expressed as:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} A_r & 0 & 0 \\ 0 & \mathrm{Re}(A) & \mathrm{Im}(A) \\ 0 & -\mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} + \begin{bmatrix} B_r \\ 2\mathrm{Re}(B) \\ -2\mathrm{Re}(B) \end{bmatrix} u \tag{97.4}$$

$$y = [\, C_r \quad \mathrm{Re}(C) \quad \mathrm{Im}(C) \,] \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} + Du \tag{97.5}$$

97.3.3 Equivalent Circuit Theory

The state-space model will be converted into corresponding spice equivalent circuit form. Now we give the principle description by the example that one system has a port and two state variables. Their state-space model can be expressed as:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

$$y = [\, c_{11} \quad c_{12} \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + du$$

From the knowledge of the state equation in circuit theory, these state variables and other elements can be replaced by the capacitor voltage, conductance, linear controlled sources and etc., so we get the equivalent circuit corresponding to the state-space model. Then, using the arrangement regularity of the matrix elements it be transformed into a Spice input net-list file for circuit simulation.

97.4 Experiment and Simulation

According to the EFT/B test standards for experimental layout, the channel input port is put on the pulse interference with the 2 kV peak and 2.5 KHZ pulse repetition frequency. The SKS-0404T EFT/B generator gives the interference signal and the output signal is measured at A/D converter output port. Because the voltage amplitude of the EFT signal is much higher, it cannot be directly connected to the oscilloscope, and thecollected voltage signal needs to be attenuated with Tektronix P6015A passive high voltage probe (attenuation ratio 1000:1). At the same time the EFT belongs to wide spectrum, an ordinary oscilloscope cannot meet the requirements. Yokogawa DL7480, a digital storage oscilloscope (maximum sampling rate 2 GSa/s, analog bandwidth 500 M) is chosen in this experiment. So we can get the output waveforms of the interference channel and can get data records.

After building the simulation models by the proposed approach we get the simulation output curve of the interference channel. The compare curve of the measured and simulation output waveform is shown in Fig. 97.2, where it can be seen that there is little error between them and proves that the method and conclusion are correct.

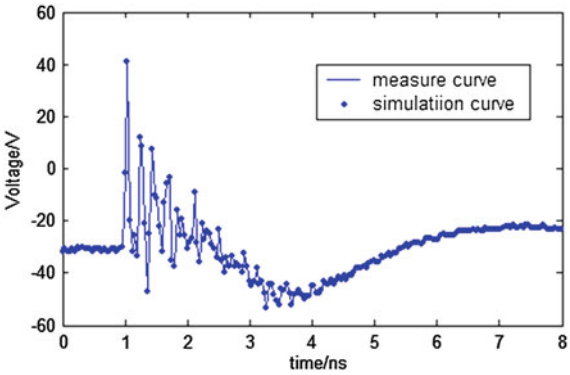


Fig. 97.2 Curves of measured output and simulation output

97.5 Conclusion

In this chapter the model for interference channel of the data acquisition system is built under the electrical fast transient interference pulse in microprocessor-based protection through the systematic method. After acquiring the input and output data both the system identification and the mathematical conversion are used to get the final state-space model, and then the equivalent circuit is exported by using Spice software. The method for modeling and some conclusion have more significance to the electromagnetic compatibility of the computer-protection devices and other system-level anti-interference design.

References

1. Su H, Gao L, Duan X, Yun B (2009) Research on the EMC of microcomputer-based relay protection. Power Syst Prot Control 37(17):97–101
2. Cheng L, Li Y, JIAO S (2004) Influence of EFT on data acquisition system of numerical protection and its countermeasure. Electr Power Autom Equip 24(1):14–19
3. Niu B, Song Z, Wang J, Geng Y, Wang J (2007) Study on electrical fast transient burst immunity of digital relay. High Volt Apparatus 43(1):14–17
4. Wang J, Liu J (2005) Research on electrical fast transient immunity for relay protection equipment. High Volt Eng 31(10):36–38
5. IEC 6100-4-4 (2004) Electromagnetic compatibility Part 4: testing and measurement technique section 4:electrical fast transient/burst immunity test (2004)
6. Zhang W, Han S (2010) Analysis of transient interference signal on switch operating system. IEEE transactions on 2010 international conference on computer, mechatronics, control and electronic engineering, vol 4, pp 143–146
7. Liu E-X, Li E, Li L-W (2005) Finite-difference time-domain macro-model for simulation of electromagnetic interference at high-speed interconnects magnetics. IEEE Trans 41(1):65–71
8. Qin X, Huang S (2006) Model and Matlab simulation of the electrical fast transient/burst disturbance in protection end control secondary circuit. Relay 34(4):17–21
9. Gustavsan B, Semlyen A (1998) Application of vector fitting to state equation representation of transformers for simulation of electromagnetic transients. IEEE Trans. Power Deliv 13(3):834–842

Chapter 98
Study of the Web Application System Requirements Analysis UML Modeling

Hong-mei Sun and Rui-sheng Jia

Abstract In order to more accurately capture and describe the web application system's requirements, to improve the software product quality and its efficiency of development, the web application system requirements analysis unified modeling language (UML) modeling was studied based on object-oriented analysis method. The UML modeling mechanism analyzed the key points of the UML modeling and process in detail, adopted an online graduation design management system as an example, given the use cases model, the static model and dynamic model results of the system requirements analysis phase. This example application accumulated the UML modeling experience of web application system requirements analysis, enriched the content of the web application research and has certain directive significance to the web application system requirements analysis UML modeling.

Keywords UML modeling · Web application · Requirements analysis · Software development

98.1 Introduction

With the rapid development of internet, the web application system is widely used because of its advantage, and more and more of the web application system software development demand was raised. In order to meet market demands,

H. Sun (✉) · R. Jia (✉)
College of Information Science and Engineering,
Shandong University of Science and Technology,
Qingdao, China
e-mail: shm0221@163.com

R. Jia
e-mail: jrs716@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_98,

improve the efficiency of software development and avoid blind development that causes unnecessary losses, it is necessary to do an analysis modeling for the web application system before the work of the software coding. Unified modeling language (UML) is a powerful standard modeling language whose application range is very wide, so it can be combined with the object-oriented requirement analysis's general method to do the system requirements analysis modeling for the web application. This paper firstly outlines the UML modeling mechanism, and then puts it into a practice of graduation design of online management system modeling. Combining UML modeling technology with the web engineering establishes the UML use cases diagram that represents online graduation design management system's use cases model. The UML class diagram describes the static structure model, and the UML sequence diagram or activities diagram describes the cases' realization process of the dynamic behavior model. The model can be helpful to think of problems, can help developers to control the complexity of the system, can greatly shorten the system development cycle, improve software reusability and the later maintenance and upgrade is also very convenient [1]. For the software system, especially the complex practical web application system, modeling research has very important practical significance and application value.

98.2 UML Modeling Mechanism Overview

UML is a powerful, universal standard modeling language that is established by the object management group (OMG); therefore it can be used on any system modeling that has the static structure and dynamic behavior [2]. Using UML and object-oriented requirement analysis method to establish web application system analysis model has three kinds, namely the use of case model, static model and dynamic model, as shown in Fig. 98.1.

Use Cases Models. A use case is an example of how to use a function of the system. From the demand perspective, a use case is an interaction between the user and computer caused by some external events. The UML use cases model is composed of a plurality of use case diagram. Use cases diagram is the foundation of all models, it indicates the functional requirement of the system and describes how the system can be used.

Static Models. Static model is used to describe the system's static structure, and class diagram may be used to signify system's static model in UML. When modeling, classes can often be selected and linked to a function which can then find out the relationship between classes, and after that establishes corresponding class diagram.

Dynamic models. Dynamic model is used to describe a system's dynamic behavior, and shows the object's interaction in different time during the running of the system. Usually use the UML sequence diagrams, activity diagrams and the state diagrams. to describe the dynamic model of the system [3].

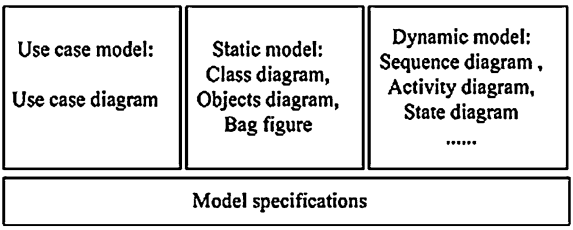


Fig. 98.1 UML model, combining with object-oriented methods of system analysis, also applied to web application system

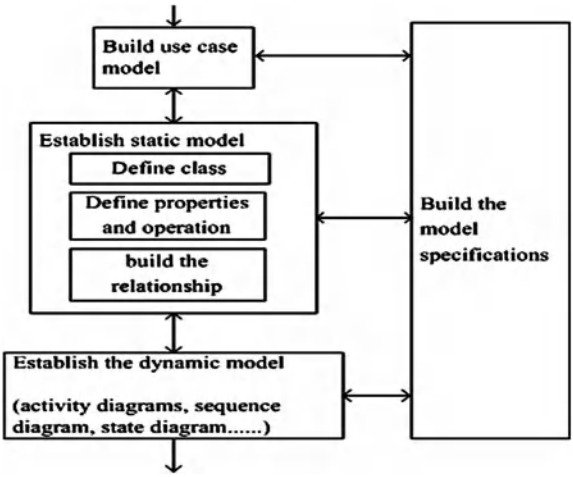


Fig. 98.2 UML modeling process, the three modeling links are complementary to each other, and so-called stage division is relative

From the application perspective, using UML modeling analysis on web system generally includes three main steps, first to carry on use case modeling, then static modeling and finally dynamic modeling. The whole process of modeling is shown in Fig. 98.2. Certainly, if in the process of modeling some models have something not right in some phase, they can go back to the related stage to supplement and perfectly complete the related model, the whole process of modeling can be allowed to repeate and iterate.

98.3 Application Example

In order to fully understand the UML modeling mechanism, below, combining with an online graduation design management system, analyzes how to complete the system's use cases modeling, the static modeling and dynamic modeling, Given all kinds of model described with UML use cases diagram, the class diagram and the sequence diagram, it analyzes the main process of the modeling and its key points, experiences modeling mechanism, summarizes the experience of an application example, enriches the research content of web application

system, analyzes modeling based on UML and promotes the development of the related research work.

98.3.1 Use Cases Modeling

The use cases model is composed of use case diagram. The basic description operators of use cases diagram have system, performer, use case and the arrow lines symbol that express the relations between them. Among them, the solid line shows the boundary of the system, outside the box means outside the system, in the box means in the system. Performers who are used to drive cases are represented with a human form mark symbols, a performer is a external person, and can also be another system. A use case is represented with an elliptic, each use case represents a function that the system provided. All kinds of relations between use cases are represented by different arrow lines. When building the use case model, first identify the system's users and related external systems, and determine the actors, and then establish the system' s use cases model according to the system's function [4].

Use the online graduation design management system as an example, its requirement description is as follows: teacher logins in his interface by using client browser, then can manage the graduation design topics, including the following operations such as add, delete, update and query, can manage students' score, can browse the results submitted by students, can browse announcement, can view other public information and can change a part of his personal information. Student user logins in his interface through client browser, then can select a topic to do from the graduation design topics, can submit his design result when his work is done, can view his score, can browse the announcement and other public information and can modify his personal information partly. After an administrator user logins in his management interface, he can manage not only the base information of teachers but also the students' and his information, can manage announcements and can view all kinds of public information. All kinds of information that built by all users through such operations as add, modify, and delete would be saved into database.

According to the description of the problem above, three kinds of actors can be identified who interact with the system they are teachers, students and administrators. They are the persons who use the system outside the system, can be determined as system's performers. Analyze each kind of actor's application scene of the system, can draw system's use case diagram as shown in Fig. 98.3. Only from the use cases diagram cannot see all the information that can be included in use cases, need further through the use cases description to describe those information who cannot reflect in graphics. Use cases description is actually the specification about how performers interact with system, can describe in words and can also use activity diagrams to describe.

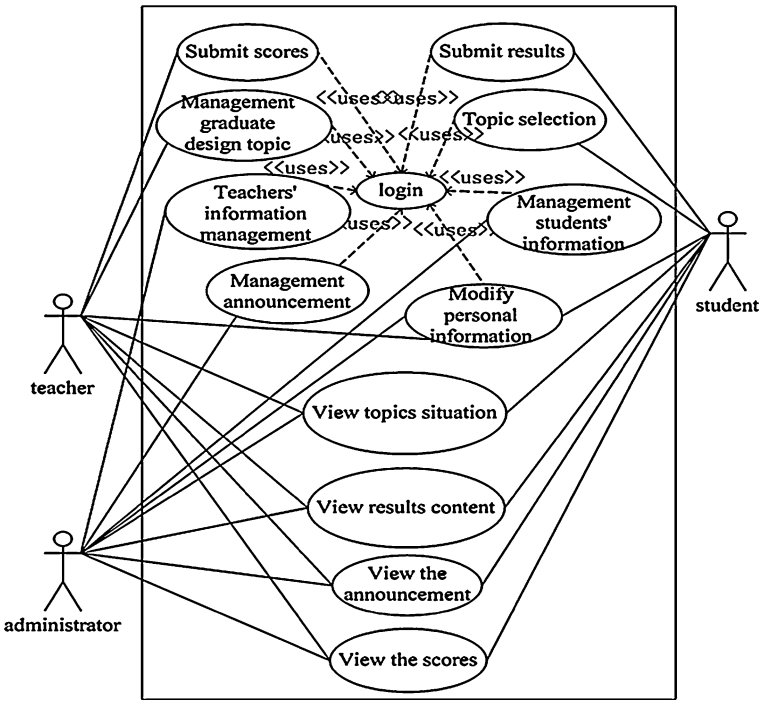


Fig. 98.3 UML use cases diagram of the online graduation design management system

98.3.2 Static Modeling

In the static modeling stage, from the system's internal structure and its static aspect to analyze and describe all kinds of entities (object and class) and the relations in their internal and between them, the established model procedure is as follows [5]:

Identify class
Identify the responsibility of the classes, Identify the system's attributes and operation
Identify the cooperators, define the relationships between classes
Review
Classes has three main versions, there are entity classes, boundary classes and control classes [6].

Boundary classes locate in the boundary of the system and outside, including all the Windows, report forms and the classes directly interact with external equipment all are boundary classes. The boundary class needed by the system can be identified through the use case diagram; the use case that directly related to each performer at least needs a window boundary class.

Control classes are those classes that control other class to work. Almost every use case will have a control class to control other class to work, such as the menu class in main interface, the class that manages the database file etc.

Entity class is the persistent object information need preserve in the computer for long-term, the persistent objects often in the form of database and other files are saved into lasting memory such as hard disk. To relational database, usually, each entity class mapping a corresponding two-dimensional table in database, the properties of the entity class mapping to the table‘s fields and each record in the table is an object of the entity class. In recognition of the entity class, usually the recognition method of combining nouns and verbs can be used, and its specific procedure is as follows [7]:

(1) Collect information: As much as possible collect relevant information, such as requirement description, cases, and related specifications, and other resources;
(2) Analyze information: Analyze those collected information, outstand the contents such as nouns, verbs, nouns phrases, verbs phrases, noun and noun phrases that can be used as a class or class attribute candidate, verbs and phrasal verbs can be used as the candidate of operation;
(3) Screening classes: According to the following principles further define the classes:

Remove redundant classes: If two classes express the same information, should retain the class that most has the ability of the description.

Remove irrelevant classes: Delete the classes that have nothing to do with the problem or the concern is not big.

Delete fuzzy class: Some initial class boundary definition not exactly, or too wide range, should be deleted.

Delete those candidate classes whose independence properties are not strong, should be class’s attribute but not class.

The description information is just about the temporary objects in realization should be cut.

Point to the online graduation design management system, according to above-mentioned method of identifying entity class, through the analysis of the system requirement description, can obtain the entity classes and their relationships described with class diagram as shown in Fig. 98.4.

98.3.3 Dynamic Modeling

The main task in dynamic modeling stage is on the basis of first two stages, to analyze system each kind of system behavior sequence state happened and interactive relationship, and all kinds of entity state change process, to describe the dynamic behavior of the system, reflecting the dynamic relationship between system internal objects.

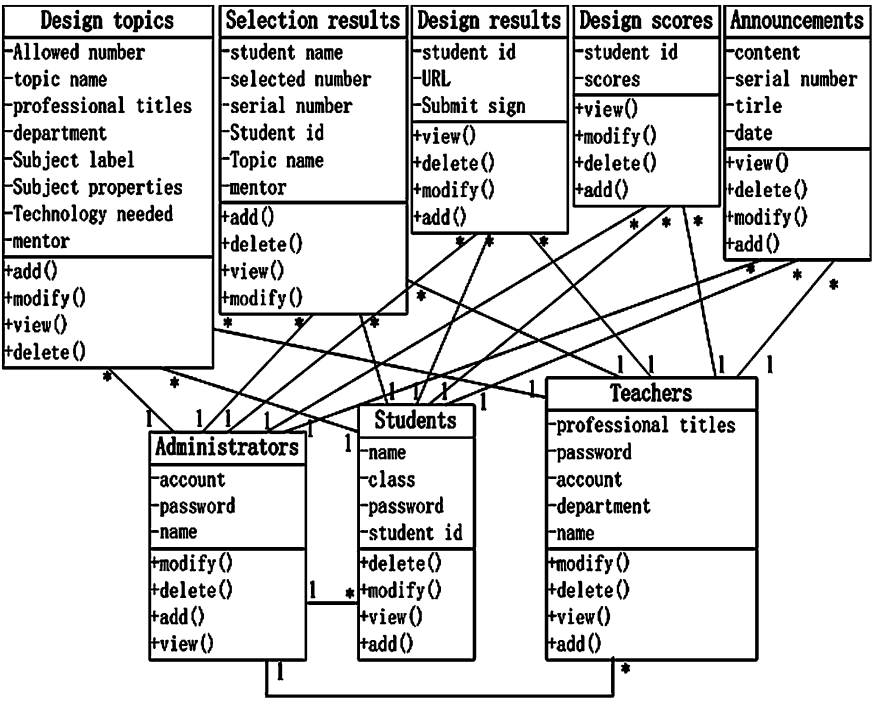


Fig. 98.4 The entity class diagram of the online graduation design management system

System dynamic models are often described by sequence diagram, cooperation diagrams, state diagrams or activity diagrams. Among them, sequence diagram is relatively commonly used, it describes the cooperation between objects, complete the dynamic process of expected behavior and it emphasize the messages sequence of time [8]. In the sequence diagram, first from left to right, according to the order of the interaction occurred place the instance objects, and then from the top to the bottom according to time order of sending messages place news between object examples. Sequence diagram can help analysts to see processing flow, such as in online graduation design management system, students choose the graduation design subject sequence diagram as shown in Fig. 98.5.

98.4 Conclusions

Using UML and object-oriented requirement analysis method to establish web system analysis model has three types, they are use cases model, static model and dynamic model, from different sides describes the system demand, helps and guides the developers to understand system needs. This paper gives the modeling process, the main method key points and application example. Through the online

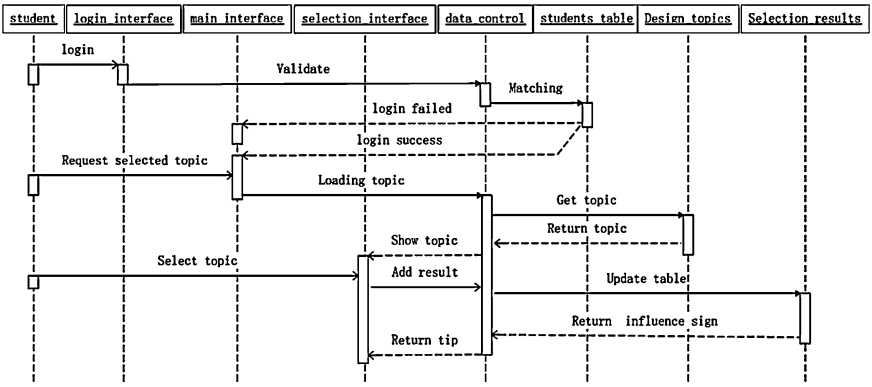


Fig. 98.5 The sequence diagram describes the case of choosing topic for student

graduation design management system UML modeling practice shows that the reference to the general process of object-oriented analysis based on UML, for the Web application system modeling, can do an effective demand analysis. The modeling process and methods used in the web application demand analysis of the system not only reduces the complexity of the system development, but also improves the system reusability and maintainability, and at the same time, improves the efficiency of system development.

References

1. Fang H-P, Chen H-P (2006) Modeling of information system based on UML [J]. Comput Eng Des 27(19):3613–3615, 3655
2. Ye J-M (2010) Software engineering [M], vol 9. Tsinghua University Press, Beijing
3. Wang X, Pan S (2004) Research on the application of UML modelling in an example [J]. J East China Shipbuild Inst (Nat Sci Ed) 18(3):55–57
4. Qian L-Q, Zhao W-G, Niu J-Y (2007) Software engineering [M], vol 3. Tsinghua University Press, Beijing
5. Pressman RS (2007) Software engineering: a practitioner's approach sixth edition [M], vol 1. Mechanical Industry Press, Beijing
6. Diao C-J, Diao Y-I (2007) UML system modeling and analysis design [M], vol 6. Mechanical Industry Press, Beijing
7. Xu J-T, Zeng L, Peng D-Z (2004) Software engineering: theory and practice [M], vol 7. Higher Education Press, Beijing
8. Sun H-M, Jia R-S (2010) Analysis and design the network aided teaching resources platform based on UML [J]. Inf Technol Informatiz 6:90–93

Chapter 99
Fault-Tolerant Control of Time-Delay Systems

Shaohua Wang

Abstract The fault-tolerant control for time-delay systems with sensor failures was investigated in this note. Attention was focused on the design of output dynamical feedback controller which guaranteed the asymptotical stability of the closed-loop systems when the sensor failed. Based on the special sensor failure model and stability criterion of time-delay systems, by introducing nonlinear transformation, the sufficient conditions for the existence of feedback controller are obtained. Moreover, the desired controller is obtained by using the cone complementary linearization iterative algorithm. Numerical examples are given to show the effectiveness of the proposed method.

Keywords Time delay · Fault-tolerant control · Cone complementary linearization · Linear matrix inequality (LMI)

99.1 Introduction

Faults may drastically change the system behavior ranging from performance degradation to instability. In order to maintain high levels of system survivability and performance, extensive researches have focused on robust fault-tolerant control of systems over the past few decades. As fault tolerance is taken into consideration, the principle goal is the maintenance of system stability under fault scenarios.

S. Wang (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: wangsh@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_99,

To the fault-tolerant control of time-delay systems, many efforts mainly focused on the state feedback methods [1]. When the state can not be obtained, in [2, 3] the static output feedback controllers were designed to solve the fault-tolerant control of time-delay systems. To the best of our knowledge, few results have been achieved on the fault-tolerant control for time-delay systems.

To the stability analysis of time-delay systems, most efforts were focused on reducing the conservatism of results [4–7]. To the controller design problem, getting the parameters of the controller is also important, which need some special functional construction and model transformations. In this paper, we consider the FTC of time-delay systems when sensor faults happen. Based on the special sensor model of failure and stability condition, by using the nonlinear transformation and the cone complementary linearization algorithm, the existence condition of time-delay system are obtained in terms of LMIs, which guarantee the asymptotically stable of the closed-loop systems even if the sensor faults occur, and then in the end, an example is given to show the effectiveness of the proposed methods.

99.2 Problem Formulation

Consider the following time-delay system:

$$\begin{aligned} \dot{x}(t) &= \boldsymbol{A}x(t) + \boldsymbol{A}_\tau x(t-\tau) + \boldsymbol{B}u(t) \\ y(t) &= \boldsymbol{C}x(t), x(t) = \varphi(t), \quad t \in [-\tau_m, 0] \end{aligned} \tag{99.1}$$

where $x(t) \in \boldsymbol{R}^{n_p}$, $u(t) \in \boldsymbol{R}^q$, $y(t) \in \boldsymbol{R}^m$ are state, input and output of the system. $\boldsymbol{A}, \boldsymbol{A}_\tau, \boldsymbol{B}, \boldsymbol{C}$ are known constant matrices, τ is the constant time delay and satisfies $0 \le \tau \le \tau_m$.

For the fault of the sensors, we introduce the sensor failure matrix $\boldsymbol{F}_\mathbf{s}$:

$$\boldsymbol{F}_\mathrm{s} \in \Theta_s = \{\boldsymbol{X} | \boldsymbol{X} = \mathrm{diag}\{f_{s1}, f_{s2}, \ldots, f_{sl}\}, 0 \le f_{si} \le \delta_s, \delta_s \ge 1, i = 1, 2, \ldots, l\} \tag{99.2}$$

$\boldsymbol{F}_\mathbf{s}$ is the diagonal matrix, $f_{\mathrm{si}} = 1$ denote the corresponding circuit is normal, otherwise the sensor failure happen; where $f_{\mathrm{si}} = 1$ denote the signal is reduced, $f_{\mathrm{si}} = 1$ denote the signal is amplified.

The aim of this paper is to design the following dynamical feedback controller:

$$\begin{aligned} \dot{x}_K(t) &= \boldsymbol{A}_K x_K(t) + \boldsymbol{B}_K y(t) \\ u(t) &= \boldsymbol{C}_K x_K(t), t \ge 0 \end{aligned} \tag{99.3}$$

to $\forall \tau \in [0, \tau_m]$, when the sensor failure happen, such that the following closed-loop system is asymptotically stable.

$$\dot{\tilde{x}}(t) = \tilde{\boldsymbol{A}}\tilde{x}(t) + \tilde{\boldsymbol{A}}_\tau \tilde{x}(t-\tau) \tag{99.4}$$

where $\tilde{x}(t) = \begin{bmatrix} x(t) \\ x_k(t) \end{bmatrix}$, $\tilde{A} = \begin{bmatrix} A & BC_K \\ B_K F_s C & A_K \end{bmatrix}$, $\tilde{A}_\tau = \begin{bmatrix} A_\tau & 0 \\ 0 & 0 \end{bmatrix}$

Lemma 1 [8] *Given any matrices* $\boldsymbol{X}$, $\boldsymbol{Y}$ *and scalar* > 0, *then*

$$\boldsymbol{X}\boldsymbol{Y}^T + \boldsymbol{Y}\boldsymbol{X}^T \le \varepsilon \boldsymbol{X}\boldsymbol{X}^T + \varepsilon^{-1}\boldsymbol{Y}\boldsymbol{Y}^T.$$

Lemma 2 [9] *For* $\forall \tau \in [0, \tau_m]$, *the system* $\dot{x}(t) = \boldsymbol{A}x(t) + \boldsymbol{A}_\tau x(t-\tau)$ *is asymptotically stable, if there exist matrices* $\boldsymbol{P} > 0, \boldsymbol{Q} > 0$, *and any matrices* $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$, *such that the following LMIs hold:*

$$\begin{bmatrix} PA + A^T P + \tau_m X + Y + Y^T + Q & PA_\tau - Y & \tau_m A^T Z \\ * & -Q & \tau_m A_\tau^T Z \\ * & * & -\tau_m Z \end{bmatrix} < 0, \quad \begin{bmatrix} X & Y \\ * & Z \end{bmatrix} \ge 0$$

99.3 Main Results

From Lemma 2, we can see that if there exist matrices $\tilde{\boldsymbol{P}} > 0, \tilde{\boldsymbol{Q}} > 0$, and any matrices $\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}}$, such that the following LMIs hold:

$$\Sigma = \begin{bmatrix} \tilde{P}\tilde{A} + \tilde{A}^T\tilde{P} + \tau_m\tilde{X} + \tilde{Y} + \tilde{Y}^T + \tilde{Q} & \tilde{P}\tilde{A}_\tau - \tilde{Y} & \tau_m\tilde{A}^T\tilde{Z} \\ * & -\tilde{Q} & \tau_m\tilde{A}_\tau^T\tilde{Z} \\ * & * & -\tau_m\tilde{Z} \end{bmatrix} < 0 \tag{99.5a}$$

$$\begin{bmatrix} \tilde{X} & \tilde{Y} \\ * & \tilde{Z} \end{bmatrix} \ge 0 \tag{99.5b}$$

then for $\forall \tau \in [0, \tau_m]$, the time-delay system (99.4) is asymptotically stable.

Let $\tilde{A}_S = \begin{bmatrix} A & BC_K \\ \delta_S B_K C & A_K \end{bmatrix}$, $B_S = \begin{bmatrix} 0 \\ \delta_S B_K \end{bmatrix}$ $C_S = [\, C \quad 0 \,]$,

$$\tilde{F} = \delta_S^{-1}(F_S - \delta_S I),$$

$$\Sigma_S = \begin{bmatrix} \tilde{P}\tilde{A}_S + \tilde{A}_S^T\tilde{P} + \tau_m\tilde{X} + \tilde{Y} + \tilde{Y}^T + \tilde{Q} & \tilde{P}\tilde{A}_\tau - \tilde{Y} & \tau_m\tilde{A}_S^T\tilde{Z} \\ * & -\tilde{Q} & \tau_m\tilde{A}_\tau^T\tilde{Z} \\ * & * & -\tau_m\tilde{Z} \end{bmatrix}.$$

We can get:

$$\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}_S + \boldsymbol{B}_S\tilde{\boldsymbol{F}}\boldsymbol{C}_S \tag{99.6}$$

For $\tilde{F}\tilde{F}^T \le I$, from Lemma 1, we can obtain:

$$\Sigma \le \Sigma_S + \varepsilon^{-1}\begin{bmatrix} \tilde{P}B_S \\ 0 \\ \tau_m\tilde{Z}B_S \end{bmatrix}\begin{bmatrix} \tilde{P}B_S \\ 0 \\ \tau_m\tilde{Z}B_S \end{bmatrix}^T + \varepsilon\begin{bmatrix} C_S^T \\ 0 \\ 0 \end{bmatrix}\begin{bmatrix} C_S^T \\ 0 \\ 0 \end{bmatrix}^T \tag{99.7}$$

So, the condition of (99.5a) is, there exists scalar $\varepsilon > 0$, such that the (99.8) holds:

$$\begin{bmatrix} \Theta & \tilde{P}\tilde{A}_\tau - \tilde{Y} & \tau_m \tilde{A}_S^T \tilde{Z} & \tilde{P}B_S & \varepsilon C_S^T \\ * & -\tilde{Q} & \tau_m \tilde{A}_\tau^T \tilde{Z} & 0 & 0 \\ * & * & -\tau_m \tilde{Z} & \tau_m \tilde{Z} B_S & 0 \\ * & * & * & -\varepsilon I & 0 \\ * & * & * & * & -\varepsilon I \end{bmatrix} < 0 \tag{99.8}$$

Then we let

$$\tilde{\boldsymbol{P}} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{P}_2 \\ \boldsymbol{P}_2^T & \boldsymbol{P}_3 \end{bmatrix}, \boldsymbol{P} \in \boldsymbol{P}^{n\times n}, \boldsymbol{P}_i \in R^{n\times n}, (i = 2, 3) \tag{99.9}$$

and P_2 is nonsingular. Construct the following transform matrices:

$$\boldsymbol{T}_1 = \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & \boldsymbol{P}\boldsymbol{P}_2^{-T} \end{bmatrix} \quad \boldsymbol{T}_2 = \text{diag}\{\boldsymbol{T}_1, \boldsymbol{T}_1, \boldsymbol{T}_1\} \tag{99.10}$$

after similitude transformation, we know that LMI (99.8) and (99.5b) are equivalent to:

$$\begin{bmatrix} \bar{P}\bar{A} + \bar{A}^T\bar{P} + \tau_m \bar{X} \\ \quad + \bar{Y} + \bar{Y}^T + \bar{Q} & \bar{P}\bar{A}_\tau - \bar{Y} & \tau_m \bar{A}^T \bar{Z} & \bar{P}\bar{B}_S & \varepsilon C_S^T \\ * & -\bar{Q} & \tau_m \bar{A}_\tau^T \bar{Z} & 0 & 0 \\ * & * & -\tau_m \bar{Z} & \tau_m \bar{Z}\bar{B}_S & 0 \\ * & * & * & -\varepsilon I & 0 \\ * & * & * & * & -\varepsilon I \end{bmatrix} < 0 \tag{99.11a}$$

$$\begin{bmatrix} \bar{X} & \bar{Y} \\ * & \bar{Z} \end{bmatrix} \geq 0 \tag{99.11b}$$

where

$$\bar{P} = \begin{bmatrix} P & P \\ P & PP_2^{-T}P_3P_2^{-1}P \end{bmatrix}, \ \bar{Q} = T_1\tilde{Q}T_1^T, \ \bar{X} = T_1\tilde{X}T_1^T, \bar{A} = \begin{bmatrix} A & B\bar{C}_K \\ \delta_S \bar{B}_K C & \bar{A}_K \end{bmatrix},$$

$$\bar{A}_\tau = \begin{bmatrix} A_\tau & 0 \\ 0 & 0 \end{bmatrix}, \ \bar{Z} = T_1\tilde{Z}T_1^T, \ \bar{B}_S = \begin{bmatrix} 0 \\ \delta_S \bar{B}_K \end{bmatrix}, \ \bar{Y} = T_1\tilde{Y}T_1^T, \ \bar{A}_K = P^{-1}P_2A_KP_2^{-1}P,$$

$\bar{C}_K = C_K P_2^{-1} P$, $\bar{B}_K = P^{-1}P_2B_K$. From (99.5b), (99.8), (99.11), we can see that the system $(\tilde{\boldsymbol{A}}_S, \tilde{\boldsymbol{A}}_\tau)$ is algebra equivalent to $(\bar{\boldsymbol{A}}, \bar{\boldsymbol{A}}_\tau)$, and the difference is the matrices $\boldsymbol{A}_K, \boldsymbol{B}_K, \boldsymbol{C}_K$ and $\bar{\boldsymbol{A}}_K, \bar{\boldsymbol{B}}_K, \bar{\boldsymbol{C}}_K$. Therefore, we can deal with the system $(\tilde{\boldsymbol{A}}_S, \tilde{\boldsymbol{A}}_\tau)$ instead of system $(\bar{\boldsymbol{A}}, \bar{\boldsymbol{A}}_\tau)$.

Let

$$S^{-1} = PP_2^{-T}P_3P_2^{-1}P - P = PP_2^{-T}\left(P_3 - P_2^TP^{-1}P_2\right)P_2^{-1}P \tag{99.12}$$

Then $\bar{P} = \begin{bmatrix} \boldsymbol{P} & \boldsymbol{P} \\ \boldsymbol{P} & \boldsymbol{S}^{-1} + \boldsymbol{P} \end{bmatrix}$.

Form $P(S + P^{-1}) = (S^{-1} + P)S > 0$, we know $\boldsymbol{J} = \boldsymbol{S} + \boldsymbol{P}^{-1} > 0$.

Construct the following transform matrices:

$$\boldsymbol{T}_6 = \text{diag}\{\boldsymbol{T}_3, \boldsymbol{T}_4, \boldsymbol{T}_5, \boldsymbol{I}, \boldsymbol{I}\}$$

where $T_3 = \begin{bmatrix} J & -S \\ I & 0 \end{bmatrix}$, $T_4 = \begin{bmatrix} I & -JS \\ I & 0 \end{bmatrix}$, $T_5 = T_3\bar{P}\bar{Z}^{-1}$, and introduce the matrices $Q_4 > 0, Q_6 > 0, Q_5$ satisfying:

$$\begin{bmatrix} Q_4 & Q_5 \\ Q_5^T & Q_6 \end{bmatrix} \le \begin{bmatrix} J^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} \begin{bmatrix} J^{-1} & 0 \\ 0 & I \end{bmatrix} \tag{99.13a}$$

after similitude transformation, the in equation (99.11a) is equivalent to:

$$\begin{bmatrix} \Xi_{11} & \Xi_{12} & A_\tau - Y_1J^{-1} & A_\tau - Y_2 & JA^T + W_C^TB^T & \tau_mL^T & 0 & \varepsilon JC^T \\ * & \Xi_{22} & PA_\tau - Y_3J^{-1} & PA_\tau - Y_4 & \tau_mA^T & A^TP + C^TW_B^T & W_B & \varepsilon C^T \\ * & * & -Q_4 & -Q_5 & \tau_mA_\tau^T & \tau_mA_\tau^TP & 0 & 0 \\ * & * & * & -Q_6 & \tau_mA_\tau^T & \tau_mA_\tau^TP & 0 & 0 \\ * & * & * & * & -\tau_mZ_1 & -\tau_mZ_2 & 0 & 0 \\ * & * & * & * & * & -\tau_mZ_3 & \tau_mW_B & 0 \\ * & * & * & * & * & * & -\varepsilon I & 0 \\ * & * & * & * & * & * & * & -\varepsilon I \end{bmatrix} < 0 \tag{99.13b}$$

where $\Xi_{11} = AJ - BW_C + JA^T + W_C^TB^T + \tau_mX_1 + Y_1 + Y_1^T + Q_1$, $W_B = \delta_S P\bar{B}_K$,

$$\Xi_{12} = A + L^T + \tau_mX_2 + Y_2 + Y_3^T + Q_2,\ X = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix}, Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix},$$

$$\Xi_{22} = PA + W_BC + A^TP + C^TW_B^T + \tau_mX_3 + Y_4 + Y_4^T + Q_3,$$

$$Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix},\ Z = \begin{bmatrix} Z_1 & Z_2 \\ Z_2^T & Z_3 \end{bmatrix},$$

$L = PAJ + P\delta_S\bar{B}_KCJ - PB\bar{C}_KS - P\bar{A}_KS,\ W_C = \bar{C}_KS.$

and the (99.11b) is equivalent to:

$$\begin{bmatrix} T_3\bar{X}T_3^T & T_3\bar{Y}T_3^T \\ * & T_3\bar{Z}T_3^T \end{bmatrix} \ge 0 \tag{99.14}$$

denoting:

$$T_3\bar{P}T_3^T = \begin{bmatrix} J & I \\ I & P \end{bmatrix},\ T_3\bar{Z}T_3^T = \begin{bmatrix} J & I \\ I & P \end{bmatrix} Z^{-1} \begin{bmatrix} J & I \\ I & P \end{bmatrix}$$

and introduce the matrix $W = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} \geq 0$, if

$$\begin{bmatrix} X & Y \\ * & W \end{bmatrix} \geq 0\ W \leq \begin{bmatrix} J & I \\ I & P \end{bmatrix} Z^{-1} \begin{bmatrix} J & I \\ I & P \end{bmatrix} \tag{99.15}$$

then we can get the (99.14).

From the mentioned above, we can get the main result of this paper.

Theorem 1 *If there exist matrices* $\boldsymbol{P} > 0$, $\boldsymbol{J} > 0$, $\boldsymbol{Q}_1 > 0$, $\boldsymbol{Q}_3 > 0$, $\boldsymbol{Q}_4 > 0$, $\boldsymbol{Q}_6 > 0$, $\boldsymbol{X}_1 \geq 0$, $\boldsymbol{X}_3 \geq 0$, $\boldsymbol{Z}_1 \geq 0$, $\boldsymbol{Z}_3 \geq 0$, $\boldsymbol{W} \geq 0$, *any matrices* $\boldsymbol{Q}_2$, $\boldsymbol{Q}_5$, $\boldsymbol{X}_2, \boldsymbol{Z}_2$, $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$, $\boldsymbol{Y}_3$, $\boldsymbol{Y}_4$, $\boldsymbol{W}_B$, $\boldsymbol{W}_C$, $\boldsymbol{L}$ *and scalar* $\varepsilon > 0$, *such that the* (99.13a)–(99.15) *hold, then there exists the dynamical feedback controller* (99.3), *for* $\forall t \in [-\tau_m, 0]$, *guarantee the asymptotical stability of the closed-loop system* (99.4), *and the parameter of the controller is given as*:

$$\begin{aligned} \bar{\boldsymbol{A}}_K &= P^{-1}(\boldsymbol{PAJ} + \boldsymbol{P}\delta_S\bar{\boldsymbol{B}}_K\boldsymbol{CJ} - \boldsymbol{PB}\bar{\boldsymbol{C}}_K\boldsymbol{S} - \boldsymbol{L})\boldsymbol{S}^{-1} \\ \bar{\boldsymbol{B}}_K &= \delta_S^{-1}\boldsymbol{P}^{-1}\boldsymbol{W}_B, \bar{\boldsymbol{C}}_K = \boldsymbol{W}_C\boldsymbol{S}^{-1} \end{aligned} \tag{99.16}$$

But we can see that the (99.13a), (99.14), (99.15) are non-convex conditions, so we let $\boldsymbol{Y}_1 = 0$, $\boldsymbol{Y}_3 = 0$, introduce parameter matrices, and use cone complementary linearization iterative algorithm [10] to get the controller parameters.

Let $V = \begin{bmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{bmatrix} > 0$,, then (99.13a) is equivalent to:

$$\begin{bmatrix} \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} & \begin{bmatrix} J^T & 0 \\ 0 & I \end{bmatrix} \\ \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} & \begin{bmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{bmatrix} \end{bmatrix} \geq 0 \begin{bmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{bmatrix} \begin{bmatrix} Q_4 & Q_5 \\ Q_5^T & Q_6 \end{bmatrix} = I \tag{99.17}$$

Let $U = \begin{bmatrix} U_1 & U_2 \\ U_2^T & U_3 \end{bmatrix} > 0$, then (99.14) is equivalent to:

$$\begin{bmatrix} \begin{bmatrix} U_1 & U_2 \\ U_2^T & U_3 \end{bmatrix} & \begin{bmatrix} J & I \\ I & P \end{bmatrix}^{-1} \\ \begin{bmatrix} J & I \\ I & P \end{bmatrix}^{-1} & \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} \end{bmatrix} \geq 0\ UZ = I \tag{99.18}$$

Let $\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_2^T & \alpha_3 \end{bmatrix} > 0, \beta = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_2^T & \beta_3 \end{bmatrix} > 0$, then (99.15) is equivalent to:

$$\begin{bmatrix} U_1 & U_2 & \alpha_1 & \alpha_2 \\ U_2^T & U_3 & \alpha_2^T & \alpha_3 \\ \alpha_1 & \alpha_2 & \beta_1 & \beta_2 \\ \alpha_2^T & \alpha_3 & \beta_2^T & \beta_3 \end{bmatrix} \geq 0 \begin{bmatrix} J & I \\ I & P \end{bmatrix} \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_2^T & \alpha_3 \end{bmatrix} = I \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_2^T & \beta_3 \end{bmatrix} = I \tag{99.19}$$

Therefore, given scalar $\varepsilon > 0$, the problem of dynamical controller is transformed to the following cone complementary linearization.

$$\begin{aligned} \text{Minimize} \left\{ tr\left(UZ + \begin{bmatrix} V_1 & V_2 \\ V_2^T & V_3 \end{bmatrix} \begin{bmatrix} Q_4 & Q_5 \\ Q_5^T & Q_6 \end{bmatrix} \right) \right. \\ \left. + tr\left(\begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_2^T & \beta_3 \end{bmatrix} + \begin{bmatrix} J & I \\ I & P \end{bmatrix} \begin{bmatrix} \alpha_1 & \alpha_2 \alpha_2^T & \alpha_3 \end{bmatrix} \right) \right\} \end{aligned} \tag{99.20}$$

subject to (99.13b), (99.15), (99.17), (99.18) and

$$\begin{cases} J > 0, P > 0, \begin{bmatrix} U & I \\ I & Z \end{bmatrix} > 0, \begin{bmatrix} V_1 & V_2 & I & 0 \\ V_2^T & V_3 & 0 & I \\ I & 0 & Q_4 & Q_5 \\ 0 & I & Q_5^T & Q_6 \end{bmatrix} \geq 0 \\ \begin{bmatrix} \alpha_1 & \alpha_2 & I & 0 \\ \alpha_2^T & \alpha_3 & 0 & I \\ I & 0 & J & I \\ 0 & I & I & P \end{bmatrix} \geq 0, \begin{bmatrix} \beta_1 & \beta_2 & I & 0 \\ \beta_2^T & \beta_3 & 0 & I \\ I & 0 & W_1 & W_2 \\ 0 & I & W_2^T & W_3 \end{bmatrix} \geq 0 \end{cases}$$

99.4 Simulation

Example Consider the following time-delay system:

$$\dot{x}(t) = \begin{bmatrix} -0.3 & 0.4 \\ 0 & -0.8 \end{bmatrix} x(t) + \begin{bmatrix} 0.2 & 0 \\ 0 & 0.1 \end{bmatrix} x(t-\tau) + \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix} u(t)$$

$$y(t) = [\,-0.1 \quad 0.1\,] x(t)$$

When $\delta_S = 2$, let $\varepsilon = 1,\ \tau_m = 1$, using the main result of this paper, we can get the dynamical output feedback controller:

$$\dot{x}_K(t) = \begin{bmatrix} -2.0869 & 0.2973 \\ 1.5235 & -0.8814 \end{bmatrix} x_K(t) + \begin{bmatrix} 0.013 \\ -0.0009 \end{bmatrix} y(t)$$

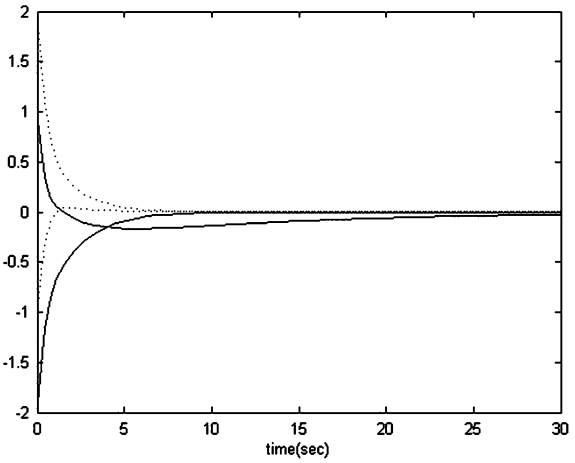


Fig. 99.1 The state trajectory of the closed-loop system

$$u(t) = [\,15.108 \quad -0.7148\,]\, x_K(t),\ t \geq 0$$

The state trajectory of the closed-loop system is given in Fig. 99.1.

99.5 Conclusion

In this paper, the fault-tolerant control of time-delay system via dynamical output feedback is investigated. Based on the special sensor model of failure and delay dependent stability criterion for nominal time-delay systems, by introducing the nonlinear transformation, the sufficient condition for the existence of feedback controller are obtained. Moreover, the desired controller is given by using the LMI and the cone complementary linearization iterative algorithm. Simulation is given to illustrate the effectiveness of the proposed method.

Acknowledgments This work is supported by the Doctoral Foundation Project of Henan Polytechnic University under Grant B2010-50, the Natural Science Project of Henan Province Office of Education under Grant 2008B510008 and Henan Provincial Open Laboratory for Control Engineering Key Disciplines under Grant KG2009-07.

References

1. Yang JJ, Wu FX, Shi ZK (2000) Robust H∞ fault-tolerant controller design for linear delay-systems with uncertainty. Control Theory Appl 17:442–444
2. Marx B, Koenig D, Georges D (2004) Robust fault-tolerant for descriptor systems. IEEE Trans Autom Control 49:1869–1876
3. Jakob S, Vincent DB (2004) Fault tolerant-control a simultaneous stabilization result. IEEE Trans Autom Control 49:305–311
4. Xu S, Lam J (2005) Improved delay-dependent stability criteria for time-delay systems. IEEE Trans Autom Control 50:384–387

5. He Y, Wang QG, Lin C, Wu M (2005) Augmented Lyapunov functional and delay-dependent stability criteria for neutral systems. Int J Robust Nonlin Control 15:923–933
6. Qian W, Cong S, Sun YX, Fei SM (2009) Novel robust stability criteria for uncertain systems with time-varying delay. Appl Math Comput 215:866–872
7. Qian W, Liu J, Sun YX, Fei SM (2010) A less conservative robust stability criteria for uncertain neutral systems with mixed delays. Math Comput Simul 80:1007–1017
8. Khargonekar PP, Peiersen IR, Zhou K (1990) Robust stabilization of uncertain linear systems: quadratic stabilization and H∞ control theory. IEEE Trans Autom Control 35:356–361
9. Gu K (2001) A further refinement of discretized Lyapunov functional method for the stability of time-delay systems. Int J Control 74:967–976
10. Young SM, Park P, Kwon WH (2001) Delay-dependent robust stabilization of uncertain state-delayed systems. Int J Control 74:1447–1455

Chapter 100
Prediction Strategy Level of Coal–Gas-Outburst Based on Expert System

Fuzhong Wang and Weizhe Liu

Abstract In order to solve the imperfect information and data of coal–gas-outburst prediction level and the problem of low prediction accuracy. Based on the analysis of factors affecting the mining face of coal–gas-outburst prediction, we used expert system theory, and established the level of coal–gas-outburst prediction model. We specifically referenced to the composition of expert systems theory on the mining face of coal–gas-outburst prediction, knowledge representation and knowledge structure, reasoning principle and the knowledge representation model of geological conditions on the prediction of outburst area. Simulation results show that the model correctly predicted 100%, basically consistent with the actual situation.

Keywords Coal–gas-outburst · Expert system · Mining face

100.1 Introduction

Gas outburst is one of the most serious disasters in coal mine production, not only it affects the normal production of coal mining enterprises, but also is a direct threat to the lives of workers. Therefore, prevention and controlling of coal–gas-outburst has become the country's need to solve the major issue. Conventional coal–gas-outburst is predicted mainly by selecting a few of the indicators that

F. Wang (✉) · W. Liu
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: wangfzh@hpu.edu.cn

W. Liu
e-mail: liuweizhe1987@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_100,

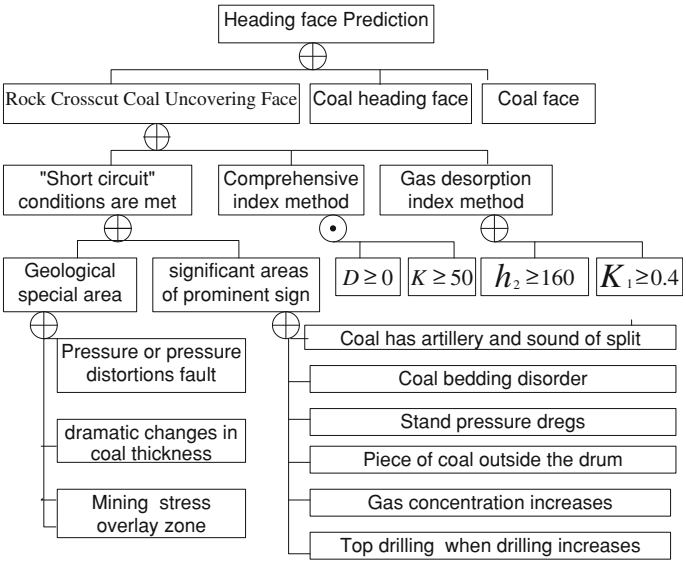


Fig. 100.1 The knowledge model of coal–gas-outburst in Rock Crosscut Coal uncovering face

affect the delineation of the threshold, inevitably increasing the subjectivity of the threshold selected, reducing the reliability of outburst prediction. Expert system can use the knowledge of human experts and problem-solving approach and the knowledge experience of an expert level to solve complex engineering problems. This article used the expert system theory and established the level of coal–gas-outburst expert system prediction model and improved the accuracy of gas prediction.

100.2 The Factors of Affecting Coal–Gas-Outburst of Mining face

The key that the establishment of coal–gas-outburst prediction expert system in mining face is to find out the correlation between outburst events and various factors. Based on the extensive collection, collation and induction about the human experts' experience and knowledge of coal–gas-outburst prediction in mining face [1], over the years we made a careful analysis and abstracted the actual data of the coal–gas-outburst in the coal mine at the eastern part of Jiaozuo and came down out the knowledge model of coal–gas-outburst in Rock Crosscut Coal uncovering face in Fig. 100.1.

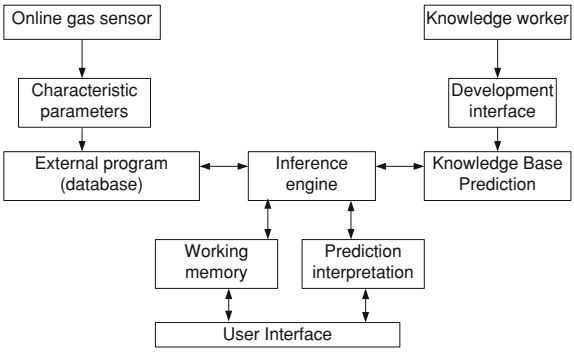


Fig. 100.2 Expert system structure

100.3 Structure of Coal–Gas-Outburst Prediction Expert System

Structure of coal–gas-outburst prediction expert system is shown in Fig. 100.2. It is composed of external program, inference engine, knowledge base, working memory, forecast explain and the user interface. Knowledge is a bridge for communication between computers and the domain expert, which allows experts use the way that they are familiar with to describe the thought process of problem-solving, without having to write complex computer programs. Inference engine based on the user questions and known facts, searching and matching the corresponding knowledge in the Knowledge Base, and based on the knowledge to infer and calculate the new facts from the known facts [2]. External programs include the database, extended discs and algorithms and support the work of the expert system. This article use the data control which is provided by VB and write code by setting some properties, realizing to query, add, modify, delete and some other operations to historical data of coal–gas-outburst. Knowledge acquisition mainly refers to talking to experts in the field of problem, or extracting the process of problem-solving by using machine learning and data mining and other knowledge discovery method from typical data sample. User interface mainly realizes to make the information between user and expert systems friendly and visualization and interactive.

100.3.1 Knowledge Representation and Knowledge Base of Coal–Gas-Outburst Prediction

The system uses integrated knowledge representation of object-oriented and visualization “body of knowledge• object blocks •components •” and knowledge representation of the “rule frame + rule body”. The model of this rule is especially suitable for describing causality of the process and non-resolution mapping, with great flexibility, it is easily added and modified on the rules. The factors of coal–

gas-outburst in mining face such as velocity of gas irradiation, consistent coefficient of coal, gas pressure of coal seam, the amount of drilling cuttings, drilling gas inrush initial rate and other geological factors, the relationship between them and coal–gas-outburst can be represented by this rule and convenient rewritten. The basic elements of knowledge representation include knowledge object, rule-based reasoning object and description of multi-technology integration [3]. Entity of rule-based reasoning implement is called the RB-Rule Body, the basic structure of knowledge is as follows:

```
ASK
  variable definition
BEGIN
  {      }
RG name of rule group
  KO feature name: target;
  …
  RULE rule 1
    IF premise itemsets THEN conclusion itemsets
    RB{conclusion itemsets =FUNCTION(premise itemsets);}
  …
END;
KO body of knowledge
{…
RULE rule name.feature name;}
```

For example, knowledge of rule-based reasoning object "gas control" is expressed as:

```
RG gas control.
KO decision-making of putting and exploration; gas emissions; gas drainage;
KO exploitation of emancipation layer: complete liberation, not fully emancipated;
RULE 1
  IF low gas concentration THEN gas emissions
    RB
    {gas emissions=gas emission quantity * discharge time;}
RULE 2
  IF high gas concentration THEN gas drainage
  RB
  {gas drainage=the average drainage volume * drainage time;}
  END:
```

The method of enabling the corresponding objective function of rule knowledge object is fast and simple :

```
KO decision-making of putting and exploration
{…
RUN RUEL decision-making of putting and exploration about gas control
…
}
```

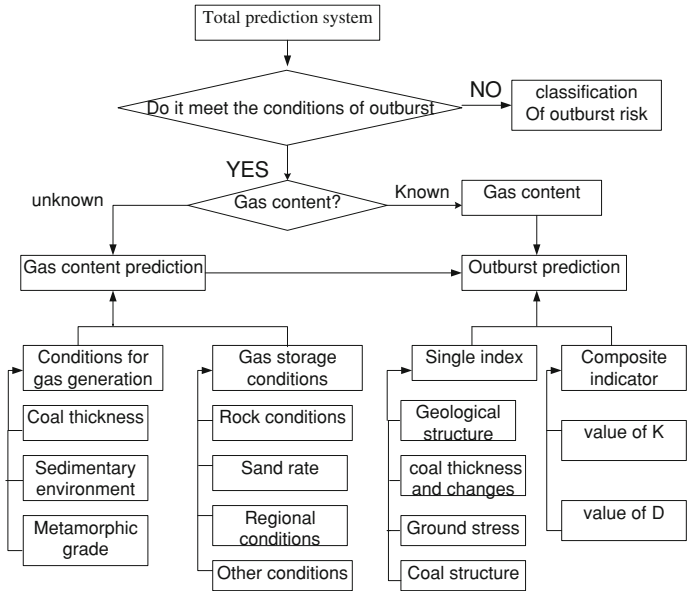


Fig. 100.3 Flow chart of prediction system

100.3.2 Reasoning Principle of Coal–Gas-Outburst Prediction Expert System

Coal–gas-outburst prediction expert system include the two major components, which are database systems and prediction systems. Database systems mainly store data and information, prediction systems primarily reason to solve. Prediction step of prediction system follows: first, we need predict the necessary conditions, if they are satisfied the necessary conditions, then the following : (1) If the values of gas content are known, we predict the gas outburst conditions only according to other geological factors and divide outburst danger section combing with the values of gas content. (2) If the values of gas content are unknown, we predict the high and low of gas content in the predicted areas according to conditions of gas occurrence and we predict the gas outburst and make the degree of risk combing with other geological factors. If they do not meet the necessary conditions, it will determine the results directly back to the prediction system. The total prediction system is shown in Fig. 100.3.

Necessary conditions of coal–gas-outburst: the beginning depth of penetration, the beginning suddenly threshold of gas pressure tectonic coal. The knowledge representation model for the necessary condition of coal–gas-outburst is as follows:

RULE A(rule name)
IF H≥the beginning depth of penetration AND
P≥the beginning depth of penetration AND
T ∈ III OR IV OR V

THEN
OUTBURST=POSSIBLE
ELSE
OUTBURST=IMPOSSIBLE
In which: *H*—coal seam depth, m;
P—gas pressure, MPa;
T—the destructive type of coal (According to the damage of coal can be divided into five categories [3]).

This rule shows: if depth of coal seam more than the beginning depth of penetration, and gas pressure exceeds the critical value, and the destruction of coal type is III or IV or V, anyone of them, then the outburst may occur; otherwise outburst does not occur.

100.3.3 The Knowledge Representation Model for the Effects Which Geological Conditions Have on Predicting the Outburst Area

According to previous studies, we know the geological conditions which have effect on the coal–gas-outburst, including geological structure of small and medium mine, coal thickness and its changes and coal structure. All these are describing to the influences that geological factors have on the outburst. In order to improve prediction results, we also use the integrated prediction indicators of gas outburst, including D and K [4].

The formula is:

$$D = \left[0.0075 \times \frac{H}{f} - 3\right](p - 0.74); K = \Delta P / f$$

In which: *D*, *K*—coal outburst indicators;
H—coal seam depth, m;
P—gas pressure, MPa;
F—consistent coefficient of coal; ΔP—Index of coal gas irradiation.

Thus, we can get the general knowledge representation model of gas outburst prediction.

RULE B(rule name)
IF earth stress =HIGH 0R
conditions of coal thickness =0K OR
(f≤threshold ANF hΔp ≥threshold)OR
thickness of soft coal≥threshold
geological structure = complex 0R
Integrated Index (D and K)=0K
THEN
geological conditions of gas outburst =OK;

Table 100.1 The results coal mine gas outburst prediction expert system in a coal mine of Jiaozuo

Outburst locations	Elevation (m)	Outburst strength (t)	Robustness coefficients of coal (f)	System prediction conclusion	Actual situation
11041 Transpotation lane	−113	13	0.36	Outburst threat area Credibility 0.73	Press out
11051 Working face	−138	30	0.38	Outburst threat area Credibility 0.78	Dump
11091 Focal mountain	−202	16	0.13	Outburst threat area Credibility 0.82	Outburst
11061 Transpotation lane	−142	94	0.09	Outburst threat area Credibility 0.95	Outburst
11091 Transpotation lane	−204	170	0.19	Outburst threat area Credibility 0.90	Outburst

100.4 Simulation Analysis

Expert systems for coal–gas-outburst prediction is an attempt of coal–gas-outburst prediction, we made a random selection about the historical data of Jiaozuo mining and made the application of the test [5], and making a simulation using matlab, the results are shown in Table 100.1. The results show that the predicted results match with the actual situation.

The reasoning process developed by system is a common reasoning control module. As long as the knowledge base, real-time database and the library of interpretation have reasonable organizations, it also can predict coal–gas-outburst of other areas and has good application prospect.

100.5 Conclusions

Based on the comprehensive analysis to the impact of the regional geological conditions of coal mine, we explore the dominant factor, screening index, through knowledge management, combing effect the factors of coal–gas-outburst and combing with expert experience and knowledge, we established the level of coal–gas-outburst prediction model. Simulation results show that the model correctly predicted 100%, in line with the actual situation.

First, the system uses integrated knowledge representation of object-oriented and visualization "body of knowledge• object blocks • components •" and knowledge representation of the "rule frame + rule body", the model of this rule is especially suitable for describing causality of the process and non-resolution mapping, with great flexibility, it is easily added and modified on the rules.

Second, we use the complete induction method to infer gas outburst and just based on the specific geological factors, predicting the conditions of the gas

outburst, and combing the gas content to divide the danger area of outburst area, this method is convenient and easy and relatively accurate.

Third, this paper established the knowledge representation model for the effects which geological conditions have on predicting the highlighted area. It fully reflects the advantages of expert systems and in particular reflect that the writing of the knowledge rules is convenient and clear and concise and easy to understand features.

Acknowledgments This work was supported by the projects of the key research funding in Henan Province (102102210203); Henan Polytechnic University, Ph.D. Foundation (B2011-064).

References

1. Zhang Z, Wu G (2001) Coal and gas outburst prediction fuzzy expert system. Guizhou Univ Technol 20(1):49–51
2. Wu B, Feng Y (2001) Expert system. Beijing Polytechnic University Press, Beijing, pp 8–10
3. Hao J (2000) Integrated expert system of coal and gas outburst area prediction. Liaoning Technical University 6:58–62
4. Wu C, Zeng Y (2004) The expert system and information storeroom of forecasting coal-gas-outburst. Coal Geol Explor 32(1):17–20
5. Hao J, Yuan C, Han D (2009) Coal and gas outburst prediction integrated expert system. Liaoning Tech Univ (Natural Sci) 19(3):240–243

Chapter 101
Development of Simulation Support Platform for Modern Tram Operation Control System

Liang Ma, Hongze Xu and Peng Zhou

Abstract The logic relationship among the subsystems of modern tram Operation Control System (OCS), and interface relationship between modern tram OCS and peripheral equipments were analyzed. On this basis, the functional requirements of OCS simulation support platform were analyzed, a design proposal was proposed, general structure and interface were designed. Upon the PC and Windows operation system, the OCS simulation support platform software was developed by VC++ 6.0 and the result has confirmed the correctness and feasible of the OCS simulation support platform.

Keywords Modern tram · Operation control system · Supporting platform

101.1 Introduction

Modern tram is an important trend of the future Urban Mass Transit. Compared with the traditional ones, modern tram has many advantages, such as medium volume, modular vehicle, low noise, low energy, and energy saving [1].

Modern tram system consists of vehicle system, track line system, traction power supply system and OCS [2]. OCS is the key subsystem to ensure the safety and to enhance efficiency.

L. Ma (✉) · H. Xu · P. Zhou
Beijing Jiaotong University, Beijing 100044, China
e-mail: 09120353@bjtu.edu.cn

H. Xu
e-mail: hzxu@bjtu.edu.cn

P. Zhou
e-mail: peng.zhou1210@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_101,

OCS is a safety criticality control system. It must meet high security, reliability, and availability in the development. From design, construction, supervision, trial operation, till it real operation, the realization of train OCS is a systematic process. It is difficult to identify the faults in it for its great mass of work, its high cost and its complexity. So its impossible to take long-term and comprehensive technical test to improve the old rail line [3]. At present, the relative research of modern tram OCS is less. So it is necessary to design Support Platform for modern tram OCS.

This paper is organized as follows. In Sect. 101.2, we analyze the functional requirements for OCS, and OCS simulation support platform. In Sect. 101.3, we first propose a model of OCS simulation support platform according to the functional requirements. Then we design general structure and interface of support platform. In Sect. 101.4, we develop the software of upon the PC and Windows operation system. Section 101.5, the relative testing is completed to confirm the correctness and feasible of OCS simulation support platform. The conclusions are provided in Sect. 101.6.

101.2 Analysis of OCS Simulation Support Platform

101.2.1 Requirements Analysis of OCS

As is shown in the Fig. 101.1, modern tram OCS consists of following subsystem: Operation Control Centre (OCC), Zone Controller (ZC), Vehicle On-Board Controller (VOBC), and Computer Interlocking (CI).

The OCC is responsible for implementing functions such as identify, tracking and displaying trains, updating passenger information system, and regulating train movement to maintain operating schedules.

The function of the ZC is to ensure the safe operation of trains in a predefined region based on the status of the wayside equipment and train operation information, and ensure the train transfer safety between the adjacent regions.

The VOBC is responsible for train location determination, generating dynamic speed profile for the safe operation of trains.

The CI is responsible for the safe implementation of traditional interlocking functions. To realize the interlocking of the signals, switches, and sections, CI interface with wayside equipment such as switch machines, signals, etc.

101.2.2 Requirements Analysis of OCS Simulation Support Platform

According to the analysis of OCS, OCS simulation support platform is composed by wayside support platform and carborne support platform.

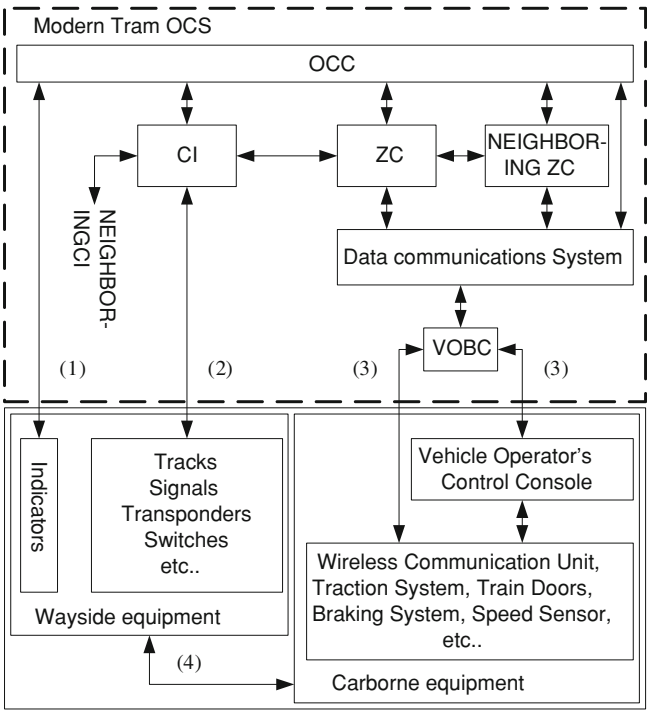


Fig. 101.1 Frame of modern tram OCS

Wayside support platform consists of following equipments: track, signal, switch, axle counting equipment, indicator, wireless access point, transponder, etc. Carborne support platform consists of following equipments: driver's control console, train operation display (TOD), wireless communication unit, transponder antenna, traction system, braking systems, speed sensor, train door, juridical recorder, etc.

1. Requirements analysis of wayside support platform. Wayside support platform provides with the status information of wayside equipment for OCS, and realizes the status changes of wayside equipment according to the command of OCS. The wayside support platform shall include the following functionalities as a minimum:

   (1) Receive control commands from CI such as switch position conversion, signal color change, simulate the relative status based on commands; Receive all messages from OCC to generate the station information.
   (2) Transmit status information to the CI, including each switch status, each track status, each signal status, and so on; Transmit indicator's status information to OCC to realize the status simulation of the wayside equipment.

(3) In view of the train location information from carborne support platform, fault information from operator, the wayside support platform simulates occupied or clearing status of tracks.
(4) According to the train location information from carborne support platform, line data information, the wayside support platform realizes the function of transponder.
(5) Fault-Inserter. Wayside support platform realizes fault injection based on the operator, the equipment including signal, switch, transponder, etc.

2. Requirements analysis of carborne support platform.
Carborne support platform simulates the operation of train, offers the status information of carborne equipment for OCS, driving information for driver. The carborne support platform shall include the following functionalities as a minimum:

(1) Receive driving commands such as accelerating, decelerating, coasting, emergency braking, etc. from driver's control console, and simulate train's operation process according to driving commands.
(2) Transmit status information to the VOBC, including train door's status, integrity, the status of traction system and braking system, etc. to realize status simulation of carborne equipment.
(3) The carborne support platform simulates driver's control console which driver realizes the necessary operates for train operation.
(4) The carborne support platform realizes the function of TOD. TOD receives train information from VOBC and shows information to the driver on the onboard TOD unit.
(5) Fault-Inserter. Carborne support platform realizes fault injection based on the operator, the equipment including TOD, wireless communication unit, transponder antenna, traction system, braking systems, speed sensor, train door, and so on.

101.3 Design of OCS Simulation Support Platform

101.3.1 Structure Design of OCS Simulation Support Platform

Figure 101.2 shows the overall model of OCS simulation support platform, which was designed according to functional requirement of OCS simulation support platform.

The main module design of OCS simulation support platform as following:

(1) TOD module.
TOD is a platform for data exchanges between VBOC and driver, with the characteristics such as the driver can set the parameters related to the tram,

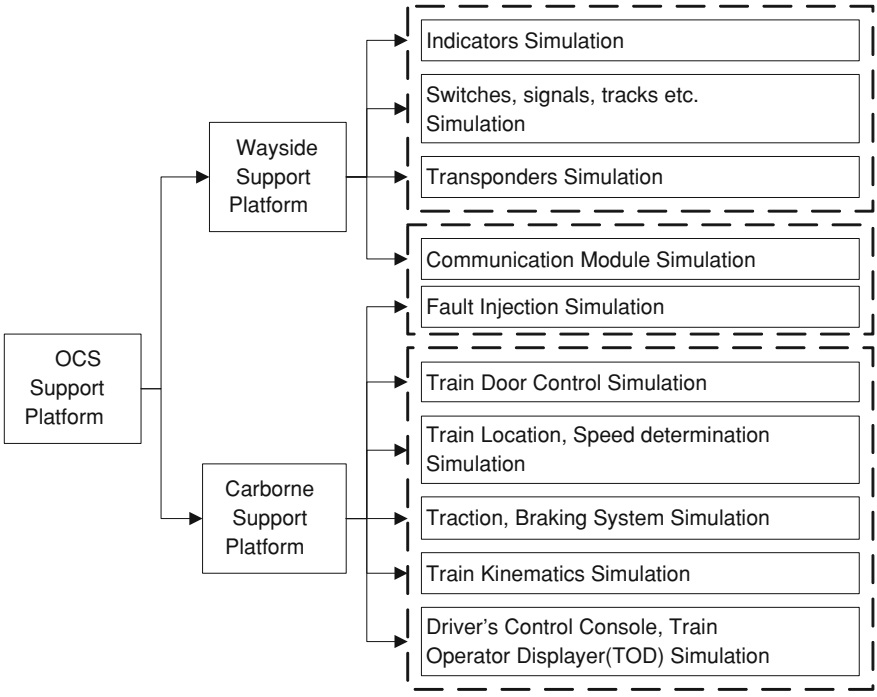


Fig. 101.2 Model of OCS simulation support platform

receiving tram and line relevant state and associated data real-time, responding timely to the commands and warnings from VOBC.

TOD module includes five sub-modules. Speed module displays the current speed, the maximum speed and the recommended speed of tram and other related information; target information module display the speed and position information to the next target point; assistant information module displays the next station, the end station, arrival time, dwell time, delayed time, and other information; information cue module displays the driver operating record and failure alarm.

(2) Driver's control console module.

The driver's control console module consists of following device: traction and braking handles (traction 4 level, coasting 1 level, braking 7 level) [4], direction handle (upward and downward direction), emergency braking button, switch request button, and train-door control button.

(3) Transponder module

The transponder module shall include the following functionalities as a minimum: first, transmit transponder ID and data information to the tram when train is through it. Second, receive dynamic information from CI system. Last, the module can simulate fault condition including cannot send information or send error information to tram.

101.3.2 Interface Design of OCS Simulation Platform

As shown in the Fig. 101.1, OCS simulation support platform interfaces with the OCC, CI, ZC, and VOBC.

(1) Interface between wayside support platform and OCC.
OCC realize information updating and status collecting for indicators via the Ethernet.
(2) Interface between wayside support platform and CI.
CI control the wayside equipment such as signal, switch, etc., collect status of signal, switch, and track via the I/O interface.
(3) Interface between carborne support platform and VOBC.
VOBC collect status of tram via I/O interface, and transmit tram's status information such as target information, station information to TOD via the Ethernet.
(4) Interface between carborne support platform and wayside support platform.
Carborne support platform transmit tram location and direction information to wayside support platform, and receive wayside information such as transponder information via the Ethernet.

101.4 Development of OCS Simulation Support Platform

The software of OCS simulation support platform consists of Logic Control Module, Human Computer Interaction Module, and Communication Module.

Logic Control Module realizes the function of sub-module based on the functional requirements of support platform. Human Computer Interaction Module shows the status of peripheral equipment, and supplies interface to Operator or Fault Injector. Communication Module completes the communications among the subsystems.

Based on the PC and Windows operation system, OCS simulation support platform software was developed using Visual C++ 6.0, Ethernet is used for data interaction among the various subsystems.

Take wayside support platform for example, main realized process of wayside support platform software can be condensed into following:

Step 1. System initialization. Initialize wayside equipment status, and establish Communication with OCC, CI, and carborne support platform.
Step 2. Transmit information of wayside equipment to the OCS. Transmit status of signals, switches, tracks, etc. to CI, and status of indicators to OCC.
Step 3. Determine whether receive commands (command maybe from OCC, CI, Fault Injector), simulate the status if receive command.
Step 4. Simulate the status of tracks and transponders by the train location and direction information.
Step 5. Determine whether the end of the procedure, if the end of the procedure to end running, or execution Step 2.

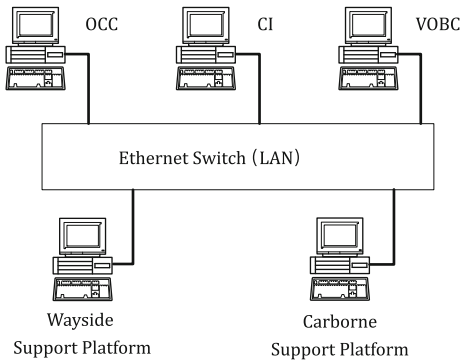


Fig. 101.3 Architecture of OCS simulation support platform

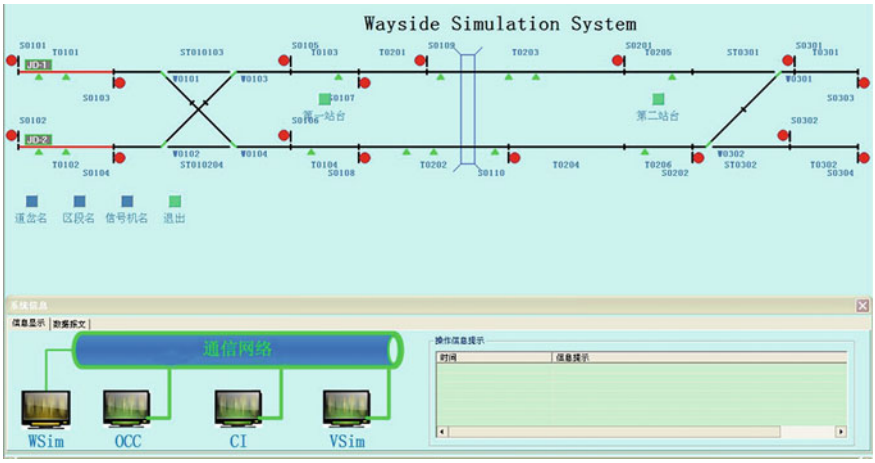


Fig. 101.4 Interface of wayside support platform

101.5 Testing of OCS Simulation Support Platform

Based on lab environment, this paper has accomplished the debugging of OCS simulation support platform, such as interfacing debugging, functions debugging, etc.

As shown in the Fig. 101.3, the peripheral testing environment of OCS simulation support platform consists of OCC, CI, VOBC, communication realized by Ethernet.

Figures 101.4 and 101.5 show interface of wayside support platform and carborne support platform. As indicated in these figures, the interface of support platform consists of following elements: the status of wayside equipment, TOD, driver's control console, interface of fault injection, and data interactions in debugging, etc.

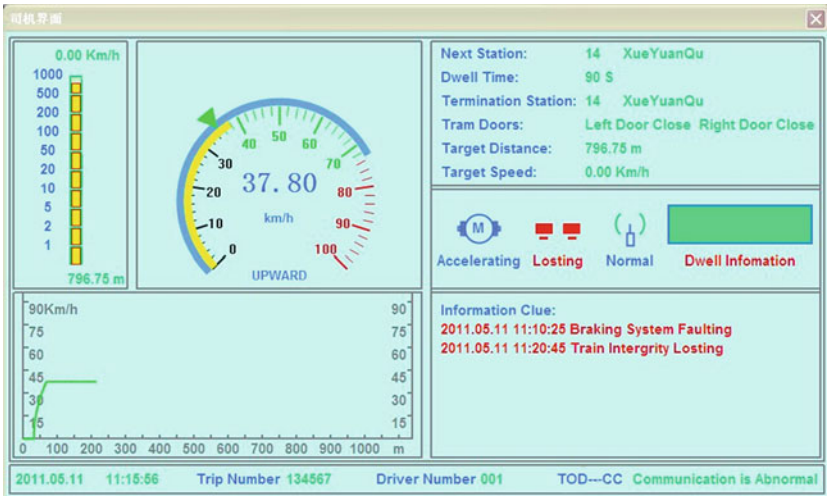


Fig. 101.5 Interface of carborne support platform

101.6 Conclusion

In this paper, by analysis the logic relationship among the subsystems of modern tram OCS, and interface relationship between modern tram OCS and peripheral equipments. For OCS simulation support platform, we first analyze the functional requirements, propose a design proposal, then design structure, and interface. On this basis, OCS simulation support platform software was developed upon the PC and Windows operation system. The test of the support platform showed that OCS simulation support platform effectively realized requirements of OCS. The support platform provides convenient research and testing platform for researchers, with a wide range of applications.

References

1. Xue M (2008) Modern trams: characteristics & development both at home and abroad. Urban Transp China 6(6):88–91
2. Yuan C (2010) Software design of vehicle operation control system for modern trams. Beijing Jiaotong University, Beijing, pp 8–10
3. Liang J (2006) The research on simulation testing platform for communications-based train control (CBTC) system. Beijing Jiaotong University, Beijing, pp 5–6
4. Li P (2007) Development of the braking control system on modern urban tramcars. Roll Stock 45(2):2–3

Chapter 102
Design of Brushless DC Motor Control System Based on Back Electric Motive Force (EMF)

Ran Dong, Hongqi Wang and Weiwei Xue

Abstract This paper discusses the Digital Signal Processor as the core to the sensorless brushless dc motor control system design method. Based on the back EMF zero crossing of brushless dc motor rotor position detection method and double closed-loop speed regulation control strategy, designs brushless dc motor sensorless control system hardware circuit and program. With the simulation and debugging experiments for the system, it shows that the control method not only simplifies the hardware system structure, but makes the motor efficient and stable operation.

Keywords Back-EMF · Brushless DC motor · Control system

102.1 Introduction

Brushless dc motor uses electronic commutation technology instead of mechanical commutation brush for electric motor with new types of mechanical and electrical integration. It has the ac motor simple structure, reliable operation, convenient maintenance, etc., and also the high efficiency dc motor, easy to control speed, good performance without the mechanical change equal advantage [1]. The traditional brushless dc motor has big volume and high cost, in the bad environment will have poor sensitivity and stability. According to position sensor brings the

R. Dong (✉) · H. Wang · W. Xue
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: odinlanna@126.com

H. Wang
e-mail: wanghq2008@foxmail.com

W. Xue
e-mail: 165230489@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_102,

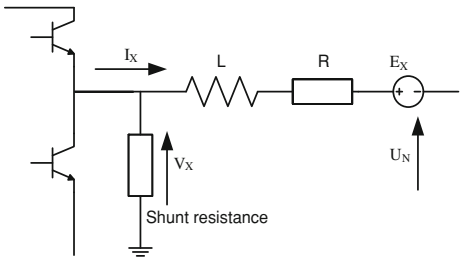


Fig. 102.1 A motor in the equivalent circuit diagram windings

drawbacks, this paper puts forward a no position sensor control technology, not only to overcome the external sensor of many of the abuses, but also widen the brushless dc motor application fields.

In recent years, sensorless control has been a more popular research subject at home and abroad. Some foreign scholars have back-EMF method, free-wheeling diode method [2]. Domestic scholars proposed have stator three harmonic methods, current method, inductive method, etc. [3]. But difficult to implement these methods, application conditions are hard and applies only to specific applications, so the applications is not very extensive. In this paper, the performance is stable back EMF zero crossing method based on DSP, designed of sensorless brushless dc motor control system.

102.2 Back EMF Control Technology

At present back EMF method is relatively commonly used a detection method position signal, the back EMF zero signal detection is actually through comparing the voltage of the three-phase winding to get back EMF indirect zero signal [4]. As shown in Fig. 102.1 is a photograph of the stator winding equivalent circuit. V_X for three-phase winding of the output voltage; The U_N for winding of star connection neutral voltage; L for phase inductance; R for phase resistance; E_X for Back EMF.

According to Fig. 102.1 established three-phase voltage of the balance equation for:

$$V_X = RI_X + L\frac{\mathrm{d}I_X}{\mathrm{d}t} + E_X + U_N \tag{102.1}$$

System by using pairwise switching modes in every moment, only two phase conduction. Set A, B interlinked electricity and C related to break, then A and B two phase equal, in opposite directions, and C phase is zero.

$$\begin{cases} U_A = Ri_A + L\frac{\mathrm{d}i_A}{\mathrm{d}t} + E_A + U_N \\ U_B = Ri_B + L\frac{di_B}{dt} + E_B + U_N \\ U_C = E_C + U_N \end{cases} \tag{102.2}$$

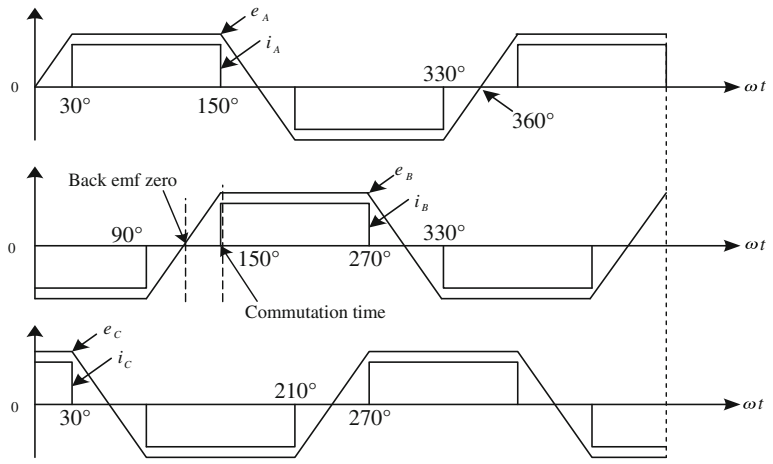


Fig. 102.2 Ideally three instead EMF e and phase I of the wave

By (102.1), (102.2) get:

$$U_N = \frac{1}{3}(U_A + U_B + U_C) \tag{102.3}$$

The (102.3) into (102.2) was:

$$E_C = U_A - \frac{1}{3}(U_A + U_B + U_C) \tag{102.4}$$

Equation (102.4) is back EMF voltage equation *C*. Similarly also can get *A*, *B* is back emf voltage of the equation.

Figure 102.2 for brushless dc motor work in pairwise energized back emf and phase of the waveform. From the figure can be seen the motor turned every 360°, each phase will have zero two back zero crossing [5]. The main idea of back EMF zero testing phase change is through testing the zero positions, delay after the zero crossing 30° inverter allows the motor to the next state.

System Control Strategy

The system uses double closed loop brushless dc motor speed control system. In order to obtain good static, dynamic performance and robustness, the speed loop and current loop all use PI regulator [6] (Fig. 102.3).

In double closed loop speed regulation system, the speed regulator's role is to adjust the speed of immunity and make no difference in the steady state; the output depends on limiting the maximum current allowed. Current regulator's role is to follow the flow from the current, over voltage protection and timely restrain disturbance.

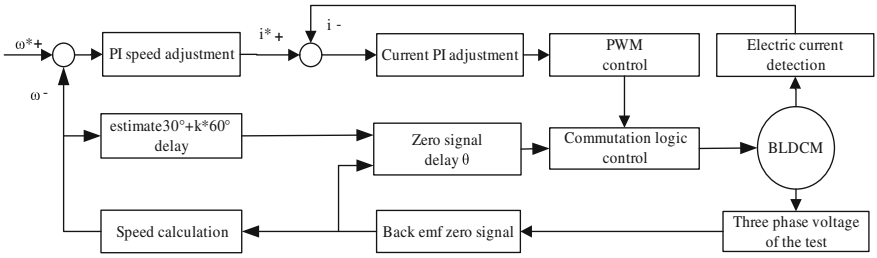


Fig. 102.3 Sensorless brushless dc motor speed control system principle diagram

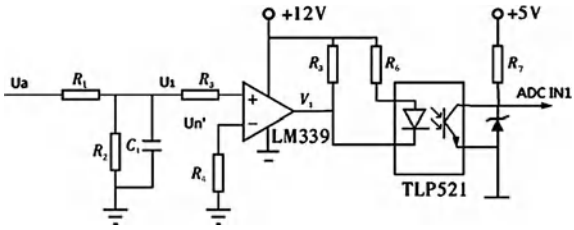


Fig. 102.4 Back EMF detection circuit

102.3 System Hardware

102.4.1 Back EMF Zero Detection Circuit

Because the electromotive force is difficult to directly measure the physical quantities obtained, detection back EMF zero crossing signal is actually detected by the three-phase winding and armature voltage between neutral point voltage comparator to obtain back EMF zero crossing signal [7]. Figure 102.4 for A phase voltage of the EMF detection circuit. Voltage U_a filtered by partial pressure, the zero-crossing detection, optical isolation into ADC IN1 channel of DSP, after the ADC conversion to get the safety voltage, and compared with neutral voltage to obtain the back EMF zero-crossing point, and this in turn trigger the six power tube to control the motor commutation.

102.4.2 Power Drive Circuit

Drive circuit using three-phase full-controlled bridge of control mode, drive chip use IR company MOSFET IR2130 [8]. It is to achieve the control part over the transmission of information processing and commutation PWM signal isolation amplifier, and to control power MOSFET tube conduction and off, so as to control the working state of the motor and speed.

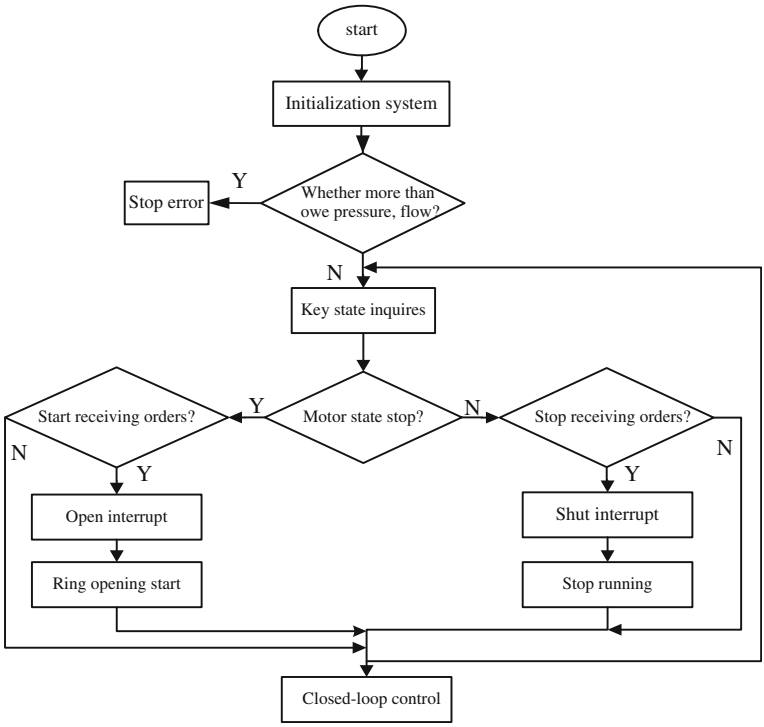


Fig. 102.5 System programming flowchart

102.4.3 Protection Circuit

The main protection system to protect DSP chip from the impact of high voltage and over current, also protects the motor drive circuit from damage. Under-voltage protection is detect input dc voltage, if the system short circuit, when sample voltage is below the set threshold, DSP PWM output pins will be set to high resistance state and blocking the PWM signal output, to protect the circuit motor and power tube purposes. The design of over-current protection is to prevent to motor starting and operation in overload, because too large current to power switch tube and for motor ontology harmful.

102.4 System Software Design

System programming design is divided into the rotor position, the open loop start, and closed-loop control of the three key steps [9].

First of a two-phase inverter turns on, and set the two-phase winding of the transient current, as the initial position of the rotor pole. Then inverter power tube

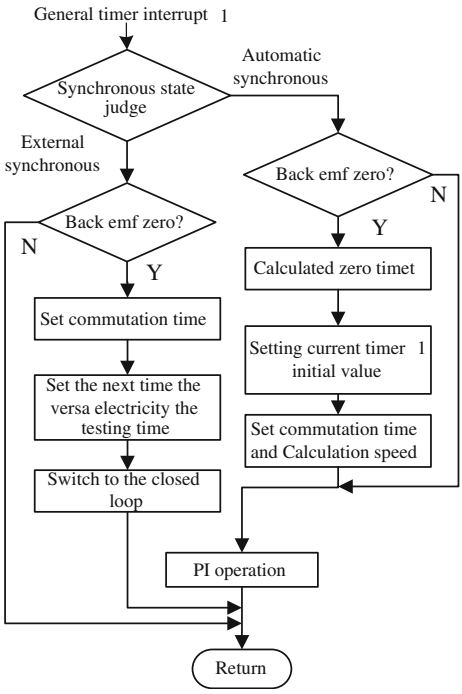


Fig. 102.6 General Timer 1 interrupt flow chart

trigger corresponding conduction, motor running into the open loops synchronous stage. If in stipulated the conditions for two times detection to the potential zero, program will from the synchronous switch to the closed loop control, or stop the synchronous phase error (Fig. 102.5).

When the motor in self synchronous stage, the motor to force change mutually to zero hour late 30° commutation, the current speed by calculating zero twice that time interval. Speed difference from current rotation speed and given speed, according to adjust the get PWM occupies empties compared, realize speed closed-loop, and drive motor stable job.

102.5.1 System Main Program Design

In the main program, perform system initialization and go to the over voltage, over current detection, and then check the button state to enter the start or stop the motor program; Otherwise, stop error occurs. The main program to the user interface cycle operation waiting for interrupt, in the interrupt service routine in position detection interrupt and commutation interrupt.

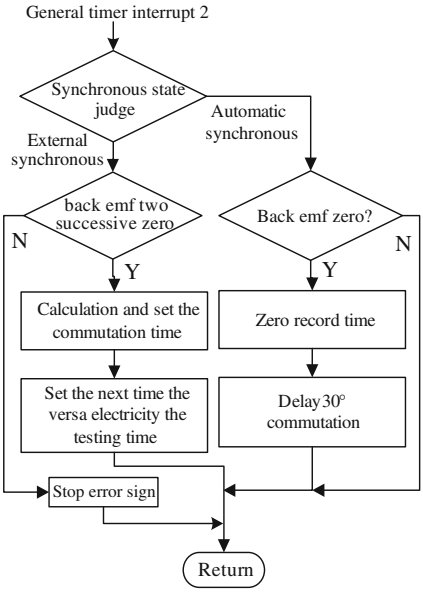


Fig. 102.7 General Timer 2 interrupts flow chart

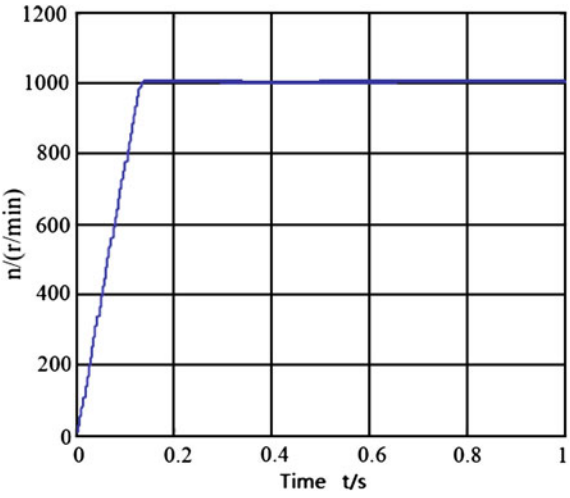


Fig. 102.8 Speed response curve

102.5.2 Position Detection Interrupts Service Routines

Position detection interrupt is responsible for back EMF zero crossing detection and judgment, and set the position of closed-loop commutation time. The interrupt event manager by the TMS320F2812 A general timer 1 interrupt to complete. Figure 102.6 for General timer 1 interrupts flow chart.

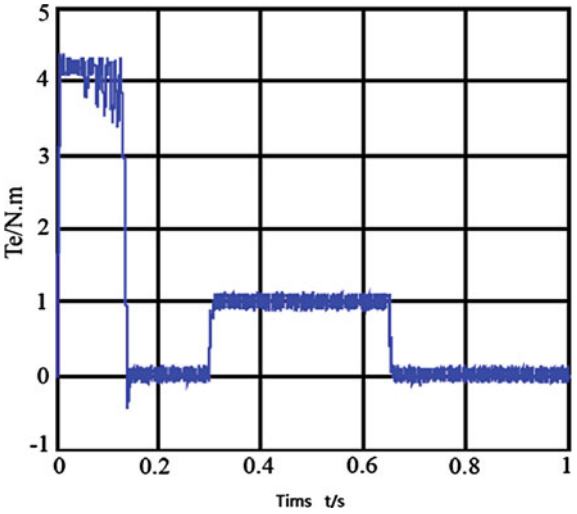


Fig. 102.9 Torque response curve

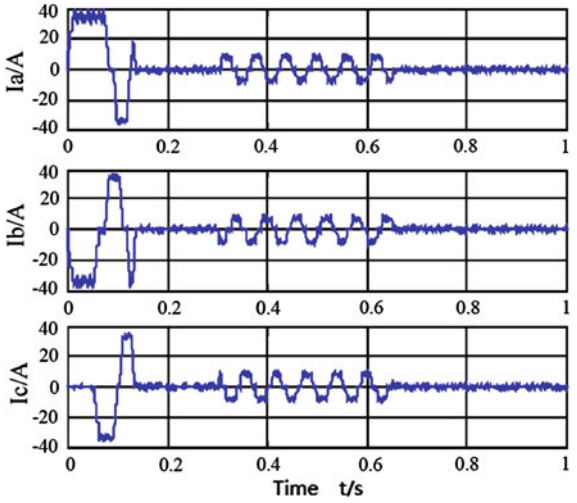


Fig. 102.10 Three phase current waveform

102.5.3 Commutation Interrupts Service Routines

According to the system design, commutation function by general timer 2 interrupt completed. The external synchronizing stage, general timer 2 as forced by change mutually controller, interrupt its cycle as commutation, to set the commutation time; since the synchronous stage, general timer 2 measurement two zero interval, and 30° angle delay timer control electric change, and now it is triggered signal by general timer 1 is match suspensions. Figure 102.7 for General timer 2 interrupt flow chart.

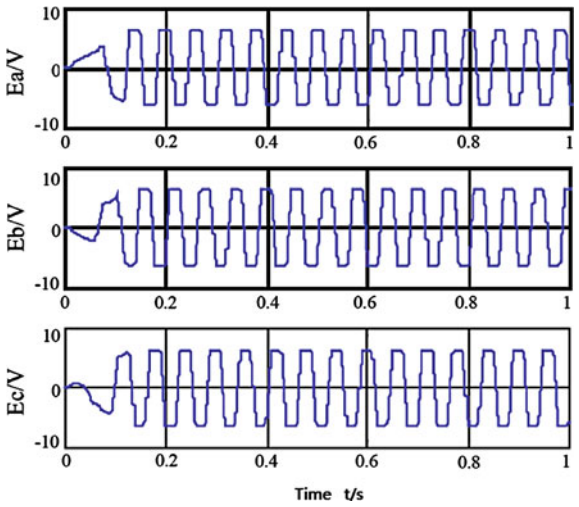


Fig. 102.11 Three-phase back EMF waveform

102.5 Simulation Results

Based on MATLAB/Simulink establish the double closed loop control system simulation. Simulation of brushless dc motor parameter is set to: the rated speed $n = 1{,}500$ r/min, rated voltage $U_N = 220$ V and rated current $I_N = 15.6$ A, the stator resistance $R = 1\ \Omega$, stator winding self-induction $L = 0.02$ L, mutual inductance $M = 0.061$ H, Moment of inertia $J = 0.005$ kg m^2, damping coefficient $B = 0.0002$ N M s/rad, Pole pairs $p = 1$. The system No-load start, when to enter the steady-state, in $t = 0.3$ s suddenly join load $T_L = 5$ N m, the t = 0.65 s remove load. Figures 102.8, 102.9, 102.10, 102.11 shows is the system speed, torque, three phase and three instead EMF simulation curve. Can be seen from the simulation waveform, the system response quickly and smoothly, phase and back EMF waveform more ideal, show good dynamic follow and stability. Figures 102.8, 102.9 show that the start stage the rotating speed is constant and not to cause major phase and torque impact, the current limiting function reference is quite effective.

102.6 Conclusion

Based on DSP chip TMS320F2812 is the core design of BLDCM control with sensorless system, and introduces the main control system of the software and hardware design in detail. Through the modeling simulation, this system can realize BLDCM starting and smooth operation,and the system operation speed is achieved.

Reference

1. Wang Y (2004) The back EMF method brushless dc motor control system research. Southeast University
2. Li W, Jiang K (2007) A brushless DC motor control system without the position sensor based on DSP. Instrumentation Technology, vol 9
3. Wang C, Zhang Q, Xiong J (2006) Design of control system of brushless DC motor based on DSP. Control Autom 7:6–7
4. Meng T (2008) Sensorless brushless DC motor control system. China Science and Technology University
5. Xie B, Ren Y (2005) Motor DSP control technique and its application. Beijing aerospace university press, Beijing
6. Shuang J, Huang Z, Chen Q, Xu G (2008) Control system based on DSP for BLDC motor with position sensorless. Electr Drive Locomot 2:41–43
7. Haoxin Zheng, Bin Zheng, Jianmin Li (2008) Design of control system of brushless DC motor based on TMS320F2812. Int Electron Elem 11:17–19
8. Lin M, Zhang Z, Lin K, Zhou G (2009) Anti-reverse rotation startup method for sensorless brushless DC motor. J Southeast Univ 25:99–202
9. Shao J, Nolan D (2005) Further: improvement of direct back-EMF detection for sensorless brushless DC (BLDC) motor drives. Appl Power Electron Conf Expo 2:933–937

Chapter 103
Analysis of Influence Factors of SRM Torque Ripple on Current Chopping Control Condition

Longsheng Yuan, Chaohui Zhao, Di Zhang, Jian Li, Feng Xiang, Shuhao Cao, Lei Chen and Huaping Li

Abstract In this paper, the influence factors of SRM torque ripple is analyzed using finite element calculation and magnetic circuit method, on condition of current chopping control. The air gap length, stator and rotor pole arc embrace and stator and rotor yoke thickness variations of switched reluctance torque ripple changes are discussed; simulation results show that: ① to the air gap length of the smaller torque ripple there is an optimum value; ② stator and rotor pole arc embrace of SRM torque ripple has an impact, turn increases the average torque with the stator and rotor pole arc embrace increasing, stator and rotor pole arc embrace in smaller numeric, torque ripple changes little, but when stator and rotor pole arc embrace are greater than certain value, with the stator and rotor pole embrace increasing, torque ripple increased largely. ③ stator and rotor yoke thickness on the SRM torque ripple has an impact, but little impact. With the stator yoke thickness increasing SRM torque ripple gets smaller; with the rotor yoke thickness increasing, SRM torque ripple gets larger.

Keywords SRM · Current chopping control · Torque ripple · Influencing factor · Finite element analysis

103.1 Introduction

Switched reluctance motor (SRM) and its drive system for simple, reliable, efficient and low cost has been widely used in the field of electric transmission for recent years. However, due to double salient structure of SRM and high torque

L. Yuan (✉) · C. Zhao · D. Zhang · J. Li · F. Xiang · S. Cao · L. Chen · H. Li
Shanghai Dianji University, Shanghai 200240, China
e-mail: yls198825@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_103,

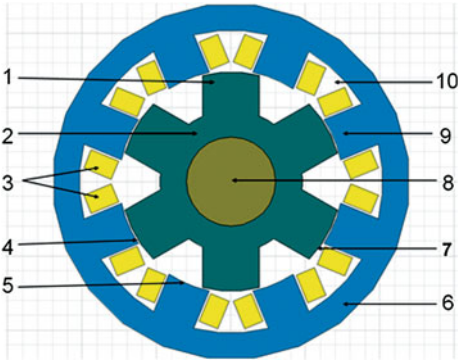


Fig. 103.1 The structure of SRM (*1* Rotor tooth *2* Rotor yoke *3* Stator winding *4* Air gap *5* Stator pole arc *6* Stator yoke *7* Rotor pole arc *8* Shaft *9* Stator tooth *10* Stator slot)

ripple, it limits its application for further use in control field [1]. Therefore, researching the torque ripple of SRM becomes a central issue, the literature [2] uses the current amplitude chopping control strategy to reduce the torque ripple. In [3] the concept of SRM torque vector control and the control strategy to suppress the torque ripple of switched reluctance motor is proposed; the literature [4] increases the width of the stator salient pole and the stator salient pole number, to reduce the torque ripple of SRM; the literature [5] used nonlinear method to calculate and analyze structural parameters of the motor under controlled conditions in the angular position of the torque ripple fluctuations.

As SRM in the constant torque region has lower speed and lower EMF, to avoid excessive current peaks chopped current control is often used; however, in the current chopping control condition,analyzing the influencing factors of SRM torque ripple did not see the report. Therefore, this paper analyzes the influencing factors of SRM torque ripple using finite element calculation and magnetic circuit method, on condition that current chopping control.

103.2 The Structure and Characteristics of SRM

From Fig. 103.1, SRM Includes salient stator and salient rotor, field winding on the stator and no windings or permanent magnet on the rotor, the only source of magnetic potential from the stator windings, this is a significant advantage of SRM. As the stator and rotor are unique salient structures, they make SRM's structure simple and stable, high reliability, fault tolerance ability, speed performance and suitable for high-speed or ultra high-speed operation. But, because of the double salient structure of SRM, the principle in work is different from that of traditional AC motor, its electromagnetic field analysis is more complex, nonlinear problem is serious and has large torque ripple and noise.

Table 103.1 The structural parameters of SRM

	Poles	Outside diameter/mm	Inside diameter/mm	Yoke thickness/mm	Pole embrace
Stator	8	120	75	9	0.5
Rotor	6	74	30	9	0.5

Table 103.2 The key performance data of SRM

Rated output power 0.55 kW	Rated voltage 300 V
Rated speed 2500 r/min	Wind friction loss 12 W
Operating temperature 75	Circuit type full-voltage

103.3 Analysis and Calculation Methods of the Motor Magnetic Field

Build differential equation of machine's interior field to ensure the region of account and the boundary condition of FEA account [6]

(1) End-effects are negligible, the machines field is uniformly distributed in axial direction, viz: the current density vector J and vector magnetism potential A has only axial component, J = Jz, A = Az;
(2) Iron core material is isotropy, and magnetization curve is humdrum, namely, the effect of hysteresis is negligible;
(3) The magnetic field is negligible in the outer shell of machine and shaft, that is: outer surface circumference of stator and the inner surface circumference of rotor are zero vector potential surface;
(4) The effect of eddy current is negligible;
(5) The magnetization of each rare-earth magnet is uniform. To rare earth permanent magnet materials, it means the remanent magnetization M is independent on work point. In addition, if alnico has uniform magnetization, its magnetism would maintain all the time, as stuffs inner coercive force HC is very big;
(6) The magnets are modeled with parallel magnetization.
From the engineering view, the outer diameter of stator is a boundary of magnet field [7]. So the calculation model of flux is shown in Fig. 103.1. For the magnetic field analysis and calculation of the SRM structure parameters are shown in Table 103.1. Switched reluctance motor is the structure of 8/6 stator/rotor pole, power converter is Full-Voltage, SRM working in the motor state, Stator and rotor steel sheet used Steel_1010, Diego of pressure coefficient is 0.95. When the system is in the simulation, current use of chopper control (Chopped Current Control, referred to as: CCC), Max Current for CCC) is 25 A, Min Current for CCC is 0 A, Opening angle is set to 0 degrees (mechanical angle, the same below) and closing angle is set to 22.5 degrees (refer Table 103.2).

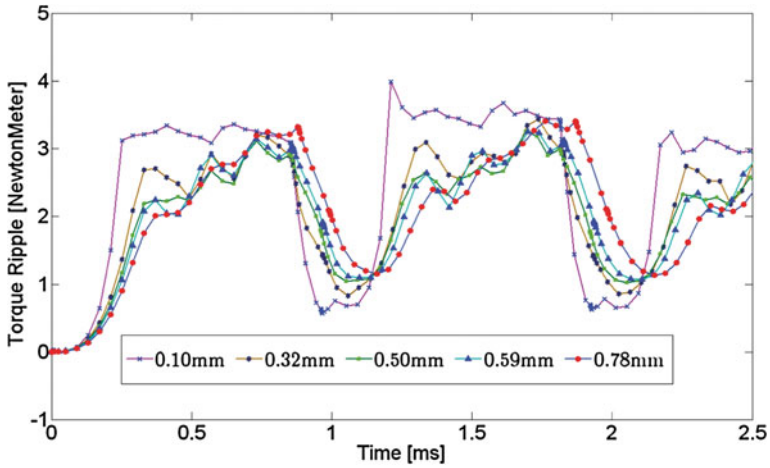


Fig. 103.2 Air gap length on the impact of SRM torque ripple

103.4 Analysis of The Influence Factors of SRM Torque Ripple

103.4.1 The Influence of Air Gap Length's Variety on SRM Torque Ripple

As seen from Fig. 103.2, when the air gap length is from 0.10 to 0.50 mm, the gap length increases, and SRM torque ripple decreases; however, when the gap length is from 0.50 to 0.78 mm, the torque ripple increases, less torque ripple corresponding to the air gap length, there exists an optimal value of 0.50 mm.

We can see from Fig. 103.2, It is not much higher for the average of the SRM torque. The air gap length is 0.1 mm, when compared with 0.5 mm. This is because, by reducing the air gap length, the permeance of air gap increases, and the main flux of motor increases. However,there is a high degree of local saturation in the core teeth, when the main magnet fluxed through the core teeth. With the main magnetic flux increasing, it will add saturation leading to local core permeability decrease. This makes the motors' main magnetic flux not to increase with the air gap length's proportional decrease; at the same time, magnetic flux leakage between the stator and rotor increases with proportional decrease in the air-gap length. The main enhancement of magnetic flux leakage effects are that, when the air-gap length is reduced to a certain value, the magnetic leakage flux of the main changes with the air-gap length becoming very small.

Figure 103.3 shows the magnetization curves of SRM when air gap length are 0.1 and 0.5 mm. As the air gap length is small, the motor iron core saturates in advance, and the motor with the load capacity decreases.

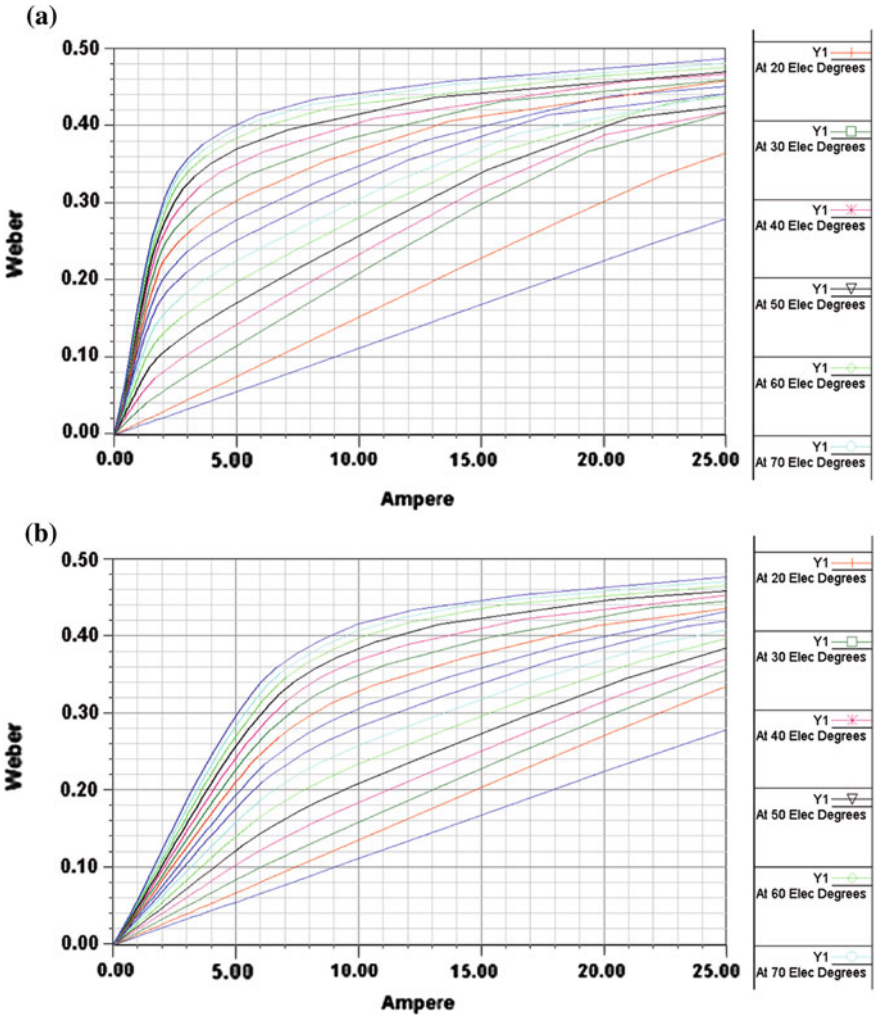


Fig. 103.3 The magnetization curve of SRM when air-gap lengths are 0.1 and 0.5 mm. **a** 0.1 mm, **b** 0.5 mm

103.4.2 The Influence of Stator and Rotor Pole Embrace Variety on SRM Torque Ripple

The stator and rotor tooth width of SRM changes with the stator and rotor pole embrace coefficient. When the stator and rotor tooth width changes, the motor output regional differences, and furthermore, the influence of the motor torque ripple and average torque changes.

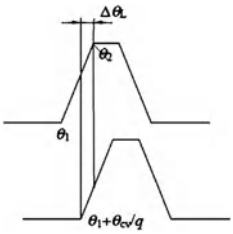


Fig. 103.4 Overlapping angle of inductance

(A) Determine the stator and rotor embrace coefficient range

As SRM is double salient structure, the stator and rotor pole embrace choice for the performance of the motor is very important. Stator and rotor pole embrace should be chosen to meet [8]:

(1) The motor at any rotor position has positive and negative direction of the self-starting capability.
(2) Reduce the minimum inductance of per phase winding when the stator and rotor is not aligned.
(3) Reduce the mutual inductance between the windings of each phase.
(4) Minimize the switching frequency of each phase winding.

In order to ensure that the motor at any rotor position has positive and negative direction of the self-starting capability, Requires adjacent phase stator and rotor pole to overlap when a phase of stator and rotor pole is in the relative position. The required overlap between two adjacent ascending phase inductance curves is shown in Fig. 103.4:

θ_{cy} is rotor displacement angle for torque angle characteristic of a cycle, $\theta_{cy} = \frac{2\pi}{Nr}$, torque angle characteristics of each phase is θ_{cy}/q ($q = N_s/2$), Therefore, the overlap angle is $\Delta\theta_L$:

$$\begin{aligned} \Delta\theta_L &= \beta_s \frac{-2\pi}{Nr} \quad (\beta_r > \beta_s) \\ \Delta\theta_L &= \beta_r \frac{-2\pi}{Nr} \quad (\beta_s > \beta_r) \end{aligned} \tag{103.1}$$

Where β_s is pole arc of stator poles and β_r is pole arc of rotor poles. So to meet the necessary conditions for the pole embrace:

$$\begin{cases} \min(\beta_s, \beta_r) & \geq \frac{2\pi}{Nr} \\ \beta_s + \beta_r & \leq \frac{2\pi}{Nr} \end{cases} \tag{103.2}$$

When the Eq. (103.2) not satisfied, SRM has only one direction of the self-starting capability, start turning of the motor depends on the position of the rotor varies.

In order to facilitate the simulation analysis, the following uses the concept of pole embrace. Pole embrace is defined as [9]: pole arc width b_p divided by the

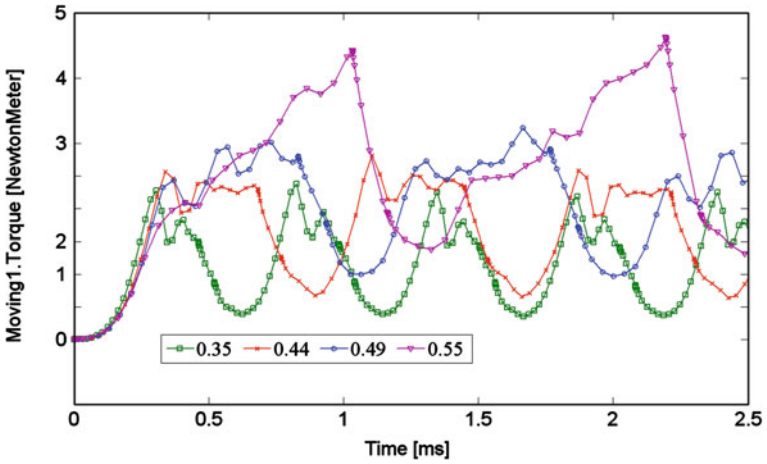


Fig. 103.5 Stator pole embrace on the impact of SRM torque ripple

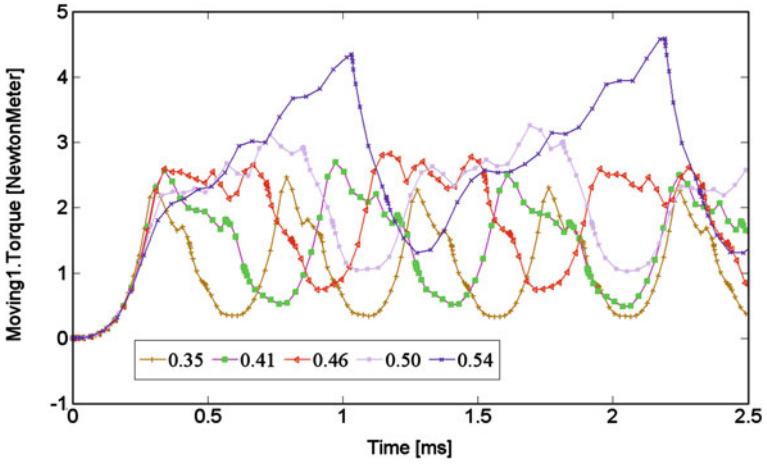


Fig. 103.6 Rotor pole embrace on the impact of SRM torque ripple

distance between the two poles τ. The following analysis also uses 8/6 stator/rotor pole SRM, so stator pole angle is 45° and rotor pole angle is 60°.

1. To determine the scope of stator embrace
   In the analysis of the stator embrace of variation, the rotor pole embrace selection has little effect on the stator analysis [10]. Select the rotor pole embrace as 0.50. At this point: $\beta r = 30°$, satisfy (2) the conditions of the $\beta s \leqq 30°$. Assumptions :$\beta_s = 30°$, to maximize the stator pole embrace:$\alpha_{\max} = 0.67$, be the smallest of stator pole embrace: $\alpha_{\min} = 0.33$.

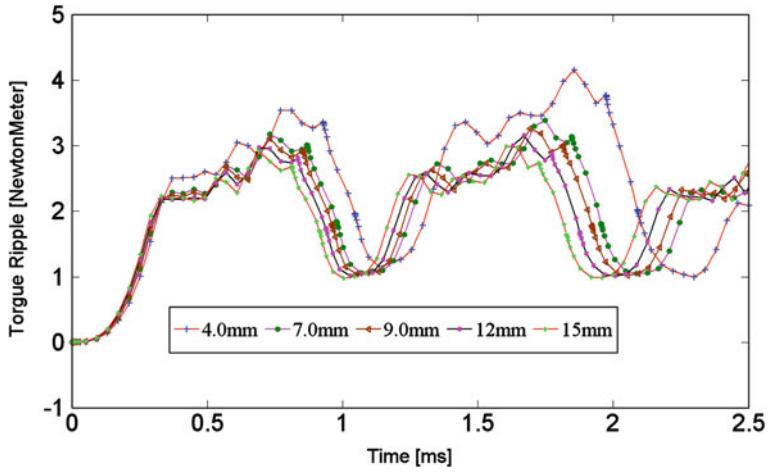


Fig. 103.7 Stator yoke thickness on the impact of SRM torque ripple

2. To determine the scope of rotor embrace
   In the analysis of the rotor embrace of variation,the stator pole embrace selection has little effect on the stator analysis [10]. Select the stator pole embrace as 0.50, namely:$\beta s = 22.5°$, the maximum rotor pole embrace is $\alpha\alpha_{\min} = 0.25_{\max} = 0.625$; the minimum rotor pole embrace is $\alpha_{\min} = 0.25$.

(B) The influence of stator pole embrace Variety on SRM torque ripple

Figure 103.5 is the curve of SRM torque ripple with the change of the stator pole embrace.

Know from Fig. 103.5, with the SRM Stator pole embrace increasing, the average torque is also increasing. However, torque ripple is almost unchanged. When the stator pole embrace is greater than 0.49, along with the increase of the average torque, torque ripple significantly increased. Therefore, considering the motor torque ripple and average torque, when the stator pole embrace is around 0.49 the motor performance is best.

(C) The influence of rotor pole embrace variety on SRM torque ripple

Figure 103.6 is the curve of SRM torque ripple with the change of the rotor pole embrace.

Know from Fig. 103.6, rotor pole embrace and stator pole embrace to the torque ripple of SRM is similar, with the SRM Stator pole embrace increasing, the average torque is also increasing; however, when the rotor pole embrace is greater than 0.49, along with the increase of the average torque, torque ripple significantly increased. Therefore, considering the motor torque ripple and average torque, when the rotor pole embraces around 0.49 the motor performance is best.

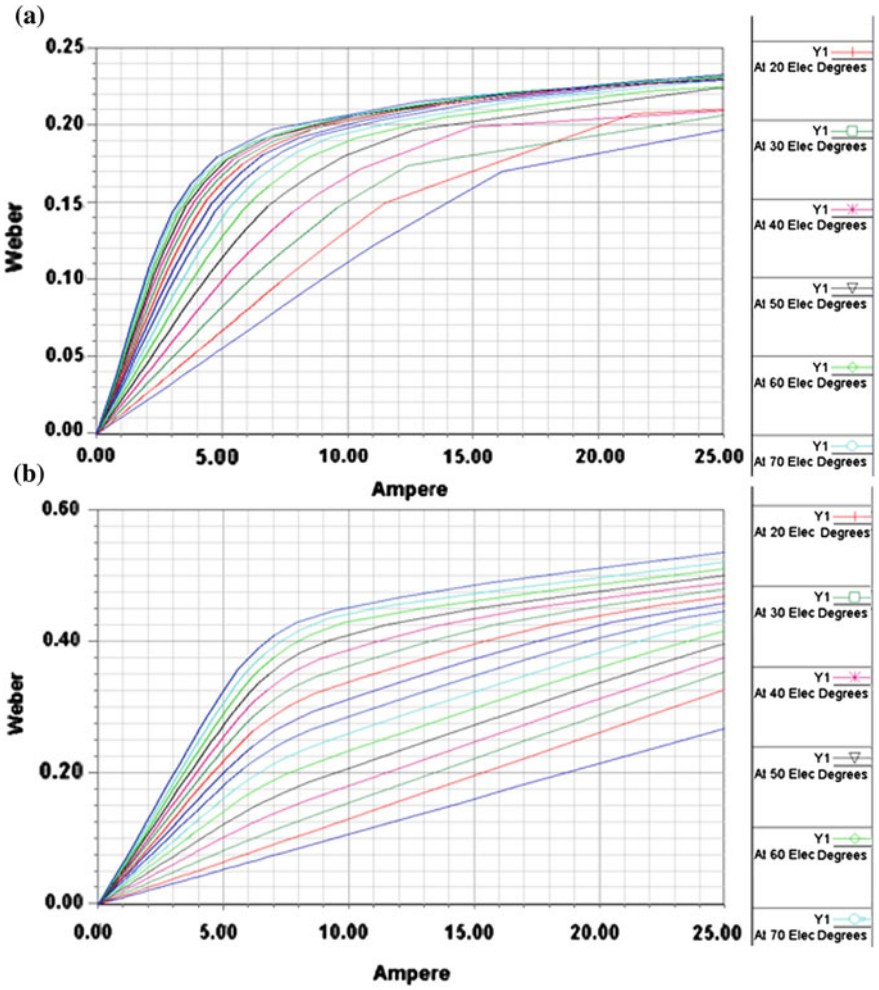


Fig. 103.8 The magnetization curve of SRM when stator yoke thickness are 4.0 and 15 mm. **a** stator yoke thickness is 4.0 mm. **b** Stator yoke thickness is 15 mm

103.4.3 The Influence of Stator and Rotor Yoke Thickness Variety on SRM Torque Ripple

(A) The influence of stator yoke thickness variety on SRM torque ripple

Figure 103.7 shows that stator yoke thickness impact on SRM torque ripple. Stator yoke thickness on the SRM torque ripple effect is relatively small.

Figure 103.8 shows that the magnetization curve of SRM when stator yoke thicknesses are 4.0 and 15 mm. When the stator yoke thickness is small, the motor of the saturation point is in advance, thereby affecting the ability of the motor with

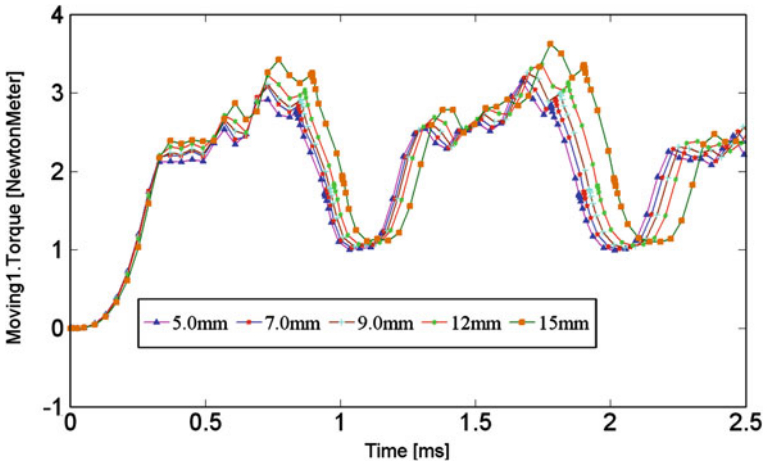


Fig. 103.9 Rotor yoke thickness on the impact of SRM torque ripple

a load. When the stator yoke thickness is large, although the motor torque ripple is relatively small, the load capacity is improved; however, the motor rated power to the timing and the stator yoke thickness increases will lead to the shorter stator teeth and smaller motor slot space, leading to losses in the motor increases and reducing the motor efficiency. so during the motor design, not only the pursuit of a single low torque rippled but also the average torque and other performance factors were ignored.

(B) The influence of rotor yoke thickness variety on SRM torque ripple

When the SRM power rating and the first air-gap length is fixed, SRM rotor yoke thickness on the the second air-gap length of SRM changes, thereby affecting the SRM motor minimum inductance,and thus affecting the rate of change of rotor position and the SRM torque and torque ripple. The second air-gap length on the SRM, theoretically hoped that the second air-gap length will be as large as possible. This will be the minimum inductance, and increase the angle change rate for inductance with the rotor position, and increase the average torque of SRM.

Figure 103.9 shows that rotor yoke thickness impacts the SRM torque ripple. With the increase of the thickness of the rotor yoke, SRM torque ripple increases, considering the average torque and torque ripple. Rotor yoke thickness of about 9.0 mm is excellent for the motor performance; on the whole, rotor yoke thickness on the SRM torque ripple effect is not large, taking into account the structural properties of the motor and the second air-gap length ratio of motor tooth width to a certain extent. Therefore, the rotor yoke thickness is selected to moderate.

103.5 Conclusion

(A) Using finite element calculation and magnetic circuit method, analyzing the influence factors of SRM torque ripple on current chopping control condition.
(B) To the air-gap length of the smaller torque ripple there is an optimum value, and the air-gap length on the motor with a load capacity also has a certain impact.
(C) Stator and rotor pole arc embrace of SRM torque ripple has an impact. Turn increases the average torque with the increase of stator and rotor pole arc embrace. When stator and rotor pole arc embrace is in smaller numeric, torque ripple changes little, but when stator and rotor pole arc embrace are greater than certain value, with the stator and rotor pole embrace increasing, torque ripple increased largely.
(D) Stator and rotor yoke thickness on the SRM torque ripple has an impact, but little impact. With the stator yoke thickness increasing SRM torque ripple gets smaller; with the rotor yoke thickness increasing, SRM torque ripple gets larger.
(E) Analyzing different structural parameters on the impact of SRM torque ripple in the current chopping control conditions, the specific design and engineering of the application of SRM has some significance. They have a certain significance for specific design and engineering applications of SRM.

References

1. Wen D, Zhou H, Yu Z (2006) Torque ripple in telligent minimization of switched reluctance motor. Micro-motor 39(3):7–11
2. Wang X, Wang X, Wang Y et al (2000) Double amplitude chopping control of switched reluctance motor. Proc CSEE 20(4):83–86
3. Zheng H, Jiang J (2002) Control of torque ripple of a switched reluctance motor based on torque vector control. Micro motor 1:15–17
4. Husain I (2002) Minimization of torque ripple in SRM drives. In: IEEE Trans Ind Eleetron on vol. 49, no.l, pp 28–39
5. Wu J, Zhan Q, Lin J (1993) Structural parameters of switched reluctance motor effect on the torque ripple. Electrotechn J 2(3):3–5
6. Zhao C, Li S, Wang X et al (2005) Influencing factor analysis of pmsm air gap flux density. Henan Agricultural University 39(3): 338–344
7. Zhao C, Zhang Z, Qin H (2010) The structure and principle of hybrid excitation motor. Science Press, Beijing
8. Wu J (2000) Switched reluctance motor design and application. Mechanical Industry Press, Beijing
9. Liu D (1992) Aviation electrical machinery. Aviation Industry Press, Beijing
10. Jian S .6/2 structure of high-speed switched reluctance motor and its implementation. [Master Thesis]. Nanjing University of Aeronautics and Astronautics

Chapter 104
Research on Control System of Belt Conveyor in Coal Mine

Li Wang and Li Zhigang

Abstract The higher demands for control system of the belt conveyor in coal mines were analyzed and a control system based on PLC was developed. The control strategies of speed regulating and power balance of multi-motor drive were discussed in detail. Considering speediness and stability of speed regulation, a control strategy with dead band was proposed. When the error is greater than the given parameter, fuzzy-parameter PID control is adopted to stable the system quickly. When the error is smaller than the parameter, fixed-parameter PID control is adopted to reduce running time. The power balance of the motors is necessary for the normal running of the belt conveyor system. A tracking method is put forward for the power balance of the motors. One motor is selected as the major motor and the other motors are tracking motors. The speed of major motor is controlled to be in accordance with the speed curve, and currents of the other motors are regulated by using fuzzy PID control. These strategies were applied in the developed system and achieved good results.

Keywords Belt conveyor · Multi-motor drive · Speed regulating · Power balance

L. Wang (✉)
College of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: wangli@hpu.edu.cn

L. Zhigang
Jiaozuo Power Plant, Jiaozuo 454000, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_104,

104.1 Introduction

Belt conveyor is one of main transport equipments in coal mine and is widely used in level workings, uphill, downhill and main haulage roadway. The belt conveyor offers advantages such as low transportation cost, large carrying capacity, free of topographic condition and convenient maintenance. With the development of mine exploit, the belt conveyor is moving in the direction of long distance, large capacity, high power, high speed and multidrive. At present, the longest distance of one single belt conveyor has achieved 30.4 km, which is used in an aluminum ore of Australia. A large conveyor belt with the capacity of 37,500 t/h and speed of 7.4 m/s has been applied in one open pit mine of Germany [1]. Compared with foreign countries, the belt conveyors used in coal mines of China have smaller capacity and lower speed. Greater disparity exists in the aspects of dynamic design, computer monitoring and intelligent level.

With the rapid development of China's coal industry, higher demands are raised for the control system of the large belt conveyor. These demands include precise control of startup, shutdown and speed regulation, comprehensive state detection and fault diagnosis, good expansibility, etc. [2]. This study focuses on the control system of the belt conveyor and presents feasible control strategy for speed regulation and power balance of multimotors.

104.2 Architecture of Electric Control System of Belt Conveyor

The belt conveyor system consists of belt unit, mechanical drive system and control system. The belt unit is composed of belt, carrier roller, tension apparatus, turnabout drum and driving drum. The mechanical drive system is composed of gear retarder, planetary transmission mechanism and motors. The control system is composed of hydraulic system and electric control system [3].

The electric control system of the belt conveyor consists of two parts: control part and detection part. Figure 104.1 shows the block diagram of the electric control system. The major task of the electric control system is to control hydraulic couplers and electromotors. The programable logic controller (PLC) FX2 N64MT is used as the control core of the system. The detection functions are realized by the various sensor modules. These modules include speed detector, deviation detector of the belt, temperature measurement module, current detector of the motors, smoke detector and coal level detecting module. The detection signals (I/O and analog) are processed and sent to the input modules of the PLC. The output modules of the PLC are configured to control the driving motors, electro-hydraulic regulating valve, sprinkler, status indicators, etc. The PLC system displays information on the spot by F940GOT, a graphic operation terminal. The module FX2N-485-BD is used for communication with upper computer, which completes the task of monitoring and fault diagnosis.

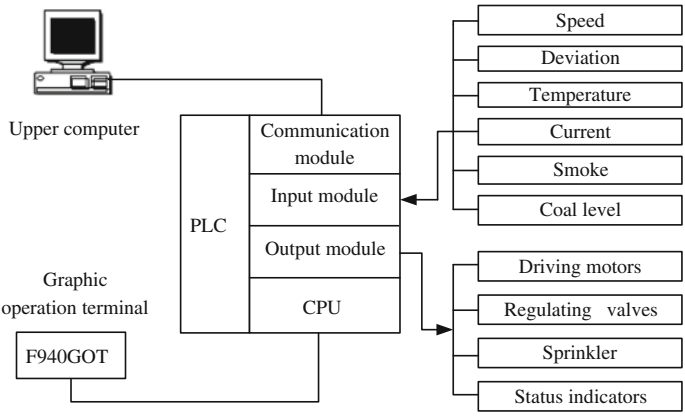


Fig. 104.1 Block diagram of control system of belt conveyer

104.3 Control Strategy of Speed Regulating

The PID control is widely applied to the industrial process because it is simple, reliable, stable and easy to implement. To this day, more than 90% of control loops adopt PID control. The conventional PID control relies too heavily on the model of the control object and robustness of the parameters is worse [4]. The control system of the belt conveyor is a nonlinear and time-variant uncertain system. So it is difficult to set up the accurate mathematical model of the system. Good control results cannot be achieved by using the conventional controller. In order to realize the speed regulating, fuzzy self-tuning parameter PID control is adopted [5]. Figure 104.2 shows the structure diagram of the control system.

The advantage of the fuzzy control lies in fast adjustment, which enhances the dynamic performance but cannot improve the steady-state characteristics. If control action is too frequent, the oscillation is likely to occur and the motor shaft is easy to be damaged. A control algorithm with dead band is a good solution. A fixed parameter is set firstly, then different control action is taken according to the relationship between the error and the parameter. When the error is greater than the parameter, fuzzy-parameter PID control is adopted to stable the system quickly. When the error is smaller than the parameter, fixed-parameter PID control is adopted to reduce running time of the control program, which can improve the execution efficiency. The control strategy is shown as the following formula.

$$|e(k)| - |e_0| = \begin{cases} > 0 & \text{fuzzy-parameter PID} \\ \leq 0 & \text{fixed-parameter PID} \end{cases} \tag{104.1}$$

where $e(k)$ is the error and e_0 is the dead band.

When $|e(k)| > |e_0|$, fuzzy-parameter PID control is taken. When $|e(k)| \leq |e_0|$, fixed-parameter PID control is taken. If e_0 is too small, the control action will be

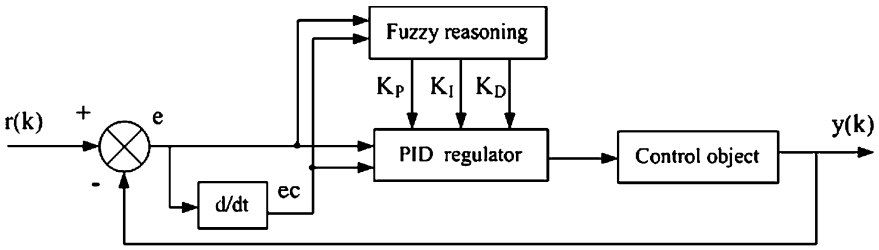


Fig. 104.2 Fuzzy self-tuning parameter PID control system

too frequent and the control object is hard to be stabled. If e_0 is too large, the longtime delay will occur.

Real-time application of fuzzy control is realized by the programs in the controller. In this study, the fuzzy control is finally realized by the programs of the PLC. There are two ways to perform the fuzzy control, one is online algorithm, another is offline algorithm. In online algorithm, sampling, fuzzification, fuzzy identification, defuzzification and output are conducted successively in real time. In offline algorithm, the fuzzy rules are written into the controller in the form of table. The task of fuzzy treatment is completed in offline mode. In the real-time control, the processing steps include sampling, table look-up and output [6]. In order to simplify the programing and improve real-time performance, the second way is adopted.

104.4 Research on Power Balance of Multi-Motor Drive

The large belt conveyor has long distance and large capacity, which demands great driving power. Within the allowable range of tension for the belts, one single motor cannot produce enough driving force. Multi-motor driving mode is adopted in the large belt conveyor, which can reduce the requirements of belt strength and power voltage. Below good case, the distribution ratio of power is same as the ratio of driving force among motors. But in reality, the distribution of actual power has greater deviation due to the influence of a variety of factors. Thus, the tensions of the belts are redistributed and the powers of the motors are out of balance. Seriously, the motors are overloaded or even damaged. The power balance of the motors is necessary for the normal running of the belt conveyor system [7].

The essence of the power balance is to guarantee that the ratio of actual power approximately equals the designed ratio. The balance is realized by the speed regulating. The speed of the motor indicates the load of the motor. If the motor is overloaded, the hydraulic motor is lubricated to drive the inner gear ring to rotate. So the drum is slowed down and the load of the motor is reduced accordingly. During the process of the belt conveyor, the power of the motor is:

$$P = \sqrt{3}UI\eta\cos\varphi \tag{103.2}$$

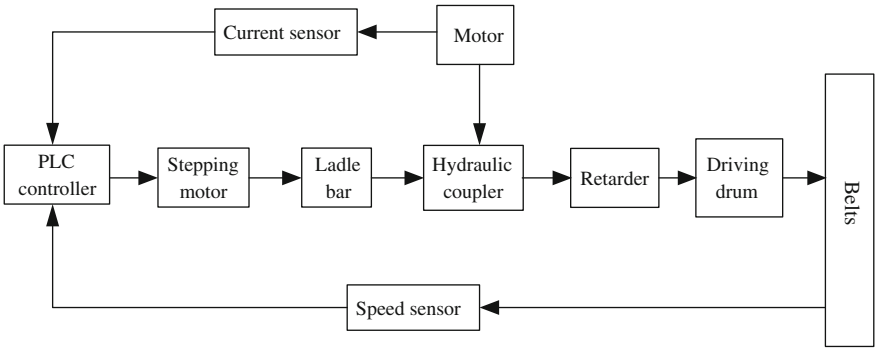


Fig. 104.3 Block diagram of the control system for power balance

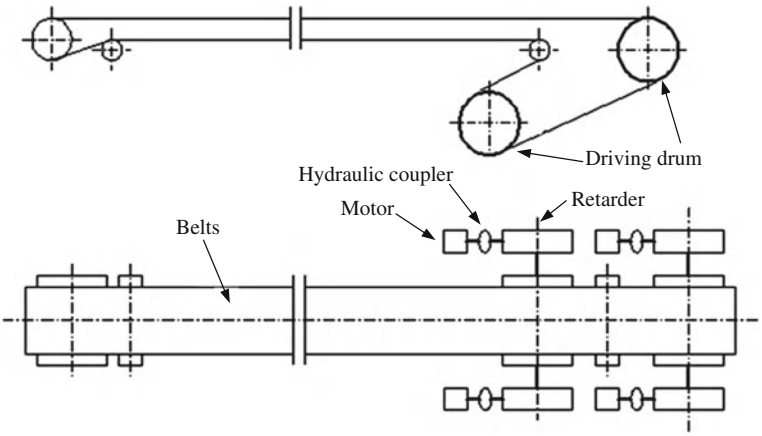


Fig. 104.4 Disposal of driving devices in belt conveyer

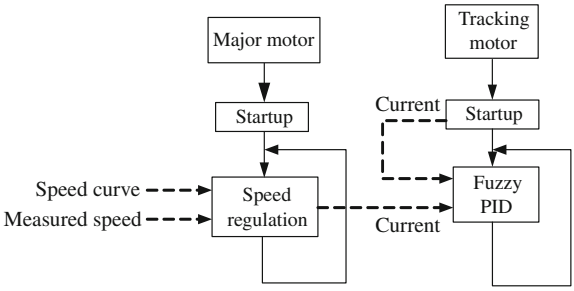


Fig. 104.5 Flow chart of power-balance

where U, η and $\cos\varphi$ remain unchanged for the same type of motors and the same power system. So the current of the motor reflects its power. The power adjustment is conducted according to the load current. The controller controls position of ladle bar in hydraulic couplers to reach power balance. For example, if the load current

is larger, the ladle bar is moved forward and the flux of the hydraulic coupler is reduced. Finally, the output power of the motor is reduced. Figure 104.3 shows the block diagram of the control system for power balance, in which the speed of the belts and the current of the motor are chosen as the control quantities. After analyzing reference current, load current and the speed, PLC controls the stepping motor to regulate the position of the ladle bar. Thus, the aim of balancing power distribution is achieved.

In this study, the belt conveyor has two driving drums and each drum is driven by two motors. The disposal of driving devices in belt conveyor is shown in Fig. 104.4. The control objects are four motors. Obviously, it is a multiple-input-multiple-output system and the four motors have strong coupling with each other.

After analysis, a tracking method is chosen for the power balance of the four motors. A motor is selected as the major motor and the other three motors are tracking motors. In order to simplify the problem, the power balance is not considered during the process of startup. After the startup is completed, the programs of power balance are invoked. Specifically, the speed of major motor is controlled to be in accordance with the speed curve, and currents of the other motors are regulated by using fuzzy PID control. The flow chart of the power balance is shown in Fig. 104.5.

104.5 Conclusions

In this paper, a control system based on PLC was developed for the belt conveyor in coal mines. The control strategies of speed regulating and power balance of multi-motor drive were discussed in detail. These strategies were applied in the developed system and achieved good results. During the start-up period, all the motors reached the steady state in about one second, which avoided the damage caused by the big starting current. When the motors were disturbed or the power distribution was off balance, the system could adjust the power distribution until a new balance was rebuild. The adjust time was about two seconds, which was fewer than the time spending in the conventional control system.

References

1. Liu X (2000) Developing trend of long belt conveyor and its control driving system. Coal Mine Mach 5(11):1–2
2. Jiang W, Han D (2008) Development situation and tendency of native mining belt conveyor. Colliery Mech Electr Technol 11(1):1–6
3. Wang F, Wang M (2007) Application of PLC and graphic operation terminal in conveyer belt. Ind Mine Autom 12(4):94–96
4. Zhang J, Wang T, Wang Z (2004) Theory and application of intelligent control. Metallurgical Industry Press, Beijing

5. Ahn, Kyoung K (2006) Position control of shape memory alloy actuators using self tuning fuzzy PID controller. Int J Control Autom Syst 4(6):756–762
6. Karasakal O (2005) Implementation of a new self-tuning fuzzy PID controller on PLC. Turkish J Electr Eng Comp Sci 13(2):277–286
7. Guo N (2010) Research on power balance of belt conveyor. China Modern Educational Equipment 7(4):54–55

Chapter 105
Research on Process Data Information System of Glass Production

Tao Zhang, Xiaowei Wang and Fashan Yu

Abstract Today, in a high degree of automation companies of the domestic glass industry, most various production sections are independent sets of distributed control system or PLC-controlled equipment. Process equipment operating status and the important production parameters are copied or printed by artificial form to be submitted to the relevant technical departments. And the relevant personnel analyzes the data and revised technical parameters. By this way a lot of manpower and material resources would be wasted, also it would be prone to error. This paper presents a Process Data Information System based on the combination of modern network technology and configuration software, which enables process data to be shared in the relevant departments within the enterprise.

Keywords Process Data Information System · Glass industry · Information system

105.1 Introduction

Now in the domestic glass industry, the various process controls are basically using either programable controllers or distributed control to achieve. Each plant

T. Zhang (✉) · X. Wang · F. Yu
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, Henan, China
e-mail: zhangtao@hpu.edu.cn

X. Wang
e-mail: proceedings@126.com

F. Yu
e-mail: yufs@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_105,

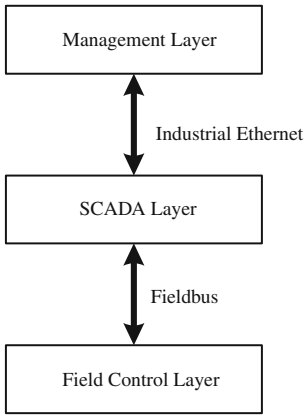


Fig. 105.1 General principle block diagram

has a number of production lines, and each line's construction time and equipment manufacturers are different and each device and system's controllers are different. The process data transmission methods and protocols are different, so that it is relatively independent as devices and systems. Process equipment operating status and production's important parameters are copied or printed by artificial form to be submitted to the relevant technical departments. Then the relevant personnel analyzes the data and revises technical parameters.

With the rapid development of the network technology and enterprise information, office automation and information sharing has become commonplace in the enterprise. The above approach has been restricted the development requirements of modern enterprises. For these reasons, we propose a Process Data Information System (PDIS) based on the combination of modern network technology and configuration software; we can rely on their existing network resources, or renovate and construct a new network for data sharing. Enterprises office automation would be brought into full play, and could improve efficiency and cost savings.

105.2 The Overall Design of PDIS

The glass industry PDIS can be divided into three layers. They are the Management layer, Supervisory Control and Data Acquisition (SCADA) layer and field control layer as shown in Fig. 105.1. Management layer communicated with SCADA layer by industrial ethernet. SCADA layer communicated with Field control layer by field bus. The design of PDIS is included hardware and software design, as follows.

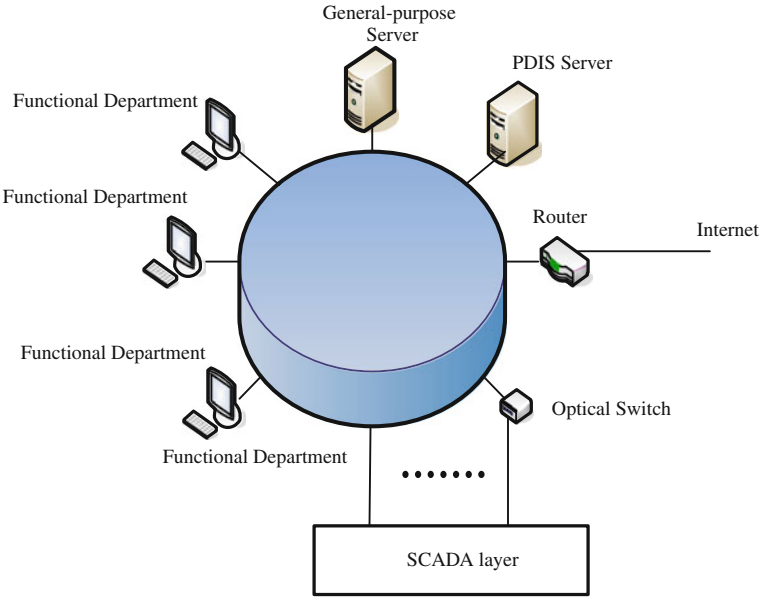


Fig. 105.2 Management layer diagram

105.3 PDIS Hardware Design

105.3.1 Management Layer Hardware Design

The major function of the management layer is that each enterprise management department, technical department and workshop supervisor and other technical personnel can browse the historical data of automatic control system, device status and on-site real-time conditions and other parameters. By analyzing the data, they could improve the production process and make plans of equipment maintenance and spare parts.

General business Office Automation system includes one general-purpose server at least, which has file sharing, WEB publishing, mail service and other functions. In the server room, there are optical switches, routers and high-speed switches and other network components. The person in charge of functional departments and corresponding computers are connected to the server via Ethernet with the router. As the process data are special and important, we install a process data sever in the server room, that is used to store the process data of the whole plant.

The circle network is used in Management layer's network as shown in Fig. 105.2. Once one node connected failure, the other node can be connected by other route. The whole network was connected in the Internet by the router and connected SCADA by optical switch and Industrial Ethernet networks.

105.3.2 SCADA Layer Hardware Design

Modification or installation of operating stations and the SCADA servers in every producing departments which are connected to management layer through the Industrial Ethernet Network. For the farther producing departments or branch factory, install fiber-optic switch to connect through optical fiber. Install the communication cards on the SCADA servers corresponding to the field controllers. In order to make SCANDA servers exchange data only with operating stations and PDIS server, we should make the appropriate settting on the routers, which helps to improve the security of the server and prevents on-site equipment misuse from other network data, as shown in Fig. 105.3.

In every workshop SCADA connected Management layer with its router. If one of these nodes is fault, the other will not be affected and will work well.

105.3.3 Feild Control Layer Hardware Design

Some of the early workshop equipment are controlled by Programable Logic Controller (PLC) and Distributed Control System (DCS), but their manufacturers and models are different, so the data transfer protocol and port are not the same. For example, one equipment is Siemens S7 series and the other is AB SLC5 series or Mitsubishi FX2 N or Q series. To transmit all equipment operation data to the SCADA server, a communication protocol which all PLC has, such as MODBUS, PROFIBUS or other bus protocol is equal, and an appropriate communication module is added. Another method is to install communication cards corresponding to the PLC communication protocol on the SCADA server, as shown in Fig. 105.4.

For example, Fig. 105.4 shows one workshop's field control layer. It had three different manufacturers PLC, Mitsubishi PLC, Siemens PLC and AB PLC. In this figure, it was used the Industrial Ethernet was used to communicate with the workshop's SCADA. So all PLC should be installed on the Ethernet Modules.

105.4 PDIS Software Design

In every layer operating system and some softwares should be installed, and some corresponding setting must be made. In this section, the following is divided into three parts to elaborate software design, as shown in Fig. 105.5.

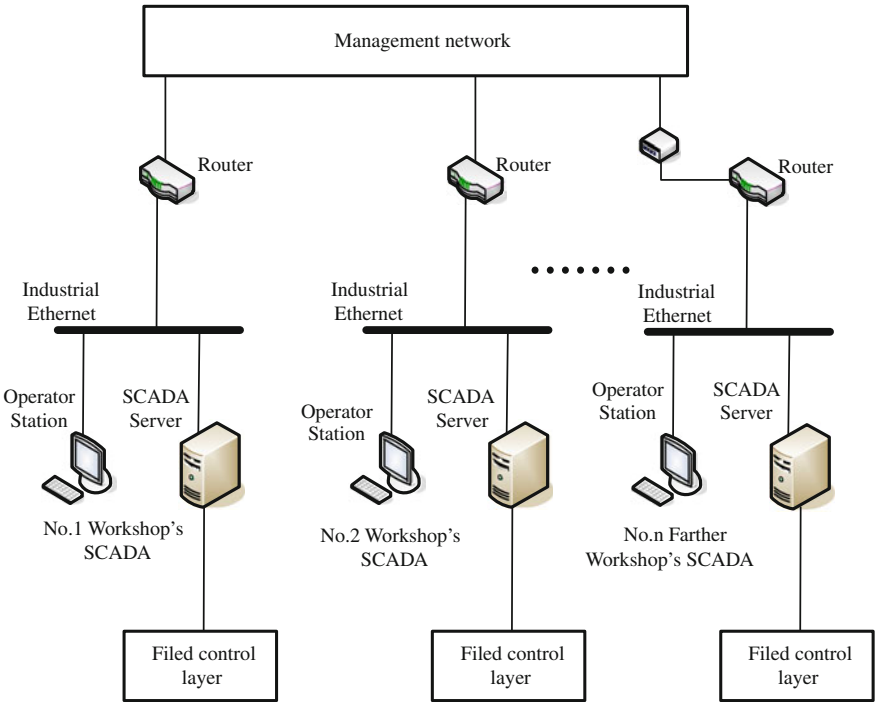


Fig. 105.3 SCADA layer diagram

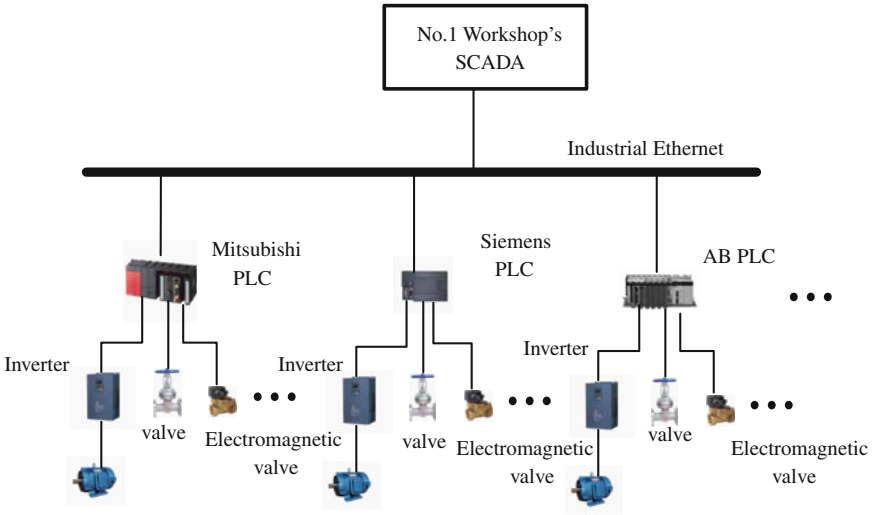


Fig. 105.4 Field control layer diagram

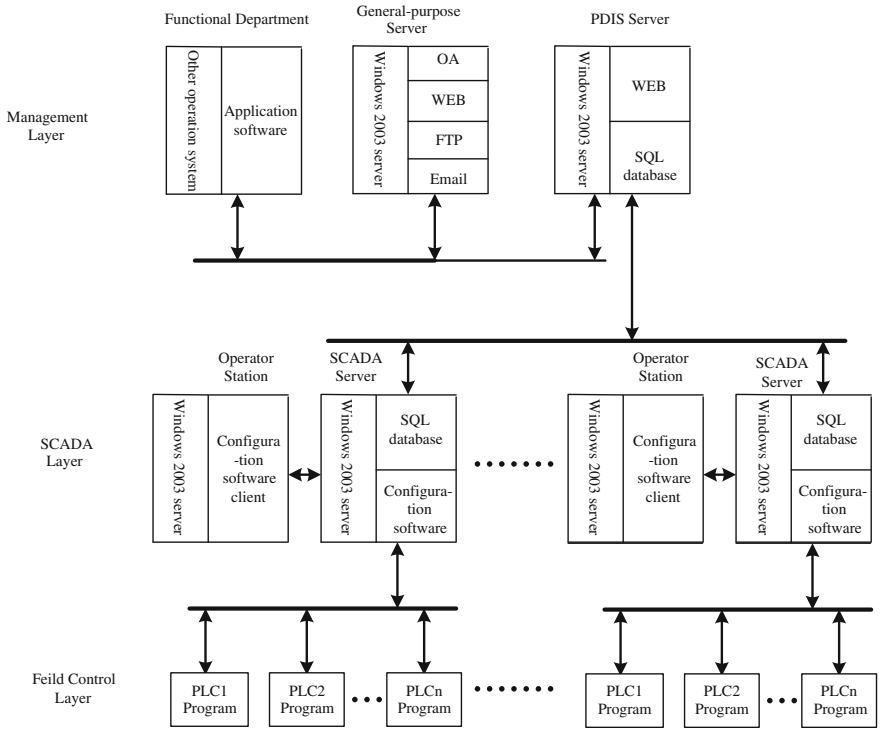


Fig. 105.5 Software diagram

105.4.1 Management Layer Software Design

On general-purpose server operating system is Microsoft Windows Server 2003 or Higher version, building Office Automation (OA), common office file sharing, FTP site server, mail services, print sharing, etc.

PDIS server is used to store each device main process parameters and the historical data. For the relevant departments accessing their required data and charts, that is the process data server software installated, configurated and designated. First of all, install Microsoft Windows Server 2003 operating system on the servers. Installed some of appropriate accessory softwares, such as IIS 6.0 (Microsoft Internet Information Services) and the WEB server configuration are installed, the Microsoft SQL server database software is installed which will beed, and build the process data database. Set to read each plant SCADA server storing the data. Classify and build the data used in each department and web pages are classified and built.

105.4.2 SCADA Layer Software Design

SCADA server operating system is Microsoft Windows Server 2003. And then install the appropriate configuration software, such as IFIX, WINCC, INTOUCH, and KingView is installed. The corresponding interface and process database is built. The Microsoft SQL server database software is installed and the historical database is built and then the historical data of the configuration software is associated with the SQL database and the historical data is stored. In the SQL database to facilitate the PDIS server to read the SCADA data.

105.4.3 Feild Control Layer Software Design

In 105.3.3 Section some PLC system has been added a new communicate module. In order to make new module communication with the SCADA system, it is necessary to make the setting of the PLC. This section mainly To makes the PLC's communication setting in the corresponding control procedures. Different PLC's have setting is different and some need programming in PLC.

105.5 Conclusion

This paper puts forward the PDIS System based on configuration software and modern network information technology. It is effective to solve the information sharing of production process data, greatly to save the manpower and material resources cost, and convenient for the management of glass enterprises to analyze data and to make decisions. PDIS System will be effectively realized the information automation management of the enterprise, and will bring great economic and social benefits.

Chapter 106
A Fast Recognition Method of Workpiece Based on Improved Coding of Orientation Run Length Coding

Li Zi

Abstract A fast recognition on workpiece is proposed in this paper. This algorithm based on object image fully takes advantage of shape features of workpiece. Orientation run length coding is used to set up length samples of part image, which is compared with sample library for recognition. Then, circumference feature is encoded for the workpiece that cannot be identified, which is adopted for second step. Test demonstrates that the workpiece recognition method proposed in this paper can distinguish fast from each other correctly and it can be applied in flexible automatic production for multispecies and small amount manufacture.

Keywords Workpiece recognition · Computer vision · ORLC · Sample matching
Introduction

106.1 Introduction

With the development of automatic in manufacture industry, an increasing number of requirements are proposed for fast recognition of workpieces on a product line, especially within the progress of computer vision technology. Computer vision technology, which has a capability of nonconnected, reliability and high precision, is generally considered as the first choice in automatic recognition of work pieces. There are already many methods about image matching and recognition [1–3] in application of computer vision technology and image processing. The orientation

L. Zi (✉)
Modern Education Technology Center, Luoyang Institute of Science and Technology, 471023 Luoyang, China
e-mail: meiguiwuxinll@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_106,

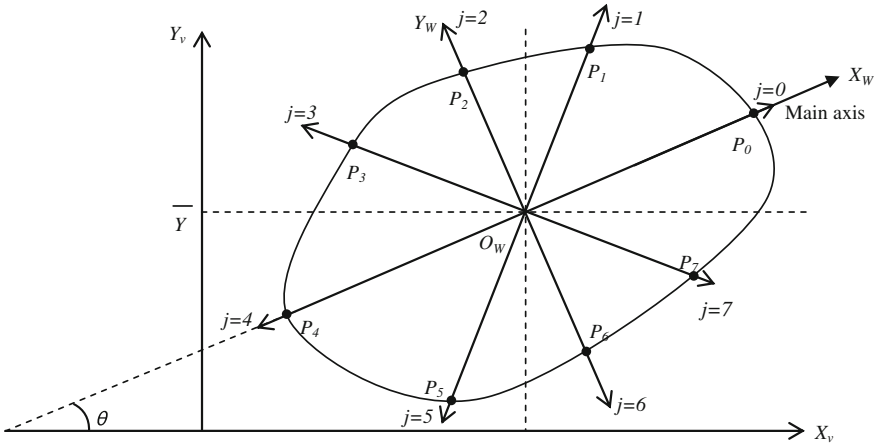


Fig. 106.1 Principle of ORLC

run length coding (ORLC) fits extraordinarily for fast and real-time identification of parts based on image recognition, especially for moving targets. Recently, some kinds of forms come out, such as coordinate samples of ORLC, length samples of ORLC and byte samples of ORLC [4, 5]. Wang achieved ORLC with BP neural network for Computer Integrated Manufacturing System [6]. However, this method will lead to wrong judgments for some workpieces. Therefore, an improved method based on ORLC is proposed in this paper, which is aimed at high identification speed and accuracy. It takes full advantage of shape feature. First, length samples of ORLC are built for matching with sample banks; second, for the unidentified objects, perimeter is used as a complemented feature which is encoded by ORLC.

106.2 A Brief Introduction of ORLC

106.2.1 ORLC

In the process of encoding for the image of workpiece, Center of image area is used as encoding center and minimum axis of inertia of image acts as criterion encoding direction [7, 8]. A detailed information on ORLC is shown in Fig. 106.1. $(\bar{X}, \bar{Y})$ is the centre of image area. X_w is the direction of minimum axis of inertia of image, which includes angle with X as θ. Line numbers of ORLC is not regulated strictly, which means we can choose $\pi/2$ angles four lines, $\pi/3$ angles six lines or $\pi/4$ angles eight lines. All of eight lines used in this paper start from Ow and each of other borders upon $\pi/4$ that spreads eight orientations.

Results of ORLC divide into three species such as coordinate samples of ORLC, length samples of ORLC and byte samples of ORLC. Coordinate samples of ORLC record coordinate figures of cross point between coding lines and workpiece image edge. Apparently, these coordinates, which include part image information of location and orientation, need uniformization for application. Length samples refer distance from Ow $(\bar{X}, \bar{Y})$ to cross point between coding lines and workpiece image edge. Byte samples are the results of interactional comparison. Although it is relatively easy, it can not reflect shape information sufficiently. Part shape information can be obtained from length samples directly. At the same time, workpiece is already processed uniformization actually if we adopt standard coding direction. Therefore, length samples are used in this paper. Length samples can be indicated as

$$\mathrm{LM} = \{\mathrm{LM0},\ \mathrm{LM1}, \ldots,\ \mathrm{LM}j \ldots,\ \mathrm{LMn} - 1\}. \tag{106.1}$$

Here, LMj is coding length in j orientation while j is the coding direction index. $0 \leq \mathrm{j} \leq \mathrm{n\text{-}1}$; n is sum of coding direction. LMj can be indicated as

$$\mathrm{LM}_j = \overline{O_w P_j}. \tag{106.2}$$

106.2.2 Determination of Minimum Moment of Axis of Inertia

The most important step of ORLC is confirming code reference direction which decides minimum moment of axis of inertia $X_{w.}$ It can be obtained from the minimum orientation through $(\bar{X}, \bar{Y})$ that corresponds the second moment of inertia product of workpiece image. The second moment of inertia of mass distribution center of object is used for simplified calculation. We calculate the second moment of inertia I_{20}, I_{02}, the second moment of inertia product I_{11} and the third moment of inertia I_{30}, I_{03} [9].

$$I_{20} = \sum_{i=0}^{n} \sum_{j=0}^{m} f(x, y)\left(x - \overline{X}\right)^2. \tag{106.3}$$

$$I_{02} = \sum_{i=0}^{n} \sum_{j=0}^{m} f(x, y)\left(y - \overline{Y}\right)^2. \tag{106.4}$$

$$I_{11} = \sum_{i=0}^{n} \sum_{j=0}^{m} f(x, y)\left(x - \overline{X}\right)\left(y - \overline{Y}\right)^2. \tag{106.5}$$

$$I_{30} = \sum_{i=0}^{n} \sum_{j=0}^{m} f(x, y)\left(x - \overline{X}\right)^3. \tag{106.6}$$

$$I_{03}=\sum_{i=0}^{n}\sum_{j=0}^{m}f(x,y)\left(y-\overline{Y}\right)^{3}. \tag{106.7}$$

Direction angle of part image can be expressed as follows

$$\theta=\frac{1}{2}\arctan\left(\frac{2I_{11}}{I_{20}-I_{02}}\right). \tag{106.8}$$

If $I_{20}<I_{02}$, Calculate I_{03} :
If $I_{03}\geq 0$,

$$\theta=\theta_0+\frac{3}{2}\pi. \tag{106.9}$$

If $I_{03}<0$,

$$\theta=\theta_0+\frac{\pi}{2}. \tag{106.10}$$

If $I_{20}\geq I_{02}$, Calculate I_{30} :
If $I_{30}\geq 0$,

$$\theta=\theta_0+\pi. \tag{106.11}$$

If $I_{30}<0$,

$$\theta=\theta_0. \tag{106.12}$$

106.3 An Improved Algorithm Based on ORLC

The algorithm design for workpiece recognition divides into two steps. First, we depend on ORLC recognition to achieve primary sort; however, it cannot distinguish all the workpieces. Then, we code circumference of the part image and compare this feature with each other for the rest of them.

106.3.1 Building Length Samples for Recognition

The generation steps of length samples about ORLC are as follows:

(1) The object image is segmented to B/W image, which is marked as F image.
(2) Then we take advantage of moment of inertia for location and orientation of F.
(3) Calculate ORLC to F and get coordinate arrays of location and orientation which consist with F.

(4) Eventually, Length samples are obtained through distance from point of intersection of coding lines and contour of workpiece to the center of figure.

Once we get length samples of objects, we can match object image immediately. The detail is as follows:

(1) Pick up status information while objects in vision system. It can be shown as location and orientation in field of CCD.
(2) Get length samples of ORLC by means of coding.
(3) Calculate matching value and then obtain conclusions.

We would like to use ORLC matching method in sign region of image processing. For a vision system, it includes all its operation target samples. Therefore, we can indicate all the ORLC length samples collection for matching and recognition by matrix form.

$$\left[LM_{T\sum}\right]_{n,W} = \begin{vmatrix} [LM_T]_{n,1} \\ [LM_T]_{n,2} \\ \vdots \\ [LM_T]_{n,k} \\ \vdots \\ [LM_T]_{n,W} \end{vmatrix}. \tag{106.13}$$

Here, $\left[LM_{T\sum}\right]n.W$ is gathering of ORLC length samples of vision system while $[LM_T]n.k$ is one of ORLC length samples which is given in formula (106.1). As indicated above, T is one object's ORLC length sample; k is its index that boundary of k is as follows:$1 \leq k \leq W$; j means direction with coding length arrays of ORLC. Theoretically, both $1 \leq j \leq n$ and n,W are positive integers greater than 1. Firstly, ORLC of object is obtained by matching with samples. Additionally, operational process is as same as building ORLC length of samples. ORLC length coding of recognition part can be expressed as follows:

$$\mathrm{LM}_{\mathrm{object}} = \{\mathrm{LM0}, \mathrm{LM1}, \ldots, \mathrm{LM}j, \ldots, \mathrm{LMn}-1\}_{\mathrm{object}}. \tag{106.14}$$

If ORLC length arrays of an object and ORLC length of a sample in the sample library of all the square of the difference between the corresponding items come to be minimum value, it can be believed that the object has reached the matching operation by the following formula express.

$$\phi_k = \sum_{j}^{n} \left(\left[\mathrm{LM}_{\mathrm{object}}\right]_{j,k} - \left[\mathrm{LM}_T\right]_{j,k} \right)^2 \tag{106.15}$$

$$\mathrm{MINIMUM}\{\phi_1, \phi_2, \ldots \phi_k, \ldots, \phi_w\}. \tag{106.16}$$

Fig. 106.2 Two images of real workpieces are shown. The left is octagon and the right one is inscribed circle

Table 106.1 ORLC length samples of two example workpieces

ORLC	Inscribed circle	Octagon
L_0	122	122
L_1	122	122
L_2	122	122
L_3	122	122
L_4	122	121
L_5	122	122
L_6	122	122
L_7	122	122

106.3.2 Disadvantage of ORLC

ORLC recognition method is not complicated and it is not connected with orientation. However, it will make error judgment to some workpiece. From Fig. 106.2a and b, you can conclude that ORLC cannot tell the difference between two part images. Therefore, we need to recommend another new character as recognition index (Table 106.1).

106.3.3 Circumference Code Adding to ORLC

In condition of Fig. 106.2, we cannot distinguish two workpieces. So, circumference coding by ORLC is considered for next step, which is offered as supplement. After scanning object pixels, we can determine this point as one of edge points if its left, right, up and down neighbors are not pixels. Then, calculate all these points of image as circumference (Fig. 106.3).

106.4 Experiment and Analysis

VS142FM CCD of Microvison and 1394 digital image acquisition card are used in workpiece recognition. Image acquisition system as shown in Fig. 106.4. The image resolution of vision system is 1392 × 1040 pixel. Gray level divides into

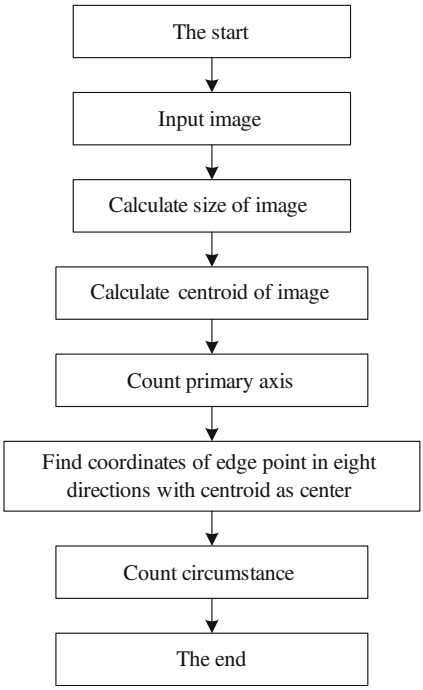


Fig. 106.3 Length and circumference coding flow chart

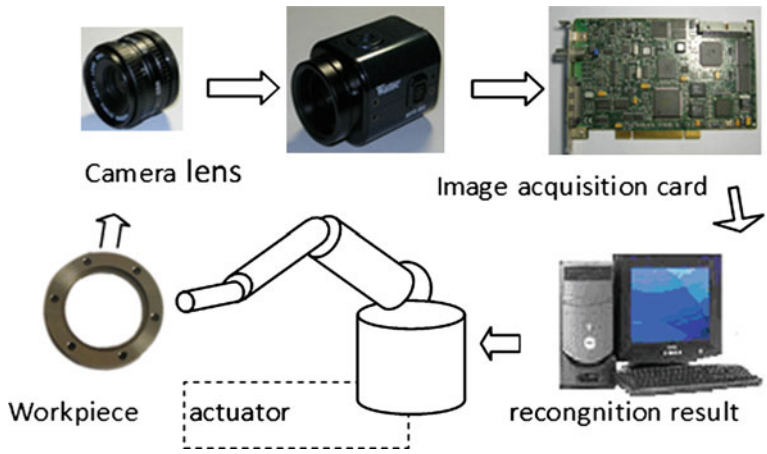


Fig. 106.4 Recognition system of workpieces

256 levels. After workpiece image is changed into (B/W) image through filtering, noise reduction and segmentation process, a recognition method based on improved ORLC, are used in the experiment. Result of recognition is sent to next actuating mechanism and then distinguishes bad ones from qualified products practically.

Table 106.2 ORLC length samples of two example workpieces

ORLC	Round	Octagon
L_0	122	122
L_1	122	122
L_2	122	122
L_3	122	122
L_4	122	121
L_5	122	122
L_6	122	122
L_7	122	122

As Fig. 106.2 shows, the two images are octagon and inscribed circle. They have the same length samples according to ORLC length image. Therefore, we propose to distinguish them with circumference that is highly efficient as Table 106.2 verified. Experimental results demonstrate that this method proposed in this paper, recognized by ORLC and circumference coding, is very efficient for workpiece recognition.

106.5 Conclusion

Computer vision technology for workpiece recognition is an effective way to improve the identification of flexible production systems and automation. Recognition algorithm takes an important part in recognition rate and speed. An improved algorithm on fast workpiece recognition is proposed through making full use of shape characteristic. At first, workpiece image is encoded by six lines ORLC in order to get length samples. Then, we match length samples of ORLC to sample library. If there is object image that cannot distuignish from others, we have to use circumference feature coded by ORLC for recognition. It means a lot to efficiency of production system. Test demonstrates that algorithm in this paper has high recognition rate and fast speed.

References

1. Wang H, Fu W, Kang Y (2008) Study of edge matching approach to workpiece image recognition. Chin J Sci Instrum 29(5):986–991
2. Yan P, He X, Yi R (2011) Automatic identification and monitoring method for workpiece based on dynamic time warping technology and power information. J Mech Eng 47(3):81–88
3. Wang G, Li J, Yi R (2006) A method of recognition and location of objects based on contours. J Henan Univ Sc Technol Nat Sci 27(6):42–45
4. Wang X, Wang F (2000) Regional geometric feature coding of finger print. Opt Technol 26(6):513–516
5. Wang X, Wang Y, Zheng Y (1995) Non-contact 2D positioning and drienting of workpieces by computer vision. J Test Meas Technol 9(2):1–6

6. Wang Y (2006) Study on a fast visual recognition method of workpiece in the context of CIMS. In: Proceedings of the 6th world congress on intelligent control and automation, pp 6895–6899
7. Wang X, Butler C (1993) Use of a machine vision system in a flexible manufacturing cell incorporating an automated coordinate measuring machine. Proc Inst Mech Eng Part B J Eng Manuf 207(B3):199–204
8. Wang X, Wang Y, Zheng Y (1997) Achievement of fast recognition for any shape workpiece use of ORLC. J Astronaut Metrol Meas 17(2):2–6
9. Wang Y (2006) Study on a fast visual recognition method of workpiece in the context of CIMS. In: Proceedings of the world congress on intelligent control and automation (WCICA), vol 2, 21–23 June 2006, Dalian, China, pp 6895–6899

Chapter 107
Equation and Graphical Solution of Characteristic Parameter Relationship of Strip Steering in Rolling Process

Xianqiong Zhao, Yilun Liu and Sheng Huang

Abstract Ideally rolling parameter is symmetrical relative to the rolling center line in the rolling process. When there is asymmetry incentive disturbance, symmetrical state will be destroyed. It brings typical quality problems like strip steering. This chapter studies transverse asymmetry, characteristic of force and deformation of rolling mill and mill bar, establishes asymmetry characteristic quantity of crucial rolling parameter, a relationship equation of asymmetry characteristic parameter of inducement and result state, online detection model which observes asymmetry characteristic parameter of rolling force and corrective model which adjust wedge roll gap, builds graphical solution of characteristic parameter relationship and equation which explains the characteristic parameter.

Keywords Rolling process · Strip steering · Asymmetry · Graphical solution · Modeling

107.1 Introduction

Ideally rolling parameter is symmetrical relative to the rolling center line in the rolling process. When there is asymmetry incentive disturbance, symmetrical state will be destroyed. Some quality problems will happen [1]. Strip steering is one

X. Zhao (✉) · Y. Liu · S. Huang
College of Mechanical and Electricity Engineering,
Central South University,
Changsha 410083, China
e-mail: csuzxq@163.com; csuzxq@qq.com

Y. Liu
e-mail: h_sheng@foxmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_107,

form of expression of rolling parameter lost symmetry relative to rolling centre line in rolling process. Currently there are many qualitative explanatory researches, corrective measures [2, 3] and numerical simulation [4]. But they lack a perfect explanation model and corrective measures against transverse asymmetry of rolling parameter which is a special breaking rolling process.

107.2 Building Asymmetry Characteristic Parameter of Strip Steering State

In order to simplify two sides, difference value is the difference between operator side (os) and drive side (ds). We build the next frame of axes: x axis is vertical rolling system; z axis is transverse rolling system and width direction of strip mill. Origin locates rolling center. y axis is height direction. Transverse asymmetry is that rolling parameter which is asymmetry relative to xoy flat.

107.2.1 Building Asymmetry Characteristic Quantity of Strip Steering Inducement

Perturbation of strip steering mostly comes from asymmetry factor of mill bar and rolling mill [1, 5]. There we extract three main perturbation and describe the characteristic parameter of asymmetry property:

1. Entry thickness of strip is asymmetry relative to width of mill bar, which is strip entry wedge. Difference thickness of two side is the characteristic parameter, which is called ΔH.
2. Location (Z) of strip is asymmetry, which is called partiality. z coordinate when particle of width centre line of mill bar enter roll is called ΔZ.
3. Roll gap (S) of rolling mill is asymmetry, which is called wedge roll gap for short. It includes initial and adjustment wedge roll gap. The difference of roll gap is characteristic parameter. Initial wedge roll gap is called ΔS_0 ($or\,\Delta G_0$). Adjustment wedge roll gap is called ΔS_s (or$or\,\Delta G_s$.) ΔG is difference which is relative to width b of mill bar. If B is hydro cylinder center distance of two sides' rolling mill, there is $\Delta G_0 = \Delta S_0 b/B, \Delta G_S = \Delta S_S b/B$.

107.2.2 Building Asymmetry Characteristic of Force and Deformation Between Rolling Mill and Mill Bar

Plastic deformation of mill bar and elastic deformation of rolling mill determine two deformation determine roll gap (S_L) under load and thickness (h) of mill bar.

Asymmetry of mill bar thickness uses difference of two sides' thickness into characteristic parameter, which is called Δh. Asymmetry of loading roll gap uses two sides roll gap difference into characteristic parameter, which is called ΔS_L (or ΔG_L), $\Delta G_L = \Delta S_L b/B$. Ignoring elasticity restoration of mill bar, there is:

$$\Delta h \approx \Delta G_L. \tag{107.1}$$

Difference deformation of two sides is ΔY_{KB}, which is asymmetry characteristic parameter of deformation of rolling mill. According to the deformation feature of mill bar, there is next relation:

$$\Delta Y_{KB} = \Delta S_L - \Delta S_S - \Delta S_0 = \Delta S_L - \Delta S. \tag{107.2}$$

Asymmetry characteristic of deformation of mill bar ΔY_{KB} correspond width of mill bar b, which is called ΔY_{Kb}. There is the next equation:

$$\Delta Y_{Kb} = \Delta G_L - \Delta G = \Delta h - \Delta G = \frac{b}{B}\Delta Y_{KB}. \tag{107.3}$$

where, $\Delta G = \Delta G_S + \Delta G_0 = \Delta S b/B$.

Deformation quantity of mill bar is rolling reduction in the direction of thickness. Difference of rolling reduction at the two sides is a characteristic parameter of deformation asymmetry, which is called ΔY_{Mb}:

$$\Delta Y_{Mb} = \Delta H - \Delta h. \tag{107.4}$$

Summation of rolling force is P. Asymmetry of rolling force is described by the difference of two sides ΔP. Force point ΔP locates in the hydro cylinder center of operator side, whose distance is $B/2$ near rolling mill center. Distribution function of rolling counter-force is called F. We suppose F and P are same. Asymmetry of rolling counter-force F is described by characteristic parameter ΔF. Force point of ΔF is close to mill bar section of operator side (near mill bar center $b/2$). Moment between ΔF against mill bar center and $f(z)$ against mill bar is equivalent. Supposing $f(z)$ is linear distribution along direction z. According to deformation characteristic of mill bar, we can deduce the next equation:

$$\Delta F = \frac{M}{6}(\Delta H - \Delta h). \tag{107.5}$$

where, M is plastic deformation coefficient of mill bar.

The meaning of asymmetry characteristic parameter is shown in Fig. 107.1.

107.2.3 Building Characteristic Quantity of Strip Steering

Strip steering of mill bar is asymmetry whose speed of mill bars are relative to rolling center. Asymmetry of entry speed V of mill bar is described by the speed

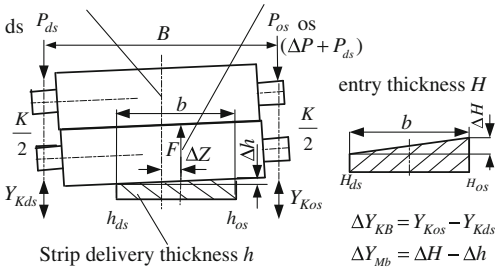


Fig. 107.1 The meaning of asymmetry characteristic parameter

difference ΔV when mill bar enters two sides. Asymmetry of exit speed v of mill bar is described by the speed difference Δv.

Metal flow of cross section is equal, which flows past rolling area in unit time. So there is the next equation under ignoring broad siding.

$$\frac{\Delta v}{\bar{v}} - \frac{\Delta V}{\bar{V}} = \frac{\Delta H}{\bar{H}} - \frac{\Delta h}{\bar{h}}. \tag{107.6}$$

According to metal deformation theory in rolling area, we can know, Δv and ΔV is always the opposite sign. Forward and backward slips exist in some kind of a proportionate relationship. Then Eq. (107.6) can be written as:

$$\begin{cases} \dfrac{\Delta v}{\bar{v}} = \alpha\left(\dfrac{\Delta H}{\bar{H}} - \dfrac{\Delta h}{\bar{h}}\right) = \alpha(\Delta H - \lambda\Delta h)/\bar{H} \\ \dfrac{\Delta V}{\bar{V}} = (\alpha - 1)\left(\dfrac{\Delta H}{\bar{H}} - \dfrac{\Delta h}{\bar{h}}\right) = (\alpha - 1)(\Delta H - \lambda\Delta h)/\bar{H} \end{cases} \tag{107.7}$$

α is partition coefficient of asymmetry characteristic between exit and entry roller speeds, $0 \le \alpha \le 1$. Building the next two characteristic quantities of strip steering which describes the degree and direction of strip steering:

① asymmetry characteristic quantity of mill bar under effective strain:

$$\Delta Y_{EM} = \Delta H - \lambda\Delta h \tag{107.8}$$

② characteristic quantity of strip steering Δ_δ

$$\Delta_\delta = -\Delta Y_{EM} = \lambda\Delta h - \Delta H \tag{107.9}$$

The quantity is the same as ΔY_{EM}, which describes the size of asymmetry rate and direction of mill bar at exit and entry. The quantity translates the asymmetry characteristic of speed into length of direction of y axis. It is order of magnitude with asymmetry characteristic parameter $\Delta H, \Delta h, \Delta Y_{Mb}$. if $\Delta v < 0, \Delta V > 0$, then $\Delta_\delta > 0$. Mill bar faces operator side to strip steering. If $\Delta v > 0$, $\Delta V < 0$, then $\Delta_\delta < 0$. Mill bar faces drive side to strip steering.

107.3 Relationship Equation of Characteristic Parameter of Mill Bar in Strip Steering

107.3.1 Relationship Equation of Asymmetry Characteristic Parameter of Force and Deformation

Rigidity K of rolling mill actually is lengthways rigidity of rolling mill. Plastic deformation characteristic of mill bar is shown by M. This chapter defines level rigidity and plastic coefficient as follows:

$$K_E = \frac{B}{b}\frac{\Delta P}{\Delta Y_{KB}} = \frac{\Delta P}{\Delta Y_{Kb}},\ M_E = \frac{b}{B}\frac{\Delta F}{\Delta Y_{Mb}} = \frac{\Delta P}{\Delta Y_{Mb}} \tag{107.10}$$

Suppose there is partiality ΔZ, partiality causes asymmetry of rolling force. At this time, Eq. (107.10) can be written by the next form:

$$\begin{cases} \Delta P = K_E \Delta \mathrm{Y_{Kb}} + \delta_{KZ}\Delta \mathrm{Z} \\ \Delta P = M_E \Delta \mathrm{Y_{Mb}} + \delta_{MZ}\Delta \mathrm{Z} \end{cases}. \tag{107.11}$$

Equations (107.3), (107.4) go into (107.11):

$$\begin{cases} \Delta P = K_E(\Delta h - \Delta \mathrm{G}) + \delta_{KZ}\Delta \mathrm{Z} \\ \Delta P = M_E(\Delta H - \Delta h) + \delta_{MZ}\Delta \mathrm{Z} \end{cases} \tag{107.12}$$

The first equation is called deformation equation of rolling mill. The second equation is called deformation equation of mill bar, where K_E, M_E, δ_{KZ} and δ_{MZ} are undetermined coefficients.

According to force analysis and relationship of deformation compatibility, we can get the next undetermined coefficient to four-rolling mill [1]:

$$\frac{1}{K_E} = \begin{bmatrix} \frac{b}{B} & \frac{3bB}{B_F^2} & \frac{3B}{b} \end{bmatrix} \begin{bmatrix} \frac{2}{K_H} & \frac{2}{B_F k_F} & \frac{2}{bk_f} \end{bmatrix}^T,\ M_{\mathrm{E}} = \frac{b}{6\mathrm{B}}M,\ \delta_{MZ} = \frac{2}{\mathrm{B}}\mathrm{P},\ \delta_{KZ} = \frac{12}{b^2}\frac{K_{\mathrm{E}}}{k_f}\mathrm{P}. \tag{107.13}$$

where

K_{H}—equivalent stiffness of stander, bearing pedestal, hydro-cylinder (kN/mm),

k_{F}—contact flatten elasticity coefficient of brace roller, work roll and mill bar (kN/mm^2),

k_{f}—contact flatten elasticity coefficient of work roll and mill bar (kN/mm^2),

BF—contact length of work roll and brace roll (mm).

107.3.2 Relationship Equation of Characteristic Parameter of Strip Steering

The direction of strip steering is the same under the disturbance of ΔG and ΔH. They can equivalent transform. There we discuss $\Delta H + \Delta Z$. Putting Eq. (107.8), (107.12) together, we can get:

$$\begin{aligned} \Delta Y_{EM} &= \Delta H - \lambda \Delta h = \frac{K_E + M_E - \lambda M_E}{M_E + K_E} \Delta H - \frac{\lambda(\delta_{MZ} - \delta_{KZ})}{M_E + K_E} \Delta Z \\ \Delta P &= \frac{K_E M_E}{K_E + M_E} \Delta H + \frac{\delta_{MZ} K_E - \delta_{KZ} M_E}{K_E + M_E} \Delta Z \end{aligned} . \quad (107.14)$$

We separately seek partial derivative towards disturbance characteristic quantity. So it has the next conclusion:

(1) Contribute which homodromous ΔH, ΔZ aim at ΔP have composition. The contributions which ΔH, ΔZ aim at ΔY_{EM} have compensation rectification.
(2) Influence weight coefficient which ΔH, ΔZ aim at ΔP magnify K_E/λ times relative to ΔY_{EM}. So ΔP is a observable value and a value which is very hypersensitive towards disturbance. It is reasonable for ΔP as observed quantity.

107.3.3 Online Testing Model of Strip Steering State

Ignoring high order small quantity of ΔZ influence term, we can deduce the relationship model between online observable quantity ΔP and characteristic quantity of strip steering Δ_δ because of Eq . (107.14):

$$\Delta_\delta = \frac{\lambda}{K_E} \Delta P - \Delta H, \ \delta = \Delta_\delta / |\Delta_\delta|. \quad (107.15)$$

The model can become the state testing model of strip steering, test the degree and direction of strip steering. The model is a much valued function. The direction of rolling force difference ΔP cannot confirm the direction of strip steering. It needs to analyze in the different condition.

107.3.4 Balance Control Model of Speed

In the ΔH, ΔZ inducement, if we judge that state is $w_a(\Delta P_a, \Delta_{\delta a})$ by Eq. (107.15), "wedge roll gap" control begins. The two-sided hydro cylinder has different adjustment quantity. Rolling becomes stable state $w(\Delta P, 0)$.

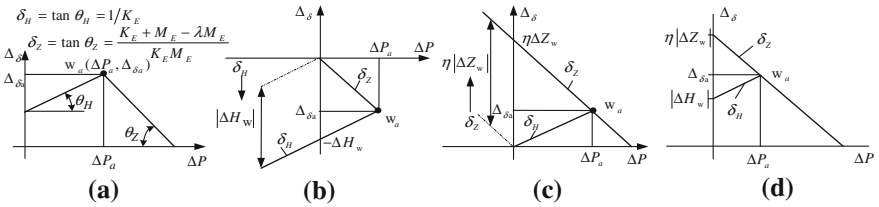


Fig. 107.2 The characteristic parameter relationship of strip steering

The new state is$w_S(\Delta P_S, \Delta_{\delta S})$. Combining Eqs. (107.9), (107.12), (107.14) can get the next equation:

$$\begin{bmatrix} \Delta P_S \\ \Delta_{\delta S} \end{bmatrix} = \begin{bmatrix} \dfrac{K_E M_E}{K_E + M_E} & \dfrac{\delta_{MZ} K_E}{K_E + M_E} & -\dfrac{K_E M_E}{K_E + M_E} \\ \dfrac{\lambda M_E}{K_E + M_E} - 1 & \dfrac{\lambda \delta_{MZ}}{K_E + M_E} & \dfrac{\lambda K_E}{K_E + M_E} \end{bmatrix} \begin{bmatrix} \Delta H \\ \Delta Z \\ \Delta G_S \end{bmatrix}. \tag{107.16}$$

If $\Delta_{\delta S} = 0$ is regulate and control target, we can deduce the balance control model of speed by Eq. (107.16). It is a correct model.

$$\Delta G_S = -\frac{K_E + M_E}{\lambda K_E} \Delta_{\delta a},\ \Delta S_S = -\frac{B}{b} \frac{K_E + M_E}{\lambda K_E} \Delta_{\delta a}. \tag{107.17}$$

107.4 The Graphical Solution of Characteristic Parameter Relationship of Strip Steering

Characteristic parameter relationship figure of strip steering can be called $\Delta_\delta - \Delta P$ figure. Combining Eqs. (107.9), (107.14), (107.15), we can get the next equation:

$$\begin{cases} \Delta_\delta = \dfrac{\lambda}{K_E} \Delta P - \Delta H = \delta_H \Delta P - \Delta H \\ \Delta_\delta = -\dfrac{K_E + M_E - \lambda M_E}{K_E M_E} \Delta P + \dfrac{\delta_{MZ}}{M_E} \Delta Z = -\delta_Z \Delta P + \eta \Delta Z \end{cases}. \tag{107.18}$$

The first equation is named line equation of δ_H. The second equation is named line equation of δ_Z.

Figure 107.2a is $\Delta_\delta - \Delta P$. The function relationship of disturbance quantity, observation quantity, adjustment quantity and characteristic quantity can quantitatively be got by the next

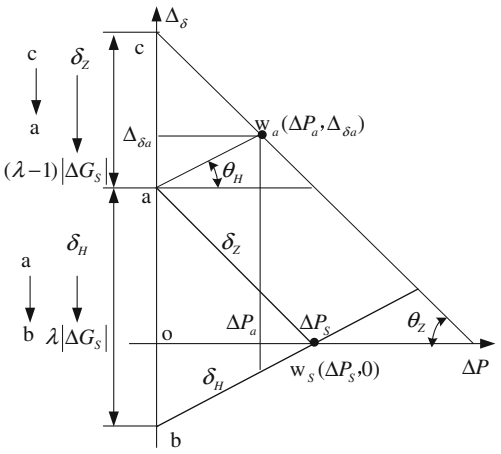


Fig. 107.3 Model graphical solution of balance control of speed

107.4.1 Disposing Disturbance Characteristic Parameter in Figure $\Delta_\delta - \Delta P$

When disturbance ΔZ is zero, line δ_Z cross origin. When disturbance ΔZ is greater than zero, line δ_Z moves down $\eta|\Delta Z|$.

When disturbance ΔH is zero, line δ_H crosses origin. When disturbance ΔH is greater than zero, line δ_H moves down $|\Delta H|$. When disturbance ΔH is less than zero, line δ_H moves up $|\Delta H|$. Figure 107.2b is the solution of work point w_a when $\Delta Z = 0$ and $\Delta H = \Delta H_w > 0$. Figure 107.2c is the solution of work point w_a when $\Delta H = 0,\ \Delta Z = \Delta Z_w > 0$. Figure 107.2d is the solution of work point w_a when $\Delta H = \Delta H_w < 0, \Delta Z = \Delta Z_w > 0$.

We talk about the figure under wedge gap disturbance. Combining Eqs. (107.9), (107.12) and (107.14), we can get the next line equation set:

$$\begin{cases} \Delta_\delta = \delta_H \Delta P + \lambda \Delta G - \Delta H \\ \Delta_\delta = -\delta_Z \Delta P + \eta \Delta Z + (\lambda - 1)\Delta G \end{cases}. \tag{107.19}$$

Therefore line δ_H moves down $\lambda \Delta G_w$ when there is disturbance $\Delta G = \Delta G_w > 0$. At the same time, line δ_Z moves up $(\lambda - 1)\Delta \mathrm{G_w}$. The node of the two lines is the new work point.

107.4.2 Graphical Solution of Balance Control of Speed

$\mathrm{w}_a(\Delta \mathrm{P}_a, \Delta_{\delta a})$, state have roll gap dispersion adjustment of $\Delta \mathrm{G_S}$. The new state is $\mathrm{w_S}(\Delta \mathrm{P_S}, \Delta_{\delta \mathrm{S}})$. It is the node of line δ_H and δ_Z in figure $\Delta_\delta - \Delta P$.

$$\begin{cases} \Delta_{\delta S} = \delta_H \Delta P_S + \lambda \Delta G_S - \Delta H \\ \Delta_{\delta S} = -\delta_Z \Delta P_S + \eta \Delta Z + (\lambda - 1)\Delta G_S \end{cases} \tag{107.20}$$

If $\Delta_a > 0$, the solution of point $\mathrm{w}_a(\Delta \mathrm{P}_a, 0)$ is shown as follows (Fig. 107.3):

107.5 Conclusion

When there is asymmetry inducement in the rolling process, rolling process will be unstable state of strip steeping and need control adjustment. Inducement is not easy testing quantity. At this time, we can test characteristic quantity of strip steering by using online testing model of strip steering. ΔP is observed quantity. Then "wedge roll gap" is adjustment quantity. We can get correct control quantity by using balance control model of speed. On the one hand there is online test; on the other hand there is control. We can use $\Delta_\delta - \Delta P$ to the graphical solution in the process.

References

1. Zhen Y, Xiao G, Zhang C (2007) The analysis of ultra-thin strip running off in tangshan steel and some correcting measures. Metall Stand Qual 12:24–27
2. Biggs DL, Hardy SJ, Brown KJ (2000) Influence of process variables on development of camber during hot rolling of strip steel. Ironmak Steelmak 27(1):55–62
3. Zhengyi J, Kiet AT (2004) Mechanics of thin strip steering in hot rolling. AIP Conf Proc 712:418–423
4. Zhao X, Liu Y, Fu Z, Lin D (2011) 3-D numerical simulation of aluminum strip steering process under asymmetric variations in hot rolling. J Hunan Univ 11(04):21–25
5. Zhao X, Huang S, Liu Y, Fu Z (2010) The modeling of trajectory deviation and deviation direction in aluminum hot tandem mill. In: 2010 International conference on digital manufacturing and automation, pp 734–737

Chapter 108
Study on the Crush Model of High-Pressure Grinding Rolls

Xianqiong Zhao, Sheng Huang, Chao Liu and Ling Deng

Abstract A high-pressure grinding roll is energy efficient crush equipment. On the basis of mass balance, energy conservation and overall balance, this paper puts forward the intersected idea due to the uneven stress in the roller surface. The grinding and comminution model of roller press is deeply investigated. It puts forward crush model using mathematic theory. The simulated specific energy consumptions and particle size distributions, compared with the experiment data, were considered good enough. The model was able to predict adequately throughput capacity, specific energy consumption and particle size distributions of the edge, center and total products.

Keywords HPGR · Bed compression · Modeling

108.1 Introduction

The first commercial application of high-pressure grinding roll (HPGR) was in 1985 and its success resulted in increasing numbers of applications since then. The classical theory is bed compression crush theory. It can be described: mineral particle is crushed under lots of bed compression condition. Every crush machine cannot reach the ideal bed compression condition because of different crush

X. Zhao (✉) · S. Huang · C. Liu · L. Deng
College of Mechanical and Electricity Engineering,
Central South University, Changsha 410083, China
e-mail: csuzxq@163.com

S. Huang
e-mail: h_sheng@foxmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_108,

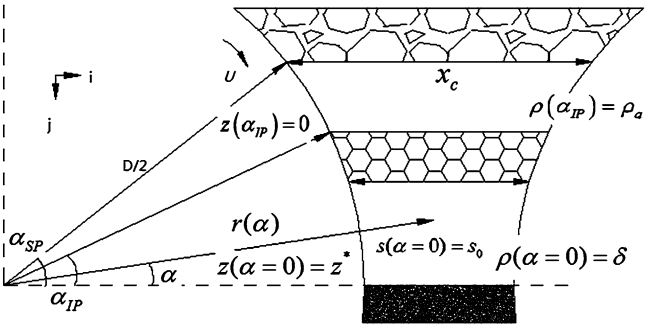


Fig. 108.1 Crush schematic diagram of HPGR

condition. Reference [1] studies the effect law which load—displacement curves, bed height and load characteristic direct crush on the basis of bed compression grinding theory. Reference [2] gets a new concept of the differential shear theory, which based on the study of process and result of the high-pressure double roller cracker for brittleness ore crush. Most of the researchers follow the crush of a single particle theory of crack propagation and energy balance principle and the experimental test data for various amendments; there is no good explanation for formation mechanism of the cake under the high-pressure crush.

108.2 Modeling

HPGR rely on pure compressive stress to squeeze broken. The breakage of ore occurred in a short time. Its energy consumption is mainly reflected by the amount of the pressure. This paper puts forward the crush model of HPGR on the basis of ore characteristics, equipment dimensions and operating conditions. The model structure is developed by Daniel and Morrell [3] on the basis of conservation of mass, conservation of energy and overall balance. So it can predict throughput, power consumption and particle size distribution of the product. Figure 108.1 is crush schematic diagram of HPGR.

108.2.1 Throughput Model

The center of mass of a band of ore with width *s(a)* is defined by the position of vector *r(a)*, diameter *D*(m), length *F*(m) and operating gap s_0(m). So:

$$r(\alpha) = \frac{1}{2}(D + s_0)\mathbf{i} + \frac{D}{2}\sin\alpha\mathbf{k}. \tag{108.1}$$

The density of the ore band at any angle a is $r(a)$. At the beginning of the particle bed compression zone, $r(a)$ is equivalent to the feed bulk density $r(a)$. The tonnage of the ore band, G_s (t/h), with $r(a)$ expressed in t/m3, is written as a function of the angle a in Eq. (108.2)

$$G_s(\alpha) = 3600\rho(\alpha)s(\alpha)LU\cos\alpha. \tag{108.2}$$

The width of the ore band as a function of the angle a can be expressed as:

$$s(\alpha) = s_0 + D(1 - \cos\alpha). \tag{108.3}$$

Under steady-state condition, if the amount of material loss is zero, i.e. the feed is equal to the product; we can get the next Eq. (108.4)

$$\rho_a D\cos^2\alpha_{IP} - \rho_a(s_0 + D)\cos\alpha_{IP} + \delta s_0 = 0. \tag{108.4}$$

We can find Eq. (108.4) is a quadratic equation of $\cos\alpha_{IP}$. Because $D = b^2 - 4ac > 0$, the equation has two roots. We can get Eq. (108.5) according to actual situation.

$$\cos\alpha_{IP} = \frac{1}{2D}\left[(s_0 + D) + \sqrt{(s_0 + D)^2 - \frac{4s_0\delta D}{\rho_a}}\right]. \tag{108.5}$$

So Eq. (108.2) can estimate throughput in any position. Specifically at the extrusion zone ($a = 0$), the throughput can be calculated as:

$$G_S = 3600\delta s_0 LU. \tag{108.6}$$

108.2.2 Power Draw Model

The force applied to the material at the particle bed compression zone is called the compression force, F (kN). The rolls operating pressure is R_P (bar). The total power draw is P (kW). Since the HPGR is operated in a choke fed condition, the applied pressure is distributed only in the upper right half of the roll. Then the projected area considered should be $\frac{D}{2}L$

$$F = 100R_P\frac{D}{2}L. \tag{108.7}$$

Then torque t (kN/m) can be written as:

$$\tau = F\sin\left(\frac{\alpha_{IP}}{2}\right)\frac{D}{2}. \tag{108.8}$$

As the power required to spin both rolls is equal to twice the torque multiplied by the rolls angular velocity, then P can be written as:

$$P = 2F \sin\left(\frac{\alpha_{IP}}{2}\right) U. \tag{108.9}$$

The energy consumption W(kWh/t)is written as the ratio between the power draw (kW) and the throughput (t/h):

$$W = \frac{P}{G_S} = \frac{F \sin\left(\frac{\alpha_{IP}}{2}\right)}{1800 \delta s_0 L}. \tag{108.10}$$

108.2.3 Particle Size Distribution Model

The HPGR is considered as a series of two size reduction stages. The first stage is single particle compression. The next stage is particle bed compression.

The single particle compression zone is located between the angle a_{SP} and a_{IP}. Particles larger than a certain size C_c are broken. The critical size C_c is got by replacing the inter particle compression angle a_{IP} in Eq. (108.3).

$$\chi_C = s(\alpha_{IP}) = s_0 + D(1 - \cos \alpha_{IP}). \tag{108.11}$$

The product of the single particle compression zone rejoins with the fraction of material of material of size lesser than or equal to the critical size C_c, forming beds of particles with a particle size distribution. Then it starts particle bed compression. According to previous sum up, the particle bed compression is broken into two areas, namely the central area and edge zone. It is also known as edge effects. The pressure profile exerted over the rolls is similar to a parabola, as it is shown in Fig. 108.2. In Fig. 108.2, rolls are divided into N_B equal portions. Because every N_B equal portion is different, every power of N_B equal portions is different and every crush rate of N_B equal portions is different. The intensive property m_{ik} (mass fraction retained by weight in size class i, in each block k) is a function of the vertical position. According to mass balance principle, the model equation consists in a system of $N' N_B$ differential equations, each one for the size class i $(i = 1, \ldots, N)$ in each block k $(k = 1, \ldots, N_B)$. As shown in Eq. (108.12)

$$v_z \frac{\mathrm{d}}{\mathrm{d}z} m_{i,k}(z) = \sum_{j=1}^{i-1} S_{j,k} b_{ij} m_{j,k}(z) - S_{i,k} m_{i,k}(z), \tag{108.12}$$

where S_{jk} is the rate of breakage of particles of size i in each block k; b_{ij} is the fraction of particles of size "i". To solve these equations the following border conditions are used:

$$m_{i,k}(z = 0) = f_i^{IP},\ m_{i,k}(z = z^*) = p_{i,k}$$

where $p_{i,k}$ is the mass fraction retained by weight in size class i, in the product of each block k. Z^* is the vertical distance from the entrance to the particle bed compression zone to the extrusion zone.

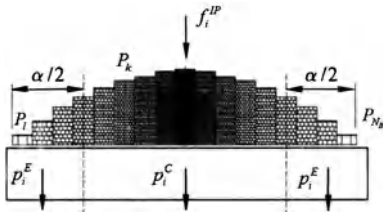


Fig. 108.2 Discretization of the roll

$$z^* = \frac{D}{2}\sin(\alpha_{IP}). \tag{108.13}$$

Equation (108.12) seems to be the batch grinding kinetic equation, which has been solved analytically by Reid [5]. The solution is written (Fig. 108.2):

$$p_{i,k} = \sum_{j=1}^{i} A_{ij,k} \exp\left(-\frac{S_{j,k}}{v_z} z^*\right). \tag{108.14}$$

where

$$A_{ij,k} = \begin{cases} 0 & i<j \\ \sum_{l=j}^{i-1} \dfrac{b_{il}S_{l,k}}{S_{i,k}-S_{j,k}} A_{lj,k} & i>j \\ f_i^{IP} - \sum_{l=1}^{i-1} A_{il,k} & i=j \end{cases}. \tag{108.15}$$

In mathematics, $\lfloor E \rfloor$ is defined as the largest integer less than or equal to E. $\lfloor E \rfloor$ is defined as the smallest integer not less than E. So the particle size distribution of the edges product p_i^E can be calculated as:

$$p_i^E = \frac{1}{E}\left[\sum_{k=1}^{\lfloor E \rfloor} p_{i,k} + (E - \lfloor E \rfloor) p_{i,\lceil E \rceil}\right]. \tag{108.16}$$

The particle size distribution of the total product p_i^{HPGR} is :

$$p_i^{\mathrm{HPGR}} = \frac{1}{N_B}\sum_{k=1}^{N_B} p_{i,k}. \tag{108.17}$$

The particle size distribution of the center product is:

$$p_i^C = \frac{1}{1-a}\left(p_i^{\mathrm{HPGR}} - a p_i^E\right). \tag{108.18}$$

Table 108.1 Data of A ore and B ore

Test	Feed	A1	A2	A3	A4	A5	A6
Operating conditions							
U (m/s)		0.66	0.38	0.66	0.95	0.66	0.65
R_P (MPa)		4.0	6.0	6.0	5.9	7.5	7.5
Size (mm)	Percent passing (%)						
45.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
31.60	97.62	100.00	100.00	100.00	100.00	100.00	100.00
22.60	57.50	98.35	98.91	99.72	99.32	100.00	100.00
15.80	36.60	84.85	96.45x	92.53	92.15	94.92	94.64
11.10	25.70	71.78	88.76	83.84	83.27	84.16	89.16
7.80	20.33	61.83	78.94	74.72	72.02	75.16	81.02
5.50	15.92	52.84	68.92	65.19	62.62	66.72	70.42
2.70	10.51	39.20	54.62	49.90	47.65	51.80	55.51
1.00	6.73	25.23	36.41	32.68	32.01	34.98	38.03
0.50	5.10	19.50	28.68	25.67	24.45	27.76	30.20
0.30	4.35	16.68	24.96	22.29	21.48	24.16	26.41
0.20	3.41	14.05	21.45	19.15	18.64	20.87	22.72
0.13	2.48	11.73	18.36	16.45	15.94	17.83	19.81
power consumption		1.43	2.04	2.01	2.03	2.41	2.45
Test	Feed	B1	B2	B3	B4	B5	B6
Operating conditions							
U (m/s)		0.66	0.38	0.66	0.95	0.65	0.65
R_P (MPa)		3.8	5.9	5.9	5.9	7.4	5.8
Size (mm)	Percent passing (%)						
31.60	96.30	100.00	100.00	100.00	100.00	100.00	100.00
22.60	65.50	100.00	100.00	100.00	100.00	100.00	99.67
15.80	43.20	94.23	96.35	96.64	97.84	96.86	96.58
11.10	31.00	83.13	85.64	85.46	88.46	88.46	86.79
7.80	22.64	70.45	74.15	74.34	77.47	77.21	75.39
5.50	17.30	58.64	64.31	64.56	66.83	67.25	65.42
2.70	11.56	42.56	48.21	48.47	50.45	50.97	49.98
1.00	7.45	26.91	30.63	30.97	33.46	33.75	33.82
0.50	5.46	20.81	23.72	24.42	26.74	26.54	26.46
0.30	4.75	17.16	20.38	21.72	23.93	23.61	23.54
0.20	3.78	15.14	17.69	18.73	19.65	19.87	20.49
0.13	2.56	12.61	14.37	15.46	18.05	17.86	17.54
0.08	1.32	11.42	13.67	13.79	17.06	15.46	15.72
Power consumption		1.37	1.92	2.04	2.07	2.37	1.99

108.3 Experiment

To test the HPGR model, we used a HPGR manufacturing company, Changsha. Its diameter is 600 mm and length is 200 mm. The test uses two kinds of ore. Data are shown in Table 108.1.

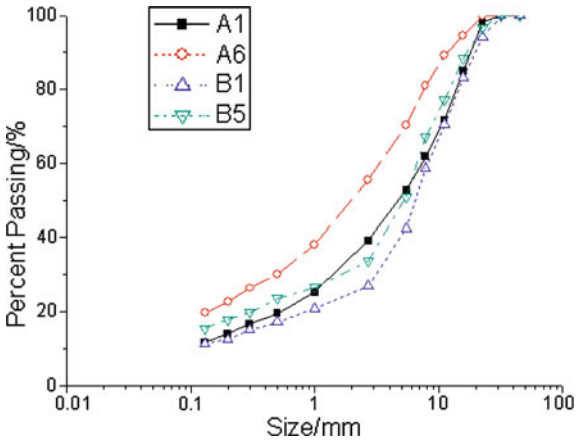


Fig. 108.3 The contrast of test data

These results show that for different operating pressures R_P, different product size distributions are obtained, as can be observed from the comparison of A1 versus A6 and B1 versus B5, under the same ore and different R_P. It is also evident that if the ore is different but the operating pressure is the same, different product size distributions are obtained. However, the effect is very is small, especially with a higher R_P, as can be observed from the comparison of A6 versus B5. With a lower R_P the differences are more evident (A1 vs. B1) (Fig. 108.3).

108.4 Conclusions

In bed compression crush, the crush of bed affects every particle crush. Model reflects the particle breakage and crush energy absorption material obey the basic principles. The smaller particle size, the greater power consumed.

In crush study, equal portions, N_B, of the roller can make throughput, power draw and the particle size more accurate.

It is also evident that if the ore is different but the operating pressure is the same, different product size distributions are obtained. However, the effect is very is small, especially with a higher R_P. With a R_P the differences are more evident.

References

1. Klymowsky R, Patzelt N, Knecht J, Burchardt E (2002) Selection and sizing of high pressure grinding rolls. In: Mular A, Halbe D, Barratt D (eds) Proceedings of mineral processing plant design, practice and control, vol 1. SME Inc, Littleton, pp 636–668
2. Daniel MJ, Morrell S (2004) HPGR model verification and scale-up. Miner Eng 17:1149–1161

3. Gan J (2008) Differential shear theory and mathematical model of high pressure double roll crush. Chin J Mech Eng 44(3):241–248
4. Li Y, Wang D (2003) High pressure material beds comminution. J Changsha Univ 17(4):36–39
5. Reid KJ (1965) A solution to the batch grinding equation. Chem Eng Sci 20:953–963

Chapter 109
Optimization of Performance Parameters of High Pressure Grinding Rolls

Mu Fusheng, Deng Ling, Liu Chao and Huang Sheng

Abstract In order to improve grinding effect of high pressure grinding rolls (HPGR), it is needed to optimize its performance parameters. According to nonlinear mapping function of artificial neural network and global optimal value search function of genetic algorithm (GA), this chapter, first, setted up the mathematical relationship between the given product size-reduction percentage M and performance parameters, such as working pressure P, roll speed V and original rolls gap S based on back propagation (BP) neural network, then, searched optimal performance parameters by GA. The result shows that using BP neural network with GA to optimize performance parameters of HPGR is accurate, fast and effective, which provides a novel approach for the selection of performance parameters.

Keywords Performance parameters · Optimization · BP neural network · Genetic algorithm

109.1 Introduction

The high pressure grinding rolls (HPGR) is a new comminution device that achieves comminution of beds of particles. Its principle of work is as follows: beds of particles are comminuted by pressure provided by hydraulic system between

M. Fusheng (✉) · D. Ling · L. Chao · H. Sheng
College of Mechanical and Electricity Engineering,
Central South University, Changsha 410083, China
e-mail: csuztg@163.com

D. Ling
e-mail: juidy123@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_109,

two rotated rolls [1]. At present, HPGR, with characteristics of energy and steel efficiency, product size-reduction fine, has been used for grinding of cement clinker, iron ores and so on.

Grinding product size distribution is one of the most important economic and technical performance indicators of grinding equipment, which reflects the work performance of crushers and directly impacts economical benefits of device users [2]. In all kinds of grinding assignments, users want to improve the fine product size-reduction percentage. Therefore, the research of product size-reduction of HPGR is a very theoretical and practical significant work. However, in process of grinding, the influence of grinding effect is varied [3–5], and the main factors are working pressure P provided by Hydraulic system, roll speed V and original rolls gap S. The traditional parameter selection mainly depends on the experience of designer, which often needs many tests and the efficiency is low. So it is necessary to establish the relationship between parameters and grinding effect and optimize the parameters, which are significant to improving effect and reducing energy consumption.

The artificial neural network can accomplish complex nonlinear mapping, and BP neural network is one of the most widely used network model [6]. Genetic algorithm (GA) has strong macrosearch ability and robustness, which can seek to the global optimal solution [7]. This chapter established the BP neural network model, that is to say, the nonlinear relationship between performance parameters and given product size-reduction percentage. Based on the trained BP network model, it took the output of network model as individual fitness value, and then searched the global optimization value of the parameters by using GA.

109.2 Establish BP Neural Model of Performance Parameters

To construct the BP neural network, it needs to define the follows: the layers of network, the number of neurons of each layer, the transfer function of each layer, the learning algorithm of network and the training and test samples.

109.2.1 Determine the Network Structure

As is known, without limiting number of hidden nodes, a three-layer network can achieve any nonlinear mapping. In this chapter, according to actual situation, a BP neural network was established which contained an input layer, a hidden layer and an output layer. The input layer has three nodes, corresponding to three input values: working pressure P, roll speed V and original rolls gap S. The output layer has one node, corresponding to objective value: given product size-reduction percentage, namely, percentage of product size-reduction that is smaller than a

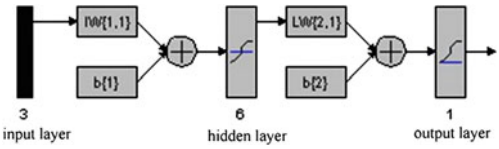


Fig. 109.1 Structure of BP network

Table 109.1 L_{16} $(4)^5$Orthogonal table

Cases	Factors		
	Working pressure P(MPa)	Original rolls gap S(mm)	Roll speed V(m/s)
Case 1	4	2	0.471
Case 2	6	4	0.628
Case 3	8	6	0.785
Case 4	10	8	0.942

certain calibration value to total product size-reduction. The nodes of hidden layer refer to the following formula:

$$l = \sqrt{m+n} + \alpha \tag{109.1}$$

where l is nodes of hidden layer, m is nodes of input layer, n is nodes of output layer and α is a constant between 0 and 10 [8]. Through cut and try method, it determined nodes of hidden layer for six. So the topological structure of BP network is 3-6-1, and the structure is shown in Fig. 109.1.

109.2.2 Select Training-Related Parameters and Samples

It selected tansig function as hidden layer transfer function, purelin as output layer transfer function and Levenberg–Marquardt back propagation algorithm as training algorithm, and sets learning rate as 0.01, mean square error as 10-4 and largest number of training as 500 epochs.

In order to make the design variables evenly distributed in variable space, it used the top three columns of orthogonal table L_{16} $(4)^5$ to arrange orthogonal test, and do the corresponding tests by the testing machine. Then it selected No. 1, 7, 12 and 14 as testing samples, the other as training samples. Table 109.1 shows the cases of factors.

109.2.3 Train BP Neural Network

After the network structure and the corresponding setting were determined, it started training. When the mean square error (MSE) reached the setting mean square error goal, training was over.

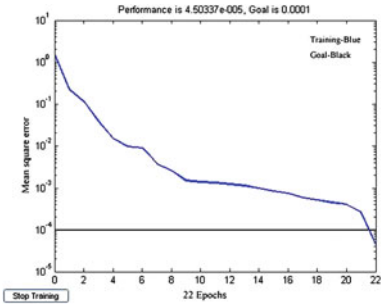


Fig. 109.2 Training process of the BP neural network

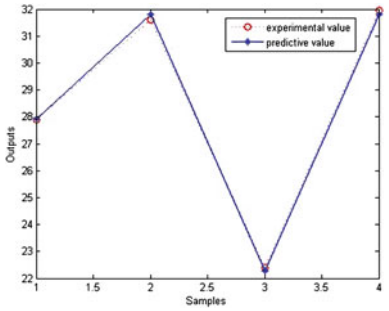


Fig. 109.3 Testing of the BP neural network

Table 109.2 BP network predictive value compared with experimental value

Nos. of sample	Predictive value	Experimental value	Error	Relative error
1	27.9215	27.83	0.0915	0.33%
7	32.2664	31.81	0.4564	1.43%
12	22.2087	22.27	0.0613	0.28%
14	31.5082	31.84	0.3318	1.04%

Figure 109.2 shows that after 22 epochs the training meets the required target error under 10^{-4}, and the training succeeds.

109.2.4 Test Network and Contrast Predictive Value With Experimental Value

Only by testing the reliability of network with experiential data that network can be used. It took No. 1, 7, 12 and 14 samples to test the neural network. Test results are shown in Fig. 109.3. The calculated error and relative error are shown in Table 109.2.

The table shows that the BP network predictive values mostly agree with the trail values, and achieve high precision. It demonstrates that the built BP neural

network has well-prediction performance. The known BP network can express the non-linear relationship between product size-reduction and performance parameters, which is helpful to select required parameters reasonably and predict the product size-reduction of the selected parameters.

109.3 Optimization Performance Parameters by Genetic Algorithm

109.3.1 Optimization Objective

Based on the former mathematical model, the relationship between performance parameters and optimization goal can be represented as:

$$m = f(P, S, V). \tag{109.2}$$

where P is working pressure, and it ranges: $0 < P \leq 10$ MPa; V is roll speed, and it ranges: $0 < V \leq 1$ m/s; S is original rolls gap, and it ranges: $0 < S \leq 10$ mm.

109.3.2 Optimization Calculation

Based on MATLAB genetic algorithm toolbox, it wrote a program. It took the binary code, the digits for each binary variables were 20, so the length of individual chromosomes was $3 \times 20 = 60$, and selected population scales as 30 and maximum number of generations as 200. The method of selection used random traverse sampling routines SUS, cross Crossing probability Px = 0.7 and took single point cross routines function xovsp, population mutation probability $P = \text{Px}/60$.

Objective function values of population after several iterations are shown in Fig. 109.4.

The evolution results of genetic algorithm are shown in Fig. 109.5.

The optimization results are $P = 8.8689$ MPa, $S = 4.9793$ mm, $V = 0.6291$ m/s and the corresponding $m = 35.1472$.

It took the optimized parameters into the former BP neural network as input, and the output forecast value $\text{m}' = 34.8169$, which is well agreed with the optimization value of genetic algorithm.

109.4 Conclusion

(1) Based on BP neural network, the nonlinear relationship established between optimization objective and design variables used in calculation and analysis of high pressure grinding roller is feasible.

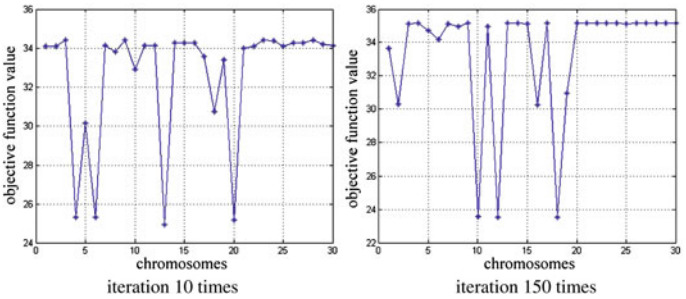
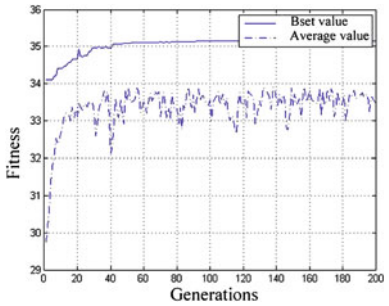


Fig. 109.4 Relationship of iteration times and objective function value

Fig. 109.5 Evolution of generations for optimization



(2) Strong global optimization ability of GA is suitable for solving the black box model established by BP neural network.
(3) Artificial Neural network and genetic algorithm are applied to optimize the grinding performance parameters of high pressure grinding roller. After optimization, the grinding effect is enhanced significantly, which provides a new way for optimization of performance parameters of high pressure grinding rolls.

References

1. Liu J (2010) Progresses in mathematical modeling and computer simulation of comminution with high-pressure grinding rolls. Metal Mine 39(8):1–6
2. Huang D, Fan X, Wu D, Yao F (2008) Crushing product size-reduction analysis of compressing crusher. Chin J Mech Eng 44(5):201–207
3. Hasanzadeh V, Farzanegan A (2011) Robust HPGR model calibration using genetic algorithms. Miner Eng 24(5):424–432
4. Lim WIL, Campbell JJ, Tondo LA (1997) The effect of rolls speed and roll surface pattern on high pressure grinding rolls performance. Miner Eng 10(4):401–419
5. Persio R, Robert H, Mike G, Bern K (2010) A preliminary investigation into the feasibility of a novel HPGR-based circuit for hard, weathered ores containing clayish material. Miner Eng 15(01):290–302

6. Zhang D (2009) MATLAB simulation and application of neural network. Publishing House of Electronics Industry, Beijing
7. Lei J (2005) MATLAB genetic algorithm toolbox and application. Publishing House of Xidian University, Xi'an
8. Shen H, Zhao X, Gao C, Qing J (2008) Determining the number of BP neural network hidden layer units. J Tianjin University of Technol 23(06):13–15

Chapter 110
Design of the Human–Machine Interface of the Control System of a Novel Controllable Mechanism Type Mechanical Excavator

Yuchen Pan, Ganwei Cai, Jinling Zhang and Hongzhou Wang

Abstract This paper introduces the design of the Human–Machine interface of the control system of a controllable mechanism type planar 3 Degree-of-Freedom mechanical excavator. Based on the kinematics model and the control principle block diagram of the excavator, the interface is established in the C++ programming language, which consists of five parts: (i) speed setting module, (ii) position setting module, (iii) driving information module, (iv) the start button and (v) the stop button. The program codes of each module are given specifically.

Keywords Excavator · Controllable mechanism · Human–Machine interface · C++

Y. Pan (✉) · G. Cai · J. Zhang · H. Wang
College of Mechanical Engineering, Guangxi University,
Nanning 530004, China
e-mail: gxu_panyuchen@163.com

G. Cai
e-mail: caiganwei@163.com

J. Zhang
e-mail: jin.ling.521@163.com

H. Wang
e-mail: gxuwanghongzhou@163.com

Y. Pan
College of Civil Engineering and Architecture, Guangxi University,
Nanning 530004, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_110,

110.1 Introduction

Excavators are widely used in surface mining, construction, and geotechnical operations, which can be divided into two types: hydraulic excavators and mechanical excavators. Hydraulic excavators are widely used because they can achieve flexible output trajectories [1]. But the hydraulic system uses fluid as the operating media with a large number of hydraulic components. Without proper maintenance, hydraulic leaks are inevitable. Mechanical excavators (also called Electric Shovel or Cable Shovel) are the excavators of choice in surface mining etc. for their durability, large breakout forces, but lack flexibility because of their structural limitations. To solve these problems, a novel mechanical excavator was introduced [2]. Theoretical model of kinematics of the proposed novel excavating mechanism was developed in Ref. [3]. The study on the control method of the excavator was carried out in Ref. [4].

Digital signal processing (DSP) is concerned with the representation of discrete time signals by a sequence of numbers or symbols and the processing of these signals, including: signal processing for control of systems, audio and speech signal processing, sonar and radar signal processing, sensor array processing, spectral estimation, statistical signal processing, digital image processing, etc.

Being regarded as an intermediate-level language, as it comprises a combination of both high-level and low-level language features, C++ is a statically typed, free-form, multi-paradigm, compiled, general-purpose programming language.

This paper introduces the design of the human–machine interface of the control system of the above-mentioned novel controllable mechanism type mechanical excavator. The configuration and kinematics model of the proposed excavator is introduced first. Then the control principle block diagram of the control system is given. Subsequently, the interface is established in the C++ programming language, which consists of five parts: (i) speed setting module, (ii) position setting module, (iii) driving information module, (iv) the start button and (v) the stop button. The program codes of each module are given specifically.

110.2 The Proposed Controllable Mechanism Type Novel Planar 3-DOF Mechanical Excavator

Referring to Fig. 110.1, the proposed new design is composed of a base (which can be attached to a traditional hydraulic excavator base such as a crawler), a bucket and two kinematic chains. One chain is essentially a planar 2-DOF five-bar controllable mechanism accomplishing complex trajectories in its working plane, of which the purpose is identical to the joint operations of the boom and the stick of the hydraulic excavator.

Referring to Fig. 110.2, the planar 2-DOF five-bar controllable mechanism (Link 1, Link 2, Link 3, Link 4 and the base) can accomplish complex trajectories

Fig. 110.1 Prototype of the new excavator

in its working plane, of which the purpose is identical to the joint operations of the boom and the stick of the hydraulic excavator. The other chain is a planar four-bar mechanism (Link 5, Link 6, Link 7 and the base) by which the bucket is actuated independently. Driven links are Link 1, Link 4, and Link 5, controlled by three servo-motors placed on the base separately. The output posture and position of the mechanism is a function of the three input variables. The position and posture of the teeth of the bucket Q, writes $Q = \{ x_q \quad y_q \quad \theta_q \}^T$, is determined by a given $M = \{ x \quad y \quad \theta \}^T$, and $M = \{ x \quad y \quad \theta \}^T$ is defined by the three actuated variables $\Theta = \{ \theta_1 \quad \theta_4 \quad \theta_5 \}^T$, say

$$A(\dot{\theta}_1 \ \dot{\theta}_4 \ \dot{\theta}_5)^T = B(\dot{x} \ \dot{y} \ \dot{\theta})^T \tag{110.1}$$

in which

$$\mathbf{A} = \begin{bmatrix} \mathbf{N} & \mathbf{0} & \mathbf{0} \\ \mathbf{P} & \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} \end{bmatrix}$$

110.3 Human–Machine Interface of the Control System

110.3.1 Control Principle of the Excavator

Based on the motion characteristics introduced in Sect. 110.2, the motion control card (DSP) can be programmed through a computer. After receiving the control

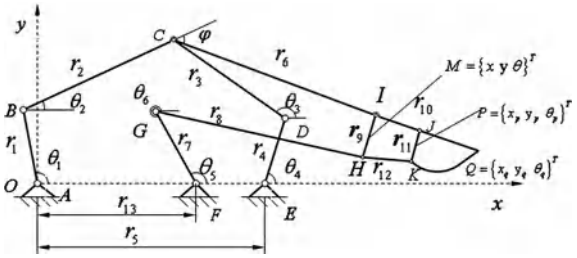


Fig. 110.2 Sketch of the new excavator

commands of the computer, the card sends out the pulse signals to the AC servo drives of the AC servo motors. Then the executive mechanism of the excavator can achieve the given output trajectories, as shown in Fig. 110.3.

The Human–Machine interface of the control system of the proposed novel excavator is designed in the C++ programming language mentioned in Sect. 110.1, which consists of five parts: (i) speed setting module, (ii) position setting module, (iii) driving information module, (iv) the start button, and (v) the stop button, which is shown in Fig. 110.4.

110.3.2 Program Codes

Program codes of the five parts of the interface are proposed specifically as follows:

(i) Speed setting module, which is used to set the initial velocities, driving velocities, accelerations.

```
int CCtrlCard::Setup_Speed(int axis, long startv, long
speed,
long add)
{
   if(startv-speed>=0){
      Result=set_startv(0, axis, startv/MULTIPLE);
      set_speed(0, axis, startv/MULTIPLE);
    }
else{
      Result=set_startv(0, axis, startv/MULTIPLE);
      set_speed(0, axis, speed/MULTIPLE);
      set_acc(0, axis, add/125/MULTIPLE);
   }
   return Result;
   }
```

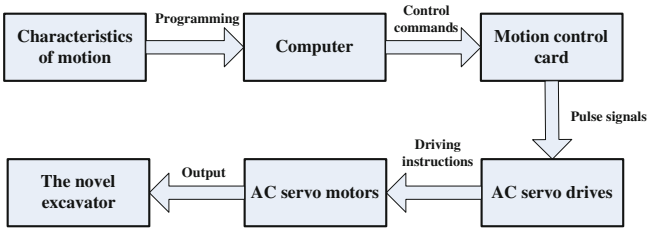


Fig. 110.3 Control principle block diagram of the Excavator

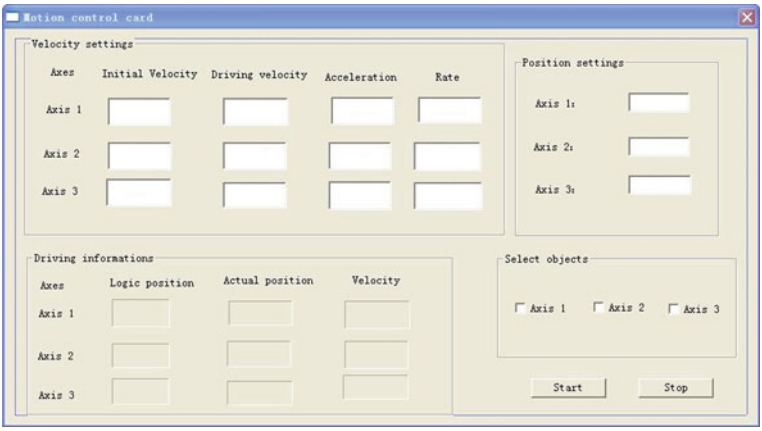


Fig. 110.4 Man-machine interface

(ii) Position setting module is used to set the step number of the servo motors.

```
int CCtrlCard::Axis_Pmove(int axis, long value)
{
Result=pmove(0, axis, value);
return Result;
}
```

(iii) Driving information module displays the real time logic positions, actual positions, and running velocities of each servo motors.

```
int CCtrlCard::Get_CurrentInf(int axis, long &LogPos,
long
&ActPos, long &Speed)
{
Result=get_command_pos(0, axis, &LogPos);
      get_actual_pos(0, axis, &ActPos);
      get_speed(0, axis, &Speed);
```

```
        Speed=Speed*MULTIPLE;
        return Result;
}
```

(iv) The start button.

```
void CVCExampleDlg::OnButtonPmove()
{
        UpdateData(TRUE);
long Startv[]={m_nStartvX, m_nStartvY, m_nStartvZ};
long Speed[]={m_nSpeedX, m_nSpeedY, m_nSpeedZ};
long Add[]={m_nAddX, m_nAddY, m_nAddZ};
if(m_bX)
{
g_CtrlCard.Setup_Speed(1, startv[0], Speed[0],
Add[0]);
g_ CtrlCard.Axis_Pmove(1, Pos[0]);
}
if(m_bY)
{
        g_CtrlCard.Setup_Speed(2, startv[1], Speed[1],
Add[1]);
        g_ CtrlCard.Axis_Pmove(2, Pos[1]);
}
if(m_bZ)
{
g_CtrlCard.Setup_Speed(3, startv[2], Speed[2],
Add[2]);
        g_ CtrlCard.Axis_Pmove(3, Pos[2]);
}
}
```

v) The stop button is used to stop the pulse signals of the control system.

```
void CVCExampleDlg::OnButtonStoprun()
{
for(int i=1; i<4; i++){
        g_CtrlCard.StopRun(i, 0);
      }
}
```

110.4 Conclusion

This paper introduced the design of the human–machine interface of the control system of a novel controllable mechanism type mechanical excavator. The configuration and kinematics model of the proposed excavator were introduced first. Then the control principle block diagram of the control system was given. Subsequently, the interface was established in the C++ programming language, which consists of five parts: (i) speed setting module, (ii) position setting module, (iii) driving information module, (iv) the start button and (v) the stop button. The program codes of each module were given specifically.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (No. 51075077), Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology Project (09–007–05S013) and Teams for Innovation in the Construction of Talent Highlands in Guangxi Institutions of Higher Learning.

References

1. Maciejewski J, Jarzgbowski A (2004) Study on the efficiency of the digging process using the model of excavator bucket. J Terramechanics 40:221–233
2. Cai GW, Pan YC (2011) China Patent 200910114447.0
3. Cai GW, Zhang Z, Pan YC (2011) Kinematics analysis and simulation of a new type of mechanical excavator with controllable mechanism. J Adv Mater Res 201–203:220–224
4. Cai GW, Wu DC, Pan YC (2011) Research on control method of a novel type of excavator based on multi degree-of-freedom controllable mechanism. J Adv Mater Res 201–203: 1949–1954

Chapter 111
The Design of Constant Pressure Water Supply System Based on Configuration Monitoring and PLC Control

Zongchao Cui, Yingli Lv and Xinjun Zhang

Abstract To describe constant pressure water supply system of speed-adjusting based on PLC control, at the same time, the real-time monitoring is managed by the configuration software of upper computer, this design is energy-saving, safety and high quality, the water supply is high reliability quality and convenient and flexible control etc., it is proved fully realize automatic constant pressure water supply, which can increase water supply quality and reduce running costs.

Keywords Configuration technology · PLC · Constant pressure · Water supply

111.1 Introduction

China is a country that is short of water and electricity. It is one of the largest water consuming countries; the current annual water consumption is about 550 billion m^3, and accounts for 12% of the world's total water consumption. For a long time, the technology of industry of our country water level is relatively backward, low degree of automatic control, saving energy and water will have a larger

Z. Cui (✉) · Y. Lv · X. Zhang
Department of Electrical Engineering, Jiyuan Vocational and Technical College, Jiyuan 459000, Henan, China
e-mail: cuizongchao@126.com

Y. Lv
e-mail: xalvyingli@163.com

X. Zhang
e-mail: zhangxinjun2006-126@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_111,

development space. For example, in the industrial water or water supply system of residents' lives, because, because the instability of water users, it will appears in the peak water, low water pressure, water supply below demand; Especially the high-rise residential users to appear too low to get water pressure of the phenomenon. The traditional solution is to use water tower, high pressure tanks or Air pressure can was water storage and pressurization. But the disadvantage is obvious: water pollution and low water pressure; it can not meet need of people at present.

For the above, this paper will combine PLC and inverter speed control technologies, it uses PID control algorithm of PLC to achieve online control and closed loop constant pressure supply water; at the same time, for the process of constant pressure water supply in real-time monitoring and dynamic display screen, the introduction of configuration techniques, and use the force control configuration software to complete.

111.2 Composition and Working Principle of This System

111.2.1 System Composition

The system consists of PLC, inverter, two three-phase water pumps, a pressure transmitter and other control equipment, as shown in Fig. 111.1 inverter realizes three-phase water pump soft start-up and frequency control, the current pressure sensor detection the water pressure signal converts to a standard pressure signal that is from 4 to 20 mA with pressure transmitter. This signal sends into the PLC through the A/D converter module, PLC to do logic operations, and then operation result will send to frequency converter which controls output frequency to change the pump speed according to the result. At the same time pressure value also from the PLC to UPPER COMPUTER, after the dynamic display is used by configuration software; from UPPER COMPUTER to PLC sends control signal to realize the pump start, stop and switching control.

111.2.2 Working Principle of System

Constant pressure water supply system used continuous closed-loop PID control, according to the given value and actual value, in accordance with system adjusting signal is sent by PLC and PID logical operation, this signal can control the output frequency of frequency converter to realize the pump unit, and also adjust pipe network water pressure. The working principle of constant pressure water supply system can show continuous single closed-loop PID control system, as shown in Fig. 111.2. When the system is working, the pressure sensor converts the pressure

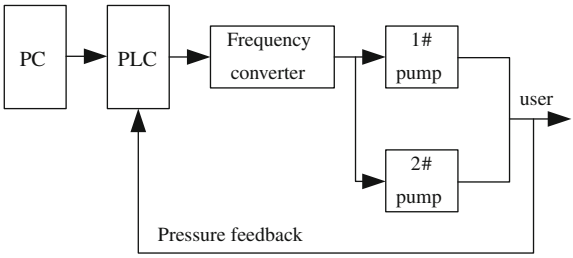


Fig. 111.1 System composition

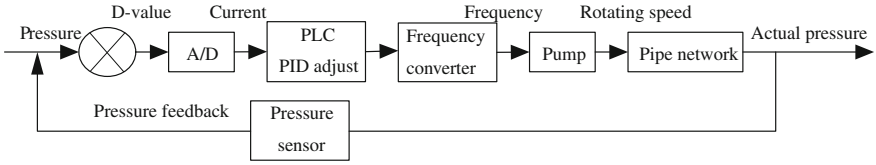


Fig. 111.2 Diagram of working principle

signal into electrical signals which is about from 4 to 20 mA by pressure transmitter in the pipe network, and then sends to the PID regulator of PLC, PID adjustor makes processing for a given and measured value in accordance with certain rules.

For example, the adjustment of the regulator to set rules proportion (P)-integral (I)-differential (D) regulation rule. And then based on the user, who set PID parameter, calculation output signal to control the working frequency of frequency converter, and adjust the water pump motor speed to realize the pipeline network flow requirements [1].

The detailed working principle: the pressure sensor sends the actual pressure after the feedback in pipe network into the comparator input, to the comparator, and comparing a given pressure, when the pressure pipeline network is insufficient, through the parameters of the operation to adjust the PID parameters, corresponding to the frequency of the inverter increases, the pump speed to accelerate, the water supply will increase, and the pipe network water pressure to rise. Conversely, if the pump speed slows down, the water supply decrease, the pipe network pressure drops to maintain constant pressure water supply and stable running.

We set the two service pumps in this water supply system, when the water is needed low, the inverter control 1 water pump to constant pressure water supply, when the water consumption gradually increasing, as No1 water pump working frequency also increase, When No1 water pump working frequency reach the highest working frequency 50 Hz, and supply pressure less than the requirement, the No1 pump will switch to the industrial frequency power supply. At the same time frequency converter is switched to 2 water pump that supplement water supply. When water consumption decrease gradually, even if No2 pump working

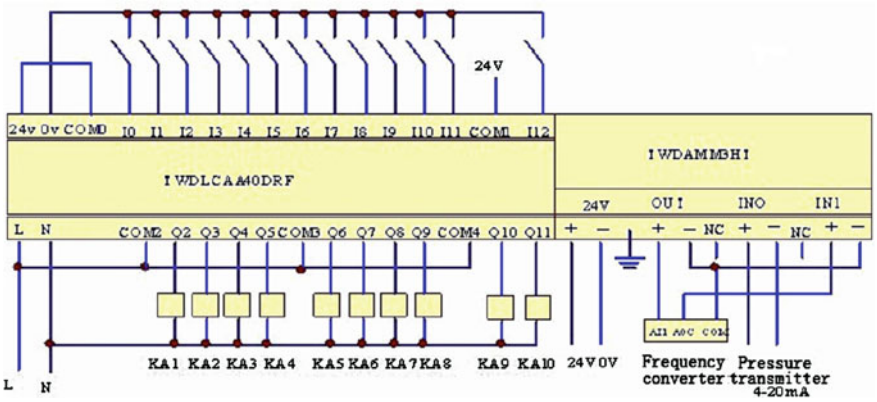


Fig. 111.3 Chart of hardware cables

frequency has been reduced to 0 Hz and water supply water pressure is too great, turn off the power supply by industrial frequency power 1 water pump, and go up the second water pump to realize constant pressure control [2].

111.3 Choice of System Hardware

This system chooses TWIDO series that model to be: TWDLAE40DRF, a body PLC, 24 inputs, 2 transistor outputs, 14 relay outputs, with Ethernet RJ-45 ports, support for Ethernet communications, with PID etc. Inverter is the Schneider ATV31, analog expansion module chooses is the two inputs and one output TWDAMM3HT type.

The system hardware connection diagram is shown in Fig. 111.3. The analog module and frequency converter is directly connected, analog input signal is two roads, the one is pressure transmitter to send a signal to the current IN0; the other road is the frequency feedback signal of frequency converter to IN1. Analog output one road voltage signal directly to send the AI1 port of frequency converter, adjust the frequency of frequency converter.

111.4 Design of System Software

In our design, through the PLC internal PID regulator to adjust the output, because PLC is a sampling control, it can calculate the control value according to the error of the sampling time. In practical application, that needs the information PID control equation converted into the PID equation of the practical application. Is the continuous formula into discrete sampling error formula, can cycle is used to

calculate the output value. So must the simulation PID discrete formula. After using differential transformation method to the simulation PID adjustor output formula for discrete difference equation, it can calculate the output value.

The ideal analog PID control following Eq. 111.1:

$$M(t) = K_C\left[e + \frac{1}{T_1}\int_0^t edt + T_D\frac{de}{dt}\right] + M_{\text{initial}} \tag{111.1}$$

A sampling period for TS, initial moment is 0, the n-th time sampling of deviation is en. The control output is Mn, and the following transformation Eq. 111.2–111.4:

$$de \approx \Delta e = e_n - e_{n-1} \tag{111.2}$$

$$dt = \Delta t = t_n - t_{n-1} \tag{111.3}$$

$$\int_0^t e(t)dt \approx \sum_{i=0}^{n} e_i T_s \tag{111.4}$$

The discretization form of analog PID adjustment is the following Eq. 111.5:

$$M_n = K_C\left[e_n + \frac{T_S}{T_1}\sum_{i=0}^{n} e_i + \frac{T_D}{T_S}(e_n - e_{n-1})\right] + M_{initial} \tag{111.5}$$

where T_S, M_n, K_C, e_n, $e_{n\text{-}1}$, and n are the sampling periods, so that adjust the output value of nth time, PID loop gain, the n-th time sampling bias, $e_n = SP_n - PV_n$ and the first $n-1$ the sampling frequency number.

It can be seen from the above formula, the integral item is the function of all error of poor items from the 1st to the current sample period; the differential item is the current sample and previous sample function; the proportional item is only a function of the current sample. In the computer, don't save the entire bad item.

As the computer from the first sample, every bias sample value must be calculated the first output value, that only need to save the former value of bias and the former value of integral item. With repetitive computer processing, the above formula can be simplified as Eq. 111.6:

$$M_n = K_C\left[e_n + \frac{T_S}{T_1}e_n + \frac{T_D}{T_S}(e_n - e_{n-1})\right] + M_X \tag{111.6}$$

where M_n, K_C, e_n, e_{n-1} and M_X are respectively the n time sampling instant, the calculated value of PID loop output, PID loop gain, the n time sampling deviation, the first $n-1$ sampling time deviation, and the before value of integral.

111.5 Monitoring Design of System

In order to be able to better display the constant pressure water supply system composition and real-time status of each valve and pump, read various water pressure and flow rate of parameters, and the alarm real-time display and fault record etc., In order to better display the constant pressure water supply system components, the configuration software of power control that type is UPPER COMPUTERAuto 6.0 as a management software of UPPER COMPUTER, via data communication between UPPER COMPUTERAuto 6.0 and PLC to realize remote systems monitoring.

Here is the design process of this monitoring system.

111.5.1 New Project

In the interface of power control project management, you can click on "New", define the project name which is "constant pressure water-supply", complete the new project creation, click on the "development system" button, and then enter the interface of "development system" that is ready for system development [3].

111.5.2 Peripheral I/O Connectivity Is Defined

In the force control, devices or program that need to exchange data with force control configuration software is used to I/O devices, these devices exchange data with UPPER COMPUTER by means of serial port or Ethernet; After I/O device is only defined, the power control can exchange data through the database variables.

Double-click "I/O device driver" item to start the commencement of the project in navigator, choose double-click "PLC", and then continue to select double-click the vendor name which is "Modicon", at the last select "MODBUS" which is ASCII & RTU TCP/IP communication [4].

Double-click the "MODBUS" appears the dialog box "I/O device definition", you input a person defined in the input box of "Device Name", and you input the address which is 1 in the input box of "Device Address", communications options "Serial Port Communication".

111.5.3 Database Variable Is Defined

The database is the core of the whole application system and the basis for building distributed applications. It is responsible for real-time data processing of the whole

power control applications, historical data storage, statistical data processing, alarm processing, data service requests processing. In the database, we manipulate the object that is point of TAG, the real-time database decides the structure of the database and allocates the database storage space according to the dictionary [5].

111.5.4 I/O Data Connection Is Established

The I/O data connection is established, that main correspond between the point parameters of database and the channel address of acquisition device, the communication is established so that each site point of operation can be responded in the monitor software of UPPER COMPUTER between each I/O points of UPPER COMPUTER and corresponding I/O address of PLC.

111.5.5 Configuration Picture Design and Animation Connection Created

Using drawing tools and libraries in software can draw the monitor picture of constant pressure water supply system.

After completion of the animation design, animation can be connected, that is established a relationship between graphics object of animation and variables, when the value of a variable changes, the graphic objects on the animation reflects dynamic changes.

111.5.6 Running

After the initial establishment of power control projects complete, enter the running phase. First you save all the configuration contents, close DBManager if not closed.

You can select menu which is "File/enter Run" in the development of the power control system which is DRAW, and then enter the system of force control running.

111.6 Conclusion

Practical operations have proved that the constant pressure water supply system has energy saving, safety and high quality of water supply quality, and many other advantages. PLC as a controller, which hardware structure is simple, low cost and

of ensure high reliability control. At the same time a combination of PLC and inverter to realize the smooth start of pump, avoid the impact of the power grid starting; in the operation of the pump can realize stepless speed regulation so that it can extend the service life of pumps and valves, this has good economic and social benefits. Also, the introduction of configuration software can dynamically display the system real-time running data and state, save fault information and performance data. This system can replace the traditional supply model to improve water quality, and realize the purpose of automatic constant pressure water supply.

References

1. Zhang W (2009) Programmable controller introduction and examples of application. China Electric Power Press, China
2. Xiao P (2008) Inverter and control technology. Mechanical Industry Press, Beijing
3. Jin B (2009) PLC-based constant pressure water supply system. Autom Technol Appl 15:121–123
4. Wang W (2009) Based on PLC and the force control configuration software to realize the temperature control system. Chem Eng Equip 21(08):82–86
5. Zhongda L, Fengxia X, Lijuan S (2008) Level control system based on fore control configuration software. Micro Comput Inf 28(11):77–78

Chapter 112
Analysis of the UHV Tie Line Active Power Peak Value of Weak Interconnected Grids Following Power Shortage Disturbance in China Power System

Feng Hong, Jinfu Chen and Xianzhong Duan

Abstract An ultra high voltage (UHV) AC tie line that connects North China and Central China power grid has been put into operation in 2009. This paper presents a method to calculate the peak value of active power oscillation of the UHV AC tie line following power shortage disturbance. The typical process of the oscillation on the UHV AC line is analyzed. It is shown that the active power oscillation on tie line of week connection grid is similar with the step response of the second-order system. The oscillation peak value is closely related to fault location, system inertia, and system inter-area damping. The method and analysis are tested in China system.

Keywords UHV AC tie line · Oscillation · Week connection system · Oscillation peak value

112.1 Introduction

Jindongnan-Nanyang-Jingmen 1000 kV AC transmission line [1, 2] which was put in operation in China on January 6th, 2009 is the unique ultra high voltage (UHV) [3–5] AC transmission line in service in the world. North and Central China power grid, two isolated power systems, were connected by the UHV AC line. Following the large disturbances, like generator tripping or dc blocking, the power flow

F. Hong (✉) · J. Chen · X. Duan
State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Electrical Power Security and High Efficiency Lab, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China
e-mail: ceeehongfeng@foxmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_112,

oscillation occurs on the UHV tie line. This is so-called "inter-area" mode phenomenon in stressed power system.

The literature [6] indicates that the large disturbances can trigger inter-area oscillation as well as small signal. Previously, researches on inter-area oscillation focus on the damping of the oscillation [7–13]. In China, the North and Central China Power Grid being connected by the UHV tie line has strong inter-area damping, and the oscillation on the tie line could persist for several swings. However, since the system is weakly connected by the UHV line with low static stability limit, large oscillation of power flow on UHV tie line could get system collapse when power flow on the UHV line power exceeds steady-state stability limit.

It is can be observed from operation data of the UHV tie line that if a disturbance, such as tripping of a generator, occurs in Central China grid (CCG), the peak value of oscillation of power flow is about 90% of the power shortage. However, the peak value is about 70% of power shortage when the same fault occurs in North China grid (NCG). Since peak value of active power flow over the UHV tie line plays a key role in stability of system, the peak value of active power flow over the UHV tie line is analyzed in this paper. The paper is organized as follows. Section 112.2 introduces background of the china power system with after UHV systems and some large disturbances of interest. A method to calculate the peak value of tie line active power of inter-area mode following large disturbances is proposed in Sect. 112.3. The proposed method is utilized to calculate the peak value of active power over UHV tie line in Sect. 112.4. Section 112.5 concludes the paper.

112.2 Case Study of Power Oscillation on UHV Tie Line

The 61$^{\#}$ unit with an output of 600 MW in Sichuan JinTang Power Plant was tripped on March 30, 2009 due to maloperation of over speed protection. The variation of active power on UHV tie line is shown as Fig. 112.1. The active power increased from 193 to 253 MW in the first swing (0–4 s). The variation in active power on UHV tie line equaled to output of the tripped unit. The active power kept around 230 MW from 4 to 20 s. Thereafter, the UHV tie line active power decreased slowly when the AGC of other generator work. It can be observed that the oscillation frequency is about 0.16 HZ and the damping ratio is about 0.15, which equals to the system "inter-area" mode.

The details of major oscillation in active power over UHV AC line due to disturbance present in CCG and NCG since 2009 are listed as Tables 112.1 and 112.2. Since the oscillation on china UHV line always attenuate, the transfer ratio is utilized to evaluate the oscillation level. The transfer ratio is percentage of max oscillation active power and the active power shortage of fault. The data listed in Tables 112.1 and 112.2 indicate that the transfer ratio is around 100% for the fault present in CCG and around 70% for the fault present in NCG.

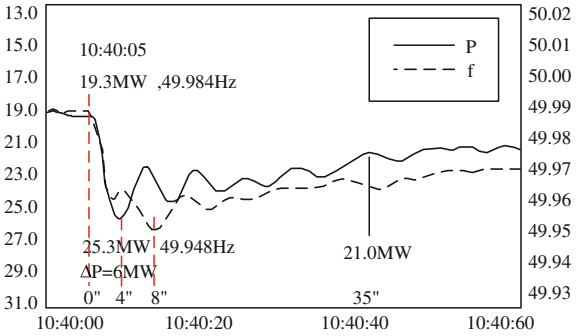


Fig. 112.1 Active power oscillation curve of the UHV tie line after tripping a generator in CCG

Table 112.1 Details of oscillation with fault present in CCG

Time	Fault type	Power shortage (MW)	UHV AC Line (MW)		Transfer ratio (%)
			Initial	Peak	
2009/3/30	Unit tripping	600	1930	2530	100
2009/5/15	Unit tripping	1400	1100	2450	96.4
2009/6/4	Load Shedding	700	1104	412	98.9
2009/7/9	DC blocking	660	1696	2477	113.2
2009/7/10	DC blocking	640	1550	2430	123.9
2009/7/23	Unit tripping	1930	1160	−600	93.3
2009/8/12	DC blocking	650	1026	1573	99.5
2009/8/16	DC blocking	670	910	1480	94.2
2009/12/1	DC blocking	1110	1470	2570	99
2010/2/26	DC blocking	950	1410	2350	99

Table 112.2 Details of oscillation with faults present in NCG

Time	Fault type	Power shortage (MW)	UHV AC Line (MW)		Transfer ratio (%)
			Initial	Peak	
2009/4/5	Unit tripping	950	760	60	73.7
2009/7/8	Unit tripping	900	1120	1782	73.6

112.3 The Peak Value of Power Oscillation of UHV AC Tie Line

112.3.1 Dynamic Phenomenon Analysis

The Inter-area oscillation could arise from either small signal disturbance or large disturbances. For the oscillation analyzed in this paper, a small group of generators close to the fault location are severely disturbed and the other generators (a large group) are disturbed as same as small signal disturbance following the large disturbance. The local oscillation is damped to synchronize soon if the system does not lose stability. The UHV tie line active power fluctuation is mainly composed by the

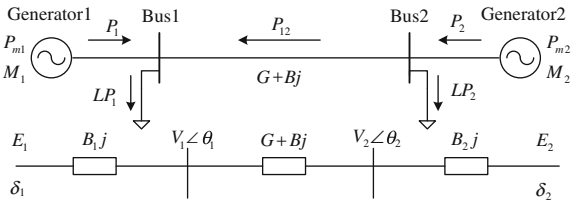


Fig. 112.2 Two-machine system

Inter-area oscillation mode. Since the CCG and NCG is interconnected weakly, the generator in each grid synchronized quickly. Therefore, two-machine system can be utilized to calculate the peak value of the active power fluctuation on tie line.

112.3.2 Equivalent Two Machine System

The Equivalent two machine system of weakly interconnected power system is shown in Fig. 112.2.

The classical model of generators is utilized in the paper. The equations of motion for the system is:

$$M_i\ddot{\delta}_i + D_i\dot{\delta}_i = \Delta P_i \quad i = 1, 2 \tag{112.1}$$

Suppose the fault occurs at time t_0, from t_0 on

$$\Delta P_i = \Delta Pm_{i0} + \Delta P_{Li} \pm \Delta P_{\text{tie}} \tag{112.2}$$

where ΔPm_{i0} denotes the value of electromagnetic mutation of generator at time t_0^+, ΔP_{Li} denotes the variation of load, ΔP_{tie} denotes variation of UHV tie line active power, M_i denotes inertia constant, and $\dot{\delta}_i$ denotes generator angle.

It should be pointed out that the power shortage (ΔP) associated with the fault equals to sum of ΔP_i according to the Power Balance. Use ΔP_{tie}, Eqs. 112.1 and 112.2 can be written as

$$\begin{cases} \ddot{\delta}_{12} = \left(\dfrac{\Delta Pm_{10}}{M_1} - \dfrac{\Delta Pm_{20}}{M_2}\right) - \left(\dfrac{\Delta P_{12}}{M_1} + \dfrac{\Delta P_{12}}{M_2}\right) - \left(\dfrac{D_1\dot{\delta}_1}{M_1} - \dfrac{D_2\dot{\delta}_2}{M_2}\right) + \left(\dfrac{\Delta P_{L1}}{M_1} - \dfrac{\Delta P_{L2}}{M_2}\right) \\ M = \dfrac{M_1 M_2}{M_1 + M_2} \end{cases} \tag{112.3}$$

Since the angle difference between the UHV bus is quite small (smaller than 15°) and the generator transient reactance is much smaller than the UHV line reactance, the tie line active power can be calculated as

$$\Delta P_{\text{tie}} = k_{\text{tie}}\Delta\delta_{12}, \quad k_{\text{tie}} = BV_1V_2 \tag{112.4}$$

where $\Delta\delta_{12}$ denotes generator angle difference.

Considering the load frequency characteristic, the load variation is

$$\Delta P_{Li} = \lambda_i(\omega_i - \omega_0) \tag{112.5}$$

and

$$\left(\frac{\Delta P_{L1}}{M_1} - \frac{\Delta P_{L2}}{M_2}\right) = \left[\frac{\lambda_1(\omega_1 - \omega_0)}{M_1} - \frac{\lambda_2(\omega_2 - \omega_0)}{M_2}\right] \tag{112.6}$$

where, λ_i is proportional to the grid generated energy P_i for both grids with similar load composition. Since the two grids are weakly connected, the power transmitted over UHV line is much smaller than grid generated energy. Therefore, $\lambda_i \propto M_i$ and Eq. 112.6 can be written:

$$\left(\frac{\Delta P_{L1}}{M_1} - \frac{\Delta P_{L2}}{M_2}\right) = \frac{\lambda_1}{M_1}\dot{\delta}_{12} \tag{112.7}$$

In equivalent system, each generator equals to a grid. Therefore, the damping coefficient is proportional to the inertia.

$$\left(\frac{D_1\dot{\delta}_1}{M_1} - \frac{D_2\dot{\delta}_2}{M_2}\right) = D\dot{\delta}_{12} \tag{112.8}$$

Using Eqs. 112.1–112.8, we obtain:

$$\ddot{\delta}_{12} = \left(\frac{\Delta Pm_{10}}{M_1} - \frac{\Delta Pm_{20}}{M_2}\right) + \lambda_1\dot{\delta}_{12}/M_1 - D\dot{\delta}_{12} - (k_{\text{tie}}\delta_{12} - k_{\text{tie}}\delta_{120})\left(\frac{1}{M_1} + \frac{1}{M_2}\right) \tag{112.9}$$

As shown in (112.4), we have:

$$\Delta\ddot{P}_{\text{tie}} = \Delta\ddot{\delta}_{12}, \quad \Delta\dot{P}_{\text{tie}} = \Delta\dot{\delta}_{12} \tag{112.10}$$

According to (112.10), Eq. 112.9 can be expressed in the form of second-order differential equation:

$$a\Delta\ddot{P}_{\text{tie}} + b\Delta\dot{P}_{\text{tie}} + \Delta P_{\text{tie}} = c \tag{112.11}$$

where $a = \dfrac{M_1M_2}{M_1 + M_2}$, $c = \left(\dfrac{\Delta Pm_{10}}{M_1} - \dfrac{\Delta Pm_{20}}{M_2}\right)\left(\dfrac{M_1M_2}{M_1 + M_2}\right)$, $b = \left(D - \dfrac{\lambda_1}{M_1}\right)\left(\dfrac{M_1M_2}{M_1 + M_2}\right)$.

The analytical solution of ΔP_{tie} can be calculated using Eq. 112.10 and the peak value of active power over UHV AC tie line can be calculated using Eq. 112.4.

112.3.3 Step Response of Second-Order System

The fault analyzed in this paper (generator tripping or dc blocking) can be viewed as a step disturbance to the system. Therefore, the active power fluctuation on UHV tie line is similar with the step response of second-order system.

The characteristic equation for the second-order system can be described as

$$s^2 + 2\xi\omega_n + \omega_n^2 = 0 \tag{112.12}$$

The two roots are

$$s_{1,2} = -\xi\omega_n \pm j\omega_n\sqrt{1-\xi^2} \tag{112.13}$$

where ω_n is natural oscillation angular frequency without damp. ξ is the damp ratio. When $\xi \in (0,1)$ the step response of second-order system is:

$$g(t) = 1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\xi\omega_n} \sin(\omega t + \theta) \tag{112.14}$$

According to (112.13), the step response of second- order system ($\xi \in (0,1)$) is composed by two parts. One is steady state component; the other one is transient component with oscillation frequency of ω. The decay speed of transient component is dominated by the envelope $1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\xi\omega_n}$. According to the characteristic of the step response of second-order system, the overshoot can be expressed as:

$$t_p = \pi/\omega \tag{112.15}$$

$$\sigma\% = \frac{g(t_p) - g(\infty)}{g(\infty)} \times 100\% = e^{-\pi/\sqrt{1-\xi^2}} \times 100\% \tag{112.16}$$

where $g(\infty)$ is equal to the value of the step function. From Eq. 112.11:

$$g(\infty) = \left(\frac{\Delta Pm_{10}}{M_1} - \frac{\Delta Pm_{20}}{M_2}\right)\left(\frac{M_1 M_2}{M_1 + M_2}\right) \tag{112.17}$$

$$g(t_p) = \max \Delta P_{\text{tie}} = \left(\frac{\Delta Pm_{10}}{M_1} - \frac{\Delta Pm_{20}}{M_2}\right)\left(\frac{M_1 M_2}{M_1 + M_2}\right)\left(1 + e^{-\pi/\sqrt{1-\xi^2}}\right) \tag{112.18}$$

The transfer ratio is:

$$\eta\% = \frac{\max \Delta P_{\text{tie}}}{\Delta P} \times 100\% \tag{112.19}$$

Equation 112.17 can be utilized to calculate the peak value of active power oscillation on tie line. The Eqs. 112.15–112.19 indicate that:

The active power fluctuation is composed of steady state component and transient component. If system has strong damping, the transient component will attenuate soon.

The peak value of the active power fluctuation is associated with grid inertia, electromagnetic mutation value, and inter-area damping ratio.

The inter-area damping ratio include the generator mechanical damping, electromagnetic damping, load frequency characteristic etc.

112.4 Case Study

112.4.1 Step Response and Fluctuation Curve

According to the analysis in Sect. 112.3, the process of the UHV tie line oscillation is similar with the process of second-order system step response. As shown in Fig. 112.3, solid line is the result of the simulation of the tie line active power after large disturbances, dotted line is the step response of a second-order system which is structured by system inter-area damping and tie line oscillation frequency. The peak value and the Decay processes are almost same.

112.4.2 Fault in Different Grid

In this part, we describe two disturbances in china grids (one in CCG, one in NCG). Then method proposed in Sect. 112.3 will be used to calculate the active power fluctuation peak value and transfer ratio.

UHV AC line disturbance test were done on December 29, 2008. The test cut 12# generator (700 MW) of three gorges power plant. Before disturbance, the active power on UHV line is 1.507 GW. The peak value of the active power fluctuation following the disturbance is 2.1415 GW. The variation of active power is 634.5 MW. The transfer ratio is 90.64%. The inter-area mode frequency is 0.15 Hz, damping ratio is 0.11. The inertia of grids is $M_1 : M_2 = 1 : 1.184$. The variation of active power is 647.5 MW and transfer ratio is 92.5% when calculated with the approach proposed in Sect. 112.3, which fit well with to the field data.

Generator with an output of 950 MW in ZouXian plant was tripped on May 5, 2009. Before disturbance, the active power on UHV line was 760 MW. During the disturbance, the peak value of the active power fluctuation was 60 MW. The variation in active power is 700 MW. The transfer ratio was 73.68%. The inter-area mode damping ratio was 0.11. The inertia constant of grids is $M_1 : M_2 = 1 : 1.366$. The active power variable quantity is 685.1 MW and transfer ratio was 72.11% when calculated with the approach proposed in Sect. 112.3. The result consists of the filed data, too.

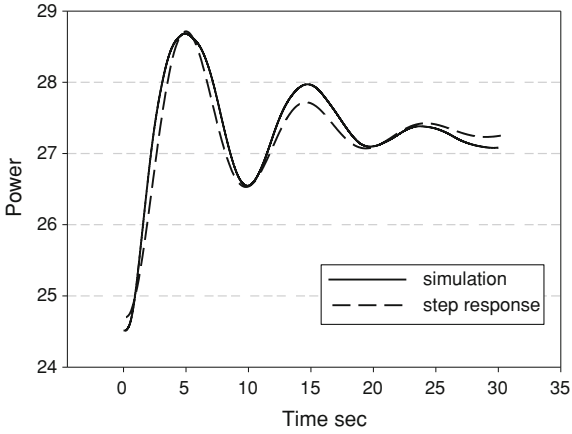


Fig. 112.3 Simulated active power of the UHV transmission line and step response of second-order system

Fig. 112.4 Faults location

112.4.3 Fault Location and Peak Value

The electromagnetic mutation of a large power grid with fault present is difficult to calculate. The electromagnetic mutation could be approximated with fault power shortage if the fault location is far away from the UHV AC line. However, this replacement will bring errors. From Eq. 112.18, the difference of electromagnetic mutation in each grid affects the peak value of active power fluctuation.

In most situations, the electrical distance is proportional to the geographical distance. Thus, the generator is closer to the fault location the active power mutation is smaller. We use simulation to analyze the relationship between fault location (Fig. 112.4) and peak value. The result is shown as Tables 112.3 and 112.4.

Derived from the theory of this paper, the distance between fault location and the UHV line placement is proportional to the transfer ratio. When the faults occur

Table 112.3 Fault in Central China Power Grid

Fault location	Cutting power (MW)	Peak value (MW)	Transfer ratio
ErTan#01	550	2894.76	0.806
HeBi05	600	2896.91	0.742
JinTang#61	600	2868.33	0.695
HanChuan	600	2792.58	0.568

Table 112.4 Fault in North China Power Grid

Fault location	Cutting power (MW)	Peak value (MW)	Transfer ratio
HuangDao5	600	1956.14	0.826
CangDong1G	600	2072.79	0.6311
ChenReG8	300	2263.57	0.6262
ZhangShan3G	600	2061.97	0.649

in CCG, the results of transfer ratio were shown in Table 112.3. Table 112.3 is sorted by distance between the fault location and UHV placement, and it can be observed that the transfer ratio is listed by descending order. Table 112.4 shows when fault occurs in NCG, the transfer ratio has the same law.

112.5 Conclusions

This paper presents a method to calculate the peak value of active power following large disturbance when system has strong inter-area damping. The following conclusions can be drawn from the analysis presented in this paper:

The inter-area oscillation is dominant by inter-area oscillation mode when system has strong damping. The peak value of the active power fluctuation is closely related to grid inertia, electromagnetic mutation value, and inter-area damping ratio. The method present in this paper can be used calculate the peak value of active power following large disturbance, like generator tripping, dc blocking etc., which can be describe by step function. The transfer ratio is associated with fault location. The transfer ratio is proportional to the distance between fault location and the UHV line placement.

References

1. Guan Z, Wang G (2005) The projects and related key techniques of ultra high voltage transmission in China. China South Power Grid Technol Res 1(6):12–18
2. Huang D, Ruan J, Yu S (2008) Overview of the recent developments of ultra high voltage AC transmission in China. IEEE Power Energy Soc Gen Meet 22(8):1–9
3. CIGRE (1983) WG 31.04 Report. Electric power transmission at voltages of 1000 kV and above: plans for future AC and DC transmission, data on technical and economic feasibility

and on general design, information on testing facilities and the research in progress. Electra 91:83–133
4. CIGRE (1988) WG 38.04 Report. Electric power transmission at voltages of 1000 kV AC or ±600 kV DC and above: network problems and solutions peculiar to UHV AC transmission. Electra 122:41–75
5. CIGRE (1994) WG 38.04 Report: ultra high voltage technology. CIGRE, Paris
6. Vittal V, Bhatia N, Fouad AA (1991) Analysis of the inter-area mode phenomenon in power system following large disturbances. IEEE Trans Power Syst 6(4):1515–1521
7. Kundur P, Wang L (2002) Small signal stability analysis: experience, achievements, and challenges. Int Conf Power Syst Technol Proc 1:6–12
8. CIGRE (1996) Task force 38.01.07 on power system oscillations. Analysis and control of power system oscillations. CIGRE, Paris Tech. Rep. no 111
9. Barbier C, Ferrari E, Johansson KE (1979) Questionnaire on electromechanical oscillation damping in power systems report on the answers, Electra 07(64):59–82, 84–90
10. Klein M, Rogers GJ, Kundur P (1991) A fundamental study of inter-area oscillations in power systems. IEEE Trans Power Syst 6:914–992
11. Liu GP, Huang Y et al (2004) Analysis of inter-area oscillations in the south China interconnected power system. Electr Power Syst Res 70:38–45
12. Bachmann U (1999) Analysis of inter area oscillations in the European electric power system in synchronous parallel operation with the Central-European networks. Proc Power Tech Conf 13(05):49–51
13. Messina AR (1995) Experience with the analysis of small signal stability in longitudinal systems: a case study with the Mexican interconnected system. J Electr Power Energy Syst 17:291–299

Part IX
Power System

Chapter 113
A New Substitution Waveform of Very Fast Transient Overvoltage in GIS

Huaying Dong, Guishu Liang, Xixiao Liu, Haifeng Sun, Xintong Yang and Xin Liu

Abstract So far there is no standard waveform of very fast transient overvoltage. Considering the amplitude, frequency and frequency weighted from a great deal of data and according to IEC 61321-1, an accurate substitution waveform of very fast transient overvoltage is presented by means of probability statistical and time–frequency analysis technique. Simulation validation is carried out while the simulated waveforms and substitution waveforms excite the transformer winding and potential transformer, respectively. Test results indicate that the proposed method is highly reliable.

Keywords Substitution waveform · Very fast transient overvoltage · Gas-insulated metal-enclosed switchgear

113.1 Introduction

Very fast transient overvoltage (VFTO) induced by switch operations in GIS system is risky to the internal insulation of the equipment and the equipments connected to GIS for the high amplitude, extra high frequency and steep front of VFTO.

H. Dong (✉) · G. Liang · X. Liu · H. Sun · X. Yang · X. Liu
Key Laboratory of Power Equipments Defense of Hebei Province,
North China Electric Power University, P.O.Box 19, Baoding 071003, China
e-mail: liuxinhust@163.com

G. Liang
e-mail: gshliang@263.net

X. Liu
e-mail: doolxx@163.com

X. Liu
e-mail: liuxinhust@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_113,

For instance, the most likely reason of the fault of secondary side protection device in one Power Company Limited (Co. Ltd.) is due to the VFTO spreading to the secondary system caused by the operation of 110 kV disconnecting switch. And VFTO causes high inter-turn voltages and internal resonance in the transformer windings, so transformer is the main casualty of VFTO. Now many manufacturers of the high voltage electrical equipment, especially power transformer, are required to provide insulation analysis results under VFTO condition. However, the waveform of VFTO has not been standardized, either in engineering circles or in academic circles. Since there is no standard waveform of VFTO so far, the designers of high voltage electrical equipment manufacturers need to use simulation method to obtain VFTO waveform and then calculate their distributions and verify the margin of the insulation design. An accurate substitution waveform of VFTO is of great importance for equipment insulation during design stage [1–5].

In this paper, the basis is the simulation VFTO waveform of the entrance of transformer in 500, 750, 1000 kV gas-insulated substation and the entrance of potential transformer in 125 kV gas-insulated substation. The VFTO substitution wave in the entrance of transformer and voltage transformer is given. In Sect. 113.2, the models of 500, 750 and 1000 kV GIS are built under various voltage levels according to the various components of the model, and circuit simulation is carried out with EMTP-ATP. The substitution waveform of VFTO is presented in Sect. 113.3. In Sect. 113.2, Simulation validation are carried out while the simulated waveforms and substitution waveforms effect the transformer winding and potential transformer, respectively.

113.2 Calculation of Very Fast Transient Overvoltage

VFTO's waveform is different in each location of gas-insulated substation, VFTO generated from disconnecting switches and circuit breakers in operation spread to both sides from the breakdown point, VFTO basic oscillation frequency is between 0.1 and 5 MHz. The very fast transient overvoltage's rise time is very short, some of its oscillation frequency is tens of MHz, the maximum may reach more than 100 MHz.

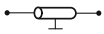
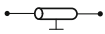
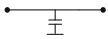
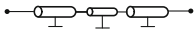
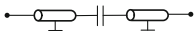
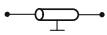
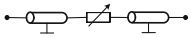
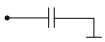
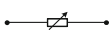
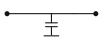
As the transmission characteristics of VFTO, the model of the distributed parameter equipments in GIS are mainly built with transmission lines method, such as the conductor within GIS, the cable inner conductor, metal shell, cable sheathing, etc. Modeling of other main elements is established with the lumped parameter [6–8]. The equivalent circuit of main elements in GIS are given in Table 113.1.

Wave impedance of cable inner conductor can be calculated by

$$Z_1 = 60 \ln(R_1/R_2) \tag{113.1}$$

where R_1 is the inner diameter of shielding conductor, R_2 is the radius of inner conductor.

Table 113.1 The equivalent circuit of main elements in GIS

Element	Equivalent circuit
GIL bus	
GIS bus	
Pillar insulator	
Earthed switch	
Closed circuit breaker	
Opened circuit breaker	
Closed disconnecting switch	
Arc period of disconnecting switch	
Transformer	
Potential and current transformer	
Lightning arrester	
Reactor	
Overhead line (terminal end)	

The wave impedance of shell (cable sheath) can be calculated by

$$Z_2 = 60\ln(2h/R_3) \tag{113.2}$$

where R_3 is the outer radius of shielded conductor, h is the height of shielding conductor.

The arc resistance of disconnecting switch can be approximately calculated by

$$R(t) = R_s + R_0 e^{-t/T} \tag{113.3}$$

where $R_s = 0.5\,\Omega$, $R_0 = 10^{12}\,\Omega$, $T = 1\text{ns}$.

The equivalent calculation models of GIS with EMTP, according to the above method, are established. And the EMTP model, which respectively belong to 750, 500 and 1000 kV GIS in different cities, were established, simulation waveforms

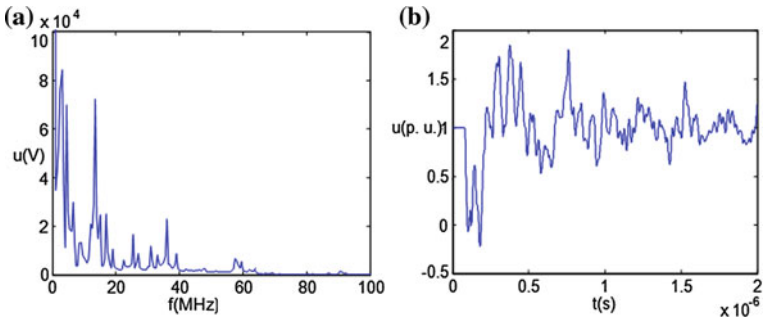


Fig. 113.1 The frequency-domain and time-domain characteristics of VFTO waveform at the entrance of transformer

of VFTO in the corresponding different operation modes are obtained. For instance, the time-domain and frequency-domain characteristics of VFTO waveform at the entrance of transformer in 500 kV GIS are shown in Fig. 113.1.

113.3 Research on Substitution Waveform of VFTO

113.3.1 The Substitution Waveform Method

Due to the short rise time and wide frequency range of VFTO, it is difficult to find out its exact substitution waveform. Reference [4, 5] put forward VFTO substitution waveform and the influence factors of VFTO waveform, but it is not desirable from the precision and expansibility.

There are five steps to obtain the proposed VFTO substitution waveform:

Step 1. Peak value

According to the process of lightning amplitude, this paper puts forward that using the corresponding amplitudes of 99% amplitude probability density [9] to obtain VFTO alternative waveform amplitude.

On the basis of "Overvoltage Protection and Insulation Coordination for AC Electrical Installations", which is the standards DL/T 620-1997 of the power industry, the probability of lightning current value carry more weight than I in China's general areas can be calculated with the following Empirical formula

$$\lg P = -I/88 \tag{113.4}$$

where I stands for lightning current amplitude, unit is kA, and P stands for the probability of lightning current value carry more weight than I.

Step 2. Frequency

According to the national standards, there are three main frequency included VFTO [10]: (1) 0.1–5 MHz, low frequency oscillation, (2) 30 MHz or so, high

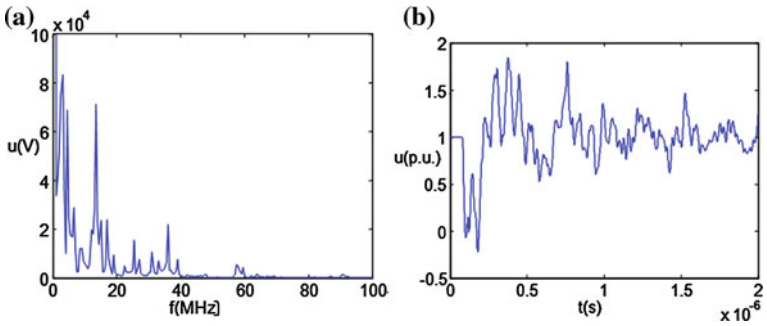


Fig. 113.2 The most serious VFTO waveform of main transformer in 500 kV GIS

frequency oscillation, (3) Up to 100 MHz above, ultra-high frequency oscillation. Find out the most serious waveform from the simulation waveforms, and find out the frequency corresponding to Maximum peak point in the vicinity of three main frequencies, and compute them as three main wave frequency of VFTO substitution waveform.

Step 3. Frequency weighting

Without considering the components of the step, the weighting factor is the ratio of the peak value of these three frequency peak values and VFTO peak value.

Step 4. The function included these three main frequency components [8]

$$A(t) = a\sin 2\pi(f_1 t) + b\sin 2\pi(f_2 t) + c\sin 2\pi(f_3 t) \tag{113.5}$$

Step 5. The function of VFTO substitution waveform

$$U(t) = Ae^{-10^6 t} \times [a\sin 2\pi(f_1 t) + b\sin 2\pi(f_2 t) + c\sin 2\pi(f_3 t)] \tag{113.6}$$

where A is peak correction factor, a, b, c are weighting factor for the frequency, f_1, f_2, f_3 are three main frequency included VFTO, $e^{-10^6 t}$ is attenuation coefficient.

113.3.2 One Example

The most serious VFTO waveform of main transformer in 500 kV GIS is shown in Fig. 113.2. The three main frequency values are $f_1 = 3$ MHz, $f_2 = 13.5$ MHz, $f_3 = 36$ MHz and corresponding peak values are 83400, 71290, 21820 V. Peak value of VFTO is 832 kV. Weighting factor values for the frequency are 8340, 7129, 2182. The function of VFTO substitution waveform is shown in (113.7)

$$U(t) = 0.025e^{-10^6 t}\times \left[8340\sin 2\pi\left(3\times 10^6 t\right) + 7129\sin 2\pi\left(13.5\times 10^6 t\right) + 2182\sin 2\pi\left(36\times 10^6 t\right)\right] \tag{113.7}$$

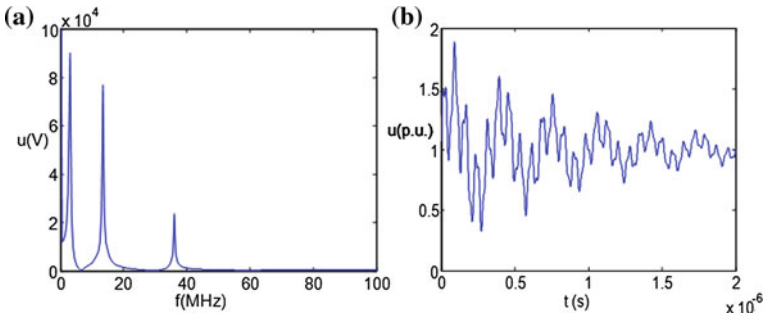


Fig. 113.3 The VFTO substitution waveform of main transformer in 500 kV GIS

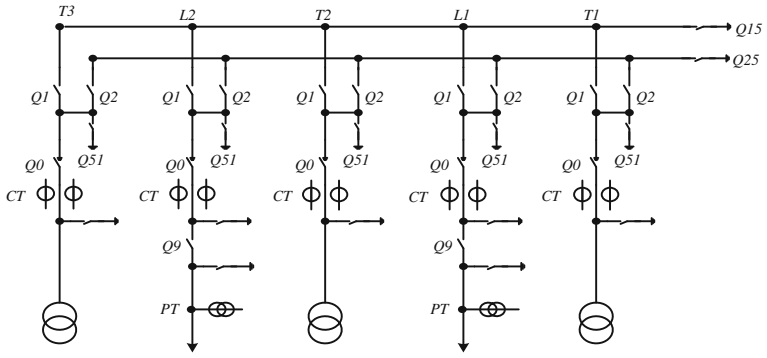


Fig. 113.4 The main electrical wiring diagram of a 125 kV GIS

Because VFTO alternative waveform does not contain the initial state, so the initial state of transformer should be added, as shown in Fig. 113.3.

113.4 Verification Through the Comparison Between Substitution Waveform and Simulation Waveform

113.4.1 The Contrast of Impact on the Transformer Windings

Simulation validation is carried out by using the simulated waveforms and the substitution waveforms exciting the transformers with different parameters, respectively. Pancake-to-ground voltage, maximum voltage between the pancakes, turn-to-ground voltage, and the maximum inter-turn voltage are calculated, in which the relative errors of the results according to the two waveforms are all less than 5%, and the difference of the errors are also small. The errors are in the range of tolerance allowance which indicates the high reliability of the proposed method.

Table 113.2 The contrast of impact on the potential transformer secondary side

Excitation	Rise time (ns)	Amplitude (p.u.)	Secondary maximum voltage (V)	Second peak–peak value (V)
Substitution waveform	9	2.55	1060	1544
VFTO	8	2.62	989	1460
VFTO	16	1.75	761	1095
VFTO	19	1.23	761	889
VFTO	25	2.12	897	1199
VFTO	30	1.17	760	884

113.4.2 The Contrast of Impact on the Potential Transformer Secondary Side

The substation is a 125 kV GIS, and outgoing line is cable, as shown in Fig. 113.4.

Establishing the electromagnetic transient model with the above GIS's electrical wiring diagram in the EMTP software, and using the method presented in Sect. 113.3, the function of the VFTO substitution waveform can be obtained as (113.8).

$$U(t) = 22.9e^{-10^6 t} \times \left[3787 \sin 2\pi\left(4 \times 10^6 t\right) + 3147 \sin 2\pi\left(9 \times 10^6 t\right) + 543 \sin 2\pi\left(41 \times 10^6 t\right)\right] \tag{113.8}$$

Using the simulation waveforms and substitution waveforms as excitation, the interference voltages at the secondary side of the 10 kV potential transformers are calculated, which are listed in Table 113.2.

From the data listed in the table it can be seen that, it can basically be covered to the interference voltage secondary side with different amplitude and rise time VFTO for interference voltage the secondary side of VFTO substitution waveform. In view of maximum voltage, the result is satisfied.

113.5 Conclusion

This paper establishes the models of GIS in 125, 500, 750, 1000 kV voltage level, and provides a great deal of simulation data. An accurate substitution waveform of very fast transient overvoltage is presented. Through the contrast of impact on the transformer windings and impact on the secondary side of the potential transformer according to the simulated waveforms and the substitution waveforms, satisfactory results are obtained. And it is of great importance for transformer

designers so that they can predict the VFTO distributions in the windings during insulation design stage.

Acknowledgments This research was supported in part by National Natural Science Foundation of China under Grant No. 50977031 and the Fundamental Research Funds for the Central Universities under Grant No. 10MG07 and 11MG36, respectively.

References

1. Haznadar Z, Carsimatovic S, Mahmutcehajic R (1992) More accurate modeling of gas insulated substation components in digital simulation of very fast electromagnetic transients. IEEE Trans Power Deliv 7(1):434–440
2. Mohana RM, Joy TM, Singh BP (2007) Transients induced on control cables and secondary circuit of instrument transformers in a GIS during switching operations. IEEE Trans Power Deliv 22(3):1505–1513
3. Smeets R (2001) Disconnector switching in GIS three-phase testing and phenomena. IEEE Trans Power Deliv 15(1):122–127
4. Yu Y, Zanji W, Chong S (2009) Determination of VFTO waveform in GIS according to simplified structure of GIS bus and circuit parameters. Power Syst Conf Expo 3:1–9
5. Chong S, Yu Y, Zanji W (2010) Modeling of GIS switching arc and its effect on VFTO waveforms. Power Syst Technol 34(7):200–205
6. Ding-xie G, Mu-hong X, Min D, Pei-hong Z Z (2007) Study on VFTO of 1000 kV GIS substation. High Volt Eng 33(11):27–32
7. Chi T (2009) Study on very fast transient overvoltage and its impacts in EHV and UHV GIS. Shenyang University of Technology, Shenyang
8. Quan G (2006) Improved time frequency distribution and its application in radar signal process. Southwest Jiaotong University, Chengdu
9. Wasserman L (2006) All of nonparametric statistics. Springer, New York
10. GB/T18134.1-2000 (2000) High-voltage testing with very fast impulses part 1: measuring systems for very fast front over voltages generated in das-insulated substations. State Administration of Machinery Industry, Beijing

Chapter 114
Ray Penetration Window Technology of Mining Explosion-Proof Instrument for Measuring Ash Content of Coal

Wen-qing Wang

Abstract In order to improve the extraction rate and reduce the gangue content in coal production, real-time data of the ash content of raw coal is needed to guide the miners in the working face to control the top-coal caving. The ray penetration window technology breaks through the rules of "the larger the mass thickness of material, the higher the amount of attenuation" and effectively solves the problem of large attenuation rate when low-energy γ-ray penetrates the explosion-proof enclosure. The application of this technology in the instrument ensures the measurement accuracy, achieves on-site monitoring of the ash content of raw coal and plays a positive role in improving the extraction rate of the mining area, and obtained remarkable economic and social benefits.

Keywords Extraction rate · Mining explosion-proof · Instrument for measuring ash content of coal · Ray penetration window · Low-energy γ-ray

114.1 Overview

Coal accounts for nearly 70% of the primary energy in China. According to the survey data about the extraction rate of coal mines all over the country, the average extraction rate is 64% in all types of coal mining area, but the average extraction

W. Wang (✉)
Beijing Polytechnic College, No.368, Shimen Road, Shijingshan District, Beijing, China
e-mail: wwq@bgy.org.cn

W. Wang
Beijing Coal Mining Electric Equipment and Technique Development Company, No.368, Shimen Road, Shijingshan District, Beijing 100042, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_114,

rate in mines is only 46% [1]. Improving the extraction rate by adopting scientific methods could not only decrease the cost of coal mines and improve the economic efficiency of coal mining enterprises [2], but also to reduce the footage under the condition of the same production and extend the lifespan of mines.

The ash content of raw coal reflects the amount of gangue in raw coal [3]. At present, there are two main methods of controlling the ash content of raw coal, one is the method of "sampling, testing and eye survey", another is online monitoring by installing common ground instrument on the conveyor of wellhead and coal preparation plant. However, there are drawbacks in the first method, such as long testing time, bad real time, big human impacts and wide fluctuation range of coal quality. The second method has been widely used, but there is also a lag in the control of top-coal caving. The actual application effect is not ideal.

A mining explosion-proof instrument is needed to measure ash content of coal. The instrument is used underground for online and real-time monitoring to meet the demand of controlling the ash content and improving the extraction rate.

114.2 Measuring Principle of Instrument for Measuring Ash Content of Coal

Coal is composed of combustible organic matter and incombustible mineral matter. The composed elements of organic matter are mainly C, H, N, O, which are called "low-Z elements" for short. Those of mineral matter are Si, Al, Fe, Ca, Mg, S, which are called "high-Z elements" for short. The ash content of coal is the percentage of oxide of the high-Z elements after complete and enough burning of coal in certain temperature [4].

The instrument for measuring ash content of coal uses ^{241}Am as the low-energy γ source, ^{137}Cs γ as the middle-energy γ source. Let two kinds of γ to penetrate the same coal seam in the same collimated beam, the intensities of both are attenuated. For narrow-beam parallel rays, they comply with the law of exponential attenuation. For ^{241}Am low-energy γ-ray, the formula is:

$$I = I_0 e^{-\mu_L \rho d} \tag{114.1}$$

where, I and I_0 are ,respectively, the count of ^{241}Am γ measured by the detector in certain time interval when there are coal and no coal, which are proportional to their flux density. μ_L is the coefficient of mass attenuation of coal for ^{241}Am γ-ray, ρd is also known as the mass thickness, ρ is the bulk density of coal, d is the thickness of coal.

Similarly, for ^{137}Cs middle-energy narrow-beam γ-ray, the formula is

$$J = J_0 e^{-\mu_m \rho d} \tag{114.2}$$

where, J and J_0 respective is the count of ^{137}Cs γ measured by detector in certain time interval when there are coal and no coal which is proportional to its flux density. μ_m is the coefficient of mass attenuation of coal for ^{137}Cs γ-ray. The mass thickness ρd in formula (114.1) and (114.2) are same because the two kinds of γ penetrate the same coal seam.

Assuming that the percentage share of quality of the high-Z elements in the measured coal is C_z, then that of low-Z elements is $(1-C_z)$, according to calculation rules of the coefficient of mass attenuation of composite material [5], the coefficient of mass attenuation of coal for ^{241}Am γ-ray:

$$\mu_L = C_z\mu_z + (1 - C_z)\mu_c = \mu_c + (\mu_z - \mu_c)C_z \tag{114.3}$$

where, μ_z is the coefficient of mass attenuation of high-Z elements for the ^{241}Am low-energy γ; μ_c is the coefficient of mass attenuation of low-Z elements for the ^{241}Am low-energy γ.

Obviously, the coefficient of mass attenuation μ_L of coal for low-energy would be larger if the amount of high-Z elements increases. Take the logarithm on both sides of (114.1) and (114.2), put the expression of μ_L in (114.3) into them and sort, then we could obtain:

$$C_Z = \frac{\mu_M}{\mu_Z - \mu_c} \cdot \frac{\ln I_0 - \ln I}{\ln J_0 - \ln J} - \frac{\mu_C}{\mu_Z - \mu_C} \tag{114.4}$$

The ash content of measured coal is recorded as A_d, and $A_d \approx 2C_Z$, so

$$A_d = \frac{2\mu_M}{\mu_Z - \mu_c} \cdot \frac{\ln I_0 - \ln I}{\ln J_0 - \ln J} - \frac{2\mu_C}{\mu_Z - \mu_C} \tag{114.5}$$

Formula (114.5) is the basic formula of the dual-energy penetration method to measure ash content of coal. For the Specific coal: μ_M, μ_z, μ_c are constant; for a given instrument for measuring ash content of coal, $1nI_0$ and $1nJ_0$ could also be determined constant.

114.3 Ray Penetration Window Technology

To accurately measure the ash content of raw coal, it is necessary to ensure that the parameters I and J in the formula (114.5) satisfy certain count values [6].

If we put instrument for measuring ash content on working surface, it is necessary that the instrument should be explosion-proof. In other words, γ-ray detector should be packed in an explosion-proof enclosure. The γ-ray from the output first penetrates the coal seam, then be absorbed partly by the explosion-proof enclosure, last be detected by the detector. A test environment for testing γ attenuation was set up, as shown in Fig. 114.1.

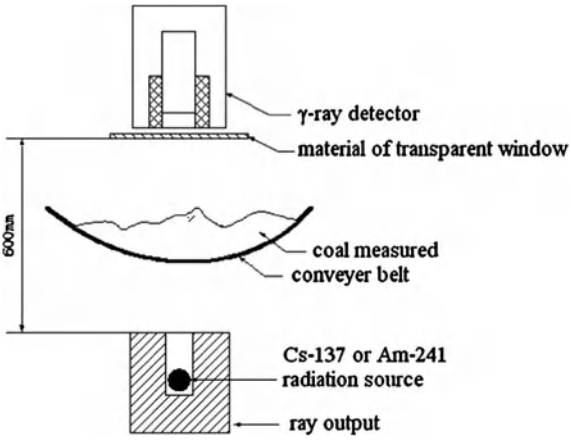


Fig. 114.1 The diagram of test environment for γ attenuation

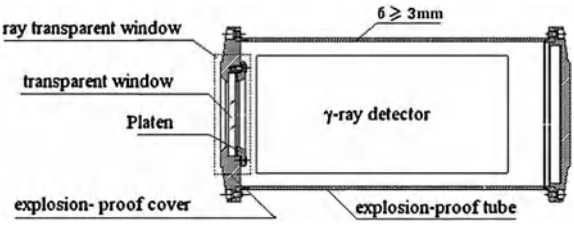


Fig. 114.2 The principle of ray penetration window technology

The aim of ray penetration window technology [7] is to reduce the attenuation rate of the enclosure for low-energy γ. We proposed a scheme that install "ray window" on the explosion-proof cover of γ-ray detector. The "ray transparent window" is composed of platen and transparent window. The structure is shown in Fig. 114.2.

The choice of transparent window material needs to meet two conditions: ① a small amount of attenuation for low-energy γ-ray; ② meet the requirements of explosion-proof.

For certain kind of material, the larger the mass thickness of material, the higher the amount of attenuation. For different materials, the attenuation not only has relation to the mass thickness, but also to the coefficient μ, the mass attenuation of material for γ-ray.

The acquisition method of the coefficient μ is different when the composition of material is different. For metal, the coefficient could be determined directly from the data table. For metal alloys with known component contents, the coefficient could be obtained by the content of the component materials and their coefficients. For nonmetal, for example the coal, the coefficient could be obtained only by the method of test because the component content could not be accurately obtained.

To find an appropriate material for transparent window which satisfies the requirement of explosion-proof, three types of material were chosen as the test samples. These materials were the nonmetal (A material), metal alloy (B material)

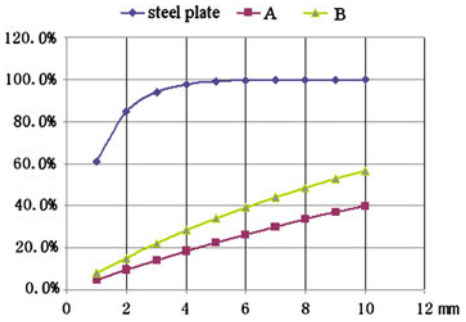


Fig. 114.3 The attenuation rates of different kinds of material of transparent window for ^{241}Am low-energy γ-ray in different thicknesses

and pure metal (A3 steel plate). Tests were conducted to test the amount of attenuation of the three above kinds of material for low-energy γ-ray. The attenuation rates in different kinds of material and different thicknesses are shown in Fig. 114.3.

The property of explosion-proof material of transparent window not only needs to meet the requirements of shock test in GB 3836.1, but also to meet that of the explosion in GB 3836.2. The thickness of transparent window material could also be different according to the different size of explosion-proof transparent window. The minimum thickness of A material is 10 mm in terms of the hole diameter which meet the requirements of explosion-proof, that of B material is 6 mm.

The mass thickness of transparent window made from 10 mm thick A material is 2.7 g cm^{-2}, that of 6 mm thick B material is 1.68 g cm^{-2}, that of 3 mm thick A3 steel plate is 2.34 g cm^{-2}. It could be seen from Fig. 114.3 that the attenuation rate of transparent window sample made from A material for low-energy γ is 40%, that of B material sample is 41.3%, that of A3 steel plate sample is 94.1%.

We could draw the following conclusions by analyzing the above test data: ① B material and A3 steel material are both metal, the amount of attenuation rate of transparent window sample for low-energy γ-ray satisfy the law of "the greater the mass thickness of sample, the higher the attenuation rate of sample for γ-ray"; ② A material is nonmetal oxide, A3 sheet plate is metal. The attenuation rate of these two kinds of material for low-energy γ-ray does not satisfy the law mentioned in ①; ③ The amount of attenuation of B material for low-energy γ-ray is close to that of A material, but the B material is not recommend to be used as transparent window material because that the national standard GB3836.2 has many limitations to it; ④ A material is suitable for being used as transparent window materials because of the smallest attenuation rate and a high penetration rate.

The transparent window made from A material not only meets the requirement of explosion-proof, but also meets that of received count rate of the low-energy γ-ray. It guarantees the measuring accuracy of the mineral explosion-proof instrument for measuring ash content of coal. The instrument is used underground to monitor the ash content online and guide the production of caving top coal in the mining area.

114.4 Conclusion

The mineral explosion-proof γ-ray detector which takes advantage of ray penetration window technology effectively solves the problem of large attenuation rate when low-energy γ-ray penetrates the explosion-proof enclosure, ensures the measurement accuracy, achieves the local monitoring of the ash content of raw coal and plays a positive role in improving the extraction rate of the mining area. The successful application of the mineral explosion-proof instrument on the working surface effectively monitors the emission amount of the waste rock, reduces the power consumption of transport and the collapse of gob area and protects the mining environment.

References

1. Shi Y (2007) Save energy start from improving coal mining rate. Land Resour 7:15–21
2. Li Y, Xie F, Jiang J (2008) Economic analysis of the best extraction rate of coal resources. Mining Res Dev 3:78–81
3. Bao X, Zheng J, Li H (2005) The study of control of caving top coal. Coal Sci Technol Shandong 3:63–64
4. Zhang Z, Zhuo Y, Lin Q (1999) Γ radiation instrument for measuring ash content of coal. Atomic Energy Press, Beijing
5. Ma Y, Zhou W, Wen F et al (1989) The study of the measurement of coal ash in the method of the transparent of low energy γ-ray. Nucl Electron Detect Technol 9:285
6. Zhang Z, Lei Z, Zhuo Y (1993) On-line measurements of coal ash. Tsinghua Univ (Nat Sci) 6:48
7. Wang W, Ren A, Tian B et al (2009) Explosion-proof γ-ray detectors and mineral instrument for measuring ash content of coal: China ZL 2009 1 0083984.3[P] 2009 05 13

Chapter 115
Cable Burning Analysis Caused by Power Harmonics

Yongle Ai, Yinghui Zhang and Yunjian Wang

Abstract This paper focuses on the two aspects that are the analysis of the cable leakage capacitance and harmonic current amplified from the point where there is harmonic current heavily existing in the power cable transmission system. Firstly, the main parameters in the cable harmonic impedance model are estimated. Secondly, the equivalent circuit based on the cable power supplying is established as well as parameters included in the equivalent circuit are estimated. Finally, the harmonic leakage current is simulated by Matlab based on harmonic current and leakage capacitance. Results show that harmonic current through the cable leakage capacitor is amplified with the coincidence cases. It will cause the cable insulation medium theat and burning due to high harmonic current. It would provide a novel way to find the reason of the cable burning by suing analytical methods and results described in this paper.

Keywords Power cable · Harmonic current · Harmonic impedance · Cable burning

Y. Ai (✉) · Y. Zhang · Y. Wang
Henan Polytechnic University, Jiaozuo 454003, Henan, China
e-mail: aiyongle@hpu.edu.cn

Y. Zhang
e-mail: zyh8407@126.com

Y. Wang
e-mail: yunjian_wang@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_115,

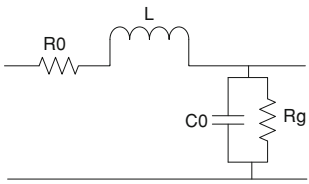


Fig. 115.1 Equivalent circuit for power cable unit

115.1 Introduction

With the rapid development of economy, various fields are growing demand for electricity. Cable is widely used in the city, the construction industry, residential areas, substations and power plants, industrial and mining enterprises and through rivers, sea and other underwater as transmission line. It has become a main mode to delive electrical energy. However, when the cable is on the state of fault, it is not only interrupting the normal power supply but may also lead to a series of vicious chain reaction, such as burning auxiliary electrical equipment and fire accidents . Cable fault types often have a low resistance grounding fault, totally cutting the point of failure, and not completely breaking the point of failure [1]. Based on cable harmonic in low resistance cases, it is discussed in this paper that harmonic currents leakage cause cable temperature to rise and burn, and in such cases provide an effective analysis approach of cable burning failure.

115.2 Circuit Model of Cable Harmonic Impedance

115.2.1 Equivalent Circuit of Cables Parallel Theory

Based on features in power supply and cables parallel theory, the cable unit equivalent circuit is shown in Fig. 115.1.

115.2.2 Calculation of Equivalent Circuit Parameters of Power Cable

For power cables, the value of inductance L is about 0.1–0.7 mH/km. It is generally to be calculate as follows:

$$L = 0.05 + 0.2 \times \ln \frac{ks}{r} \text{ (mH/km)}, \tag{115.1}$$

where k—Coefficient for the parallel structure of the cable, $k = 1.26$; s—Distance of the center conductor (mm); r—Conductor radius.

Table 115.1 Commonly used dielectric constant of insulating materials

Material name	ε	$t_g\ \delta$
Adhesive-impregnated paper insulation	4.0	0.01
Oil-filled cable insulation	3.3–3.7	0.004–0.0045
Rubber insulation (ethylene-propylene, butyl)	3.0–4.5	0.04–0.05
PVC insulation	8.0	0.1
Polyethylene insulation	2.3	0.001
XLPE insulation	2.5	0.008

The calculated value of inductance L is about 0.1–0.7 mH/km. Since the equivalent value of L is relatively small, the actual analysis can be regarded as a short circuit. Resistance wire conductor of power cable can be calculated using the available formula (115.2),

$$R = \rho \frac{L_c}{S} \tag{115.2}$$

where, ρ—resistivity, such as (ambient temperature is 20°C); L_c—cable length (m); S—cable conductor cross-sectional area (mm^2).

The equivalent capacitance C_0 of cable core wire relative to the metal shield can use formula (115.3) to calculate,

$$C_0 = \frac{55.7\varepsilon}{\ln \frac{d_2}{d_1}} (\mu F/km), \tag{115.3}$$

where ε—dielectric permittivity of insulating medium; d_1—conductor wire diameter; d_2—cable insulation diameter.

Commonly used dielectric constant of insulating material (relative) [2] is shown in Table 115.1.

The calculated value C_0 is about 0.1–0.7 μF/km. Since the power cable insulation resistance R_g to ground is $10^9\ \Omega$, it can be equivalent to open. The main harmonic components existing in the power system are 2–20 orders [3], which can cause the harm of power cable. The cable equivalent capacitance of cable core wire relative to the metal shield should be carefully considered as the main parameters. Namely, power cable harmonic impedance model is the equivalent capacitance C_0 of cable core relative to the metal shield. The equivalent circuit is shown in Fig. 115.2.

115.3 Harmonic Amplification Principles

For power supply system by power cable, especially coal mine power cable, schematic diagram of the power supply system is shown in Fig. 115.3. A constant current source is to replace the harmonic source because external impedance

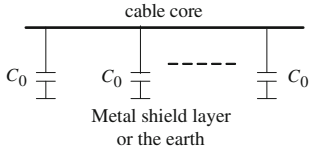


Fig. 115.2 Harmonic impedance model power cable

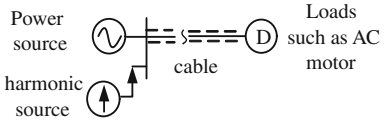


Fig. 115.3 Cable power supply system

changes and the harmonic currents have not changed much. Equivalent circuit of the power supply system is shown in Fig. 115.4. R_s and L_s are the equivalent resistance and inductance of the power part. R_l and L_l are the equivalent resistance and inductance of the cable supply load. I_n is the n-order harmonic current source.

The equivalent circuit of the cable supply power system is shown in Fig. 115.4. The equivalent circuit [4] can continue to be simplified and shown in Fig. 115.5. In Fig. 115.5, the inductance L' and resistance R' in the series branch are the equivalent circuit of the branch $R_s - L_s$ of power supply circuit and the load branch $R_l - L_l$ constituted parallel circuit branches in Fig. 115.4. Thus, by calculation, equivalent value of the $R^{'}$ and $L^{'}$ can be shown as follows

$$R' = \frac{(R_s R_l - \omega^2 L_s L_l)(R_s + R_l) + \omega^2 (R_s L_l + R_l L_s)(L_s + L_l)}{(R_s + R_l)^2 + \omega^2 (L_s + L_l)^2} \tag{115.4}$$

$$\mathrm{L}' = \frac{(R_s L_l + R_l L_s)(R_l + R_s) - (R_s R_l - \omega^2 L_s L_l)(L_s + L_l)}{(R_s + R_l)^2 + \omega^2 (L_s + L_l)^2} \tag{115.5}$$

For example, power cables models: YBF450-8-type special explosion-proof fan (Frequency Conversion) is the three-phase asynchronous motor, whose main parameters are:

Motor power: 220 kW; Rated voltage: 6 kV; Rated current: 27.5 A;

Efficiency: 94.5%; Power Factor: 0.89; Rated frequency: 50 Hz.

Thus, by using data supplied from motor, it can be estimated the stator resistance and the inductance of stator side. Power factor can be considered as 0.5.

$R_l = 2.2\,\Omega$; $L_l = 0.012\,\mathrm{H}$.

The equivalent impedance of the power bus is $Z_S = R_S + jX_S$. Usually $R_S \ll X_S$, for the analysis of convenience, R_S can be ignored. So it is similar to get $Z_S \approx jX_S$. For example, the cable bus parameters are given as follows:

Rated voltage: 6 kV; Rated frequency: 50 Hz;

Short circuit capacity: 124 MVA; Short circuit current: 6462.3 A.

According to short-circuit capacity of the system, the approximate short-circuit impedance can be estimated (system equivalent impedance)

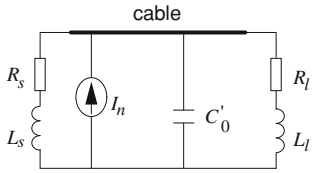


Fig. 115.4 Equivalent circuit of the power cable supply

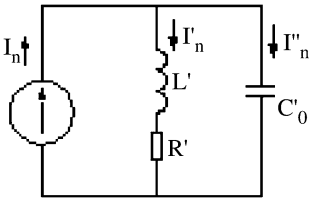


Fig. 115.5 Equivalent circuit

$$X_S = Z_d = \frac{S_d}{3I_d^2} = \frac{124 \times 10^6}{3 \times 6462.3^2} = 0.99\,\Omega \quad L_s = \frac{X_s}{2\pi f} = \frac{0.99}{314} = 0.0032\,\mathrm{H}$$

Application of Eqs. (115.4) and (115.5) may get that $R' = 0.0762\,\Omega$, $L' = 0.0031\,\mathrm{H}$

$$I_n'' = \frac{X_{L'} + R'}{X_{L'} + R' - X_{C_0'}} I_n = \frac{2\pi nfL' + R'}{2\pi nfL' + R' - \frac{1}{2\pi nfC_0'}} I_n \tag{115.6}$$

In equation (115.6), n—the number of harmonics, for example, $n = 5, 7\ldots$

Power system exists the main harmonics such as 5th, 7th, 11th and 13th harmonics. For example, the 5th harmonic amplitude is in the range of 20–30 A.

By Eq. (115.3), it can calculate range of the C_0. If the cable length is 10 km, the cable distribution of capacitance $C_0^{'}$ will change from 0.1 to 1.4 μF.

By Eq. (115.6), cable leakage current I_n'' with the 13th harmonic current I_{13} and the leakage capacitor $C_0^{'}$ simulation diagram are shown in Fig. 115.7. Figure 115.6 is simulation diagram of the cable leakage current $In^{''}$ with 19-time harmonic current I_{19} and leakage capacitor $C_0^{'}$ (Fig. 115.6).

Figure 115.6 can be seen that the circuit exists for the 13th harmonic, when its current is in the range of 20–30 A, the cable capacitive current leakage is in change between 40 and 100 A, indicating that the 13th harmonic current in the cable leakage capacitance to be amplified. With the increase in the number of harmonics, harmonic current amplification problems become increasingly serious, which can be seen to verify from Fig. 115.7. It can be found from Fig. 115.7 that alhough the 19th harmonic currents varied only between 1 and 2 A, the cable leakage current can change between 100 and 700 A. The capacitor leakage negative current in the Figs. 115.6 and 115.7 is only the current direction.

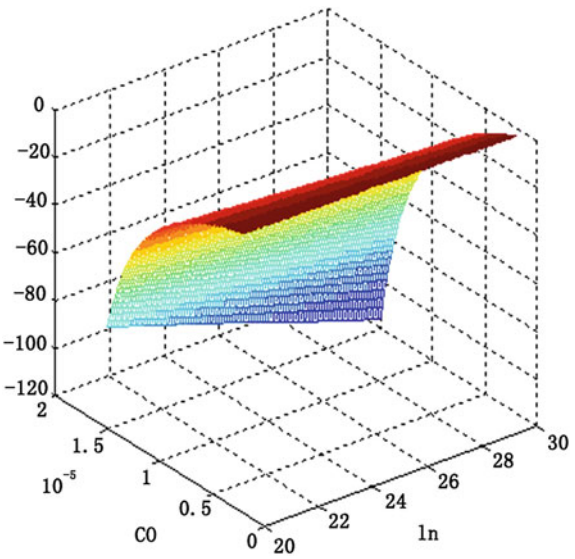


Fig. 115.6 Capacitor leakage current is along with change of the 13th harmonic current to change

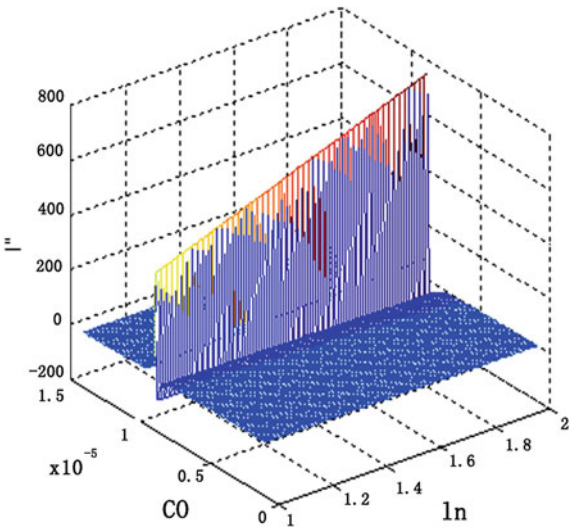


Fig. 115.7 Capacitor leakage current is along with change of the 19th harmonic current to change

115.4 Thermal Effect of Harmonic Leakage Current

Harmonic leakage current is the main component of the cable partial discharge current, so the size of harmonic leakage current is an important manifestation of the cable insulation condition. In the partial discharge process, the ionization out of the electronics, under the effect of the electric force, when they hit the insulation wall, they will have enough power to break chemical bonds of polymer insulating

materials, resulting in the phenomenon of fission and causing branches to be generated [5]. Uneven distribution of leakage current is in the small harmonic impedance of the cable local region, harmonic leakage current will collect in the region. In the Fig. 115.8, according to $P = I^2R$ the thermal effect of current, simulation results of the leakage current can be estimated in the leakage current region. When insulation medium is heat to high temperature, it will be burnt.

115.5 Conclusion

This paper is focused on the cable parameters to analyze and consider leakage capacitance of the cable as the main parameters. The cable-powered equivalent circuit diagram and parameter are proposed. Through the harmonic leakage current analysis and simulation, the conclusions are as follow:

(1) When the leakage capacitance of the cable and system parameters are in the coincidental case, the harmonic current can be amplified. When it flows into the cable capacitance, the higher frequency harmonics, the greater amplify factor.
(2) Harmonic leakage current causes the cable insulation medium heat. Especially in high-order harmonic currents in dielectric heating, the temperature will be high and result in insulating material to be burned.

References

1. Chuanqing S (2002) Questions and answers for power cable installation and operation, vol 1. China Electric Power Press, Beijing, pp 14–17
2. Li G (2007) Power cable design and construction manual, vol 14. China Electric Power Press, Beijing, p 108
3. Ren Z (2003) Coal mine power harmonic analysis and governance, vol 3. China University of Mining Press, Xuzhou, pp 38–40
4. Guanyuan Q (1999) Circuit-4 version, vol 2. Higher Education Press, Beijing, pp 32–35
5. Rihong J (1997) XLPE power cable lines, vol 5. China Electric Power Press, Beijing, pp 98–101

Chapter 116
Insulation Status Mobile Monitoring for Power Cable Based on a Novel Fringing Electric Field Method

Meng-yu Liang, Song-yi Dian and Tao Liu

Abstract Water tree formed in the insulation layer of underground power cable may cause aging breakdown of cables. Focusing on mobile monitoring for insulation status of these cables, this paper proposes fringing electric field (FEF) method to detect probably formed water tree. With regard to the disturbance expansion caused by applying traditional interdigital FEF sensor in the water tree detection, we design an arch substrate-based interdigital FEF sensor, and build its mathematical model by finite element method. Finite element analysis is carried out to compare the electric field distribution of the above two kinds of sensors in insulation layer. Effectiveness of the electric field penetrability is analyzed by varying finger width and electrode separation of the proposed sensor. The above effectiveness validation also promotes to optimize the parameters of the sensor. Simulation results demonstrate the sensor's design is reasonable and the electric field penetrability of the novel sensor is remarkably improved compared with traditional interdigital FEF sensor.

Keywords Water tree · Fringing electric field · Interdigital sensor · Insulation status monitoring · Underground power cable

M. Liang · S. Dian (✉) · T. Liu
School of Electrical Engineering and Information Technology,
Sichuan University, Chengdu 610065, China
e-mail: scudiansy@scu.edu.cn

M. Liang
e-mail: arsenal630919@yahoo.com.cn

T. Liu
e-mail: favmoon@hotmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_116,

116.1 Introduction

Underground power cables are used extensively in today's power delivery and its reliability and long-term performance is a major concern in the power industry. Crosslinked polyethylene (XLPE) is used most widely in insulation material of cables, because of its excellent mechanical, heat and electrical properties. However, after put into service, surrounding factors will affect serviceability and lifetime of cables. In order to ensure high quality of electric power supplied to customers, it is necessary to estimation of the aging status of installed cables.

One of the primary causes of aging breakdown for XLPE cable is water tree, which may be formed in the insulation due to the combined effect of water and intense electric field. Various methods have been proposed for the diagnosis of water tree, such as DC leakage current method, residual charge method and insulation impedance method. All of these methods are researches on identity, which can not locate water tree in insulation layer accurately. In literature [1], a new and innovative mobile monitoring method is proposed to replace manual inspection. In this method, fringing electric field (FEF) sensor is applied to measure the presence of water in the bulk of the insulation. At present, mostly applications of fringing electric field dielectrometry method are on monitoring of moisture concentration in transformer pressboard.

In this paper, a novel fringing electric field method is adopted in mobile monitoring to diagnose water tree. An arch substrate based interdigital FEF sensor is proposed to improve the accuracy of water tree detection. Mathematical model of the novel sensor is built by the finite element method. We also discuss the effectiveness on electric field penetrability by finite element simulation, and optimize the sensor's parameters including the finger width and electrode separation.

116.2 Principle of Mobile Monitoring

Insulation status mobile monitoring of underground power cables is based on multi-sensor information fusion technology. Monitoring system depends on the real-time processing of information that collected by sensors, estimates the status of insulation. This system also depends on a robot platform which can move along underground power cables. We have developed the robot platform with such function for insulation status mobile monitoring [2]. The framework of the robot platform, as shown in Fig. 116.1, is designed as a novel shrimp rover-based mobile robot. Hourglass-shaped neoprene wheels can keep the platform travel along the cable, and climb over obstacles stably. The measurement subsystem of the robot platform includes four modules: water tree detection module, partial discharge detection module, overheating monitoring module, and video monitoring module.

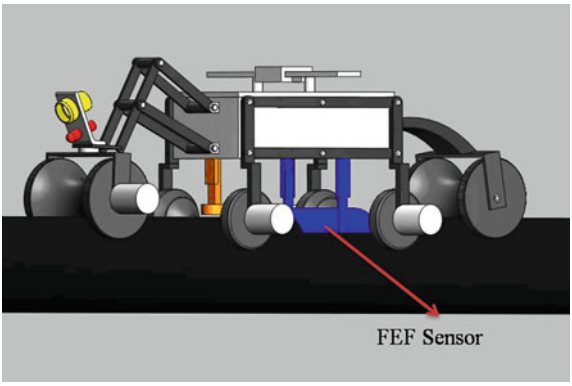


Fig. 116.1 Shrimp rover-based mobile robot for monitoring underground power cable

The inspection robot can perform monitoring tasks either autonomously or under remote control. When under autonomous mode, infrared sensor works continuously, take temperature readings as robot travels along the cable. If general scan detects a hot pot existing, the robot will stop and take detail measurement. FEF sensor and acoustic sensor are driven by a balanced mechanism to contact with cable surface. While detection accomplished, sensor data will be saved in control board to make judgment on aging status. Within the communication distance, robot works under remotely operation, transmit information back to the host computer, and can be host to measure insulation status of appointed position.

Diagnostic sensors chosen for mobile monitoring must be able to make detection without destroying physical structure of cables. Sensors must be small and lightweight that can be easily assembled into the platform, and must be have high precision that can make correct judgment in real-time. Therefore, fringing electric field method could be used as an effective way to measure the aging status of cable insulation layer in mobile monitoring. Dielectric properties of sample are associated with physical characteristics in the method. Presence of water in bulk of the insulation layer will change the permittivity of insulation material. And the measurement of dielectric properties of outer layer can provide valuable insights into insulation status.

116.3 Mathematical Model

A FEF sensor can be visualized as a parallel-plate capacitor whose electrodes open up to provide one-sided access detection. Electric field lines pass through the material under test, therefore, capacitance and conductance between two electrodes depends on dielectric properties of the sample. And the performance of FEF sensor will affect precision of water tree detecting.

A traditional interdigital FEF sensor, which used for nondestructive detection, is presented in Fig. 116.2. The drive and sensing electrodes are designed as

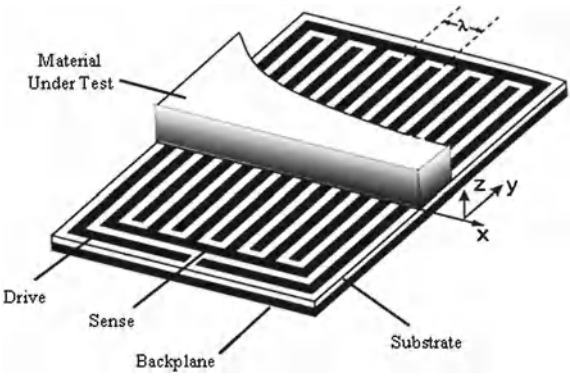


Fig. 116.2 Schematic diagram of interdigital FEF sensor detection [3]

coplanar fingerlike pattern that repeated several times, produce periodic varietal electric potential difference along the surface in the x direction. The penetration of fringing quasi-static electric fields above the interdigital electrodes is proportional to the spacing between the centerlines of sensing and driven fingers. The interdigital electrodes enlarges capacitance of sensor, improves the electric field distributing, so that it can guarantee better accuracy. However, in the case of cable monitoring, the traditional FEF sensor contact with cable surface lack-of-fit. When temperature and humidity change, the variety of atmosphere dielectric property will interference the detection result.

A novel arch substrate based interdigital FEF sensor is designed to overcome this problem, it reduce the gap between sensor and cable, and provide contact surface as large as possible. Because lacking of characteristic and geometrical features of cable insulation, 2D electric field distribution model is not convenient to express surface capacitance of the novel sensor. Whereas, potential distribution calculation is still follow the law of charge conservation (Fig. 116.3).

Finite element method (FEM) is chosen to build model of the novel sensor. FEM is based on the idea of partitioning bounded domains in into a number of small, non-overlapping subdomains, the finite elements. After the scalar potential distribution is figured out by FEM, electric charge quantity on sensing electrode can be calculated. If it is known quantity of the potential difference between driven and the sensing, the capacitor can be calculated in the following way:

$$C = \frac{Q}{U} \tag{116.1}$$

When the space is assumed without free charge, the sensor model based on Poisson equation:

$$\nabla[\varepsilon_0 \varepsilon(x, y) \nabla \phi(x, y)] = 0 \tag{116.2}$$

is satisfied everywhere in the space. In this equation, ε_0 is the permittivity of vacuum, $\varepsilon(x, y)$ is the relative permittivity distribution on cross section of sensor, $\phi(x, y)$ is electric potential distribution. As division element shown in Fig. 116.4, the potential distribution can be obtained as follows:

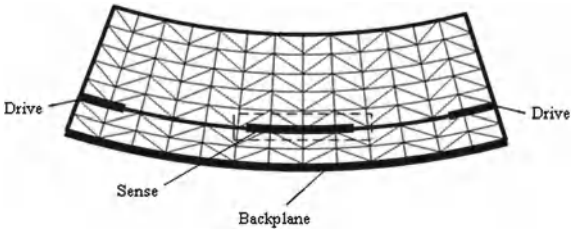


Fig. 116.3 The simplified model for novel interdigital FEF sensor

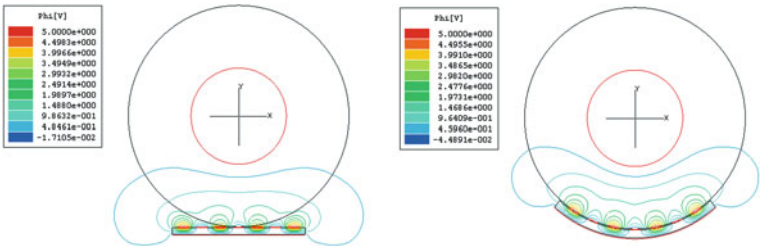


Fig. 116.4 Equipotential line in insulation layer produced by FEF sensors

$$\phi = \frac{\varepsilon(x,y)}{2S}\left[\sum_{p=1}^{3}(a_p + b_p x + c_p y)\phi_p\right] \tag{116.3}$$

where, a_p, b_p, c_p are the coefficient depending on vertex of triangle element, and ϕ_p is the electric potential of vertex. From the electromagnetic field theory, we obtain:

$$E(x,y) = -\nabla\phi(x,y) \tag{116.4}$$

By Gauss theorem, we obtain electric charge quantity on sensing electrode:

$$Q = \int_{\Gamma} \varepsilon_0\varepsilon(x,y)Eds = -\int_{\Gamma} \varepsilon_0\varepsilon(x,y)\nabla\phi(x,y)ds \tag{116.5}$$

where Γ is the camber surrounding the sensing. Assume N is the number of electrodes in the array, we substitute (116.5) into (116.1) to yield, then:

$$C = N\frac{Q}{U} = -\frac{N\varepsilon_0}{U}\int_{\Gamma} \varepsilon(x,y)\nabla\phi(x,y)ds \tag{116.6}$$

116.4 Finite Element Analyse

The software package Maxwell 2D by Ansoft Corp was used for the finite element simulations. Electrical field is assumed uniform along the surface in the y direction of the sensor in Fig. 116.2. Therefore, entire field distribution can be obtained through analysis of the plane perpendicular to the y direction.

2D AC Conduction Field Solver is adopted to solve the model. Drawing the simplified model at Cartesian coordinate, we set XLPE insulation layer with dielectric constant 2.3 and conductivity 10.8 S/m. Material of electrodes is copper and substrate is Teflon. The signal source for the driving electrode is alternating current with amplitude of 5 V. The sensing electrode and the backplane are connected to ground. With the automatic elements distribution in the entire area accomplished, denser distribution is performed in the area near the electrode to get an accurate solution.

The simulation results of equipotential lines in the insulation layer, which caused by sensors with varied substrate shape, are shown in Fig. 116.4. Obviously, the equipotential line produced by traditional sensor is almost parallel to plane. The two fingers close to the cable surface have better penetration capability, while the outside fingers lead to the electric field mostly in the air between the sensor and the cable, with weaker penetration capability. In the figure which shows equipotential lines produced by the novel interdigital FEF sensor, we can get observation that four fingers of the sensor posses approximately equal penetration capabilities. Equipotential lines which around sensing fingers are in sharper shape, and are denser in the insulation layer. Meanwhile, equipotential lines along the edge of the novel sensor have uptrend with a certain degree arch. Thus, covering area of the electric field expands, which results in more accurate detection results.

116.5 Parameter Optimization

In order to conduct further research into effect of the novel FEF sensor in water tree detection of cables, electrode separation and finger width are varied respectively. Effectiveness of the electric field penetrability is analyzed with FEM simulation, which leads to the optimization of parameters.

When comparing the electric field strength in the insulation layer, we apply normalized method to set aside influence caused by driving voltage. Attenuation ratioη is used to represent variation of the field strength, which is defined as follow:

$$\eta = (E/E_0) \times 100\% \tag{116.7}$$

where E is field strength on the radius of the cable in vertical direction, E_0 is the field strength of the intersection point of that radius and insulation surface. After solve electric field by Maxwell, we fit the data at a coordinate with MATLAB.

Keeping finger width as 3 mm, sensors with separation constant 2.0, 2.5, 3.0, 3.5, 4.0 mm are chosen for simulation and obtained respective electric potential strength on the circumference of the circle with radius 23 mm. Curves of electric potential and field density are shown in Fig. 116.5. We can see from the plot that with separation increases, magnitude of electric potential on the circumference

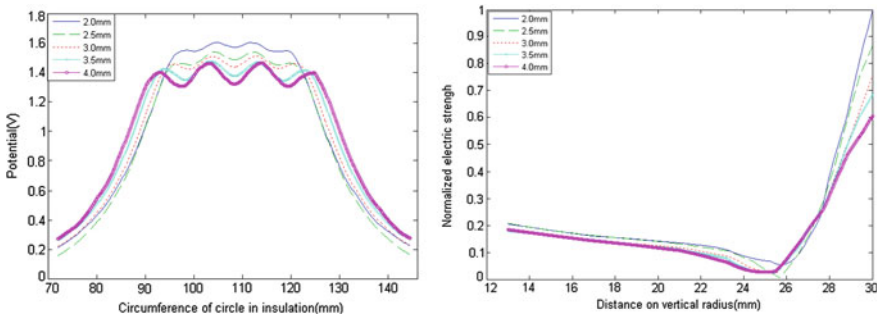


Fig. 116.5 Electric potential and field density curves with variation of electrode separation

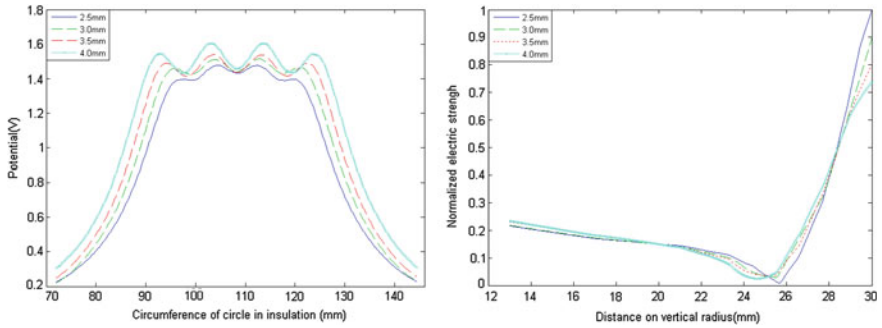


Fig. 116.6 Electric potential and field density curve with variation of electrode finger width

decreases gradually, whereas covering area of electric field expands. The normalized field density curves on vertical radius shows that the magnitude of the electric field strength nearest to the sensor decreases while attenuation speed slows down.

Keeping separation as 3 mm, sensors with finger width 2.5, 3.0, 3.5, 4.0 mm are chosen for simulation and obtained respective electric potential strength on the circumference of the circle in the cable insulation layer with radius 23 mm. Curves presented in Fig. 116.6 shows, while finger width increases, magnitude of electric potential on the circumference increases, the covering area of electric field grows as well; hence, the penetrating capability of sensor augments. The normalized electric field strength on the vertical radius caused by sensors of varied finger width shows, with finger width increasing, field strength at surface of the insulation layer decreases and attenuation speed slows down gradually. Due to the demand that the size of the sensor should be small enough in mobile monitoring, the finger width is limited in a certain degree. Based on available area of practical application, we should pick interdigital electrodes with the largest possible finger width.

116.6 Conclusion

This paper designed a novel interdigital FEF sensor to provide an innovative way for insulation status measurement. By finite element simulation, the electric potential and the field distributing are analyzed to compare the novel FEF sensor with the traditional one. It is proved that the novel sensor is superior in water tree detection since it can contact cable surface tightly. Parameters including finger width and electrode separation are optimized in simulation, which will instruct the physical design of the proposed sensor and water tree detection experiment while it is applied in the developed robot platform in our future work.

References

1. Jiang B, Sample A, Wistort R (2005) Autonomous robotic monitoring of underground cable systems. In: Proceedings of international conference on advanced robotics, pp 673–679
2. Dian SY, Liu T, Liang Y, Liang MY, Zhen W (2011) A novel shrimp rover-based mobile robot for monitoring tunnel power cables. In: Proceedings of 2011 international conference on mechatronics and automationm, pp 887–892
3. Mamishev A, Kishore SR, Yang F (2004) Interdigital sensors and transducers. In: Proceedings of the IEEE, pp 808–845
4. Jiang B, Mamishev A (2004) Robotic monitoring of power systems. IEEE Trans Power Delivery 19:912–918
5. Mamishev A, Du Y, Lesieutre BC (1999) Development and applications of fringing electric field dielectrometry sensors and parameter estimation algorithms. J Electrostat 46:109–123
6. Werelius P, Tharning P, Eriksson R (2001) Dielectric spectroscopy for diagnosis of water tree deterioration in XLPE cables. IEEE Trans Dielectr Electrl Insulation 8:27–42
7. Sun Q, Shi TM, Liu J (2009) The methods of capacitance calculation between the electrodes of sensor in ect system. China Instrumentation 72–75
8. Chen DY, Li MZ, Zheng GB (2006) Analysis and parameter optimizing of multi-electrode capacitance transducers based on finite element method in electrical capacitance tomography system. Chinese J Electr Devices 29:427–430

Chapter 117
Application of Bacteria Foraging Algorithm on Harmonics Optimal Control for Inverter

X. J. Li, Q. Pan and K. F. He

Abstract In the process of converter driving mine hoist, the output voltage contains a lot of harmonics. Aiming at this problem, Harmonics optimal control based on bacterial foraging algorithm is developed in this study. Establishing the objective function to minimize total harmonic distortion (THD) with certain constraints, using bacterial foraging algorithm to solve a series of nonlinear transcendental equation to get the optimal switching time in a certain frequency range and to determine the witching time of the power devices. Compared with other literature, the results show that the application of bacterial foraging algorithm can more effectively suppress the harmonic amplitude in theory, and the total harmonic distortion is smaller.

Keywords Bacterial foraging algorithm · Switch time · Harmonics

117.1 Introduction

Alternating Current (AC) variable frequency speed regulation system with high efficiency, wide range and high precision, and other characteristics, widely used in mine hoist. However, due to the use of the high-power diode rectifier and

X. J. Li · Q. Pan (✉) · K. F. He
Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiantan 411201, China
e-mail: panqing_86@163.com

X. J. Li
e-mail: hnkjdxlxj@163.com

K. F. He
e-mail: hkf791113@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_117,

transistors inverse in converter, the inverse usually adopted Sinusoidal Pulse Width Modulation (SPWM). SPWM is using sinusoidal signal to modulate carrier led to the harmonic generation associated with the carrier and the signal wave. The harmonic of the output voltage was too large, which resulted in not only increasing the waveform deviation factor and harmonic loss of the power supply network and reducing the output efficiency of converter, and mechanical load vibration is also growing. Moreover, the running speed variation of mine hoist has bad influence on the whole AC variable frequency speed regulation system, especially it will cause an obvious fluctuation of the network voltage in the case of long time and repetitive running of the mine hoist [1].

With the increase of switching frequency of the sine pulse width modulation technology, the waveform can be improved better. The high-frequency SPWM technology have been widely used in small and medium-sized inverter. However, for the large capacity current type of inverter, the higher switching frequency is, the greater switching loss is. When the switching frequency is much higher, switch loss gets larger, it is not advisable. Especially in some high-power switching devices, it is not suitable for improving the waveform by increasing switching frequency. In this case, the optimization PWM technology is the most appropriate. The improvement of the waveform is by choosing the appropriate switching times, does not need high switching frequency.

The principle of optimization PWM technology is: according to a specific optimization objective, it calculates the switch mode in a certain frequency range in the offline state to minimize the target function, then stores the results and outputs it through the table look-up or other methods. However, the calculation of switching times for SPWM is a series of nonlinear equations and it is very difficult to solve. In recent years, the solutions are widely used, such as Newtonian algorithm, Homotopy algorithm, Genetic algorithm, Particle swarm optimization algorithm and so on. The bacterial foraging algorithm is a new bionic optimization algorithm, it is proposed by studying the behavior of swallowing the food of *Escherichia coli*. The algorithm has the advantage of group intelligence, parallel search and can easy to jump out of the local minimum point, fast convergence speed [2, 3]. Bacterial foraging algorithm is firstly applied to optimize a three level voltage pulse width modulation inverter in this paper, it can more effectively suppress the harmonic amplitude in theory, and the total harmonic distortion is smaller.

117.2 Harmonic Mathematical Model

117.2.1 Inverter Output Voltage Harmonic Mathematical Model

Three-phase bridge inverter circuit is widely used, Fig. 117.1 shows the voltage source three-level inverter circuit. U_{dc} is the DC voltage; the inverter output voltage is U_{AN}, L, C are the filter inductor and filter capacitor, U_m is the output voltage after filtered

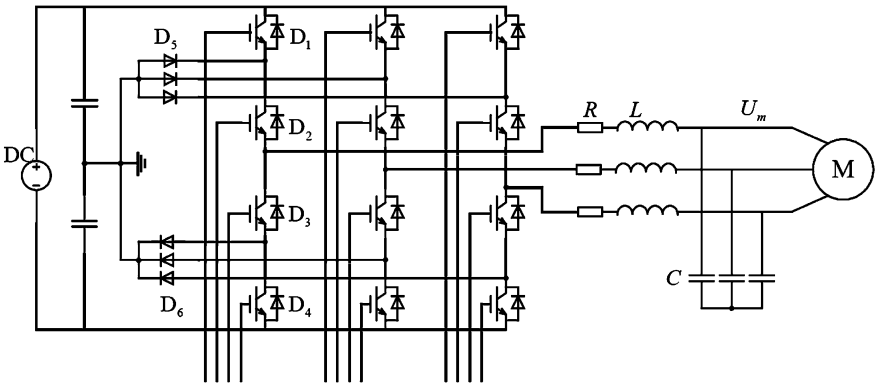


Fig. 117.1 Three-level high-voltage inverter system

Figure 117.2 is a three-level SPWM waveform. It can be seen: N is the number of switching times in a quarter-period, that is $\delta_1, \delta_2, \ldots, \delta_N$, which is symmetrical of $\pi/2$ on interval $[0, \pi]$ and is inversely symmetrical of π on interval $[0, 2\pi]$. The inverter output voltage waveform is an odd function, but also odd harmonics function, the other switching angles can use the N switching angles to represent. Once given the first N pulse switch time of quarter-period, the output voltage waveform of A-phase can be uniquely determined, B, C phase voltage lag $2\pi/3$ in order. The output voltage of three-level inverter and the filter can be uniquely determined. Evidently, these angles must satisfy the basic constraint:

$$0 < \alpha_1 < \alpha_2 < \alpha_3 < \cdots < \alpha_N < \frac{\pi}{2} \tag{117.1}$$

The Fourier representation of the A-phase voltage can be expressed as:

$$u_A(\alpha) = \sum_{k=1,3,5\ldots}^{\infty} U_{Ak}\sqrt{2}\sin(k\alpha) \tag{117.2}$$

where U_{Ak} is the RMS value of the harmonic components of A-phase voltage u_A:

$$U_{Ak}(\delta) = \frac{4U_{\mathrm{dc}}}{\sqrt{2}k\pi}\sum_{i=1}^{N}[(-1)^{i-1}\cos k\alpha_i] \tag{117.3}$$

In the three-phase systems, the triplen harmonics in the phase voltages will be canceled in the line–line voltages, so the output line voltage RMS value of fundamental and harmonic is:

$$U_{Ak}(\alpha) = \frac{4\sqrt{3}U_{\mathrm{dc}}}{\sqrt{2}k\pi}\sum_{i=1}^{N}[(-1)^{i-1}\cos k\alpha_i], \quad k = 1, 5, 7, 11, 13\ldots \tag{117.4}$$

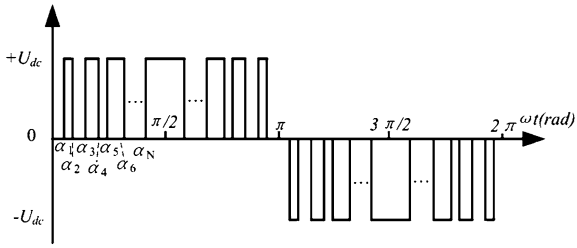


Fig. 117.2 Three-level SPWM output voltage waveform

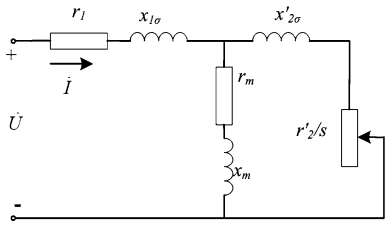


Fig. 117.3 The equivalent circuit of motor

117.2.2 Mathematical Model of Harmonic After Filtering

The basic consideration of the specific optimization objective is not based on the inverter itself, but the inverter and the load to be considered as a whole system, then get the best results by controlling. Figure 117.3 shows the equivalent circuit load motor.

Equivalent impedance can be expressed as:

$$Z_{\text{eq}} = r_1 + jx_{1\sigma} + \frac{(r_m + jx_m)(\frac{r'_2}{s} + jx'_{2\sigma})}{r_m + jx_m + \frac{r'_2}{s} + jx'_{2\sigma}} = R_d + j\omega L_d \tag{117.5}$$

Furthermore, from Fig. 117.3 we can get the filter output phase voltage values, which is:

$$U_{APk}(\alpha) = \frac{4U_{\text{dc}}}{\sqrt{2}k\pi}\sum_{i=1}^{N}[(-1)^{i-1}\cos k\alpha_i]H(k\omega), \quad k = 1, 5, 7, 11, 13\ldots \tag{117.6}$$

where:

$$\begin{aligned} H(k\omega) &= \left\{\dot{U}_m\right|/\dot{U}_{AN} \\ &= \{(R_d^2 + k^2\omega^2 L_d^2)/[(R_d + R - k^2\omega^2 L_d RC - k^2\omega^2 R_d LC)^2 \\ &\quad + (k\omega L_d + k\omega L + k\omega R_d RC - k^3\omega^3 L_d LC)^2]\}^{1/2} \end{aligned} \tag{117.7}$$

The RMS value of the filter output line voltage is:

$$U_{ALk}(\alpha) == \frac{4\sqrt{3}U_{\mathrm{dc}}}{\sqrt{2}k\pi}\sum_{i=1}^{N}[(-1)^{i-1}\cos k\alpha_i]H(k\omega), \quad k = 1,5,7,11,13\ldots \tag{117.8}$$

Let be $\alpha = \omega t$, the formula (117.8) can be written as:

$$U_{ALk}(t) = \frac{4\sqrt{3}U_{\mathrm{dc}}}{\sqrt{2}k\pi}\sum_{i=1}^{N_2}[(-1)^{i-1}\cos k\omega t_i]H(k\omega), \quad k = 1,5,7,11,13\ldots \tag{117.9}$$

117.2.3 The Establishment of the Objective Function

The objective function is to find the switching times to make the fundamental output voltage $U_{AL1} = U_{\mathrm{ref}}$, and minimize the THD. THD defines the relationship between the total harmonic voltage and the fundamental voltage [4]. Therefore, we have:

$$\eta_{THD} = \sqrt{\sum\nolimits_{k=2}^{k=50} U_{ALk}^2}/U_{AL1} \tag{117.10}$$

For N is number of pulses per quarter-period, let $F(t)$ be the objective function; then, we can write the problem as:

$$\begin{aligned} &\min: F(t) = \left(\frac{1}{U_{AL1}}\right)\sqrt{\sum_{m=1}^{8}\left(U_{AL(6m-1)}(t_i)^2\right) + \left(U_{AL(6m+1)}(t_i)^2\right)} \\ &\quad i = 1,2,3\ldots N, \\ &\quad k = 6m \pm 1, m = 1,2,3\ldots,8 \\ &\quad s.t.: 0 < t_1 < t_2 < t_3 < \cdots < t_N < \frac{\pi}{2\omega} \end{aligned} \tag{117.11}$$

U_{AL1} is the fundamental voltage; U_{Alk} is the total harmonic voltage; $k = 6\ m \pm 1$ is the highest harmonic number, and the THD is computed throughout the 50th.

117.3 Principle and Process of BFA

BFA is proposed by studying the behavior of swallowing food of *Escherichia coli*: Bacteria find and access to food through surging and rolling which is the two trends movement in order to maximize the obtained energy in unit time.Based on

foraging theory, natural selection will usually eliminate the individuals with poor foraging ability, and the left has a good foraging strategy, because these individuals have a greater reproductive capacity. Each generation is a natural choice, after many generations, the poor foraging strategies may be eliminated or re-engineering, then evolve into a good strategy [5]. Bacteria foraging algorithm optimization process:

(1) Initialize the population, assuming that the solution space is the environment for the bacteria. In the environment, randomly generate S independent bacteria, and make it evenly distributed, the location coordinates of the bacteria are the independent solutions, such as

$$P^i_{j,G} = (p^i_{j,t_1,G}, p^i_{j,t_2,G}, \ldots, p^i_{j,t_N,G}) \tag{117.12}$$

N is the number of control variable which need to optimized, that is the dimension of solution space; $i = 1, 2, \ldots S$; $j = 1, 2, \ldots N_c$; $G = 1, 2, \ldots, G_{\max}$; N_c is number of the tendency operations; $G_{\max}$ is the maximum iteration number; $P^i_{j,G}$ is the position of i bacteria, in the G generation and after the j tendency operation in the solution space;

(2) Evaluate the location of the bacteria, that is, calculate the fitness with equation (117.11).
(3) Randomly generate a unit vector, according to (117.12) update the location of each bacteria, if the fitness improve and the number progress is less than N_s, go ahead and update the fitness function value. Otherwise, take a forward position of the bacteria and calculate the next bacteria.
(4) If the tendency operation number $j < N_c$, return (117.3) for the next tendency operation, Otherwise, for the copy operation. Get sum of the fitness to each bacteria after the N_c times tendency operations as a health function, to measure the obtained food of the bacteria, then eliminate $N_{re}/2$ bacteria which get the least food, another $N_{re}/2$ bacteria reproduce. The total number of bacteria unchange.
(5) If the copy operation number $k < N_{re}$, return(117.4). Otherwise the copy operation end, the migration operations began, according to the migration probability P_{ed}, re-initialize the location of the selected bacteria, the elite retained.
(6) If the migration operations number $l < N_{ed}$, return(117.5). Otherwise, end the migration operations, and output results.

Table 117.1 Results obtained by BFA and other methods [5]

No.	1	2	3	4	5	6	7
S.t.(BFA)	0.2006	0.8006	1.2503	1.7096	3.2433	3.2502	4.4899
S.t. (PSO) [5]	0.31829	0.77961	2.36056	3.36250	4.25566	4.66052	4.70548
S.t. (NS) [5]	1.11649	1.41733	2.24786	2.79931	3.40900	4.12390	4.161432
S.t. (GA) [6]	2.49225	2.77028	3.33031	3.78028	3.80002	3.99711	4.22536

Table 117.2 Fundamental value, harmonic value and total harmonic distortion

U_{ALk}	BFA	PSO	NS	GA
$k = 1$	1.00052	0.999222	1.000000	1.004393
$k = 5$	0.00674	0.007551	−0.000000	0.044009
$k = 7$	0.00257	0.002334	−0.000005	−0.004009
$k = 11$	−0.00102	0.001063	−0.010226	−0.002698
…	…	…	…	…
$k = 49$	0.00110	−0.001913	−0.003083	−0.000083
η_{THD}	0.0241	0.0390	0.2200	0.0976

Table 117.3 The total harmonic distortion in a frequency range

f(Hz)	30	35	40	45	50
N	13	11	9	7	7
f_s	780	770	720	630	700
η_{THD}(BFA)	0.0423	0.0115	0.0195	0.0138	0.0041
η_{THD} (PSO)[5]	0.1193	0.0926	0.0996	0.0584	0.0390
η_{THD} (NS)[5]	0.2572	0.3814	0.3329	0.3206	0.2200

117.4 Application

Bacteria foraging algorithm optimize switching time by Matlab programming. Given parameters according to [6]. The optimization process is as follows:

(1) Calculate the fundamental voltage, and the corresponding equivalent load $R_d + j\omega L_d$ with the output frequency f.
(2) according to (117.11), calculate the switching times $t_1 \sim t_N$ under the frequency of 50 Hz.
(3) Let be the switching frequency $f_{\mathrm{switch}} = fN \leq 800$ Hz, obtain the switching times of each frequency and the total harmonic distortion.

In this paper, bacteria foraging algorithm is used to calculate the switching times at the output frequency $f = 50$ Hz. The switching times show in Table 117.1 respectively calculated by BFA algorithm and methods in literature [6]. Table 117.2 is fundamental value and harmonic value of the output voltage respectively calculated by BFA algorithm and methods in literature [6]. Table 117.3 is the total harmonic distortions at each frequency respectively calculated by BFA algorithm and methods in literature [5]. From the results,

compared to other algorithms, the average time of switching times and the total harmonic distortion optimized by bacterial foraging algorithm are smaller. Through comparison and analysis, the switching times optimized by BFA algorithm can be better suppressed the amplitude of inverter output harmonics voltage and effectively reduce the total harmonic contents, and can meet the requirements of engineering practice.

117.5 Summary

Bacterial foraging algorithm approach is first used to find the switching time and improve SPWM control characteristics of high-voltage inverter equipment. By comparing the results with other methods, the switching times optimized by BFA algorithm can better suppress the amplitude of inverter output harmonics voltage and effectively reduce the total harmonic contents, and meet the requirements of engineering practice.

Acknowledgments CEEUSRO special plan of Hunan province (2010XK6066), Aid program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province, Industrial Cultivation Program of Scientific and Technological Achievements in Higher Educational Institutions of Hunan Province (10CY008), Natural Science Foundation of Hunan Province Key Project (09JJ8005), are gratefully acknowledged.

References

1. Li XJ, Pan Q, He KF (2011) Option model calibration using a bacterial foraging optimization algorithm. submitted to J Comput
2. Ying C, Shao ZB (2005) Application of bacterial foraging technique trained artificial and wavelet neural networks in load forecasting. J Shenzhen Univer Sci Eng 25–2:153–157
3. Zhou YL (2010) Biologically-inspired algorithms for financial modelling. Comput Eng Appl 46–20:16–21
4. Huang H, Ji YC (1997) A hybrid genetic algorithm and bacterial foraging approach for global optimization. Proc CSEE 17(5):344–347
5. Wang YL (2010) A hybrid genetic algorithm and bacterial foraging approach for global optimization. Coal Mine Mach 31(11):59–62
6. Lou HB, Mao CX (2007) A hybrid genetic algorithm and bacterial foraging approach for global optimization. Proc CSEE 27(33):108–112

Chapter 118
Research on Low-Voltage Early Warning System of Coalmine Power Grid

Tieying Zhao and Na Wang

Abstract Research is proposed on coalmine power grid safety-running in case to insure coalmine normal production. According to the Low-voltage problem now still lying in the coalmine power grid, a method combined fuzzy comprehensive evaluation with gray comprehensive evaluation is proposed, selects power grid evaluation indicators, sets the impact judgment hierarchical matrix of power grid operation, judges the coalmine power grid state in the future by those index value changes, and uses different colors to indicate different power grid states, gives early warning signals if the state is below a certain level. Simulation results show that this method can accurately determine the possible development trend of power grids.

Keywords Coalmine power grid · Low-voltage early-warning · Fuzzy comprehensive evaluation · AHP

118.1 Introduction

Coal is the primary energy supply of China, and its proportion in China's energy production and consumption is now more than 70% [1]. The main power supply of coal mine is electricity, so its power grid state has a direct effect on the safe production of the coal mine. The total power of coal production equipment have

T. Zhao (✉)
Department of Mechanical and Electronic and Information Engineering,
China University of Mining and Technology, Beijing 100083, China
e-mail: hnjzzty@163.com

N. Wang
Department of Electrical Engineering, Jiyuan Vocational and Technical College,
Jiyuan 454650, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_118,

increased sharply, equipment, heavy-duty concentrated expression in the large-capacity, high-power, high reliability, high voltage. Total installed shearer capacity reaches to 2,390 kW, and installed scraper conveyor capacity is 3,000 kW. Total installed capacity of major equipment at the work face has been more than 8 MW, if the belt conveyor is counted in, the total capacity of a comprehensive work face will be more than 10 MW [2]. Constrained by the underground space of coal mine, the major coal mine power grid network uses a single-ended form of laying-like, the grid has long power supply distance and greatly changed loads. Thus, low voltage will happen at the end of the feed-lines frequently. Low voltage of coal mine power grid causes great losses of electricity and economy; it also has brought great harm to the safe operation of equipment underground [3].

Therefore, it is necessary to implement research on early warning monitoring system of coal mine. We can adjust the coal mine grid operation mode in time through real-time voltage status analysis and voltage trend prediction, and take necessary measures to ensure the quality of its electric power supply, which is one of the safe insurances and reliable measures of coal mine production.

118.2 Factors Affecting Coal Mine Power Voltage

Coal mine power grid is a part of the power system; it belongs to the end user. There is no generating equipment in the coal mine power grid, and it neither produces electricity nor feeds back electricity to the power system. Its main function is to achieve electricity distribution and use at the whole coal mine, so its electricity flows in one direction, that is, always flows from the power system to coal mine power system grid. As a result, we can take it as a load site in power system analysis.

Coal mine power grid has always adopted a single-ended radial form of one end supplying power and one end using electricity. A single coal mine power grid is an interior power supply system of the whole power system, so the power system which has been connected with coal mine power grid can be considered as an infinite large-scale system. During the coal mine power grid running process we need not consider the problem of power balance at both ends of the grid, namely the frequency stability problem.

In the actual operation of mine power grid, the main parameter affecting its stable operation is supply voltage; therefore, coal mine power grid early-warning research focuses on grid voltage stability and possible voltage trend analysis.

Power system analysis calculation shows that amplitude of the voltage across the circuit is mainly determined by the vertical difference component $\triangle U$, while voltage phase angle difference is mainly determined by the horizontal component δU [4].

$$\Delta U = \frac{PR + QX}{U} \tag{118.1}$$

$$\delta U = \frac{PX - QR}{U} \tag{118.2}$$

Among which, P, Q is the active and reactive power of the line end, U_j is the voltage of the line end.

$\triangle U$ is mainly connected with voltage value and power of the line end. In the case that voltage of first end has been determined, $\triangle U$ is mainly decided by power of the line end, greater P, Q value is, greater the corresponding voltage will drop, greater the loss of line voltage will have. Coal power grid lines, mostly of cable lines, reactive power transmission along those lines will cause more obvious voltage drop, and that lower voltage of the end, in turn, increases voltage drop, brings much bad effects on normal operation of equipment at the end of the line. There are a large number of reactive power load in the mine power grid, therefore, these load has an important impact to the voltage supply of coal power grid during normal coal power grid operation.

The analysis above can show that line end voltage is closely related to the reactive power load level transmission line carries in the mine power grid. By analyzing load composition associated with voltage and other voltage-effected factors, we can predict the line voltage level, then, analysis and forecast the entire grid running state.

118.3 Mine Voltage Warning System Framework

In order to meet the complex environment of coal mines and the special requirements of power supply, coal power warning system is from two aspects respectively for running on-line analysis and early warning of the mine power grid: static voltage stability analysis and early warning, early warning results analysis and processing. System structure is shown in Fig. 118.1.

With the development of power system automation technology, many studies using power system automation technology are applied to coal mine remote power monitoring, coal power has achieved the remote monitoring of the whole power, which can real time check the various parts power data of coal power network [1, 5, 6]. Based on the existing mine power remote monitoring system SCADA (EMS), to establish the mine voltage early warning system and getting real-time data from the coal mine monitoring system carry on data classified and processing through the information analysis and pre-processing module. Then, the system respectively will be required information into the voltage stability judgment module and the voltage development trend analysis module to related judgment, and judging the results will be convened the data analysis module, which has a comprehensive judgment for the results to get the power integrated running and the corresponding control measures. Finally, the early warning results and control

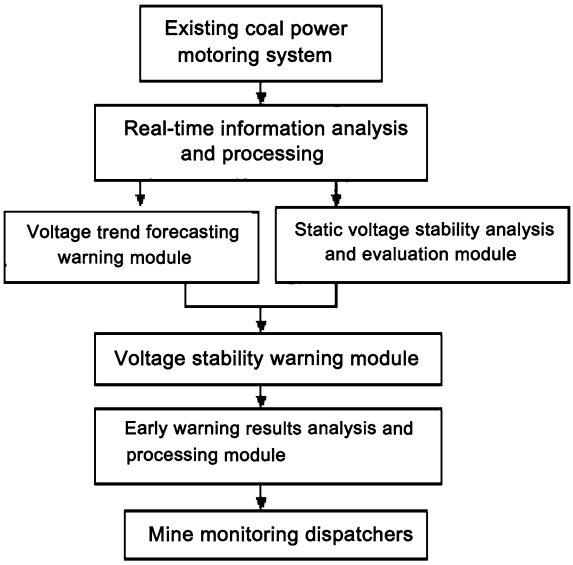


Fig. 118.1 Mine voltage warning system framework

measures will be displayed to the dispatcher as a reference of the adjustment power.

Voltage warning analysis has two parts: a voltage comprehensive assessment and a voltage possible development trends judgment.

118.3.1 Voltage Comprehensive Assessment

The voltage comprehensive assessment module uses the gray fuzzy comprehensive evaluation to analyze the current operating parameters of various power and other factors of affecting voltage level, using the analytic hierarchy process (AHP) to analyze various factors for the degree of influence of voltage [7–9]. After the future several time, the module will be make a judgment for voltage level roughly.

In the comprehensive assessment module of voltage level, to establish the mine voltage level judgment hierarchical structure first of all uses the analytic hierarchy to establish the mine voltage level judgment hierarchical structure [10, 11].

According to the different voltage levels for the power affect of coal power actual operation, the voltage level is divided into six different levels from high to low, adopting different colors respectively, which is black, red, orange, yellow, blue, green to expresses the different level of tension of the voltage levels. Analyzing the main factors of impact voltage stability, setting the index set of impact voltage, and according to the various factors for the voltage state level that have the different influence, which can establish the importance judgment matrix e of voltage, calculating the importance weights of the indicators.

Setting different fuzzy membership functions, the selected voltage impact indicators carry on fuzzy processing, which can be indicators fuzzy membership matrix. Then the methods of using fuzzy synthetic and the weight vector synthesis of the index importance can get the voltage rating values and the corresponding color judgment value of voltage level.

118.3.2 Warning Result Analysis and Judgments Processing

In this paper, the results of coal power warning system use the expert system to analyzing and processing. In the coal mine power warning systems, expert system is located in the results processing and analysis of the whole system, which are mainly on the front of the voltage judgment and to determine the voltage trend warning analysis system. Then, according to warning signal combination results of two different, giving the appropriate and reasonable control measures for coal power structure to carry on reasonable adjustment.

118.4 Case Study

In this paper, an example of a mine in Henan Province, collecting data to establish coal mine power warning system, and then using power system simulation software to simulate the mine power system state. Now, selecting three different states of the coal supply network to simulate: (1) State 1, real-time power status; (2) State 2, after the power warning system analysis, expert system will provide the power state of system control measures adjusted; (3) State 3, not to take any control measures, the power operation state after a period of time. The power real-time state parameter is provided by the monitoring system; If power is adjusted, the state parameters provided by the expert system provide system adjustment method to be calculated; If power isn't adjusted, selecting the state parameters are based on the Mine historical data and combined with load analysis and forecasting, selecting the appropriate power parameters to establish the simulation model for power state simulation. Finally, through three kinds comparative analysis above to verify the results of early warning system.

118.4.1 Simulation Results

Using Power World to established coal power system simulation diagram, and choosing Newton–La rapid flow calculation method to calculate respectively the trend distribution of power network under three states. Flow calculation shows that

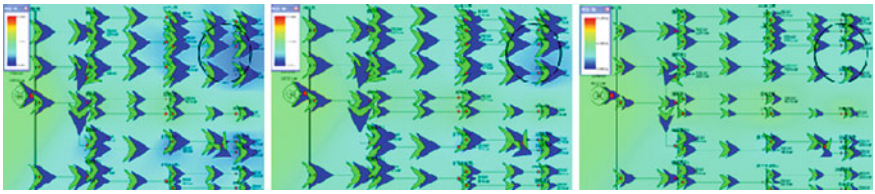


Fig. 118.2 Voltage contour line

Table 118.1 Total line loss

Power state	Active power (kW)	Reactive power(kvar)
State 1	410.66	819.68
State 2	275.08	487.33
State 3	418.64	1117.42

some of the line voltage levels appear some bias, using Power World system voltage contour line drawing tool to draw the system voltage amplitude contour. The voltage contours is shown in third part of Fig. 118.2, which is respectively state 3, state 1, and state 2 from left to right.

The arrows in the figure represent the system's active and reactive power and direction. We can see its main power distribution and mobility of the system from the figure. The deeper color is in the figure shows the lower voltage levels, which as shown in Fig. 118.2 the regional part of the circle. By comparing the voltage contour map in Fig. 118.2 we can see: current voltage is slightly low, and grid running state is as normal. If we do not take necessary adjustment, voltage is likely to fall down, and the color will become deeper, which will impact other work devices in the grid; if the grid is to take measures to adjust its structure, re-distribute load and make necessary reactive power compensation and other measures, voltage supply will be improved to meet the requirements of all operation equipment in the grid.

The total line losses as is shown from Table 118.1.

The current running power lines are larger losses, especially the reactive losses; the results that the voltage of line end is relatively lower. At the same time, the line power loss is also a certain proportion, the power losses of the whole coal are a larger, and these inevitably lead to a poor economic power network.

If the system is in accordance with the method of expert system provided for the power restructuring, the line losses will be greatly reduced, the power losses will be reduced, to reduce the production cost of coal, to bring some economic benefits for coal produce.

If not make any adjustments for power, after a period of time, with the load increased gradually, line losses will gradually increase, mine is no guarantee operating economy. At the same time, excessive reactive power makes the lower end voltage of the line, and part of the devices may not function properly, normal production of coal mines will be brought a big security risk.

118.5 Conclusion

This paper combines fuzzy comprehensive judgment and gray evaluation method to carry on the mine power supply status warning analysis, and using different colors indicate different warning value of the power grid state, to reflect the power actual running status of the visual image, using expert system to process judgment of warning results, which provide the appropriate adjustment method as the operator reference, and taking advantage of the field expert knowledge to improve the accuracy and reliability of warning system deal with the problem. But this method requires a lot of data to support and an advance certain data preparation calculation, also it needs to select the appropriate voltage factor indicator. Otherwise, it will bring about a certain impact for judging the accuracy of the results.

References

1. Sun JP (2010) Review and prospect of technologies of automation and informatization of coal mine. Ind Mine Autom (6):26–30
2. Tao X, Shang Y (2010) Modern mine safe and efficient fully mechanized coal face power technology. Coal J 35(11):1930–1934
3. Shi SC (2008) Transform and upgrade of under coal mine power supply voltage. Coal Technol 27(5):32–34
4. Xia DZ (2004) Power system analysis. China Electric Power Press, Beijing
5. Jin ZG (2005) Designing scheme of integrated automation for transformer substation of mining area underground. Ind Mine Autom 2(4):67–69
6. Zhang JY, Chen H et al (2007) Application and research of embedded technologies in coal mine monitor systems. Aeronaut Comput Tech 37(4):107–109
7. Ding HB, Yang YL, Yueji Li et al (2007) Anti-interference problem of mine monitor systems. J Shanghai Dian ji Univ 10(4):276–278
8. Tang HZ, Peng JC (2003) Research on synthetic and quantificated appraisal index of power quality based on fuzzy theory. Power Syst Technol 27(12):85–88
9. Li QH, Zhou L, Zhang F et al (2008) Comprehensive evaluation of forewarning grade of voltage state and tendency in power systems based on fuzzy theory and analytic hierarchy process. Power Syst Technol 32(4):40–45
10. Zhang Y, Xiong XF, Zhou J (2008) Research on sequence of maintenance of circuit breakers based on grey fuzzy comprehensive evaluation. Power Syst Technol 32(8):21–24
11. Bu G, Zhang Y (2002) Grey fuzzy comprehensive evaluation based on the theory of grey fuzzy relation. Syst Eng Theory Pract 22(4):141–144

Chapter 119
Online Correction of Voltage Stability in the Case of Communication Interruption of Slack Node

Jun-dong Duan and Fan Zhou

Abstract The development and applications of wide-area measurement system (WAMS) provide great potential for online assessment of voltage stability in large-scale power system. By using the WAMS to collect the measured value of local voltage and current vector directly, we can realize the online estimate and instruction of the voltage stability degree of key node. On the basis of the above reasons, this paper further researches the phase failure problem of slack node, puts forward that taking the voltage phase of the node, which has the nearest electrical distance from the slack node, as the reference phase, can also realize online correction of the voltage stability of the key node. The simulations show that this method is suitable for the surveillance of online static voltage stability of node and has certain feasibility.

Keywords Wide-area measurement system · Power system communication interruption · Voltage stability · Electrical distance

119.1 Introduction

Since the reform and opening up, with the high-speed development of national economy, the social demand for power is growing rapidly. The power system is more and more complicated, developing to the big unit, high voltage and long-distance transmission, to form the massive power grid. It is significant to improve

J. Duan (✉) · F. Zhou
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: moon-two@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_119,

economic benefit and protect the environment for rational utilization of natural resources, but it also brings some new safety operation problems to the power system. Especially since the 1970s, a series of power outage at home and abroad, which led to widespread and long time power outages and caused great economic losses and the disorder [1, 2] of social life, are mostly due to the voltage fluctuation. The voltage stability problem attracted universal attention of the electric power industry and academia all over the world.

Currently, most methods are to take some of the voltage stability index [3–6] to indicate the stability degree of the node voltage. As the load impedance modulus can be measured, impedance mode index has been used in the power system voltage stability analysis and online voltage stability surveillance to some degree. The key of this kind of problems is to calculate the equivalent impedance of the key nodes of the power grid side, it is difficult to get the measurement data of the whole grid simultaneously online in SCADA/EMS, and data refresh need a long time, so we can do analysis online only according to several typical operation mode and state of the system, which has some defects. With the development and applications of wide-area measurement technology, it has provided new data acquisition and measurement for the power system. The system realizes vector measurement, and the data refresh faster than SCADA/EMS, so it provides a good technology platform for estimating equivalent impedance of the key nodes of the grid side.

Duan et al. and Guo [7–9] use the measured value of local voltage, current vector collected by WAMS to put forward a kind of index of voltage stability, which can judge the stability degree of the key nodes voltage online. The calculating for local voltage phase of this index uses the voltage phase of slack node as the reference phase, however Liu and Kai [1] analyzes from the perspective of scheduling and gets that the communication interrupt, out of getting correct data are one of the reasons of the further expansion of the accidents. On the basis of the above reasons, this paper further researches the phase failure problem of slack node, puts forward that taking the voltage phase of the node, which has the nearest electrical distance from the slack node, as the reference phase, can also realize online correction of the voltage stability of the key node.

119.2 Online Correction Method of Key Node

Duan et al. [8] analyzed the simplified model of node in Fig. 119.1b, deduced and proved with some examples that using the WAMS can get equivalent impedance modulus of the power grid online:

Denote the measured value of voltage vector of key node at time t as $\dot{V}_t$, the corresponding measured value of load current vector as $\dot{I}_t$; denote the measured value of voltage vector of key node at time $t-1$ as $\dot{V}_{t-1}$, the corresponding measured value of load current vector as $\dot{I}_{t-1}$. If the voltage and current changes of key node are mainly due to the load change of the node, the equivalent impedance of the power grid side is approximated as

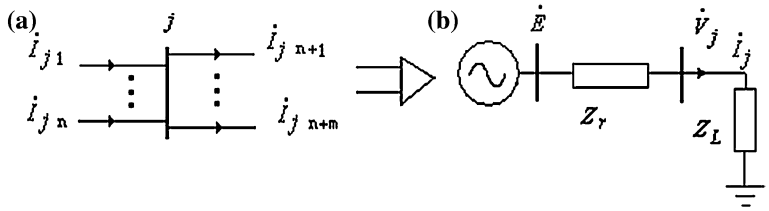


Fig. 119.1 Thevenin equivalent of node *j*

$$Z_{r\cdot t} = -\frac{\dot{V}_t - \dot{V}_{t-1}}{\dot{I}_t - \dot{I}_{t-1}} \tag{119.1}$$

.

The impedance modulus is approximated as

$$|Z_{r\cdot t}| = \frac{|\dot{V}_t - \dot{V}_{t-1}|}{|\dot{I}_t - \dot{I}_{t-1}|} \tag{119.2}$$

Use the ratio of load impedance modulus $|Z_{L\cdot t}|$ and equivalent impedance modulus of the power grid side $|Z_{r\cdot t}|$ as the index of the voltage stability degree of the node [7, 8]. Namely:

$$k = \frac{|Z_{L\cdot t}|}{|Z_{r\cdot t}|} = |\frac{\dot{V}_t}{\dot{I}_t}|/|\frac{\dot{V}_t - \dot{V}_{t-1}}{\dot{I}_t - \dot{I}_{t-1}}| \tag{119.3}$$

The calculating for measured value of voltage vector in formula (119.1), (119.2), (119.3) uses the voltage phase of the slack node as the reference phase.

119.3 Online Correction of Key Node in the Case of Communication Interruption and Phase Failure of Slack Nodes

As the power system become much more dependent on communication network, the reliability of information and communication system will be an important factor that impacts the stability of modern power system. A key communication equipment malfunctions will make the whole power system go dead, and lead to the lost of controllability and monitoring of the power system [10, 11]. Due to the slack node is the reference node of the whole power grid, other nodes collect data from the slack node. If communication breaks down at the slack node, it will cause phase failure of slack node, we are also unable to utilize above method to realize online correction of the voltage stability of the key node.

On the basis of the above reasons, this paper puts forward using voltage phase of auxiliary node to replace voltage phase of slack node as the reference phase,

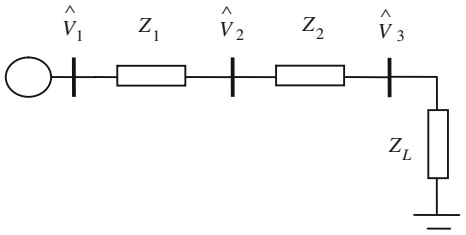


Fig. 119.2 Equivalent three-node system

use voltage phase of auxiliary node to revise voltage phase of local node, the revised voltage phase of local node is described as follows:

Denote the measured value of voltage vector of key node at time t as $\hat{V}_t$, the corresponding measured value of load current vector as $\hat{I}_t$; denote the measured value of voltage vector of key node at time $t-1$ as $\hat{V}_{t-1}$, the corresponding measured value of load current vector as $\hat{I}_{t-1}$; if the voltage and current changes of key node are mainly due to the load change of node, the equivalent impedance of the key nodes of the power grid side is approximated as:

$$Z_{r\cdot t} = -\frac{\hat{V}_t - \hat{V}_{t-1}}{\hat{I}_t - \hat{I}_{t-1}} \tag{119.4}$$

The impedance modulus is approximated as

$$|Z_{r\cdot t}| = \frac{|\hat{V}_t - \hat{V}_{t-1}|}{|\hat{I}_t - \hat{I}_{t-1}|} \tag{119.5}$$

The reference phase of voltage phase in formula (119.4), (119.5) is not unique, the voltage phase angle of key node varies with the voltage phase change of different reference nodes.

We can get an index of the voltage stability degree of key node by the formula (119.3). Namely:

$$k^* = \frac{|Z_{L\cdot t}|}{|Z_{r\cdot t}|} = \left|\frac{\hat{V}_t}{\hat{I}_t}\right| \Big/ \left|\frac{\hat{V}_t - \hat{V}_{t-1}}{\hat{I}_t - \hat{I}_{t-1}}\right| \tag{119.6}$$

Simplify the power system, view from the load node,equate the power system to three-node system, as shown in Fig. 119.2:

In Fig. 119.2, $\hat{V}_1$ is the system equivalent potential (node 1 is the slack node), Z_1 and Z_2 are system equivalent impedance, $\hat{V}_2$ is the voltage of auxiliary node, $\hat{V}_3$ is the load voltage, Z_L is the load impedance.

Denote node 3 as the key node, use the voltage phase of slack node as the reference phase, from formula (119.3) we can get:

$$k = \frac{|Z_L|}{|Z_1 + Z_2|} \tag{119.7}$$

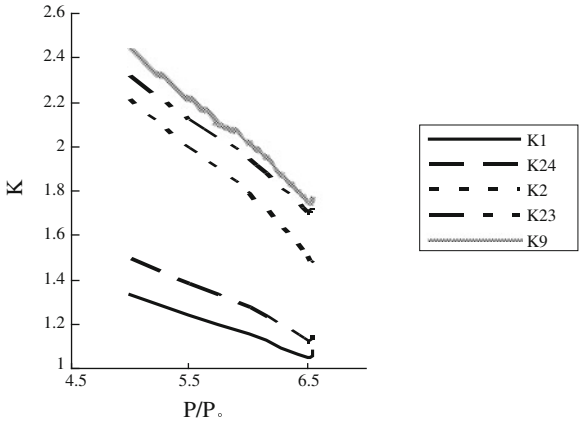


Fig. 119.3 Curves of K

Use the voltage phase of auxiliary node as the reference phase, from formula (119.6) we can get:

$$k^* = \frac{|Z_L|}{|Z_2|} \tag{119.8}$$

From formula (119.7), (119.8) we can get:

$$d = \frac{|Z_1 + Z_2|}{|Z_2|} \tag{119.9}$$

d represents the proximity degree of value k^* and value k, d is getting smaller when value k^* is more closer to value k. When Z_1 is kept to a minimum, value k^* is the closest to value k, the equivalent impedance from the slack node to the key node is the closest to the equivalent impedance from the auxiliary node to the key node at this time, namely, the electrical distance from this auxiliary node to the slack node is the closest. When communication interruption and phase failure of the slack node happen, we can use the voltage phase of the auxiliary node which has the nearest electrical distance from the slack node as the reference phase.

119.4 Examples

Figure 119.3 uses the eight generators 36 nodes of PSASP to simulate, choose node 16 as the research node, change the load of node 16, keep the power factor of node 16 and other nodes constant (power factor is 0.9), the slack node bears the increase of the load, keep the growing of the load until the computing become non-convergent. Figure 119.3 shows that curves between k and power increment:

Figure 119.4 uses the eight generators 36 nodes of PSASP to simulate, choose node 16 as the research node, change the load of node 16, keep the power factor of

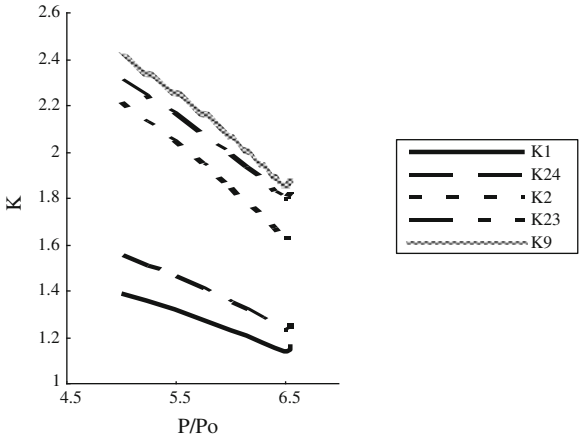


Fig. 119.4 Curves of K

node 16 and other nodes constant (power factor is 0.9), the slack node and G3 bear the increase of the load according to a certain proportion, keep the growing of the load until the computing become non-convergent.

In Figs. 119.3, 119.4, k_1 is obtained using the voltage phase of slack node as the reference phase; k_{24}^* is obtained using the voltage phase of node 24 as the reference phase; k_2^* is obtained using the voltage phase of node 2 as the reference phase; k_{23}^* is obtained using the voltage phase of node 23 as the reference phase; k_9^* is obtained using the voltage phase of node 9 as the reference phase.

The simulation shows the variation curves between value k and power increment when the load increase. The variation results of curves show that the change trends of value k choosing the voltage phase of slack node or other auxiliary nodes as the reference phase respectively are consistent. When the voltage reaches static stability limit and power increment reaches limit, the value k tends to one, which shows the voltage stability degree of the key node approximately in the current condition. When communication interruption and phase failure of the slack node happen,we can use the voltage phase of the auxiliary node which has the nearest electrical distance from the slack node as the reference phase. The value k_{24}^* is the closest to value k_1, so we can use the voltage phase of node 24 which has the nearest electrical distance from the slack node as the reference phase. And we can sort the auxiliary nodes which can replace the slack node by proximity degree between the value k^* and value k, the order from good to bad is: the node 24, the node 2, the node 23, the node 9.

119.5 Conclusion

From the theoretical analysis and simulation results we can see, taking the voltage phase which has the nearest electrical distance from the slack node as the reference phase can settle the phase failure problem of slack node. When losing the voltage

phase of slack node, taking the voltage phase of the auxiliary node, which has the nearest electrical distance from the slack node as the reference phase can also realize online correction of the voltage stability of key node.

References

1. Liu Y, Xie K (2003) Analysis on blackout of interconnected North America Power Grid occurred on August 14 from the viewpoint of power system dispatching. Power Syst Technol 28(8):10–15
2. Hu XH (2003) Rethinking and enlightenment of large scope blackout in interconnected North America power grid[J]. Power Syst Technol 27(9):2–6
3. Taylor CW (1994) Power system voltage stability. McGraw-Hill Inc, NewYork
4. Zhou S, Zhu L, Guo X et al (2004) Power system voltage stability and its control[M]. China Electric Power Press, Beijing
5. Lu B, Wang C, Guo Z et al (2002) Judgement of static voltage stability in power system. J Harbin Institute Technol 34(2):225–227, 231
6. Liu Z (1999) Voltage stability evaluation based on node load impedance analysis. Proc CSEE 19(11):64–68
7. Duan J, Sun Y, Yin X (2009) Voltage stability's online prediction using WAMS. High Volt Eng 35(7):1748–1752
8. Duan J, Guo Z (2005) Research of on-line assessment of voltage stability. Proc CSEE 25(25):199–202
9. Duan J, Guo Z (2006) A new method for on-line determination of the capability curves of voltage stability. Proc CSEE 26(04):113–118
10. Xing NZ, Yan HF (2007) Research on the reliability of electric power telecommunication system[J]. Telecomm Electr Power Sys 28(176):26–30, 38
11. Bo SU, Qi MM (2010) A method of reliability evaluation for power communication network[J]. Telecomm Electr Power Syst 31(215):5–8, 14

Chapter 120
Transient Analysis and Time Selection of Capacitor Switched Device of Reactive Power Compensation

Weifeng Yin

Abstract Based on the transient analysis of the compensation capacitor switching, studying the effect of the time of capacitor's input and removal for grid stability. And based on the stationary solution unchanged, focusing on the general solution of the circuit-order homogeneous linear differential equations and analysis of the effect for the switching current shocks. The results show that the phenomenon of over-current can be avoided. The results of ATP simulation show that, the proposed method can effectively reduce the effect of capacitor switching for grid, improve power quality, and extend the life of the capacitor.

Keywords Capacitor switching · Grid impact · Moment choice

120.1 Introduction

In the power system, Reactive power is the power transfer strength between the load and power supply [1]. For the sinusoidal voltage $u(t) = \sqrt{2}\text{Usin}\omega t$ and sinusoidal current $i(t) = \sqrt{2}\text{Isin}(\omega t + \varphi)$, Reactive power actually represents the phase difference between current and voltage, which Caused the power work on the load and then load deliver it to power.

Compensation of Shunt Capacitor is a common means of reactive power compensation. However, due to the changing of power grid equipment and parameters, we must cut and put in the capacitor frequently according to the actual power demand. So, the research of the transient process and the time choice of the capacitor's input and removal is very significant to reduce the impact on power, extend the life of circuit breakers and capacitors and other key aspects.

W. Yin (✉)
China Coal Technologies Group Corporation,
Beijing 100013, China
e-mail: jphstudy@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_120,

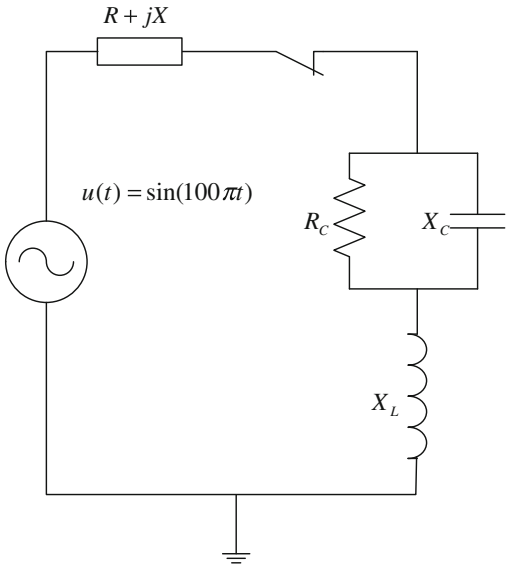


Fig. 120.1 Equivalent circuit diagram of Capacitor

120.2 Transient Process Analysis of Capacitor's Input and Remove

Take three-phase symmetry for an example, then the middle point voltage is zero, the power grid can be predigested to a single-line diagram, shown in Fig. 120.1.

(1) Capacitor XC. This is a major part of the capacitor, it will inject reactive power to the nodes of system when it connect into grid, the size of reactive power is defined by the value of capacitor and the node-voltage.
(2) Capacitor discharge resistor RC. The RC is used to cut off the capacitor and releases the residual voltage in time. the discharge time constant is determined by the resistance values and capacitor.
(3) Current-limiting inductor X_L. It is used to prevent the rush current when the capacitors connected. The value of current-limiting inductor should be designed by the value of short-term maximum allowable current.

120.2.1 Analysis of the Transient Process When the Connection of Capacitor

When the capacitor is connected into grid after removal, if the difference between the grid voltage and capacitor voltage is larger, it will have a great current shocks

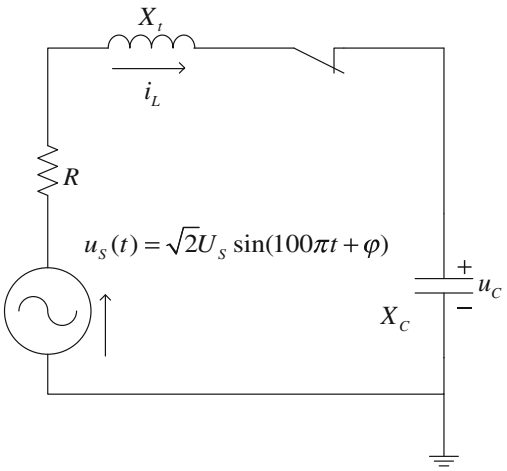


Fig. 120.2 Simplified circuit of capacitor

due to the voltage on the capacitor cannot change suddenly, which is likely to damage thyristor, or take frequency effect to the power grid [2].

According to Fig. 120.1, the discharge resistance RC is more larger than the capacitor resistance, so RC can be ignored. Then combining the limiting inductor with the transmission line impedance, we can simplify the circuit, shown in Fig. 120.2.

By analysis the circuit of Fig. 120.2, and combination of the knowledge of second-order linear differential equation, we can obtain the following equation:

$$\frac{\mathrm{d}^2 u_C}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}u_C}{\mathrm{d}t} + \frac{1}{LC}u_C = \frac{1}{LC}u_S \tag{120.1}$$

The solution of this second-order linear differential equations have two parts, the first is the general solution of the corresponding second-order homogeneous linear differential solution, the second is the particular solution that is the steady-state solution for the circuit. Using the knowledge of phaser arithmetic, we can get steady-state voltage of the ends of capacitor X_L.

$$\dot{U}_C = \dot{U}_S \frac{jX_C}{R + j(X_t + X_C)} \tag{120.2}$$

Zero-input response of the circuit should have a general solution in the form of index [3],

$$u_{C0}(t) = Ae^{\mathrm{pt}} \tag{120.3}$$

$$p_{1,2} = \begin{cases} -\alpha \pm \sqrt{\alpha^2 - \omega_0^2} (\alpha \geq \omega_0 > 0) \\ -\alpha \sqrt{\omega_0^2 - \alpha^2} (\omega_0 \geq \alpha > 0) \end{cases} \left(\alpha = \frac{R}{2L}, \omega_0 = \frac{1}{\sqrt{LC}} \right) \tag{120.4}$$

The overlay of the two results is the full response of the circuit.

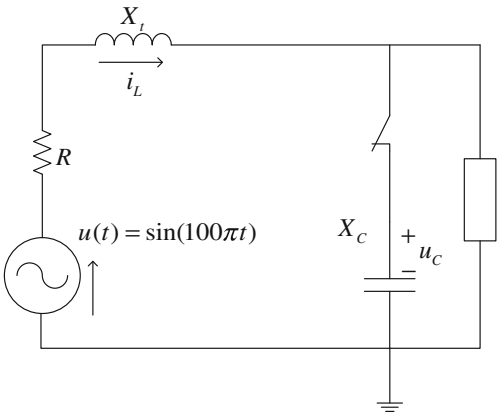


Fig. 120.3 Simplified circuit of the state of capacitor's resection

120.2.2 Analysis of the Transient Process When the Remove of Capacitor

The transient process of removing is much simpler than the input. As shown in Fig. 120.3, because the inductor current in transmission line is not sudden changed, based on the node's current law, current will instantly transfer to the load when the capacitor is cut down, then the whole system come into the transition process and soon it will be stable [4]. So if it is expected to minimize the system impact caused by removal of capacitors, the capacitor current must put to minimize or zero, in this case the current transferred to the load will be very small or zero, the transition time is relatively stable and short, the capacitors can be removed smoothly and achieve steady-state.

120.2.3 Choosing the Smallest Impact Time of Capacitor Switching

In fact, taking the impact of capacitor switching, the particular solution of steady state cannot be considered. Because no matter what kind of transient response, the particular solution of steady state will not be changed, so it is important to reduce the general solution [5]. In other words, if the general solution can be minimized as long as possible, the current shock can be utmost reduced.

From the discussion, in order to reduce the current impact caused by capacitor switching, the zero-moment state of the above equation should be adjusted, that is to say the undetermined coefficient A of the general solution must be minimal. Adjust the zero point is to adjust the moment that put into capacitor. It is visible that as long as we make $\varphi + \theta = k\pi$ we can avoid the large current impact. Where φ expresses corresponding power phase when capacitor switching.

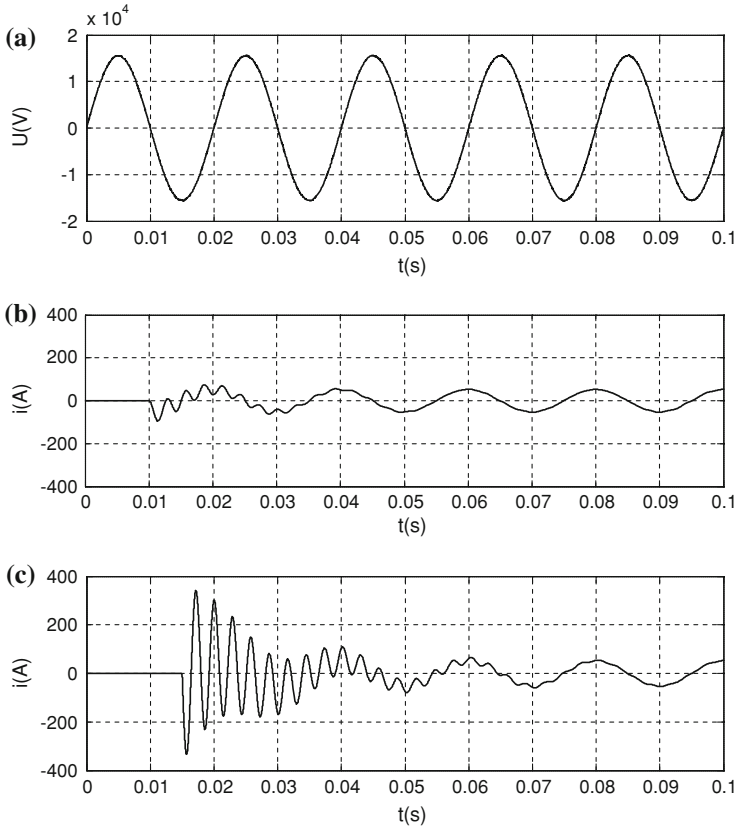


Fig. 120.4 The current impact of different input moment of capacitor. **a** Supply voltage. **b** Put in When the voltage is zero. **c** Put in when the voltage is not zero

120.3 Experimental Verification

ATP is a simulation software that used for power system electromagnetic transient analysis. It can calculates the impact of the power system that caused by system variables based on time. Usually the trapezoidal rule of integral is used to solve the differential equations of the system components in the time domain, the non-zero initial state can be automatically determined by the steady-state vector solution, or by the user to input for certain components. In the experiment, we selected a group of actual operation 10 kV system parameter, $X_t = 6.28\Omega$, $R = 3\Omega$, $X_C = 302.5\Omega$, We can obtain the following waves by simulation (Fig. 120.4).

It is Visible that the current impact will be three times steady-state current when the capacitor not put in at zero voltage moment. Although switch the capacitor on zero voltage will take decay process (due to the initial value of inductor current is zero), but almost no current impact.

120.4 Conclusion

Based on the analysis of the transient process with capacitor input and removal, this paper discusses how to reduce power shocks by adjusting the zero-moment state. Through the ATP simulation of 10 kVpower grid, the results show that the adjustment of the switching moment can effectively prevent over-voltage and over-current.

(1) When capacitor with zero voltage invested into grid, if the access point's voltage is equal to zero-voltage at the steady state, the impact to the power grid may be minimal.
(2) When the capacitor is removed from the grid, if the current in capacitor is zero, in this case, it can effectively shorten the transitionprocess time, and make the minimum impact to the grid.

References

1. Xie X, Jiang Q (2006) Flexible AC transmission systems: principles and applications
2. Zhou Y, Zhao J (2000) Software for automatic control of voltage and reactive power in substation and its applications. Autom Electr Power Syst 9:56–59
3. Yu Z, Li Y, Zhan H (2007) Linear algebra and analytic geometry. Tsinghua University Press, Beijing
4. Wu G-X, Jiang H, Zhang P-M (2009) Development of intelligent compound switch for reactive power compensation capacitor switching. Electrotech Electric 28–52
5. Kang M, Zhang J, Chen J (2004) Applications of static VAR compensators in distribution power systems. Int Electr Powr China 6:37–41

Chapter 121
Non-Parametric Statistical Inference for Interval-Valued Steady-State Availability of Electric Power Equipment Based on Markov Renewal Process

Xianjun Qi, Xinhong Shi and Jiayi Shi

Abstract The interval-valued steady-state availability (SSA) of electric power equipment is necessary for the imprecise reliability assessment of power system. This paper presented a method of non-parametric statistical inference to get the interval-valued SSA. First the analytical expression of SSA based on Markov renewal process was proposed. Then the point estimations of each equipment's SSA were obtained through non-parametric estimation. And the interval-valued SSA of this type of equipment was obtained by the Bias-corrected and Accelerated Bootstrap confidence interval estimation. Finally, a case study illustrated the effectiveness of the above method. The interval-valued SSA can cover the true availability well under a given confidence level. So it lays a foundation for the imprecise reliability assessment of power system.

Keywords Interval-valued steady-state availability · Non-parametric statistical inference · Markov renewal process · Bootstrap · Imprecise reliability

121.1 Introduction

Power system reliability is a measure of the ability of a power system to deliver electricity to all points of utilization at acceptable standards and in the amount desired [1]. The input data in the reliability assessment mainly includes the probability information of random variables which reflect the equipment reliability

X. Qi (✉) · X. Shi · J. Shi
School of Electrical Engineering and Automation,
Hefei University of Technology, Hefei, China
e-mail: qxj_216@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_121,

behavior. Its accuracy is important for the credibility of the assessment. According to the completeness of random variable probability information, reliability assessment can be divided into three levels: precise reliability assessment, quasi-precise reliability assessment and imprecise reliability assessment.

Since only part of the random variables' probability information can be obtained, it is difficult to achieve precise reliability assessment. At present, quasi-precise reliability assessment is widely used in power system reliability assessment, where the probability distribution and correlation (independent or perfect correlation) of random variables are assumed first, and the approximate probability information is achieved through parameter estimation. However, some problems exist in the quasi-precise reliability assessment:

Instead of using the real probability information, we use the assumed probability distribution and estimates of parameters, which may cause deviation of the assessment. Thus we may wonder the credibility of the assessment results.

The different parameter estimation results from kinds of statistical inference methods and the errors in statistical data, etc. will cause great uncertainty in the results of reliability assessment.

Some new components such as new energy generators, flexible AC transmission devices and so on, may be lack of or even without failure statistical data. In this case, how to carry out parameter estimation effectively?

The purpose of imprecise reliability assessment is to analyze the reliability of the system only with available information or minimum of assumptions, without additional assumptions such as the assumption of complete probability information. Imprecise reliability only appeared and developed in the last ten years, and some important results have been achieved in this field [2–5]. But it has not yet been applied and researched in power system reliability assessment. In order to execute the imprecise reliability assessment of power system, firstly we need to build the analysis and estimation model of availability based on incomplete probability information, and this paper focuses on this. In the paper, the analytical expression of SSA is derived by using Markov renewal process theory and its interval value is estimated based on Bootstrap confidence interval estimation.

121.2 Definition of Availability and States Merging

121.2.1 Definition of Availability

Availability can be defined by qualitative method and quantitative method. This paper makes a quantitative analysis and estimation on the availability, and the quantitative expression of availability is as follows.

Suppose an equipment has $N+1$ states, in other words, the state space is $\mathbf{S} = \{0, 1, \ldots, N\}$, which can be divided into the available state set $\mathbf{S_A} = \{0, 1, \ldots, K\}$ and the unavailable state set $\mathbf{S_U} = \{K+1, K+2, \ldots, N\}$.

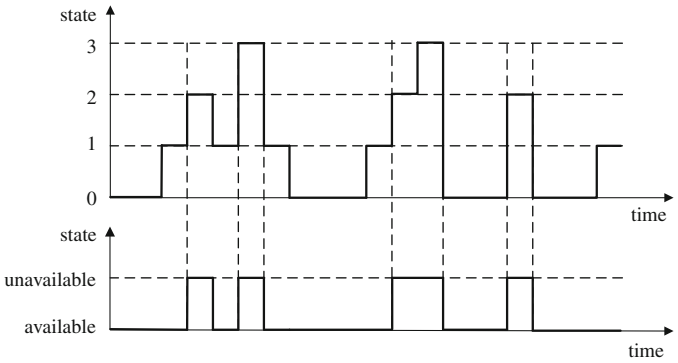


Fig. 121.1 Illustration of states merging

$P_i(0)(i = 1, \ldots, N)$ is the initial distribution, and let $s(t)$ denote the equipment's state at a specified time t.

Instantaneous availability. The instantaneous availability at time t is defined as the probability that equipment is in the available state at time t.

$$A(t) = P\{s(t) \in \mathbf{S_A}\} = \sum_{j=1}^{K} P_j(t). \tag{121.1}$$

Steady-state availability. If the limit of $A(t)$ exists, then A is called the steady-state availability (SSA).

$$A = \lim_{t \to \infty} A(t). \tag{121.2}$$

121.2.2 States Merging

There may be many states in the equipment's running process, while we only need to consider the available and unavailable state sets during the analysis and calculation of availability. So it is necessary to merge many states into the two states. The process of states merging is illustrated with an example: given one type of equipment have four running states. Its available state set and unavailable state set are separately $\mathbf{S_A} = \{0, 1\}$ and $\mathbf{S_U} = \{2, 3\}$. Then the process of states merging is showed in Fig. 121.1.

121.3 SSA based on Markov Renewal Process

The sojourn time in each state in a continuous-time Markov process is exponentially distributed. If the distribution of sojourn time is generalized, we will get the Markov renewal process, also called as Semi-Markov process. In the availability

analysis of equipment based on Markov renewal process, the following assumptions are made.

The sojourn time in available state, denoted as TA, has a general continuous distribution function $F(t)$; and the sojourn time in unavailable state, denoted as TU, obeys a general continuous distribution function $G(t)$.

TA and TU are mutually independent.

According to the theory of Markov renewal process, the SSA can be obtained [6].

$$A = E(\mathrm{TA})/(E(\mathrm{TA}) + E(\mathrm{TU})). \tag{121.3}$$

Note that if a repairable equipment with two states is considered, the above TA and TU are separately TTF and TTR, and their mathematical expectations $E(\mathrm{TA})$and $E(\mathrm{TU})$ are separately mean time to failure (MTTF) and mean time to repair (MTTR).

121.4 Non-Parametric Statistical Inferences for SSA

Availability is essentially a probability and interval probability is one kind of imprecise probability, so interval probability is adopted to obtain the interval-valued SSA in the paper.

Suppose we have recorded the running states of N electric power equipment of the same type, and have obtained the sample sequences of two-state stochastic process by the states merging method. The stochastic process sample sequence of the ith equipment is denoted as $\{\mathrm{TA}_{ik}, \mathrm{TU}_{ik}\}(i = 1, \ldots, N;\ k = 1, \ldots, M_i)$. We can get the interval-valued SSA through the following steps: (1) Based on the sample sequences of each equipment, the point estimates $\hat{A}_i$ of each equipment are obtained; (2) Considering $\hat{A}_i$ as the random samples of SSA of the same type equipment, the interval-valued SSA of this type equipment is obtained through re-sampling technique and Bootstrap confidence interval estimation.

121.4.1 Point Estimation of SSA

The point estimate of the SSA for the *ith* equipment is

$$\begin{aligned}\hat{A}_i &= \frac{1}{M_i}\sum_{k=1}^{M_i}\mathrm{TA}_{ik} \Bigg/ \left(\frac{1}{M_i}\sum_{k=1}^{M_i}\mathrm{TA}_{ik} + \frac{1}{M_i}\sum_{k=1}^{M_i}\mathrm{TU}_{ik}\right) \\ &= \sum_{k=1}^{M_i}\mathrm{TA}_{ik} \Bigg/ \left(\sum_{k=1}^{M_i}\mathrm{TA}_{ik} + \sum_{k=1}^{M_i}\mathrm{TU}_{ik}\right).\end{aligned} \tag{121.4}$$

The point estimate of SSA for this type of equipment is

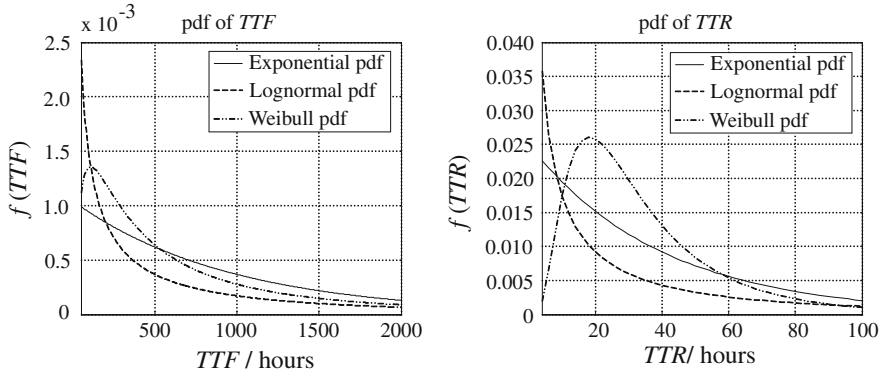


Fig. 121.2 Probability density functions of TTF and TTR

$$\hat{A}=\frac{1}{N}\sum_{i=1}^{N}\hat{A}_i. \tag{121.5}$$

121.4.2 Bootstrap Confidence Interval Estimation for SSA

Bootstrap is an effective method to estimate the variance and distribution of statistics, and can be used to construct confidence intervals. The construction methods of the Bootstrap confidence interval include Basic Bootstrap interval, Bootstrap-t interval, Percentile Bootstrap interval, BC*a* interval (which stands for bias-corrected and accelerated interval), Studentized Bootstrap. The BC*a* interval has both advantages of covering property and range-preserving. So it is adopted in the paper, and the concrete solving steps can be found in [7].

121.5 Case Study

For the convenience of illustrating the above method, a repairable generator unit with two states is considered. Suppose MTTF and MTTR of this type of generator unit are 960 and 40 h separately. So the true SSA is 0.96. Six simulative generator units 'sample sequences with 6 years' simulation time are formed and the following three different cases are considered. The probability density functions of TTF and TTR are shown in Fig. 121.2.

Case 1. TTF and TTR are both exponentially distributed.

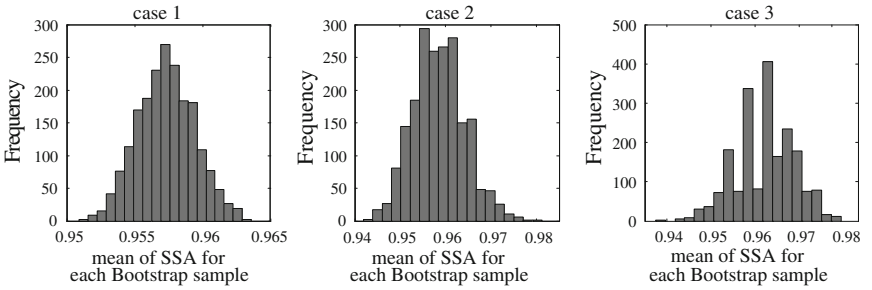


Fig. 121.3 Frequency histogram of Bootstrap sampling

Fig. 121.4 Bootstrap confidence intervals of SSA

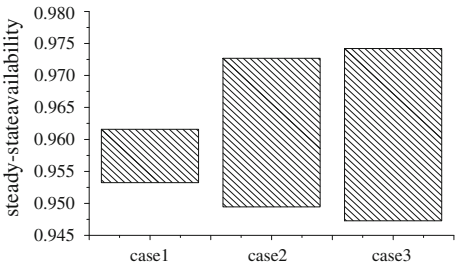
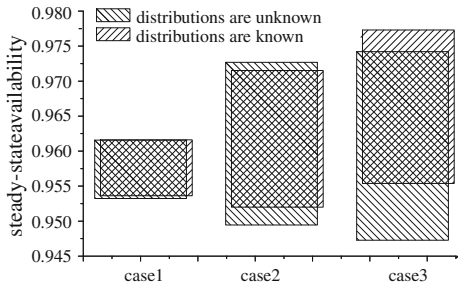


Fig. 121.5 Effects of probability information completeness on interval-valued SSA



Case 2. TTF and TTR are both lognormally distributed.
Case 3. TTF and TTR are both weibully distributed.

121.5.1 Bootstrap Confidence Intervals of SSA

After 2000 Bootstrap iterations, the frequency histogram of Bootstrap sampling is shown in Fig. 121.3.

The 95% Bootstrap confidence intervals of SSA for this type of generator unit under the three cases are shown in Fig. 121.4.

121.5.2 Effects of Probability Information Completeness on Interval-Valued SSA

If the distribution type and parameters of TTF and TTR are all unknown, point estimates of MTTF, MTTR and A_i of each generator unit can be obtained through non-parametric statistical method. Then the interval-valued SSA of this type equipment is got through BC*a*. The length of interval value may be larger for lack of probability information. While if all the distributions are clear and only the parameters are unknown, point estimates of MTTF, MTTR and A_i of each generator unit can be obtained through parametric statistical method. And the length of interval-valued SSA through BC*a* will be reduced for more probability information. The effects of probability information completeness on SSA interval value are shown in Fig. 121.5. And moreover, if the distribution type and parameters are all known, then the interval-valued SSA will evolve into a point value, that is, precise probability.

121.6 Conclusion

SSA is essentially a probability. Since the incompleteness of random variables probability information and the randomness of sampling, point estimates of availability are hard to accurately describe the real availability, interval probability is used to describe the impreciseness. The analytical expression of SSA and non-parametric estimation for statistical function of distribution function based on Markov renewal process make the proposed method no longer limited to the assumption of a specified probability distribution, and therefore can be applied more broadly. The BC*a* Bootstrap confidence interval estimation is characterized by the properties of coverage performance and range-preserving, so the interval-valued SSA can cover the true availability well under a given confidence level. The interval-valued SSA lays a foundation for power system imprecise reliability assessment.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant 51007017, in part by Special Funds for Doctors in HFUT under Grant GDBJ2009-057, and Science Research and Development Funds of HFUT under Grant 2009HGXJ0011. We herewith acknowledge with thanks.

References

1. Saman N, Singh C (2007) Reliability assessment of composite power system using genetic algorithms. Computational Intell Reliab Eng (SCI) 39:237–286
2. Aughenbaugh JM, Herrmann JW (2009) Information management for estimating system reliability using imprecise probabilities and precise Bayesian updating. Int J Reliab Saf 58(1): 35–56

3. Choi HH, Seo J (2009) Safety assessment using imprecise reliability for corrosion-damaged structures. Comput-Aided Civ Infrastructure Eng 24(4):293–301
4. Wagenknecht M, Gocht U (2007) On the reliability of multistate systems with imprecise probabilities. In: 5th conference of the European-society-for-fuzzy-logic-and-technology, Ostrava CZECH REPUBLIC. New dimensions in fuzzy logic and related technologies, vol I, Proceedings pp 473–476
5. Utkin LV, Coolen FPA (2007) Imprecise reliability: an introductory overview. In: Computational Intelligence in Reliability Engineering (SCI) 40, Springer-Verlag, Berlin, Heidelberg, pp 261–306
6. Birolini A (2010) Reliability engineering: theory and practice. Springer, Berlin, pp 162–182
7. Rizzo ML (2007) Statistical computing with R. Chapman & Hall/CRC, London, pp 201–206

Chapter 122
Simulation Study of Fault Location Based on EMD Algorithm for Small Current to Ground System

Xiaowei Wang, Lu Shan, Shu Tian, Tao Zhang and Shuai Wang

Abstract In this paper, it established the single-phase to ground fault model of small current to ground system, and simulated the model, according to decomposed by EMD algorithm for transient zero-mode power signal of every detection point along the line installed. It utilized the high-frequency IMF components got to subtract backward first-order, according to analysis the differential waveforms of every detection point, it has the differential waveforms similarity degree of two detection point located the same side of fault point is higher, while the differential waveforms similarity degree of two detection point located the two sides of fault point is lower. In addition, if the fault point was center, the differential waveform symmetry of two sides of fault point is higher. Besides, it has shown that the characteristic information of first-order differential waveform was not varied by the transition resistance's change. So it utilized the characteristic information of differential waveform to provide a new idea for fault location of small current to ground system.

Keywords Simulation · Empirical mode decomposition · Waveform similarity · Fault location

X. Wang (✉) · S. Tian · T. Zhang · S. Wang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan Province, China
e-mail: proceedings@126.com

L. Shan
Department of Mechanic and Electrical Engineering, Jiaozuo University, Jiaozuo 454000, Henan Province, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_122,

122.1 Introduction

In recent years, the fault location and fault selection of small current to ground system has not been solved. In addition to their own complexity of small current to ground system, and its modeling and simulation's right or wrong constrained the successfully solved these problems. Therefore, if it these problems wanted to be solved better, it should be established accurately the model for the actual system, and simulated to get the voltage, current and power component of every fault style, and then it is filtered out to the transient components of fault occurred as the basis for follow-up analysis.

Based on the accurate model and simulate for single-phase to ground fault of small current to ground system, it is decomposed the transient zero-mode power signal measured by every detection point installed along the line with Empirical Mode Decomposition (EMD) algorithm, and utilized the high-frequency Intrinsic Mode Function (IMF) components to subtract backward first-order. Finally, according to analyze the first-order differential waveforms of every detection point, it could get clear that the first-order differential waveforms's similarity of the first IMF component of two detection point located the same sides was higher, while the two points located the two sides of fault point was low. Therefore, it could utilize the waveform similarity to achieve accurately the fault location by appropriate algorithm.

122.2 Transient Zero-Mode Power Features

It installed the detection devices along the line named A, B, C and D, supposed the fault occurred in BC section, and its equivalent zero-mode network was shown in Fig. 122.1.

Because the distance of AB section is short, the capacitive current to ground relative to the zero-mode capacitive current sum of nonfault lines was very small, it was negligible, therefore, the zero-mode power of two ends for AB section were approximately equal, that was p_{0A} equaled to p_{0B}, and their waveforms were similar [1]; also, the zero-mode power of two ends for CD section were approximately equal, that was p_{0C} equaled to p_{0D}, it was shown in Fig. 122.2. When the fault occurred, it could produce a virtual power at the fault point, the zero-mode current from the fault point flow from the fault point to upstream line, that was toward the bus; and the other part of zero-current flow from the fault point to downstream line, that was deviated from the bus. Therefore, it ignored the impact of public zero-mode voltage, the initial polarity of transient zero-mode power detected from B and C were reverse, and the two waveforms were very different, so it was not similar, it was shown in Fig. 122.2b and c. From the point of view of frequency, the amplitude of upstream signal was large, and the transient main resonant frequency was lower, while the amplitude of downstream signal was

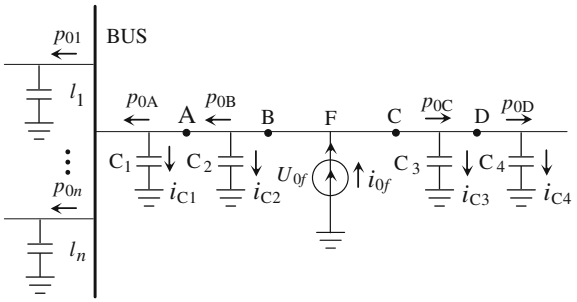


Fig. 122.1 Equivalent circuit of zero-mode network U_{0f} was virtual zero-mode voltage source; C_1, C_2, C_3 and C_4 were capacitance of AB, BF, FC and CD

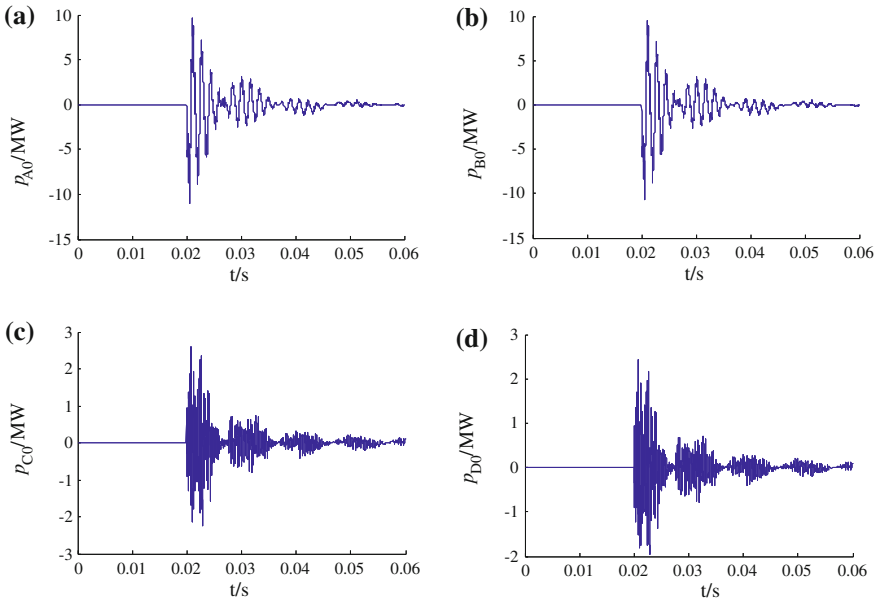


Fig. 122.2 Zero-mode power of every point (**a**) A point (**b**) B point (**c**) C point (**d**) D point

small, and its frequency were high, so the transient zero-mode power signal located at the upstream and downstream of fault point has different frequency components, and their waveforms were quite different [2].

122.3 Basic EMD Theory

IMF calculated by EMD should meet the following two conditions: (1) the extreme point number of signal data and the zero point number were equaled or differed to one at most. (2) The mean value of two envelope formed by connecting the local

maximum and local minimum was equaled to zero at any point. The decomposition process of EMD was extracted IMF components' process [3–6]. It met the IMF condition could make the instantaneous frequency got by Hilbert transform with the actual physical meaning, and the specific calculation method of EMD was shown as follow:

First, according to the maxima and minima of signal $s(t)$ to calculate the average value of upper envelope $l_1(t)$ and the lower envelope $l_2(t)$.

$$m_{11} = \frac{1}{2}[l_1(t) + l_2(t)] \tag{123.1}$$

And then, calculate the difference of $s(t)$ and m_{11}, it was:

$$d_{11} = s(t) - m_{11} \tag{123.2}$$

If d_{11} was not IMF, then d_{11} was as a new $s(t)$, to repeatedly calculate the (123.2) k times, that was:

$$d_{1k} = d_{1(k-1)} - m_{1k} \tag{123.3}$$

In (123.3), d_{1k} was the k times selection data; $d_{1(k-1)}$ was the $k-1$ times selection data; m_{1k} was the average value of upper envelope and lower envelope; utilized the S_D to determine the results whether or not the IMF components are met.

$$S_D = \sum_{t=0}^{r} \left| \frac{\left| d_{1(k-1)}(t) - d_{1(k)}(t) \right|^2}{d^2_{1(k-1)}(t)} \right| \tag{123.4}$$

The value of S_D was equaled between 0.2 and 0.3. When d_{1k} met the S_D value, then

$$c_1 = d_{1k} \tag{123.5}$$

c_1 was a IMF signal, then the following calculation was

$$r = s(t) - c_1 \tag{123.6}$$

Put r as a new $s(t)$, repeat upper calculation process, and then got the second IMF c_2, the third IMF c_3, $\cdots$, up to $r(t)$ was basic monotonic trend or $|r|$ was small, it could be regarded as measurement error, the final were:

$$s(t) = \sum_{i=1}^{n} c_i + r \tag{123.7}$$

Equation (123.7) showed that the completeness of decomposition of EMD.

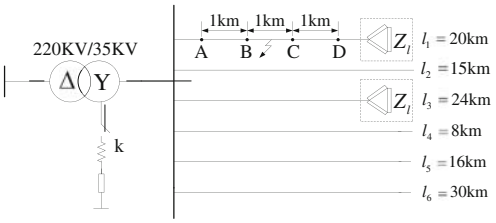


Fig. 122.3 Simulation model of small current to ground system

122.4 ATP Simulation and Analysis

The line adopted distributed parameter model, and the parameter as follows:

Positive sequence impedance: $Z_1 = 0.17 + j0.38$ Ω/km; positive sequence admittance to the ground: $b_1 = j3.045$ us/km; zero sequence impedance: $Z_0 = 0.23 + j1.72$ Ω/km; zero sequence admittance to the ground: $b_0 = j1.884$ us/km; $l_1 = 20$ km; $l_2 = 15$ km; $l_3 = 24$ km; $l_4 = 8$ km; $l_5 = 16$ km; $l_6 = 30$ km.

Transformer parameters as follows:

Δ/Y connection; primary voltage: 220 kV; secondary voltage: 35 kV; leakage impedance of primary voltage: $Z_{1\delta} = 0.4 + j12.2$ Ω; leakage impedance of secondary voltage: $Z_{2\delta} = 0.006 + j0.183$ Ω; steady-state magnetizing current: 0.672A; main flux: 202.2 Wb; magnetizing impedance: 400 kΩ; rated capacity of transformer: $S_N = 40000$ KVA; no-load loss: 35.63 KW.

It used the impedance to replace the load, and the value was: Zl = 400 + j20 Ω.

Sampling frequency $f = 1$ MHz. When the switch K was open and close, it is indicated, respectively, the ungrounded system and arc suppression coil grounding system. The arc suppression coil is over compensation state, and compensation degree was 8%, if the series impedance of coil was the 10% of its inductance value, so according to calculating, the value of series resistance was equaled to 121.35 Ω. Utilized the transient information, so it has proved that whether or not the arc-suppression coil, the results was little influence. So it simulated the single-phase to ground fault of resonant neutral grounding system for example in this paper, the fault occurred in BC section, as shown in Fig. 122.3.

When the initial condition is 0° and 1000 Ω, the transient zero-mode power signal of every detection point was decomposed by EMD algorithm to get the IMF1 components in 1/4 cycle, it was shown in Fig. 122.4. According to the EMD decomposition theory, the first IMF components were the most high-frequency components, it contained a wealth of fault information, so the study was based on the IMF1 components to analyze in the paper.

It can be seen from Fig. 122.4, although the most high-frequency IMF1 components were not ruled and changes greatly in 1/4 cycle, but generally speaking, the IMF1 amplitude located at fault occurred was largest. Therefore, it filtered effectively the IMF1 components of transient zero-mode power after fault occurred to ensure the rationality of calculation, but it could also reduce the data operations

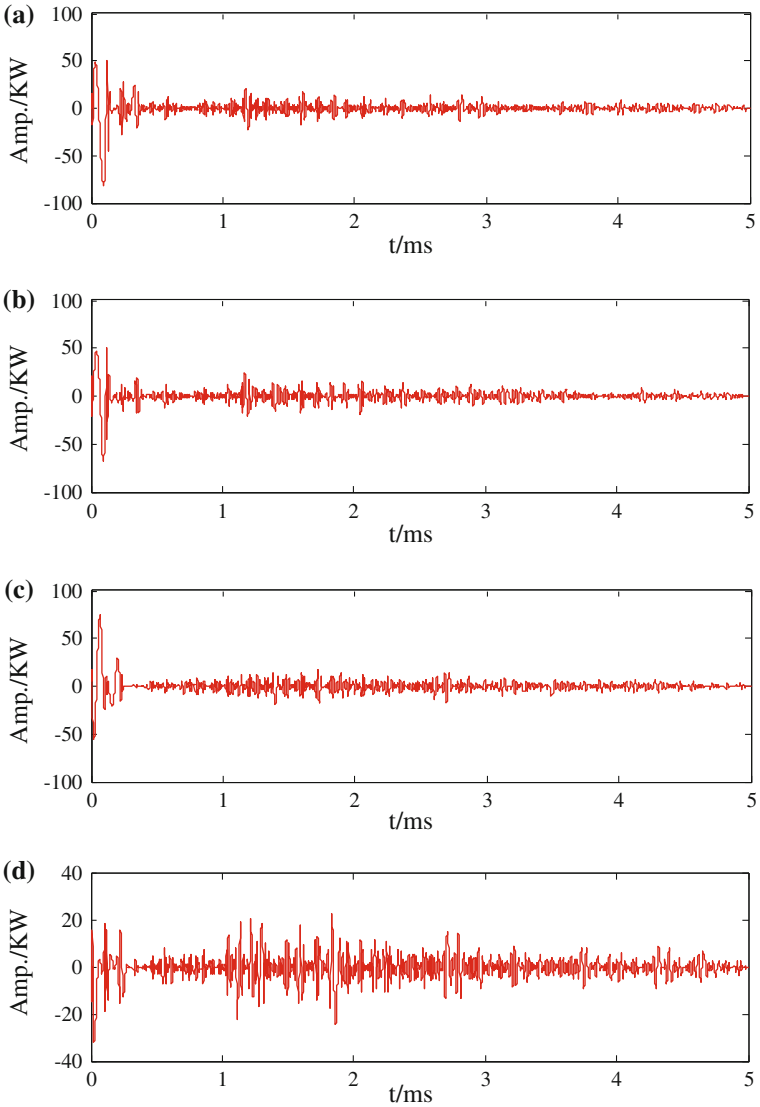


Fig. 122.4 IMF1 components of every detection point **a** IMF1 of A point **b** IMF1 of B point **c** IMF1 of C point **d** IMF1 of D point

size, so it was useful to achieve the real-time fault location for the power system. Therefore, it selected 200 points data to achieve the first-order differential calculation after the fault occurred in the paper, through according to analyze the similarity of first-order differential waveforms for every detection point to provide a new ideas for the fault location of small current to ground system.

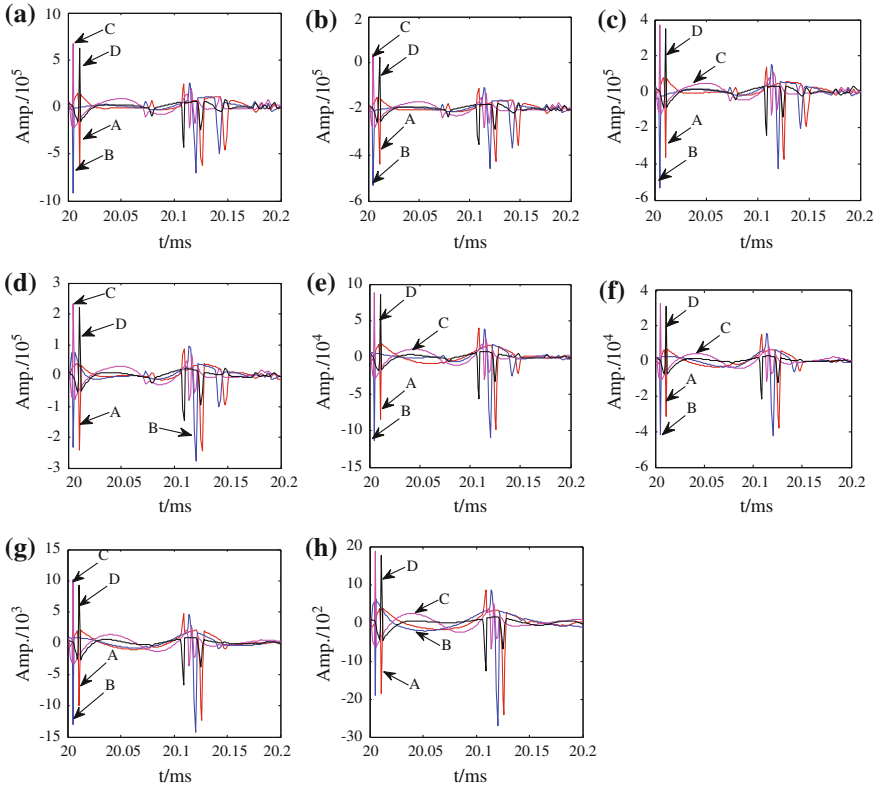


Fig. 122.5 IMF1's first-order differential waveforms of different transition resistance **a** 5 Ω, **b** 20 Ω, **c**100 Ω, **d** 200 Ω, **e** 500 Ω, **f** 1000 Ω **g** 2000 Ω, **h** 5000 Ω

Due to space limitations, it only gave the IMF1 components' first-order differential waveforms of the 0°and different transition resistance, it were shown in Fig. 122.5.

As can be seen from Fig. 122.5, when the transition resistance was changed from 5 to 5000 Ω, there was a clear difference between before the fault point and after the fault point for the IMF1 components's first-order differential waveforms of transient zero-mode power. When the fault occurred in 0.02 s, the initial polarity of IMF1 components's first-order differential waveforms located at A and B of the upstream fault point were negative, and the amplitude were larger; But initial polarity of IMF1 components's first-order differential waveforms located at C and D of the downstream fault point were positive, and the amplitude were also larger. As can also be seen from Fig. 122.5, the initial polarity of A and D for the IMF1 components's first-order differential waveforms were different, but generally speaking, the symmetry of two waveforms is higher, the symmetry of B and C for IMF1 components's first-order differential waveforms were also higher. In addition, as can also be seen from Fig. 122.5, with the transition resistance's change,

the corresponding IMF1 components's first-order differential waveforms were not changed, it only the amplitude was decreased, it also has relation with bigger transition resistance. Therefore, the characteristic information reflected from the Fig. 122.5 could also provide novel ideas for fault location of small current to ground system's single-phase to ground fault.

122.5 Conclusion

Because the fault location problem is difficult for small current to ground system, it established the distribution network model of 220/35 kV to simulate the single-phase to ground fault occurred. It utilized the detection devices installed along the lines to detect the transient zero-mode power signal after the fault occurrence, and used EMD algorithm to decompose the transient zero-mode power signal of every detection device, it based on the high-frequency IMF1 components obtained, and according to acquire the first-order differential waveforms of IMF1 for every detection point to observe, it found that the initial polarity of the IMF1 components first-order differential waveforms for A and B which located at upstream lines were negative, and the similarity was higher, but the initial polarity of the IMF1 components first-order differential waveforms for C and D which located at downstream lines were positive, and the similarity was also higher, while the similarity of waveforms got by two sides of fault point was lower. Besides, with the transition resistance's increase, these feature information was not changed much, therefore, it utilized these feature information to build excellent fault location algorithm, from analysis above, it was not affected with transition resistance. According to modeling and simulation in the paper, it provided a novel thinking for fault location problem of small current to ground system.

References

1. Ma S, B Xu, Gao H et al (2008) An earth fault locating method in feeder automation system by examining correlation of transient zero mode currents. Autom Electric Power Syst 32(7):48–52
2. Tian S, Wang X, Wang J (2010) A new method of fault location for small current to ground system. Electr Power 43(12):14–18
3. Li T, Zhao Y, Ji X et al (2005a) Application of HHT method for analysis of fault signal in electric power system. Trans China Electrotech Soc 20(6):87–91
4. Li T, Xie J, Zhang F et al (2007) Application of HHT for extracting model parameters of low frequency oscillations in power systems. Proc CSEE 27(28):79–83
5. Li T, Zhao Y, Li N (2005b) Apply empirical mode decomposition based Hilbert transform to power system transient signal analysis. Autom Electr Power Syst 29(4):49–52
6. Shu H (2009) Application of electrical power signal. Scientific Publication, Beijing

7. Tian S, Wang X, Wang J (2011) Comparative research on fault location by transient zero-module current and transient zero-module power based on correlation analysis. Power Syst Technol 35(4):206–211
8. Shu H, Xu L, Peng S et al (2008) Correlation analysis for faulty feeder detection in resonant earthed system. Electr Power Autom Equip 28(9):6–9

Chapter 123
Methods of Harmonic Source Identification and Detection in Distribution Network

Zhiguo Hu, Junwei Tian and Yongle Ai

Abstract In order to create Green Grid and supply the electricity consumers with safe, liable and clean power energy, the accurate and practical harmonic source location methods are urgently needed. Therefore this paper concludes various harmonic source location methods based on the direction of power flow, and explains the disadvantages and advantages of the methods. At last, this paper studies the principles and realization of $i_p - i_q$ algorithm. Moreover, in the Matlab/Simulink, the model is established to stimulate. This paper further analyzes the principles and characteristics of $i_p - i_q$ algorithm.

Keywords Harmonic source identification · Harmonic detection · Algorithm

123.1 Introduction

With the development of modern industrial technology, power electronic devices and high-power nonlinear loads are widely used, which makes the harmonic pollution get worse. In order to effectively control harmonic, we should clarify the harmonic source distribution and the harmonic state in the power system because it is the most important issue of harmonic analysis and control.

Harmonic source localization is to determine by measuring the voltage, current or power values of the common connection point [1, 2]. If the system side is the main harmonic source, then it bears the main responsibility for the voltage and

Z. Hu (✉) · J. Tian · Y. Ai
School of Electrical Engineering and Automation,
Henan Polytechnic University, Henan 454000, China
e-mail: jeffhzg@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_123,

current distortion. For the user side, it is same. This paper summarizes the harmonic source localization methods based on the direction of power flow, analyzes harmonic source criterion for the active power direction method and the reliability of the power direction method [3, 4]. Finally this paper studies the realization of $i_p - i_q$ algorithm.

123.2 Location Method Based on Power Flow

123.2.1 Critical Impedance Method

This method is to determine the side of main harmonic source by analyzing harmonic reactive power produced by the system equivalent voltage source $E_u\angle 0$, calculating impedance x required to absorb reactive power completely and then comparing it with half of harmonic impedance of both sides. If $x > X/2$, the amplitude of voltage source of system side is larger, if $x < X/2$, then the voltage source of user side has larger amplitude [5].

The coefficient of harmonic impedance (CI) is:

$$X = -\frac{Q}{I^2} = -\frac{E_u}{I}\sin\theta, \quad \mathrm{CI} = 2\frac{Q}{I^2}. \tag{123.1}$$

When $|\mathrm{CI}| > X_{\max}$, the system side is the main harmonic source; when $|\mathrm{CI}| < X_{\max}$, the user side is the main harmonic source. $X_{\max}$ and $X_{\min}$ are all possible maximum and minimum values for X.

In theory, this method solves principle defects of the active power direction method, and to a certain extent, remedies the shortage of reactive power direction method, but this method also has its own drawbacks. In practice, this method needs to predict the approximate impedance information to start the algorithm, and the characteristic of system harmonic impedance should be the same everywhere, so a large error is produced in actual use. In addition, the method ignores the impact of resistance in the design process, and supposes that harmonic impedance is uniformly distributed in the system, as a result, it will be affected to some extent in practical applications.

123.2.2 Discrimination Method of Simultaneous Measurement

Equation (123.2) can be obtained by making: $a = \frac{E_u}{X_u + X_c}$, $b = \frac{Q}{P}$

Then:

$$E_c \sin\delta = \frac{P}{a} E_c \cos\delta = E_u - \frac{Q}{a} \tag{123.2}$$

So: $E_c^2 = \left(\frac{P}{a}\right)^2 + \left(E_u - \frac{Q}{a}\right)^2$

By simplifying and rearranging:

$$E_c^2 - E_u^2 = E_c\left[\left(1 + b^2\right)E_c \sin^2 \delta - 2E_u b \sin \delta c\right]. \tag{123.3}$$

Because E_u and E_c are harmonic open circuit voltage amplitudes of PCC of both sides, and both are positive, we just need to know the polarity of $(1 + b^2)E_c \sin^2 \delta - 2E_u b \sin \delta$ to determine the polarity of $E_c - E_u$.

According to the reactive and active power of PCC and the phase angle difference of equivalent harmonic voltage source on both sides of PCC, the harmonic source can be located by the following criterions:

If $\frac{1-b^2}{2} > bctg\delta$ and $E_c - E_u > 0$, the user side is the main harmonic source;

If $\frac{1-b^2}{2} < bctg\delta$ and $E_c - E_u < 0$, the system side is the main harmonic source.

This method improves the measurement accuracy, but it requires measuring the value of the phase angle difference which is difficult to get. With the development of the GPS, the value of the phase angle difference can be measured accurately, which provides direction for the discrimination method of simultaneous measurement.

123.3 Harmonic Current Detection Based on $i_p - i_q$

123.3.1 Realization of $i_p - i_q$ Algorithm

The $i_p - i_q$ algorithm is based on the theory of H.Akagi instantaneous power, and improved from the p–q method [6]. It is not needed to detect three-phase voltage of power system. And when the voltage generates distortion, high frequency harmonics in the current, reactive power and negative sequence current also can be detected accurately. The principle of harmonic current detection based on $i_p - i_q$ algorithm is shown in Fig. 123.1:

$$C = \begin{bmatrix} \sin \omega t & -\cos \omega t \\ -\cos \omega t & -\sin \omega t \end{bmatrix}, \quad C_{32} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}$$

$$CC^{-1} = E \quad C_{32}C_{23}^{-1} = E \tag{123.4}$$

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = C_{32} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \begin{bmatrix} i_p \\ i_q \end{bmatrix} = CC_{32} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \begin{bmatrix} i_{af} \\ i_{bf} \\ i_{cf} \end{bmatrix} = C_{23}C^{-1} \begin{bmatrix} \bar{i}_p \\ \bar{i}_q \end{bmatrix} \tag{123.5}$$

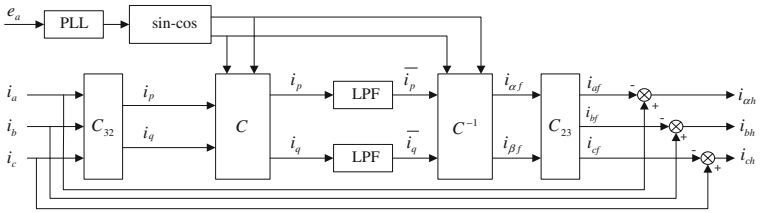


Fig. 123.1 Schematic of harmonic current detection based on $i_p - i_q$

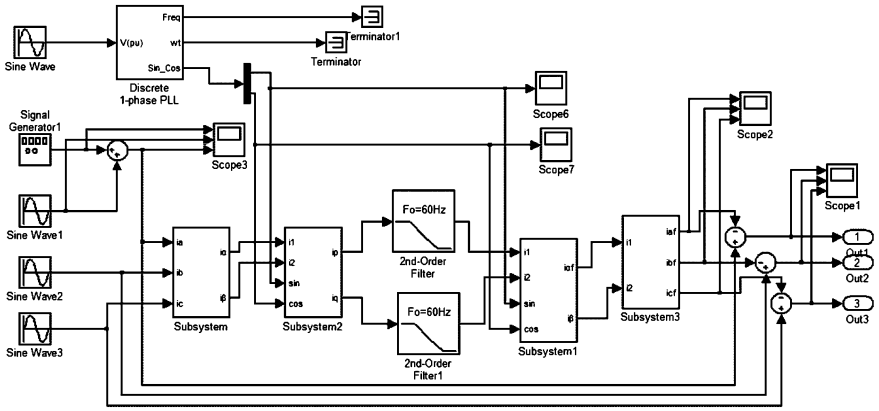


Fig. 123.2 Harmonic current detection model based on $i_p - i_q$

e_a is the A-phase voltage signal. The function of $\sin \omega t$ and $\cos \omega t$ are generated by PLL and cosine function generator. The matrix C_{32} is multiplied by three-phase current to get two-phase instantaneous currents i_α and i_β. Positive sequence active and reactive current $\bar{i}_p$, $\bar{i}_q$ corresponding to the fundamental can be drawn when i_p, i_q flow through Low-pass filter (LPF). We can obtain three-phase fundamental positive sequence current i_{af}, i_{bf}, i_{cf} by inverse transform. The difference between three-phase current i_a, i_b, i_c and three-phase fundamental positive sequence current is the value of three-phase harmonic and load current i_{ah}, i_{bh}, i_{ch}.

123.3.2 Simulation Study Based on Simulink

As is shown in Fig. 123.1 the harmonic current detection based on $i_p - i_q$ algorithm is implemented. Cosine function generator, low pass filter, the current module and multiplication module in Matlab are used to build a Simulink model for the implication of the harmonic current detection, which is shown in Fig. 123.2.

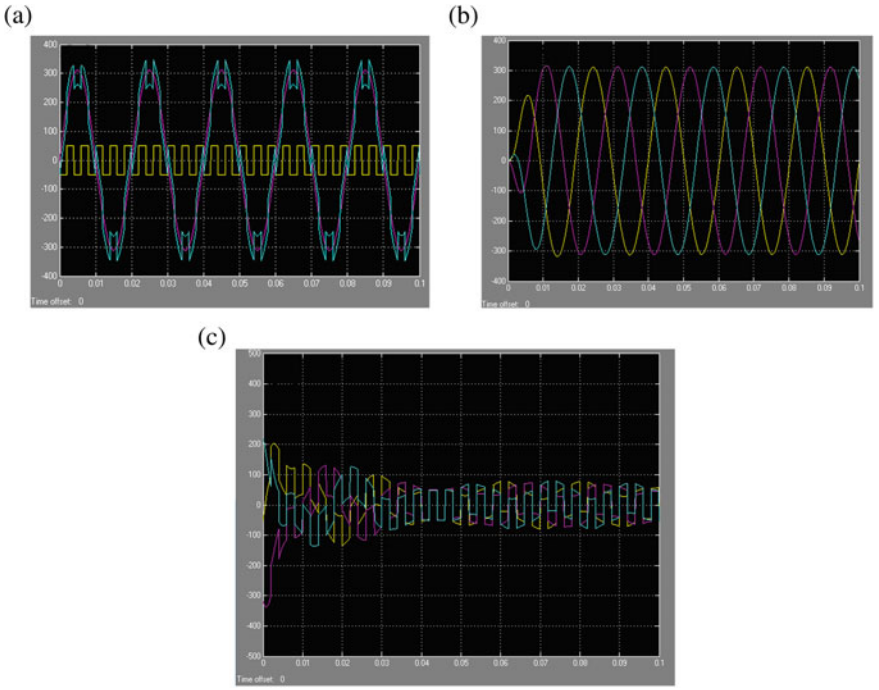


Fig. 123.3 Simulation results based on $i_p - i_q$ **a** Three-phase current containing harmonic. **b** Three-phase fundamental positive sequence current. **c** Detection of harmonic current waveforms

Suppose that the expression of frequency current is:

$$i_a = 220\sqrt{2}\sin(100\pi t)A;,\ i_b = 220\sqrt{2}\sin\left(100\pi t - \frac{2}{3}\pi\right)A;$$

$$i_c = 220\sqrt{2}\sin\left(100\pi t + \frac{2}{3}\pi\right)A$$

Harmonic current is square wave with the amplitude of 50 A and the frequency of 25 Hz. The simulation results are shown in Fig. 123.3.

It can be seen from the simulation results that the harmonic current detected is a square wave with the amplitude of 50 A and the frequency of 250 Hz. This method can detect the harmonic components more accurately. However, $i_p - i_q$ algorithm needs to use phase-locked loop (PLL) and LPF, resulting in a certain deviation delay between harmonic detected by $i_p - i_q$ algorithm and ideal harmonic, affecting the detection accuracy.

123.4 Conclusion

First, this paper concludes the principle of various location methods based on the direction of power flow, and explains the disadvantages and advantages of the methods. Second, this paper studies harmonic and reactive current detection ($i_p - i_q$) based on the theory of instantaneous reactive power, and the principles and realization of $i_p - i_q$ algorithm. Moreover, the model is established in the Matlab/Simulink to stimulate. This paper further analyzes the principles and characteristics of $i_p - i_q$. Because $i_p - i_q$ algorithm needs to use PLL and LPF, resulting in a certain deviation delay between harmonic detected by $i_p - i_q$ algorithm and ideal harmonic, the effect is less than the real-time detection when the deviation delay is too large. It is necessary to improve it. Currently, many improved harmonics and reactive current detection methods based on $i_p - i_q$ algorithm have been proposed one after another, and they can also accurately detect fundamental active and reactive current in the case of three-phase grid voltage distortion and asymmetry, which will be my future research directions.

References

1. Xu Z, Hou S, Lv H et al (2006) Power system harmonics source localization review. Electr Eng Energy 25(3):64–67
2. Jing K, Tian L (2008) The limitation study of power direction method in harmonic source detection. In: Twenty-fourth annual meeting proceedings (volume) of power system and automation in colleges of China. Beijing, China agriculture university, 1992–1994
3. Li C, Pan W, Bai X (2004) The comparison of harmonic source identification method in distribution network. Northeast Electr Power Technol 10:11–13
4. Moradloo M, Tabrizi MA (2008) A new method for identification of main harmonic source based on the superposition and critical impedance methods. IEEE Power Symp 10(13):1–6
5. Wang JP, Lin C (2009) Harmonic of grid detection based on virtual instrument. J Wu Zhou Coll 19(3):72–76
6. Lu X, Xu M (2010) The study of high harmonic current detection method based on the theory of Instantaneous reactive power[J]. Electr Meas Instrum 47(530):9–12

Chapter 124
Research on Wavelet Neural Network for Fault Location in Power Distribution Network

Zhang Yujun and Wang Xiaowei

Abstract The distribution network own characteristics such as: complicated topology structure and variable operation mode, so the fault location problem of the distribution network has not been solved. Based on the analysis of the single-phase ground of the system, this paper combines wavelet analysis and neural network. The method not only utilizes time–frequency features and zoom capabilities of wavelet analysis in order to pick up the characteristics of component, but also take advantage of the ability of non-linear fitting of neural network. These studies establish mapping between the characteristic component and fault location, thereby, resolving successfully the problem of the precision positioning, the thought of fault location has fully theoretical basis. The data of EMTP/ATP simulation shows that the scheme has high precision, adaptable and reliable performance.

Keywords Distribution network · Neural network · Fault location · Wavelet analysis

124.1 Introduction

At present, people desire a higher quality of electricity supply, so the study of distribution network fault location become increasingly active, many scholars made a substantial amount of research with the traveling wave fault location in

Z. Yujun (✉) · W. Xiaowei
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zyj@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_124,

distribution network. It is difficult to apply in complicated structure of distribution network for *A* principle-based traveling wave, because the wave refraction and reflection are vulnerable to bus, branch point and other factors. Traveling wave principle of the double-end achieve a certain application because of resisting the impact of these factors [1], but it can only locate to the trunk lines, and has a smaller positioning scope for many branch network. Reference [2] put forward a method which based on transient characteristics by the time series of wavelet neural network to achieve straight lines with single-phase ground fault location, but it is hard to extract high-frequency information in fault time; Ref. [3] used time difference between zero-mode and line mode traveling-wave when they transmitted in line, and utilize a law which Lipschitz index of the first wave of zero-mode varied with transmission distance. This fault location algorithm adopts the time–frequency ability of wavelet analysis and the non-linear fitting ability of BP neural network, but the simulation test based on the extinction theory of power frequency, so the practical applications have yet to be further improved; Ref. [4] made use of transient voltage and current of characteristic frequency band to calculate reactive power, and determine the fault segments by the different direction of transient reactive power before and after in fault location. However, it is hard to ascertain the optimum detection point and the effective area, thus affect the quality of fault location. In [5], it determined the fault segments by judging the relevance of transient zero current of adjacent FTU, this method demanded strictly signals which is collected by two FTU to keep pace with each other, thereby it has to increase the GPS timing device.

Application of wavelet theory are generally limited to the scope of small-scale, the main reason is that large-scale application need a large number of cost both wavelet - based structure and storage, while the neural network is fit for dealing with large-scale problem, so it is possible to make neural networks combine with wavelet analysis. As described above, this paper based on the analysis of wavelet time–frequency to extract information of transient fault, and take full advantage of the good non-linear fitting ability of neural networks to achieve a mapping between the fault characteristic component and fault location. The thought has sufficient theory evidence; therefore it can realize precise positioning.

124.2 Wavelet Neural Network

124.2.1 The Construction of Wavelet Neural Network

At present, there are two main combination ways between wavelet analysis and neural network.

Noncompact combination

The wavelet analysis serves as a means of preprocessing, and it provides the input vectors for neural network.

Compact combination

Wavelet analysis and neural network integrated directly, using wavelet function and scale function directly as the driving function of neurons. The wavelet neural network which we usually said is the latter form.

According to [6], it derived wavelet function as driving function for neural networks and the general driving function are the same in nature. Besides, it has verified that the wavelet network is better than ordinary in the area of function approximation. But at the same time, this conclusion pointed out that if the scale, displacement and the initial value of weight is not right, may be the wavelet network will not converge.

For the function space $L^2(R)$, suppose ψ is mother wavelet which has well localized, then the partial wavelet transform of a signal $f(t)$ is

$$c_{a,b} = |a|^{-\frac{1}{2}} \int_R f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \tag{124.1}$$

It has given components of a certain frequency band. Supposed the signal sampling sequence of limited time $f(t_i)$ is equal to $f(i\Delta)$, $t_i \in [t_1, t_2]$, therein, $i = 1, \ldots, N$. Then the formula (124.1) can be calculated discretely.

$$c_{a,b} = \frac{t_2 - t_1}{N} a^{-\frac{1}{2}} \sum_{i=1}^{N} f(i\Delta) \overline{\psi_{a,k}(a^{-1}i\Delta - a^{-1}b)} \tag{124.2}$$

According to the theory of wavelet analysis, with the changes of parameters a and b, the above-mentioned wavelet transformation can be achieved from the local, non-steady-state signal to extract the various frequency components of different frequencies. Provided select the appropriate wavelet and a number of sufficient item, the complete information was recovered by the wavelet transform formulas, that is, to form another expression for fault transient signal. However, only one specific frequency bands are sensitive to the changes of fault location [7], it is not necessary to calculate all frequency bands, the benefit of it is not only reflects significantly the characteristic components of fault location, but also eliminates the influence of non-fault line distance variety factors, and to achieve high performance devices of fault location, besides, it can also reduce the complexity of modeling.

The target problem is fault location, it is unknown for fault location which components can best reflect. So make the scale parameters which can reflect frequency components and the displacement parameters which can reflect location as a whole part weight space of wavelet neural network, and to adjust in accordance with objectives problem, thus to achieve a specific target adaptive wavelet transform with transient signal. In the light of the above mentioned, to construct a wavelet neural network model which is suitable for small current grounding system single-phase ground fault, as shown in Fig. 124.1.

In Fig. 124.1, the wavelet neural network model include: input layer, wavelet transformation layer, hidden layer and output layer. According to the number of

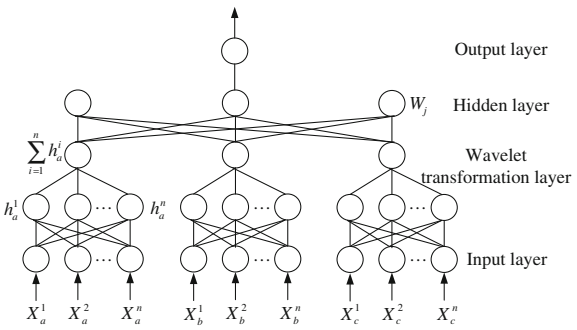


Fig. 124.1 The model of wavelet neural network for fault location

imported electrical quantities, input layer is divided into several groups, every node of group input the amount of its corresponding fault transient electrical time sequence in turn. Output layer contains a single neuron which is characterized by normalized value of distance, the distance is from fault point to device installed location, and its value reflect the location of fault.

124.2.2 The Algorithm of Wavelet Neural Network

For this paper, the wavelet neural network adopts BP algorithm to adjust whole weight space, including scale parameter a and displacement parameter b.

Forward recursion

In formula (124.1), if $a > 0$, then the signal $f(t)$ is dispersed to $f_N(i\Delta)$, so the formula (124.1) can be changed to:

$$W_\psi f(a,b) = a^{-\frac{1}{2}} \sum_{i=0}^{N} [f_N(i\Delta)\Delta]\overline{\psi}(a^{-1}i\Delta - a^{-1}b) \tag{124.3}$$

The symbol N stands for the total number of sequence points, T is for the time window width of signal and equal to $(N-1) \times \Delta$.

To prevent the boundary effect of transformation [2], the realization of fault location when the time window may be set at 1.5 cycle: before fault 1/4 cycle and after fault 5/4 cycle. The forward recursion of wavelet neural network model is as follows (suppose only a electrical quantity)

① Input time-series

$$X_1 = [X_0^1, X_1^1, \ldots, X_N^1]$$

The input node g_0^1 of adaptive wavelet transformation layer is:

$$g_0^1 = \left(a_0^1\right)^{-\frac{1}{2}} \sum_{i=0}^{N} [X_K^1 \times \Delta]\overline{\psi}\left(a_0^{-1}i\Delta - a_0^{-1}b_0^1\right) \tag{124.4}$$

Similar formula (124.4), to calculate the output vector of adaptive wavelet transformation layer

$$G^1 = [g_0^1, g_1^1, \ldots, g_{J-1}^1]$$

Therein, J is the number of all nodes for adaptive layer.

② To calculate the input vector of hidden layer.

$$\overrightarrow{U} = [u_0, u_1, \ldots, u_{K-1}]$$

$$\overrightarrow{U^T} = \overrightarrow{W_i^{2,1}} \cdot \overrightarrow{G^{1^T}}$$

Therein, K is the number of nodes for hidden layer.

And then, to calculate the output vector of hidden layer.

$$\overrightarrow{O} = f_2(\overrightarrow{U})$$

③ To calculate the output of network

$$y = f_3(I)\,,\ I = \overrightarrow{W^{3,2}} \cdot \overrightarrow{O^T}$$

Weight adjustment

When the neurons function is differentiable continuously, it can be strictly calculated by gradient method. To suppose the output error E_P of sample p is equal to $(d-y)^2$, then the weight adjustment of every layer is:

$$\begin{cases} \delta_P^3 = (d_P - y_P)_3^{'}(I) \\ \Delta_P w_j^{3,2} = \alpha \cdot \delta_P^3 \cdot o_{Pj} (j = 0, 1, \ldots, K) \\ \Delta_P w_{ij}^{2,1} = \alpha \cdot (\delta_P^3 i^{3,2}) \cdot f_2^{'}(u_i) \cdot g_{Pj}^1 \\ \delta_{Pi}^2 = (\delta_P^3 \cdot w_i^{3,2})_2^{'}(u_i) (i = 0, 1, \ldots, K-1; j = 0, 1, \ldots, J) \end{cases} \tag{124.5}$$

Therein, α is learning step, the threshold of every node is viewed as constant importation weight, and it equals to one.

According to equivalent error δ_{Pi}^2 of hidden layer, to adjust two parameters of adaptive wavelet function:

$$\Delta a_i^1 = \alpha \left\{ \sum_{K=0}^{K-1} \delta_{Pk}^2 \cdot w_{ki}^{2,1} \right\} \cdot \frac{\partial g_i^1}{\partial \alpha_i^1} \tag{124.6}$$

$$\Delta b_i^1 = \alpha \left\{ \sum_{K=0}^{K-1} \delta_{Pk}^2 \cdot w_{ki}^{2,1} \right\} \cdot \frac{\partial g_i^1}{\partial b_i^1} \tag{124.7}$$

Therein, $i = 0, 1, \ldots, J-1$

The adjustment of weight space should ensure that the parameter a is greater than zero. In order to determine the range of parameter b, the time window centers as well as the border of wavelet function are confined in the signal window. It can

Table 124.1 Input vectors of the locating neural network modules

Fault location module	Basic location layer	Generalization layer	Correction layer
Single-phase short circuit	U_f,I_f,I_0	U_f,I_f,I_0	I_M,U_M,D_{Mf}
Two-phase short-circuit	U_{fd},I_{fd},I_2	U_{fd},I_{fd},I_2	I_M,U_M,D_{Mf}
Two-phase ground	U_{fd},I_{fd},I_2,I_0	U_{fd},I_{fd},I_2,I_0	I_M,U_M,D_{Mf}
Three-phase short-circuit	U_A,U_B,U_C,I_A,I_B,I_C	U_A,U_B,U_C,I_A,I_B,I_C	I_M,U_M,D_{Mf}

increase the momentum term and adaptive learning rate and other modified methods in application.

124.2.3 The Determination of Input Vectors for Fault Location Module

The structure of wavelet neural network for fault location can be divided into the basic location layer, generalization layer and correction layer [8]. The other reasons of designing correction layer are for alleviating the every module fitting capacity of the basic location layer, and also it rectifies the basic fault location layer. Obviously, it is important to ensure the accuracy of fault location.

In this paper, based on fault analysis and combined with the requirement of rapid location, it determines the input vectors of every module. As shown in Table 124.1.

124.2.4 Construction of Wavelet Neural Network Studying Sample

In this Paper, taking the system which is from [2] for example, Mexican hat wavelet is choose over wavelet mother function in wavelet transform layer. The simulation system structure is shown in Fig. 124.2. Positive sequence impedance: $Z_1 = (0.17 + j0.37)\ \Omega$/km; Positive sequence conductance: $b_1 = 3.044\ \mu$s/km; Zero sequence impedance: $Z_0 = (0.23 + j1.71)\ \Omega$/km; Zero sequence conductance: $b_0 = 1.884\ \mu$s/km. The way of ground is overcompensation which is equal to 7.5%.

The studying fault mode cluster which is formed by fault location and ground resistance is shown in Fig. 124.2.

Choosing the fault location $k\Delta l$ along the entire line, Δl is the distance change step, $k = 0, 1, \ldots, L/\Delta l$;

Fault transition resistance is $R_f (= 0, 10, 20, 30, 50, 100, 200, \text{and } 300\,\Omega)$.

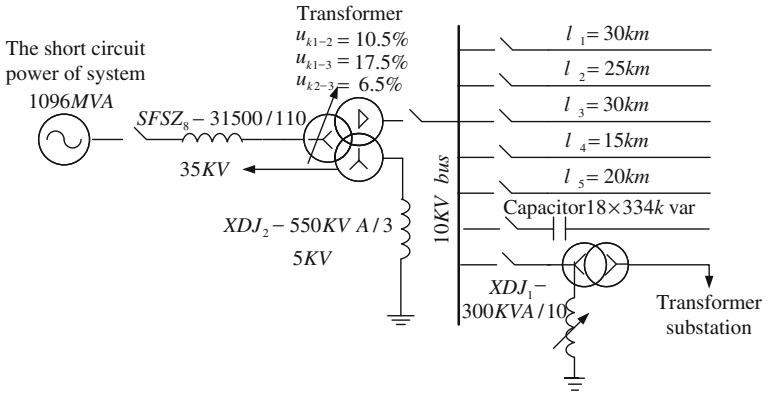


Fig. 124.2 The structure of non-direct-ground system

When $\Delta l = 0.5$ km, the combination of two factors can form $41 \times 8 = 328$ studying fault mode, to obtain the raw data by the application of EMTP/ATP simulation program. In order to enhance the positioning accuracy of near bus terminal at the time of practical training, it can appropriately reduce the distance change step at the range of near bus.

124.2.5 Analysis of Fault Location Result

In order to assess the actual positioning results of neural network better after learning, it should be tested by non-study samples.

Choosing fault location

$$(k\Delta l + 0.5)\ \text{km},\ \ \Delta l = 1,\ \ k = 0, 1, \ldots, 19;$$

Fault transition resistance

$$R_f (= 1, 30, 80, 150, \text{and } 250\Omega)$$

To form the fault testing mode, and the number is $20 \times 5 = 100$. In accordance with the pretreatment method to constitute input vectors cluster, and get the result of fault location by wavelet neural network forward recursion. Specific conditions were shown in Table 124.2.

Based on the analysis of Table 124.2, wavelet neural network after studying and training has better fits the corresponding relationships between input vectors and fault location. For the sample of neural network testing, the error kept basically below 1%, and have better positioning results. The accurate location scheme have two sides of reason, on the one hand, although the face to user directly, single-supply systems and the flexible operating way, but the majority of operating parameters are known, besides, this fault location scheme lower the fuzziness of operating parameters to some extent. On the other hand, the chief fault location principle of wavelet neural

Table 124.2 Results of fault location of single-phase-to-earth fault on non-direct-ground system

Ideal output/km	Actual output/km	Absolute error E/km	Relative error c/%
0.500 0	0.401 0	0.099 0	0.495
1.500 0	1.441 1	0.058 9	0.294
2.500 0	2.389 6	0.110 4	0.551
3.500 0	3.299 1	0.200 9	1.005
4.500 0	4.566 0	0.066 0	−0.033
5.500 0	5.379 0	0.121 0	0.605
6.500 0	6.673 2	0.173 2	−0.866
7.500 0	7.431 5	0.068 5	0.343
8.500 0	8.445 5	0.054 5	0.273
9.500 0	9.541 0	0.041 0	−0.050
10.500 0	10.627 3	0.127 3	−0.637

network lies in the extraction and application of transient high-frequency information while it reflects the release of capacitive current from fault point and system. Therefore, the wavelet neural network after studying has reliable fault location, higher precision and other distinguishing features.

124.3 Conclusion

Based on the analysis of single-phase ground fault characteristics of small current grounding system, this paper utilizes the superior time–frequency localization features and zoom capabilities for non-stationary random signal of wavelet theory to extract the transient characteristics component at the time of fault, and then put wavelet function which serve as driving function directly of the neurons into the neural network training. Because of the existing multi-dimensional wavelet neural network is not suitable for the analysis of post-fault transient electrical quantities, so it extends the concept of wavelet neural network and adopts the wavelet neural network model which fit for the analysis of signal time series. According to the good non-linear fitting ability of neural network, to establish the mapping between fault component and fault location, thereby, it obtains the aim of precise fault location. This fault location thought has sufficient theory evidence, and achieving the organic integration between the feature extraction and mapping. The results of simulation by EMTP/ATP indicated the small current grounding system have high precision, strong adaptability and reliable performance.

References

1. Yongduan X, Xu B, Jing L et al (2006) Traveling wave based fault location in 10KV automatic blocking and continuous railway power lines. Autom Electr Power Syst 30(5):68–73
2. Zhaoning Z, Peng M, Weiyong Y et al (2001) The fault location method for to-earth fault with time series wavelet neural networks system. Proc CSEE 21(6):66–71

3. Fan Z, Zhencun P, Shanshan M et al (2007) Fault location algorithm based on wavelet transform and neural network in distribution network. Autom Electr Power Syst 31(22):83–86
4. Sun B, Sun T, Xue Y et al (2008) Single phase to ground fault section location based on transient signals in non-solidly earthed network. Autom Electr Power Syst 32(3):52–55
5. Ma S, Xu B, Gao H et al (2008) An earth fault locating method in feeder automation system by examining correlation of transient zero mode currents. Autom Electr Power Syst 32(7):48–52
6. Zinging L, Xiaoru W, Zhengyou H et al (2002) Analysis and comparison of function approximation ability based on wavelet transformation, neural network and wavelet network. Autom Electr Power Syst 26(20):39–44
7. Zhang Q (1997) Using wavelet network in nonparametric estimation. IEEE Trans Neural Netw 8(2):227–236
8. Wang X, Zhang J, Li Z (2005) Application of distributed wavelet neural network based on multi-information in electric power system fault locating. Electr Power Autom Equip 25(11):51–55

Chapter 125
Study on Preventing Override Trip System for High-Voltage Power Grid of Coal Mine Underground Based on the Ethernet

Yu-mei Wang and Yi-du Guo

Abstract This thesis analyzes the reason of power grid of underground Override Trip incident and the existing solution, puts forward a plan of the grid of coal mine underground interlocking protection system based on switched Ethernet, makes designs for software and hardware and makes network real-time characteristic simulation. The results show that the interlocking protection system can satisfy the requirement of quick-break protection quick-action and selectivity and prevent override trip fault.

Keywords The power system of coal mine underground override trip switched Ethernet intelligent interlocking device

125.1 Introduction

The high-voltage grid of coal mine underground often uses unilateral power multilevel main type system, by many short cable (100–1500 m) to form series radial trunk-type network. Because the impedance of distribution circuit is small, make the beginning and end of sections of line's short-circuit current magnitude have smaller difference, to distinguish the short-circuit protection by current amplitude cannot meet the requirements of vertical selective [1], typically. When the lower line short-circuit current exceeds higher-level quick-break protection set value, the higher-level quick-break protection startup ahead causes Override Trip,

Y. Wang (✉) · Y. Guo
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: doraemonwq@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_125,

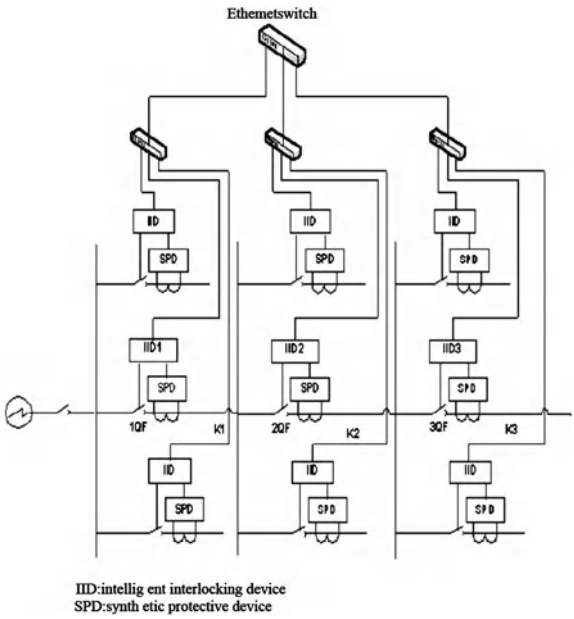


Fig. 125.1 The grid of underground coal mine interlocking protection system based on switched ethernet

causes underground a big electricity cut, the coal mine production system failures, threat to the mine and personal safety [2]. To ensure the reliability of the mine power system, the technical solutions of Web Communication to avoi override trip increases in recent years. The literature [3] uses the RS485 serial network be able to achieve quick-break protection selectivity, but cannot meet the protection quick-action requirement because of low communication speed. Differential protection scheme in the literature [4] can be applied in the grid of underground Multi-cascade structure, the wiring is very complicated. Literature [5] designs the Interlocking protection network based on CAN bus; still cannot meet real-time requirement because of low speed of CAN bus. This paper puts forward the interlocking protection scheme based on switched Ethernet, has advantages of high communication speed, low time-lapse, high reliability, can satisfy the requirement of quick-break protection, quick-action and selectivity and prevent override trip fault.

125.2 The Design of Grid of Underground Coal Mine Interlocking Protection System Based on Switched Ethernet

Ethernet is used gradually in industrial control domain in recent years by the advantage of high communication speed, good compatibility, easy to expand and so on. But the initial Ethernet is not designed for industrial control, it is unable to

Table 125.1 The relevance logic of each intelligent interlocking device

Fault position	IID number	Fault is detected	Receive interlocking	Delay time (ms)	Send interlocking
K3	3	yes	0	10	2
	2	yes	1	20	1
	1	yes	2	30	–
K2	3	no	0	–	–
	2	yes	0	10	1
	1	yes	1	20	–
K1	3	no	0	–	–
	2	no	0	–	–
	1	yes	0	10	–

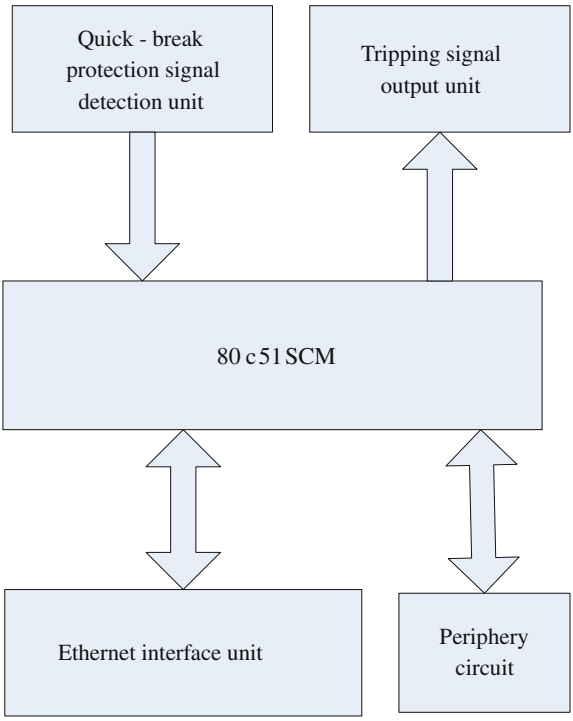


Fig. 125.2 Intelligent interlocking device structure

satisfy the need of the control signal transmission low latency and certainty. Switching technology overcomes the disadvantage of the traditional shared Ethernet uncertainty and bad real-time transmission, has obtained the widespread application in industrial control domain [6]. Switched Ethernet through connecting each network segment by switch, can form multiple data channels between ports at

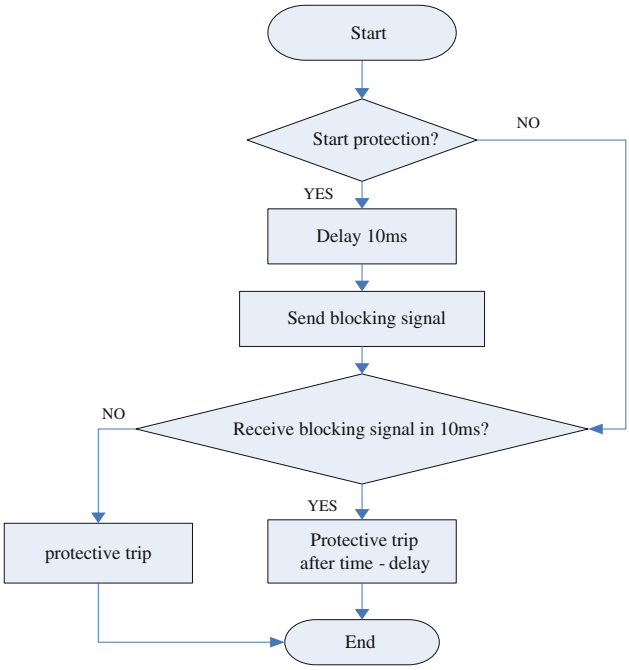


Fig. 125.3 The intelligent interlocking device program flowchart

the same time, the data transmission between ports is not limited by Media Access Control protocol, provides the point-to-point connection between source and destination, the port has a monopoly on the bandwidth and communicate rapidly and in real-time. So, we use the switched Ethernet in the scheme, the transmission medium is optical fiber, have long data-transmitting distance and resistance of the electromagnetic interference.

The design based on tree structure switched Ethernet has to add the intelligent interlocking device on each synthetic protective device, as the composition of the interlocking system. It is shown in Fig. 125.1.

If a fault is detected, the intelligent interlocking device delays 10 ms to activate the quick-break protection (except the last stage), and send interlocking signals to all upper-level intelligent interlocking device over the network at the same time. Each intelligent interlocking device delays 10 ms to activate the quick-break protection when receiving one interlocking signal, to organize all levels of protection action time, prevent override trip fault and have backup protection reliability well. Table 1 shows the relevance logic of each intelligent interlocking device.

The key of scheme is the reasonable and reliable data transmission networks, able to transmit interlocking signals timely and accurately in the early 10 ms to organize all levels of protection action time.

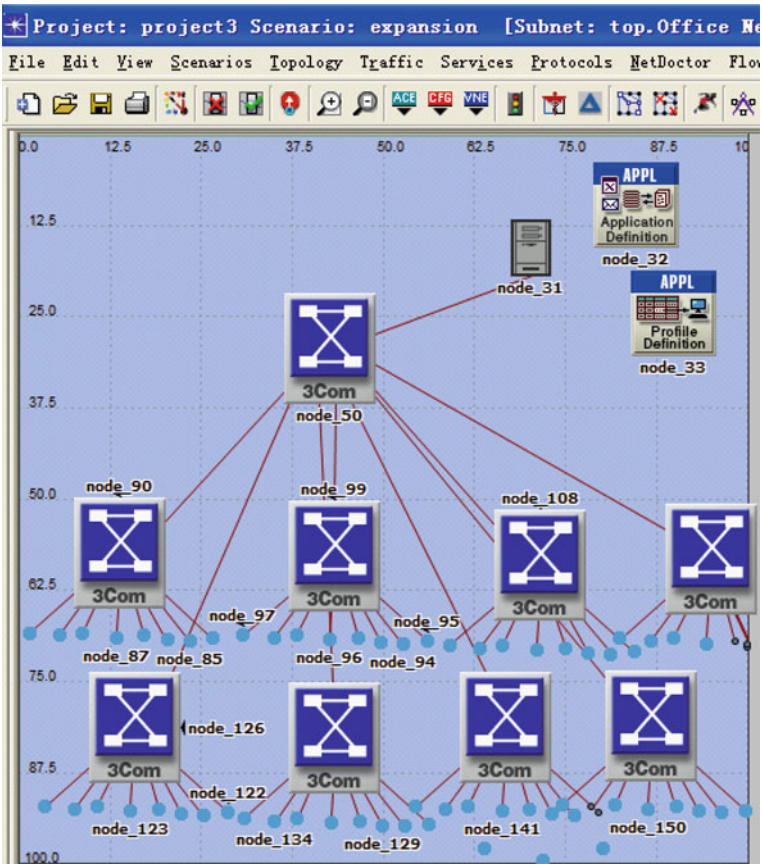


Fig. 125.4 The simulation network model

125.3 The Design of Intelligent Interlocking Device With Ethernet Support

(1) Hardware Design

The intelligent interlocking device consists of 80c51 SCM, periphery circuit, quick-break protection signal detection unit, trip signal output unit and Ethernet interface unit, as shown in Fig. 125.2.

The quick-break protection signal detection unit's function is to adjust the quick-break protection signal and convert it to digital signal, which is sent to the 80c51 SCM. The SCM signals through the corresponding treatment in order to achieve the purpose of sending interlocking and tripping signal.

The Ethernet interface unit uses RTL8019AS as the Ethernet controller chip, accepts data packets from the Ethernet and analyzes them. It can also send data by Ethernet.

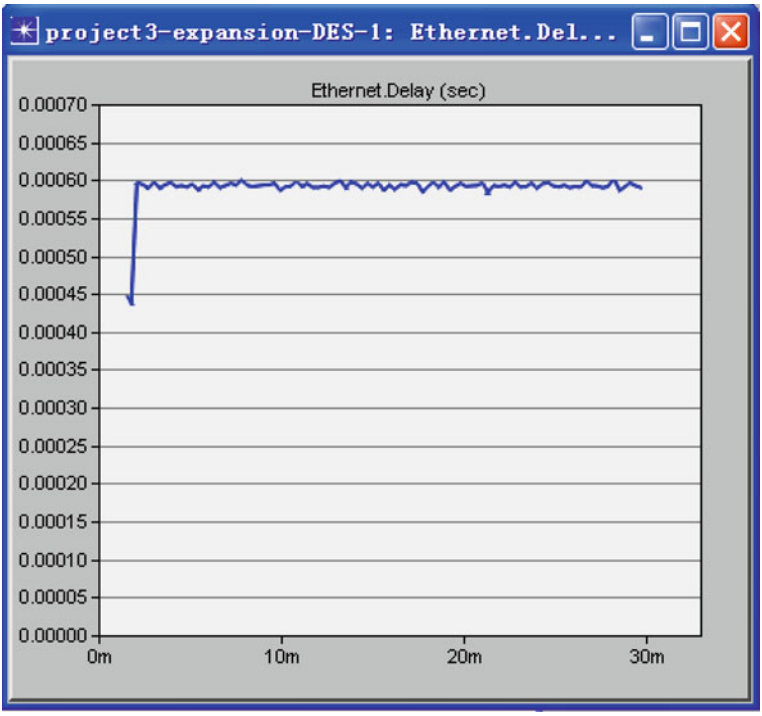


Fig. 125.5 The information of communication delay

(2) Software Design

The intelligent interlocking device delays to activate the quick-break protection send interlocking and tripping signal by programing. The program flowchart is shown in Fig. 125.3.

125.4 Real-Time Characteristic Emulate of Switched Ethernet

The high-voltage grid of coal mine Underground is based on many short cable to form series more radial trunk-type network, as the composite protection device are scattered, the scheme is composed of the three tier tree network system structures, and the OPNET Emulation Programs is used to simulate the behavior of the real time. The simulation network model as shown in Fig. 125.4, the number of second stage switch is eight, and there are eight nodes each connected to second stage switch.

The simulation experiment configures data service in network by Application tool, the length of data frame is 256bytes and the transmit periods are 0.01 s [7]. Get the information of communication delay as shown in Fig. 125.5.

The figure shows the communication delay of network is very stabilized, the maximum delay does not take more than 10 ms, and can meet the time limit of section interlocking system.

125.5 Conclusion

The system structure is not altered in the scheme, the intelligent interlocking device in series in the tripping signal circuit of synthetic protective device, it has advantages of easy transformation, simple connection structure and high reliability. With the development of techniques, the development trend of the relay protection is computerization, networking, intelligent, protect, control, measuring and communication integrate. The intelligent interlocking device will be integrated into synthetic protective device to achieve concentrated data transfer, strong real-time, quick transmission speed and big capacity of the Ethernet will become more obvious advantages.

References

1. Wang Y, Ma X, An W (2009) A scheme of adaptive protection system for high-voltage power grid of coal mine underground based on independent substation. Ind mine autom 10:5–8
2. Wu W, Chen B, Gao Y (2008) Research of grade-skip trip of underground power grid and its suggestion of resolution. Ind mine autom 6:136–138
3. Ning C (2005) Research of underground coal mine high-voltage grid power-off protection. Coal technology (24)9
4. Feng J, Feng Q (2006) Study of chain-type over-current protection for cascaded lines. Relay 34(8):80–83
5. Zhang G, Ma X, Zhang C (2009) Research on avoiding override trip system of underground high-voltage distribution network. Colliery mech electr technol 2:56–58
6. Gong L (2008) Research on real-time performance for switched industrial ethernet. Taiyuan Univ Sci Technol 07
7. H Jun (2007) Build real-time reliable and secure industrial ethernet.Ind Control Comput 16(10)

Chapter 126
Research on the Intelligent Safety Assessment System of Colliery Power Distribution System

Duan Jun-dong, Guo Li–li and Cao Kai

Abstract Safe and reliable power supply is an important guarantee of safety production in coal mines. The intelligent safety assessment of a colliery power distribution system can effectively improve the safety and reliability of the coal power grid. According to the characteristics of coal mine production, combined with the practical situation of coal mines, this paper has proposed the intelligent safety assessment index of the colliery power distribution system, given the intelligence assessment and warning display of the security of the system by using the flow data of the various operation states, and given the safe decreasing load order of colliery power distribution system, which provides the reasonable load adjustment scheme. Then, a colliery power distribution system as an example is given to show the rationality of the intelligent safety assessment.

Keywords Coal mine safety production · Colliery power distribution system · Power supply safety of coal mine · The safety index · Intelligent safety · Assessment

D. Jun-dong (✉) · G. Li–li · C. Kai
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: jundongd@hpu.edu.cn

G. Li–li
e-mail: dlxtguo@126.com

C. Kai
e-mail: caokai.pds@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_126,

126.1 Introduction

In recent years, safety accidents in coal mines have been happening frequently. There were more than 70 mine disaster accidents in China in 2010 [1], which caused major economic losses and serious injuries to our country and enterprises. To improve safety of coal mine production is essential. Safe and reliable power supply is an important guarantee of safety production in coal mines. The coal mine power supply automation level determines the safety and reliability of the colliery power distribution system. The intelligent safety assessment of the colliery power distribution system can improve the safety and reliability of the colliery power distribution system and effectively improve the coal mine safety production level.

The intelligent safety assessment of the colliery power distribution system is the intelligent assessment of the safety of coal mine power supply network. With the characteristics of big users and distribution network, this paper assesses colliery power distribution system as a distribution network. The power system security assessment is divided into the transmission network safety assessment and the distribution network safety assessment. The voltage grade is lower, the load is lighter, and the dynamic element is less, so that there is no obvious transient process in the distribution network. The distribution network usually has the characteristics of closed-loop design and opened-loop operation. Based on these, the safety analysis of the distribution network mainly refers to the static analysis, not to copy the criterion and analysis control method of the transmission network. At present, the research on distribution network safety assessment is rare. Reference [2] proposed that the dissipated variable structure network model is expanded by a new data structure of bus node, based on which, a simplified model of distribution networks considering substations is proposed. Reference [3] proposed the severity index of fault and the strength index of the network structure to analyze distribution network security. Coal power supply networks often use double circuit and closed-loop design, opened-loop operation, multi-voltage sub-peer and many cable lines, so the distribution network safety assessment is referable [4]. But possible accident types of coal mine power supply are multi-faceted and complex, and many loads in coal mine power supply are the primary load, if blackouts cause major economic losses or are life-threatening. There must be an index which can reflect the real-time network information as the standard of the assessment.

According to the characteristics of the coal power supply network, through the establishment of coal mine power network contingency, this paper has proposed the intelligent safety assessment system of the colliery power distribution system, and given safety assessment and dynamic display by using the changes of assessment index, proposed in this paper. This paper further discusses the load adjustment in poor safety and overload.

126.2 Distribution Network Safety Assessment System

According to the role in the power system, power network is divided into transmission network and distribution network. The traditional safety assessment of a distribution network is used to analyze contingency to a given operation mode, and warn about the accident threat to the grid security, evaluate the safety level of the distribution network, and find out the weak links in the system operation. Most traditional safety assessments of distribution network are about the grid strength and severity of fault, this method is referable for complex coal mine networks. This paper proposes the intelligent safety assessment system of colliery power distribution systems through repeated contingency analysis to a coal mine, which can give a certain real-time warning about coal mine power supply safety [5–7].

Reference [3] proposed the severity index of fault and the strength index of the network structure to analyze the distribution network security, and an example is given.

Let S_{FL} be lost load capacity, S_{SL} be load capacity, N_{PC} be all losses after the accident, N_{SC} be the number of users for the system, the loss rate of load after the accident is calculated as follows:

$$\rho_{\mathrm{FL}} = \frac{S_{\mathrm{FL}}}{S_{\mathrm{SL}}} = \sum_{i=1}^{N_{\mathrm{FC}}} S_{\mathrm{FL}}^{i} / \sum_{j=1}^{N_{\mathrm{SC}}} S_{\mathrm{SL}}^{j} \tag{126.1}$$

Let r_{FL}^{i} be the grade factor of the removed user i, r_{SL}^{j} be the grade factor of the user j, loss rate of users can be expressed by:

$$\rho_{\mathrm{FC}} = \sum_{i=1}^{N_{\mathrm{FC}}} r_{\mathrm{FL}}^{i} / \sum_{j=1}^{N_{\mathrm{SC}}} r_{\mathrm{SL}}^{j} \tag{126.2}$$

Let N_B be the number of system branches, λ_i be the failure rate branch i, L_i be the length of the branch i, C_F^i be a single incident severity index of the branch i, T_{FL}^{i} be the fault restoration (last) time, ω_1 and ω_2 be weight coefficient of energy loss rate and user loss rate, $\rho F = \omega_1 \cdot \rho\mathrm{FL} + \omega_2 \cdot \rho\mathrm{FC}$, the severity index of fault can be calculated by the following:

$$C_{SF} = \sum_{i=1}^{N_B} C_F^i / N_B = \sum_{i=1}^{N_B} \left(\lambda_i \cdot L_i \cdot T_{FL}^i \cdot \rho_F^i\right) / N_B \tag{126.3}$$

Let K_i be the K value of branch i happens accident, every possible accident, there is a corresponding K value is found in every distribution network, N_B be the number of system branches, the strength index of the network structure is calculated as:

$$K' = \left[\sum_{i=1}^{N_{\mathrm{B}}} (K_i + 1) / C_F^i\right] \Big/ \left(\sum_{j=1}^{N_B} C_F^{j^{-1}}\right) \tag{126.4}$$

The whole safety index of system can be expressed by:

$$S_S = K'/C_{\mathrm{SF}} \tag{126.5}$$

126.3 Intelligent Safety Assessment System of Colliery Power Distribution System

126.3.1 Proposing the Intelligent Safety Assessment Index of Colliery Power Distribution System

The purpose of the safety assessment of a colliery power distribution system is to assess whether branch is overload (the branch load is more than the maximum load regulation in normal operation) when system is normal or after the load of some fault branch is decreased. The change of branch load would cause the change of bus voltage, current, and power of the branch, so this paper defines the safety index using the bus voltage, current, and power of the branch [8–10]:

$$K_u = U/U_a \tag{126.6}$$

$$K_P = P/P_a \tag{126.7}$$

$$K_i = I/I_a \tag{126.8}$$

All three parameters exist in inner link, therefore the intelligent safety assessment of colliery power distribution system just needs to choose one of them. In this paper, the current index is an example.

$$K = I/I_a \tag{126.9}$$

where I denotes branch current or through on both sides of the current transformer, I_a denotes branch current ratings or transformer on both sides of the ratings that may take over current protection setting value or the current value of branch allows.

The value K reflects the size of the system safety degree. Synthesizing wire maximum carrying current and static security analysis results, it can be considered: while $K < 0.85$, the system in the security condition; while $0.85 < < K < 1$, some branch of system in the poor security condition; while $K > > 1$, some branch of system in overload operation condition (Fig. 126.1).

126.3.2 Process of Intelligent Safety Assessment

Analyze the coal mine power supply network, make the system diagram of coal mine power supply network, calculate the flow data of the various operation states

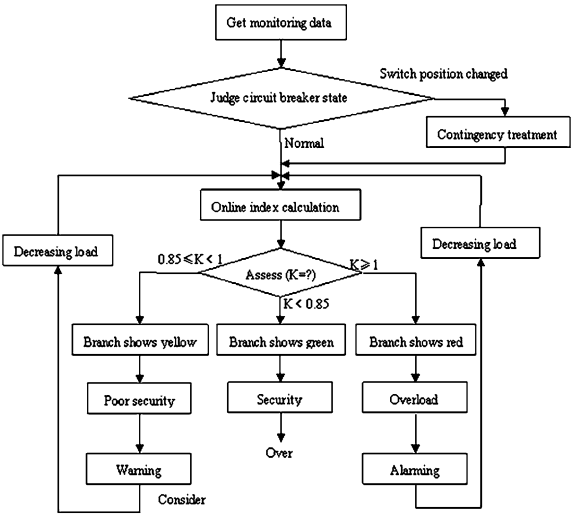


Fig. 126.1 The flowchart of intelligent safety assessment system

using Newton iteration method, and establish the coal mine power network contingency set. System can be online assessed by using contingency set. Evaluation processes are as follows: first, get the monitoring data, judge the circuit breaker switch state. Normal to judge online indexes, if it appears that switch position is changed to shift to contingency treatment, then to judge online indexes, paint in the corresponding line, give warning display of poor securely branch and the related branch user information table, give alarm display of the overload lines and the related branch user information table.

While $K < 0.85$, all branches are in the security condition, green; while $0.85 \ll K < 1$, part of the branches are in poor safety operation status, yellow and warning display; while $K \gg 1$, part of the branches are in overload operation status, red and alarm display. When branch warning or alarm, load can be adjusted according to the load information from the branch user information table, until the warning or alarm disappears and branch shows green. The following is the flowchart of the intelligent safety assessment system:

126.4 Load Adjustment of Branch Warning and Alarming

126.4.1 Establishing the Branch User Information Table

While the wire is in overload operation condition, it will overheat to burnt out the insulating layer or lead to other fire accidents. Many loads in coal mines are primary loads, causing blackouts that are major economic losses or even life-threatening. To ensure safety in coal mine production, generally, branch cannot be

long-term in no-security condition. Branch in poor safety operation state needs to prevent or adjust to the security condition by decreasing load according to the actual situation; branch is in overload operation condition that generally needs to adjust to poor safety operation state by decreasing load, then process according to the actual situation of the poor security condition. Coal mine loads generally fall into under Level 3, to ensure the safety of primary load by removing tri-grade load or secondary load in the load adjusting process. At this time, a branch user information table which can provide the basis for the safety decreasing load is needed. The branch user information table should include load type, level, working status, and information about related branch. From the load, is defined the branch directly connected with the load as primary, secondary, tertiary and so on.

126.4.2 Load Adjustment While Branch Warning and Alarm

While branch warning shows that branch is in poor security condition, it needs to strengthen the guard, or decrease load until warning branch into safe operation state according to the load information shown in the branch user information table. While branch alarming shows that alarm branch is in overload condition, it needs to decrease load until alarming branch into poor security operation state according to the load information shown in the branch user information table, then process load according to the actual situation of the poor security condition.

The following is usually the decreasing load order. The tri-grade load (such as the dining room substation, Workers village substation, etc.) of primary can be generally removed, then the secondary load (such as underground production links power lighting equipment, production and living water pump, spare fire pump, wellbore heat preservation equipment, mining water pump, automatic switch, railway transport equipment of electrical concentration, miner's charging equipment, worker hospital operation room, etc.) of primary branch, and then the tri-grade load and the secondary load of the secondary branch, the tri-grade, secondary load of the higher level of branch in turn.

Warning and alarming can be eliminated from low to high in turn. If all branches are in safety state after the adjustment of the lower branch load, they need not adjust again; if its superior branch is still in poor safety or overload operation state, continue to adjust load from low to high in turn, until all branches are into safe operation state.

The branch load is adjusted in a nearby principle, while branch warning or alarming is first considered with decrease load of the tri-grade and secondary load of the lowest level branch in all warning or alarming branches. If the branch is still in the warning or alarming state after adjusting, it can be considered to decrease the load of the lower branch until entering safety status; if its superior branch still appears branch alarming or warning after adjusting, adjust the load to their superior warning or alarm branch in the nearby principle.

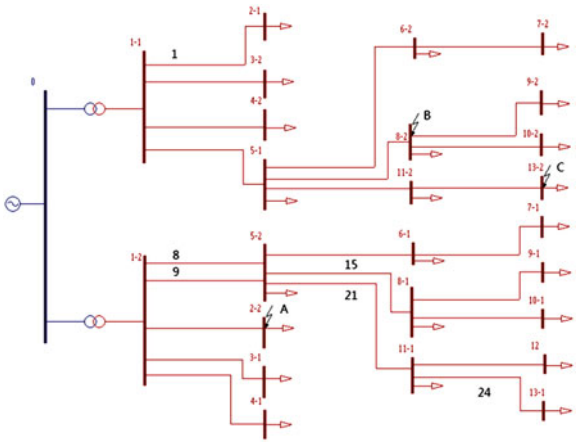


Fig. 126.2 The system diagram of some colliery power distribution system in normal operation

Because of the characteristics of coal mine safety production in coal mines, load level will change in special circumstances. When overload appears, the sequence of decreasing load should be adjusted according to the load level changes in the actual situation. Usually, mine fan, lift personnel hoist, mine main drainage equipment, mine gas drainage, and ventilation equipment belong to primary load, and level of part load changes in special circumstances. For instance, the importance of primary load disappears, which can be used as secondary load after all personnel withdraw from mine; the importance of part loads disappears when people are trapped at the bottom of a mine, but fan and drainage equipment are still used as primary load; the importance of drainage equipment disappears, which can be used as secondary load when drainage equipment is flooded.

126.5 Simulation of the Intelligent Assessment System of Some Colliery Power Distribution System

In order to prove the feasibility of the intelligent safety assessment, some colliery power distribution system as an example is given as follows. Figure 126.2 is the system diagram of some colliery power distribution system in normal operation.

where i denotes the bus $i(0,\ 12)$, $i-1$ denotes the east end of the bus $i(1, 2, \ldots, 11, 13)$, $i-2$ denotes the west of the bus $i(1, 2, \ldots, 11, 13)$. 1, 8, 9, 15, 21, 24 denote the numbers of branches. In addition to bus 12, other connection mode is single bus section, load switches to another side bus when a side line fault.

All breaches show green after load switch when the point A fault in Fig. 126.2, system is in a safe state. Part breaches show yellow after load switch when the point B fault in Fig. 126.2, as shown in Fig. 126.3.

Part branches show red after load switch when the point C fault Fig. 126.2, as shown in Fig. 126.4.

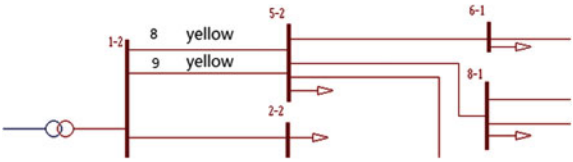


Fig. 126.3 The part area shown *yellow* after load switch when the point *B* fault

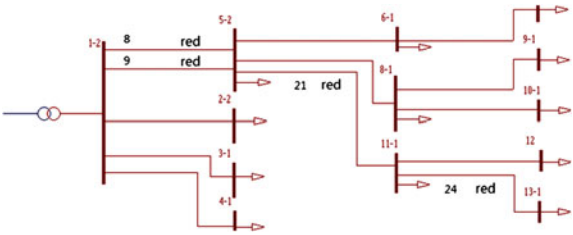


Fig. 126.4 The part area shown *red* after load switch when the point *C* fault

Table 126.1 The safety index value of the relevant lines in different operation states

	K_1	K_8	K_9	K_{15}	K_{21}	K_{24}
Before the point *A* fault	0.09	–	–	–	–	–
After load switch when the point *A* fault	0.18	–	–	–	–	–
Before the point *B* fault	–	0.82	0.82	0.15	–	–
After load switch when the point *B* fault	–	0.85	0.85	0.2	–	–
Before the point *C* fault	–	0.82	0.82	–	0.69	0.52
After load switch when the point *C* fault	–	1.14	1.14	–	1.61	1.41

The security of various operation states is assessed by using the index defined in this paper as follows. $K_i(= 1, 8, \ldots, 24)$ denotes the safety index K value of line i, the results are as follows (Table 126.1):

System in safe state need not adjust after load switch when the point *A* fault. The load can be adjusted in the poor security state according to the information of the branch user information table after load switch when the point *B* fault. The load can be adjusted in the overload state according to the information of the branch user information table after load switch when the point *C* fault. The safety index values of the relevant lines before and after load switch when the point *B* fault is given as follows (Table 126.2).

After load switch when the point *B* fault, branches 8 and 9 give warning, which are at the same level, load is adjusted in nearby principles according to branch user information table. The warning disappears after adjusting and load adjusting is complete.

After load switch when the point *C* fault, branches 8, 9, 21, 24 appear alarming, first adjust branch 24 from low to high, alarming of branch 24 and 21 disappeared after adjustment; branch 8 and 9 appear warning, adjust the load in nearby principles, warning disappear after adjustment, load adjustment is complete (Table 126.3).

Table 126.2 The safety index value of the relevant lines load before and after load switch when the point *B* fault

	K_8	K_9	K_{15}
Before load switch when the point *B* fault	0.85	0.85	0.2
After load switch when the point *B* fault	0.8	0.8	0.16

Table 126.3 The safety index value of the relevant lines load before and after load switch when the point *C* fault

	K_8	K_9	K_{21}	K_{24}
Before load switch when the point *C* fault	1.14	1.14	1.61	1.41
After load switch when the point *C* fault	0.84	0.84	0.82	0.8

126.6 Conclusion

Through the example analysis it is proved that the proposed intelligent safety assessment system can give intelligence assessment and warning display according to the contingency set. The intelligence safety assessment system of colliery power distribution system is real-time monitoring, evaluation, not security warning, load adjustment for the coal mine power network, which provides power for coal mine security production, reduce the production accidents from the power supply.

References

1. Chinese mine disaster accidents list in 2006–2010. http://www.docin.com/p-154188480.html
2. Liu J, Xu J, Dong H (2003) Security analysis of distribution networks[J]. Autom Electr Power Syst 27(17):26–30
3. Liu W, Guo Z (2003) Research on security indices of distribution networks[J]. Proc CSEE 23(8):85–90
4. Duan J, Sun Y, Yin X (2009) Voltage stability's online prediction using WAMS[J]. High Volt Eng 35(7):1748–1752
5. Duan J, Guo Z (2005) Research of on-line assessment of voltage stability[J]. Proc CSEE 25(25):199–202
6. Sidhu TS (2000) Contingency screening for steady-state security analysis by FFT and artificial neural networks[J]. IEEE Trans Power Syst 15(1):421–426
7. Ding P, Li Y, Xu D, Tian F, Yan J (2010) Improved algorithm of fast static state security analysis of power systems[J]. Proc CSEE 30(31):77–82
8. Damir N (2008) Emerging technologies in support of smart grids[A]. In: IEEE PES general meeting[C], Pittsburgh
9. Wang Z, Li H, Li J, Han F (2009) Assessment index system for smart grids[J]. Power Syst Technol 33(17):14–18
10. Yang D, Li Y, Rehtanz C, Liu Z, Luo L (2009) Study on the structure and the development planning of smart grid in China[J]. Power Syst Technol 33(20):13–20

Chapter 127
Research on Power Quality Data Transfer Format Based on XML in Power Quality

Lu Zhao, Guohai Liu and Yue Shen

Abstract Strength, compatibility, economy and interactivity are the main developmental features of Smart Grid in China. Aiming at the data exchange and sharing problem, Extensible Markup Language (XML) are introduced based on the research of Power Quality Data Transfer Format (PQDIF) in power quality management system. To study this research deeply, PQD files are transformed into XML files to solve these issues, and it describes the elaborate notes and attributes of XML files. In addition, now taking mass data needed to be stored into account, two open data compression methods will also be introduced below. The generated XML files have superiority on scalability, readability, compression and other aspects absolutely, which will be used widely in power quality data monitoring system.

Keywords PQDIF · XML · Power quality · Data exchange

127.1 Introduction

In recent years, the rapid development of the Internet technology has provided an effective way for the power quality monitoring system network that is to build a power grid monitoring network based on the network platform for real-time

L. Zhao (✉) · G. Liu · Y. Shen
College of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu, China
e-mail: zhaolu19898@126.com

G. Liu
e-mail: ghliu@ujs.edu.cn

Y. Shen
e-mail: shen@ujs.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_127,

monitoring and disseminating power quality information. On this basis, the Power Quality Data Transfer Format (PQDIF) comes out, which analyzes and describes the operation status of power grid in a relatively simple way, and conducts a comprehensive monitoring and analysis for multiple monitoring points of the whole network. It plays an important role as a bridge for data transmission and sharing of the power industry, and the issues of vast amounts of data, data multi-attribute and data sharing appear in power quality monitoring can be solved. Our country has drawn up national standards for the power quality concept according to the IEEE [1], so it is an urgent work to establish an efficient power quality system on the network platform.

For the goal of data conversion and information exchange, this paper describes the PQDIF recommended by the IEEE, and proposes a new data exchange model using the Extensible Markup Language (XML) with characteristics of open, scalable and easy-to-network transmission combined with power quality data transfer format, then verifies its feasibility and practicality. The result meets the future requirement to build multi-function power quality management system.

127.2 The PQDIF

PQDIF is standardized by the IEEE to provide a stable target for third-party developers [2, 3], and it can provide a platform-neutral way to interchange Power Quality data between instruments and software packages. Its characteristics are simple, extensible and flexible; even more it has very good compatibility for multiplying data sources and data processing ability. A PQDIF file is made up of a set of records which are logically related data sources, monitor settings and observations. With each record, there is a set of elements which define the contents of the record. Because each element is tagged, there is a separation between the physical structure and logical structure.

127.2.1 The Physical Structure of the PQDIF

The physical structure can be divided into two sections to describe, the structure of the whole file and the internal structure of records. A PQDIF file is made up of a series of linked records. The linked nature allows manipulations to the record structure without having to rewrite the entire file. In Fig. 127.1, there is a linked record. As shown in Fig. 127.2a–c, it can delete a record by adjusting the links to skip around it, and insert a new record—even if it is physically at the end of the file, and reorder a record by adjusting the links.

The internal record structure is like Table 127.1. Each record is made up of a header and a body. The header provides the high-level information, such as the signature and the record type tag. The body of each record is composed of a

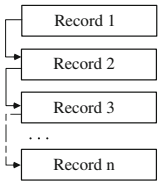


Fig. 127.1 A linked record

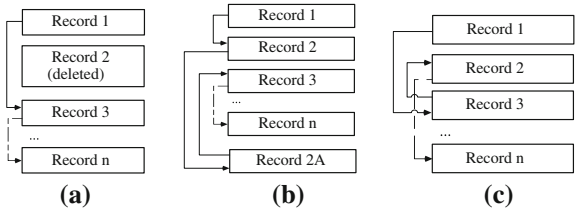


Fig. 127.2 Change the linked record

Table 127.1 The internal record structure

Header
–Signature
–Record type tag
–Size of the body
–Link to next record
Body
–A self-contained block of data

collection of elements. There are three types of all the elements, scalar, vector and collection. The scalar represents a single data value and the vector represents an array of data values and the collection contains other elements. In order to be able to identify a specific element, each one is given a tag. The tag is actually a globally unique identifier (GUID). Any computer in the world can generate a GUID and assure that it is absolutely unique.

127.2.2 The Logical Structure of the PQDIF

The logical format defines what the tags are, and what element type and primitive data type(s) are expected for each. In the high-level logical format, the records are strung together in an implicit hierarchy, based on the record type tags. Then in the low-level logical structure, there are standard tags and hierarchy definitions. The file starts with a container record, which owns all the number of data source records. Each data source record owns Monitor Settings and Observation records. Figure 127.3 shows the general overview of the logical structure. Since collections can contain other collections, the logical format must also specify this hierarchy.

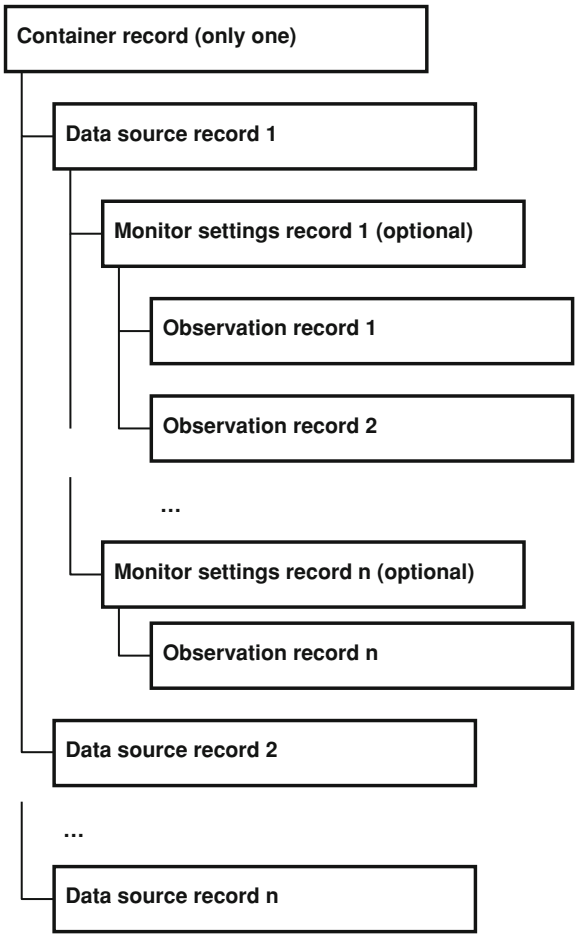


Fig. 127.3 The logical structure of PQDIF

127.2.3 Data Compression in PQDIF

Since the increase of monitoring index, especially a variety of transient process, such as impulses, swells and sags, it is necessary to adopt the high-speed data acquisition system to reflect the complete characteristics of the data. Also, the increase of sampling frequency has a strict requirement for the storage space. Therefore, under the limited storage capacity of circumstance, it needs to adopt data compression to meet the requirement of high sampling frequency and the certain storage space. PQDIF is a binary format, and it is fairly compact even uncompressed. Here are two compression methods, ZLIB and PKZIP.

a. ZLIB: it provides the library which data compression uses, and is developed by Jean-loup Gailly and Mark Adler. First Version0.9 was released on May 1, 1995. It supports the DEFLATE algorithm, one type of the LZ77 variant.

b. PKZIP: PKZIP is compression software which is introduced by PKware. The main function of PKZIP is to build compressed archive file for the specified file. Archive file can also be adjusted and repaired by PKZIP if needed.

127.3 The XML Technology

127.3.1 The Overview of the XML

XML is a markup language created by the World Wide Web Consortium (W3C). It is a standard, self-describing way of encoding both text and data. The greatest advantage of XML is that it can exchange data across a wide range of hardware, software and operating systems to achieve a variety of formats data exchange seamlessly. It gives users the power to deliver structured data from a wide variety of applications. XML is the Standard Generalized Markup Language (SGML) as well as HTML [4]; however, HTML provides a set of predefined tags, XML provides a standard, and then the user can customize the new markup language using the standard, and define specific tags. In XML the appearance of the data is specified by the application that is displaying the file or by associated style sheet file. Furthermore, it is extensible because users are free to create tags, but these tags are ignored by web browsers. As a result, XML is rapidly becoming the universal language for representing data on the internet, it is human-readable and applicable to many third-party analysis and publishing tools, and can adapt well to the needs to heterogeneous data.

XML is an ordered data file, for it is defined as a markup language rather than the data model. In order to use XML data in the application, the W3C defined the Document Object Model (DOM), which gets XML data represent to some structured data. In the DOM it defines a set of interfaces to access XML elements and attributes, and the program can use those interfaces to access or update the style, structure and content of XML file. XML has many coding tools such as DOM4J and JDOM, and the DOM4J is an open code of XML parser package produced by dom4j.org [5]. The definition of a specific markup language is called a schema. Also, XML has received a lot of research results in data compression technology. Therefore, with the extensibility provided by XML, many more attributes can be inserted to accommodate other power quality disturbances, events and analysis results.

127.3.2 The Principle of the PQDIF with XML

XML is very similar to the description of the PQDIF in these aspects about layered structure, mark and expansibility. First, from the perspective of the logical structure, the binary PQDIF file uses tags to mark each record, which is same as

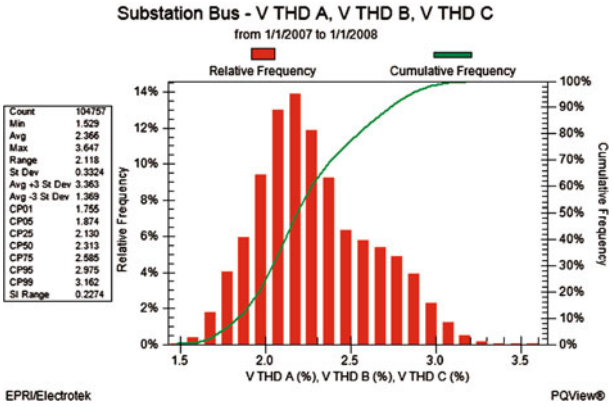


Fig. 127.4 Histogram of voltage THD

that of the presentation of XML. Second, we need to define XML representation of the physical structure of PQDIF. The format of XML elements are as follows:

```
<xsd:complexType element name=''tag1''>
 <xsd:sequence>
  <xsd:element name=''tag2'' type=''v1''></xsd:element>
 </xsd:sequence>
</xsd:complexType>
```

Among them, "tag1", "tag2" are the tag names of the elements, and "tag2" is defined as "tag1" child objects. "v1" is the attribute value, the object of elements and attributes all can be expanded. The PQDIF specification defines names for all tags and ids, for readability and convenience, XML representation follows the regulation. For example, in an observation record, "tag1" can denote "tagRecObservation", and "tag2" can denote one of its child objects, such as "ID_QT_PHASOR". Besides because the type of the PHASOR class is string, then "v1" is the declaration of string. The rest of the type of the definition is like this. So, making full use of XML technology in power quality fields, whether from the data transmission, share information or data compression, have important development prospects.

127.4 The Application of Power Quality Data with XML

Based on the above idea and principle, in recent years, to express power quality data effectively, there appears many models such as the model about comprehensive SOAP-based PQ data exchange on the internet. PQDiffractor is also in accordance with the above principle, and its emergence proves that the PQD file conversion for XML file is feasible and useful.

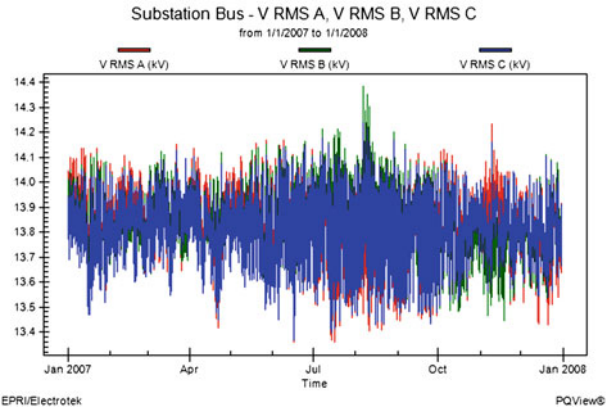


Fig. 127.5 Trend chart of voltage RMS

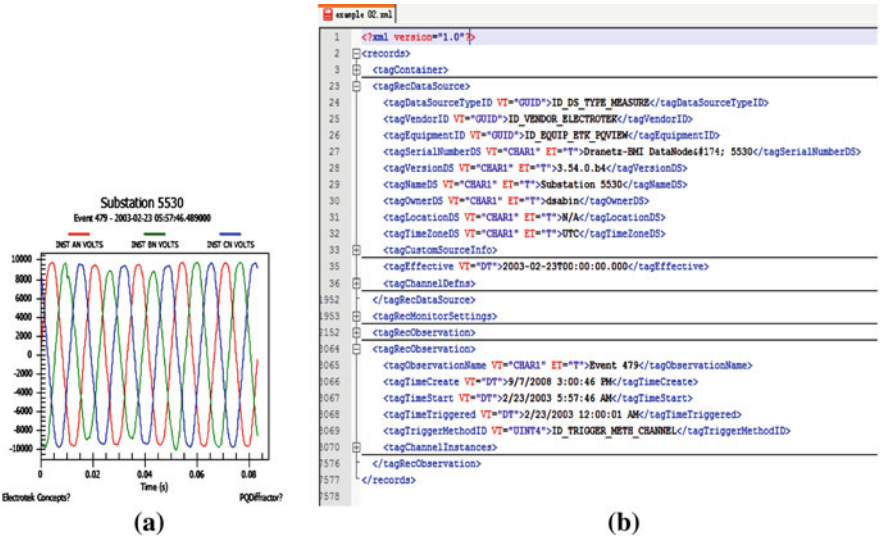


Fig. 127.6 A voltage sag measurement stored in a PQDIF file and the XML file

127.4.1 PQDiffractor

PQView is a multi-component software system for building and analyzing databases of power, power quality, and energy measurements. It integrates data from microprocessor relays, digital fault recorders, power quality monitors, smart meters and SCADA historians into an open relational database. PQDiffractor is a PQDIF file viewer utility for browsing, diagnosing and converting PQDIF files. Its functions are formidable, and the main outstanding functions are as follows:

Create interactive charts from observations with channels of any quantity type include value logs, waveforms, phasors, response, X–Y, X–Y–Z, flash density, cumulative probability frequency, histogram and 3D histogram.

Convert from native binary PQD to XML. This allows us to read or browse a PQDIF file using a text or XML parser. The example of this function will be introduced below.

Convert from native binary PQD to XML, exporting just the structure and definitions.

Determine numerous PQView compatibility problems.

127.4.2 Examples

PQDiffractor can carry out many functions, so this section presents some charts and necessary statements. The pictures display charts showing analysis of steady-state measurements, interval measurements and snapshot measurements. Figure 127.4 is a histogram. It displays one year of voltage total harmonic distortion (THD) measurements with statistics for Phase A, B and C. Figure. 127.5 is a trend chart. It displays one year of rms voltage measurements for Phase A, B and C.

Figure 127.6a is an example of PQD file, and Fig. 127.6b is the XML representation of the measurement according to the PQD file. It illustrates a voltage sag measurement stored in a PQDIF file with channel instances. As shown we choose three channel instances (VA, VB and VC). From Fig. 127.6b we can clear that it contains the data source, monitor settings and observations records for the measurement. The definitions of all the data are stored in the data source records, such as the three phase voltage, current, waveform and the maximum and minimum values. Each data source record is defined as a corresponding "tagChannelDefns". It describes the type, name, and format of the data in detail.

127.5 Conclusion

This paper is to solve the issues of multiple data sources, huge amounts of data and data sharing about power quality data. PQDIF meets all of our goals. If handled properly, PQDIF can become a widely accepted format. With appropriate software tools, using the PQDIF with XML appears to solve a number of technical issues for the power quality research. It is an urgent need to achieve the resource sharing between measuring instruments and software analysis programs. Besides, storing more data need more efficient compression method. The trend of opening and data sharing information is not only the demand of power quality monitoring but also a necessary condition of the development of smart grid, and it has laid a solid foundation for the realization of information, interaction and automation of smart grid.

Acknowledgments This project is funded by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (10KJB470003, 11KJB470005), the priority academic program development of Jiangsu higher education institutions (PAPD) and the major project subject to tender of the key research base of the college philosophy social science in Jiangsu province (2010-2-9).

References

1. IEEE Recommended Practice for the Transfer of Power Quality Data, IEEE Std 1159.3-2002
2. Li MQ, Dai YX (2006) Application of XML technology in sharing of power quality data[J]. Microcomput Inf 22:177–179
3. Zhang J, Xiao XY (2006) The design and implementation of data exchanging structure based on XML of open access power quality management system [J]. Power Syst Technol 30: 264–267
4. Ma YM, Wang Yu, Li JJ et al (2009) Accessing interface program exploitation of PQDIF date files based on visual C++ and database technology [J]. Electr Power Sci Eng 25:72–74
5. Extensible Markup Language XML, http://www.w3.org/XML/

Chapter 128
Design of Energy Management Terminal in Microgrid

Luo Yan, Deng Wei and Pei Wei

Abstract This chapter introduces an energy management terminal (EMT) designed for a microgrid. The EMT acts as a central online control and monitor unit in the microgrid. It can acquire parameters of measuring points in the microgrid regularly by communicating with other terminals or measuring instruments installed in the measuring points all over the microgrid. The EMT is developed in embedded system based on ARM micro processor. The software is programmed by C++ with QT programming language for computation and GUI display. The EMT is used in one of the smart grid demonstration project and proved to be useful.

Keywords Energy management terminal · Microgrid · ARM · QT

Introduction

With increasing development of green power generation, clean energy power station can provide large amount of power supply to local users. Microgrid as a concept is proposed by Lasseter in 2001 [1, 2], and research on it is last for years

L. Yan (✉) · D. Wei · P. Wei
Department of Advanced Electric Power System Technology,
Institute of Electrical and Engineering,
Chinese Academy of Sciences, Beijing, China
e-mail: luoyan@mail.iee.ac.cn

D. Wei
e-mail: dengwei@mail.iee.ac.cn

P. Wei
e-mail: peiwei@mail.iee.ac.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_128,

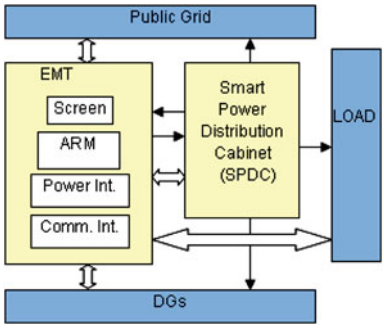


Fig. 128.1 Total design

all over the world [3–6]. In this chapter we see microgrid as a local grid that integrates clean energy power stations, storage batteries, local loads, and control units. Microgrids in remote areas can provide a large amount of green power supply to local users. In peak time of power generation, if local load is not sufficient to consume all the electrical energy, storage batteries are charged. When the climate is not good enough to provide sufficient power, such as at night or with no wind, storage batteries can provide power to users together with the public grid. In such means all the power demand can be fulfilled with utilizing green power to the largest extent. On the other hand, features of green power generations are different. Therefore it is necessary for a microgrid to integrate all kinds of power station and to provide a stable power supply with no regards of the difference.

To monitor real-time situation in a microgrid for effective control, it is necessary to have a device that can monitor the situation of the microgrid. Our research is based on such consideration and provides a design solution named energy management terminal (EMT), which is a device with function of data sampling and processing, communication, display, and local control. The EMT is integrated with a smart power distribution cabinet (SPDC) as an Energy Router in the project named "Research and Demonstration on Distributed Smart Grid Technology".

128.1 Total Design

As shown in Fig. 128.1, the EMT works together with a SPDC as a bridge among the public grid, distributed generations, and the load. The EMT is the central control and monitor unit and the PDC is an action unit to provide power supply to the loads, and can switch the power source between the public grid and the DGs according to the EMT's command. The EMT integrated with SPDC is named as Energy Router, which means the equipment can control the route from different power supplies to the load with the object of saving traditional unrecyclable energy recourses to the largest extent.

The EMT includes an ARM microchip, a display screen, power interfaces, and communication interfaces. It is developed by the ARM microchip-based embedded system [8, 9]. An ARM micro processor is used as the core chip, which is a high performance, low power consumption, low cost, small size, and steady solution for industry use. Functionally, the hardware includes an ARM chip, display screen, power interfaces, communication interfaces, etc. There are two kinds of interfaces equipped in the EMT namely power interfaces and communication interfaces. The power interfaces are used to connect to SPDC, which is a high voltage terminal for power connection among the public grid, distributed generations, and the load. Communication interfaces are used as low voltage terminal for data communication between the EMT and the public grid, distributed generations, and the power distribution cabinet. There are several kinds of communication interfaces, such as RS485 for sampling signals from the bus, Ethernet for communication with upper computer, and RS232 for uploading program and debugging.

The software is developed by C++ GUI QT 4 programming in Fedora Linux environment. The software has functions of data sampling, screen display, background processing, and data communication, which is introduced in Sect. 128.4 in detail.

128.2 Hardware Design

The hardware of the EMT is composed of three circuit boards, namely the base board, the core board, and the signal conditioner board. FLYSUN PXA270 power monitor board is used as the base board. It is embedded with PXA270 ARM10 processor produced by MARVELL Co. Ltd. The ARM10 core has a dominant frequency of 520 MHz, and can use ARM V5TE command set. It has a full feature set of MMU. With merits of high performance, low power consumption, low cost, and small size, it is widely used in embedded systems of different areas, such as communication, military application, aeronautics and aerospace, industrial control, and transportation. The FLYSUN PXA270 power monitor board is specially developed for power monitor. The diverse communication interfaces can be applied to different industrial application for data sampling, and communication.

The core board is composed of a CPU and a 32–128 M SDRARM, a 32–64 MB NorFlash as the memory for saving boot program and operating system, and a 128 MB–1 GB NandFlash as the memory for saving application program or data.

The signal conditioner board includes a 10/100 M self adaptive Ethernet interface, a USB Host (USB1.1/2.0, 12 Mb/s), a CAN 2.0a/b interface, a RS485 serial communicate interface, several digital I/Os (5 output I/Os and 4 input I/Os), a TFT LCD interface to connect to Sharp 10” crystal screen, a RTC, a JTAG interface, and a GPS interface [7].

The board is based on arm-linux core and the root file system is specially designed for PXA270. The compiler is arm-linux-gcc-3.4.6 that can compile application program, Linux core, and the root file system.

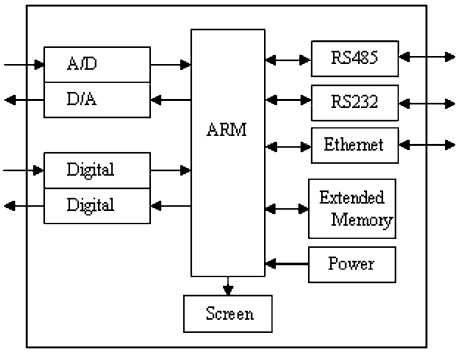


Fig. 128.2 Hardware infrastructure

Hardware system of EMT is composed of analog data sampling/convert module, digital input and output module, communication module, display module, and power supply module. The structure of the EMT hardware system and interfaces is roughly shown in Fig. 128.2.

128.3 Software Design

The software is developed by C++ with QT 4 programming language. Main function of the software includes initializing, data acquisition and processing, GUI display, and communicating with other terminals and upper computer. The software flow chart is shown in Fig. 128.3.

For data acquisition and communication, a Round-Robin scheduling algorithm is adopted. When the main program starts the thread, it will first initialize the RS485 and Ethernet interface, which are two kinds of interfaces for the communication. For every 0.2 s, the program sends a query to the bus through RS485 interface, and then reads from it to check if the captured dataset is the one replied by the demanded device. Devices connected with storage batteries, PV panels, wind turbines, and other power measurement devices [10, 11] are queried by sequence. In each round, the program will wait for 0.3 s after the query, if a new dataset from the demanded advice is captured, the program will enter a mode of data processing, precision adjustment, computation, and display. The captured dataset contains values of three- and single-phase voltage, current, active and inactive power, apparent power, power factor, etc. The display function is realized in QT framework by utilizing slot and signal mechanism. Values displayed in the screen are part of the dataset, and the full dataset is sent to upper computer through Ethernet for further processing.

Note that the values of dataset in the current round of data sampling will not be changed and remains to be the same as the former ones if new updated data are not captured in the current round.

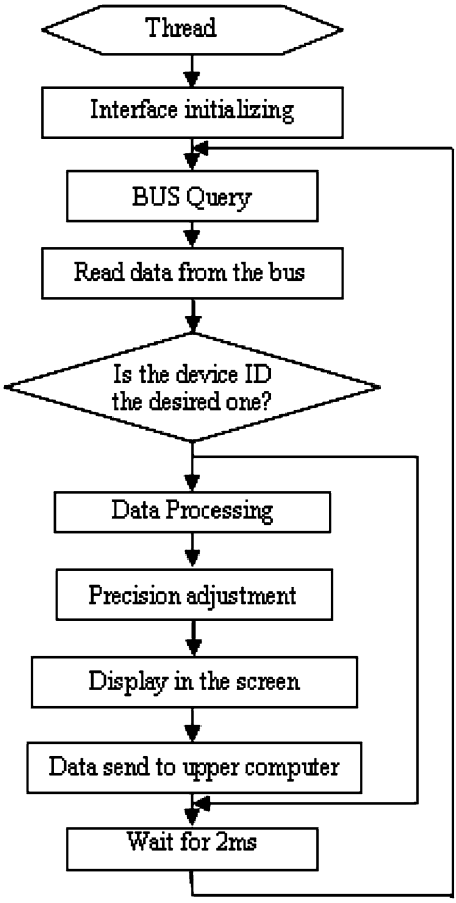


Fig. 128.3 Software flow chart

Data acquiring from the bus is conducted through RS485 interfaces. To display the dataset and other flag information in the screen, C++ GUI programming with QT4 framework is adopted. In QT4, Qdialog and its subclass is adopted to create two kinds of objects namely Label and LineEdit. The Label object is to show a picture, which includes static label and dynamic label, and the LineEdit is an object to place the current values of data in text format.

In QT Designer, we first create a dialog and then place blank objects of Labels and LineEdits. For static icons, pixmap feature of the Label object is set to show the desired image files. For dynamic icons that need to be changed according to the current situation such as arrows and switches, Signal and Slot mechanism of QT is adopted.

Signals are defined in head files such as:

signals: void label_1(QPixmap);

To emit a signal, the command of emit is used as shown below, where kai.jpg is the image file of an opened switch.

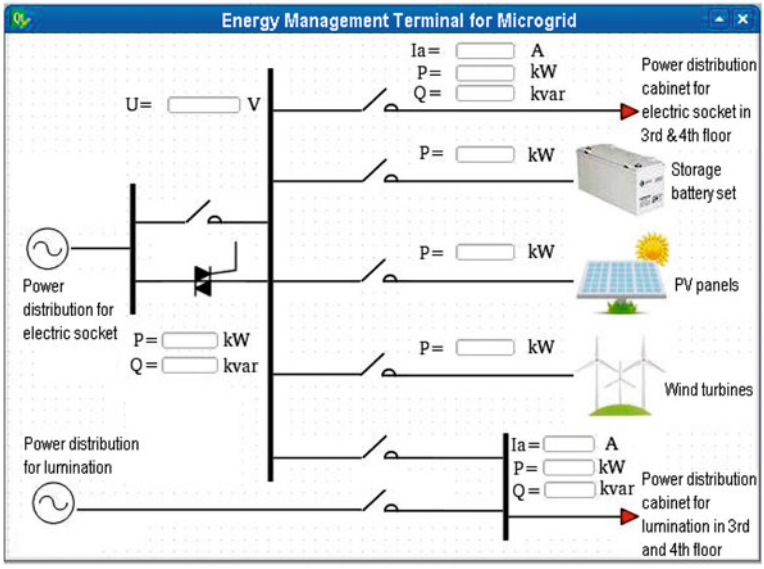


Fig. 128.4 Display screen

Fig. 128.5 Energy router

emit label_1(QPixmap(:/kai.jpg));

Slots are defined as private slots or we just use the public ones.

In main.cpp, the slot is connected with a signal, as shown below, where threadmanage is the name of the thread, and dialog is an object created for the main display dialog.

QObject::connect(threadmanage, SIGNAL(label_1(QPixmap)), dialog->label_1, SLOT(setPixmap(QPixmap)));

By the connection, whenever the emit command of a signal is executed, the connected slot will change the icon of label_1 to show the image file of kai.jpg.

For data display in the main display dialog, LineEdit object is used. An array named as UIPQvalue is adopted to save all the data, including voltage, current, active, inactive power etc., and then set it as a signal in the head file as shown below:

signals: void sendUIPQvalue(double *);

The emit command in the threadmanage.cpp file is defined as below:

emit sendUIPQvalue(UIPQvalue);

Then a function named as setTextForCalResult is defined in the .cpp file of the Qdialog, where the text of each LineEdit object is set according to the captured data. In main file, the connection between signals and slots is defined as below:

QObject::connect(threadmanage,SIGNAL(sendUIPQvalue(double*)),dialog, SLOT(setTextForCalResult(double*)));

The final screen is shown in Fig. 128.4.

128.4 Case study

In the smart grid demonstration project, a microgrid is built in our office building as shown in Fig. 128.4. The power generation comes from 20 kW PV panels, two 2 kW wind turbines, and the public grid. A 50 kW storage battery set is used as the energy storage equipment. The power demand of lighting and electric socket in third and fourth floor is treated as the load of the microgrid. The EMT is used in the Energy Router as described in Sect. 128.2. Figure 128.5 is a picture of Energy Router consists of the EMT (left side) and a Power Distribution Cabinet (right side).

Acknowledgments This research and demonstration project is sponsored by Beijing Municipal Science and Technology Commission.

References

1. Lasseter RH (2001) Microgrids (distributed power generation). IEEE power engineering society winter meeting, Columbus, Ohio, February 2001, vol 01, pp 146–149
2. Lasseter RH (2002) Microgrids. IEEE power engineering society winter meeting, New York, vol 01, pp 305–308

3. Pratt RG (2004) Transforming the U.S. electricity system. IEEE PES power systems conference and exposition, vol 3, pp 1651–1654
4. European Commission (2006) Vision and strategy for European electricity networks of the future for more information, European Smart Grids Technology Platform, ISBN 92-79-01414-5
5. Morozumi S (2006) Overview of micro-grid research and development activities in Japan. A symposium on micro-grids, Montreal, June 23
6. Tao L, Schwaegerl C et al (2011) From laboratory microgrid to real markets—challenges and opportunities. 8th international conference on power electronics, Korea, 30 May–3 June
7. Beijing Fei Xu Science Technology Co. Ltd., FS270-User Manual (In Chinese)
8. Sanhong Z (2010) Design and implementation of information interface in microgrid for distributed generation, Master Thesis, Chinese Academy of Science (In Chinese)
9. Wei D et al (2007) Design and implementation of data acquisition, communication and monitoring system for photovoltaic power station in microgrid. ISES solar world congress 2007, 18–21 Sept, Beijing, vol 3, p 1538
10. Shan dong Li Chuang Technology Co. Ltd., EDA9133A Integrated Supervisory Electric Power Control Module (In Chinese)
11. Shan dong Li Chuang Technology Co. Ltd., EDA9150A 16-channel Digital Input Module User Manual (In Chinese)

Chapter 129
Study on Low-Voltage Switchgear Reliability Based on Fuzzy Fault Tree

Jingqin Wang, Junlan Nie and Shuyu Xu

Abstract In view of the inaccurate factors for the probability of equipment failure, the reliability analysis of the low-voltage switchgear system is analyzed by fuzzy fault tree theory. According to the analysis of water pump circuit system of low-voltage switchgear using trapezoidal fuzzy number, we can obtain the fuzzy probability of return route failure and the fuzzy importance degree of bottom event. Combined with the fuzziness of probability, the project error has widespread application prospect in the reliability project.

Keywords Fuzzy fault tree · Trapezoidal fuzzy number · Fuzzy importance degree · Reliability analysis · Low-voltage switchgear

129.1 Introduction

In 1965, the USA cybernetics expert Professor L. A. Zadeh promoted the ordinary set to the fuzzy set, and the fuzzy mathematics discipline came into being. The application of mathematics was expanded from precise definition "either this or that" to the "both this and that". Then a series of project application disciplines

J. Wang (✉) · S. Xu
Hebei University of Technology, Tianjin 300130, Hebei, China
e-mail: jqwang@hebut.edu.cn

S. Xu
e-mail: shuyuxu@126.com

J. Nie
Yanshan University, Qinhuangdao 066004, China
e-mail: niejll3@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_129,

came into being. The basic idea of the fuzzy fault tree theory is that, the characteristic quantity of the top event and the importance degree of the bottom event are calculated, being used the method that defines, constructs and analyses the fault tree as well as the method and knowledge of the arithmetic fuzzy thinking about the fuzziness. In the fault diagnosis, the diagnostic process is carried on by fault tree analysis (FTA) and the diagnostic operation by trapezoidal fuzzy.

There are many uncertainty factors in the reliability engineering, only if these uncertainty factors get enough consideration, the reasonable outcome can be got to the system analysis. There are two style types for the uncertainty factors: one of these is randomness that is the event itself has certainty connotation, there is chanciness between the condition and the event only for the occurred condition is insufficiency. Another one is fuzziness or possibility that is the notion of the event itself is fuzzy, the uncertainty of the event is because of the fuzziness of the extent. Being used probability statistics, the random uncertainty can be handled; of what the traditional reliability theory made a study is this uncertainty. The fuzziness is another type of uncertainty different to the randomness. It is the transition of the objective event's difference that is the event is unclear and occurs in the event's some states, phenomenon, parameter and their relationship. To some uncertainty events such as having the phenomenon that the factors have more subjective influence, the data is not complete and so on, it is suitable to use fuzzy method what is created by L. A. Zadeh that is based on trapezoidal fuzzy [1, 2].

129.2 Fuzzy Set Theory

Fuzzy set theory is used to deal with inexactitude or fuzzy problem. Its root idea is to make the absolute membership in the classical set more flexible or called fuzzier [3]. It is expressed in the characteristic function that, the grade of membership of the element set A is not limited in 0 or 1 and it can take any value from 0 to 1, this value reflects the degree of element x belonging to the set A.

Definition Set up U as a domain made up of objects, and then defines a fuzzy set $\tilde{A}$ that in the domain U as a membership function:

$$\mu\tilde{A}(X) : U \to (0,\ 1), \quad X \in U \tag{129.1}$$

It parallels elements in U to the field of real number (0, 1), marked as $\tilde{A} = \int_{X\in U} \mu\tilde{A}(X)/_X \mu\tilde{A}(X)$ is gone by the name of degree that element x in the domain U belonging to the fuzzy set. The more larger value of the $\mu\tilde{A}(X)$, the more deeply that X belongs to $\tilde{A}$. In order to write easily, the membership function is simply marked as $\mu A(x)$.

The definition of the Trapezoidal fuzzy number: Set up the domain U as the field of real number, q declared as the one-dimensional trapezoidal fuzzy number

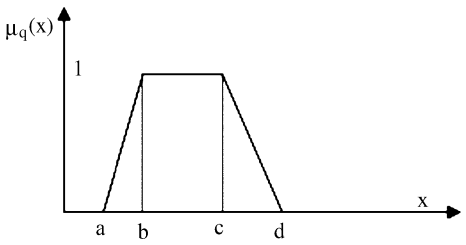


Fig. 129.1 The trapezoidal fuzzy number

and $\mu_q(x)$ as its membership function (Fig. 129.1). The membership function is expressed by the formula:

$$\mu_q(x) = \begin{cases} \frac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \frac{x-c}{d-c} & c \le x \le d \end{cases} \tag{129.2}$$

129.3 The Reliability Analysis of Fuzzy Fault Tree

129.3.1 The Fuzzy Fault Tree Analysis Based on Failure Rate Being Fuzzy Number

In the traditional reliability theory, the relationship between the reliability and the failure rate [4]:

$$R(t) = e^{-\int_0^1 \lambda(t)\mathrm{d}t} = \exp\left(-\int_0^1 \lambda(t)\mathrm{d}t\right) \tag{129.3}$$

The above formula is the reliability function's general expression. When $\lambda(t)$ a constant value is ($\lambda(t)$ is numeric constant), the reliability formula:

$$R(t) = e^{-\lambda t} \tag{129.4}$$

Then, the relationship between the unreliability and the failure rate:

$$F(t) = 1 - e^{-\lambda t} \tag{129.5}$$

When the failure rate is a fuzzy number, the relationship between the fuzzy unreliability and the fuzzy failure rate is:

$$\overline{F(t)} = 1 - e^{-\bar{\lambda} t} \tag{129.6}$$

For only the fuzzy fault tree analysis is discussed here, in order to make it expedient, marked as $F(t) = 1 - e^{-\lambda t}$.

When set up the failure rate of the component as trapezoidal fuzzy number:

$$\mu_q(x) = \begin{cases} \frac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \frac{x-c}{d-c} & c \le x \le d \end{cases} \tag{129.7}$$

(a, b, c, d) in (129.7) is the parameter of trapezoidal fuzzy number, make $R(t) = e^{-\lambda t}$ as the reliability function, based on extension principle:

$$\mu_{-\lambda t}(x) = \begin{cases} -\frac{x+at}{t(b-a)} & -bt \le x \le -at \\ 1 & -ct \le x \le -bt \\ \frac{x+ct}{t(c-d)} & -dt \le x \le -ct \end{cases} \tag{129.8}$$

The membership functions of the fuzzy reliability:

$$\mu_{e^{-\lambda t}} = \begin{cases} -\frac{\ln x+at}{(b-a)t} & -bt \le x \le -at \\ 1 & -ct \le x \le -bt \\ \frac{\ln x+ct}{t(c-d)} & -dt \le x \le -ct \end{cases} \tag{129.9}$$

In the classical fault tree analysis theory, the whole minimal cut sets of the fault tree are assumed: C1, C2,…, CNk, when the minimal cut sets do not intersect with each other, the probability of the top event:

$$P_T = F_s(t) = \sum_{j=1}^{Nk} \left(\prod_{i \in Cj} F_i(t) \right) \tag{129.10}$$

129.3.2 The Quantitative Analysis of The Fuzzy Fault Tree

Not all the components in the system have the same importance; some components' breakdown will make the whole system breakdown, while others will not. Generally, the contribution that a component or a minimal cut set's breakdown tends to the top event called importance degree. It is the function of the system structure, the component's reliability and the time. Having the data of the each component's quantitative importance degree, the weak link of the system design will be got, and then the performance of the system will be improved. In the reliability engineering, sometimes mass data will not be got, the probability of the top event is fuzzy and the system's analysis is carried on by the fuzzy importance degree.

Link the structure function of the system as $\Phi(X) = \Phi(x1, x2, \ldots xn)$, the fuzzy probability of the top event as Pxi, the membership function as $\mu_{p_{x^i}}, \mathrm{i} = 1, 2, \ldots, \mathrm{n}$, the fuzzy probability of the top event as $P(Px1, Px2, \ldots, Pxn) \equiv PT$ and it is membership function as μ_{pT}, based on extension principle [5, 6]:

$$\mu_{p_T}(y) = \begin{cases} \text{SUP}\min[\mu_{p_{x_i}}, \ldots\ldots\mu_{p_{x_n}}] \text{ if } \phi^{-1}(y) \neq \phi \\ x \in \phi^{-1}(y) \\ 0 \qquad\qquad\qquad \text{Other} \end{cases} \tag{129.11}$$

Mark: the failure fuzzy probability of the top event as $P(Px1,\ldots, Pxi-1, 1, Pxi+1,\ldots, Pxn) = PTi1$, when the component i in case of $xi = 1$, as $P(Px1,\ldots, Pxi-1, 0, Pxi+1,\ldots, Pxn) = PTi0$, when the component i in case of $xi = 0$, then the fuzzy importance of the component i is defined as:

$$FI_i = E[P_{Ti1} - P_{Ti0}] = \frac{\int_0^1 x\mu_{p_{T_{i1}}}(x)\mathrm{d}x}{\int_0^1 \mu_{p_{T_{i1}}}(x)\mathrm{d}x} - \frac{\int_0^1 x\mu_{p_{T_{i0}}}(x)\mathrm{d}x}{\int_0^1 \mu_{p_{T_{i0}}}(x)\mathrm{d}x} \tag{129.12}$$

Modified Fuzzy Importance:

$$FI_{ic}\left(P_{Ti}^1, P_{Ti}^0\right) = \left|a_{Ti}^1 - a_{Ti}^0\right| + \left|b_{Ti}^1 - b_{Ti}^0\right| + \left|c_{Ti}^1 - c_{Ti}^0\right| + \left|d_{Ti}^1 - d_{Ti}^0\right|. \tag{129.13}$$

129.4 Case Analysis

Low-voltage switchgear water pump circuit system offer 380 AC voltage for the water pump through Breaker, AC contactor and so on. The fault tree is shown as Fig. 129.2.

According to the fault tree of the water pump circuit with low-voltage switchgear, fuzzy probability of bottom events and the probabilistic formula of top events, get the fuzzy importance degree of the bottom event. At the same time, we list the probability importance degree of the bottom event which is calculated with the classical approach to compare, the results are listed in Table 129.1.

The mathematical signification of fuzzy importance degree is: the fuzzy importance degree of the component i is the mathematical expectation of the difference between the top events' fuzzy probability when the component i value 1 and 0. In physical sense, fuzzy importance degree is the "rose" of the system's fuzzy unreliability result in the unreliability of the component i. The dimensions of its "rose" reflects the "contribution" and reflection of component i in the whole system's unreliability, in another words, it reflects the importance of component i in the whole system's reliability. Here "rose" is used quotation marks, for the system's fuzzy unreliability is a fuzzy probability, and the dimension of its rose is a fuzzy probability too. It is more appropriate to describe it with mean value. According to above calculation, the result of the probability importance degree calculated with classical approach is close to its calculated trapezoidal fuzzy number. The reason is that though the probability of bottom event happened is a trapezoidal fuzzy number, the signification of trapezoidal fuzzy number is based on probability statistics and empiric value.

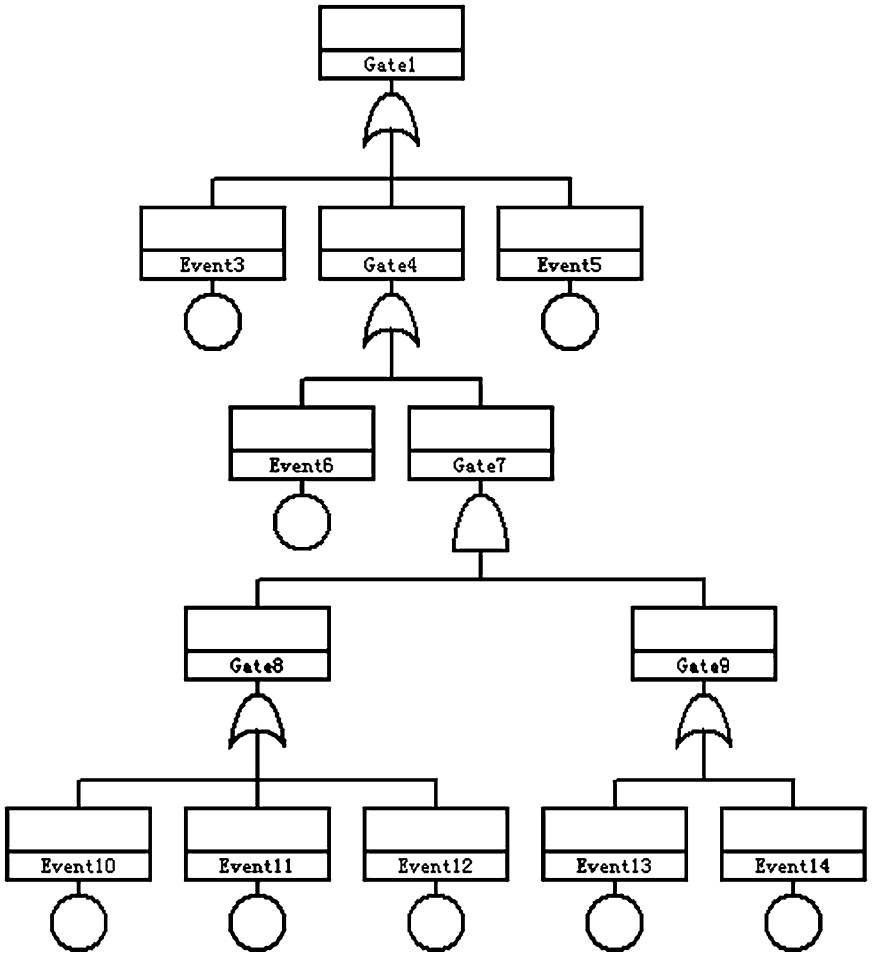


Fig. 129.2 Fault tree of the water pump circuit

Table 129.1 Importance degree of low-voltage switchgear

Event	Trapezoidal fuzzy importance degree	Probability importance degree
Event3	(0.452, 0.506, 0.543, 0.592)	0.507
Event5	(0.276, 0.314, 0.361, 0.401)	0.339
Event6	(0.276, 0.314, 0.361, 0.401)	0.339
Event12	(0.235, 0.268, 0.290, 0.324)	0.284
Event10	(0.015, 0.018, 0.020, 0.024)	0.020
Event11	(0.015, 0.018, 0.020, 0.024)	0.020
Event13	(0.033, 0.036, 0.039, 0.041)	0.040

129.5 Conclusion

According to the above all, when fuzzy set theory be used to analyze the system based on fault tree, in which uses fault rate as the fuzzy number, it could overcome the disadvantages such as classical approach hardly evaluated accurately better. This technique not only reflects the fuzziness of the probability, but also admits some error to a certain extent of the evaluation in the probability. At the same time, it combined some experiment data with experiences of the engineer and technicians so can be widely used in the reliability project.

Acknowledgments This work supported by the National Natural Science Foundation of China (No. 51077039) and Specialized Research Fund for the Doctoral Program of Higher Education, China (No. 20070080006).

References

1. Singer D (1990) Fault tree analysis based on fuzzy logic. Comput Chem Engineer 3:259–266
2. Shi Z, Xingui H, Yu J (1999) A method of reliability analysis based on fuzzy number arithmetic operations. J Syst Eng Electron 9:74–77
3. Wang Y, Yu W, Zhuang Z (2000) Study on fuzzy fault tree analysis method based on fault rate as a fuzzy numbers. J Syst Eng Theory Pract 12:102–107
4. Lu J, Wang J (2004) Reliability theory and application of low voltage protection electrical apparatus. Machinery Industry Publishers, Beijing (in press)
5. Li Q, Lu T (2000) Study on fuzzy importance degree analysis method. Fuzzy Sys Math 1:89–93
6. Lian J, Hua X, Li A (2002) Determination of fuzzy probability on event. J Taiyuan Heavy Mach Inst 2:152–155

Chapter 130
An Innovative Protective Algorithm for UHV Power Transformer

Xiang-li Deng, Chuan-qi Wang, Zhe Zhang and Xiang-gen Yin

Abstract Transformer differential protection is influenced seriously by harmonics in inrush current and the operation time is long. In order to solve this problem, this paper introduces a T-type equivalent model of transformer protection. The parameters of the model can be calculated uniquely and timely at steady state. A new transformer protection algorithm is presented. The algorithm is based on the balanced equations of T-type equivalent model and the variation of the measured excitation inductance. With EMTP simulation tool, a simulation platform for the first UHV power transmission system in China is established. Situations such as energizing, external faults, internal faults and normal operations in the system are covered in the simulation. The operation time of the proposed protective scheme is short and the scheme has high sensitivity.

Keywords Inrush current · On-line least-squares identification · UHV transformer protection · Measured excitation inductance

X. Deng (✉) · C. Wang · Z. Zhang · X. Yin
School of Electrical and Electronics Engineering,
Huazhong University of Science and Technology,
Wuhan 430074, Hubei, China
e-mail: dengxl989@hotmail.com; dengxiangli@dongfang-china.com

C. Wang
e-mail: chuanqiw@sina.com

Z. Zhang
e-mail: zz_mail2002@163.com

X. Yin
e-mail: xgyin@mail.hust.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_130,

130.1 Introduction

Large UHV transformers are essential and important elements of UHV power transmission system. An UHV transformer's reliability and security have been considered to be a key issue for electric utilities. When an UHV transformer experiences a fault, it is necessary to take the transformer out of service as soon as possible. Accordingly, higher demands are imposed on UHV power transformer protective relays than on EHV power transformer. The requirements include prompt reaction during internal fault and security during external fault [1]. A number of researchers have long been engaged in search of innovations to discriminate between internal faults and inrush [2–6]. When transformer is energizing, the algorithms, which employ harmonic components to discriminate between magnetizing inrush and internal faults have to wait to trigger until harmonic components decay completely if a slight turn-to-turn fault exists.

The algorithms, which use the ratio of harmonic components or unsymmetrical waveform of differential currents to identify magnetizing inrush, are widely used in digital transformer protective relays [2–5]. They only rely on currents with at least one cycle data window.

Recently, the algorithms based on transformer models have been reported in the literature [7–12]. Combination of current and voltage is used to identify magnetizing inrush. These algorithms do not rely in the presence of harmonic components to identify magnetizing inrush and have further advantages. The Algorithms of references [7, 8] rely on the measured inductance of magnetizing branch of transformer to identify inrush. Reference [9] relies on unbalanced equations of equivalent circuit of transformer to identify internal and external faults. The Algorithms of references [10–12] identify equivalent circuit parameters and rely on the variation of parameters to identify internal faults.

However, problems associated with the algorithms based on transformer models have not been solved yet. As discussed in [10], at steady-state, On-Line Least-Squares Identification Algorithm (OLSIA) is invalid because the estimation array of |ATA| is 0. The Algorithms of [11, 12] are not sensitive enough to detect the slight turn-to-turn faults in transformer.

In attempt to solve the above problems, this paper introduces a new algorithm. The algorithm is sensitive to slight turn-to-turn faults and detects inner faults fast due to its less-than-half-cycle fault detecting time. It does not rely on harmonic components to identify inrush. To accomplish this task, a two-parameter-winding model is established. Based on the model, parameters are estimated by OLSIA in steady-state operation condition.

This paper consists of two parts. In the first part, based on the analysis of convergence property of OLSIA, a T-type equivalent circuit model is proposed. The parameters can be calculated in a unique and real-time manner. In the second part, a digital algorithm, which is either sensitive to detect slight turn-to-turn faults or heavy terminal faults of transformer, is proposed.

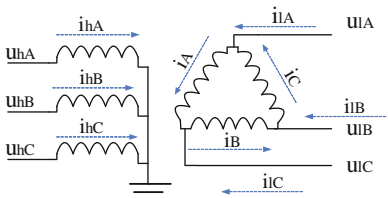


Fig. 130.1 Y0/Δ winding diagram for transformer

130.2 T-Type Equivalent Circuit of Transformer

130.2.1 Convergence of OLSIA for Transformer Model

As illustrated in Fig. 130.1, the transformer model for Phase A is described by Eq. (130.1):

The simplified stable equation of three windings of transformer is as Eq. (130.1):

$$\begin{cases} R_h - k_{21}R_l = (1-k_{12})U\cos\alpha \\ L_{h\sigma} - k_{21}L_{l\sigma} = \dfrac{(1-k_{12}U\sin\alpha)}{\omega} \end{cases} \tag{130.1}$$

In the case with OLSIA, it is clear from Eq. (130.1) that the parameters of R_h, R_l, $L_{h\sigma}$, $L_{l\sigma}$ cannot be obtained when transformer operates stably. If there are only R_h and $L_{h\sigma}$ in equation, the stable values of parameters are estimated with good accuracy and convergence.

130.2.2 T-Type Equivalent Circuit of UHV Transformer Model

As illustrated in Fig. 130.2, the main transformer of UHV transformer system is an autotransformer. The value of short-circuit impedance is x_k, which is provided by manufacturer. The equation is given by (130.2).

$$R_h i_L + L_{h\sigma}\frac{\mathrm{d}(i_h(1+k_{21}) - k_{21}i_L)}{\mathrm{d}t} = u_h - u_L - k_{21}u_L + k_{21}L\frac{\mathrm{d}(i_h - i_L)}{\mathrm{d}t} \tag{130.2}$$

where, $L = \frac{x_k}{\omega}$, $R_l = \frac{R_h}{k_{21}}$

As the values of R_1, L_1 illustrated in Fig. 130.3 can be estimated with OLSIA by Eq. (130.2), inductance L_k is calculated in Eq. (130.3).

$$u_h = i_h R_1 + L_{h\sigma}\frac{\mathrm{d}i_h}{\mathrm{d}t} + L_k\frac{\mathrm{d}(i_h + i_l)}{\mathrm{d}t} \tag{130.3}$$

Where: $L_k = \frac{L_m L_s}{L_m L_s}$, $i_h + i_l$ is differential current.

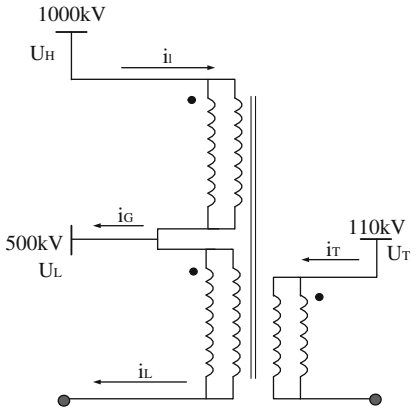


Fig. 130.2 Winding connection diagram for UHV main transformer

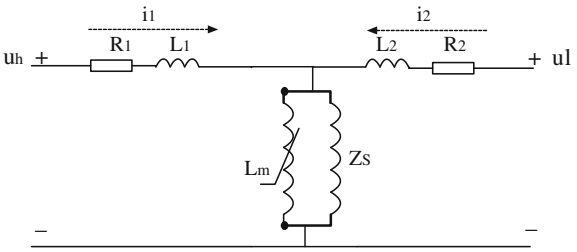


Fig. 130.3 Transformer equivalent model for turn-to-turn faults

130.3 Transformer Protection Criteria

The values of resistance and leakage inductance can be estimated by the recursive OLSIA algorithm in real time. The inductance value L_k of magnetizing branch can be computed by Eq. (130.3). Therefore, all parameters of transformer model are determined. The parameters estimated by Eq. (130.2) can be substituted into different phase circuit equations to form error functions. Taking phase A as an example, its parameters can be substituted into the equation of the phase C and the equation is given by (130.4):

$$\begin{aligned}\delta = {} & u_{hA} - u_{hB} - u_{lA}{}' + u_{lB}{}' - \frac{x_k}{\omega}\frac{\mathrm{d}i_{lA}{}'}{\mathrm{d}t} \\ & - R_h(i_{hA} - i_{hB} - i_{lA}{}') + L_{h\sigma}\frac{\mathrm{d}(i_{hA} - i_{hB} - i_{lA}{}')}{\mathrm{d}t}\end{aligned} \tag{130.4}$$

If the average value $\sum \delta > \sum \delta_{zd}$, the fault is identified as a heavy internal fault and the protective relay issues trip command. A counter of D_k is adopted to reflect the dots of "significant rate of change" for the inductance L_k. If $|L_k(k+1) - L(k)| > \mathrm{d}L_{kzd}$, D_k is increased by 1. If Eq. (130.5) is satisfied, turn-to-turn fault is classified. The logic of the algorithm is described in Fig. 130.4.

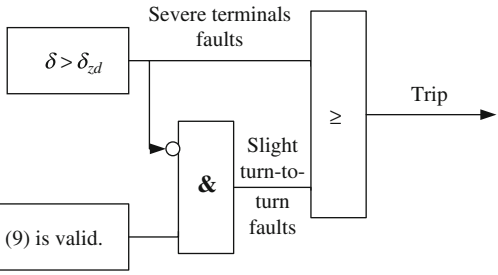


Fig. 130.4 Transformer protection tripping logic based on OLSIA

$$\begin{cases} \frac{2}{N}\sum_{k=1}^{N/2}|L_k(k)| < L_{zd} \\ D_k < D_{zd} \end{cases} \tag{130.5}$$

130.4 Simulation Results

130.4.1 Simulation Model

Figure 130.5 is the dynamic simulation system of UHV power transmission system established in HUST laboratory according to the first economic UHV transmission system operated in China. The parameters of simulation model used by this paper are derived from this dynamic simulation system as discussed in [13]. The simulation is carried out with program EMTP. The transformer simulated in the system is a three-single-autotransformer group of T1 illustrated in Fig. 130.5 with the winding connection of star-grouded/star-grouded/delta. The parameters of the single transformer model are described as below:

rated capacity: 1,000 MVA
rated voltage of high winding: 1,050 kV
rated voltage of low winding: 525 kV
rated voltage of tertiary winding: 110 kV
short-circuit impedance of high winding: 18%
low winding: 62%
tertiary winding: 40%.

130.4.2 Simulation Results

Assumption of the errors of OLSIA are made as 20% for winding (because the winding resistances are too small to be identified accurately) and 10% for the errors of winding leakage inductances. The setting of unbalanced output of equivalent circuit model equations is $\sum|\delta_{ZD}| = 0.4$, the counter setting of the rate

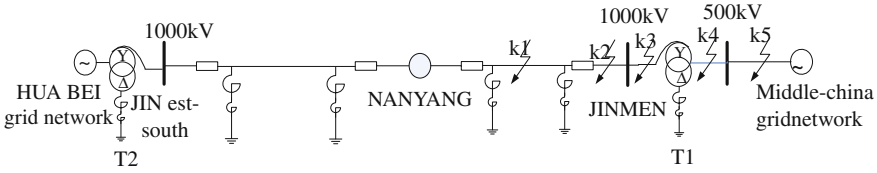


Fig. 130.5 UHV power transmission system model

Table 130.1 Values of transformer protection criteria in various cases

Operation conditions	Number	Fault type	$\sum \lvert\delta\rvert$	D_k	$\sum \lvert L_k\rvert$	Result	Detect time (ms)
Normal operation	1	No	0.205–0.255	1	5.1e4	Normal	
Energizing	2	No	0.0002–0.005	9	1,400–3,200	Normal	
Switch onto internal faults	3	Turn-to-turn 2.2%	0.054–0.062	3	0.53–0.88	Abnormal	10
	4	4.58%	0.053–0.066	3	0.55–0.87	Abnormal	10
	5	9.1%	0.052–0.064	3	0.5–0.8	Abnormal	10
	6	Single phase grouded	2.8–3.5	2	0.65–1.05	Abnormal	
	7	Phase to phase	2.9–3.5	2	1.2–2.1	Abnormal	
	8	Turn-to-turn 2.2%	0.11–0.136	1	7.6–12.8	Abnormal	10
Internal fault	9	4%	0.11–0.135	1	3.5–5	Abnormal	10
	10	7.58%	0.102–0.128	1	1.97–3.15	Abnormal	10
	11	9.08%	0.102–0.1277	1	1.85–3.19	Abnormal	10
	12	Single phase grouded	0.6–0.82	0	1.7–2.5e9	Abnormal	10
	13	Phase to phase	1.740–2.453	1	6.6–5.1e8	Abnormal	10
External fault	14	Single phase grouded	0.204–0.255	0	220–432	Normal	

of change of magnetizing branch inductance is $D_{ZD} = 7$, and average values setting of magnetizing branch inductance is $\sum |L_{ZD}| = 100$. The simulation results are illustrated in Table 130.1.

When UHV transformers are energized, the low voltage windings and tertiary windings are fed by zero current. Therefore, current i_L in Eq. (130.2) is 0. The item associated with the resistance in Eq. (130.2) dose not affect unbalanced output of equivalent circuit model equations. It is observed (line 2 in Table 130.1) that the

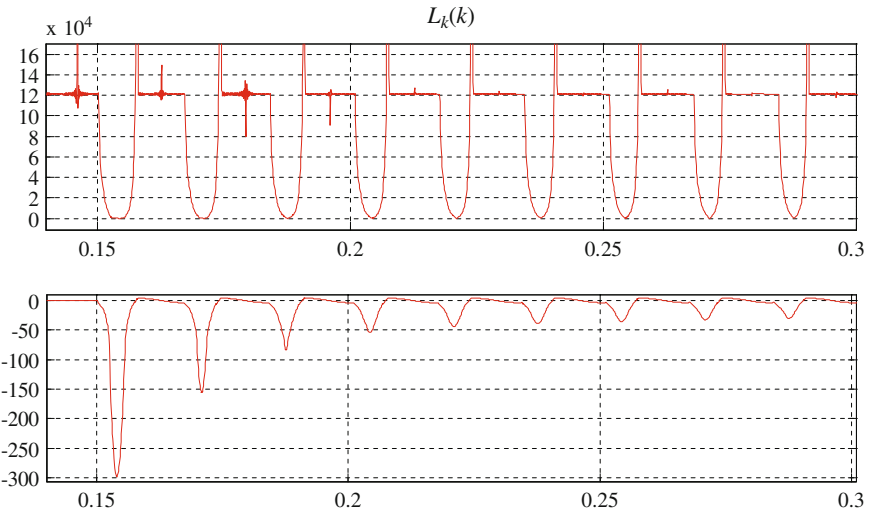


Fig. 130.6 Values of inductance of magnetizing branch of transformer

value of $\sum |\delta|$ reaches the smallest one of 0.0002, which is far less than the setting with no trip command being issued. From line 2 in Table 130.1, it can also be observed that the value of $\sum |L_k|$ is far more than the setting without trip command being issued as well.

Figure 130.6 shows when the sharp waveform occurs in the inrush current, B/H curve of corn works in linear saturation zone. The measured magnetizing branch inductance value is small and near to 0. It also shows when the discontinuous current occurs in inrush current waveform, B/H curve of corn works in unsaturated zone and the measured magnetizing branch inductance value is much bigger.

When transformer operates normally, the low voltage winding and the tertiary winding are fed by load currents. Errors of resistance in Eq. (130.2) dose affect unbalanced output of equivalent circuit model equations much more severely. From line 1 in Table 130.1, it is observed that unbalanced output can be 0.255, which is less than the setting and no trip command is issued.

When turn-to-turn faults occur (including switching onto turn-to-turn faults and turn-to-turn faults when transformer operates normally) because the unbalanced outputs of equations are less than the setting, the fault cannot be identified if the algorithm only relies on unbalanced output criteria (seen from line 3 through line 5 and line 9 through 11 in Table 130.1).

However, differential currents, used by the transformer protective relay, flow through magnetizing branch of T-type circuit. Magnetizing inductance can be measured correctly. The measured values of magnetizing inductance are far less than the setting, and thus the turn-to-turn faults can be identified reliably.

When the transformer is switched onto turn-to-turn faults, the item associated with resistance in Eq. (130.2) is 0 and the maximum value of the unbalanced output is only 0.066 (seen from line 4). When the turn-to-turn faults occur in normal

condition, the resistance item in circuit equation is not 0. The unbalanced output can reach 0.136 (seen from line 8), however, is still far less than the tripping setting.

130.5 Conclusion

This paper analyzes the convergence of equations of equivalent circuits when transformer operates in stable states and presents a T-type equivalent circuit model of transformer with parameters being calculated on-time. This paper also introduces a new algorithm to detect severe and slight faults in transformer based on the T-type equivalent circuit. The algorithm is capable to detect any faults in less than 10 ms. It has been shown in simulation that the new algorithm can detect any internal faults in transformer and does not issue wrong trip commands in situations such as energizing, normal operation and extern fault. Extensive experimental results validate the proposed model and algorithm and feasibility for field applications.

References

1. Lui Z (2005) Ultra-high voltage grid, 3rd edn. China Economic Publishing House, Beijing
2. He B, Xu X (1998) Protection based on wave comparison. Proc CSEE 18(6):395–398
3. Zhang X, He B (2005) A new method to identify inrush current by phasor comparison. Proc CSEE 25(19):43–47
4. Jiao S, Liu W (1999) A novel scheme to discriminate inrush current and fault current base on integrating the waveform. Proc CSEE 19(8):35–38
5. Xu Y, Wang Z, Yang Q (2004) Research on novel transformer protection based on the characteristics of voltage and differential current. Proc CSEE 24(2):61–65
6. Lin X, He Z et al (2001) Discussions on some aspects of sampling value differential current protection. AEPS 10(10):27–33
7. Wang Z, Xu Y, Wang X et al (2003) A novel scheme based on flux restraint theory used in distinguishing inrush currents for UHV transformers. Proc CSEE 23(12):52–58
8. Zong H, Jin H, Zhu Z et al (2001) Transformer inrush detected by the variation of magnetizing impedance. Proc CSEE 21(7):91–94
9. Wang Z, Xu Y, Wang X et al (2003) Study on the novel transformer protection principle based on the transformer model. Proc CSEE 23(12):54–58
10. Suonan J, Jiao Z, Kang X (2008) Algorithm to identify leakage inductances of power transformer with Y-Delta connection. Proc CSEE 28(13):84–90
11. Xiong X, Deng X, You B (1999) Transformer protection using parameter identification method. AEPS 23:18–21
12. Suonan J, Kang X, Song G (2007) Survey on relay protection using parameter identification. Proc CSU-EPSA 19(1):14–20
13. Zeng L, Lin X et al (2010) Modeling and electromagnetic transient simulation of UHV autotransformer. Proc CSEE 30(1):91–97

Chapter 131
Partial Discharge Detection in XLPE Cable Joint Based on Electromagnetic Coupling Method

Rijun Dai, Fangcheng Lv and Heming Li

Abstract Aiming at the partial discharge (PD) characteristics of XLPE cable, a set of PD detection system which bases on the wideband current sensor is developed. The amplitude-frequency characteristics of the wideband current sensor and its influencing factors such as iron core material, permeability, turns of the coil and integrating resistor are analyzed by experiments. Finally, four typical partial discharge models are designed, and the PD detection system is utilized to detect them. The sensor signals corresponding to different partial discharge models are analyzed.

Keywords XLPE cable · Partial discharge · Electromagnetic coupling

131.1 Introduction

High voltage XLPE cable has been widely used in power system because of its good electrical properties, simple structure and short manufacturing cycle. Operation experience and research achievements show that Partial Discharge (PD) is the main reason of insulation breakdown to the XLPE cable. The cable joint is weak link of the insulation structure in cable lines [1], therefore, PD detection of the XLPE cable joint is the important means to find potential failures timely [2].

Currently electrical and non-electrical measurements are two main methods for cable PD detection [3–5]. There are many kinds of electrical measurements such as

R. Dai (✉) · F. Lv · H. Li
Hebei Provincial Key Laboratory of Power
Transmission Equipment Security Defense, North China Electric Power University,
Baoding 2071003, China
e-mail: electricpower@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_131,

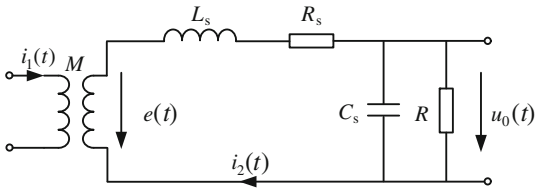


Fig. 131.1 The equivalent circuit of the wideband current sensor

capacitance coupling method, electromagnetic coupling method [6] and UHF method. The electromagnetic coupling method uses the wideband current sensor, which can be easily installed at the cable terminal or the earth line nearby the transition joint, what's more, there is no direct contact between measuring circuit of the sensor and high voltage cable, which can suppress the noise better [7].

In this paper, a set of PD detection system which bases on wideband current sensor is developed. The amplitude-frequency characteristics and its influencing factors of the wideband current sensor are analyzed by experiments. Finally, four typical partial discharge models are developed, and the PD detection system is utilized to detect them. The sensor signals corresponding to different partial discharge models are analyzed.

131.2 Wideband Current Sensor

The wideband current sensor is the Rogowski coil with high frequency core. The equivalent circuit is shown in Fig. 131.1 [8].

In Fig. 131.1 L_s and R_s represent self-inductance and equivalent resistance of the coil, C_s is the stray capacitance, M is the mutual inductance of the coil and R is the self-integrating resistance. R and L_s form the self-integrating circuit and $u_0(t)$ can be got from R. The sensitivity $G(s)$ of the Rogowski coil is shown in (131.1), in which N is the turns of the coil.

$$G(s) = R/N. \tag{131.1}$$

As the equivalent circuit of current sensor is similar to parallel resonance loop of high frequency small-signal [9], we can use the theory to get the low limiting frequency f_L, upper limiting frequency f_H and working frequency band f_B as

$$f_L = \frac{1}{2\pi}\frac{R+R_S}{L_S+RR_SC_S} \approx \frac{1}{2\pi}\frac{R+R_S}{L_S} \tag{131.2}$$

$$f_H = \frac{1}{2\pi}\frac{L_S+RR_SC_S}{L_SR_SC_S} \approx \frac{1}{2\pi}\frac{1}{RC_S} \tag{131.3}$$

$$f_B = f_H - f_L = \frac{1}{2\pi}\left(\frac{L_S+RR_SC_S}{L_SR_SC_S} - \frac{R+R_S}{L_S+RR_SC_S}\right) \tag{131.4}$$

Table 131.1 Parameters of the current sensors

No.	Core material	Relative permeability, μ	Turns, N	Resistance, R (kΩ)
1	Ni–Zn	200	10	1
2	Ni–Zn	100	10	1
3	Ni–Zn	40	10	1
4	Mn–Zn	2,000	10	1
5	Ni–Zn	200	10	2
6	Ni–Zn	200	20	1

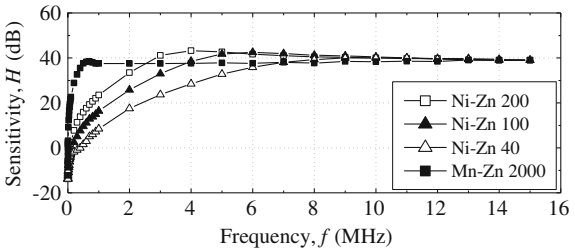


Fig. 131.2 Amplitude-frequency characteristics of sensors 1–4

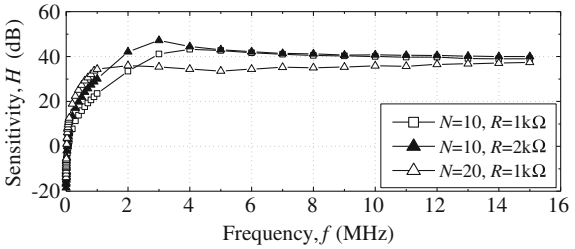


Fig. 131.3 Amplitude-frequency characteristics of sensors 1, 5 and 6

As can be seen from (131.1) to (131.4), the frequency characteristics of the coil is determined by L_s, R and the turns N, while L_s is determined by the iron core material and relative permeability μ. As the PD signal is weak and has a large frequency range, ferrite materials are chosen as the core material, then there will be an optimum R and N which make the working frequency band wide and the sensitivity high. In this paper, six current sensors are designed. The parameters are shown in Table 131.1.

A HP 33120A signal generator is used to produce a current of which the frequency is adjustable from 10 Hz to 15 MHz and make it go through the center of sensor. We can get the amplitude-frequency characteristics of each coil (shown in Figs. 131.2, 131.3). Figure 131.2 shows the sensors 1–4, while sensors 1, 5 and 6 are shown in Fig. 131.3.

Figure 131.2 shows that the low limiting frequency of Mn–Zn coil is about 0.5 MHz while the low limiting frequencies of Ni–Zn coils are more than 2 MHz, and the low limiting frequency of the sensor will rise with the decrease

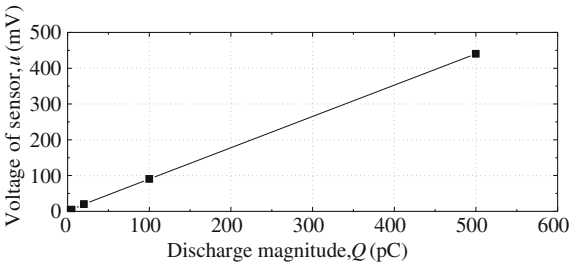


Fig. 131.4 Calibration curve under different discharge magnitudes

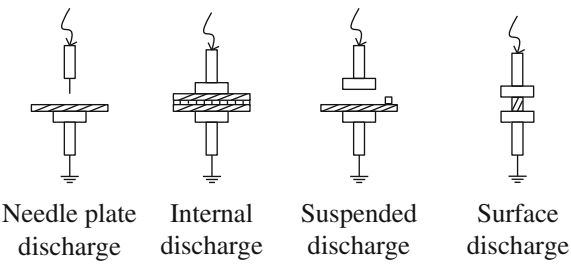


Fig. 131.5 Four typical discharge models

of permeability. Theoretical and experimental studies have shown that the frequencies of PD pulse signals in XLPE cable are mainly within 100 MHz, while frequencies of noise interference are under 1 MHz [10, 11]. It can avoid interference of the noise effectively and improve ability of anti-interference and SNR of the sensor by choosing Ni–Zn coil [12].

Figure 131.3 shows that the sensitivity of sensor rises with the increase of R, however, the working frequency band decreases obviously from (131.2) and (131.3). Figure 131.3 also shows that the sensitivity will descend with the increase of N. Therefore, the Ni–Zn ferrite which permeability is 200 is chosen as the iron core, the number of the turns is 10 and self-integrating resistance is 1 kΩ, which ensures the high sensitivity and wide working frequency band.

The calibration unit of LDS-6 PD detector of LDIC Company is utilized to calibrate the current sensor. Figure 131.4 shows the calibration curve of the current sensor.

131.3 Experiment and Result Analysis

131.3.1 PD Model Design and Experiment System

According to the PD characteristics of XLPE cable joint, four typical discharge models are designed: needle plate discharge, internal discharge, suspended discharge and surface discharge. The structures are shown in Fig. 131.5 and the experiment system is shown in Fig. 131.6.

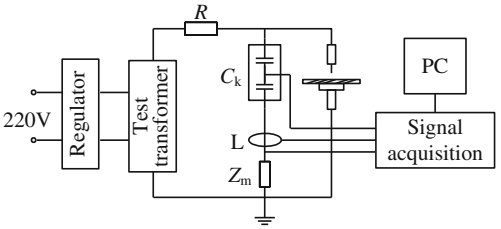


Fig. 131.6 PD experiment system

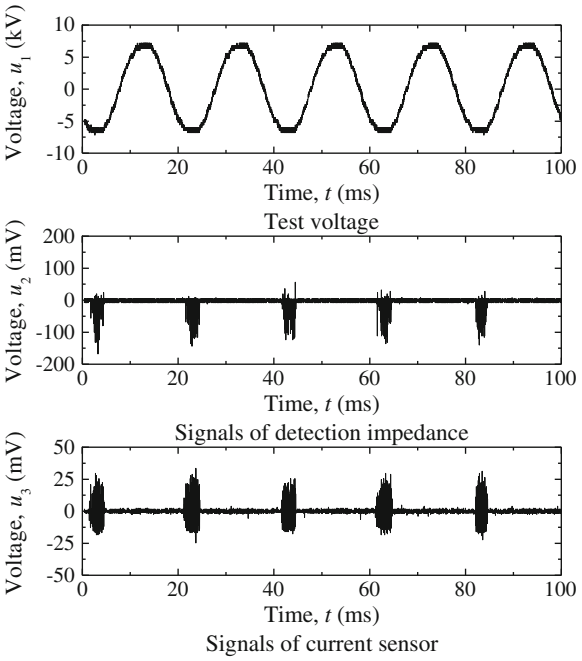


Fig. 131.7 Results of needle plate discharge

In Fig. 131.6, R is current-limiting resistance, C_k is coupling capacitor, L is the current sensor and Z_m is the detection impedance of LDS-6 PD detector.

131.3.2 Experiment Results and Analysis

Needle plate discharge. As shown in Fig. 131.7, applied voltage (peak) is about 7.2 kV and discharge magnitudes are about 25 pC. Discharges gather at the negative half-cycle of the peak, while little discharge at the positive half-cycle.

Internal discharge. As shown in Fig. 131.8, applied voltage (peak) is about 11.6 kV and discharge magnitudes are about 100 pC. Discharges which are acicular appear at both positive and negative half-cycle and gather at the rising edge.

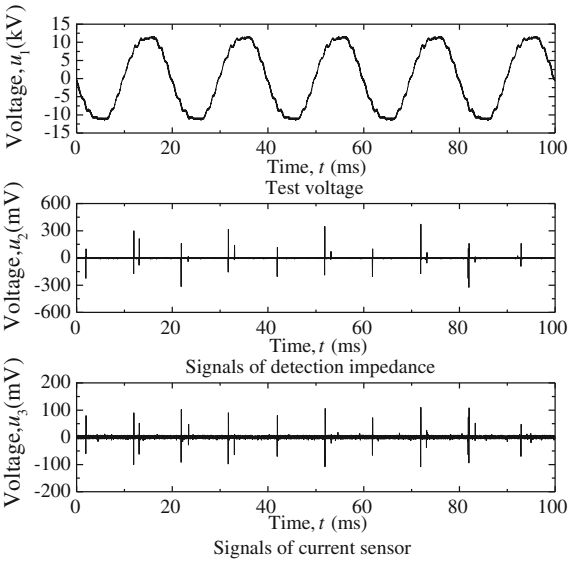


Fig. 131.8 Results of internal discharge

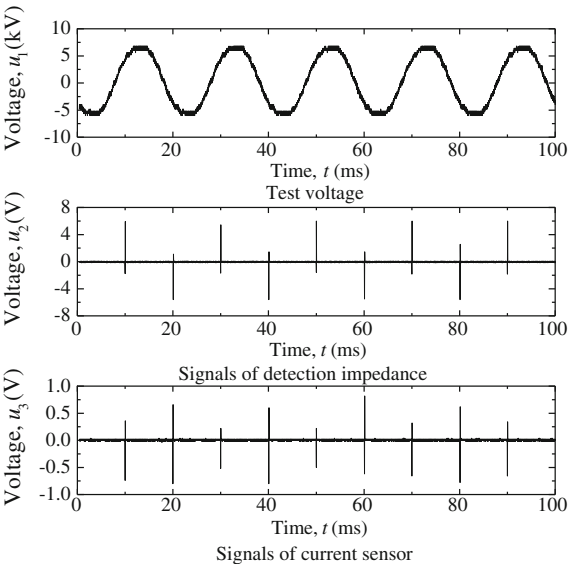


Fig. 131.9 Results of suspended discharge

Suspended discharge. Figure 131.9 shows the results of suspended discharge, of which discharge magnitudes are much higher than others, about 800 pC and obvious sizzle sound can be heard while the applied voltage is about 7.0 kV. Meanwhile, discharges appear at both positive and negative half-cycle.

Surface discharge. As shown in Fig. 131.10, the applied voltage (peak) is about 8.2 kV and discharges at the rising edge of 150 pC appear at both positive and negative half-cycle. Unlike the internal discharge, discharges of surface discharge are tufted at both positive and negative cycle.

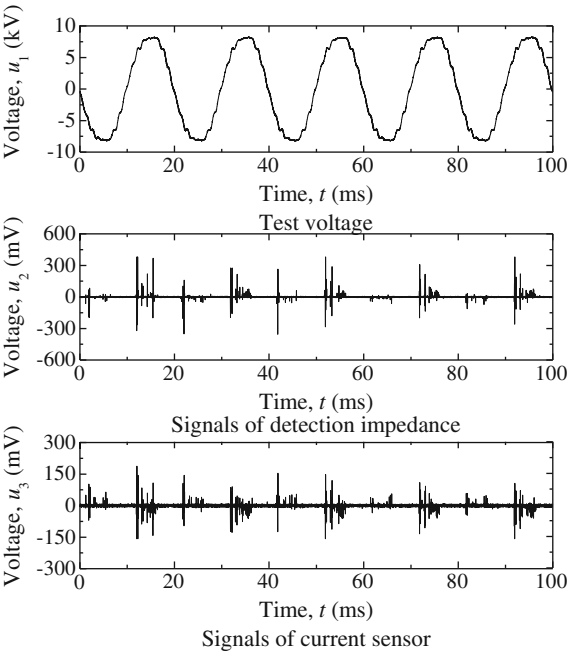


Fig. 131.10 Results of surface discharge

131.4 Conclusion

A set of PD detection system which bases on wideband current sensor is developed. The wideband current sensor designed has many advantages such as wideband of the measurement, high sensitivity, strong anti-interference ability and high operational security. Based on the current sensor, fast and accurate detection of PD in cables can be realized.

The results of four typical discharge models show that we can extract the PD signals effectively and observe the discharge information through the wideband current sensor and then realize online monitoring and pattern identification to XLPE cable joint.

References

1. Luo J, Qiu Y, Yang L (2003) Operation fault analysis of XLPE power cable above 10 kV. High Volt Eng 29(6):14–16
2. Mashikian MS (2002) Preventive maintenance testing of shielded power cable systems. IEEE Trans Ind Appl 38(3):736–743
3. Li Y, Chen Z, Lv F et al (2003) Pattern recognition of transformer partial discharge based on acoustic method. Proc CSEE 23(2):108–111
4. Krivda A (1995) Automated recognition of partial discharge. IEEE Trans Dielectr Electr Insul 2(5):796–821

5. Xiao Y, Yu W (2005) Present status and prospect of research of on-line partial discharge monitoring system in GIS. High Vol Eng 31(1):47–49
6. Duan N (2003) Study of the wideband electromagnetic coupling method used to detect partial discharge in power cables. Xi'an Jiaotong University, Xi'an
7. Zhu H, Feng J, Luo J (2004) Study on PD HF detection technology of XLPE power cable. High Volt Eng 30(136):75–76
8. Boczar T (2001) Identification of a specific type of PD from acoustic emission frequency spectra. IEEE Trans Dielectr Electr Insul 8(4):598–606
9. Lemke E, Strehl T, Russwurm D (1999) New developments in the field of PD detection and location in power cable under onsite condition. The 11th international symposium on high voltage engineering, London, 22–27 Aug 1999, vol 5, pp 106–111
10. Ahmed NH, Srinivas NN (1998) On-line partial discharge detection in cables. IEEE Trans Dielectr Electr Insul 5(2):181–188
11. Luo J, Feng J, Yuan J et al (2001) Study on detection of partial discharge in XLPE cable at higher frequency. Power Syst Technol 25(12):42–45
12. Cheng Y, Li W, Xie H et al (1998) Study on the ultra-wideband sensor for partial discharge. High Volt Eng 24(1):9–11

Chapter 132
Fast Unrestrained Differential Protection Based on Phaselet Algorithm

Chuan-qi Wang, Xiang-li Deng, Zhe Zhang and Xiang-gen Yin

Abstract The misoperation of differential protection caused by saturation of CT would take place during external fault of transformer. In order to solve this problem, this paper describes a fast unrestrained differential current protection algorithm based on phaselet algorithm, which relies only on sampling data of one-fourth circle. Based on the relationship between phaselet algorithm and full circle Fourier algorithm, this paper induces the linear relationship between the two algorithms. The magnitude-frequency characteristics of phaselet are analyzed and the fast protection criteria is proposed. The feasibility of the proposed algorithm is validated by the static testing and dynamic testing. The relay based on the proposed algorithm has been applied in field. The algorithm proposed has fast operation time and can be influenced by CT saturation.

Keywords Unrestrained differential current protection · Phaselet algorithm · CT saturation · Transformer protection

C. Wang · X. Deng (✉) · Z. Zhang · X. Yin
School of Electrical and Electronics Engineering, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China
e-mail: dengxl989@hotmail.com; dengxiangli@dongfang-china.com

C. Wang
e-mail: chuanqiw@sina.com

Z. Zhang
e-mail: zz_mail2002@163.com

X. Yin
e-mail: xgyin@mail.hust.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_132,

132.1 Introduction

Unrestrained Differential Relay (UDR) is a kind of overcurrent relay, which reflects differential currents flowing through transformer. The prompt tripping action takes place when there is inner fault in transformer without any delay caused by CT saturation detecting logic.

The tripping action would be delayed and heavy harm to the body of transformer would be caused during severe inner fault in transformer if extra CT saturation detecting logic is added in UDR. Thus, prompt tripping is the basic demand for UDR.

However, in 500 kV substation, the high voltage busbar is connected as the type of one-and-half breaker. When severe faults take place in high voltage busbar, busbar differential relay can issue tripping command; however, the misoperation would be caused by UDR due to transformer terminal CT saturation and severe blackout would take place. The misoperation cases of UDR have ever been reported in substations of HV (35–110 kV) caused by transformer terminal CT saturation. Thus, the influence of CT saturation over UDR should be considered in the design of protective element of UDR. A new fast algorithm is needed by UDR, which is affected neither by CT saturation nor by delay.

In order to promote the safety and reliability of power transformer, many algorithms have be proposed in Ref. [1, 2]. Phaselet algorithm has such advantages as short data-window, fast speed to complete calculation of fault quantities and is focused on widely by many researchers (the reference of [3–6]). The characteristic and accuracy of phaselet algorithm are analyzed in the reference [3] and is applied in transmission line differential current relay. However, there is no application information of phaselet in some solid software and hardware platform in it. The reference [4] focuses on the application in line protective relay and the reference [5] focuses on the application of phaselet algorithm in ratio differential protection of transformer relay. However, this paper focuses on the application of phaselet algorithm in UDR for transformer.

This paper analyzes linear relationship between phaselet algorithm and full-wave Fourier algorithm (FWFA). Based on the linear relationship, calculation method of linear relationship array is induced and the amplitude-frequency characteristic of phaselet algorithm is analyzed. The fast unrestrained differential protective criteria is proposed based on the result of phaselet algorithm and the feasibility of application is validated by static and dynamic testing.

132.2 Unrestrained Differential Protective Criteria Based on Phaselet Algorithm

132.2.1 Phaselet Algorithm and Full-Wave Fourier Algorithm (FWFA)

The amplitudes of current and voltage are calculated by FWFA in digital protective relay. They are used for completion of protective criteria. Full circle data

window is needed by FWFA in order to complete protective criteria, and the algorithm is described in Eq. (132.1):

$$\begin{cases} A = \frac{2}{N}\sum_{k=1}^{N} i_k \times \cos\left(\frac{2\pi}{N}k\right) \\ B = -\frac{2}{N}\sum_{k=1}^{N} i_k \times \sin\left(\frac{2\pi}{N}k\right) \end{cases} \tag{132.1}$$

in which N is the sampling dot number of one circle and i_k is the input signal. The phaselet is defined as the summation of the sampled input signal multiplied by sinusoidal quantities or cosine quantities. Phaselet algorithm is the general circumstance of FWFA. Half- or full-circle data window is adopted by traditional FWFA while part of sampling data window is adopted by phaselet algorithm. So, it is uncompleted FWFA and Half- or full-circle data window is not needed. However, there is a linear relationship between phaselet algorithm and FWFA.

Assuming that No. q phaselet is given by the equation (132.2):

$$\begin{cases} a = \sum_{k=pq-q+1}^{pq} i_k \cos\left(\frac{2\pi}{N}k\right) \\ b = \sum_{k=pq-q+1}^{pq} i_k \sin\left(\frac{2\pi}{N}k\right) \end{cases} \tag{132.2}$$

in which p is the number of sampled values. There are N/p phaselets in one circle.

Assuming the input signal is sinusoidal, the phaselet algorithm and FWFA are calculated with the initial angle of 0 and 90 degree, respectively and the two constant array can be obtained below.

$$a_{\text{Arry}}\begin{bmatrix} a_0 & a_{90} \\ b_0 & b_{90} \end{bmatrix},\ A_{\text{Arry}} = \begin{bmatrix} A_0 & A_{90} \\ B_0 & B_{90} \end{bmatrix}$$

in which a_{Arry} is the constant array of phaselet algorithm with initial calculation angle of 0 and 90 degree and A_{Arry} is the constant array of FWFA.

Thus, calculation relation array between phaselet and FWFA is given by:

$$K_{\text{Arry}} = A_{\text{Arry}} \times a_{\text{Arry}}^{-1}$$

During the course of data processing in digital relay, the real and image part can be calculated by phaselet with Eq. (132.2) using sampling data. The corresponding real and image part calculated by FWFA are given by Eq. (132.4):

$$\begin{bmatrix} Y_{\text{Rcir}} \\ Y_{\text{Icir}} \end{bmatrix} = K_{\text{Arry}} \times \begin{bmatrix} a \\ b \end{bmatrix}$$
$$\text{Mag} = \sqrt{Y_{\text{Rcir}}^2 + Y_{\text{Icir}}^2} \tag{132.4}$$

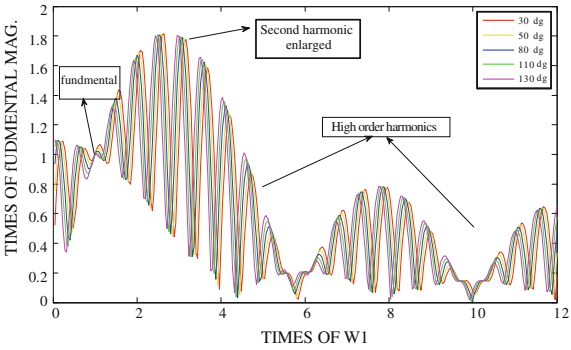


Fig. 132.1 Magnitude-frequency characteristics of phaselet

in which Mag is the magnitude of FWFA. Thus, the linear relationship between one-fourth circle phaselet algorithm and FWFA can be deduced by Eq. (132.4).

When the CT saturation is caused by the out-zone fault of transformer, there exists linear CT transferring time of about 5–6 ms from initiation of fault through CT saturation. After 6 ms, the output current waveform of CT would have been distortional and the calculation error would be caused if the FWFA were used directly. In order to void calculation error, the phaselet algorithm with 5 ms data window is adopted to calculate the real and image part of differential current, the magnitude of differential current calculated by FWFA is obtained using the linear relationship array between phaselet algorithm and FWFA, and the magnitude can be used by UDR and no misoperation occurs. Because only the prosaturation data is adopted to calculate the magnitude of differential current and the magnitude can reflect the real differential current in transformer, no calculation error can be caused by FWFA and no misoperation is caused by UDR.

132.2.2 Magnitude-Frequency Characteristic of Phaselet Algorithm

It is important for magnitude-frequency characteristic of filter algorithm to be used in protective algorithm. The capacity, with which the valuable frequency part is extracted and invaluable frequency part is suppressed, can be reflected in the magnitude-frequency characteristic of phaselet algorithm. Thus, characteristic can be used as important criteria to judge performance of filter algorithm. The magnitude-frequency characteristic of phaselet algorithm is drawn in Fig. 132.1.

The analysis on the magnitude-frequency characteristic focuses on two aspects: the calculation accuracy influenced by different initial angles in input signal and the suppressing degree to high frequency components.

It is shown in Fig. 132.1 that the fundamental components can be filtered out by phaselet algorithm without being influenced by different initial angles of input signal and the enlargement factor is 1 because the different magnitude-frequency curves of different initial angles insect at one dot.

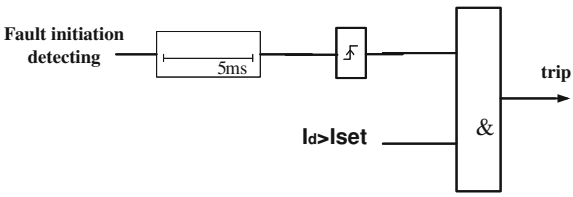


Fig. 132.2 Fast unrestrained differential protection criteria based on phaselet technique

There is no calculation error for phaselet algorithm theoretically if only fundamental frequency signal is injected. The harmonic components more than five order are suppressed better and suppressed below 0.7 time of fundamental frequency components. Sixth and tenth harmonic components are almost removed; however, the second and third harmonic components are enlarged to 1.8 times of fundamental.

132.2.3 Fast Unrestrained Differential Protective Criteria Based on Phaselet Algorithm

The CT saturation can occur at time of 5–150 ms after external fault initiation, the CT saturation cannot be caused by the transient fault current at once, and there is linear transferring time about 5 ms. If the unrestrained differential protective criteria could be completed in 5 ms, misoperation caused by CT saturation would be avoided. The logic scheme is shown in Fig. 132.2.

It is shown in Fig. 132.2 that there are two parts in the protective logic: half-circle Fourier element and quarter-circle phaselet algorithm element. In order to avoid misoperation of UDR cursed by stable CT saturation, the half-circle Fourier element is needed. Its setting is higher than that of phaselet algorithm element and is bigger than the most large external fault current in power system. At the time of 5 ms after fault initiation, the quarter-circle phaselet algorithm element is triggered once, then, after that it is blocked. Thus, it is not influenced by CT saturation and the misoperation caused by severe CT saturation during external faults can be avoided.

It is shown in Fig. 132.2 that there are two parts in quarter-circle phaselet algorithm element: fault initiation detecting element and phaselet algorithm element, which can be calculated only once in 5 ms.

When fault initiation detecting element operates, it shows that one new fault is initiated, the time of 5 ms is delayed and the output of rising-edge block is effective. At this time, if the magnitude calculated by phaselet algorithm is bigger than the UDR setting, tripping command is triggered and if it is lower, the whole logic is reset and waits for the coming of the next fault.

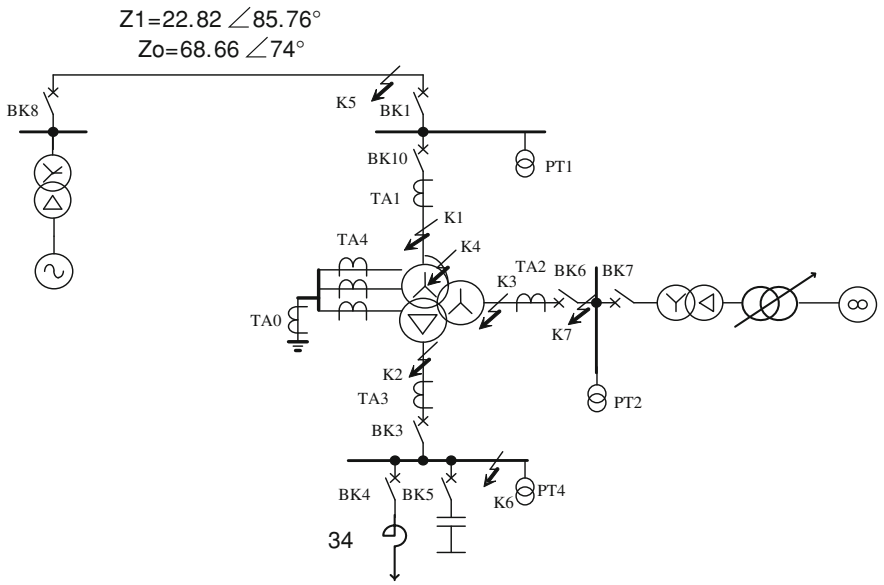


Fig. 132.3 The connection diagram of dynamic testing for transformer

132.2.4 Hardware Platform Development for Digital Relay

The hardware platform of DF3630 includes protective DSP board, HMI board, auxiliary CT board, auxiliary relay board and signal relay board. The protective DSP board is divided into two modules: core DSP module and peripheral module. Because of standard plug interface, the core DSP board can be plugged into different peripheral modules, thus, the number of different category hardware platform is reduced (many products share the same core DSP board) and the reliability is increased. In the core DSP board, TMS320C6713 DSP chip of TI is adopted and the 8 MB of SDRAM is adopted to store executive code of protective relay and dynamic local variables. During reset process of relay, the executive code of relay is read from 2 MB FLASH to SDRAM and is executed. In the case of blackout in substation, the fault record data stored in 2 MB NVRAM can be read to analyze the accident in power system.

132.3 Dynamic Testing

The DF3630 transformer protective relay based on fast unrestrained differential criteria have past dynamic testing in DFE company testing centre, HUST dynamic LAB and KETOP LAB, respectively. The system connection diagram in HUST dynamic LAB is shown in Fig. 132.3.

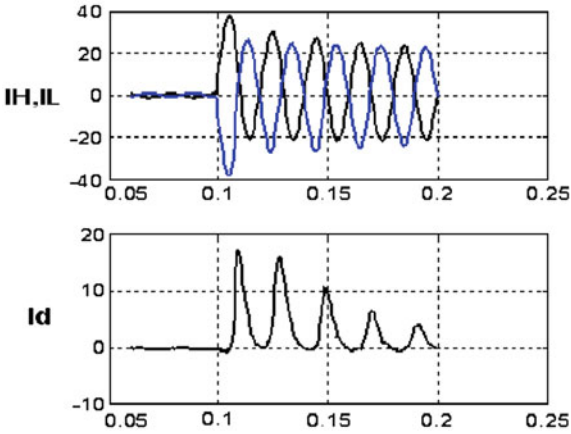


Fig. 132.4 Current waveforms when CT is saturated during out zone fault in middle voltage side

Table 132.1 The tripping time of fast unstrained differential protection during in zone fault in high voltage side

No.	Phase	Operation time
1	AN	16
2	BN	16
3	CN	17
4	ABN	17
5	BCN	16.5
6	CAN	16.5
7	AB	16
8	BC	16
9	CA	14.5
10	ABCN	16

The transformer simulated in the system is a three-single-autotransformer group illustrated in Fig. 132.3 with the winding connection of star-grouded/star-grouded/delta. The parameters of the single transformer model are described as below:

rated capacity: 1,000 MVA
rated voltage of high winding: 1,050 kV
rated voltage of low winding: 525 kV
rated voltage of tertiary winding: 110 kV
short-circuit impedance of high winding: 18%
low winding: 62%
tertiary winding: 40%.

The dynamic testing waveform is shown in Fig. 132.4. Testing data are shown in Table 132.1.

It is shown in Fig. 132.4 that differential current and fault current would not occur at the same time during the CT saturation caused by severe external fault in middle voltage side of transformer (K7). There is a linear transferring time after

the fault initiation and after that, differential current is increased at once and then decreased. The linear transferring time of CT can be used by phaselet algorithm to finish fault detecting. It is shown in Fig. 132.4 that the fast unrestrained differential protection is not affected by the CT saturation caused by external fault and no trip command is triggered

The fast UDR operation time is shown in Table 132.1 during various faults taken place in high voltage side of transformer of K4. The operation time in Table 132.1 includes algorithm detecting time and the auxiliary relay operation time. Because slow auxiliary relay is included in DF3630 transformer relay and the time of relay contact closing is about 10 ms, it is known from Table 132.1 that the fault detecting time of fast UDR is about 4.5–7 ms and it can complete fault detecting before CT saturation.

132.4 Conclusion

This paper analyzes the linear relationship between phaselet algorithm and full-wave Fourier algorithm. It proposes the method of linear relationship array calculation. A fast unrestrained differential protective criteria is proposed as well, which is not influenced by CT saturation caused by external fault. The fault detecting is completed in linear transferring time of 5 ms before CT saturation. If there is no fault in transformer, the setting of unrestrained differential element is increased and half-wave Fourier algorithm is used to detect fault. All protective criteria are developed in DF3630 transformer relay platform and have past static and dynamic testing. The transformer relay has been applied in field. All the analysis and testing results have been validated in the application field.

References

1. He B, Xu X (1998) Protection based on wave comparison. Proc CSEE 18/6:395–398
2. Zhang X, He B (2005) A new method to identify inrush current by phasor comparison. Proc CSEE 25/19:43–47
3. Wen M, Chen D (2003) Simple analysis of phaselet algorithm. Autom Electr Power Syst 27(3):42–44
4. Ha H, Zhang B (2005) New phaselet-based algorithm of fast distance relay. Autom Electr Power Syst 29(18):39–44
5. Zhao Y, Lu Y (2008) Transformer adaptive elliptical restraint differential protection based on phaselet algorithm. Proc CSEE 28(7):84–90
6. Cheng D, Zhang Z, Yin X (2000) Digital protective relay (1st edn.). CEPP 7:59–62

Chapter 133
Analysis of Electric Wire Fracture Based on Infrared Image Diagnosis

Yanzhou Sun, Xue Zhou and Xinwei Niu

Abstract First the infrared images which are collected in Luoyang substation are analyzed. The types and eliminating methods of the defects are determined. Second, Visual Basic 6.0 combining with MatrixVB is used to design an infrared image processing system software. The system realizes the filter for infrared image processing, grayscale processing and edge detection conveniently in actual application. Using this system, the substation's infrared fault diagnosis are convenient and can be used for qualitative analysis.

Keywords Infrared image processing · Wire fracture · Visual Basic

133.1 Introduction

Substation is an important part of power system. If the malfunction appeared, the normal operation of substation would be affected. More seriously, a great loss and accidents would be brought. So the fault diagnosis of the substation is particularly important. At present, the infrared diagnostic technique is one of the important means of the fault diagnosis technique. Because in actual condition the infrared fault diagnosis will be affected by many factors or only certain areas in the images

Y. Sun (✉) · X. Zhou
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, Henan, China
e-mail: sunyz@hpu.edu.cn

X. Niu
Department of Substation Operation,
Henan Luoyang Power Company,
Luoyang, Henan, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_133,

need to be emphasized, so the infrared image which is collected should be processed to make the infrared image more in line with the requirement of the people [1, 2].

The paper analyzes the infrared images which are collected in Luoyang substation of Henan province. The types and eliminating methods of the defects are determined. Second, Visual Basic 6.0 combining with MatrixVB is used to design an infrared image processing system software. The system realizes the filter for the infrared image processing, grayscale processing and edge detection conveniently in actual application. For the software system and the image acquisition system are independent, it could realize the secondary development of the image. Thus the image processing could be realized more convenient.

133.2 Infrared Diagnostic Analysis of Phase Line Side in Substation

On July 31st, 2010, the temperature of the main transformer of Luoyang Liming substation in Henan province showed anomaly. After the effective measures to control the temperature of the main transformer were taken, the substation equipment of liming substation were detected by infrared detection. Finally we found the side of Limu 1st line's A phase fever, as shown in Fig. 133.1.

Using HX8600 infrared thermography device, the infrared image of the side of Limu 1st line's A phase is shown in Fig. 133.2.

According to the Fig. 133.2, diagnostic analysis could be done: the fever part was the side of Limu 1st line and the temperature was too high. According to the method of the defects types' determination, for 173.4°C is higher than 130°C, so the emergency measures should be taken and cut power to deal with.

We can see the disconnection defects of the line could be discovery and eliminate promptly by electrician to ensure the normal operation of the power system stability.

133.2.1 Filtering Processing of Substation's Infrared Image

Due to the external factors (such as environmental temperature, air quality, the distance between the cameras and targets) interference, the collected infrared image would be affected. The affect of external factors most cases with noise exists in the form of the infrared image. Normally, the noise can be divided into random noise interference or stable and sustainable noise. The size and distribution of the random noise's interference is random. The condition of image interference shows scattered and granular form. The distribution of sustainable noise is uniform and continues for a long time. So the sustainable noise causes long influence to the image [3, 4].

Fig. 133.1 Wire fracture of transmission line

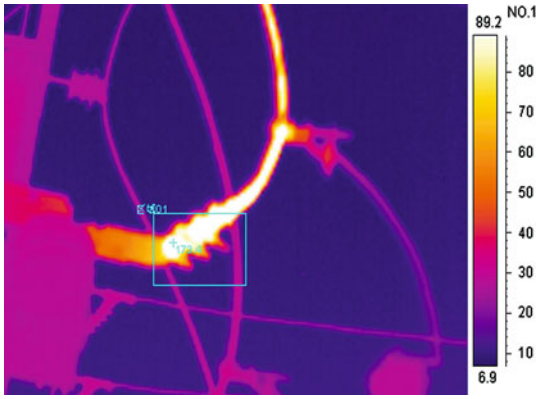


Fig.133.2 The infrared image of the side of the line

In order to obtain clean and real images and to reduce the noise negative effects, we use the noise filter to process the collected image.

Using the infrared image processing's software system of the substation to do filtering processing. As shown in the Fig. 133.3a, the image contains disturbed noises. After the methods of low pass filtering processing, the image is Fig. 133.3b. Compare with Fig. 133.3a and b, the infrared image after processing is close to the actual image basically. The image could be more clearly identified at fault point and the filtering effect is obvious.

133.2.2 Gray-Scale Processing of Substation's Infrared Image

In the process of the scanning, for the scanning system or the photoelectric conversion system and so on, in the image there often appears some drawbacks such as uneven, not enough contrast, lack of administrative and blur detail and so on. Using grayscale transform can modify the image pixels in the collection system to make the whole image much more uniform [5–7].

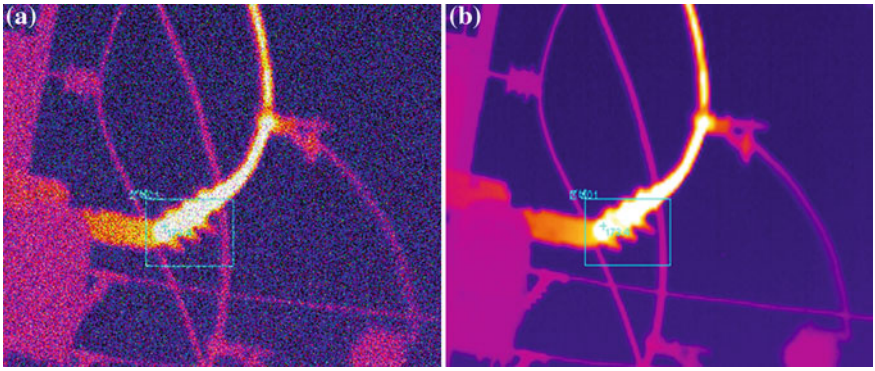
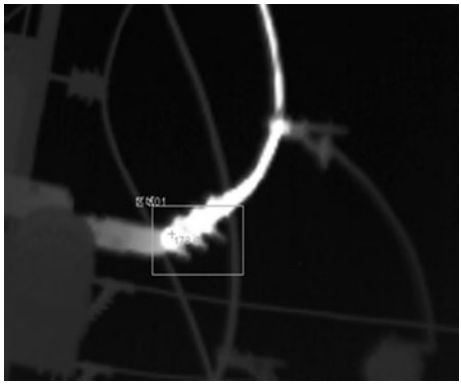


Fig. 133.3 The figure of filtering processing **a** including noise image **b** the image after filtering processing

Fig. 133.4 Gray image



The most functions in the MATALAB are processing on the gray image. Combined with the MATALAB could realize a lot of image processing function. So gray-scale processing for the collected image is very necessary. Used the function rgb2gray (RGB), after the gray-scale processing of Fig. 133.3b, the image is shown in Fig. 133.4.

According to the Fig. 133.4, we can see that the gray image is clear and we can get the gray image conveniently by the system.

133.3 Edge Detection of Substation's Infrared Image

In the time of the image's analysis and understanding, edge detection is first thing. At present edge detection is one of the most active subjects in machine vision research field. Edge detection plays a very important role in engineering application [8].

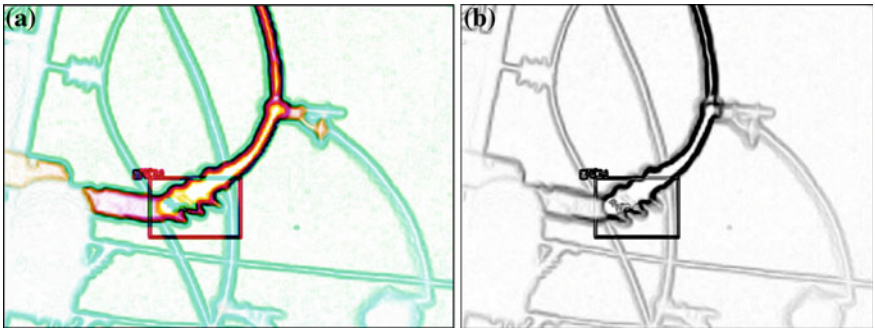


Fig. 133.5 Edge detection figure **a** edge information map of RGB image **b** edge information map of the gray image

Edge detection of Infrared image is a kind of infrared image enhancement. The processing of edge detection is emphasized the adjacent edge which the value of tonal pixel are larger. Image edge usually refers to the place which the image tonal mutation or the boundary of features. After edge detection, not only the image can be more clearly and show different phenomenon's border and features boundary, but also the linear images should appear which is better to identify the terrain and delineate the distribution [9].

The software system of this paper can realize edge detection of the infrared image conveniently. The results of edge detection are shown in Fig. 133.5. The Fig. 133.5a is the edge detection of Fig. 133.3b. The Fig. 133.5b is the edge detection of Fig. 133.4.

According to Fig. 133.5, the edge detection figure can be acquired quickly by using the software system. The edge information of original image shows well. The image has intense administrative levels and visual effect, which provide important information for image recognition.

133.4 Summary

The code of infrared image formats is very special, bad commonality and more difficult for second development. These are the problems for infrared image processing. The infrared image processing system which the paper designed has the following features:

(1) With quoting MatrixVB, the software system realizes the infrared image processing well and use simply, conveniently and quickly.
(2) For the infrared thermography is very expensive in the market, the matched infrared image processing software could cause consumers extra burden. So in order to get rid of the dependency of the matched software, the paper designs a software system just to meet this requirement. And the system can also make secondary development for infrared image without others' distractions.

(3) By using the software system, the original image can be better reproduced and the information blocked off from the original image can also be displayed to make the information of the infrared image more abundant.

According to the above, the paper analyzes the infrared images which are collected in Luoyang substation. The types and eliminating methods of the defects are determined.

An infrared image processing system software designed and the results of image processing acquired conveniently. The system makes the substation's infrared fault diagnosis more accurate and convenient and be helpful for qualitative analysis of the substation's fault diagnosis.

Acknowledgments It is a project supported by the Henan Higher Education Excellent Teacher Program and China Natural Science Foundation (50974053).

References

1. Xipu Z (1992) The development of power system's infrared diagnostic technique. J North China Power Technol 1:19–23
2. Jinyu X, Liu N, Hao J et al (2005) Electric power equipment condition monitoring technology research situation and development. Relay 33:80–84
3. Li H, Li J, Li P (2008) The application of infrared thermal imaging technology in the power plant. J Mod Electr Technol 31:184–188
4. Ruijun R (2010) The infrared diagnostic of 6 kV distribution line defect. Oil Gas Field Surf Eng 4:49–53
5. Wang B, Chen Y, Yin Q (2004) The application analysis of high-pressure electric equipment infrared diagnostic. J Yunnan Power Technol 1:24–51
6. Zhang D, Wangxia, Liqian A (2003) A kind of infrared thermal imaging files and its image processing method. Opt Technol 3:334–341
7. Kailun J, Hu X, Zhao Y (2009) The discuss of improve electrical equipment infrared diagnosing technical performance. J Guizhou Power Technol 12:51–53
8. Xu B, Zhao J (2008) The application of several edge detection algorithm in transformer faults diagnosis of infrared image processing. Mech Electr Eng Technol 37:109–111
9. Tao T (2004) The method of comprehensive edge detection and regional growth of infrared image segmentation. Photoelectr Proj 10:49–53

Chapter 134
Sentiment Analysis of Text Using SVM

Yong Yang, Chun Xu and Ge Ren

Abstract In recent years, development with the internet, information processing turns more and more important for us to get useful information. Text Categorization, the automated assigning of natural language texts to predefined categories based on their contents, is a task of increasing importance. Therefore, an approach of text emotional classification based on support vector machine was proposed. The system includes four parts: word segmentation, establishment emotional word database, training the model of classification, and testing. The system is actually used for classification of the reader's emotions that classify the text to achieve the purpose of forecast readers' emotion.

Keywords Text · Sentiment analysis · Classification · Support vector machine · Word segmentation

Y. Yang (✉)
Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Science, Urumqi, China
e-mail: Yangyong1900@163.com

Y. Yang · G. Ren
College of Computer Science, XinJiang Normal University, Urumqi, China
e-mail: RenGe@163.com

C. Xu
College of Computer Science and Technology, Xinjiang University of Finance and Economics, Urumqi, China
e-mail: XuChun@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_134,

134.1 Introduction

The booming Internet has spread its influence on people throughout all aspects of social life. In recent years, online communities, blogs, and multiple kinds of forums has offered people a broader platform for information exchange with the conversion of most Internet users from passively receiving information to initiatively creating information. The large number of Internet information has expressed people's emotional coloring as well as their tendencies, such as "Happy, Angry, Sorrow, and Joy". In-depth popular opinions on certain incident or product can be obtained through emotional analysis of text, whereas the rapid expansion of information amount has made it impossible to accomplish text emotional analysis merely by manual methods, for which sake text emotional analysis in terms of computer technology has been introduced [1].

The second part of this article elaborates on the methods adopted in this article during emotional information extraction phase, which include measurement algorithm (TF-IDF) and classification algorithm Support Vector Machine (SVM); the third part details the methods specific to the accomplishment of text emotion analysis in this article; the fourth part lists the specific experiment as well as its results; the fifth part comes the conclusion of this article.

134.2 Algorithm Description

134.2.1 SVM Classification Algorithm

Training methods and classification algorithms are core section in classification system. Currently there are a variety of training algorithms and classification algorithms based on vector space model, such as SVM, neural networks, the average maximum entropy method, close to recent methods, and Bayesian K, etc. [2]. This article takes SVM algorithm during the specific system implementation.

Establish the sample set (x_i, y_i), $i = 1, \ldots, n\, y \in \{-1, +1\}$ is the category label. The general form of linear discriminant function as well as separating hyperplane equation in d-dimensional space is expressed as $g(x) = w \cdot x + b$ and $w \cdot x + b = 0$, normalize the discriminant function to ensure all samples meet $|g(x)| \geq 1$ and the nearest sample to separate hyperplane meeting $|g(x)| = 1$ with the classification interval equal to $2/||w||$ and make sure that the separating intervals are maximum equivalent to minimizing $\|w\|$ $\left(\text{or } \|w\|^2\right)$ Requiring the classification hyper-plane to correctly classify all samples means that the classification hyper-plane should meet:

$$y_i[(w * x_i) + b] - 1 \geq 0, i = 1, \ldots n \tag{134.1}$$

Therefore, the hyperplane which meets formula (134.1) while minimizing $\|w\|^2$ value is called optimal separating hyperplane, and the training samples on H1, H2

corresponding to the optimal separating hyperplane is defined as support vectors. w^*, the solution to w, can be gained from optimization theory, likewise, b^*, the solution to will be obtained by using support vector machines, after which optimal classification function is gained

$$f(x) = \text{sgn}\{(w^* \cdot x) + b^*\} \tag{134.2}$$

Substitute the optimal classification function to the input sample type.

134.2.2 Text Description

As for text expression, this article uses vector space model (VSM), the basic idea which is :Each text has some individual properties which are represented by a number of concept words for content expression, whereas each property can be seen as a dimension in concept space, and these individual properties are regarded as text feature item, therefore the text is seen as a collection of such items, which releases the necessity to consider the complex relationships among paragraphs, sentences, and words in text structure Hereby the text can be represented as a form like d = (t1, *W*1;t2, *W*2;……;tn, *W*n), among which ti and *W*i refer to the feature item and correspondent weight [3–5]. The feature item weight *W*i is commonly measured by TF-IDF formula which is defined as follows:

$$W_i = \frac{tf_{ij} \times \log(N/n_i + 0.01)}{\sqrt{\sum_{t \in \check{d}} \left[tf_{ij} \times \log(N/n_i + 0.01)\right]^2}} \tag{134.3}$$

Here tf_{ij} refers to the frequency of feature item *ti* in text *d* with *N* being the total text number, n_i being the number of text in which *ti* has appeared in text set, and the denominator refers to normalizing factor [6].

Obviously, feature items with large weight refer to those words found sufficient frequent in one text while rarely seen in other texts in the whole text set, which will play an important role in text classification.

134.2.3 Feature Extraction

A large number of words are necessary for one text, in response; the dimension of vector space used for text expression is also quite large (may be up to 10,000 above). Therefore, it is mandatory to implement dimension reduction for the following two reasons [7]: 1. Enhancement in program efficiency and operation speed. 2. Each of all the words (may be up to ten thousands above) contributes differently to text classification in that those common words which are available in each category contribute less whereas those found frequent in certain specific

category while rarely seen in other categories are able to make substantial contribution, for which sake in each category it is suggested that those less expressive words be removed so as to screen out the feature item set specific to that category for classification accuracy, and feature scoring function constructed based on mutual information, information gain, expected cross entropy, and text evidence is regarded as the common extraction method.

The criterion in which mutual information of words and categories for feature item extraction is considered as one effective method, the basic procedures of which are described as below:

Step 1. The collection of feature items shall contain all words appeared in this category during initial case

Step 2. Calculate the mutual information between words and categories for each word.

$$\log\left(\frac{P\left(W|C_j\right)}{P(W)}\right) \times \exp\left(\frac{N_j}{N}\right) \tag{134.4}$$

In (134.4), $P\left(W|C_j\right)$ refers to the frequency of the feature item W from training corpus in category C_j, $P(W)$ represents the frequency of the feature item in training corpus. N_j is the text number of the feature item W appeared in category C_j, N refers to the text amount for the feature item W appeared in training corpus.

Step 3. Sort all words in this category as per the mutual information amount calculated above.

Step 4. Extract a certain number of words as the feature item.

Finally, 1,000 terms are selected as the feature item in each category.

134.3 Functional Accomplishment

This article has initially accomplished the structure diagram specific to system as per the above mentioned solution to text classification system, which is listed as below Fig. 134.1:

In the pre-processing procedure for training text, text data will undergo words separation, removal of stop words and parameter statistics (including words frequency, relevant text number and weight) in sequence so as to get data represented by text vector model whereas feature item extraction will be realized by using above mentioned mutual information method for words. The training data generation means processing each text into text vector data expressed by feature item and making correspondent adjustment as per the SVM interface specification selected. Pre-processing and data generation of testing text refers to the word separation and expression of feature item vector model process for one unknown text, classification number of the unknown text which is in test will be generated by classification and output with associated statistical accuracy available.

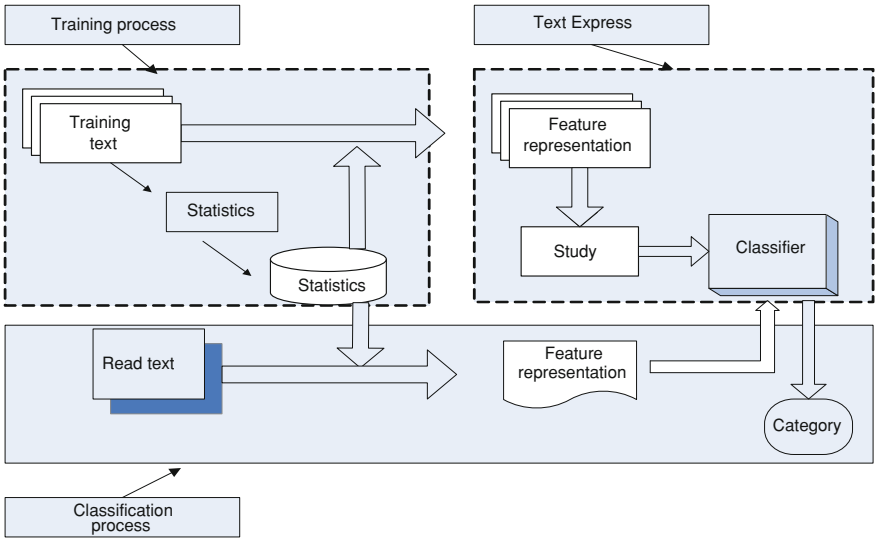


Fig. 134.1 Structure of text classification system

134.4 Analysis and Evaluation on Results

Evaluation index for text theme classification including recall ratio and precision ratio and F value are used in this experiment for emotional classification of text in this article. Precision ratio refers to the percentage of the text that matches the result of manual classification in all tested text; recall ratio is the percentage of the text that matches classification system in all text as per manual classification result. Both precision ratio and recall ration shall be taken into consideration comprehensively since they represent two different aspects in classification quality, therefore one kind of new evaluation index is introduced, which is named as testing value F. Set a1 and a2 as the correct text number in all positive text and negative text determined by classifier respectively; b1 and b2 as the positive text and negative text number determined by classifier respectively; c1 and c2 as positive text and negative text number respectively. It can be obviously seen that $c_1 + c_2 = b_1 + b_2$. The calculation formula for evaluation index is shown in Table 134.1:

Obviously, $F1 = F2 = F$ since $c_1 + c_2 = b_1 + b_2$ All text in corpus are composed of 800 network article and blog posts generally taken from multiple portals and BBS, which are classified into five categories: Funny, Angry, Sorrowful, Novel, and Boring. The method for training set and testing set selection is as below: Divide these classified corpus into five equal sets with one set selected as an open test set and others as training set and closed training set, likewise, each set will be selected as the open training set in turns, run classification algorithm

Table 134.1 Calculation formula for evaluation index

Category measurement	Positive	Negative	Total
Precision ratio	$pp = \frac{a_1}{b_1} \times 100\%$	$pn = \frac{a_2}{b_2} \times 100\%$	$F2 = \frac{a_1 + a_2}{b_1 + b_2} \times 100\%$
Recall ratio	$RP = \frac{a_1}{c_1} \times 100\%$	$RN = \frac{a_2}{c_2} \times 100\%$	$F2 = \frac{a_1 + a_2}{c_1 + c_2} \times 100\%$
F value	$FP = \frac{2 \times RP \times PP}{RP + PP}$	$FN = \frac{2 \times RP \times PP}{RP + PP}$	$F = \frac{2 \times F1 \times F2}{F1 + F2}$

Table 134.2 Experiment results

Category	Recall ratio in closed test (%)	Precision ratio in closed test (%)	F value in closed test (%)	Recall ratio in open test (%)	Precision ratio in open test (%)	F value in open test (%)
Positive	81.08	81.10	81.18	80.33	80.40	80.56
Negative	82.39	81.78	81.08	81.17	81.26	82.11
Total	81.11	81.42	81.25	81.09	81.12	81.20

(total five times classification operations) to get the average value. The experimental results are shown in Table 134.2.

134.5 Conclusion

This article has come up with one solution plan for text emotional system as per SVM classification method, which is based on classification algorithm specific to supervised study model which is different from that to unsupervised study model. This article has introduced the algorithm model for systematic solution and come up with the associated structural plan as well as the implementation method. Meanwhile this article has also described the experiment procedures and the results while conducting further analysis of problems encountered in experiment and proposed recommendations for future work.

Acknowledgment Sponsor acknowledgment: Natural Science Fund of Xinjiang Uygur Autonomous Region. No:2010211A22 Natural Science Fund of China. No: 31040050.

References

1. Yu H, Hatzivassilogou V (2003) Towards answering opinion questions: separating facts from op inions and identifying the polarity of opinion sentences. In: Proceedings of the 2003 conference on empirical methods in natural language processing, Association for Computational Linguistics, Morristown, NJ, pp 129–136
2. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the conference on empirical methods in natural language processing, pp 79–86

3. Dave K, Lawrence S (2003) Mining the peanut gallery: opinion extracts on and semantic classification of product reviews. In: Proceedings of the 12th international conference on knowledge capture. ACM Press, New York, pp 70–77
4. Zhang W, Yoshidab T, Tang X (2010) A comparative study of TF*IDF, LSI and multi-words for text classification [J]. Expert Systems with Applications 38(3):2758–2765
5. Kamps J (2002) Visualizing word net structure. In: Proceeding of the first global word net conference, India, pp 182–186
6. Xu Y, Wang B, Li J (2008) An Extended Document Frequency Metric for Feature Selection in Text Categorization [J]. Lect Notes Compu Sci 4993(20):71–82
7. Pei-yu L, Yu-zhen Y, Jing Z (2011) Text representation combining syntax in vector space model. Adv Inf Sci Serv Sci 3(7):251–259

Part X
Detection Measurement Technologies

Chapter 135
Research of Condition Monitoring and Fault Diagnosis System for Induction Motor

Jiang Yongying

Abstract Continuous on-line condition monitoring provides information on the energized induction motor's response to normal and emergency conditions, which makes it is possible to evaluate current condition of induction motor and detect abnormality. It allows corrective measures to be taken to prevent upcoming failure. Condition monitoring can reduce the downtime of the processes and increase the maximum interval between failures, thus minimizing the number and cost of unscheduled maintenances, which can reduce breakdown maintenance and increase production system availability and effectiveness.

Keywords Condition monitoring · DSP · Induction motor · Fault diagnosis

135.1 Introduction

The increasing complexity of production management system generates a need for monitoring of more and more parameters. Equipment reliability may be significantly improved through the effective prediction of equipment degradation. Condition monitoring technologies allow achieving this goal by minimizing downtime through the integrated planning and scheduling of repairs indicated by condition monitoring techniques [1]. Efficient production management system assumes having reliable information about the managed system. The only way to maintain such information at the management station is a continuous monitoring

J. Yongying (✉)
School of Mechanical Engineering, Wenzhou University,
Wenzhou 325035, Zhejiang, People's Republic of China
e-mail: yyjiang2003@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_135,

of the system parameters which affect management decisions. Today a large number of condition monitoring techniques and methods have been extensively used in industry because of their contributions to effective maintenance.

The ability to continuously on-line monitor the condition of energized equipment enables operation and maintenance personnel with a means to determine the operational status of equipment, to evaluate present condition of equipment, timely detection of abnormal conditions, and initiate actions preventing upcoming possible forced outages. The consequences of such faults are serious enough to justify the efforts to build a monitoring system to protect electric facilities from disaster [2].

Continuous monitoring will identify potential problem areas that can lead to substantial equipment damage. Early alarms are given to notify operation and maintenance personnel of a potential problem. Further deterioration of the condition could lead to a recommendation to take the equipment out of service for repair and maintenance [3].

135.2 Choice of Maintenance Strategies of Production Management System

If the operator or equipment maintenance staff wants to know whether the equipment could function well or meets their requirement, they must gather many equipment data and judge the data according to their experience. The maintenance strategies could be divided into three kind of categories, which lists as below:

135.2.1 Failure Maintenance

In this kind of strategy [4], the equipment is used until it could not function well. In this situation, the maintenance staff must be on job for 24 h a day. If the equipment does not work properly, the maintenance staff must react immediately. Usually the equipment could not be fixed instantly, and even it could not be repaired. So, it needs lots of spare equipment. This strategy will increase production system maintenance cost.

135.2.2 Preventive Maintenance

Traditionally, this kind of strategy [5] is the most common category in factory. The equipments stop at fixed time interval, perform the maintenance work, and change some components. However, the staff might not need to stop the equipment or change component when the maintenance interval arrives. This strategy leads over-maintenance causes unnecessary cost, and could not handle the unexpected failure.

135.2.3 Equipment Condition-Based Predictive Maintenance

The maintenance job performs according to the condition of the equipment. This method provides a mechanism that prevents over-maintenance; detects the occurrence of equipment failure and fixes it instantly; arranges the maintenance time in advance. The data gathered from equipments needs to be interpreted by experienced maintenance staff, and the equipment condition could be determined correctly. The monitoring system speeds up the time to find out the equipment condition, and prevents the incorrectness due to personal carelessness. The condition of equipment could be detected more efficiently and correctly by this system.

135.3 Architecture of Condition Monitoring System

In this system, its communication network of equipment adopts Controller Area Network (CAN) Field Bus, and microprocessor of monitoring module adopts Digital Signal Processor (DSP) [6], and the communication network between the monitoring module and monitoring center of company adopts Local Area Network (LAN) [7]. The monitoring center of company is connected with higher monitoring center of head company by way of Internet.CAN Field Bus entirely realizes digital communication and can supervise 110 sets of equipment in real time [8]. It can collect lots of information relevant to running status of the equipment and transmit them to monitoring center of company, which makes for upper maintenance of the equipment. It can carry out fault diagnosis to equipment and giving a alarm timely. At the same time this system can online monitor equipment run status. In this monitoring system, all monitoring centers are connected with each other closely. So all information of whole system can share together, which makes for management and maintenance of monitoring system. Architecture of condition monitoring system is shown in Fig. 135.1.

135.4 Monitoring of Induction Motor

Induction motors are the “workhorses” of industry and are the most widely used electrical machine [9]. Because of its simple structure and high reliability, induction motor is used for many purposes such as: pumps, blowers, fans, compressors, transportation, etc. Induction motor is the most typical and representative equipment in modern production system. So induction motor is used to illustrate work process of monitoring system based on condition and diagnosis in detail.

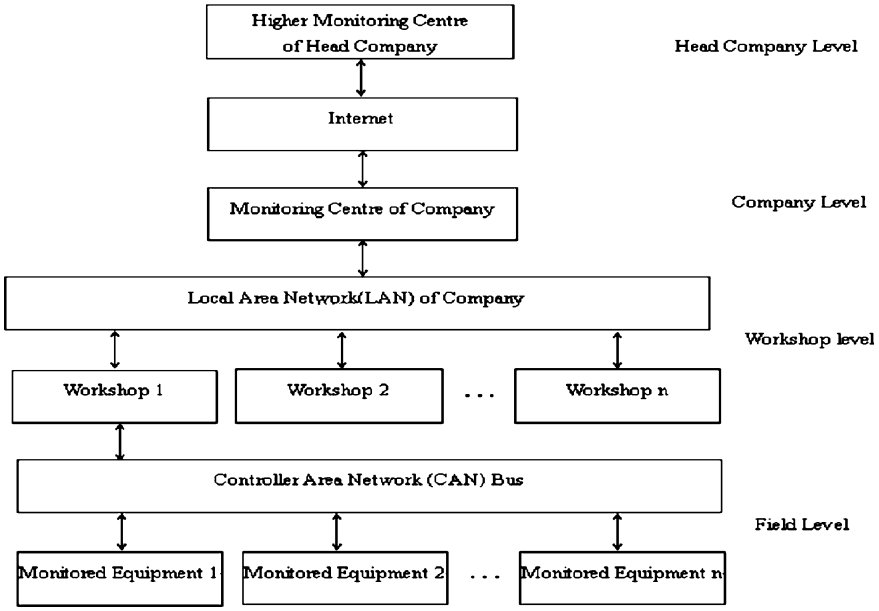


Fig. 135.1 Architecture of condition monitoring system

135.4.1 Importance of Monitoring Induction Motor

Owing to the thermal, electrical and mechanical stresses, mechanical and electrical failures are unavoidable in induction motors. Early detection of abnormalities in the motor will help to avoid expensive failures. Operators of electric drive systems are under continual pressure to reduce maintenance costs and prevent unscheduled downtimes that result in lost production and loss of financial income. The online monitoring of induction motors is becoming increasingly important.

The modern industry has widely used reliability-based and condition-based maintenance strategies [10] to reduce unexpected failures and downtime. These techniques can increase the time between planned shutdowns for standard maintenance and reduce maintenance and operational costs. The operation of the machine in unsafe condition must also be avoided.

135.4.2 Hardware Architecture of Motor-Monitoring Module

Microprocessor of motor-monitoring module adopts DSP. DSP offers marked advantages over conventional devices in terms of speed, portability, versatility, and is insensitive to user error. In addition to these features the DSP-based scheme provides facilities for real time filtering, digital signal processing, and control if

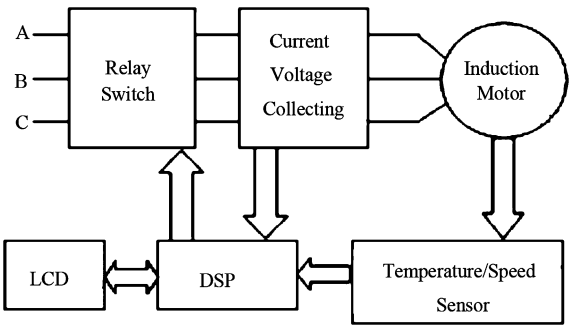


Fig. 135.2 Construction of motor-monitoring module

needed. So it is now possible to rapidly and accurately measure the necessary parameter determination data, calculate the parameters, use them in any control algorithm and/or store the information for any latter analysis.

The construction of motor-monitoring module is shown in Fig. 135.2. It could be divided into four levels, which are listed as below.

135.4.2.1 Data Acquisition Level

There are two ways to gather the motor data. First, motor data (such as temperature, current, voltage, and speed, etc.) gathered from sensor or data acquisition device, and is fed to the monitoring system. Second, the system gets the equipment data from the equipment operating database which stores the online equipment data.

135.4.2.2 Data Processing Level

The monitoring system processes the equipment data for the detection and diagnosis purpose, and derives the feature indexes or characteristic values.

135.4.2.3 Fault Defection and Diagnosis

The monitoring system determines the equipment condition by the feature indexes and characteristic values derived from the former level. The inference system in this level can determine the equipment operating status. If the equipment in bad statns the ECMS could locate the faulty component and sends the messages to the equipment maintenance staff and equipment manager to perform associated jobs.

135.4.2.4 Maintenance Strategy Level

If the failure of equipment occurred, the monitoring system could conclude one or many maintenance strategy and guide the maintenance staff to do a repair job or perform it automatically. With this mechanism, the maintenance time can be reduced; equipment effectiveness and product quality will be improved.

135.4.3 Monitoring of Motor Operational Parameters

The monitoring system on-line collects motor operational parameters (such as temperature, current, voltage, and speed, etc.), compares those parameters with their range when motor functions well and predicts motor running condition. The monitoring steps are listed as follows:

(1) Collecting the motor data listed above in 1 year or getting the data specification from manufacturer
(2) Constructing the monitoring limits for motor operational parameters.
(3) Save the model into the maintenance and diagnosis database.
(4) Performing on-line data collection and compare the on-line data with the limits saved in the maintenance and diagnosis database.
(5) If the parameter value passes beyond a limit, the monitoring system will send associated alarms and list the corresponding maintenance strategies

135.5 Conclusion

Condition monitoring is used for increasing machinery availability and machinery performance, reducing consequential damage, increasing machine life, reducing spare parts inventories, and reducing breakdown maintenance. An efficient condition monitoring scheme is capable of providing warning and predicting the faults at early stages. This monitoring system can monitor operational status of production system in real time, predict fault before big damage comes and implement predictive maintenance ahead of time. Thus it can reduce maintenance cost and increase availability and effectiveness of production system.

Acknowledgments The author would like to thank the support on the education department project (Y201017103) granted by the education department of Zhejiang Province, People's Republic of China.

References

1. Habetler TG, Harley RG (2006) Non-stationary motor fault detection using recent quadratic time-frequency representations. IEEE Ind Appl Soc 5:2333–2339
2. Habetler TG, Harley RG (2005) Diagnosis of rotor faults in brushless DC (BLDC) motors operating under non-stationary conditions using windowed fourier ridges. IEEE Ind Appl Soc 1:26–33
3. Nandi S, Toliya HA (1999) Condition monitoring and fault diagnosis of electrical machines—a review. Conf Ind Appl Conf 1:197–202
4. Bemieri G, Betta CL (1997) Real-time re-configuration of multi-DSP measurement stations. In: Proceeding IMEKO TC-4 international symposium Glasgow (Scotland), pp 253–256
5. Baccigalupi A, Bemieri A, Pietrosanto A (1997) A digital-signal-processor-based measurement system for on-line fault detection. IEEE Trans Instrum Meas 46:731–736
6. Jeffrey HN (2001) Electric utility consideration for circuit breaker monitoring. In: Proceedings of the 2001 IEEE/PES transmission and distribution conference and exposition, 28 Oct–2 Nov 2002, Atlanta, Georgia, pp 1094–1097
7. Koch D, Garson R (1994) Square D Type FB4 SF6 circuit breaker contact resistance. Minutes of the sixty-first annual international conference of Doble clients, 1994, pp 1145–1150
8. Gockenbach E, Werle P, Wassenberg V, Borsi H (2001) Monitoring and diagnosis systems for dry type distribution transformers. 7th international conference on solid dielectrics (ICSD), Eindhoven Netherlands, June 2001, pp 2101–2108
9. HAlm-Owoo AK, OSuen K (2002) Applications of fault detection and diagnostic techniques for refrigeration and air conditioning: a review of basic principles. In: Proceeding of the institution of mechanical engineers, ∼01.216, number 3, pp 121–132
10. Haves P, Khalsa SK (2000) Model-based performance monitoring: review of diagnostic methods and chiller case study. ACEEE 2000 summer study on energy sciences in buildings, efficiency and sustainability, pp 589–595

Chapter 136
Mismatch Detection of Distributed Target in Non-Gaussian Clutter

Tao Jian

Abstract The mismatch detection of radar-distributed target is analyzed when the signal may not lie in alignment with the look direction, in non-Gaussian clutter modeled as a spherically invariant random vector. Based on the generalized likelihood ratio test design procedure, a distributed target detector is obtained, by utilizing a priori target scatterer density. It is showed that, the GLRT-based detector can reject the unwanted signal effectively. Moreover, the effects of the correlated fluctuating scatterers and of the additive thermal noise on the detection performance are also evaluated.

Keywords Radar distributed target · Mismatch detection · Non-Gaussian clutter · Thermal noise

136.1 Introduction

A high-resolution radar (HRR) can resolve a target into a number of scatterers, which is referred to as a so-called distributed target (or range-spread target) [1, 2]. As some target energy spreads into the contiguous range cells, the adaptive detection strategies for point-like targets may fail for distributed targets [1]. It is also known that, the background clutter may no longer be modeled accurately as a Gaussian random variable (RV) for HRR [2]. At the higher range resolution, the radar system receives target-like spikes clutter, which can be suitably modeled by

T. Jian (✉)
Research Institute of Information Fusion,
Naval Aeronautical and Astronautical University,
Yantai 264001, China
e-mail: iamjiantao@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_136,

a spherically invariant random vector (SIRV) [3]. An increasing interest has been directed toward the distributed target detection schemes optimized under non-Gaussian clutter.

However, note that, in practical application, the signal direction vector is usually mismatched with the steering vector [4]. In this work, the mismatch detection of distributed target is analyzed when the signal may or may not lie in alignment with the look direction.

136.2 Problem Formulation

It is assumed that data are collected from N sensors and deal with the problem of detecting the presence of a distributed target across K range cells $\mathbf{z}_t$s, $t = 1, \ldots, K$. Moreover, the clutter-dominant environment is considered, and the internal noise is ignored. Hence the problem of detecting a distributed target in additive clutter can be formulated in the binary hypotheses test:

$$\begin{aligned} \mathrm{H}_0 &: \mathbf{z}_t = \mathbf{c}_t, \qquad t = 1, \ldots, K \\ \mathrm{H}_1 &: \mathbf{z}_t = \alpha_t \mathbf{p} + \mathbf{c}_t, \quad t = 1, \ldots, K \end{aligned} \tag{136.1}$$

where $\mathbf{p}$ denotes the normalized signal direction vector, and the unknown parameters α_ts, $t = 1, \ldots, K$ account for both the target and the channel effects. Note that, for the uniform linear array, $\mathbf{p} = \left(1, e^{j\phi_1}, e^{j2\phi_1}, \ldots, e^{j(N-1)\phi_1}\right)^{\mathrm{T}} \Big/ \sqrt{N}$, where ϕ_1 denotes a constant phase shifting and $(\cdot)^{\mathrm{T}}$ represents the transpose.

Based on the SIRV distribution [3], the N-dimensional clutter vector $\mathbf{c}_t$ can be given by

$$\mathbf{c}_t = \sqrt{\tau_t} \cdot \boldsymbol{\eta}_t, \ t = 1, \ldots, K \tag{136.2}$$

where $\boldsymbol{\eta}_t = (\eta_t(1), \eta_t(2), \ldots, \eta_t(N))^{\mathrm{T}}$; $\eta_t(n)$s, $n = 1, \ldots, N$ are zero-mean complex circular Gaussian RVs; and the texture component τ_t is a semipositive real RV with unknown probability density function (PDF) f_τ. Moreover, $\boldsymbol{\eta}_t$ and τ_t are assumed to be independent. And the positive definite covariance matrix $\boldsymbol{\Sigma}$ is defined as

$$\boldsymbol{\Sigma} = E\left\{\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{\mathrm{H}}\right\}, \ t = 1, \ldots, K \tag{136.3}$$

To simplify the analysis, only one equivalent scatterer is supposed to occupy one resolution cell. The scatterers of a distributed target usually occupies only a fraction of K range cells, while the echo amplitudes of range cells occupied by target scatterers are significantly greater than those of range cells with clutter only. Let h_0 ($h_0 \leq K$) denote the number of equivalent scatterers; and Θ_{h_0} denote index set of cells with target equivalent scatterers.

Accordingly, the binary hypotheses test (136.1) can be simplified as

$$\begin{aligned} \mathrm{H}_0 &: \mathbf{z}_t = \mathbf{c}_t, \quad t \in \Theta_{h_0} \\ \mathrm{H}_1 &: \mathbf{z}_t = \alpha_t \mathbf{p} + \mathbf{c}_t, \quad t \in \Theta_{h_0} \end{aligned} \tag{136.4}$$

Based on the previous assumptions, the joint conditional PDF of $\mathbf{z}_t$s with $t \in \Theta_{h_0}$ under each hypothesis is given by

$$\prod_{t\in\Theta_{h_0}} f(\mathbf{z}_t|\mathrm{H}_0, \tau_t) = \prod_{t\in\Theta_{h_0}} \frac{1}{\pi^N \tau_t^N \det(\boldsymbol{\Sigma})} \times \exp\left[-\frac{1}{\tau_t}\mathbf{z}_t^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\mathbf{z}_t\right] \tag{136.5}$$

and

$$\prod_{t\in\Theta_{h_0}} f(\mathbf{z}_t|\mathrm{H}_1, \tau_t) = \prod_{t\in\Theta_{h_0}} \frac{1}{\pi^N \tau_t^N \det(\boldsymbol{\Sigma})} \times \exp\left[-\frac{1}{\tau_t}(\mathbf{z}_t - \alpha_t\mathbf{p})^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}(\mathbf{z}_t - \alpha_t\mathbf{p})\right] \tag{136.6}$$

where $\det(\cdot)$ denotes the determinant and $(\cdot)^{\mathrm{H}}$ implies the conjugate transpose.

136.3 Mismatch Detection

According to the Neyman–Pearson criterion, the optimum solution to the hypotheses testing problem (136.1) is the likelihood ratio test, but the uniformly most powerful test does not exist [5]. Herein, the design procedure based on generalized likelihood ratio test (GLRT) [6] is used, which can be begun as follows

$$\frac{\max\limits_{\tau_t}\max\limits_{\alpha_t}\prod\limits_{t\in\Theta_{h_0}} f(\mathbf{z}_t|\alpha_t, \tau_t, \mathrm{H}_1)}{\max\limits_{\tau_t}\prod\limits_{t\in\Theta_{h_0}} f(\mathbf{z}_t|\tau_t, \mathrm{H}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} T \tag{136.7}$$

The maximum likelihood (ML) estimates of α_t and τ_t in (136.7) can be expressed as

$$\hat{\alpha}_t = \mathbf{p}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\mathbf{z}_t / \left(\mathbf{p}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\mathbf{p}\right) \tag{136.8}$$

$$\mathrm{H}_0 : \hat{\tau}_t = \mathbf{z}_t^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\mathbf{z}_t / N \tag{136.9}$$

and

$$\mathrm{H}_1 : \hat{\tau}_t = (\mathbf{z}_t - \hat{\alpha}_t\mathbf{p})^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}(\mathbf{z}_t - \hat{\alpha}_t\mathbf{p}) \Big/ N \tag{136.10}$$

By replacing the unknown parameters with their ML estimates under each hypothesis, the GLRT-based statistic is denoted as [2]

$$\lambda = -N \sum_{t \in \Theta_{h_0}} \ln(1 - w_t) \underset{H_0}{\overset{H_1}{\gtrless}} T_1 \tag{136.11}$$

where the quantity w_t is defined by

$$w_t = \frac{\left|\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{z}_t\right|^2}{\left(\mathbf{z}_t^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{z}_t\right)\left(\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{p}\right)} \tag{136.12}$$

In sort ascending, the order statistics of w_ts, $t = 1,\ldots,K$ are denoted as $0 \le w_{(1)} \le \cdots \le w_{(k)} \le \cdots \le w_{(K)} \le 1$. According to (136.11), by exploiting the h_0 largest values of $w_{(t)}$s, $t = 1,\ldots,K$, the GLRT-based statistic can also be given by [2]

$$\lambda = -N \sum_{k=K-h_0+1}^{K} \ln\left(1 - w_{(k)}\right) \tag{136.13}$$

As to the signal direction vector $\mathbf{p}$, the steering vector of array is assumed as $\mathbf{q} = \left(1, e^{j\varphi_2}, e^{j2\varphi_2}, \ldots, e^{j(N-1)\varphi_2}\right)^{\mathrm{T}} \Big/ \sqrt{N}$, where ϕ_2 denotes a constant phase shifting. For the direction-mismatched case, $\phi_2 \neq \phi_1$. For notational convenience, let

$$A_{\mathbf{p}} = \sqrt{\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{p}} \tag{136.14}$$

and

$$A_{\mathbf{q}} = \sqrt{\mathbf{q}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{q}} \tag{136.15}$$

For the direction-matched case, the maximum signal-to-clutter power ratio (SCR) is denoted as

$$\mathrm{SCR}_{\mathbf{qq}} = \sum\nolimits_{t=1}^{K} |\alpha_t|^2 \mathbf{q}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{q} \Big/ Kb \tag{136.16}$$

where b indicates the mean of τ_t. However, the direction-mismatched case will result in the SCR loss. When the array is steering in the direction corresponding to $\mathbf{q}$, the SCR is

$$\mathrm{SCR}_{\mathbf{pq}} = \frac{\sum_{t=1}^{K} |\alpha_t|^2}{Kb} \frac{\left|\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{q}\right|^2}{\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{p}} \tag{136.17}$$

Herein, the inner product measure of distance $\cos^2\theta$ can be used to relate the SCR to the maximum SCR as follows [4]

$$\cos^2\theta = \frac{\mathrm{SCR}_{\mathbf{pq}}}{\mathrm{SCR}_{\mathbf{qq}}} = \frac{\left|\mathbf{p}^{\mathrm{H}} \boldsymbol{\Sigma}^{-1} \mathbf{q}\right|^2}{A_{\mathbf{q}}^2 A_{\mathbf{p}}^2} \tag{136.18}$$

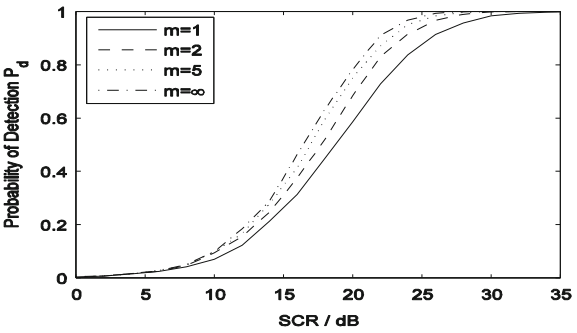


Fig. 136.1 P_{d} versus SCR of GLRT-based detector for $P_{\mathrm{fa}} = 10^{-4}$, $K = 15$, $h_0 = 3$, $N = 2$, $L = 1$, $m = 1, 2, 5, \infty$, $\rho = 0$, $\cos^2\theta = 1$

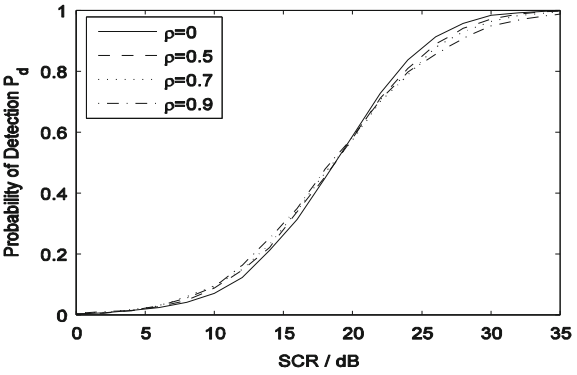


Fig. 136.2 P_{d} versus SCR of GLRT-based detector for $P_{\mathrm{fa}} = 10^{-4}$, $K = 15$, $h_0 = 15$, $N = 2$, $L = 1$, $m = 1$, $\rho = 0, 0.5, 0.7, 0.9$, $\cos^2\theta = 1$

The parameter $\cos^2\theta$ accounts for the performance loss resulting from the direction-mismatched case. The direction-matched case means $\cos^2\theta = 1$ and $\mathbf{q} = \mathbf{p}$; while the totally direction-mismatched case means $\cos^2\theta = 0$.

136.4 Performance Assessment

In this section, the detection performance of the GLRT-based detector is analyzed as well as their rejection performance for unwanted mismatched signal.

For the underlying mixing distribution f f_τ, the gamma distribution is adopted with the following PDF

$$f_\tau(x) = (L/b)^L x^{L-1} e^{-(L/b)x} \Big/ \Gamma(L), \quad x \geq 0 \tag{136.19}$$

where $\Gamma(\cdot)$ is the gamma function, b indicates the mean of the distribution and L controls the deviation from Gaussian statistics. The smaller L is, the more spikier will appear in clutter.

The detection probability P_{d} of the GLRT-based detector depends on α_ts, $t = 1,\ldots,K$, $\mathbf{q}$, $\boldsymbol{\Sigma}$ and b only through the input SCR defined as [2]

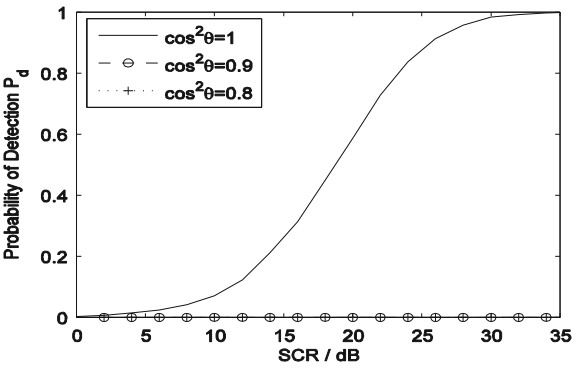


Fig. 136.3 P_{d} versus SCR of GLRT-based detector for $P_{\mathrm{fa}} = 10^{-4}$, $K = 15$, $h_0 = 3$, $N = 2$, $L = 1$, $m = 1$, $\rho = 0$, $\cos^2\theta = 1, 0.9, 0.8$

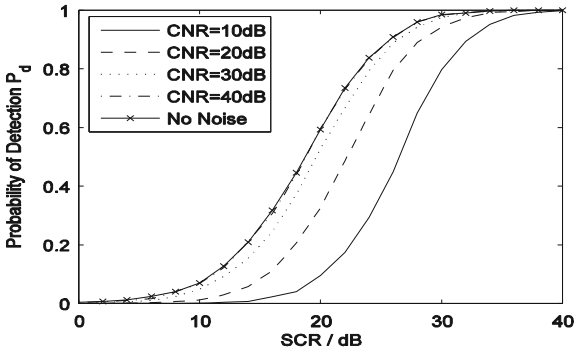


Fig. 136.4 P_{d} versus SCR of GLRT-based detector for $P_{\mathrm{fa}} = 10^{-4}$, $K = 15$, $h_0 = 3$, $N = 2$, $L = 1$, $m = 1$, $\rho = 0$, $\cos^2\theta = 1$, $CNR = 10$, 20, 30, 40 dB,∞

$$\mathrm{SCR} = \mathrm{SCR}_{\mathbf{qq}} \tag{136.20}$$

The performance assessment is carried out resorting to Monte Carlo simulations. In order to limit the computational burden, it is assumed $P_{\mathrm{fa}} = 10^{-4}$ and $K = 15$.

It is customary to model the $|\alpha_t|^2 \mathrm{s}, t = 1, \ldots, K$ as Chi-squared RVs. In addition, it would be significant to evaluate the impact of correlation degree among the target scatterers on the performance. To this end, it is assumed that the $|\alpha_t|^2 \mathrm{s}, t = 1, \ldots, K$ are drawn from an exponentially correlated random sequence with one-lag correlation coefficient ρ, and the covariance between $|\alpha_h|^2$ and $|\alpha_k|^2$ is given by [7]

$$cov(|\alpha_h|^2, |\alpha_k|^2) = \overline{A_h^2} \cdot \overline{A_k^2} \cdot \rho^{|h-k|^m}, \; h, k \in \{1, \cdots, K\} \tag{136.21}$$

where $\overline{A_t^2}$ denotes the mean square values of $|\alpha_t|$. Moreover, the integer number m rules the depth of the fluctuation. Special cases are $m = 1$, corresponding to a Rayleigh-distributed amplitude, and $m = \infty$, corresponding to a non-random amplitude. The coefficient ρ rules the depth of correlation: $\rho = 1$ corresponds to the total dependence while $\rho = 0$ corresponds to independence.

Firstly, in Fig. 136.1, the effects of different fluctuation degrees ($m = 1$, 2, 5, ∞) on the detection performance are evaluated for the GLRT-based detector. It is

observed that, the detection performance improves as the fluctuation degree weakens, while the performance gain increases as the SCR increases.

Secondly, in Fig. 136.2, the effects of different correlations ($\rho = 0$, 0.5, 0.7, 0.9) between scatterers on the detection performance are evaluated. It is implied that, the curves of different correlations crosses at $P_d = 0.6$. More precisely, in the range of high values of P_d, the detection performance decreases as the correlation degree enhances, while the performance increases with enhancing correlation in the range of low values of P_d.

Furthermore, the effects of different direction-mismatched cases on the detection performance are evaluated for $\cos^2\theta = 1, 0.9, 0.8$ in Fig. 136.3. It is indicated that, for all match cases with $\cos^2\theta \leq 0.9$, the GLRT-based detector cannot detect the target, which shows that the GLRT-based detector can reject the unwanted signal effectively.

Finally, note that, according to (136.1), the development of the GLRT-based detector was based on the assumption of clutter-dominated disturbance, which means a large clutter-to-noise ratio (*CNR*). Thus, the additive thermal white noise was neglected. However, it is still necessary to assess the effect of noise on the detection performance for the GLRT-based detector. Figure 136.4 refers to the detection performance for $CNR = 10$, 20, 30, 40 dB and no noise cases. It shows that, the detection performance is degenerated as the noise increases, especially for low SCR. Moreover, the noise can completely be neglected for $CNR \geq 40$ dB.

136.5 Conclusions

In this work, the mismatch detection of radar-distributed target is addressed when the signal may not lie in alignment with the look direction, in non-Gaussian clutter modeled as a SIRV. Based on the GLRT, a distributed target detector is introduced. And then the inner product measure of SCR loss is given. The experimental results show that, the GLRT-based detector can reject the unwanted signal effectively. As the correlation degree enhances, the detection performance decreases in the range of high values of detection probability, while the performance increases in the range of low values. In addition, the detection performance improves as the fluctuation degree of target amplitudes weakens.

Acknowledgments This work was supported by National Natural Science Foundation of China (61102166) and Scientific Research Foundation of Naval Aeronautical and Astronautical University for Young Scholars (HYQN201013).

References

1. Wehner DR (1995) High-Resolution Radar. Artech House, Boston
2. He Y, Jian T, Su F, Qu CW, Gu XF (2010) Novel range-spread target detectors in non-Gaussian clutter. IEEE Trans Aerosp Electron Sys 46(3):1312–1328

3. He Y, Jian T, Su F, Qu CW, Ping DF (2010) CFAR assessment of covariance matrix estimators for non-Gaussian clutter. Sci China Inf Sci 53(11):2343–2351
4. Robey FC, Fuhrmann DR, Kelly EJ, Nitzberg R (1992) A CFAR adaptive matched filter detector. IEEE Trans Aerosp Electron Sys 28(1):208–216
5. Bose S, Steinhardt AO (1996) Optimum array detector for a weak signal in unknown noise. IEEE Trans Aerosp Electron Sys 32(3):911–922
6. Van Trees HL (2001) Detection, estimation, and modulation theory, part I: detection, estimation and linear modulation theory. Wiley, New York
7. Conte E, De Maio A, Ricci G (2001) GLRT-based adaptive detection algorithms for range-spread targets. IEEE Trans Signal Process 49(7):1336–1348

Chapter 137
Relaxed Hybrid Forecasting and its Application to Railway Passenger Turnover

Xuejun Chen and Suling Zhu

Abstract Conventional hybrid forecasting model has been widely used in various forecasting problems, but the sum of weights is limited to 2. An improved hybrid model named as relaxed hybrid model is proposed in this study, where the weights are relaxed to positive real data. The weights are searched by particle swarm optimization algorithm with compress factor technique. The relaxed hybrid model is employed to railway passenger turnover forecasting. The forecasting results show that our proposed model is an effective model for nonlinear time series forecasting.

Keywords Relaxed hybrid model · Forecasting · Railway passenger turnover · CFPSO

137.1 Introduction

Transportation is an important trade, which is critical for people's daily life. There are many passengers every day in China, and some are tourists. Archer and Morley well documented the benefits of accurate forecasts in the tourism forecasting

X. Chen
Gansu Meteorological Information and
Technique Support and Equipment Centre,
Lanzhou 730020, People's Republic of China
e-mail: xuejunchen1971@163.com

S. Zhu (✉)
School of Mathematics and Statistics,
Lanzhou University, Lanzhou 730000,
People's Republic of China
e-mail: zhusuling02@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_137,

literature [1]. Actually, the accurate forecasting is also important for the transportation. So the forecasting with higher precision and rational planning is essential for transportation management. Figure 137.1 shows descriptive statistics pie chart in China transportation, and it shows that highway and railway are important components in China transportation. There are many forecasting techniques for transportation, such as gray model (GM), ARMA, Neural network and SVM. However, none of these methods can yield the results with desired accuracy for all forecasting problems because of their drawbacks, just as Moghram and Rahman concluded [2, 3]. Therefore we propose an improved hybrid model based on the conventional model for railway forecasting.

Both theoretical and empirical results suggest that combining different methods can be an effective and efficient way to increase forecasting precision [4]. Hybrid model is one kind of combining techniques which has been widely applied in various domains, such as price forecasting of electricity markets [5, 6], stock market [7], power load [8] and celebrity endorsement effects on consumer attitudes [9]. So we apply the hybrid model to transportation forecasting. In order to overcome the drawbacks of the conventional hybrid model, we propose one improved hybrid model named as relaxed hybrid model because the coefficients are relaxed to positive data. The weights are determined by particle swarm optimization algorithm with compress factor (CFPSO) method. Particle swarm optimization (PSO) [10] is an effective method for the optimization problems, and it is used in many domains.

Because the railway passenger turnover has an increasing trend with the time, we apply the Gray model to forecast it. The difference between actual data and the forecast shows seasonal fluctuations, so we apply seasonal ARIMA model to forecast the difference series. The forecasts for railway passenger turnover are the weighted sum of the forecasts of GM and seasonal ARIMA.

137.2 The Forecasting Techniques

137.2.1 The Relaxed Hybrid Model

The hybrid model is one conventional error correction technique. Generally speaking the time series suffers from trend and nonlinear feature. However, there exists no one model which is suitable for all forecasting problems. The hybrid model is designed for nonlinear forecasting with higher precision compared with the single model. Assume that $y_t(t = 1, 2, \ldots, N)$is the historical time series data, N is the sample number, $\hat{L}_t$is the forecasting data for time t from one prespecified model and $e_t = y_t - \hat{L}_t$ is the forecasting residual or difference. The weighted hybrid model can be described as:

$$\hat{y}_t = \hat{a}\hat{L}_t + \hat{b}\hat{E}_t \tag{137.1}$$

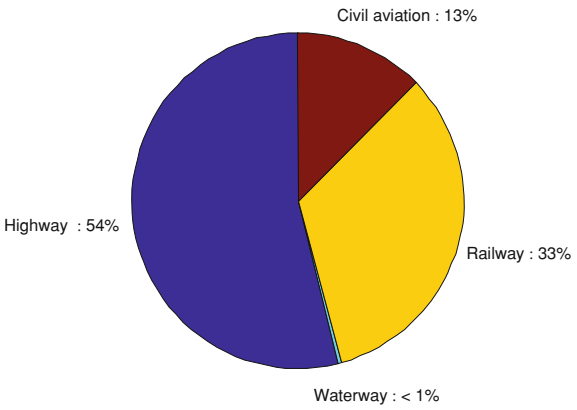


Fig. 137.1 Descriptive statistics pie chart in China transportation

where $\hat{E}_t$ is the forecasting data of the residual, the weights satisfy $\hat{a} + \hat{b} = 2$ and $\hat{y}_t$ is the forecasting value obtained by the weighted hybrid model. In the conventional hybrid model, the weight coefficients are $\hat{a} = \hat{b} = 1$. For different time series, the nonlinear and linear (trend) make different contributions to the forecasting results, so the weighted hybrid model is an effective method in data analysis. In order to provide more effective forecasts and rational hybrid technique, weight coefficients of the proposed hybrid model are relaxed to positive real data, so it is called relaxed hybrid model whose weights are searched by CFPSO.

137.2.2 Particle Swarm Optimization Algorithm with Compress Factor

Inspired by the flocking behavior of the birds, Eberhart and Kennedy [10] firstly introduced particle swarm optimization (PSO) in 1995. The PSO algorithm works by initializing a flock of birds randomly over the searching space, where every bird is called as a "particle" [10, 11]. In the D-dimensional space, each particle has a position vector $X_i = (x_{i1}, x_{i1}, \ldots, x_{iD})$ and a velocity vector $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$, where n is the number of particles in the swarm. The best position visited by the total particle swarm (gbest) is denoted as vector $P_{\text{gb}} = \left(p_{g1}, p_{g2}, \ldots, p_{gD}\right)$ and the best position visited by the ith particle (pbest) is denoted as $P_{\text{ib}} = (p_{i1}, p_{i2}, \ldots, p_{iD})$.

The original PSO algorithm can be described as follows [11]:

$$v_{id}(t+1) = v_{id}(t) + C_1 r_1 [p_{id}(t) - x_{id}(t)] + C_2 r_2 \left[p_{gd}(t) - x_{id}(t)\right] \tag{137.2}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) 1 \le i \le n 1 \le d \le D \tag{137.3}$$

The formula of CFPSO about velocity can be described as:

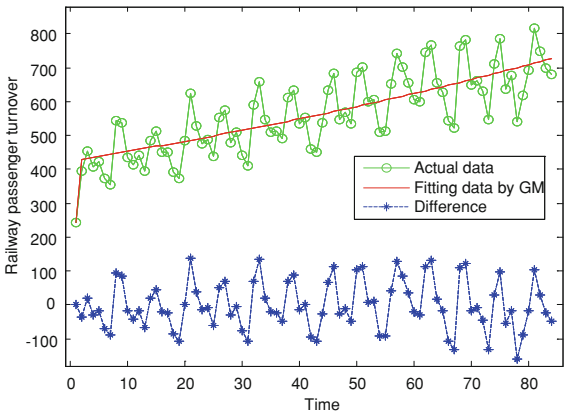


Fig. 137.2 Actual data, fitting data of GM and the difference

$$v_{id}(t+1) = \varphi\{v_{id}(t) + C_1 r_1[p_{id}(t) - x_{id}(t)] + C_2 r_2[p_{gd}(t) - x_{id}(t)]\},$$
$$\varphi = 2\Big/\left|2 - C - \sqrt{C^2 - 4C}\right| C = c_1 + c_2 \tag{137.4}$$

where c_1 and c_2 are positive constants, which are commonly set to be $c_1 = c_2 = 2$, r_1 and r_2 are random numbers between zero and one. For the CFPSO, $C = c_1 + c_2 > 4$. In this paper, $c_1 = 2.8$ and $c_2 = 1.3$.

137.2.3 Weights Determination

Let the weight coefficients of relaxed hybrid model be a particle which is a D-dimensional vector ($D = 2$). The detailed procedures of weights determination are as:

Step 1. Randomly initialize position and velocity of particle.

Step 2. Calculate the adaptive degree by $F_i = \sum_{t=1}^{n} \left|\frac{y_t - \hat{y}_{it}}{y_t}\right|$. Save the current location $x_{i,d}(t)$ and F_i to pbest, and save the best location and adaptive degree of pbest to gbest.

Step 3. Generate new particles for the next generation. The new particles will be generated using Eqs. (137.2–137.4).

Step 4. Compare the adaptive degree for every particle with its experienced best position. If the new particle is better than the past one, it saved as the current best position.

Step 5. Compare the adaptive degree of all the current pbest and gbest, update gbest.

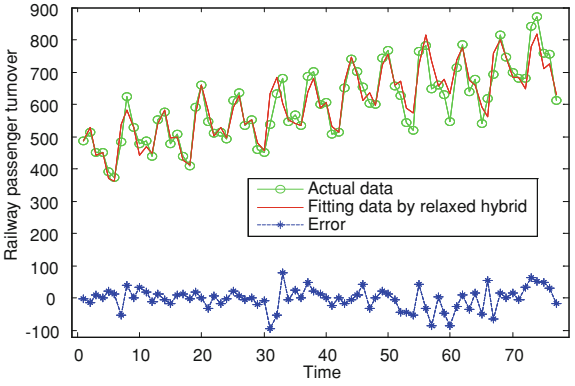


Fig. 137.3 The fitting and forecasting results of relaxed hybrid model

Step 6. Check whether the termination criterion is satisfied. If the termination criterion [the maximum iteration number (100)] is reaching, continue to Step7. If the criterion is not satisfied, go back to Step 3.

Step 7. Terminate the searching process and output the results.

137.3 Case Study

137.3.1 Data Set and Accuracy Criterion

The data about the monthly railway passenger turnover is downloaded from http://202.201.7.22:82/. The gray model is selected to forecast it, and Fig. 137.2 shows the actual data, fitting values and residuals. It can be seen that the forecasts of GM reveal the trend of passenger turnover, but the fluctuations caused by different months are neglected. Therefore, the seasonal ARIMA (SARIMA) is selected to correct the forecasting residuals. The weights in the relaxed model are searched by CFPSO. Two popular forecasting precision indexes mean absolute error (MAE) and mean absolute percentage error (MAPE) are applied in this study and they are defined as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\%, \quad MAE = \frac{\sum_{t=1}^{n}|y_t - \hat{y}_t|}{n}$$

where y_t and $\hat{y}_t$ are the actual data and the forecasting value at time t, respectively.

The accuracy improvement (AI) is defined as:

$$\mathrm{AI} = \frac{S - S_h}{S} \times 100\%$$

Table 137.1 The fitting and forecasting errors of GM, hybrid and relaxed hybrid

Model	Fitting			Forecasting		
	MAPE (%)	MAE	AI (%)	MAPE (%)	MAE	AI (%)
GM	10.32	59.7858	61.09	10.06	74.2198	45.84
Hybrid	9.51	53.9864	56.91	9.08	64.135	37.32
Relaxed	3.96	23.2604		5.19	40.2002	

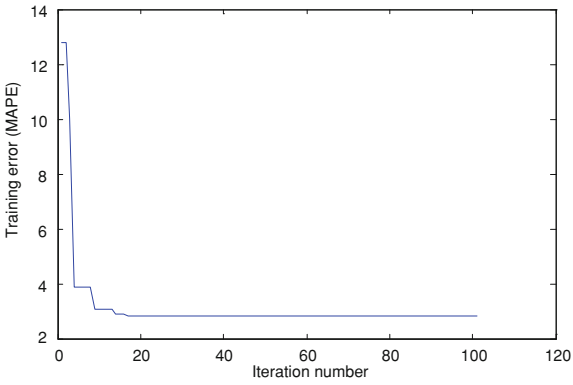


Fig. 137.4 Trend of the fitting error with the increasing iteration number of relaxed hybrid model by CFPSO

where S_h is the MAE for the relaxed model, S represents the mean absolute error obtained from the single model, such as hybrid and GM. It can be concluded that if AI > 0, the relaxed hybrid model performs better and if AI < 0, the relaxed model does not overcome the deficiency of single model.

137.3.2 The Application of Relaxed Hybrid Model

Figure 137.3 shows the fitting and forecasting results of relaxed hybrid model, and we can conclude that the relaxed hybrid model provides more accurate forecasting results than GM. More specifically, the changes of the time series obtained from relaxed hybrid model are consistent with the actual. In order to explicitly show the fitting and forecasting ability, Table 137.1 lists the fitting and forecasting errors. For the fitting procedure, the MAPE of GM, hybrid and relaxed hybrid model are 10.32%, 9.51% and 3.96%, respectively. From the MAE index we can see that the relaxed hybrid model is the best one. The accuracy improvements (AI) over GM and hybrid models are 61.09% and 56.91%, respectively. The MAPE of forecasting results are 10.06%, 9.08% and 5.19%, respectively, so the relaxed hybrid is the best one.

Furthermore, the forecasting improvements over the GM and hybrid model are 45.84% and 37.32%, respectively. So the relaxed hybrid model is better than GM and hybrid. Figure 137.4 shows trend of the fitting error with the increasing

iteration number of relaxed hybrid model by CFPSO, and it can be seen that the fitting error is a constant after the 19th iteration.

137.4 Conclusion

In this study, we propose an improved hybrid model for railway passenger turnover forecasting. The fitting and forecasting results show that the proposed relaxed model is more effective than the conventional hybrid. In order to provide more effective forecasting technique, we will study the adaptive hybrid model based on the statistical learning models.

Acknowledgments This work is financially supported by the Natural Science Research Project of Education Department of Henan Province (2011A110018). The authors would like to thank the reviewers' suggestions.

References

1. Dharmaratne GS (1995) Forecasting tourist arrivals in Barbados. Ann Tour Res 22:804–818
2. Moghram l, Rahman S (1989) Analysis and evaluation of five short-term load forecasting techniques. IEEE Trans Power Syst 4:1484–1494
3. Zhu SL, Wang JZ, Zhao WG, Wang JJ (2012) A seasonal hybrid procedure for electricity demand forecasting in China. Appl Energy (in press)
4. Krogh A, Vedalsby J (1995) Neural network ensembles cross validation and active learning. Adv Neural Inf Process Syst 7:231–238
5. Amjady N, Keynia F (2008) Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. Int J Electri Power Energy Sys 30(9):533–546
6. Dong Y, Wang JZ, Jiang H, Wu J (2011) Short-term electricity price forecast based on the improved hybrid model. Energy Convers Manag 52(8–9):2987–2995
7. Wei LY, Chen TL, Ho TH (2009) A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market. Expert Syst Appl 36(8):11108–11117
8. Wu Q (2010) Power load forecasts based on hybrid PSO with Gaussian and adaptive mutation and Wv-SVM. Expert Syst Appl 37(1):194–201
9. Sheu JB (2009) A hybrid dynamic forecast model for analyzing celebrity endorsement effects on consumer attitudes. Math Comput Model 52:1554–1569
10. Eberhart RC, Kennedy J (1995) Particle swarm optimization. In: Proceeding of IEEE international conference on neural network, Perth, pp 1942–1948
11. Han F, Ling QH (2008) A new approach for function approximation incorporating adaptive particle swarm optimization and a priori information. Appl Math Comput 205:792–798

Chapter 138
Application and Research on Extended Kalman Prediction Algorithm in Target Tracking System

Zheng Zhang, Xiaowei Liu, Guangyou Yang and Jiang Min

Abstract Target tracking is a critical important part in robot controlling system and it usually cannot get satisfied results because of the complex condition. The Kalman Forecast Algorithm on the camera platform which has two degrees of freedom has been researched in this paper. The motion model has been built based on the particular characteristic of the target tracking system. An extend kalman filter algorithm has been made up on this paper for this nonlinear model to dynamically compensate the error in line procession. The results show that the new algorithm can get fine effect in control of noise, time and punctual. And it meets the requirement of real-time target tracking. Therefore, a great value of application is showing up.

Keywords Extended kalman prediction · Target tracking · Camera

138.1 Introduction

With the growth used in industry, military, medical, family and robots have made great economic benefits. And the technology of target tracking could be a critical important project, which has strict requirements such as high speed, high

Z. Zhang (✉) · X. Liu · G. Yang · J. Min
School of Mechanical Engineering, Hubei University of Technology, Wuhan, China
e-mail: zhangzheng_hb@sina.com

X. Liu
e-mail: Liuxiaowei1987@126.com

G. Yang
e-mail: pekka@126.com

J. Min
e-mail: jmzpxzz@sina.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_138,

precision, perfect intelligence, low cost and so on. This paper is based on the camera platform which has two degrees of freedom and been widely used in robot's eye sight model and other sides. The technologies concerts have been realized in FPGA.

As the algorithm of target tracking is complicated and has obvious error, many kinds of filter algorithms have come up to improve the effect. Kalmen filter is becoming widely used because of simple model, little data, and easy been brought up.

Most of the target tracking system may be nonlinear, so it will make linearization error by the linear of Kalman filter, addition to the random noise of target's unknown movement and noise of observation, the filter may generate larger error [1].

An extend Kalman filter is researched in this paper to estimate the noise of observation and target's unknown movement. Linear error of observation equation has been dynamically compensated and the impact of observation error has been reduced. The Matlab simulation is processed for the algorithm and a good result has been obtained.

138.2 Model

A simple and perfect mathematical model is needed to obtain a satisfactory accuracy of prediction of the state of the moving target. Kalman filter is commonly used in the dynamic system to estimate error of linear minimum variance, so the next state of target can be optimal predicted and the filter is stable [2]. The traditional Kalman filter algorithm consists mainly of state equation and observation equation.

$$\begin{aligned} X(k+1) &= \Phi(k+1)X(k) + \Gamma(k+1)w(k) \\ Y(k+1) &= H(k+1)X(k+1) + v(k+1) \end{aligned} \tag{138.1}$$

The $X\,(k+1)$ is a $n \times 1$ dimensional vector of original data in time $k+1$, Φ is the $n \times n$-dimensional system state transition matrix, Γ is the $n \times n$-dimensional system noise transfer matrix, $Y\,(k+1)$ is the $m \times 1$ dimensional vector of observations, H is the $m \times n$-dimensional observation matrix, $w\,(k)$ is $n \times 1$ dimensional dynamic random noise vector at k time, $v\,(k+1)$ is the $m \times 1$ dimensional observation noise vector at the $k+1$ time.

The linear minimum variance estimation of system state $X\,(j)$ based on k times previous observations $Y\,(1)$, $Y\,(2)$, $Y\,(3)$, …, $Y\,(k)$ is the orthogonal projection of $X\,(j)$ in $Y\,(k) = [y\,(1), \ldots, y\,(k)]T$, that is: $X\,(j \mid k) = E\,[X\,(j) \mid Y\,(k)]$.

When $j = k$, $X\,(k \mid k)$ is the optimal filter estimate: $X\,(k \mid k) = E\,[X\,(k) \mid Y\,(k)]$.

After getting observations $Y\,(k)$, considering of the estimation error, we can get the form of Kalman filter iteration equations:

$$\hat{X}(k|k) = \Phi(k,k-1)X(k-1|k-1) + K(k)[Y(k) - H(k)\Phi(k)(k-1) \quad X(k-1|k-1)] \tag{138.2}$$

$$K(k) = P(k,k-1)H^T(k)[H(k)P(k,k-1)H^T(k) + R(k)]^{-1} \tag{138.3}$$

$$P(k|k-1) = \Phi(k,k-1)P(k-1|k-1)\Phi^T(k,k-1) + \Gamma(k,k-1)Q \quad (k-1)\Gamma^T(k,k-1) \tag{138.4}$$

$$P(k|k) = [I - K(k)H(k)]P(k|k-1) \quad k = 0,1,2,\ldots \tag{138.5}$$

K (k) is Kalman gain matrix, P (k | k–1) is the prediction error covariance matrix. If $k = n$, the optimal estimation $\hat{X}$ (n | n) of true sequence is obtained according to Eq. (138.2).

Based on $\hat{X}(n \mid n)$, the optimal forecast of system state vector X ($n + m$) for linear discrete visual target tracking system can be obtained. And $m \geq 1$ that the optimal valuation of X ($n + m$) is obtained based on observations y (1), y (2), y (3), …, y (n):

$$\begin{aligned} X(n+m|n) &= \Phi(n+m,n)E[X(n)\,|Y(1),Y(2),Y(3),\ldots,Y(n)\,] \\ &= \Phi(n+m,n)X(n|n) \end{aligned} \tag{138.6}$$

After $\hat{X}1$, then we can get $\hat{X}2$ [3].

Next, we analyze the system model. Suppose the coordinates of the camera is the origin, the moving object is a particle, the current position of the target can be observed in any coordinate. From the current coordinates, the next moment coordinate data of target can be predicted, so the system can make a camera get offset angle that is predicted in advance go to the location. Now the speed of camera rotation does not depend on the target's motion, thereby we can eliminate a better camera lag phenomenon with a certain amount in advance [4].

The forecast error will exist in the course of predicting, we can continue to obtain new observations to fix it step-by-step. Combining with the first set of data, prediction $\hat{X}1$ is obtained and the deviation of $\hat{X}2$ is estimated after the second set of observations and so on. The Kalman filter algorithm has better prediction in the linear model case. The extended Kalman filter algorithm is used in the system to reduce error. According to the above analysis, we can establish target tracking model based on prediction algorithm of Kalman, which is shown in Figs. 138.1 and 138.2.

Polar coordinates are used in the model. We assume that the camera is located in the origin of coordinates. When the target has gone into the tracking range of the camera, the camera is estimating the location of the target at initial time through

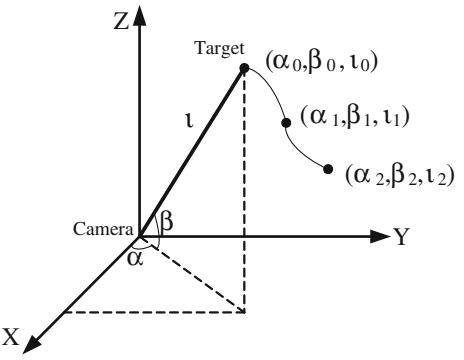


Fig. 138.1 The target tracking model based on Kalman prediction algorithm

the acquired data that is the target's distance $\hat{l}_0$, azimuth $\hat{\alpha}_0$, pitch $\hat{\beta}_0$ and speed weight Vx, Vy, Vz of various axial.

Dynamic model of system:

$$X(k+1) = \Phi X(k) + B(k)S(k) + w(k) \tag{138.7}$$

The original state vector of the relative movement of the target and the camera: $X = [x, y, z, Vx, Vy, Vz]T$

The x, y and z are the moving target's coordinates in the X, Y, Z direction;

$$X = \begin{bmatrix} l \\ \alpha \\ \beta \\ Vx \\ Vy \\ Vz \end{bmatrix} = \begin{bmatrix} \sqrt{x^2+y^2+z^2} \\ \arctan\left(\frac{y}{x}\right) \\ \arctan\left(\frac{z}{\sqrt{x^2+y^2}}\right) \\ Vx \\ Vy \\ Vz \end{bmatrix}, \Phi = \begin{bmatrix} 1 & 0 & 0 & T & 0 & 0 \\ 0 & 1 & 0 & 0 & T & 0 \\ 0 & 0 & 1 & 0 & 0 & T \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathrm{B} = \begin{bmatrix} -\frac{T^2}{2} & 0 \\ 0 & -\frac{T^2}{2} \\ -T & 0 \\ 0 & -T \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$S = [ax, ay, az]T$ is acceleration of the observatory, which is the coercion of system. And w is the dynamic noise.

T = tk + 1–tk is the frame interval of corresponding image sequence, $W(k) = [w1(k), w2(k), \ldots, w6(k)]$T is model noise.

Measurement equation is description of the relationship between motion measurements and the target's state.

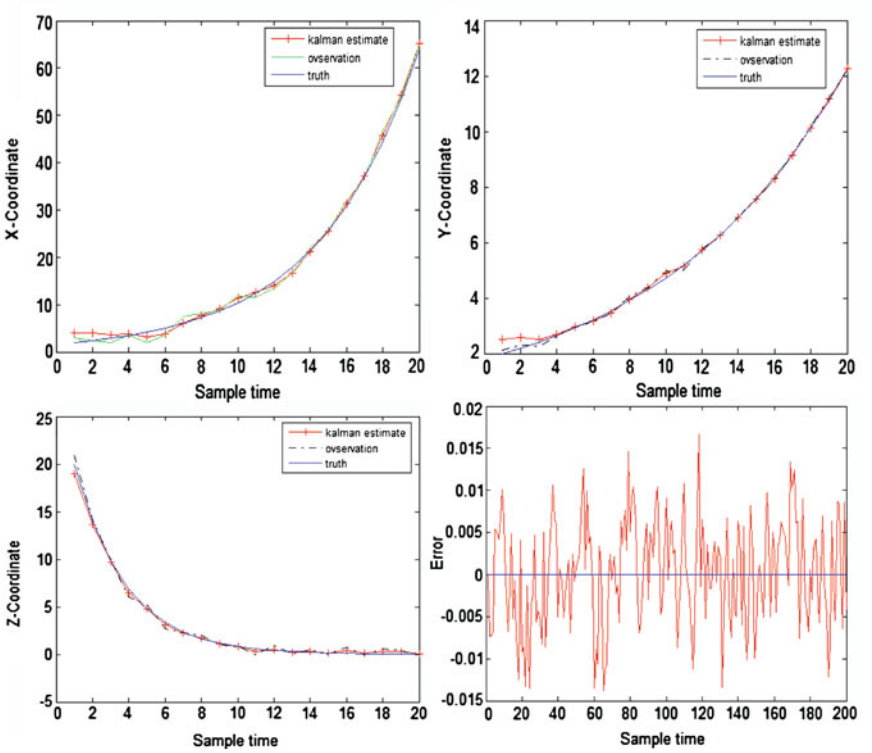


Fig. 138.2 The axial position of each prediction and estimation speed error

Observation model of system [5],

$$Y(k+1) = H(X(k+1)) + V(k+1) = \arctan(y/x) + V(k+1) \qquad (138.8)$$

Mean of $X(y)$ is zero and variance is δ_θ^2.

138.3 EKF (Extended Kalman Filter)

The system state equation is linear, while the measurement equation is nonlinear in rectangular coordinates. Therefore, the observation matrix of measurement equation is linearzed,

$$L(k) = \frac{\partial H[X(k+1)]}{\partial [X(k+1)]^T} = \left[\frac{\partial\theta}{\partial x} \quad \frac{\partial\theta}{\partial y} \quad \frac{\partial\theta}{\partial z} \quad \frac{\partial\theta}{\partial Vx} \quad \frac{\partial\theta}{\partial Vy} \quad \frac{\partial\theta}{\partial Vz}\right] \qquad (138.9)$$

Extended Kalman filter equations:
State prediction value: $X(k+1|k) = \Phi X(k|k)$
State error covariance: $P(k+1|k) = \Phi P(k|k) \Phi^T + Q(k)$
Filter gain:

$$K(k+1) = P(k+1|k) \ L^T(k) \ [L(k) P(k+1|k) L^T(k) + R(k+1)]^{-1}$$

State error update of covariance matrix:

$$P(k+1|k+1) = [1 - K(k+1)L(k)]P(k+1|k)$$

Predicted value update of the state [6]:

$$X(k+1|k+1) = X(k+1|k) + K(k+1)[Y(k+1) - L(k)X(k+1|k)] \quad (138.10)$$

By the 7 and 8th compositions of the nonlinear equations, iterative calculation can be carried out.

138.4 Results

In order to verify the correctness and validity of the model, simulation has done for maneuvering target tracking in three-dimensional space based on MATLAB. We assume that the target's measurement distance is L, azimuth is α, elevation angle is β and the sampling period is T, then the measurement equation could be as follows.

$$Y(k+1) = H(X(k+1)) + v(k+1) \quad (138.11)$$

The target's initial coordinates $X_0 = 5$, $Y_0 = 5$, $Z_0 = 3$ m; Initial velocity $V_x = 1$, $V_y = 2$, $V_z = 0$ m/s; Sampling period $T = 1$ s; sampling time $t = 15$ s.

The simulation results are shown. The results of the simulation analysis show that: because of the system noise and measurement noise, the target coordinates of the curve observed from the real coordinates of the target curve has some deviation from the extended Kalman filter is a clear error convergence results. After several operations, when the observation error further reduced, the Kalman filter prediction step updates the data will be more close to real data.

138.5 Conclusion

For the system random noise and measurement noise, Kalman filter has high accuracy and real-time reaction. In this paper, a target tracking model of three-dimensional space is established based on extended Kalman filter. With information have been observed, the best forecast of the next occurrence of the target

coordinates can be given. Simulation results show that the correctness and validity of the camera tracking model has good application prospects 2.

Acknowledgments This paper is supported by the national natural science fund project of Hubei province (number: 2010CDB02504), and the national natural science fund major project of Hubei province (number: 2010CBB0800).

References

1. Julier SJ (2004) Unscented filtering and nonlinear estimation. Proc IEEE 92(3):401–422
2. Guo W, Gao X (2007) Kalman prediction algorithm based on three-dimensional head tracking. Shanxi University of Technology, 3(25):121–125
3. Qu S, Shi Z (2001) Image-based vehicle tracking. Flight Dyn 13(l2):014–020
4. Xiao Q, Lei B (2006) Kalman filter-based camera tracking. Xi'an Institute of Science 2(2):056–061
5. Haritaolu I, Harwood D et al (1999) Active outdoor surveil-lance. In: Proceedings, international conference on image analysis and processing, pp 1096–1009
6. Shen Y, Gong H, Yan X (2007) Application of adaptive Kalman filter in target tracking system. Comput Simul 11(24):143–148

Chapter 139
Novel Hexagonal Opto-Electronic Detector and the Corresponding Object Recognition Criterions Used for Multi-Object Detection and Tracking

Yong Song, Kai Zhang, Xiang Li, Bangzhi Kang, Fuzhou Shang and Qun Hao

Abstract In this paper, in order to achieve multi-object detection and tracking, a novel hexagonal opto-electronic detector and the corresponding object recognition criterions are proposed. First, a hexagonal detector with 360° field of view is proposed. Second, to achieve the detection and tracking of multi-object with high speed, two object recognition criterions of the proposed detector, which are called object signal width criterion and horizontal scale ratio criterion, respectively, are proposed. Finally, the simulation experiments of the proposed hexagonal detector and the two criterions have been carried out, while the results show that the detection and tracking of multi-object can be achieved with high speed, which indicates that the proposed method will offer significant advantages in photoelectric detection, computer vision, virtual reality, augment reality, etc.

Keywords Linear CCD · Opto-electronic detector · Detection and tracking

Y. Song (✉) · K. Zhang · X. Li · B. Kang · F. Shang · Q. Hao
School of Opto-electronic, Beijing Institute of Technology, Beijing 100081, China
e-mail: yongsong@bit.edu.cn

K. Zhang
e-mail: 20040491@bit.edu.cn

X. Li
e-mail: baiya@bit.edu.cn

B. Kang
e-mail: 20904106@bit.edu.cn

F. Shang
e-mail: 2120100657@bit.edu.cn

Q. Hao
e-mail: qhao@bit.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_139,

139.1 Introduction

There are numerous cases in research, development and manufacturing where full-circle 360° photo-electric detection system is needed. In the research of such fields, Cao, Oh have acquired half spheral field of view by using spheral optical system [1, 2]. The resolutions of the proposed systems, however, are low due to the reason that all the information concentrates in a piece of circular image; Jarvis obtained the ability of panorama detection by placing cone-shaped reflector on CCD camera [3]. However, due to the distortion brought by cone-shaped reflector, the burden of image process is increased greatly; Ishiguro captured the panoramic images of indoor scenery by using rotation cameras [4]. A panoramic imaging system based on rotation cameras have been proposed by Thierry and Ryad [5]. Due to the use of mechanical rotation device, these systems have the characteristics of complicated structure and lower efficiency.

In this paper we have presented a novel hexagonal opto-electronic detector and the corresponding object recognition criterions used for multi-object detection and Tracking. First, a detector with hexagonal structure, which has 360° field of view, has been developed. To achieve the detection and tracking of multi-object in 360° field of view with high speed, object recognition criterions of object signal width criterion (OSWC) and horizontal scale ratio criterion (HSRC) are proposed. Consequently, an algorithm based on these criterions, which are suitable for the proposed hexagonal detector, can be developed. Finally, simulation experiments have been carried out to verify the feasibility and validity of the proposed technology.

139.2 The Hexagonal Detector

The hexagonal detector is composed of six opto-electronic detection units. Each unit, which has 60° field of view, is composed of optical lens, linear CCD, etc. The optical axis of each unit is perpendicular to the linear CCD, and intersects with the central pixel of the corresponding linear CCD. The axis of optical system of each linear CCD goes through the center of hexagonal detector, and the angle between the axes of every two neighboring linear CCDs is 60°. As the result, a hexagonal detector, which has 360° field of view, has been developed, which is shown in Fig. 139.1.

It is assumed that the output signal of the kth linear CCD at the time of t is $I_k(x, t)$. Connect the output signal of six pieces of linear CCD in series, then, the spatial continuous detection signal of hexagonal detector at t can be expressed as

$$I_O(x,t) = \sum_{k=1}^{6} I_k(x,t) \qquad 1 \le x \le kN. \tag{139.1}$$

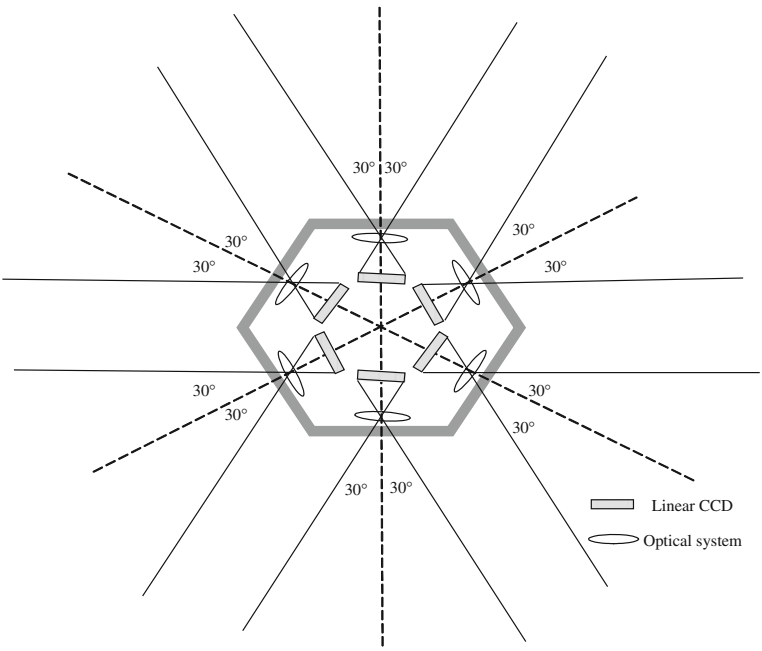


Fig. 139.1 The structure of hexagonal detector

Where k ($k = 1, 2, \ldots, 6$) is the series number of linear CCD in hexagonal detector. N is the number of active pixels of a linear CCD. x ($x = 1, 2, \ldots, 6\ N$) is the serial number corresponding to the object of the whole active pixels of six linear CCDs. As a result, the movement of objects in 360° field of view can be described as the variation of one-dimension data of linear CCDs.

In the proposed hexagonal detector, the field of view of each linear CCD is 60°. It is assumed that there is a moving object of n in the field of view corresponding to a linear CCD, and the serial number of the central pixel of the signal width in this linear CCD is i. Thus, the angular position of $\theta_n(t)$ of the object can be expressed as:

$$\theta_n(t) = \frac{360}{60N}x = \frac{6}{N}[(k-1)N + i]. \tag{139.2}$$

It is assumed that $\theta_n(t_1)$ and $\theta_n(t_2)$ are the angular positions of object n in angular coordinate system at the time of t_1 and t_2, respectively. As a result, the angular velocity of the moving object, which is denoted as ω_n, can be expressed as:

$$\overrightarrow{\omega}_n = \frac{\overrightarrow{\Delta\theta}}{\Delta t} = \frac{\theta(t_2) - \theta(t_1)}{\Delta t}. \tag{139.3}$$

In Eq. (139.3), in the case of $\Delta\theta > 0$, the moving direction of object is considered as the positive direction, otherwise it is considered as the negative direction.

139.3 Algorithm of Object Recognition and Tracking

In our study, self-adaptive background subtraction is used for achieving moving object detection. After the background refreshment is achieved, the object image of $D(k)$, which does not include the background, can be achieved.

139.3.1 Object Recognition Based on OSWC

It is assumed that $D[x,(t-1)]$ and $D(x, t)$, which contain moving objects with different sizes, are the object signals captured by hexagonal detector at the time of $t-1$ and t, respectively, there are object signals with the signal width of $Xn, t-1$ and Xn, t in $D[x,(t-1)]$ and $D(x, t)$, respectively. If the following condition, which is expressed as Eq. (139.4), is satisfied, $Xn, t-1$ and Xn, t are considered as the signals of the same object.

$$\left|X_{n,t} - X_{n,t-1}\right| < T_N. \tag{139.4}$$

T_N is the threshold, which represents the maximum variable of the signal width of the same objects.

In the case of several objects are overlapped in the field of view of hexagonal detector, the output signal corresponding to the objects is a signal with single width. It is assumed that the object image of $D[x, (t-1)]$ contains two moving objects of n and $n+1$, which have the signal width of $X_{n,\,t-1}$ and $X_{n,\,t}$, respectively and move towards each other. The objects of n and $n+1$ overlap in the signal of $D(x, t)$ at the time of t. As a result, the output signal corresponding to the objects in $D(x, t)$ is a signal with single width. Furthermore, the signal width of the objects, which is expressed as $X_{n',\,t}$, satisfies the condition of Eq. (139.5).

$$X_{n,t-1} + X_{n+1,t-1} > X_{n',t} > \left|X_{n,t-1} - X_{n+1,t-1}\right|. \tag{139.5}$$

On the other hand, the coordinates of object n and $n+1$ in $D(x,t)$, which are expressed as $C_{n,t}$ and $C_{n+1,\,t}$, respectively, can be determined by Eq. (139.6). The following is the description of the algorithm in the case of two objects overlap: (139.1) Firstl, confirm that there is no object signal of n and $n+1$, which have the signal widths of $X_{n,t-1}$ and $X_{n+1,\,t-1}$, respectively, in the object signal of $D(x, t)$ by determining the signal widths; (139.2) Confirm that the object signal of $D(x, t)$ contains a object signal of $X_{n',\,t}$, which can satisfy the condition of Eq. (139.5); (3) According to Eq. (139.6), $C_{n,\,t}$ and $C_{n+1,\,t}$ are calculated, respectively.

$$\begin{cases} C_{n,t} = C_{n',t} - \frac{1}{2}\left(X_{n',t} - X_{n,t-1}\right) \\ C_{n+1,t} = C_{n',t} + \frac{1}{2}\left(X_{n',t} - X_{n+1,t-1}\right) \end{cases}. \tag{139.6}$$

139.3.2 Object Recognition Based on HSRC

In the case of there is little difference among the object signal widths, the validity of the method described in Sect. 139.3.1 is limited. As a result, new criterion should be developed.

It is assumed that the object image of $D[x, (t-1)]$ contains an object signal, which has the signal width of $X_{n,t-1}$. The distance between the left edge of the object signal and the maximum point is $X_{\max,t-1}$, and the distance between the left edge of the object signal and the minimum point is $X_{\min,t-1}$. Similarly, the object image of $D(x,t)$ also contains an object signal, which has the signal width of $X_{n,t}$. In $D(x,t)$, the parameters corresponding to the $X_{\max,t-1}$ and $X_{\min,t-1}$ of $D[x,(t-1)]$ are $X_{\max,t}$ and $X_{\min,t}$, respectively. Then, the new criterion called HSRC can be expressed as:

$$T = T_1 + T_2 = k_1 \left| \frac{X_{\max,t}}{X_{\max,t-1}} - \frac{X_{n,t}}{X_{n,t-1}} \right| + k_2 \left| \frac{X_{\max,t}}{X_{\min,t}} - \frac{X_{\max,t-1}}{X_{\min,t-1}} \right| \le T_D. \qquad (139.7)$$

where T_1 is the difference between the width ratio of the maximum points ($X_{\max,t}/X_{\max,t-1}$) and the width ratio of the whole object signals ($X_{n,t}/X_{n,t-1}$) in continuous object images, T_2 is the difference between the width ratio of the maximum point to the minimum point in $D[x,(t-1)]$ ($X_{\max,t-1}/X_{\min,t-1}$) and it in $D(x, t)$ ($X_{\max,t}/X_{\min,t}$), k_1 and k_2 are the adjustment coefficient of T_1 and T_2, respectively, T_D is the threshold used for object recognition.

139.4 Simulation Experiment

The simulation experiment device is composed of two linear CCD cameras, imaging card, computer and moving objects, as shown in Fig. 139.2. In our experiment, object *a* and *b* have irregular outline, and the size of object *a* is bigger than it of object *b*. The moving images, which contain the moving object of *a* and *b*, were captured by camera A and camera B. Image data were inputted to the computer by using the image card. Finally, the image data were connected in series and processed by the software based on the proposed algorithms, which were discussed in Sect. 139.3.

139.4.1 Object Recognition Based on OSWC

In the case of the distance between the detector and the object keeps constant fundamentally, two objects, which have different sizes, move towards each other in the field of view of hexagonal detector. Figure 139.3 shows the variation of signal width of the two objects in object images, which have the sampling interval

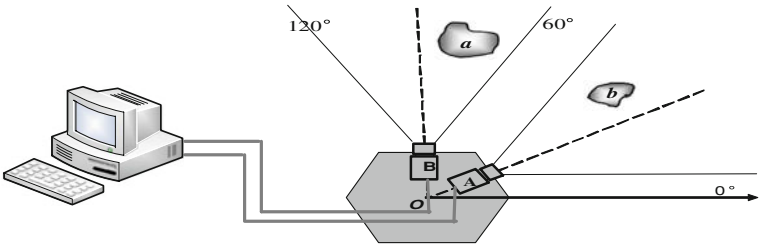


Fig. 139.2 Experiment device (*top view*)

Fig. 139.3 The variation of signal width

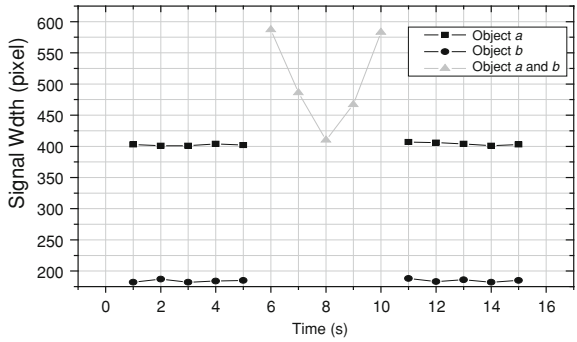
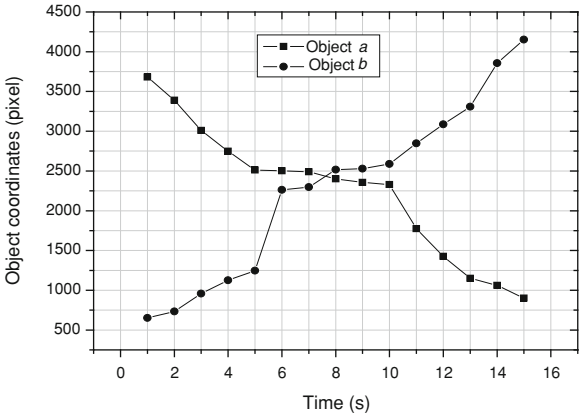


Fig. 139.4 The object coordinates of the two objects



of 1 s. It can be seen from Fig. 139.3 that in the case of $t \in [1, 5]$ or $t \in [11,15]$, there is little variation in the signal width values of the two objects. In case of $t \in [5, 10]$, the signals of the two objects are replaced by a new signal. The width of the new signal decreased in the case of $t \in [6, 8]$, and it reaches the minimum of 410 at $t = 8$. In contrast, the width of signal increased in the case of $t \in [8, 10]$, and it reaches the maximum of 583 at $t = 10$. The object coordinates of the two objects, which are determined by using Eqs. (139.5) and (139.6), are shown in Fig. 139.4.

Table 139.1 Parameters determined by the algorithm based on HSRC

Object	Time	Measurement result			Calculated result					
		X_{max}	X_{min}	X_n	$X_{max,t}/X_{max,t-1}$	$X_{n,t}/X_{n,t-1}$	$X_{max,t}/X_{min,t}$	T_1	T_2	T
a	t_1	213	350	822	–	–	0.61	–	–	–
	t_2	224	371	833	1.05	1.01	0.60	0.04	0.01	0.05
	t_3	277	440	974	1.23	1.17	0.63	0.06	0.03	0.09
	t_4	325	524	1100	1.17	1.13	0.62	0.04	0.01	0.07
	t_5	338	573	1122	1.04	1.02	0.59	0.02	0.03	0.05
	t_6	352	607	1140	1.04	1.02	0.58	0.02	0.01	0.03
b	t_1	399	223	413	–	–	1.79	–	–	–
	t_2	487	269	502	1.22	1.21	1.81	0.01	0.02	0.03
	t_3	507	285	517	1.04	1.03	1.78	0.01	0.03	0.04
	t_4	588	334	626	1.16	1.21	1.76	0.05	0.02	0.07
	t_5	717	410	745	1.22	1.19	1.75	0.03	0.01	0.04
	t_6	739	418	775	1.03	1.04	1.76	0.01	0.01	0.02

139.4.2 Object Recognition Based on HSRC

In the case of objects approach or leave away from the hexagonal detector, the criterion of HSRC, which is expressed as Eq. (139.7), is used for multi-object recognition. In our experiments, the two objects approach the detector central along radial direction. The parameters determined by the algorithm based on HSRC are listed in Table 139.1.

From Table 139.1, conclusions are therefore deduced as follows: (1) the values of T_1 of the two objects are less than 0.06, indicating that the width ratio of the maximum points ($X_{max,t}/X_{max,t-1}$) are close to the width ratio of the whole object signals ($X_{n,t}/X_{n,t-1}$); (2) the values of T2 of the two objects are less than 0.03, indicating that the width ratio of the maximum point to the minimum point in $D[x,(t-1)]$ ($X_{max,t-1}/X_{min,t-1}$) are close to that in $D(x, t)$ ($X_{max,t}/X_{min,t}$); (3) the values of T of the two objects are both less than 0.1, indicating that the proposed criterion of HSRC be useful for recognizing multi-object in the determination and tracking based on hexagonal detector.

139.5 Conclusion

Some important conclusions are therefore deduced as follows: (1) the moving images of multi-object can be captured in 360° field of view by using the proposed hexagonal opto-electronic detector, which has the characteristics of simple structure and low cost; (2) the detection and recognition of multi-object can be achieved with high speed by using the proposed criterions of OSWC and HSRC; (3) the methods demonstrate significant advantages in the detection and tracking of multi-object of photo-electric detection, computer vision, virtual reality, augment reality, etc.

Acknowledgments This paper is supported by National Natural Science Foundation of China (60801050), Scientific and Technological Innovation Project of Beijing Institute of Technology (3040012241002) and Basic Research Foundation of Beijing Institute of Technology of China (3040012211112).

References

1. Cao ZL, Oh SJ, Hall EL (1986) Dynamic omnidirectional vision for mobile robots. J Robot Sys 3(1):5–17
2. Oh SJ, Hall EL (1987) Guidance of a mobile robot using an omnidirectional vision navigation system. Proc SPIE Mobile Robot II 852:288–300
3. Jarvis RA, Byrne JC (1988) An automated guided vehicle with map building and path finding capabilities. In: 4th ISPP, pp 497–504
4. Ishiguro H, Yamamoto M, Tsuji S (1992) Omnidirectional stereo. IEEE Trans Pattern Anal Mach Intell 14(2):257–262
5. Maniere T, Benosman R, Gastaud C, Devars J (1998) Vision system dedicated to panoramic three dimensional scene reconstruction. J Electron Imag 7(3):672–676

Chapter 140
The Translation Method to Calculate the Radar Detection Probability Under Noise Jamming

Wang Hui-ting, Gao Bin and Zhang Man-zhi

Abstract According to the reasoning for the signal-jamming-ratio, the function relationship formula of radar detection probability under non-jamming and jamming conditions is found. In virtue of the JNR, An attenuation coefficient is introduced to denote the translation magnitude compared with the radar detection probability curve under non-jamming conditions. The simulation results for Marcum and Swerling model show that this translation method makes the detection probability calculation more convenient and maneuverable.

Keywords Noise jamming · Detection probability · Attenuation coefficient

140.1 Introduction

Every country's army is increasingly paying more attention to radar noise jamming [1]. It is the foundation of jamming effectiveness for the whole electronic warfare system to evaluate radar noise jamming effects, so the influence from

W. Hui-ting (✉)
College of Electronic & Information Engineering,
Chongqing University of Science and Technology,
Chongqing, China
e-mail: wht1979322@163.com

G. Bin
School of Electronics and Information Engineering,
Beijing University of Aeronautics and Astronautics,
Beijing, China
e-mail: melinshe@163.com

G. Bin · Z. Man-zhi
Unit 95478 of P.L.A, Chongqing, China
e-mail: manzhi023@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_140,

noise jamming on radar detection probability must be investigated to obtain the value of detection probability conveniently and exactly. The correlative jamming calculation formula of detection probability was acquired based on JNR. According to the attenuation coefficient, a translation method to calculate the radar detection probability was presented with noise jamming environment.

140.2 The Relationship Formula Formation of Detection Probability

To analyze the jamming effect coming from jammer to radar, we firstly consider signal power of usable echo that radar receives without jamming environment. Using basic radar principle, we can get following expression:

$$S = \frac{P_t G_t G_r \sigma \lambda^2}{(4\pi)^3 R_t^4 L_r}$$

where P_t is the transmitter power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, σ is the target radar cross section, λ is the transmitter wavelength, R_t is the slant range, L_r is the system losses. The receiver thermal noise N_0 is equal to $k_{\mathrm{T0}} B_r F$, where k is Boltzmann's constant, T_0 is the temperature, Br is the receiver bandwidth, and F is the receiver noise figure. The receiver noise figure is measured relative to a reference temperature $T_0 = 290$ K (approximately room temperature), and the factor k_{T0} becomes 4×10^{-21} W/Hz.

When the detection of the radar signal is limited by an external noise source, such as a deliberate noise jammer rather than by receiver noise, the parameters of importance in determining range performance are slightly different from the traditional radar range equation. The receiver noise power per unit bandwidth is now that determined by the jammer rather than the receiver noise figure. When a radar is performing search mission and jamming power enters from a particular direction via the sidelobes, the jamming power density on radar receiver Pjr equals to:

$$P_{jr} = \frac{P_j G_j G_{rj} \lambda^2}{(4\pi)^2 R_j^2 L_j} \cdot \frac{B_r}{B_j}$$

Where P_{j} is the jammer power, G_j is the jammer antenna gain, G_{rj} is the receiver antenna gain in the jammer direction, B_r is the receiver bandwidth, R_j is the range between radar and jammer, L_j is the jammer losses, B_j is the jammer bandwidth.

So, when a jammer disturbs a radar, the Signal-to-Jamming S/J on the radar receiver input port is as in (140.1):

$$\frac{S}{J} = \frac{S}{N_0 + P_{jr}} = \frac{P_t G_t G_r \sigma \lambda^2}{(4\pi)^3 R_t^4 L_r} \Bigg/ \left[K T_0 B_r F + \frac{P_j G_j G_{rj} \lambda^2}{(4\pi)^2 R_j^2 L_j} \cdot \frac{B_r}{B_j} \right]. \tag{140.1}$$

From above formula, we can learn that the effect of noise jamming is decided by the S/J ratio on radar antenna input port.

According to radar system theory, the radar detection probability is mainly affected by signal noise ratio (SNR) on radar receiver input port. When the SNR (or SJR) changes, probability distribution of receiver output amplitude will change too.

Equation (140.1) is divided by N_0 up and down, we get:

$$\frac{S}{N_0 + P_{jr}} = \frac{S/N_0}{1 + P_{jr}/N_0}$$

where P_{jr}/N_0 is the jamming-noise-ratio JNR. When the False-alarm probability is confirmed, the detection probability is the function of SNR. So, the radar detection probability can be rewritten as (140.2) under no jamming signal if we assume that the technique system of radar and jammer, the probability of false alarm and target characteristics are all defined:

$$P_d = f(\text{SNR}) \tag{140.2}$$

The function $f(x)$ denotes the relationship formula between the probability of detection and SNR based on Marcum model [2] or Swerling model [3, 4]. So, after considering the noise jamming, the probability of detection can be written:

$$P_d^{'} = f\left(\frac{\text{SNR}}{1 + \text{JNR}}\right)$$

where we create a new conception of attenuation coefficient which indicates the weakened degree of radar's SNR because of noise jamming. The attenuation coefficient η is defined as as follows:

$$\eta = 1/(1 + \text{JNR})$$

We can calculate the attenuation coefficient η firstly, and then utilize function $f(x)$ to carry parallel remotion manipulation with η_{dB} (η_{dB} denotes relevant dB value of η) units, so the curve of the detection probability with noise jamming can be achieve conveniently. This is our translation method to calculate the radar detection probability under noise jamming.

140.3 Simulation Applications

If we assume that a search radar is interfered with support jamming by a noise jammer using radio signal in broad bands, then the probability of detection under noise jamming can be given by above method based on Marcum model [5] and Swerling model [6] according to different tactics scenes, radar detection styles and RCS characteristics of target fluctuation.

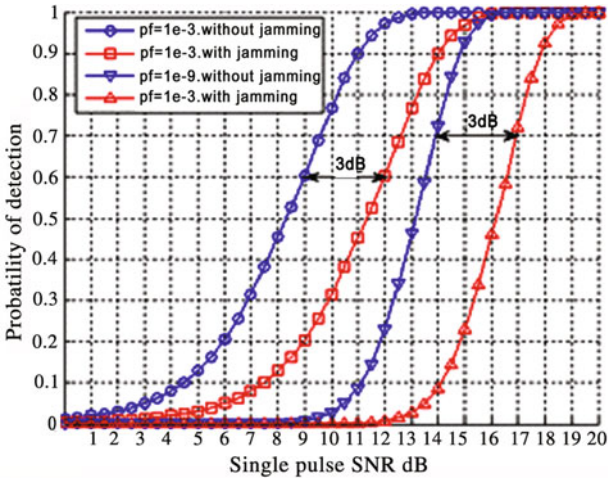


Fig. 140.1 Comparison curves of Pd for single-pulse before and after interference when JNR = 0 dB

140.3.1 The Detection Probability for Marcum Model Under Noise Jamming

Figures 140.1 and 140.2 show two curves for single-pulse detection probability. In Fig. 140.1 we can know that the two detection probability curves have 3 dB horizontal interval under the same false-alarm probability Pf when the radar JNR equals to zero dB. The reason can be explained by the attenuation coefficient in the detection probability formula after noise jamming. When JNR equals to zero dB, the attenuation coefficient equals to 3 dB, so this make the curve of detection probability move 3 dB rightwards. The same is Fig. 140.2, and the only difference is the value of the attenuation coefficient which equals to 6.2 dB when JNR equals to 5 dB.

From above analysis, we can learn the curve of detection probability with noise jamming can be attained to refer to the curve without jamming. We only use JNR to calculate the attenuation coefficient and its dB value η_{dB}, then move the curve without jamming η_{dB} dB rightwards.

Figures 140.3 and 140.4 show two curves of detection probability for coherent and noncoherent integration radar with noise jamming and without jamming environment. Where the number of integrated pulses np equals to 10, the false-alarm probability Pf $= 10^{-9}$, and JNR equals to 0 and 5 dB respectively. From Figs. 140.3 and 140.4 we can see the remotion amplitudes of two group curves equal to relevant η_{dB} values of their JNR. This validates above conclusion.

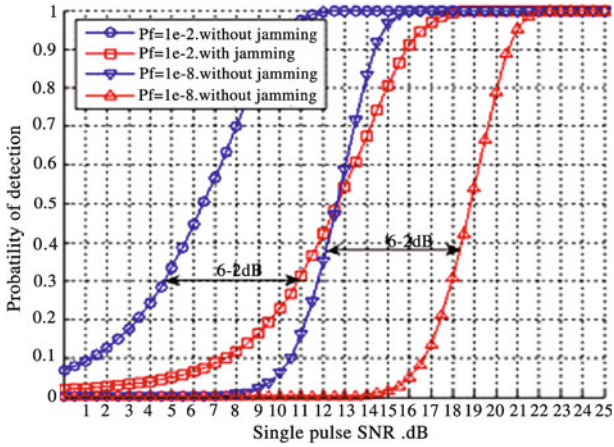


Fig. 140.2 Comparison curves of Pd for single-pulse before and after interference when JNR = 5 dB

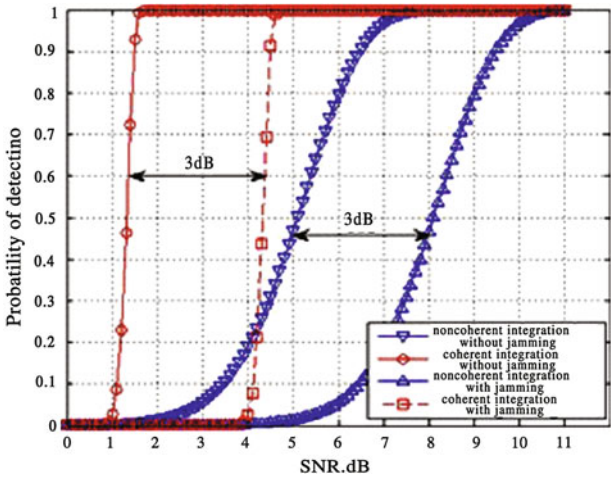


Fig. 140.3 Comparison curves of Pd for coherent and noncoherent integration when JNR = 0 dB and np = 10

140.3.2 The Detection Probability for Swerling Model Under Noise Jamming

Figures 140.5 and 140.6 show a group curves of detection probability for Swerling to four distinct cases of target RCS fluctuation with noise jamming and without jamming environment. Where the number of integrated pulses np

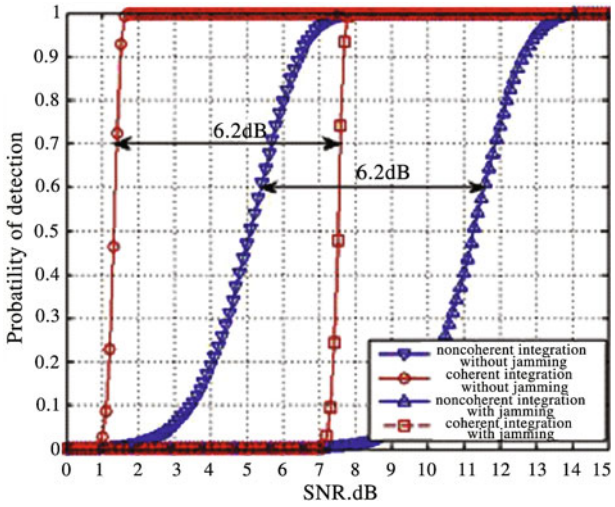


Fig. 140.4 Comparison curves of Pd for coherent and noncoherent integration when JNR = 5 dB and np = 10

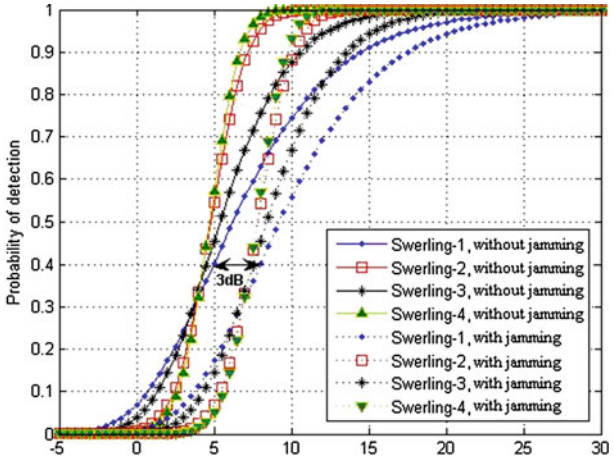


Fig. 140.5 Comparison curves of Pd for Swerling targets before and after interference when JNR = 0 dB

equals to 10, the false-alarm probability Pf = 10–9, and JNR equals to 0 dB and 5 dB respectively. From Figs. 140.5 and 140.6 we can see the remotion amplitudes of two group curves equal to relevant η_{dB} values of their JNR. This validates above conclusion.

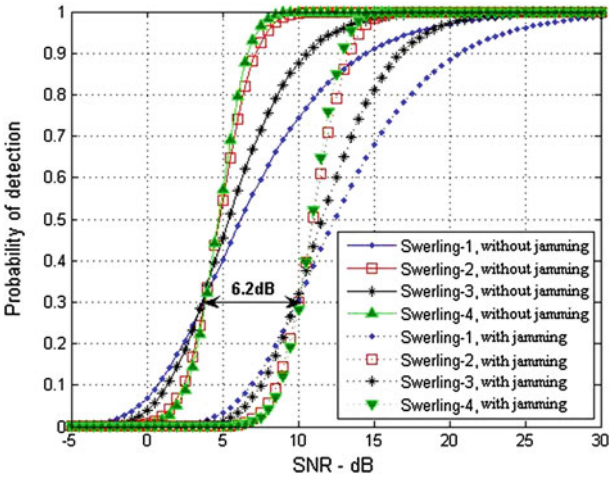


Fig. 140.6 Comparison curves of Pd for Swerling targets before and after interference when JNR = 5 dB

140.4 Conclusions

Noise jamming can depress the probability of detection greatly and the quantificational method to measure the accurate change in radar's detection probability is very important. Because the detection probability can be calculated, for a given threshold, if the statistical properties of the background interference are defined. The simplest case is for random (Gaussian) noise, when the envelope of signal plus noise follows the Rician distribution [7]. In this case the corresponding equations for coherent and noncoherent detection are mature. The North approximation is often used to calculate detection probability, as it is both simple and accurate. According to above convenient algorithms, using the attenuation coefficient and correlative movement manipulation on the curve without noise jamming, the translation method to calculate the radar detection probability under noise jamming is presented reasonedly.

References

1. Schleher DS (1999) Electronic warfare in the information age. Artech House Inc
2. Marcum JI (1960) A statistical theory of target detection by pulsed radar. In: IRE Transactions on Information Theory, pp 59–267
3. Swerling P (1960) Probability of detection for fluctuating targets. IRE transactions on information theory, pp 269–308
4. Swerling P (1965) More on detection of fluctuating targets. IRE transactions on information theory, pp 459–460

5. Mahdi QS (2010) 3D simultaneous multi-beams radar processing by using planner array antennas. In: 2010 Loughborough Antennas & Propagation conference, 8–9 Nov 2010, Loughborough, pp 625–628
6. Kostic A, Rancic D (2003) Radar coverage analysis in virtual GIS environment. In: Telsiks 2003-4th international conference on telecommunications in modern satellite, cable and broadcasting services, Ni, pp 721–724, 1–3 Oct
7. Awadallah RIS, Gehman JZ, Kuttler JR et al (2004) Newkirk, modeling radar propagation in three-dimensional environments. Johns Hopkins Apl Technical Digest 25(2):101–111

Chapter 141
Application of Compact-Type Wavelet Neural Network in Fault Diagnosis of Frequency Converter

Xin Wang and Yunyun Zhang

Abstract The modified wavelet packet algorithm and the principle and algorithm flow of the compact-type wavelet neural network (CWNN) were put forward. For common faults of six pulse AC–AC frequency converter, the fault feature vector is extracted using the modified wavelet packet. The fault feature vector is used as the input feature vector of the CWNN. Simulation results indicate that the CWNN can accurately classify faults. The convergence rate of the CWNN obviously surpasses the BP neural network, and the fault diagnosis of six pulse AC–AC frequency converter is efficiently accomplished.

Keywords Compact-type wavelet neural network · Fault diagnosis · Wavelet packet · AC–AC frequency

141.1 Introduction

The compact-type wavelet neural network (CWNN) derives from the wavelet decompose, and it is the extension of a kind artificial neural network based on the wavelet analysis. The wavelet analysis and the neural network fuse directly, and the wavelet function or the scale function acts as the excitation function. It has been proved that the CWNN inherits the advantages of the wavelet transform and the neural network. The CWNN can adaptively adjust network parameters through training. Therefore, the CWNN has the strong approximation ability and the fault

X. Wang (✉) · Y. Zhang
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: wangxd@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_141,

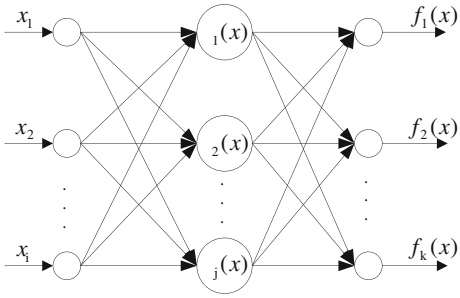


Fig. 141.1 The structure of the CWNN

tolerance ability. Exploring the efficient learning algorithm of the CWNN and its application in fault diagnosis is necessary [1–3]. In this paper, a modified CWNN to diagnose the thyristor open circuit fault of six pulse frequency converter was put forward. The simulation results prove its efficiency.

141.2 The Modified Wavelet Packet Algorithm

In the multi-resolution analysis, $L^2(R) = \bigoplus_{j\in Z} W_j$ indicates that the Hilbert space $L^2(R)$ is decomposed into the orthogonal sum of all subspace $W_j(j \in Z)$ by the multi-resolution analysis according to different scale factor j, where W_j is the wavelet subspace of the wavelet function $\psi(t)$. The wavelet subspace W_j is further subdivided according to the binary fraction to improve the frequency resolution [4].

The decomposition formula of the signal from the scale $j+1$ to j can be defined as

$$\begin{cases} d_l^{j,2n} = a_{j,2n} \sum_k \left(h_{2k+1-2l} d_{2k}^{j+1,n} + h_{2k-2l} d_{2k+1}^{j+1,n} \right) \\ d_l^{j,2n+1} = b_{j,2n+1} \sum_k \left(g_{2k+1-2l} d_{2k}^{j+1,n} + g_{2k-2l} d_{2k+1}^{j+1,n} \right) \end{cases} \tag{141.1}$$

where $a_{j,2n}$ and $b_{j,2n+1}$ are selection factors. $a_{j,2n} = 1$ when the subspace U_j^{2n} is focused, but in other cases $a_{j,2n} = 0$. $b_{j,2n+1} = 1$ when the subspace U_j^{2n+1} is focused, but in other cases $b_{j,2n+1} = 0$.

If $a_{j,2n} = 1$ and $b_{j,2n+1} = 1$ when $0 < j < M-1$, this algorithm is called the coefficient dichotomy wavelet packet algorithm.

141.3 The Principle and Algorithm of the CWNN

The basic idea of the CWNN is that the wavelet function replaces the hidden layer function of the conventional neural network. At the same time, the corresponding weight from the input layer to the hidden layer and the hidden threshold respectively

Table 141.1 Feature vectors of training samples

No.	$\lvert\gamma_3^*\rvert$	E_{21}^*	E_{30}^*	E_{31}^*	E_{40}^*	E_{41}^*	E_{42}^*	E_{43}^*	State
1	0.0332	0.0022	0.7109	0.0080	0.7032	0.0065	0.0008	0.0071	F_1
2	0.1510	0.0027	0.7147	0.0136	0.6991	0.0149	0.0013	0.0123	F_2
3	0.2171	0.0026	0.7145	0.0131	0.6993	0.0146	0.0012	0.0119	F_3
4	0.1211	0.0019	0.7115	0.0053	0.7026	0.0078	0.0007	0.0046	F_4
5	0.0107	0.0018	0.7106	0.0064	0.7036	0.0059	0.0007	0.0056	F_5
6	0.1610	0.0022	0.7115	0.0083	0.7026	0.0078	0.0016	0.0067	F_6
7	0.1024	0.0024	0.7118	0.0084	0.7022	0.0084	0.0017	0.0066	F_7
8	0.1754	0.0024	0.7132	0.0114	0.7007	0.0114	0.0009	0.0105	F_8
9	0.1206	0.0024	0.7138	0.0117	0.7001	0.0122	0.0009	0.0108	F_9
10	0.0836	0.0027	0.7131	0.0102	0.7009	0.0107	0.0016	0.0085	F_{10}
11	0.0257	0.0026	0.7133	0.0104	0.7006	0.0111	0.0018	0.0085	F_{11}
12	0.0617	0.0027	0.7126	0.0084	0.7014	0.0100	0.0022	0.0062	F_{12}
13	0.1212	0.0028	0.7123	0.0087	0.7018	0.0096	0.0022	0.0064	F_{13}

instead by the scale function and the translation function of the wavelet function [5]. The basic structure is shown in Fig. 141.1.

In Fig. 141.1, x_i is the number i input sample of input layer, $\varphi_j(x)$ is the wavelet function, $f_k(x)$ is the number k output of the output layer, w_{ji} is the weights between the number i neuron of the input layer and the number j neuron of the hidden layer, w_{kj} is the weight between the number j neuron of the hidden layer and the number k neuron of the output layer. Neurons excitation function of the hidden layer is

$$\varphi_j(x) = \varphi\left[\left(x - b_j\right)/a_j\right] \tag{141.2}$$

Where ϕ is the wavelet function, a_j and b_j are respectively the scale function and the translation function [6–8]. $f_k(x)$ is defined as:

$$f_k(x) = \sum_{j=1}^{M} w_{kj}\varphi_j(x) = \sum_{j=1}^{L} w_{kj}\varphi\left[\left(\sum_{i=1}^{L} w_{ji}x_i - b_j\right)\Big/ a_j\right] \tag{141.3}$$

In this paper, the CWNN adopts the steepest descent method which has an adaptive learning rate and additional momentum factors to revise the weight and the threshold. The transfer function of the hidden layer adopts the wavelet Mexican Hat. The transfer function of the output layer adopts a linear function.

141.4 Diagnosis of Simulation Signal Using Wavelet Neural Network

In all types of faults of the frequency converter, it is common that only a thyristor open-circuit fault in three phases, that is, U, V and W. This paper builds a six pulse

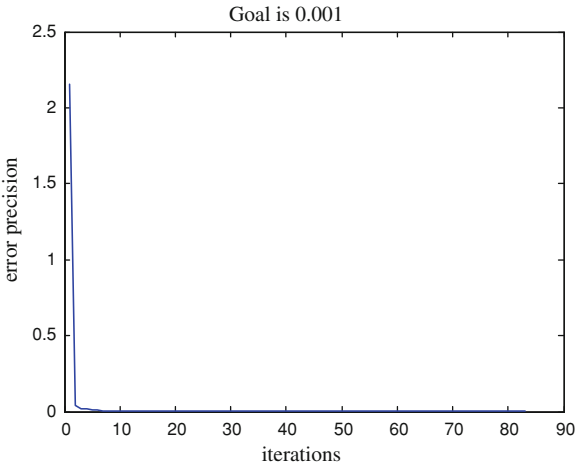


Fig. 141.2 Training error curve of the CWNN

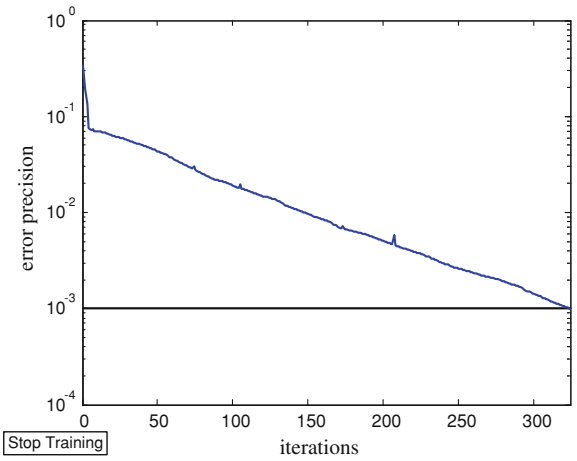


Fig. 141.3 Training error curve of the BP neural network

frequency control of motor speed system in Simulink to carry out the fault simulation experiment. The system adopts three phase asynchronous motor, and adopts one-third frequency and one-third volt control mode. The U phase as an example, one thyristor is switched off in turn to simulate the thyristor open-circuit fault, and then the output current is collected to do fault diagnosis.

This paper adopts an improved wavelet packet algorithm to extract the signal feature, and then constructs a feature vector to input the CWNN. The specific fault diagnosis procedure is as follows:

(1) The wavelet packet decomposition in four levels is done by using the wavelet db4 for the sampled current signal of every fault state, respectively signal features of all frequency bands are extracted in every level.

Table 141.2 Diagnosis results of test samples

No.	$\lvert\gamma_3^*\rvert$	E_{21}^*	E_{30}^*	E_{31}^*	E_{40}^*	E_{41}^*	E_{42}^*	E_{43}^*	State
1	0.0334	0.0023	0.7106	0.0076	0.7029	0.0063	0.0008	0.0072	F_1
2	0.1508	0.0027	0.7148	0.0134	0.6989	0.0151	0.0012	0.0122	F_2
3	0.2170	0.0025	0.7146	0.0130	0.6995	0.0142	0.0014	0.0117	F_3
4	0.1213	0.0020	0.7118	0.0051	0.7027	0.0076	0.0006	0.0043	F_4
5	0.0105	0.0019	0.7107	0.0065	0.7035	0.0056	0.0009	0.0580	F_5
6	0.1609	0.0022	0.7118	0.0081	0.7026	0.0076	0.0015	0.0066	F_6
7	0.1024	0.0025	0.7116	0.0086	0.7023	0.0081	0.0019	0.0063	F_7
8	0.1755	0.0021	0.7133	0.0115	0.7009	0.0112	0.0007	0.0102	F_8
9	0.1208	0.0025	0.7139	0.0116	0.7002	0.0121	0.0008	0.0109	F_9
10	0.0835	0.0029	0.7130	0.0103	0.7011	0.0105	0.0018	0.0083	F_{10}
11	0.2590	0.0027	0.7131	0.0105	0.7005	0.0113	0.0016	0.0084	F_{11}
12	0.6190	0.0027	0.7129	0.0084	0.7017	0.0100	0.0020	0.0062	F_{12}
13	0.1211	0.0026	0.7120	0.0086	0.7016	0.0095	0.0023	0.0061	F_{13}

(2) The energy of each frequency band is calculated, and the frequency bands whose energy is obvious change are found out. They are E_{21}, E_{30}, E_{31}, E_{40}, E_{41}, E_{42} and E_{43}.

(3) The feature vector T is constructed by using the absolute value of output current skewness of the U phase and seven band energy in (141.2) to construct the feature vector T.

$$T = [|\gamma_3| \; E_{21} \; E_{30} \; E_{31} \; E_{40} \; E_{41} \; E_{42} \; E_{43}] \tag{141.4}$$

T^* is the normalization of the vector T.

$$T^* = \left[\left|\gamma_3^*\right| \; E_{21}^* E_{30}^* \; E_{31}^* E_{40}^* E_{41}^* \; E_{42}^* \; E_{43}^*\right] \tag{141.5}$$

(4) T^* is taken as the input vector input of the CWNN to do fault diagnosis.

In this paper, the CWNN has three layers. The number of input nodes is 8. The number of output nodes is 13. The number of hidden nodes is 17. The error goal is set as 0.001.

Neural network output mode is as follows:

F_1: normal (1000000000000).
F_2: U_{VT1} open circuit fault (0100000000000).
F_3: U_{VT2} open circuit fault (0010000000000).
F_4: U_{VT3} open circuit fault (0001000000000).
F_5: U_{VT4} open circuit fault (0000100000000).
F_6: U_{VT5} open circuit fault (0000010000000).
F_7: U_{VT6} open circuit fault (0000001000000).
F_8: U_{VT7} open circuit fault (0000000100000).

Table 141.3 Diagnosis results of the neural network output

F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}
0.9901	0.0155	0.0148	0.0154	0.0390	0.0414	0.0163	0.0447	0.0238	0.0453	0.0610	0.0295	0.0228
0.0493	0.9812	0.0114	0.0099	0.0312	0.0538	0.0047	0.0139	0.0097	0.0022	0.0132	0.0183	0.0207
0.0225	0.0130	0.9832	0.0048	0.0267	0.0605	0.0015	0.0038	0.0043	0.0001	0.0032	0.0054	0.0107
0.0206	0.0095	0.0206	0.9288	0.0221	0.0391	0.0007	0.0011	0.0018	0	0.0007	0.0026	0.0046
0.0150	0.0073	0.0099	0.0040	0.9832	0.0227	0.0003	0.0003	0.0010	0	0.0001	0.0007	0.0033
0.0075	0.0050	0.0052	0.0018	0.0148	0.9882	0.0001	0.0001	0.0004	0	0	0.0002	0.0015
0.0020	0.0039	0.0041	0.0010	0.0123	0.0200	0.9869	0	0.0002	0	0	0.0001	0.0007
0.0018	0.0030	0.0053	0.0006	0.0117	0.0170	0	0.9880	0.0001	0	0	0.0001	0.0007
0.0010	0.0021	0.0050	0.0003	0.0097	0.0117	0	0	0.9879	0	0	0	0.0003
0.0003	0.0015	0.0037	0.0002	0.0058	0.0136	0	0	0	0.9867	0	0	0.0001
0.0003	0.0010	0.0032	0.0001	0.0038	0.0130	0	0	0	0	0.9884	0	0.0001
0.0003	0.0007	0.0016	0.0001	0.0036	0.0038	0	0	0	0	0	0.9940	0.0001
0.0002	0.0006	0.0012	0	0.0025	0.0036	0	0	0	0	0	0	0.9894

F_9: U_{VT8} open circuit fault (000 0000010000).
F_{10}: U_{VT9} open circuit fault (0000000001000).
F_{11}: U_{VT10} open circuit fault (0000000000100).
F_{12}: U_{VT11} open circuit fault (0000000000010).
F_{13}: U_{VT12} open circuit fault (0000000000001).

where U_{VT1}–U_{VT12} denote 12 thyristors. One set of feature vectors of training samples for each state is shown in Table 141.1.

Figure 141.2 shows the training error curve of the CWNN. In the training process of the CWNN, the selections of momentum factor and learning rate are very important. The value of momentum factor is always around 0.9. In this paper it is 0.925. This paper uses several different learning rates to train the CWNN, and then selects the best one. We can get that the CWNN reaches the target after 80 iterations under suitable initial parameters. Figure 141.3 shows the training error curve of the BP neural network on the same network parameters. It can be seen that the BP neural network reaches the target after 328 iterations, and the diagnostic accuracy of the BP neural network is lower than that of the CWNN.

After the network training finishes, the fault diagnosis model which meets the requirements will exist in the network in the form of network structure, weights, and threshold. It stores knowledge in each neuron distribution. It has high learning efficiency and strong fault tolerance. It is rich in association and has good reasoning ability. Even if fault features of the input parts aren't obvious and the information is wrong, it can also make comprehensive judgments and give more ideal results. When fault feature vectors are inputted to the CWNN, it can accurately diagnose the fault of corresponding thyristor.

Test samples are tested by using the trained CWNN, and part of test samples are shown in Table 141.2. Diagnosis results of the neural network output are shown in Table 141.3. We can see that diagnosis error is very small. We can judge that the training is successful and the diagnosis accuracy is high. Through a large amount of data are tested, the fault diagnosis accuracy of the CWNN in this paper reaches 96%.

141.5 Conclusion

Aiming at common faults of six pulses AC–AC frequency converter, current signals are decomposed with the improved wavelet packet and the frequency spectrum characteristics of current signals are extracted to construct the state feature vector. By using the trained CWNN, the test data can diagnosis the fault. It can be seen that the CWNN in this paper is very good at fault classify, and its convergence rate and diagnosis accuracy are superior to the BP neural network under the same conditions. Therefore, the CWNN is more effectively for fault diagnosis of AC–AC frequency converter.

Acknowledgments This work was supported in part by Research Foundation for Innovative Persons of Henan Province (No.2008HASTIT022).

References

1. Chen W, Pan C, Yun Y et al (2008) Fault diagnosis method of power transformer based on improved wavelet neural network algorithm. Chin J Sci Instr 29:1489–1493
2. Zheng H, Chen X, Li Z et al (2002) Implementation and application of a neural network fault diagnosis system based on wavelet transform. Trans Chin Soc Agric Mach 33:73–77
3. Cai G, Jia L, Yang J et al (2010) Improved wavelet neural network based on hybrid genetic algorithm applicationin on fault diagnosis of railway rolling bearing. Int J Digit Content Technol Appl 4:135–141
4. Xin W, Jian H (2007) Study in the extraction of fault signal symptom based in wavelet packets. Chin J Electr Devices 30:999–1002
5. Saravanan N, Ramachandran KI (2010) Incipient gear box fault diagnosis using discrete wavelet transform (DWT) for feature extraction and classification using artificial neural network (ANN). Expert Sys Appl 37:4168–4181
6. Peng T, Ma Q (2010) Application of wavelet neural network in rolling bearing fault diagnosis. Comput Eng Appl 46:213–215
7. Xu Q, Meng X, Han X, et al (2008) Gas turbine fault diagnosis based on wavelet neural network. In: Proceedings of the 2007 international conference on wavelet analysis and pattern recognition, pp 738–741
8. Cui B (2010) Open-circuit fault diagnosis of transistor in inverter based on wavelet neural network. In: Proceedings of the 29th Chinese control conference, pp 4019–4023

Chapter 142
Study of Fluorescence Sensor Temperature Compensation Based on BP Neural Network

Qiongfang Yu, Qiongxia Yu and Aihua Dong

Abstract Fluorescent sensor can convert the information of molecular recognition into fluorescent signal perceived by the outside world. Because fluorescent sensor has the features of high selectivity, high sensitivity, low production cost, not destroying the samples, and easy to realize the online testing, it has wide application prospects. Because the temperature has some influence on the fluorescence characteristics, it is very important to prevent the temperature effect on the measurement accuracy of the fluorescent sensor by using BP network for temperature compensation. In this paper, first the characteristics of fluorescence sensor and some knowledge of the fluorescence are described; next the BP model and BP neural network structure are introduced; then the LMBP algorithm is derived; finally, the principle of BP neural network compensation for sensor characteristics is illustrated, and the LMBP network structure is designed.

Keywords Fluorescent sensor · BP neural network · LMBP algorithm

Q. Yu (✉) · Q. Yu · A. Dong
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, China
e-mail: yuqf@hpu.edu.cn

Q. Yu
e-mail: abprwm@126.com

A. Dong
e-mail: dah@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_142,

142.1 Introduction

Fluorescent sensor identifies the type and concentration of the measured substances by measuring the characteristic fluorescence spectroscopy of the measured substances. This kind of sensor is most attractive because its measuring accuracy and sensitivity is higher, and its development is rapid. It also has the function of special signal processing. Fluorescent sensors have broad application prospect in biological analysis, clinical diagnosis, environmental monitoring, etc. Therefore, now fluorescent sensors have become a hot spot in the sensor research field.

Back propagation (BP) network is proposed by the team of scientists that is led by Rumelhart and McCelland in 1986, which is a multilayer feedforward network whose training is according to the error back propagation algorithm. It is one of the most widely used neural network model. BP network can learn and store large numbers of mapping relations of input–output model, and it doesn't have to reveal the mathematical equations of this mapping relationship in advance. Its learning rule is using the steepest descent method. Through constantly adjusting the weights and thresholds of the network, error quadratic sum of the network is minimized. The topological structure of BP neural network model is composed of input layer, hidden layer and output layer.

Temperature has influence on the characteristics of fluorescence quantum yield, fluorescence intensity, and so on, which will affect the measurement precision. Therefore, it is very necessary to use BP network to compensate the temperature of fluorescent sensors to improve its measurement precision.

142.2 Fluorescent Sensor

Some chemicals absorb energy from the outside and store it to enter into the excitation state. When they return to the ground state from the excitation state, the excess energy radiates in the form of electromagnetic radiation, i.e., giving out light, which is called fluorescence. Molecules or atoms which can produce fluorescence give out light when they receive energy, but once the energy supply ceases, fluorescence phenomenon disappears [1].

Temperature has some effect on the molecular diffusion, activation, intramolecular energy conversion. Generally, when temperature decreases, the fluorescent quantum yield and fluorescence intensity will both increase. As shown in Fig. 142.1, when the temperature rises, fluorescence quantum yields of sulfuric acid solution of the uranyl sulfate and glycerite of rhodamine B both decrease; however, the capacities of spectrum absorption and light absorption have little change. If the molecular structure changes with temperature, influence on the spectrum absorption is much greater. For example, when the temperature of isobutanol solution of rhodamine B increases, not only the quantum yield decreases, but also the spectrum absorption has a significant change.

The main reason that the fluorescence intensity of the solution increases with the increasing temperature is the effect of intramolecular energy conversion.

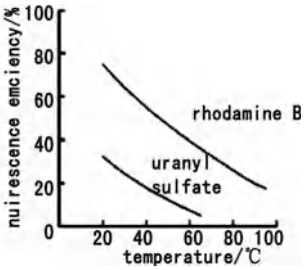


Fig. 142.1 Relation of fluorescence quantum yield and temperature

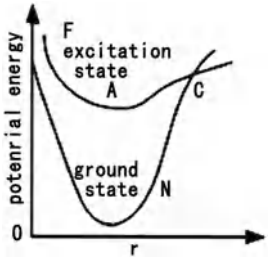


Fig. 142.2 Potential energy curve of internal energy conversion

The potential energy curves of polyatomic molecules in the ground state and excitation state may intersect or tangent at one point, as shown in Fig. 142.2.

When the excited molecules receive additional heat and move along potential energy curve to the intersection C, the excitation energy converts to the vibration energy of the ground state when converting to the potential energy curve NC, and followed by loss of vibrational energy through oscillation relaxation. The relation curve of fluorescence intensity and temperature of the solution can be expressed by the following equation, that is:

$$(F_0 - F)/F = ke^{-E/RT} \tag{142.1}$$

where: E is the additional heat that must be obtained when the excited molecules transferring to the ground state curve,that is the excitation heat energy.

The size of E is amount to the height of A–C in Fig. 142.2. E is usually 16.8–30.4 kJ obtained by tests, and it is about two times larger than the vibration energy obtained by the infrared spectrum of the molecule.

$$Y_F = k_f/(k_f + \sum k) \tag{142.2}$$

where: Y_F is the fluorescence quantum yield; k_f is the fluorescence emission rate constant; Σk is the rate constant sum of non-radiation deactivation processes of various single-molecules.

When there is no quencher in the solution, it can be seen from the formula (142.2) that the size of fluorescence quantum yield has relation with the relative rate of the radiation and non-radiation processes. The rate of radiation process does not change with temperature, therefore, changes in fluorescence quantum yield reflects the rate changes of non-radiation processes. With the decrease of the

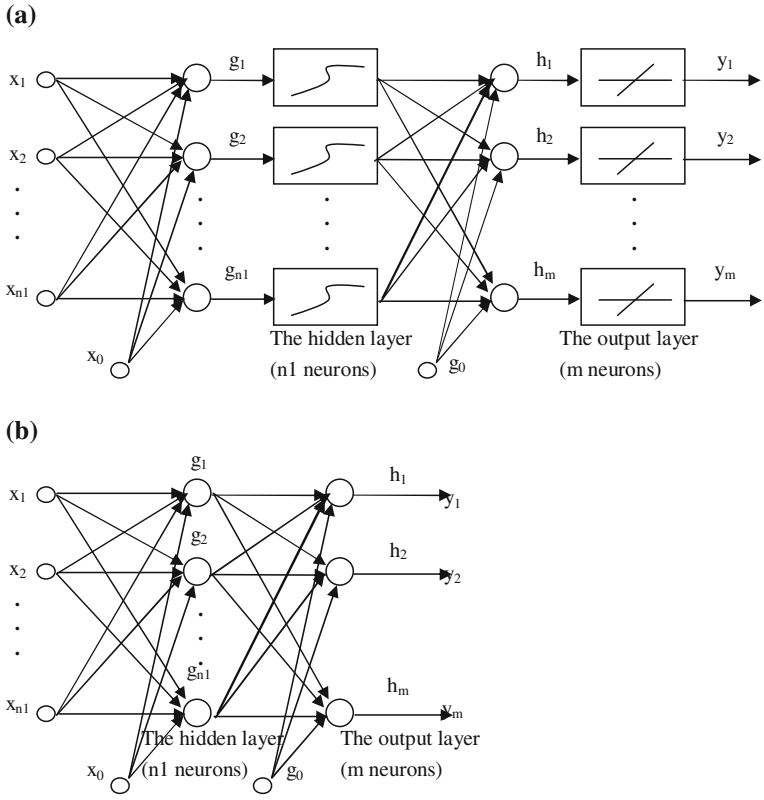


Fig. 142.3 The BP network with a hidden layer and an output layer

solution temperature, the medium viscosity increases, so that the opportunities that the fluorescent substance molecules collide with the solvent molecules will reduce.

When there is quencher in the solution, if the fluorescence quenching effect is caused by the collisions between fluorescent substance molecules and quencher molecules, the fluorescence intensity will increase with the increasing temperature.

142.3 BP Network

BP network usually has one or several hidden layers. Neurons in hidden layer use the S-type function, and neurons in output layer use the linear transfer function. If the last layer of BP network uses the S-type activation function, the output of the entire network will be limited to a smaller range; if the last layer of BP network uses the linear activation function, the output of the entire network can take any value. Figure 142.3 describes a BP network with one hidden layer, and the network in Fig. 142.3a is usually simplified to the network in Fig. 142.3b.

142.4 LMBP Algorithm (Levenberg–Marquardt Algorithm)

In mathematics and computer, Levenberg–Marquardt algorithm (LMA) provides a numerical solution in the parameter space of a function for the function minimization problems, and the solution is generally nonlinear. Most of these minimization problems appear in the least squares curve fitting and nonlinear optimization [2, 3].

Newton method of the optimal performance function $F(W)$ is [4]:

$$w_{t+1} = w_t - A_t^{-1} g_t \tag{142.3}$$

where: w_{t+1} is the amendment weight in the $t+1$ time; w_t is the weight in t time; $A_t = \nabla^2 F(W)|_{W=W_t}$, $g_t = \nabla F(W)|_{W=W_t}$; $F(W)$ is the quadratic sum of the error function $e(w)$, that is:

$$F(W) = \sum_{i=1}^{n} e_i^2(w) = e^T(w)e(w) \tag{142.4}$$

.

Then the jth gradient component of $F(W)$ is:

$$[\nabla F(W)]_j = \frac{\partial F(W)}{\partial W_j} = 2\sum_{i=1}^{n} e_i(w)\frac{\partial e_i(w)}{\partial w_j} \tag{142.5}$$

Then the gradient can be written in the form of matrix:

$$\nabla F(W) = 2J^T(W)e(w) \tag{142.6}$$

where, $J(W) = \begin{pmatrix} \frac{\partial e_1(w)}{\partial w_1} & \frac{\partial e_1(w)}{\partial w_2} & \cdots & \frac{\partial e_1(w)}{\partial w_n} \\ \frac{\partial e_2(w)}{\partial w_1} & \frac{\partial e_2(w)}{\partial w_2} & \cdots & \frac{\partial e_2(w)}{\partial w_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial e_N(w)}{\partial w_1} & \frac{\partial e_N(w)}{\partial w_2} & \cdots & \frac{\partial e_N(w)}{\partial w_n} \end{pmatrix}$ is the Jacobian matrix.

Here the Hessen matrix needs to be calculated. Define the Hessen matrix as:

$\nabla^2 F(W) = \begin{pmatrix} \frac{\partial^2 F(W)}{\partial w_1^2} & \frac{\partial^2 F(W)}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 F(W)}{\partial w_1 \partial w_n} \\ \frac{\partial^2 F(W)}{\partial w_2 \partial w_1} & \frac{\partial^2 F(W)}{\partial w_2^2} & \cdots & \frac{\partial^2 F(W)}{\partial w_2 \partial w_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 F(W)}{\partial w_n \partial w_1} & \frac{\partial^2 F(W)}{\partial w_n \partial w_2} & \cdots & \frac{\partial^2 F(W)}{\partial w_n^2} \end{pmatrix}$ then the tjth element of Hessen matrix is:

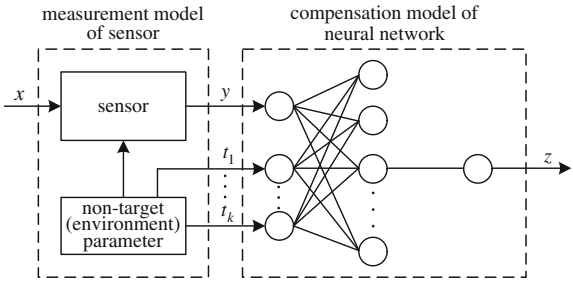


Fig. 142.4 Schematic of sensor's characteristic compensation

$$\left[\nabla^2 F(W)\right]_{t,j} = \frac{\partial^2 F(W)}{\partial w_t \partial w_j} = 2\sum_{i=1}^{n}\left\{\frac{\partial e_i(w)}{\partial w_t}\frac{\partial e_i(w)}{\partial w_j} + e_i(w)\frac{\partial^2 e_i(w)}{\partial w_t \partial w_j}\right\} \tag{142.7}$$

So Hessen matrix can use Jacobian matrix to be expressed as:

$$\nabla^2 F(W) = 2J^T(W)J(W) + 2S(W) \tag{142.8}$$

where: $S(W) = \sum_{i=1}^{n} e_i(w)\nabla^2 e_i(w)$, if $S(W)$ is very small, formula (142.9) can be approximately expressed as:

$$\nabla^2 F(W) \cong 2J^T(W)J(W) \tag{142.9}$$

Substitute (142.10) and (142.7) in (142.4), then the Gauss–Newton algorithm formula is obtained:

$$\begin{aligned} w_{t+1} &= w_t - \left[2J^T(W_t)J(W_t)\right]^{-1} 2J^T(W_t)e(W_t) \\ &= w_t - \left[J^T(W_t)J(W_t)\right]^{-1} J^T(W_t)e(W_t) \end{aligned} \tag{142.10}$$

But the problem of (142.10) is the matrix $H = J^T J$ may irreversible, and a reversible matrix $G = H + \mu_t I$ is used to improve, where μ is the adjustment factor, and I is the unit matrix. Then the Levenberg–Marquardt algorithm formula is obtained:

$$w_{t+1} = w_t - \left[J^T(W_t)J(W_t) + \mu_t I\right]^{-1} J^T(W_t)e(W_t) \tag{142.11}$$

142.5 Temperature Compensation of the Fluorescence Sensor

142.5.1 Principle of BP Neural Network for Sensors' Characteristic Compensation

The schematic of using BP neural network to improve the sensor's output characteristic is composed of the sensor model and neural network model two parts [5, 6], as shown in Fig. 142.4.

In the schematic, the sensor model is:

$$y = f(x, t_1, t_2, \ldots, t_k) \tag{142.12}$$

where: x is the objective parameter to be measured; $t_1, t_2, \ldots, t_k$ are k objective parameters; z is the output through the compensation of BP network.

The function relation model expressed in formula (142.12) is quite complex. It is hard to describe in specific function, but the complex nonlinear relationship described by the BP neural network approximation can be used. Taking the sensor output measured by the objective parameters and the outputs of various non-target parameters sensitive elements as the input of BP network, the output z processed by BP network is the objective parameter to be measured x that eliminates the effects of various non-target parameters.

142.5.2 Design of LMBP Network Structure

This design uses the LMBP algorithm, and uses a three-layer BP network with an input layer, a hidden layer, and a output layer. Neurons in the hidden layer use the S-tangent transfer function, and the neurons in the output layer use the linear transfer function.

References

1. Wang Y (2003) Optoelectronics and optical fiber sensor technology. National Defence Industry Press, Beijing
2. Yang L (2009) Study of Vibrating wire sensor temperature compensation. Changsha University of Science and Technology, Changsha
3. http://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt_algorithm
4. Yingjie Lv, Changhua Hu (2006) Fault prediction method and its application based on LMBP neural network. Mech Sci Technol 25(1):28–30
5. Yang D (2006) The research of sensor temperature compensation technique based on BP neural network. Northeastern University, Shenyang
6. S Luo (2000) Theoretical basis of large-scale artificial neural networks. China Railway Press, Beijing, pp 18–32

Chapter 143
Temperature Sensitivity Analysis of LPFG by New Transfer Matrix Method

Guodong Wang, Rui Wang, Yunjian Wang and Suling Wang

Abstract The temperature sensitivity of long period fiber grating is analyzed by new transfer matrix method. The new transfer matrix method can analyze the modes coupled between the core mode and multiple cladding modes. Compared to the usual method used, such as solving the coupled mode equation by the fourth order adaptive step size Runge–Kutta algorithm, the new transfer matrix method has a faster calculation speed. The theoretical results are exactly in agreement with the method of solving the coupled mode equation.

Keywords Temperature sensitivity · Long period fiber grating · New transfer matrix

143.1 Introduction

Long period fiber gratings (LPFG) have found many applications in optical telecommunications such as mode converters [1], rejection filters [2], gain-flattening filters for erbium-doped fiber amplifiers [3], and optical fiber sensors for strain [4],

G. D. Wang (✉) · Y. J. Wang · S. L. Wang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China
e-mail: wgd@hpu.edu.cn

Y. J. Wang
e-mail: yunjian_wang@163.com

S. L. Wang
e-mail: tjslwang@163.com

R. Wang
School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China
e-mail: quantumwgd@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_143,

temperature [5], and refractive index measurements [6] because of their capability of coupling power between core and cladding modes at the resonant wavelengths [7–10].

LPFG represent important alternatives to the use of FBGs in many sensor application, and show many of the same advantages in terms of an intensity-independent output to encode the measurement. In addition, the wavelength shifts of some of the key spectral features are often greater than FBGs, thereby offering potentially a higher sensitivity. The temperature sensitivity of LPFG can be obtained by analyzing the variation of spectral under a certain temperature changes. The common methods used to analyze the spectral characteristics of LPFG and fiber Bragg gratings include transfer matrix method (TMM) and solving the coupled mode equations (SCME) [11–13]. The traditional TMM is able to analyze the uniform and non-uniform LPFG and fiber Bragg gratings when only two modes being considered [14, 15]. The SCME method could obtain the spectrum of the uniform and non-uniform structure LPFG when multiple modes are considered. In this paper, a new transfer matrix method about LPFG with multiple cladding modes coupled is proposed and applied to analyse the temperature sensitivity of LPFG. This new TMM can analyze the modes coupled to both uniform and the non-uniform between the core mode and the multiple cladding modes. And the new TMM is simple to implement, almost always sufficiently accurate, and generally faster than that of SCME.

143.2 New Transfer Matrix

The coupled-mode equations of LPFG are given as [14]

$$\begin{cases} \dfrac{\mathrm{dA^{co}}}{\mathrm{d}z} = jk_{01-01}^{\mathrm{co-co}}A^{\mathrm{co}} + j\sum\limits_{v}\dfrac{m}{2}k_{1v-01}^{\mathrm{cl-co}}A_v^{\mathrm{cl}}e^{-j2\delta_{1v-01}^{\mathrm{cl-co}}z} \\ \sum\limits_{v}\left[\dfrac{\mathrm{dA_v^{cl}}}{\mathrm{d}z} = j\dfrac{m}{2}k_{1v-01}^{\mathrm{cl-co}}A^{\mathrm{co}}e^{j2\delta_{1v-01}^{\mathrm{cl-co}}z}\right] \end{cases} \tag{143.1}$$

where A^{co} is the amplitude for the core mode, A^{cl} is the amplitude for the cladding mode HE_{1v}, k_{01-01}^{co-co} is the coupling constant for core-mode–core-mode, k_{1v-01}^{cl-co} the coupling constant for core-mode–cladding-mode.

$$\delta_{1v-01}^{\mathrm{cl-co}} = \frac{1}{2}\left(\beta_{01}^{\mathrm{co}} - \beta_{1v}^{\mathrm{cl}} - \frac{2\pi}{\Lambda}\right) \tag{143.2}$$

where β_{01}^{co} and is β_{1v}^{cl} the propagation constant of the core mode and cladding mode, and Λ is the period of grating.

If defined S_v as

$$S_v = A_v^{\mathrm{cl}}e^{-j2\delta_{1v-01}^{\mathrm{cl-co}}z} \tag{143.3}$$

following equations have been obtained as

$$\begin{cases} \dfrac{dA^{co}}{dz} = jk_{01-01}^{co-co}A^{co} + \sum\limits_{v} j\dfrac{m}{2}k_{1v-01}^{cl-co}S_v \\ \sum\limits_{v}\left[\dfrac{dS_v}{dz} = j\frac{m}{2}k_{1v-01}^{cl-co}A^{co} - j2\delta_{1v-01}^{cl-co}S_v\right] \end{cases} \tag{143.4}$$

then the matrix form of formula (143.4) can be expressed as

$$\begin{bmatrix} \frac{\mathrm{dA^{co}}}{\mathrm{d}z} \\ \frac{\mathrm{dS_1}}{\mathrm{d}z} \\ \vdots \\ \frac{\mathrm{dS_v}}{dz} \end{bmatrix} = F \begin{bmatrix} A^{\mathrm{co}} \\ S_1 \\ \vdots \\ S_v \end{bmatrix} \tag{143.5}$$

where

$$F = j\begin{bmatrix} k_{0,1-01}^{co-co} & \frac{m}{2}k_{1,1-01}^{cl-co} & \frac{m}{2}k_{1,2-01}^{cl-co} & \cdots & \frac{m}{2}k_{1,v-1-01}^{cl-co} & \frac{m}{2}k_{1,v-01}^{cl-co} \\ \frac{m}{2}k_{1,1-01}^{cl-co} & -2\delta_{1,1-01}^{cl-co} & 0 & \cdots & 0 & 0 \\ \frac{m}{2}k_{1,2-01}^{cl-co} & 0 & -2\delta_{1,2-01}^{cl-co} & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 & \vdots \\ \frac{m}{2}k_{1,v-1-01}^{cl-co} & 0 & \vdots & 0 & -2\delta_{1,v-1-01}^{cl-co} & 0 \\ \frac{m}{2}k_{1,v-01}^{cl-co} & 0 & 0 & \cdots & 0 & -2\delta_{1,v-01}^{cl-co} \end{bmatrix} \tag{143.6}$$

So, the transmission matrix of the uniform LPFG can be expressed as

$$T = e^{FL} \tag{143.7}$$

where L is the grating length.

For the non-uniform LPFG the grating can be divided into hundred segments, and every segment could be considered as uniform. The transmission matrix of the ith segment can be expressed as

$$T_i = e^{FL_i} \tag{143.8}$$

where L_i is the length of ith segment. The total transmission matrix of the grating can be expressed as

$$T = T_0\, T_1 \ldots T_N \tag{143.9}$$

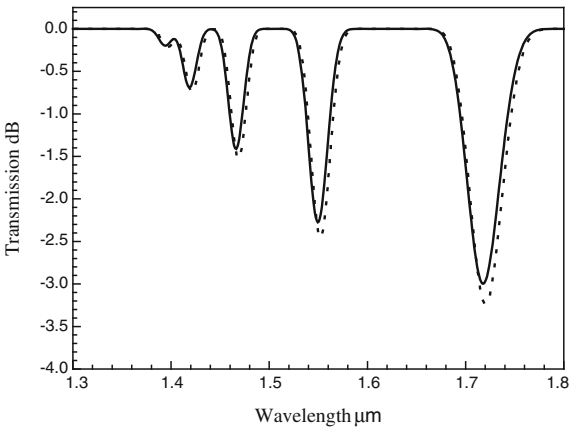


Fig. 143.1 Theoretically calculated transmission spectrum by new TMM (*solid line*) and by SCME (*dotted line*) through a Blackman apodization grating

143.3 Temperature Sensitivity Analysis

In this section we analyze the temperature sensitivity of LPFG by using the new TMM described above. The fiber considered in this paper has a step-index profile and a three layers structure. The parameters of the fiber are given as: the core radius $r_1 = 2.625\,\mu m$, the cladding radius $r_2 = 62.5\,\mu m$, the core index $n_1 = 1.458$, the cladding index $n_2 = 1.45$ and the air index $n_3 = 1.0$.

First, in order to prove the correctness of the new TMM, a Blackman apodization LPFG is analyzed. The transmission spectrum of this grating is shown in Fig. 143.1 (solid line). The grating parameters are given as: grating period $\Lambda = 470\,\mu m$, grating length $L = 25\,mm$, the peak induced-index change is 0.0001. The five main dips seen in the spectrum correspond to coupling to the $v = 1, 3, 5, 7, 9$ cladding modes. For comparison, the transmission spectrum of this grating is calculated again by the SCME, which is illustrated in Fig. 143.1 (dotted line). From this figure we could get the conclusion that the new TMM is exactly enough on analyzing the transmission spectrum of LPFG.

Compared to the SCME method, the advantage of the new TMM is that it could analyze the transmission characteristic of non-uniform LPFG with a faster speed. In order to compare the calculating speed of these two methods, the computing time is illustrated in Fig. 143.2 when different cladding modes have been considered. We could get the conclusion that the calculating speed using the new TMM is faster than that by SCME.

When the ambient temperature varies, the grating period and the effective refractive index will be changed and can be shown as

$$\Lambda(\Delta T) = \Lambda(T_0)(1 + \alpha) \tag{143.10}$$

$$n_{co}^{eff}(\Delta T) = n_{co}^{eff}(T_0)(1 + \xi_{co}) \tag{143.11}$$

$$n_{cl,m}^{eff}(\Delta T) = n_{cl,m}^{eff}(T_0)(1 + \xi_{cl}) \qquad m = 1, 2, \ldots \tag{143.12}$$

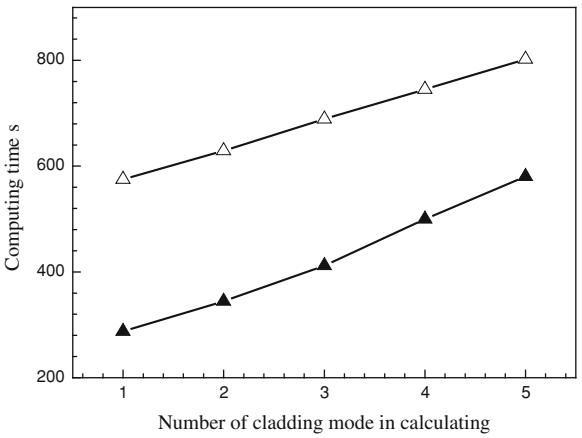


Fig. 143.2 The computing times for new TMM (*filled triangle*) and SCME (*white triangle*)

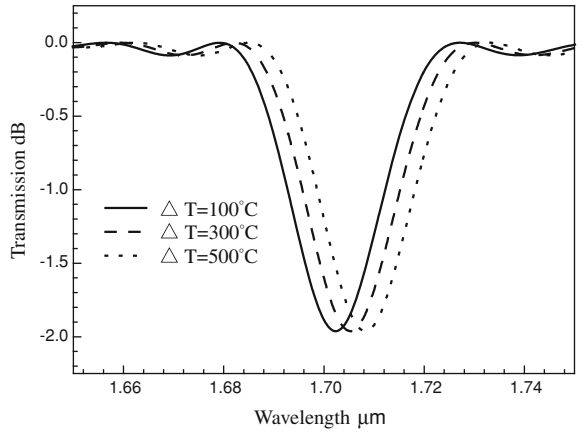


Fig. 143.3 The transmission spectral of LPFG when ambient temperature varied

where T_0 is the room temperature, α is the coefficient of thermal expansion of fiber, ξ_{co} and ξ_{cl} are the thermo-optical coefficient of core and cladding material.

Figure 143.3 shows the transmission spectral of LPFG when ambient temperature varied. The single resonance that is visible with in the wavelength range plotted is associated with coupling to the $v = 9$ cladding mode. This figure shows that the coupled wavelength will shift to the longer wavelength when the ambient temperature rose. The more details on the wavelength shift about more cladding mode is illustrated in Fig. 143.4. From these figure we can find that the larger the ambient temperature rise, the farther wavelength shift.

143.4 Conclusion

In this paper a new transfer matrix method about long period fiber grating with multiple cladding modes is proposed. The new method can be used to analyze the temperature sensitivity of long period fiber grating. Compared with the usually

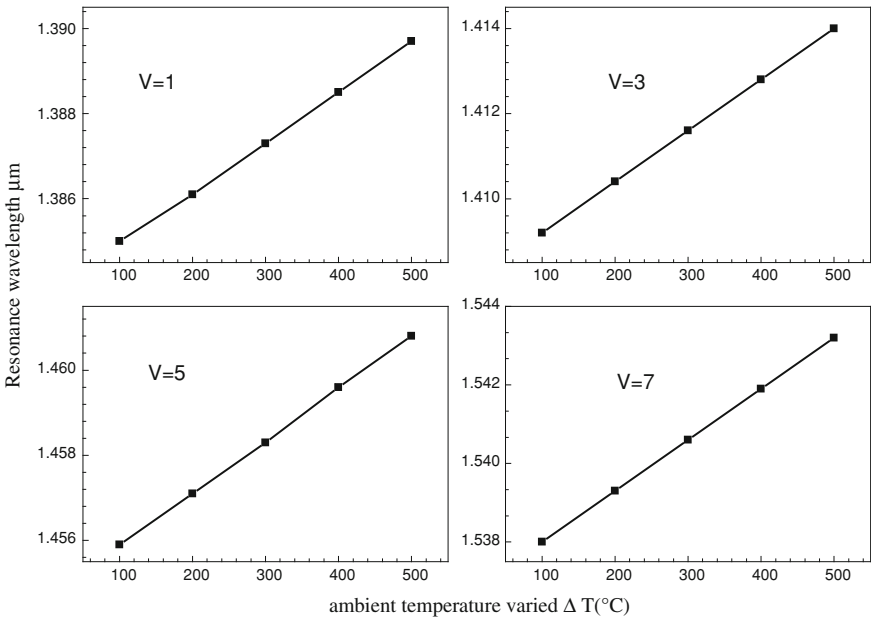


Fig. 143.4 The resonance wavelength of LPFG when ambient temperature varied

used method, such as solving the coupled mode equation by the fourth order adaptive step size Runge–Kutta algorithm, the new transfer matrix method is of a faster calculation speed. The theoretical results are exactly in agreement with the method of solving the coupled mode equation.

Acknowledgments This work is supported by the Program of the National Natural Science Foundation of China (Grant No. 61040016), by the Open Foundation from Henan Provincial Open laboratory of Control Engineering Key Disciplines China (No. KG2009-16) and by the Doctor Foundation from Henan Polytechnic University China (No. 648393).

References

1. Andermahr N, Fallnich C (2010) Optically induced long-period fiber gratings for guided mode conversion in few-mode fibers. Opt Express 18(5):4411–4416
2. Chen K, Sheng Q, Dong X (2004) Band-rejection and band pass filters based on mechanically induced long-period fiber gratings. Microw Opt Technol Lett 42(1):15–17
3. Ni N, Chan CC, Tan KM, Tjin SC, Dong XY (2007) Broad-band EDFA gain flattening by using an embedded long-period fiber grating filter. Opt Commun 271:377–381
4. Kang J, Dong X, Zhao C, Qian W, Li M (2011) Simultaneous measurement of strain and temperature with a long period fiber grating inscribed Sagnac interferometer. Opt Commun 284:2145–2148

5. Venugopalan T, Sun T, Grattan KTV (2010) Temperature characterization of long period gratings written in three different types of optical fiber for potential high temperature measurements. Sens Actuators A 160:29–34
6. Martinez-Rios A, Monzon-Hernandez D, Torres-Gomez I (2010) Highly sensitive cladding-etched arc-induced long period fiber gratings for refractive index sensing. Opt Commun 283:958–962
7. Harhira A, Guay F, Daigle M, Lapointe J, Kashyap R (2010) Long-period fiber gratings fabricated with a CO_2 laser beam and phase mask. Opt Commun 283:4633–4638
8. Eggen CL, Lin YS, Wei Tao, Xiao Hai (2010) Detection of lipid bilayer membranes formed on silica by double-long period fiber grating laser refractometry. Sens Actuators B Chem 150:734–741
9. Nam SH, Lee J, Yin SZ (2011) Control of resonant peak depths of tunable long-period fiber gratings using over coupling. Opt Commun 284:961–964
10. Jiang M, Feng D, Sui Q (2009) Characteristic research on mechanically induced long-period fiber gratings. Chin Opt Lett 7(2):112–114
11. Wang G-D, Xie B-B (2011) Improving the performance of chirped fiber grating with cladding being etched sinusoidal function. Optik 122:557–559
12. Dong X, Feng SC, Ou X, Lu S, Pei Li (2009) Add/drop channel filter based on two parallel long period fiber gratings coupler. Optik 120:855–859
13. Shao L-Y, Laronche A, Smietana M, Mikulic P, Bock WJ, Albert J (2010) Highly sensitive bend sensor with hybrid long period and tilted fiber Bragg grating. Opt Commun 283:2690–2694
14. Erdogan T (1997) Cladding-mode resonances in short- and long-period fiber grating filters. J Opt Soc Am A 14:1760–1773
15. Erdogan T (1997) Fiber grating spectra. J lightw technol 15:1277–1294

Chapter 144
Coal−Rock Interface Detection Using Digital Image Analysis Technique

Jiping Sun and Bo Su

Abstract Aiming at the defects in traditional coal−rock interface detection (CID) system, a new CID model was developed using the digital image analysis technique. First, image features of coal and rock were analyzed in detail and feasibility study of the detection method based on vision was conducted. Second, some rock and coal samples were collected from typical mines in China, and image acquisition was carried out in the laboratory within a simulated environment. Third, the color image of the sample was enhanced by a method based on RGB color-components, in which a new image was synthesized from the color-components with weighted sum method and the weighted coefficients were generated using back-propagation (BP) neural network. Finally, the synthesized gray image was segmented by using a rapid region-based segmentation algorithm and a ratio presenting content of the rock in the coal seam was calculated as the output of the CID system. Study results indicate a very low detection error. The proposed algorithm in this paper is valid for all types of coal mines. The model is required to be calibrated with respective image data taken from the working face.

Keywords Coal−rock interface detection (CID) · Image analysis · BP neural network · Region-based segmentation

J. Sun (✉) · B. Su
State Key Laboratory of Coal Resources and Safety Mining,
China University of Mining and Technology (Beijing),
Beijing 100083, China
e-mail: sjp@cumtb.edu.cn

B. Su
College of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_144,

144.1 Introduction

On the coalface of mines, shearer equipped with automatic control system is one of the critical machines realizing automatic mining. In the ideal case, the shearer is capable of automatically tracking the boundaries of a coal seam and extracting only the coal. The critical element providing automatic seam tracking control for the shearer is coal−rock interface detection (CID) system. The CID system measures the properties of coal and the adjacent formation and then makes intelligent decisions, such as identifying what type of geologic material is being cut (coal or rock), locating where the coal−rock interface is, and determining the thickness of the coal remaining below the roof or above the floor. Figure 144.1 shows the schematic diagram of CID. With the aid of CID system, the height of the cutting drum is automatically adjusted according to the formation of the coalface, which yields significant increases in mining productivity, worker safety and quality of coal extracted.

Researches of CID method have been the major subject in the research of mining technology. Until now, more than 30 types of CID have been put forward in China and abroad beginning in the early 1960s. These methods include natural gamma radiation, radar, infrared radiation, motor power, memory cutting, machine vibration, sound pressure and dust. Each method has limited success under specific geological conditions and general method has not been discovered. The mature methods include natural gamma radiation and memory cutting. In China, only 20% of the mines are appropriate to adopt the method of natural gamma radiation. The memory cutting is semi-automatic, not a real automatic method [1, 2].

At present, the adjustment of drum height is completed by manual operation in most mines of China. The operator of the shearer judges whether the coal or the rock is cutting via both visual and auditory information. But because of poor visibility and big noise on the coalface, it is difficult for the operator to make accurate judgments. As a result, the shearer inevitably cuts the top rock or the bottom rock, causing great increase in rock content, transportation volume and abrasion of cutting tooth. In a word, the automatic CID is one of key techniques of coal face automation and the related researches are strongly valuable both in theory and application [1].

Along with the development of computer vision technology and pattern recognition technology, some analysis and recognition systems based on vision have been applied in coal mines. These applications include detection of coal level in coal bunker, miner face image detection, recognition of empty trains, identification of coal and rejects on conveyor belt, video monitoring of rock-seam and image segmentation of coal flotation froth. The study of CID based on vision is still at an early stage and there are no relevant reports and literature. Due to harsh environment on the coalface, the quality of captured images is vulnerable to factors such as illumination, dust, spray and vibration. Therefore, image analysis and target identification are more difficult than on the ground.

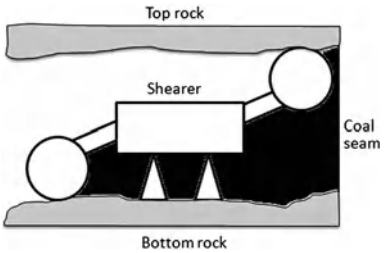


Fig. 144.1 Schematic diagram of CID

This study has investigated the potentiality of the digital image analysis technique for CID in coal mines. Coal and rock samples were collected from different types of coal seams which were selected according to classification of coal. The study was conducted in the laboratory scale in which the sample images were captured in a simulated environment. The images were then preprocessed, segmented and some useful features were extracted to distinguish rock from coal. An image analysis method based on these features was proposed for identification of coal−rock interface.

144.2 Image Features of Coal and Rock

Image features of the object are the premise and foundation of image analysis. Several obvious differences in images features were found to distinguish rock from coal. The image properties of coal and rock are outer shape of chemical composition and molecular structure. These properties include color, luster, texture, fracture shape, etc.

In the national standard "Chinese classification of coals", the coals in Chinese mines are classified into three types, namely anthracite coal, bituminous coal and brown coal. The bituminous coal is further divided into coking bituminous coal and general bituminous coal. Anthracite coal has the highest coalification degree, black color, metallic luster and conchoidal fracture. Brown coal has the lowest coalification degree, brown to brownish-black color, trace of wood on the profile. Bituminous coal has the coalification degree between anthracite coal and brown coal, grayish-black to black color, obvious banded structure [3].

The main types of rocks adjacent to the coal seam include shale, mudstone and sandstone. The shale has leafy or laminated bedding and is lackluster and fragile. The colors of the shale vary in a wide range because of different components. The shale containing organic matter is grey–black to black and that containing iron is red–brown. The sandstone is cemented by sands and the sand content is over 50%. Its color is drab or red. The mudstone is a sedimentary rock solidified form mud and clay. The particle of mudstone is smaller than 0.0039 mm, which will not be seen with the naked eye. The mudstone has horizontal bedding and different colors for different kinds.

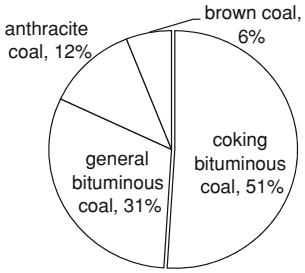


Fig. 144.2 Production ratio of coals in China

In general, there are obvious differences in image features between a type of coal and a type of rock. In a specific working face, the types of coal and rock remain unchanged, so the image features of coal and rock have good stability during the mining process. A CID system is expected to be developed using their images features.

144.3 Sample Collection and Image Acquisition

The study focuses on representative types of coal and rock. In China, production ratio of various coals is presented in Fig. 144.2. The production of anthracite coal and bituminous coal accounts for 94% of the total production [3]. So, anthracite coal and bituminous coal are selected as the main research objects. In order to get representative samples, the research group collected some samples from the mines located in China's Shanxi and Henan provinces. Anthracite coal samples were collected from Longshan Mine of Anyang, 1/3 coking coal (of coking bituminous coal) from No. 10 Mine of Pingdingshan, long flame coal (of general bituminous coal) from Dongzhouyao Mine and Madaotou Mine of Shanxi. The corresponding rock samples were collected in the roof and floor of the coal seam, including shale, mudstone and sandstone.

The image acquisition of the collected samples was carried out in the laboratory within a simulated environment. The experimental setup for this study consists of an experimental shelf, an illumination system, a digital camera, an image acquisition card and a personal computer. An aluminum plate treated with Teflon serves as base board, which greatly reduces reflection. The digital camera (SSC-ET185P, SONY, China) is placed over the base board to capture sample images. A bowl-shaped LED light tube is mounted to provide uniform and diffused lighting. The light tube and the lens are coaxial. The external image acquisition card (MV-U2000, Microvision, China) is used to transfer the images to the computer. A schematic diagram of the image acquisition setup is presented in Fig. 144.3. The images acquired are of 720 $\times$ 576 pixels in size. Figure 144.4 shows some of the images, which were taken during the study.

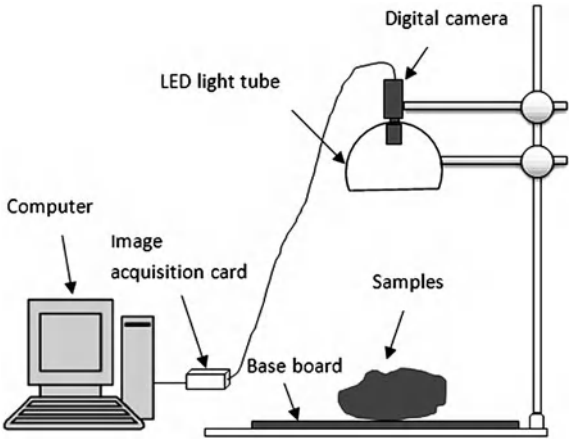


Fig. 144.3 Schematic diagram of the image acquisition setup

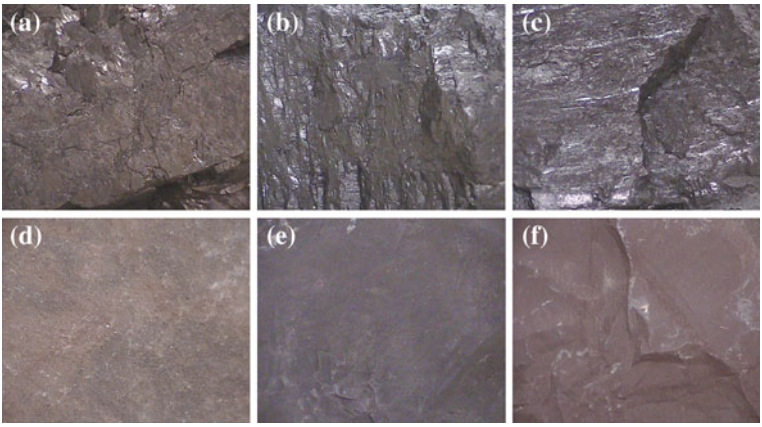


Fig. 144.4 Sample images of different coals and rocks taken from the laboratory **a** Anthracite coal **b** 1/3 Coking coal **c** Long flame coal **d** Sandstone **e** Shale **f** Mudstone

144.4 Image Analysis

The images captured were processed to extract the meaningful attributes for CID. The steps involved for the image analysis include image enhancement, image segmentation and feature extraction.

144.4.1 Image Enhancement

Individual coal and rock have obvious difference in the image features. But the difference is reduced at working face because of the impact of environmental

factors, such as geological structure, coal dust, illumination, etc. The difference in image features needs to be enhanced as much as possible for better segmentation and recognition. This paper presents a special image enhancement method based on RGB color-components. A new image is synthesized from the color-components with weighted sum method. The weighted coefficients are generated using BP neural network.

An original sample image F (color, 24 bit) is separated into three component images F_R, F_G, F_B (gray, 8 bit). A new image F_N, which has more obvious differences between coal and rock, is generated by

$$F_N = T_R * F_R + T_G * F_G + T_B * F_B \tag{144.1}$$

where T_R, T_G, T_B are the weighted coefficients and range from 0 to 1.

For each component image $f(x, y)$, five primary features were extracted. One of the features is histogram average (C_1), a statistical characteristic based on gray histogram. The other four features are based on gray-level co-occurrence matrix (GLCM), including angular second moment (C_2), contrast (C_3), entropy (C_4) and inverse difference moment (C_5). Suppose the gray-level range is [0, L–1]. GLCM and histogram array of $f(x, y)$ are denoted $m(i, j)$ and $h(z)$, where i, j and z are integers in the range [0, L−1]. The GLCM matrix is the average of four matrixes with different directions (0, 45, 90 and 135°). The five features (C_1, C_2,…, C_5) are obtained by the following formulas. For each color image, a total of fifteen features (C_{R1}, C_{R2},…, C_{R5}, C_{G1}, C_{G2}…, C_{G5}, C_{B1}, C_{B2}…, C_{B5}) are extracted.

$$C_1 = \frac{\sum_{z=0}^{L-1} zh(z)}{\sum_{z=0}^{L-1} h(z)} \tag{144.2}$$

$$C_2 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} m^2(i,j) \tag{144.3}$$

$$C_3 = \sum_{i=1}^{L-1}\sum_{j=1}^{L-1} [(i-j)^2 \times m(i,j)] \tag{144.4}$$

$$C_4 = -\sum_{i=1}^{L-1}\sum_{j=1}^{L-1} m(i,j) \times \log_{10} m(i,j) \tag{144.5}$$

$$C_5 = \sum_{i=1}^{L-1}\sum_{j=1}^{L-1} m(i,j)/[1 + (i-j)^2] \tag{144.6}$$

The architecture of the BP neural network model used in the study is shown in Fig. 144.5. The fifteen features were used as the input nodes, the three nodes were used as the hidden layer and two nodes (coal type and rock type) were used as the outputs. The outputs of three nodes in the hidden layer are the weighted coefficients in formula (144.1). Before training the model, the coefficients are initialized

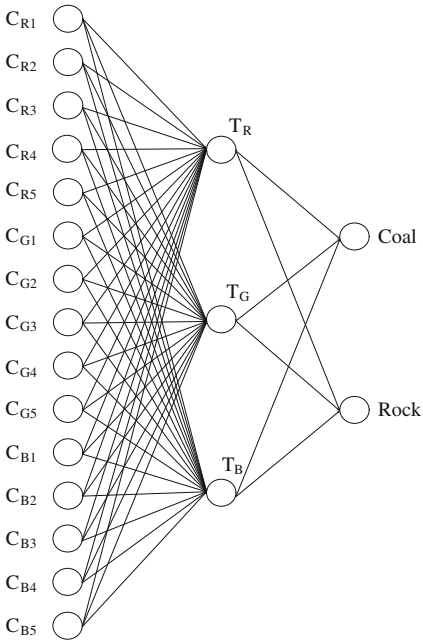


Fig. 144.5 BP neural network architecture for coal−rock classification

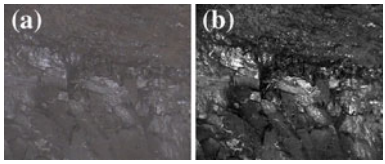


Fig. 144.6 Original image and enhanced image of a coal−rock sample **a** original image **b** enhanced image

as $T_R = 0.229$, $T_G = 0.587$ and $T_B = 0.114$, which are from the fixed formula of converting color image to gray image. The other weights of the network are initialized with random values. During the learning phase, the network is presented with a set of features and corresponding type of the sample (coal or rock). The weights and the three coefficients are updated during the training process. After completion of the training, the model can be used for prediction of coal and rock. The final values of the three coefficients are used to enhance the original image.

Figure 144.6 shows an example of the original color image and its enhanced image obtained by BP neural network and formula (144.1). Enhanced image shows better differential capacity than the original image.

144.4.2 Image Segmentation

In the following steps, the enhanced image is subdivided into constituent regions by the technique of image segmentation. No segmentation technique is 100%

accurate to identify the objects [4]. In this study, a rapid region-based image segmentation algorithm is proposed. The following steps are discussed below.

(1) Grayscale morphological top-hat/bottom-hat transformation.

The grey image (enhanced image) was performed with top-hat and bottom-hat transformations [5]. A disc-shaped structuring element with a radius of 25 pixels was used for these morphological operations. The resulted image was obtained by adding the grey image with top-hat image and subtracting the bottom-hat image. This step results in a significant improvement to maximize the contrast between the coal and rock.

(2) Calculate threshold of segmentation.

Otsu automatic thresholding technique was used to calculate the threshold value. Otsu thresholding is a simple yet effective global automatic thresholding method for binarizing grayscale images such as foregrounds and backgrounds. Because Otsu threshold operates on histograms, it is quite fast [5]. In this step, the Otsu thresholding segmentation was not applied to the image and only the threshold value T was calculated for the region-based image segmentation.

(3) Divide the original image into blocks.

Image block operation improves the efficiency of segmentation. In this step, the original image is divided into non-overlapping and equal rectangular blocks, $m \times n$ blocks. Each rectangular block is an image block.

(4) Classification of each block.

The mean gray degree of each block image, $G_{ij}(i \in [1, m],\ j \in [1, n])$, is calculated and compared with threshold value T. If $G_{ij} \geq T$, all the pixels inside the block is set to 1. If $G_{ij} < T$, all the pixels inside the block is set to 0. Thus the blocks are divided into two classes according to their gray characteristics.

(5) Feature extraction for coal−rock interface.

The ratio of the blocks marked with 1 among all the blocks represents content of rock in the coal seam.

The larger the ratio, the greater content of the rock. The ratio is taken as the output of the CID system. The shearer adjusts the drum height according to the value and varying tendency of the ratio.

144.5 Conclusions

In this paper, a coal−rock interface detection model based on vision was developed by mapping the image features with content of the rock in the coal seam. For conducting this study, broken rock and coal samples were collected from typical

anthracite mines and bituminous mines in China. An image acquisition setup was built up in the laboratory and the images of the samples were taken. The color image of the sample was enhanced by the method based on RGB color-components. The synthesized gray image was segmented by using a rapid region-based segmentation algorithm. Finally a ratio representing content of the rock in the coal seam was calculated as the output of the CID system. The study results indicate a very low detection error. Hence, the effectiveness of the image processing technique is demonstrated in CID. The success of this initial trial study on a laboratory basis is suggesting for the actual field implementation of this strategy in the CID system of the shearer. The vision-based model is sensitive to the environmental factors and the model is a case-specific model. Therefore, the same model could not be useful for the other coal mines. But the proposed algorithm is valid for all types of coal mines. The model is required to be calibrated with respective image data taken from the working face.

References

1. Sun J (2011) Study on identified method of coal and rock interface based on image identification. Coal Sci Technol 2:77–79
2. Mowrey GL (1991) Promising coal interface detection methods. J Min Eng 43(1):134–138
3. Chen P (2007) Character, classification and application of China's coal, 2nd edn. Chemical Industry Press, China
4. Graham DJ, Reid I, Rice SP (2005) Automated sizing of coarse-grained sediments: image-processing prodedures. Math Geol 37(1):1–28
5. Gonzalez RC, Woods RE (2007) Digital image processing, 2nd edn. Publishing House of Electronics Industry, China

Chapter 145
Study on Ultrasonic Detecting Technology for Drill Collar Defects

Lu Yang, Yanhua Zhang and Gang Zhang

Abstract In order to improve the detecting accuracy of drill collar defects and reduce omission and misjudgment, a multi-probe synthesized detecting method was proposed, which is applicable to detect defects in large diameter thick-walled pipe. Hardware schematic diagram was provided, ultrasonic signal processing methods suitable for engineering application were proposed, and location and type of internal defects were displayed through a graphical interface. It makes a beneficial exploration for the automation, digitization and visualization of the drill collar ultrasonic detecting.

Keywords Ultrasound · Drill collar · Transverse wave · Longitudinal wave

145.1 Introduction

Drill collar is the major part of the drilling string substructure. It has thick wall and high rigidity, mainly plays the role of making pressure on the head of drill and preventing slant of well. Due to processing craft limitation of drill collar, crackle, inclusion and stoma defects may be produced so as to form the hidden security danger, even resulting in accidents of the drill collar fracture during oil field drilling. Therefore, an automatic on-line testing system, which is used to evaluate drill collar internal quality, is required to ensure the production quality, service life and meet the needs of mass production.

L. Yang (✉) · Y. Zhang · G. Zhang
National Key Laboratory for Electronic Measurement Technology,
North University of China, Taiyuan, Shanxi, China
e-mail: tyyanglu@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_145,

At present, the commonly used NDT methods include penetrating testing, magnetic particle testing, X-ray testing, eddy testing, and ultrasonic testing. For thick-walled pipe, penetrating testing and magnetic particle testing have the obvious deficiency: penetrating testing can only detect defects of which surface is open. Magnetic particle testing is limited to detect the surface and certain depth of subcutaneous defects [1, 2]. X-ray testing also has inadequacy that high cost, limited detection thickness, long detection cycle, low detection sensitivity (especially for defects whose separation surface is vertical to ray), requiring protection, etc. [3, 4]. Eddy testing can only defect magnetic materials. Its detection depth is limited, it is easily influenced by testing instrument and conditions [5]. Therefore, ultrasonic testing method is used for drill collar defect detection in this paper. The main advantage of the ultrasonic testing is: ultrasonic testing is harmless to the human body, it can detect thicker materials and identify locations and quantities of the defects. Besides, it has higher testing speed, lower cost, and higher sensitivity to the dangerous defects of large area [6, 7].

The ultrasonic testing system in this paper includes ultrasound probe, the hardware circuit, the data acquisition card and comprehensive analyzing software. After the system starting, the comprehensive analyzing software sends signals to hardware circuit, drill collar is controlled in spiral movement and launching circuit launches ultrasonic pulse signal. The data acquisition card acquires defects signals into computer and signals are analyzed by comprehensive analyzing software. After testing, testing results are shown.

145.2 Design of Ultrasound Probe Box and Detection Principle

145.2.1 Ultrasound Probe Box

The structure of ultrasonic testing probe box is shown in Fig. 145.1, it comprises of three ultrasonic probes. Probe A and B are vertical to the top surface of probe box, and probe C has an angle of 15° with the top surface of probe box. Prove A: longitudinal wave vertical detection, mainly detects circumference defects that are parallel to pipe axis. Prove B: transverse wave and longitudinal wave circumference detection, mainly detects radial defects that are parallel to pipe axis. Prove C: transverse wave axial detection, mainly detects radial defects that are vertical to pipe axis. Three probes work together to finish testing tasks. Entire box is fixed on the surface of drill collar by mechanism equipment. Before testing, get the water pipe into water penetration hole and pour water into probe box through water tank. When water flows out continuously from water overflow hole (stating the box is full of water), testing could be begun.

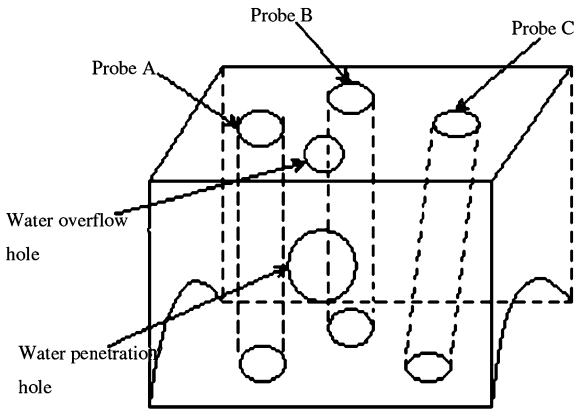


Fig. 145.1 Prove box

145.2.2 Detection Principle

The principle of longitudinal wave vertical detection: when defects are small, defect wave and the bottom wave appear at the same time, the size of the defects can be evaluated according to the height of defects wave. When defects are big enough, the bottom wave will disappear, area size of defects can be measured by the half wave height method.

The principle of transverse wave and longitudinal wave circumference detection: transverse wave detects inner wall and longitudinal wave detects external wall. The eccentricity is calculated according to the principle which is shown in Fig. 145.2.

$$r/R = \sin \beta_s \tag{145.1}$$

$$\sin \alpha_L / C_{L1} = \sin \beta_S / C_{S2} \tag{145.2}$$

$$\sin \alpha_L = x/R \tag{145.3}$$

In Fig. 145.2, α_L is incident angle of longitudinal wave, C_{L1} is speed of longitudinal wave in water, β_s is refraction angle of transverse wave, C_{S2} is speed of transverse wave in steel. Eccentricity x can be obtained according to above three formulas. According to the calculated eccentricity, place the probes to ensure the accuracy of testing for internal groove, external groove of standard defects and inner hole defects. The signals of internal groove defects and inner holes are formed by transverse wave, the signals of external groove defects are formed by longitudinal wave.

The principle of transverse wave axial detection: the key of the method is to calculate the included angle from probe to the top surface. In order to detect defects with transverse wave, the included angle needs to be greater than the first critical angle that is

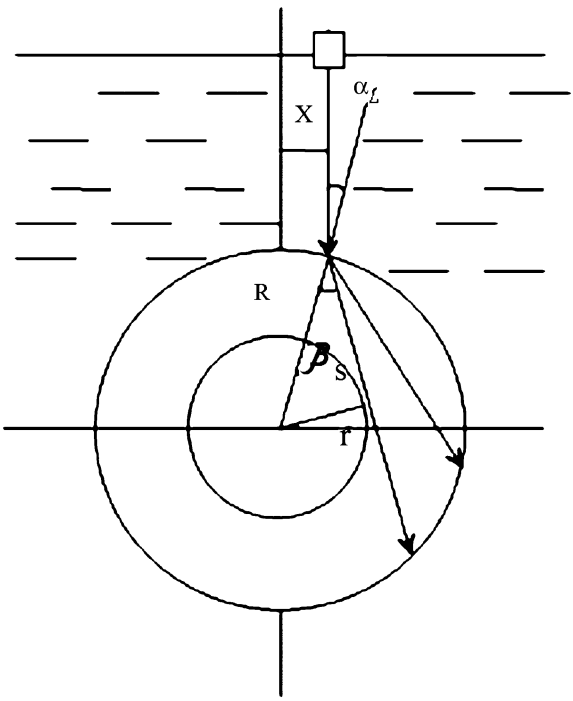


Fig. 145.2 Calculation of eccentricity

$$\alpha \geq \sin^{-1} \frac{C_{L1}}{C_{L2}} \tag{145.4}$$

After the ultrasonic testing to various types of drill collars, eventually 15° is selected as testing angle for good testing effect.

145.3 System Hardware Design

The system hardware design is shown in Fig. 145.3.

Power supply module mainly provides the voltages of ±5 and ±12 V.

Triggering and controlling module partly realize important parameters adjustment of the driving source, such as repeat frequency and the output pulse width, etc.

Launching circuit is triggered and produces high-frequency narrow pulse which is added to probe to stimulate piezoelectric chip vibration and produce ultrasound in work piece. Ultrasound will be reflected when meets defects or undersurface, and be transformed to electrical signal by piezoelectric chip when it returns to probes.

The function of ultrasonic receiving circuit is to make weak signal that ultrasonic probe produced filtered and amplified. The signal is sent to adjustable gain

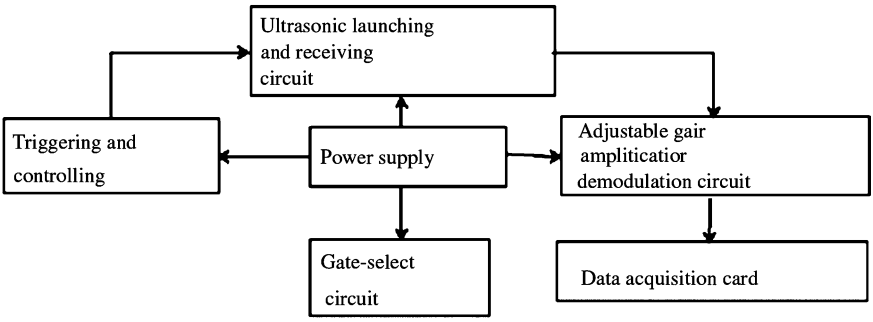


Fig. 145.3 Hardware diagram

amplification demodulation circuit to be further amplified and transformed to demodulation signal.

At the same time, the gate-select circuit will choose the useful part of the demodulation signal and filter out useless waveform. The data acquisition card acquires the demodulation signal and sends it to computer to analyze for processing.

145.4 Software Design and Signal Processing

System software includes the functions of automatic control, ultrasonic signal data acquisition, digital signal processing, types and location recognition of defect and storage and loading of the testing data, etc. Run software and click setting button to set the external diameter, length of drill collar detected, sampled points and the gate wide and height for each channel. After parameters setting, click starting testing button, the system software will appear the inscription "is testing…", which shows that the system is working. When testing is finished, defects will be identified on the corresponding position in the software. When the mouse moves to the position of defects and click double the left mouse button, the corresponding waveform will be displayed in the below window. This software can also storage the acquired data.

Signal processing algorithm is the core of the software. Whether the algorithm is good imposes a direct influence on the accuracy of the testing, and the complexity degree of the algorithm also influences the computer processing speed. So the software here uses the following concise and practical signal processing methods.

(1) *The translation of the data*. Ultrasonic of three probes launched to drill collars is not in the same transverse surface. The data of 3 channels needs to be translated to align to show defects position accurately.

(2) *Binary height filter*. The acquired data is binary processed, that is, setting a gate height and comparing acquired data with gate height. The data above the height is set to 1 and below is set to 0.
(3) *Width filter*. All kinds of interferences are inevitable in the entire automatic ultrasonic testing process, such as the rotation of the electric motor, smoothness of the work piece and water flowing speed, so not all value "1" can be considered defects. Effective way is to record the number of value that is higher than the threshold and set up a threshold width value. Only when the recorded number of "1" is greater than threshold width value, take it as defect. Less than the threshold width value, take it as interference. The choice of gate height value and gate wide value is very important, at the same time they are also contradiction. If the gate height is too low and the gate wide is too narrow, error judge will be caused. If the gate height is too high and the gate wide is too wide, omission will be caused. So all kinds of work pieces need to be tested repeatedly to set a suitable gate.
(4) *Mean filter*. Due to the influence of other mechanical equipments and smoothness of drill collars, large quantities of clutter wave will be brought in when ultrasonic testing. So 5 points mean filter will be adapted to the acquired data for smooth processing and burr eliminating. It is very helpful to recognize defects position and types.

145.5 Conclusion

After the detection of drill collars (outside diameter 124 mm, wall thickness 70 mm, length 9 m) containing artificial internal and external standard defects (length 20 mm, width 1 mm, depth for 2 and 1.5 mm, respectively) and internal standard hole (diameter for 1.5 mm), the results are shown in Fig. 145.4. Results show that length and depth measuring is accurate, and linearity is good.

To prove the repeatability of the system, defects depth measuring and omission, repeated tests are made on the other six standard samples of the drill collars, respectively. Confirm the sensitivity of discard according to standard samples. Panel testing is done for products in order to determine the misjudgment rate. Testing time for single product is determined through the small amount continuous panel testing. Testing results are as follows:

(1) Omission disappears and repeatability is consistent after repeated testing to A, B, C, D, E, F six types of the standard samples.
(2) Omission rate is 0 and the misjudgment rate is less than 1% after testing to 100 A drill collars, 100 D drill collars, 150 B, C, E and F drill collars, respectively.
(3) Testing time of drill collars is 193 s, meeting online production time.

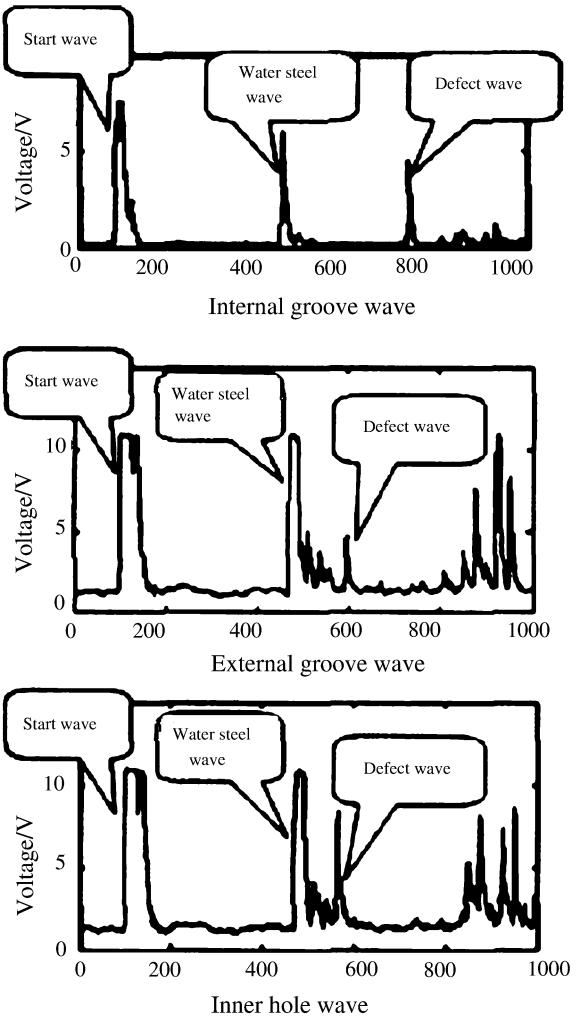


Fig. 145.4 Testing waveforms

References

1. Kasai N, Takada A, Fukuoka K, Aiyama H, Hashimoto M (2011) Quantitative investigation of a standard test shim for magnetic particle testing. NDT&E Int 9(5):421–426
2. Massa GM (2006) Finding the optimum conditions for weld testing by magnetic particles. Non-destruct Test 2(1):16–26
3. Guanghan Peng (2008) X-ray hardening correction for ICT in testing work piece. Spectrosc Spectr Anal 28(6):293–297
4. Östman E, Persson S (2008) Application of X-ray tomography in non-destructive testing of fibre reinforced plastics. Mater Des 5(9):142–147
5. Chen H, Wu X (2011) Ferromagnetic material pulsed eddy current testing signal modeling by equivalent multiple-coil-coupling approach. NDT&E Int 3(2):163–168

6. Beck KH (1993) Ultrasonic refraction angles for inspection throughout the total wall thickness of tubes and pipes. Mater Eval 51(5):875–882
7. Martínez-Martínez J, Benavente D (2008) Multivariate statistical techniques for evaluating the effects of brecciated rock fabric on ultrasonic wave propagation. Int J Rock Mech Min Sci 45(4):609–620

Chapter 146
On-Line Safety Monitoring for Disk Braking System of Rope-Less Elevator

Hongwei Zhang, Fashan Yu, Xinhuan Wang and Youfeng Luo

Abstract The rope-less elevator driven by permanent magnet linear synchronous motor (PMLSM) is a revolutionary new technology for high-rise buildings. In the practical operation, PMLSM inevitably encounters various kinds of turbulences which directly affect the rope-less elevator's safe and steady operation. The function and structure of disk braking system of rope-less elevator is introduced in this chapter. The influencing factors of disk braking system are analyzed, and the research on safety monitoring system technique is made. Combining with practical situation, it is concluded that physical quantity being inspected, the on-line monitoring model is established with modern control method. The monitoring system is developed by programable logic controller to achieve on-line momtoring and alarm of given physical quantity, storage and fetch of important parameter, and search of operating trouble record so as to realize the on-line monitoring safety of braking system.

Keywords Rope-less elevator · Permanent magnet linear synchronous motor · Braking system · On-line monitoring

146.1 Introduction

Many skyscrapers have been built so far in the world, where the high-speed elevators have been developed in order to meet the demand. The typical elevators employ the rope-hoisted method. With the continuous increase in building height,

H. Zhang (✉) · F. Yu · X. Wang · Y. Luo
School of Electrical Engineering and Automation,
Henan Polytechnic University, No. 2001 Century Avenue,
Jiaozuo 454000, Henan, China
e-mail: zhanghw@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_146,

the rope weight may exceed the limit of the strength of the rope itself if the rope length becomes lager than 1,000 m. Therefore rope-less elevators are required to realize a high-rise skyscraper elevator system with sufficient transport capacity [1–3]. Permanent magnet linear synchronous motor (PMLSM) has the characteristics of larger thrust, high efficiency, high power factor and energy conservation, is the ideal driving source for rope-less lifting system. It drives the elevator car directly without wire rope and breaks the limitations of rope-hoisted method.

Supported by the National Natural Science Foundation of China, Excellent Youth Foundation of Henan Province, etc., we have been carrying out search work in the modeling, driver, safe operation control strategy of PMLSM, and constructed a rope-less elevator system driven by PMLSM with multi-segment primary. However, PMLSM inevitably encounters various kinds of turbulences in the practical operation, which directly affect the rope-less elevator's safe and steady operation. It is the key problem that puzzles the designer and the user. If the motor is out-of-step or power fail in operation, the mover and the elevator car are dropped down quickly by gravity, which will cause major security accident. Therefore, designing effective and accurate safety protection system to make sure the safe operation of rope-less system driven by PMLSM has always been the problems that people want to resolve [4].

In this chapter, we described the structure and features of rope-less elevator. The design of safety protection system for rope-less elevator was proposed. The function and structure of disk braking system of rope-less is introduced in the chapter. The influencing factors of disk braking system are analyzed and the research on safety monitoring system technique is made.

146.2 Description of Rope-Less PMLSM Elevator

146.2.1 Configuration of Ropeless Elevator

Figure 146.1 shows the basic configuration of the rope-less elevator experimental apparatus. It has been driven by a single-sided PMLSM. The attracting force adsorbs the mover on the track, which can make the elevator car stable. The specification of which is shown in Table 146.1.

146.2.2 Safety System for Rope-Less Elevator

According to the features of PMLSM rope-less elevator, quadruple protections are designed: energy consumption brake, fail-safe disk brake, safe gear and bottom buffer. In the normal parking state, the system starts disk brake at service floors. Elevators can be positively held during stop and prevent the elevator drop down.

Fig. 146.1 Rope-less elevator driven by single-sided PMLSM

Table 146.1 Specification of rope-less elevator

Name	Value
Driving mode	Single sided PMLSM
Rated power	1.5 (kW)
Rated load	260 (kg)
Mass of cage	70 (kg)
Mass of rated load	190 (kg)
Movable height	4 (m)
Rated velocity	0.3 (m/s)

Once encountering operation faults, such as out-of-step accident, power fail and mechanical fail, the disk brake and energy consumption braking are both started. Safe gear over speed protection is taken as a backup protection, which can effectively prevent the phenomenon that elevator drops down accelerated.

146.3 Design of Disk Braking System

The brake system is normally designed to operate with a controlled braking force to ensure the same retardation levels during all operating situations, regardless of the direction of travel, speed, load or other factors. This greatly improves the safety performance of the rope-less elevator. Normally, there are two braking

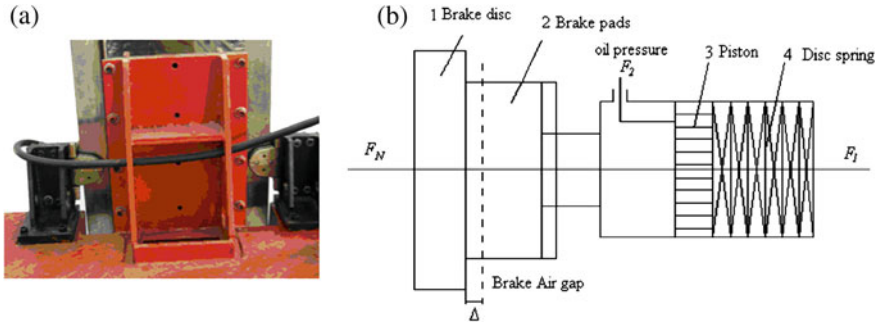


Fig. 146.2 **a** Disk brake system for rope-less elevator **b** sketch diagram of disk brake working

modes: stopping from creep speed to full stop level (service braking), controlled braking during an emergency situation (emergency stop/safety braking). In case of emergency, the butterfly spring will generate positive braking pressure and arrest the fall of the elevator car. During normal stop the hoist speed is reduced by the motor drive system. When the hoist has decelerated to creep speed and the conveyance is located about 0.5 m from the final stop level, the hydraulic pressure in the brake units is reduced to obtain a smooth light contact on the brake disks. The pressure is then reduced to zero in two steps before the conveyance reaches the final stop level. The hydraulic fail-safe disk brake is shown in Fig. 146.2a. The working principle is shown in Fig. 146.2b.

With the increase of the oil pressure, the thrust F_2 generated by oil increased, which pushed the piston compressed disk spring. Spring force will be overcome, which can make the brake pads leave the brake disk. When open the solenoid valve, oil pressure decreases; disc springs forces will move the brake pads to the brake disk to produce the positive braking force and arrest the fall of the elevator car. The hydraulic and electric diagram is shown in Fig. 146.3.

146.4 Design of Monitoring System of Disk Brake

146.4.1 Braking System Fault

(1) Brake does not open. The reason is no oil or hydraulic oil shortage.
(2) The brake cannot brake. The reason is hydraulic or brake damage caused by stuck.
(3) Long braking time, braking distance is long slide, a small braking force.
(4) Uneven brake shoe wear, wear too fast.
(5) Brake slowly.

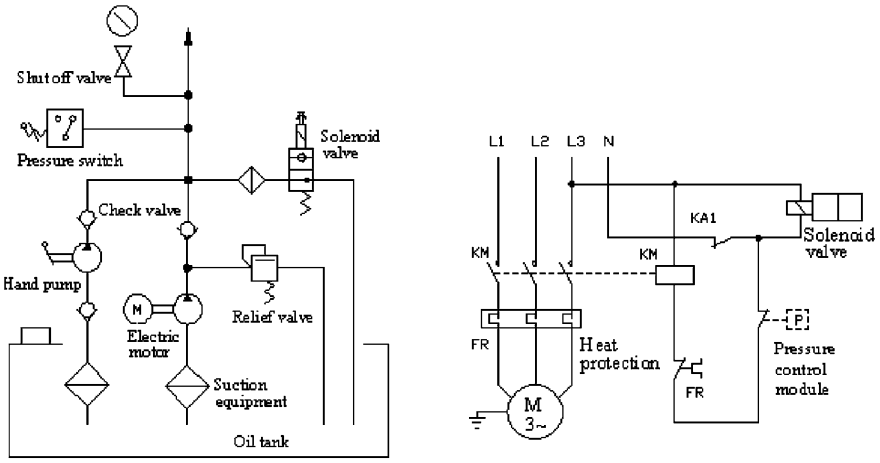
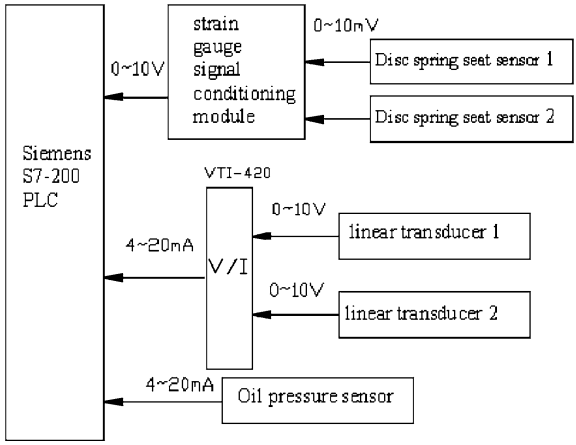


Fig. 146.3 Hydraulic and electric diagram of fail-safe disk brakes

Fig. 146.4 The scheme of brake monitoring system



146.4.2 Scheme of Brake Monitoring System

According to the factors of ropeless elevator disk braking system. The air gap influences the hydraulic brake unit's clamping force and is very important for the reliable performance of the rope-less brake system. If the air gap increases, the clamping force will reduce while lowering spring package life time. The brake monitoring system is therefore provided with electronic sensors to monitor the air gap of each brake caliper half. With a resolution of 0.1 mm, the measurement is very accurate. So the brake monitoring system needs to measure the air gap, oil pressure and disk spring pressure. The monitoring scheme of brake monitoring system is shown in Fig. 146.4.

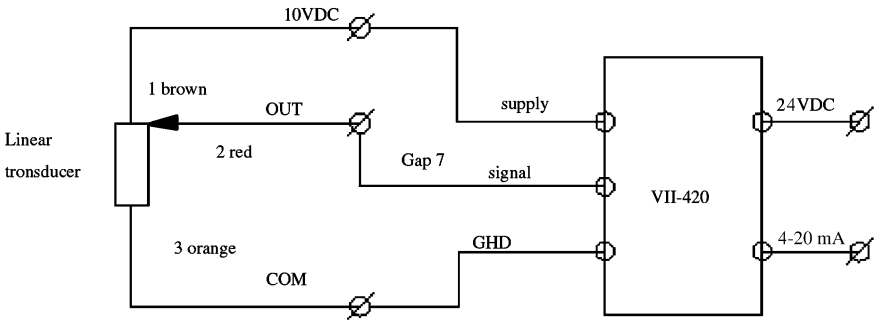


Fig. 146.5 Schematic diagram of linear transducer conditioning circuit

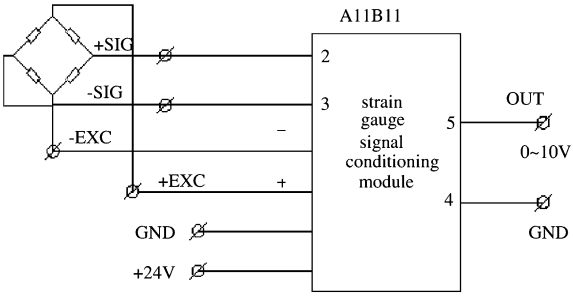


Fig. 146.6 Schematic diagram of gauge signal sensor conditioning circuit

The control and functionality of the brake system is supervised and maintained by an Siemens PLC installed in the elevator cage. It integrates the necessary hardware and software modules. It is provided with a battery back-up to allow full operation during a short power failure. The control system monitors status of the brake equipment, faults and air gaps of the brake calipers

(1) Air gap

The device uses non-contact linear transducer to achieve real-time air gap of brake shoe clearance. The sensor is installed by qualified personnel on the bracket in consideration of all relevant safety regulation. First, store the value s_0 when the brake closes. After a period of time t after brake opens, measure the new value s_t of linear transducer. The difference between the two values is the brake air gap. The schematic diagram of signal conditioning circuit is shown in Fig. 146.5.

$$\delta_t = s_t - s_0. \tag{146.1}$$

(2) Disc spring pressure

Disk spring pressure is measured by strain gauge sensor, which is installed in the disk spring seat. The output signal of strain gauge sensor is 0–10 mV. It needs

to process by isolated strain gauge signal conditioning module. The schematic diagram of disk spring pressure signal conditioning circuit is shown in Fig. 146.6.

146.5 Conclusion

A rope-less elevator driven by PMLSM with multi-segment primary has been made and tested. The function and structure of disk braking system of rope-less elevator is introduced in the chapter. The influencing factors of disk braking system are analyzed, and the research on safety monitoring system technique is made. The brake disk system provides operators numerous benefits including:

(1) Improved safety.
(2) Reduced maintenance.
(3) Controlled braking at emergency stop.
(4) Brake system status and operation monitoring information.

References

1. Chevallier S, Jufer M, Perriard Y et al (2005) Linear motors for multi mobile systems Yamaguchi, Hitoshi. Conf Rec—IAS Annu Meet 3:2099–2106
2. Ahmet O, Ender K, Norio T et al (2010) Design and implementation of a linear motor for multicar elevators. IEEE/ASME Trans Mechatron 15(5):685–693
3. Yamaguchi H, Osawa H, Watanabe T (1996) Brake control characteristics of a linear synchronous motor for ropeless elevator. Int Workshop Adv Motion Control, AMC 2:441–446
4. Wang F, Yuan S (2008) Intelligent fault diagnosis research for permanent magnet linear synchronous motor. Proceedings of the world congress on intelligent control and automation (WCICA) 1951–1955

Chapter 147
Underground Emergency Escape System Based on the Internet of Things

Hui Li, Yingpei Sun and Li Zhang

Abstract To have dynamic emergency escape instructions to guide the trapped workers, underground emergency escape system based on the internet of things is studied. It uses nodes to monitor the concentration of gas in coal mines. The nodes, distributed in the mine, are the component units of the wireless network. If the concentration exceeds the warning line, it broadcasts an emergency signal to the whole network and calculates the next safe evacuation direction through the data collected by real-time algorithm and displays the direction on the LED dynamically. By then, the network shows a number of security paths pointed by all the individual nodes connected together. So it most likely avoids the dangerous areas consisted by disaster nodes and guides people to flee the danger zone.

Keywords WSN · The internet of things · Emergency escape system

147.1 Introduction

Among a number of major mining accidents in the country, the excessive [5] gas in the mine is the dominant element, but at present, at home and abroad, the research of gas monitoring system based on wireless sensor network [1, 2] only can monitor gas, but could not guide people to escape. Such systems only provide an alarm

H. Li (✉) · Y. Sun · L. Zhang
School of Electrical Engineering and Automation,
Henan Polytechnic University, 454010 Henan, China
e-mail: li20022004@hpu.edu.cn

Y. Sun
e-mail: xiaribingcheng329@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_147,

without the instructions which guide people to escape, so they could not provide security [3] guidance to the trapped workers.

Thus, if developed an emergency escape guidance system based on wireless sensor network [1, 2], when excessive [5] gas and other emergency occurs [4–6], the system can combine mine structure with the location of excessive [5] gas and other information, quickly indicate the best routes of escape and provide alarm information which the trapped workers according to in order to flee quickly from the disaster site orderly and detect the location of excessive [5] gas, when mine disaster occurs [6], it can help external aid workers accurately locate the disaster site and organize rescue work more effectively, so to some extent, it can solve the difficult questions of emergency escape and emergency relief when the complex mine emergency occurs [6]. This kind of emergency escape guidance system can curb the development of group of people injury or dead malignant in gas explosion accidents [3–6] and reduce casualties and property losses, so it is of great social benefits and economic benefits.

147.2 System Architecture

The system is underground emergency escape system based on the internet of things. It monitors the real-time concentration of mine gas through the sensors of the node of wireless sensor network [1, 2]. When the event of excessive [5] gas occurs, the node closer to the gas will detect the danger and broadcast an emergency signal to the whole network. Each node calculates the next safety [3] walking direction of miners through real-time algorithm according to the receiving information and display dynamically on the LED. Then, these paths will most likely avoid the dangerous areas consisted by disaster nodes, and guide people to escape danger zone.

The system uses ZigBee technology to build wireless sensor network [1, 2], the entire system including the sensor nodes built in the underground and the daemon device where sensor nodes used for the overall environmental monitoring covered. The sensor node includes sink nodes and common node. Common node communicates with backstage device through Sink node, each common connection wirelessly between nodes. Sink node and as described in the background monitoring device communicates with each other through a wired connection. Underground emergency escape guidance systems need to achieve the overall functional structure shown in Fig. 147.1.

The design of wireless sensor nodes of the system can be divided into four main modules: Master node module, instruction module, sensing module and power source module. These four modules are connected together by connecting lines to form a unified whole. These wireless sensor network [1, 2] nodes are the basic building blocks of WSN, the node performance will directly affect the entire wireless sensor network [1, 2] performance, so the node design is very important.

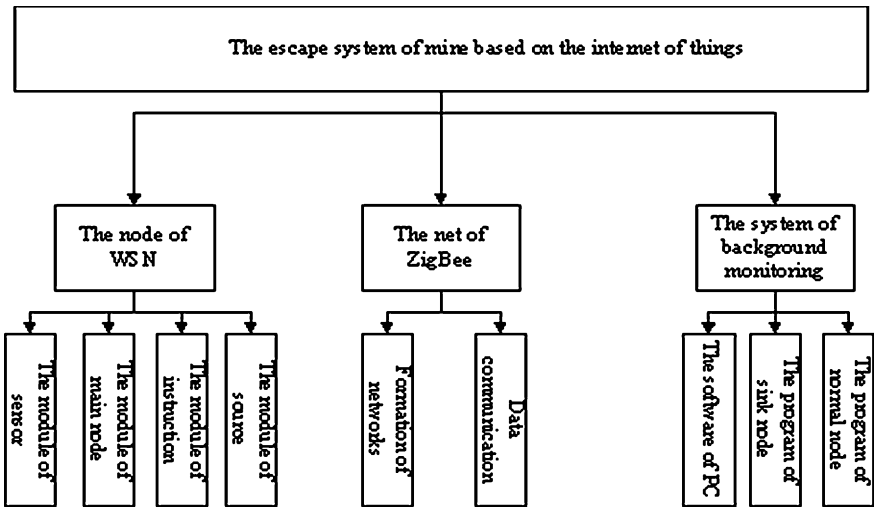


Fig. 147.1 Function diagram of emergency escape system

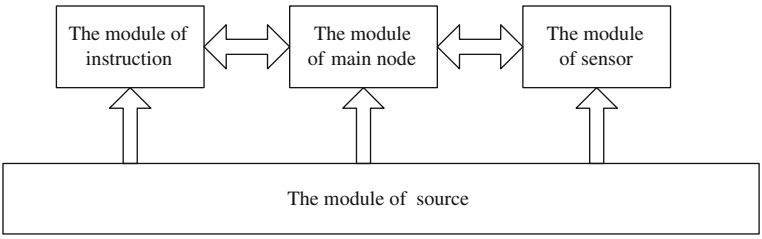


Fig. 147.2 Node architecture

In order to meet the sensing and directive function, the four modules are interconnected as shown in Fig. 147.2.

147.3 System Modules

147.3.1 Master Node Module

The master node module participating in the construction and the correspondence of the wireless network of ZigBee, the core module of this system, is mainly responsible for the self-composition of wireless sensor network [1, 2] and the operation of algorithm to guide escape and display sensing module and

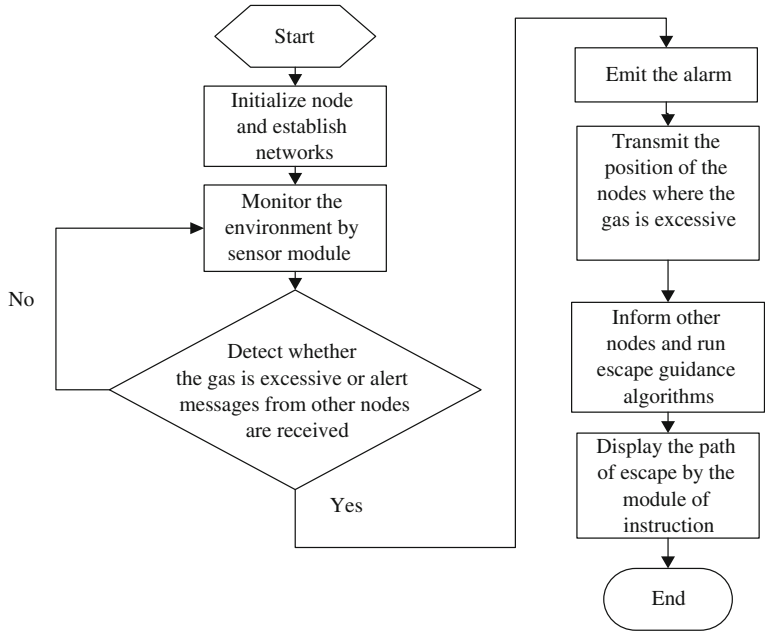


Fig. 147.3 Flow chart of system

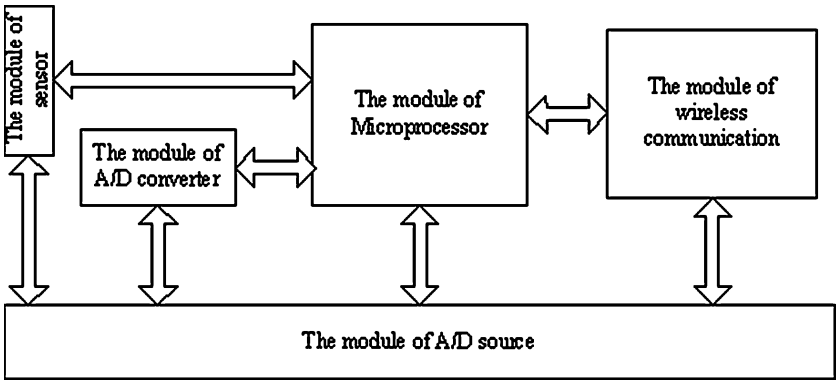


Fig. 147.4 The structure of master node

the crucial correspondence duties and so on. The master node module sensing changes of the environment and detects whether gas is excessive [5] through the sensing module. When gas exceeds [5] the allowed figure it instructs this way through the instruction module. The whole process flow chart is shown in Fig. 147.3 and the structure of master node is shown in Fig. 147.4.

147.3.2 Instruction Module

LED module pointing the escape path to the users in the escape guidance system so as to guide the users rapidly escape from the scene where gas exceeds the allowed figure [5]. The system will use LED display to demonstrate the guidance way. It is not only convenient to guide path where people can escape but also can save the cost. In the system the modules and the master node linkage, according to the instruction of the master node, LED module can display the escape path dynamically and intelligently.

The node instruction module how demonstrate the escapes path specifically is assigned by the master node module. It needs to provide the appropriate interface to communicate with the master node.

147.3.3 Emergency Escape Guidance Algorithms

After completing the initialization process of the node, the system starts guide escape algorithm when the gas information sensed by node or an emergency received.Algorithm measures the degree of risk for each node with weight, the higher the weight the more dangerous areas, in which the path of being expedient of higher value guidance will point to gradually low point right value guidance, until guide to the recent export of the node. When gas exceeds [5] the allowed figure, the weights of all nodes need to be dynamically updated for guiding people to the safest exit. The node receiving the emergency packet first updates the distance to the gas node; then determine whether the node is in the danger zone. If in the danger zone adjusts its weight according to the method of renewal power value adjustment of the node, then determine the node whether a local minimum node and adjust the weight; then the node continues to broadcast emergency information packets; finally node chooses minimum dangerous neighbors as the direction of guidance.

147.3.4 Sensing Module

The sensing module in sensor nodes is to collect the concentration of mine gas and send the information which is collected and treated to controller module. Therefore, when selecting the sensor of the sensor node not only the costs such as the condition but the size and the power consumption and other factors also need to be considered.

TGS2442 is very sensitive to CO, so it is ideal for CO detection. In the situation of CO existence, the conductivity of sensor increases with the concentration of gas increasing, using simple pulse voltage circuit in 1 s loop drive, you can convert

output signals which are corresponding to the gas concentration to conductance rate. TGS2442 sensor features low power consumption, small size, the CO selectivity, high sensitivity, inhibition of alcohol sensitivity, long life, low cost, little affected by temperature.

147.3.5 Power Source Module

The node energy consumption mainly comes from three modules: master node module, instruction module and sensing module. The master node module is the node functional control center and the data processing center, networking in ZigBee networks and wireless communications process will consume energy. The instruction module has responsibility to demonstrate escape way. The display part is demonstrated by the LED lattice module but the LED lattice in demonstration time, needs to consume the massive energies. Simultaneously when master control chip AT89S51 movement also needs to consume the partial energies. Sensing module is mainly responsible for automatic detecting two kinds of events periodic and the aperiodic, its energy consumption total quantity may summarize simply for the single sampling consumption energy product with the sampling number of times.

The design of nodes' energy supply module needs to design energy-saving for each node module. Regarding the master node module and the instruction module, may consider from two aspects the software and hardware. In the hardware, may from low-power devices, the timely sleep and idle down-technology and so on to save the energy consumption. In software, needs to design the more appropriate algorithm. The algorithm is more complex and the more article instructions, so the consumption of energy is bigger. However, the algorithm is the validity, reliability and complexity of the paradox, effective and reliable algorithms tend to have higher complexity. Simple algorithm validity, reliability may not meet the application requirements in specific implementations. This requires a balance between the demands.

In addition, when designs the power source we should consider the explosion-proof mechanism and the design of reserve power source. When mine shaft gas exceeding [5] the allowed figure needs to limit power consumption or have the breakdown needs power to be cut we should ensure power source supply of escapes system is normal.

147.3.6 Backstage Management System

Management system realizes a series of functions such as the guidance algorithm establishment and the node gas monitoring using the wireless sensor network [1, 2] technology, the serial port communications technology and the visual programing

techniques and many other technologies. The guidance algorithm establishment including the initialization condition, the node disposition and the parameter establishment three functions, mainly does the initialization disposition work for the entire escapes guidance system when gas exceeds [5] the allowed figure. Node gas monitoring demonstrates the node arrangement in the building in graph way. When gas exceeding [5] the allowed figure occurs it may demonstrate the escape direction deferring to the guidance algorithm and provide the function of the node information inquiry for monitoring the management of the background system running.

147.4 Conclusion

This paper proposes an emergency escape guidance system based on wireless sensor network. When disaster event occurs in the mine, it can quickly provide an internal alarm and intelligently indicate the best escape routes; simultaneously it not only helps the exterior personnel to locate the disaster place accurately, but also is the effective means which exterior rescue groups use to organize the rescue work effectively. The system plays an important role in avoiding the occurrence of adverse events; meanwhile it plays a key role in evacuating the staff reasonably after the disaster. Therefore, the application prospect of the system is very prominent and its social meaning is bigger than its economic significance far.

Acknowledgments This work is supported by the Research Foundation of Education Bureau of Henan Province (Grant No 2009B510008) and Doctorate Program of Henan Polytechnic University (Grant No B2009-27).

References

1. Jang W, Healy W, Skibniewski M (2008) Wireless sensor networks as part of a web based building environmental monitoring system. Autom Constr 17:729–736
2. Stallings W (2005) Wireless communications and networks. 2nd edn. Pearson/Pretice Hall, New Jersey
3. Li JSM, Chow WKN (2003) Numerical studies on performance evaluation of tunnel ventilation safety systems. Tunn Undergr Space Technol 18:435–452
4. Pomroy WH, Carigiet AM (1995) An analysis of underground coal mine-fire incidents in the United States from 1978 through 1992. In: Proceedings of the seventh US mine ventilation symposium, Lexington KY U S A pp 5–7
5. Simcox S, Wilkes NS, Jones IP (1992) Computer simulation of the flows of hot gases from fire at King's Cross underground station. Fire Safety J 18:49–82
6. Wala AM, Dziurzynski W, Tracz J, Wooton D (1995) Validation study of the mine-fire simulation model. In: Proceedings of the seventh US mine ventilation symposium, Lexington pp 5–7

Chapter 148
Real-Time Detection System of State of Train Wheels Based on LabVIEW

De En, Ningning Wang, Xiaobin Wang, Jieyu Feng and Ningbo Zhang

Abstract Virtual instrument (VI) is the computer technology and traditional instrumentation technology product of the combination, and is an important direction of development of the instrument. LabVIEW is a graphical programing language based on virtual instrument software development tools. This paper uses the LabVIEW 2010 point-by-point analysis of the database, sets up the state of the train wheels real-time detection system. It can carry on the examination in the train normal operation situation to the wheel's condition. As can be seen from the simulation, the test results are precise and rapid, with high practical value.

Keywords LabVIEW · Real-time · Detection

D. En (✉) · N. Wang · X. Wang · J. Feng · N. Zhang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: ende@hpu.edu.cn

N. Wang
e-mail: aning1578@163.com

X. Wang
e-mail: final-dragon@163.com

J. Feng
e-mail: jieyumujin@163.com

N. Zhang
e-mail: mazhutianxia@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_148,

148.1 Introduction

For a long time, China's railway equipment at a low rate, high efficiency, high-intensity state of motion, especially with the speed of the train again to improve, more emphasis on the burden of the train wheels. The wheels of the train is in good condition related to the status of the safe operation of trains, in order to ensure road safety, improve transport capacity, railway departments must often detect the state of train wheels. The traditional detection method is the wheels railway workers struck by a hammer, to judge by the sound of different wheels for damage. However, this manual approach is clearly too slow, easy to false detection, and must be in the train stop to test. For now the current situation, domestic and international efforts are to focus on aspects of the wheels of production or repair the defect detection, online testing is relatively small, based on this situation, this paper introduces a virtual instrument technology and design ideas, Virtual instruments in LabVIEW 2010 development environment, using the LabVIEW 2010 point-by-point analysis of the database, research the real-time detection system of the train wheels.

148.2 LabVIEW Introduction

LabVIEW [1, 2] is a line with icons instead of text to create applications for graphical programing language. Traditional textual programing language based on the sequence of statements and command decision-making process of the reception order, and LabVIEW data flow programing is a way, the block diagram the flow of data between the nodes determines the execution order. It uses the icon expression function, and lines to show data flows. LabVIEW provides a lot of looks with traditional instruments (such as oscilloscopes, multimeters) similar to the controls, can be used to easily create user interface. The user interface is called the front panel in LabVIEW; uses icons and segments, may carry on the control through the programming to front panel's object, which is the graphical source code, also known as G code. LabVIEW graphical source code in a way is similar to the data flow diagram, also known as the block diagram code. Virtual instrument technology is that users in general-purpose computer platform, according to the test task needs to define and design the test function, and its essence is to make full use of computers to implement and extend the functionality of traditional instruments. "Software is the instrument" reflects the essential characteristics of virtual instrument technology. Produced by National Instruments NI-LabVIEW is the most successful, most widely used virtual instrumentation software development system, it is a language based on G 32-bit compiled graphical programming language, and its graphical interface can easily be virtual instrument development.

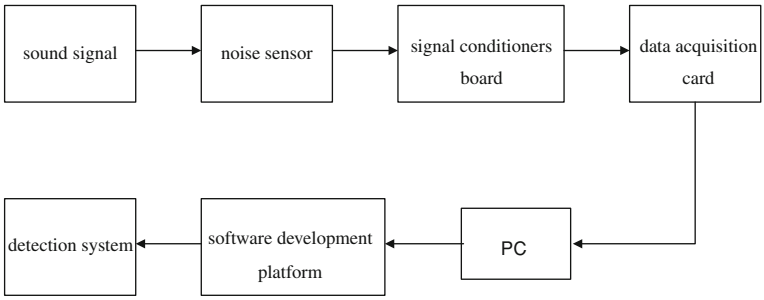


Fig. 148.1 Detection system block diagram

148.3 System Structure

The design of the detection system [3] consists of sensors, signal conditioning board, data acquisition and processing, and software systems, user interface design, etc., the basic framework of the detector is shown in Fig. 148.1. The signal noise in the system by the noise sensor into signal conditioners circuit, signal conditioning devices will send the noise signal into the sensor for sampling the noise signal acquisition cards, data acquisition card for conditioning by the signal after sampling, and sampling results into the computer by the software to complete the signal analysis. Hardware part of course is the basis of the whole system, but a good software design and human–machine interface is an essential part of the system. Virtual instrument panel design resembles a real instrument, users simply select the appropriate button measurements can be observed.

The overall system software develops by the LabVIEW unique graphical language, and is consist of data acquisition, signal analysis, waveform display, data logging, cross the line alarm modules. System software and modules constitute the basic block diagram is shown in Fig. 148.2.

148.4 Detection Principle

Lays aside a detector on the railway rail, will detect a signal with noise. Train wheel signals generated in the running include both low and high frequency components [4], as shown in Fig. 148.3. Low-frequency component is the sound of the train when it is of the normal operation, as shown in Fig. 148.3 left. Good and bad wheels are the same low-frequency components. The maximum low-frequency components of the detector corresponding to the top of the wheel just after a moment, there are two peaks because of the train wheels are usually very close by the two wheels as a pair exists, the corresponding low-frequency components of the minimum or close to the wheel leave time. The high frequency

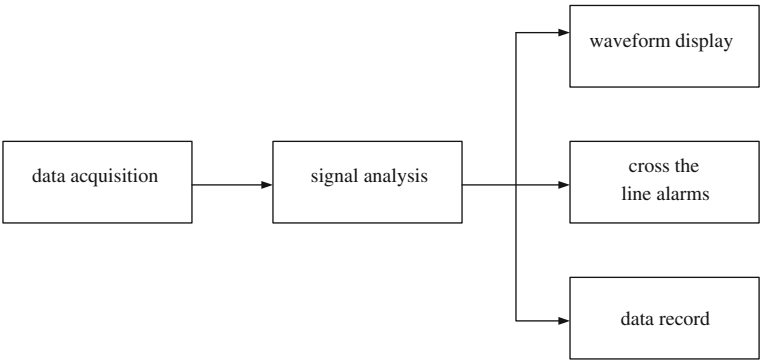


Fig. 148.2 Software system to modules

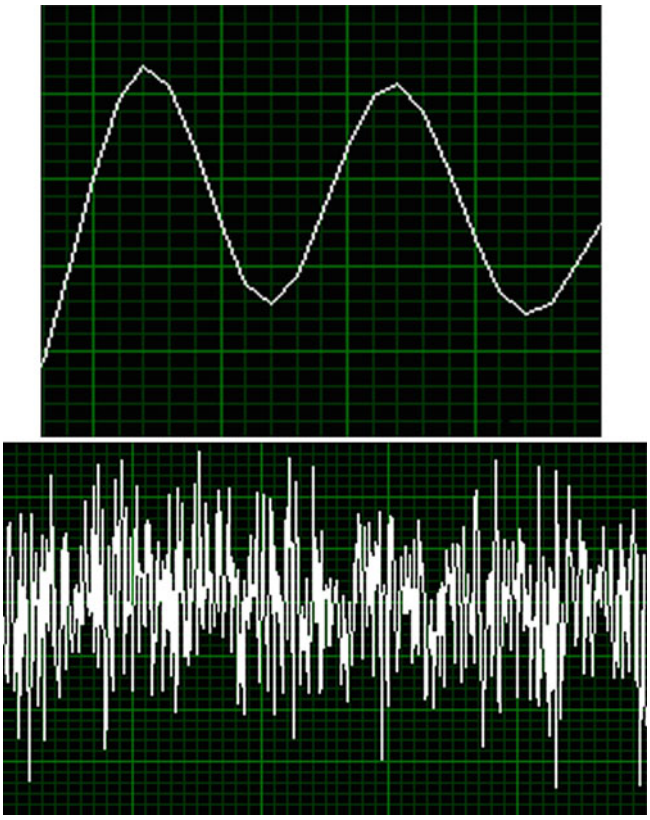


Fig. 148.3 The low frequency and high frequency components of the wheel signal

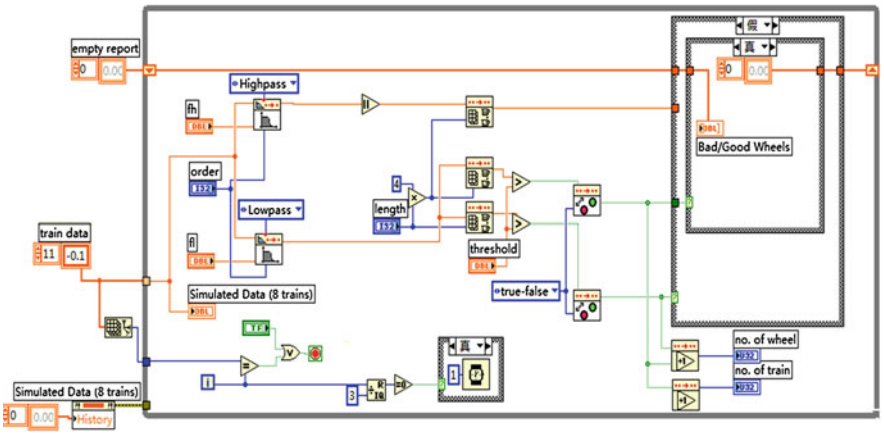


Fig. 148.4 Front panel block diagram

component reflects the extent of damage to the wheel, the greater the energy of high frequency components, that is, the larger the wheels the greater the extent of damage.

148.5 LabVIEW Programing Method

Regarding the speed 60–70 km train, each second several hundreds to several thousand sampling rates can obtain the enough sampling information [5], meets the analysis need. Real-time separates the high frequency component and the low frequency component point-by-point through the Butterwort filter function. Separately makes the peak detection to the high frequency signal and the low-frequency signal. (ArryMax & MinPtByPt.vi) completes the following work separately with three peak detection function.

(1) Each wheel to determine the maximum value of high-frequency energy, thus the extent of damage on behalf of wheels;
(2) Determine whether the end of each train;
(3) Determine whether each of the wheels end. Threshold value is set to 15, it means that when low-frequency signal amplitude is lower than 15, that the wheels have been away from the detection equipment.

Known sampling rate, and then running through the train's speed and the minimum distance between wheels, can be calculated to detect whether the end of the sampling length (length); Similarly, according to wheels size and the minimum distance between the wheels, can be calculated high frequency components of the sampling depth of wheels. Two Boolean Crossing PtByPt.vi functions were used to determine the single wheels or whether the end of the train signals.

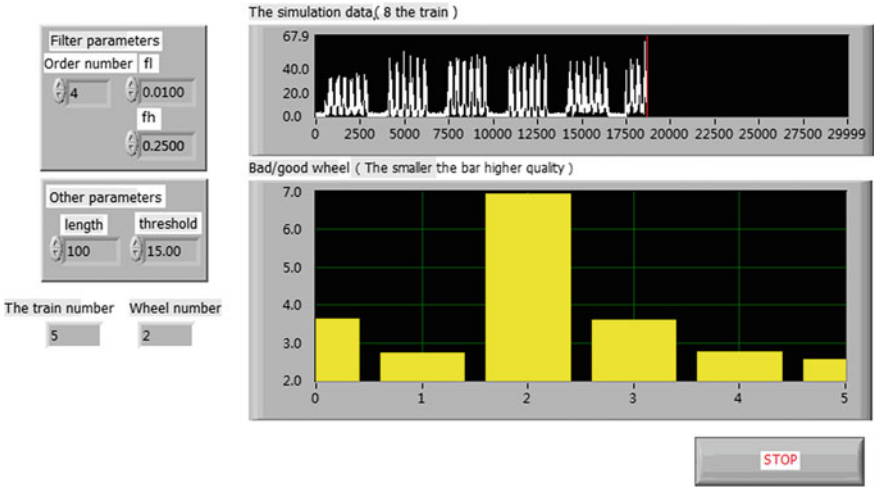


Fig. 148.5 Simulation results

Program front panel diagram in Fig. 148.4.

148.6 Simulation Results

The simulation data acquisition process replaces by the simulated data array, the simulation altogether eight trains tractor, each tractor has six pairs of wheels, the results in histogram display, the scope is higher, represents wheel's damaged condition to be bigger. The simulation result is shown in Fig. 148.5.

148.7 Conclusion

LabVIEW is a revolutionary graphical development environment, can be used to create a generic application, it is called a universal programing language, which is in testing, measurement, automation and other fields have a greater advantage because LabVIEW offers a large number of tools and functions for data acquisition, analysis, display and storage. It also provides a number of commonly used in automated test and measurement field of graphic control, which allows users in a very short period of time to complete a full set of connections from the instrument, data collection to analysis, display and storage of automated test and measurement systems, so it is widely used in automotive, communications, aerospace, semiconductor, electronic design manufacturing, process control and biomedical fields.

This virtual instrument of train wheels on the state of real-time detection system is among the LabVIEW measurement in automation applications. Experimental studies have shown that the virtual instrument technology development and applications in China have great significance.

Acknowledgments This paper is supported by the followings: (1) National Natural Science Foundation of China (41074090) (2) Key Technologies R & D Program of Henan Province (022102210360) (3) Control Engineering Key Discipline Open Foundation Program of Henan Province(KG2009-12) (4) Program of He'nan Educational Committee (2009B4800004)

References

1. Chen X, Zang Y (2007) LabVIEW8.2 programming from basic to familiarity [M]. Tsinghua University Press, Beijing
2. Travis J, Kring J (2008) LabVIEW university practical course. Electronics Industry Press, Beijing
3. Wang H (2006) Automobile tire intelligent detection system [J]. Urban public transp 4
4. Wang P, Li K (2008) Track multi-channel FSK signal detection system software design. J Instrum users 2
5. Ni Y, Wang C (1998) Digital signal processing-principles and implementation. Shanghai Jiaotong University Press, Shanghai

Part XI
Sensors and Signal Processing

Chapter 149
Distributed Optimal Fusion Filter for Multi-Sensor Systems with Finite Consecutive Packet Dropouts

Haixia Chen and Shuli Sun

Abstract The distributed fusion filtering problem is studied for systems with multiple sensors of different finite consecutive packet dropout rates. The phenomenon of packet dropouts is described by a Bernoulli distributed random variable. By introducing a set of new variables, the system is transferred into a random parameter system by using state augmentation method. The filtering error cross-covariance matrix between any two local filters is derived. The distributed fusion state filter is obtained by applying the optimal matrix-weighted fusion estimation algorithm in the linear minimum variance sense. A simulation example shows the effectiveness of the proposed algorithms.

Keywords Multi-sensor system · Finite consecutive packet dropout · Distributed fusion filter · Cross-covariance matrix · Projection theory

149.1 Introduction

In recent years, networked control systems (NCSs) have received widespread attention [1, 2]. In NCSs, data are transmitted through the communication channels from the sensor to the controller and/or from the controller to the actuator. Due to the variability of transmission channels and unreliability of communication, data

S. Sun (✉)
School of Electrical Engineering, Heilongjiang University,
Harbin 150080, China
e-mail: sunsl@hlju.edu.cn

H. Chen
School of Mathematics Science, Heilongjiang University,
Harbin 150080, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_149,

packets may be lost. It will lead to instability and the decline of the system performance.

At present, a lot of research work mainly concentrated on single sensor system with packet loss. The systems with random delays and packet losses can be described through random parameters [3−10]. For systems with multiple packet dropouts, the H_2 filter is presented by linear matrix inequality method [4], the linear minimum variance estimators are presented by the innovation analysis method [5], and the optimal full-order and reduced-order estimators are also presented [6]. In [7], the optimal linear filter with sensor and control packet dropouts is also studied. Moreover, the information fusion estimators are investigated for multi-sensor systems with different packet loss rates, including the optimal centralized and distributed information fusion estimators [8]. In [9], the linear estimators are presented for single-sensor system with finite consecutive packet dropouts. However, so far, multiple sensors are not considered.

On the basis of the results about single sensor in [9], this paper will further study the information fusion filter for multi-sensor systems with different packet dropout rates. We will present the distributed fusion filter by the matrix weighted fusion algorithm in the linear minimum variance sense [10].

149.2 Problem Formulation

Consider the discrete linear stochastic system with finite multiple packet dropouts

$$x(t+1) = \Phi x(t) + \Gamma\omega(t) \tag{149.1}$$

$$z_i(t) = H_i x(t) + v_i(t) \tag{149.2}$$

$$\begin{aligned} y_i(t) = {} & \xi_i(t)z_i(t) + (1-\xi_i(t))\xi_i(t-1)z_i(t-1) + \cdots \\ & + (1-\xi_i(t))(1-\xi_i(t-1))\cdots \\ & (1-\xi_i(t-d_i+1))\ z_i(t-d_i)\ \ i=1,2,\ldots,L \end{aligned} \tag{149.3}$$

where $x(t) \in R^n$ is the state, $z_i(t) \in R^{m_i}$ is the sensor output, $y_i(t) \in R^{m_i}$ is the output received by the filter, $\omega(t) \in R^r$ is the input white noise, $v_i(t) \in R^{m_i}$ is the sensor observation noise, Φ, Γ and H_i are constant matrices with suitable dimensions, and $\xi_i(t)$ is a Bernoulli distributed random variable with probability $\text{prob}\{\xi_i(t) = 1\} = \alpha_i,\ 0 \le \alpha_i \le 1$, $\text{prob}\{\xi_i(t) = 0\} = 1 - \alpha_i$ and is uncorrelated with other random variables. $d_i > 0$ is the number of the largest packet dropout. L is the number of sensors.

Model (149.3) [9] means that the number of consecutive packet dropouts in data transmission is bounded by d_i. The latest packet received previously will be used if the current packet is lost. Furthermore, it also describes the possible delay d_i.

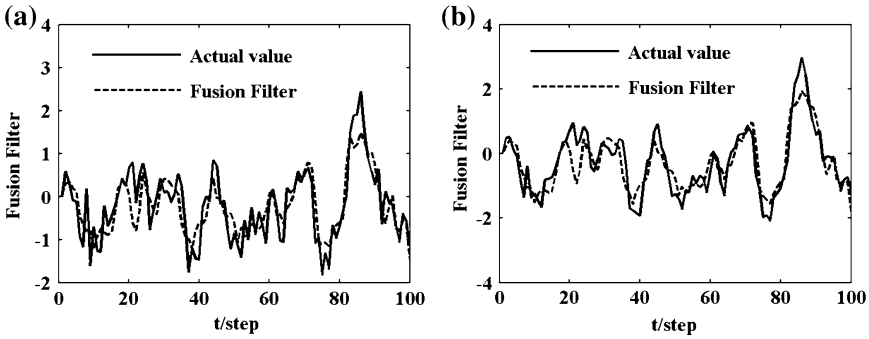


Fig. 149.1 Distributed fusion filter weighted by matrices. **a** The first state component. **b** The second state component

Introducing the new variables

$$\begin{aligned} Y_{i,k}(t) &= \xi_i(t)z_i(t) + (1-\xi_i(t))Y_{i,k+1}(t-1),\ k = 1, 2, \ldots, d_i - 1 \\ Y_{i,d_i}(t) &= z_i(t) \end{aligned} \tag{149.4}$$

Substituting (149.2) into (149.4) and $y_i(t) = Y_{i,1}(t)$. Letting $X_i(t+1) = \left[x^{\mathrm{T}}(t+1), Y_{i,1}^{\mathrm{T}}(t), Y_{i,2}^{\mathrm{T}}(t), \ldots, Y_{i,d_i}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$ and $W_i(t) = \left[\omega^{\mathrm{T}}(t), v_i^{\mathrm{T}}(t)\right]^{\mathrm{T}}$, where the superscript T denotes the transpose, we have the following system equivalent to (149.1)–(149.3) by augmentation:

$$X_i(t+1) = \tilde{\Phi}_i(t)X_i(t) + \tilde{\Gamma}_i(t)W_i(t) \tag{149.5}$$

$$y_i(t) = \tilde{H}_i(t)X_i(t) + \xi_i(t)v_i(t) \tag{149.6}$$

where

$$\begin{aligned} \tilde{\Phi}_i(t) &= \Phi_{0i} + \xi_i(t)\Phi_{1i} + (1-\xi_i(t))\Phi_{2i}, \tilde{\Gamma}_i(t) = \Gamma_{0i} + \xi_i(t)\Gamma_{1i}, \tilde{H}_i(t) \\ &= \xi_i(t)H_{0i} + (1-\xi_i(t))H_{1i} \end{aligned}$$

and

$$\Phi_{0i} = \begin{bmatrix} \Phi_n & 0_{n\times d_i m_i} \\ 0_{(d_i-1)m_i\times(n+d_i)m_i} & \\ H_i & 0_{m_i\times d_i m_i} \end{bmatrix}, \Phi_{1i} = \begin{bmatrix} H_i & 0_{n\times(n+d_i m_i)} \\ & 0_{(d_i-1)m_i\times d_i m_i} \\ & 0_{m_i\times(n+d_i m_i)} \end{bmatrix},$$

$$\Phi_{2i} = \begin{bmatrix} & 0_{n\times(n+d_i m_i)} & \\ 0_{(d_i-1)m_i\times(n+m_i)} & & I_{(d_i-1)m_i} \\ & 0_{m_i\times(n+d_i m_i)} & \end{bmatrix}, \Gamma_{0i} = \begin{bmatrix} \Gamma & 0_{n\times m_i} \\ & 0_{(d_i-1)m_i\times(m_i+r)} \\ 0_{m_i\times r} & I_{m_i} \end{bmatrix},$$

$$\Gamma_{1i} = \begin{bmatrix} 0_{n\times(m_i+r)} \\ 0_{(d_i-1)m_i\times r} \quad I_{(d_i-1)m_i\times m_i} \\ 0_{m_i\times(m_i+r)} \end{bmatrix}, \quad \begin{aligned} H_{0i} &= [H_i \quad 0_{m_i\times d_i m_i}] \\ H_{1i} &= [0_{m_i\times n} \quad I_{m_i} \quad 0_{m_i\times(d_i-1)m_i}] \end{aligned} \tag{149.7}$$

The work of this paper is carried out based on the following assumptions.

Assumption 1 $\omega(t)$ and $v_i(t)$ are correlated white noises with zeros mean and $\mathrm{E}[\omega(t)\omega^{\mathrm{T}}(t)] = Q_\omega$, $\mathrm{E}[\omega(t)v_i^{\mathrm{T}}(t)] = S_{\omega v_i}$, $\mathrm{E}[v_i(t)v_j^{\mathrm{T}}(t)] = R_{ij}$, where $R_{ii} = R_i$, we get

$$S_{ij} = \mathrm{E}\begin{bmatrix} W_i(t) & v_j^{\mathrm{T}}(k) \end{bmatrix} = \mathrm{E}\left\{ \begin{bmatrix} \omega(t) \\ v_i(t) \end{bmatrix} v_j^{\mathrm{T}}(k) \right\} = \begin{bmatrix} S_{\omega v_j} \\ R_{ij} \end{bmatrix} \delta_{tk} \tag{149.8}$$

$$Q_{W_{ij}} = \mathrm{E}\begin{bmatrix} W_i(t) & W_j^{\mathrm{T}}(k) \end{bmatrix} = \mathrm{E}\left\{ \begin{bmatrix} \omega(t) \\ v_i(t) \end{bmatrix} \begin{bmatrix} \omega^{\mathrm{T}}(k) & v_j^{\mathrm{T}}(k) \end{bmatrix} \right\} = \begin{bmatrix} Q_\omega & S_{\omega v_j} \\ S_{\omega v_i}^{\mathrm{T}} & R_{ij} \end{bmatrix} \delta_{tk} \tag{149.9}$$

where $S_{ii} = S_i$, $Q_{W_{ii}} = Q_{W_i}$, E is the expectation, δ_{tk} is the Kronecker delta function.

Assumption 2 The initial state $x(0)$ with mean μ_0 and covariance matrix P_0 is independent of $\omega(t)$, $v_i(t)$ and $\xi_i(t)$, $i = 1, 2, \ldots, L$.

Based on the received measurements $(y_i(t), y_i(t-1), \ldots, y_i(0))$, $i = 1, 2, \ldots L$, our objective is to find the distributed matrix weighted fusion optimal Kalman filter $\hat{x}_o(t|t)$ such that it satisfies the unbiasedness and optimality in the linear minimum variance sense. Note that the value of $\xi_i(t)$ at each time instant is unknown but the probability α_i, and satisfies $0 \le a_i \le 1$, $i = 1, 2, \ldots, L$.

149.3 Distributed Fusion Filter

Before giving the filter of system (149.5), we first introduce some lemmas.

Lemma 1 For the random variable $\xi_i(t)$ and (149.6), the following results hold:

$$\begin{aligned} &\mathrm{E}[\xi_i(t)] = \alpha_i, \mathrm{E}[\xi_i^2(t)] = \alpha_i,\ \mathrm{E}[(1-\xi_i(t))^2] = 1-\alpha_i, \mathrm{E}[\xi_i(t)(1-\xi_i(t))] = 0, \\ &\mathrm{E}[\xi_i(t)(1-\xi_j(t))] = \alpha_i(1-\alpha_j), i \ne j, \mathrm{E}[(\xi_i(t)-\alpha_i)(\xi_j(t)-\alpha_j)] = 0, \\ &\mathrm{E}[(1-\xi_i(t))(1-\xi_j(t))] = (1-\alpha_i)(1-\alpha_j) \end{aligned} \tag{149.11}$$

$$\begin{aligned} &\bar{\Phi}_i = \mathrm{E}[\tilde{\Phi}_i(t)] = \Phi_{0i} + \alpha_i\Phi_{1i} + (1-\alpha_i)\Phi_{2i}, \\ &\bar{\Gamma}_i = \mathrm{E}[\tilde{\Gamma}_i(t)] = \Gamma_{0i} + \alpha_i\Gamma_{1i}, \bar{H}_i = \mathrm{E}[\tilde{H}_i(t)] = \alpha_i H_{0i} + (1-\alpha_i)H_{1i} \end{aligned} \tag{149.12}$$

$$\begin{aligned}\tilde{\Phi}_i(t)-\bar{\Phi}_i&=(\xi_i(t)-\alpha_i)\breve{\Phi}_i,\ \tilde{\Gamma}_i(t)-\bar{\Gamma}_i=(\xi_i(t)-\alpha_i)\Gamma_{1i}\\ \tilde{H}_i(t)-\bar{H}_i&=(\xi_i(t)-\alpha_i)\breve{H}_i\end{aligned}\tag{149.13}$$

where

$$\breve{\Phi}_i=\Phi_{1i}-\Phi_{2i},\breve{H}_i=H_{0i}-H_{1i}\tag{149.14}$$

$$\begin{aligned}Q_i=&\mathrm{E}[\tilde{\Gamma}_i(t)W_i(t)W_i^{\mathrm{T}}(t)\tilde{\Gamma}_i^{\mathrm{T}}(t)]=\Gamma_{0i}Q_{W_i}\Gamma_{0i}^{\mathrm{T}}+\alpha_i\Gamma_{0i}Q_{W_i}\Gamma_{1i}^{\mathrm{T}}+\alpha_i\Gamma_{1i}Q_{W_i}\Gamma_{0i}^{\mathrm{T}}+\alpha_i\Gamma_{1i}Q_{W_i}\Gamma_{1i}^{\mathrm{T}},\\ Q_{ij}=&\mathrm{E}[\tilde{\Gamma}_i(t)W_i(t)W_j^{\mathrm{T}}(t)\tilde{\Gamma}_j^{\mathrm{T}}(t)]\\ =&\Gamma_{0i}Q_{W_{ij}}\Gamma_{0j}^{\mathrm{T}}+\alpha_j\Gamma_{0i}Q_{W_{ij}}\Gamma_{1j}^{\mathrm{T}}+\alpha_i\Gamma_{1i}Q_{W_{ij}}\Gamma_{0j}^{\mathrm{T}}+\alpha_i\alpha_j\Gamma_{1i}Q_{W_{ij}}\Gamma_{1j}^{\mathrm{T}}\end{aligned}\tag{149.15}$$

$$\begin{aligned}\bar{S}_i&=\mathrm{E}[\tilde{\Gamma}_i(t)W_i(t)v_i^{\mathrm{T}}(t)\xi_i(t)]=\alpha_i(\Gamma_{0i}+\Gamma_{1i})S_i,\\ \bar{S}_{ij}&=\mathrm{E}[\tilde{\Gamma}_i(t)W_i(t)v_j^{\mathrm{T}}(t)\xi_j(t)]=(\alpha_j\Gamma_{0i}+\alpha_i\alpha_j\Gamma_{1i})S_{ij}\end{aligned}\tag{149.16}$$

Proof From the distribution of $\xi_i(t)$, (149.11) can be easily verified. (149.12) follows directly from taking expectation on (149.6). (149.13) and (149.14) follows directly from (149.6) and (149.12). Substitution (149.6) (149.8) and (149.9) into the equation, we obtain (149.15) and (149.16).

Lemma 2 Under Assumptions 1 and 2, the state covariance matrix of system (149.5) satisfies the following recursion:

$$q_i(t+1)=\alpha_i(1-\alpha_i)\breve{\Phi}_i q_i(t)\breve{\Phi}_i^{\mathrm{T}}+\bar{\Phi}_i q_i(t)\bar{\Phi}_i^{\mathrm{T}}(t)+Q_i\tag{149.17}$$

where $q_i(t)=\mathrm{E}[X_i(t)X_i^{\mathrm{T}}(t)]$ with the initial value $q_i(0)=\begin{bmatrix}P_0+\mu_0\mu_0^{\mathrm{T}} & 0\\ 0 & 0\end{bmatrix}$

Proof It follows directly from (149.5) and Lemma 1.

149.3.1 Local Optimal Kalman Filter

In the following, we will give the local optimal filter.

Lemma 3 [9] Under Assumptions 1 and 2, the ith local sensor subsystem (149.5) has the optimal recursive one-step predictor and filter

$$\hat{X}_i(t|t)=\hat{X}_i(t|t-1)+K_i(t)\varepsilon_i(t)\tag{149.18}$$

$$\hat{X}_i(t+1|t)=\bar{\Phi}_i\hat{X}_i(t|t-1)+L_i(t)\varepsilon_i(t)\tag{149.19}$$

$$\varepsilon_i(t)=y_i(t)-\bar{H}_i\hat{X}_i(t|t-1)\tag{149.20}$$

$$K_i(t) = P_i(t|t-1)\bar{H}_i^{\mathrm{T}} Q_{\varepsilon_i}^{-1}(t) \tag{149.21}$$

$$L_i(t) = [\alpha_i(1-\alpha_i)\breve{\Phi}_i q_i(t)\breve{H}_i + \bar{\Phi}_i P_i(t|t-1)\bar{H}_i^{\mathrm{T}} + \bar{S}_i]Q_{\varepsilon_i}^{-1}(t) \tag{149.22}$$

$$Q_{\varepsilon_i}(t) = \alpha_i(1-\alpha_i)\breve{H}_i q_i(t)\breve{H}_i^{\mathrm{T}} + \bar{H}_i P_i(t|t-1)\bar{H}_i^{\mathrm{T}} + \alpha_i R_i \tag{149.23}$$

$$\begin{aligned} P_i(t+1|t) &= \alpha_i(1-\alpha_i)[\breve{\Phi}_i - L_i(t)\breve{H}_i]q_i(t)\,[\breve{\Phi}_i - L_i(t)\breve{H}_i]^{\mathrm{T}} + Q_i \\ &+[\bar{\Phi}_i - L_i(t)\bar{H}_i]P_i(t|t-1)[\bar{\Phi}_i - L_i(t)\bar{H}_i]^{\mathrm{T}} - \bar{S}_i L_i^{\mathrm{T}}(t) - L_i(t)\bar{S}_i^{\mathrm{T}} + \alpha_i L_i(t) R_i L_i^{\mathrm{T}}(t) \end{aligned} \tag{149.24}$$

$$P_i(t|t) = P_i(t|t-1) - K_i(t)Q_{\varepsilon_i}(t)K_i^{\mathrm{T}}(t) \tag{149.25}$$

where $q_i(t)$ is computed by (149.17), $\varepsilon_i(t)$ is the innovation with covariance $Q_{\varepsilon_i}(t)$, $P_i(t+1|t)$ and $P_i(t|t)$ are the prediction and filtering error covariance matrices of the ith local sensor subsystem, $K_i(t)$ and $L_i(t)$ are the filtering and prediction gain matrices. The initial values are $\hat{X}_i(0|-1) = [\mu_0^{\mathrm{T}}\ 0]^{\mathrm{T}}$ and $P_i(0|-1) = \begin{bmatrix} P_0 & 0 \\ 0 & 0 \end{bmatrix}$.

149.3.2 Cross-Covariance Matrix

Theorem 1 For the ith and jth sensor subsystems of system (149.5) with Assumptions 1 and 2, the cross-covariance matrices of optimal one-step prediction and filtering errors between them are computed by

$$\begin{aligned} P_{ij}(t+1|t) =& [\bar{\Phi}_i - L_i(t)\bar{H}_i]P_{ij}(t|t-1)[\bar{\Phi}_j - L_j(t)\bar{H}_j]^{\mathrm{T}} \\ &+ \alpha_i\alpha_j L_i(t)R_{ij}L_j^{\mathrm{T}}(t) + Q_{ij} - \bar{S}_{ij}L_j^{\mathrm{T}}(t) - L_i(t)\bar{S}_{ji}^{\mathrm{T}} \end{aligned} \tag{149.26}$$

$$P_{ij}(t|t) = [I_n - K_i(t)\bar{H}_i]P_{ij}(t|t-1)[I_n - K_j(t)\bar{H}_j]^{\mathrm{T}} + \alpha_i\alpha_j K_i(t)R_{ij}K_j^{\mathrm{T}}(t) \tag{149.27}$$

Proof Substituting (149.6) into (149.20), the innovation $\varepsilon_i(t)$ can be rewritten as

$$\varepsilon_i(t) = (\xi_i(t) - \alpha_i)\breve{H}_i X_i(t) + \bar{H}_i\tilde{X}_i(t|t-1) + \xi_i(t)v_i(t) \tag{149.28}$$

Subtracting (149.19) from (149.5) and using (149.28), we have the prediction error equation

$$\begin{aligned} \tilde{X}_i(t+1|t) =& [\bar{\Phi}_i - L_i(t)\bar{H}_i]\tilde{X}_i(t|t-1) \\ &+ [\xi_i(t) - \alpha_i][\breve{\Phi}_i - L_i(t)\breve{H}_i]X_i(t) + \tilde{\Gamma}_i(t)W_i(t) - \xi_i(t)L_i(t)v_i(t) \end{aligned} \tag{149.29}$$

We easily obtain (149.26) by computing $P_{ij}(t+1|t) = \mathrm{E}[\tilde{X}_i(t+1|t)\tilde{X}_j^{\mathrm{T}}(t+1|t)]$. It follows from (149.5) and (149.28) that the filtering error equation is given by

$$\tilde{X}_i(t|t) = [I_n - K_i(t)\bar{H}_i]\tilde{X}_i(t|t-1) - [\xi_i(t) - \alpha_i]K_i(t)\breve{H}_i X_i(t) - \xi_i(t)K_i(t)v_i(t) \tag{149.30}$$

Substituting (149.30) into the filtering error covariance matrix equation $P_{ij}(t|t) = \mathrm{E}[\tilde{X}_i(t|t)\tilde{X}_j^{\mathrm{T}}(t|t)]$, so we have the filtering error covariance matrix (149.27).

149.3.3 Information Fusion Filter Weighted by Matrices

Based on Lemma 3 we have the local filters $\hat{X}_i(t|t)$ and covariance matrix $P_i(t|t)$, from Theorem 1 we have the filtering error cross-covariance matrix $P_{ij}(t|t)$.

From (149.5) we have the local optimal filter and error covariance matrix as follows

$$\hat{x}_i(t|t) = [I_n\, 0 \ldots 0]\hat{X}_i(t|t) \tag{149.31}$$

$$p_{ij}(t|t) = [I_n\, 0 \ldots 0]P_{ij}(t|t)[I_n\, 0 \ldots 0]^{\mathrm{T}} \tag{149.32}$$

Theorem 2 For system (149.1)–(149.3) with Assumptions 1 and 2, we have the distributed weighted fusion optimal filter as follows [10]

$$\hat{x}_o(t|t) = \sum_{i=1}^{L} \mathrm{A}_i(t)\hat{x}_i(t|t) \tag{149.33}$$

the optimal weighted matrix $A_i(t)$, $i = 1, 2, \ldots, L$

$$[A_1(t)\, A_2(t) \ldots A_L(t)\,]^{\mathrm{T}} = \Sigma^{-1}(t|t)e\left(e^{-1}\Sigma^{-1}(t|t)e\right)^{-1} \tag{149.34}$$

where $e = [I_n \ \ldots \ I_n]^{\mathrm{T}}$ is $nL \times n$ dimension matrix, $\Sigma(t|t) = p_{ij}(t|t)$, $i, j = 1, 2, \ldots, L$ is $nL \times nL$ dimension matrix.

The corresponding variance of the weighted fusion optimal filter is computed by

$$P_o(t|t) = (e^{\mathrm{T}}\Sigma^{-1}(t|t)e)^{-1} \tag{149.35}$$

also we have $P_\circ(t|t) \leq P_i(t|t)$, $i = 1, 2, \ldots, L$.

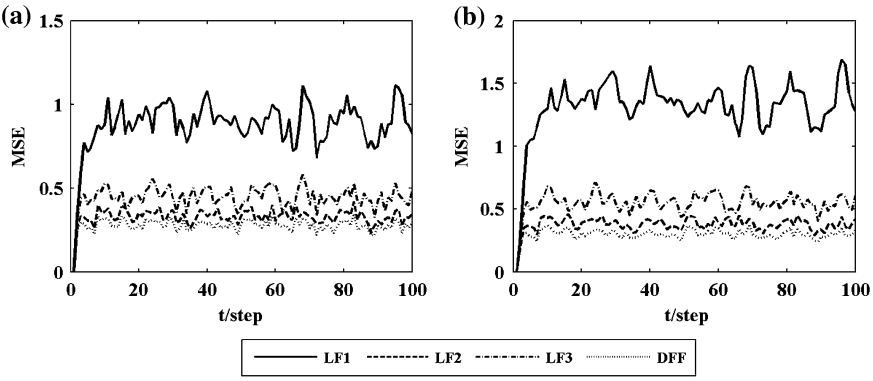


Fig. 149.2 Comparison of accuracy for local filters and distributed fusion filter. **a** MSE of the first state component. **b** MSE of the second state component

149.4 Simulation Research

Consider a stochastic systems with three sensor

$$x(t+1) = \begin{bmatrix} 0.8 & 0 \\ 0.9 & 0.2 \end{bmatrix} x(t) + \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix} \omega(t) \tag{149.36}$$

$$z_i(t) = H_i x(t) + v_i(t) \tag{149.37}$$

$$\begin{aligned} y_i(t) = {} & \xi_i(t) z_i(t) + (1 - \xi_i(t)) \xi_i(t-1) z_i(t-1) + (1 - \xi_i(t)) \\ & \times (1 - \xi_i(t-1)) z_i(t-2) \end{aligned} \tag{149.38}$$

where the measurement matrices $H_1 = [1 \quad 0]$, $H_2 = [1 \quad 1]$ and $H_3 = [1 \quad 2]$, $\omega(t) \in R^r$ is the input white noise with variance $Q_w = 1$, $v_i(t) \in R^{m_i}$, $i = 1, 2, 3$ are the sensor observation noises, and satisfy the relation $v_i(t) = c_i \omega(t) + \eta_i(t)$, where the white noises $\eta_i(t)$ with mean zero and variances $Q_{\eta_1} = 8$, $Q_{\eta_2} = 2$, $Q_{\eta_3} = 10$ are independent of $\omega(t)$, and the correlated coefficients $c_1 = 0.2$, $c_2 = 1$, $c_3 = 0.8$. $\xi_i(t)$, $i = 1, 2, 3$ are the Bernoulli distributed random variables with $\alpha_1 = 0.3$, $\alpha_2 = 0.6$, $\alpha_3 = 0.7$. Our aim is to find distributed matrix-weighted fusion optimal filter $\hat{x}_o(t|t)$ in the linear minimum variance sense.

In the simulation, we set the initial values $x(0) = [0 \quad 0]^{\mathrm{T}}$ and $P_0 = 0.1 I_2$. The performance of optimal fusion filter $\hat{x}_o(t|t)$ weighted by matrices is shown in Fig. 149.1, where the solid curves denote the real value and the dashed ones denote the fusion filter values. Comparison of the mean square error (MSE) for all local filters (LF) and distributed fusion filter (DFF) are shown in Fig. 149.2 by 200 times Monte-Carlo test. We see that the fusion filter has higher accuracy than any local filter does.

149.5 Conclusion

In this paper, the distributed fusion estimation problem for systems with multiple sensors of different finite consecutive packet dropout rates is investigated. The filtering error cross-covariance matrix is derived between any two local filters. The distributed matrix-weighted fusion state filter is obtained by applying distributed matrix-weighted fusion estimation algorithm in the linear minimum variance sense.

Acknowledgments This work was supported by Natural Science Foundation of China under Grant NSFC-60874062, Key Project of Chinese Ministry of Education under Grant 209038, Program for New Century Excellent Talents in University under Grant NCET-10-0133, 1154-NCET-01, Program for High-qualified Talents under Grant Hdtd2010-03.

References

1. Yaz YI, Yaz EE, Mohseni MJ (1998) LMI-based state estimation for some discrete-time stochastic models. In: Proceedings of the 1998 IEEE international conference on control applications, Trieste, pp 456–460
2. Hespanha JP, Naghshtabrizi P, Xu Y (2007) A survey of recent results in networked control systems. Proc IEEE Spec Issue Netw Control Syst 95:138–162
3. Wang Z, Yang F, Ho DWC, Liu X (2006) Robust H-infinity filtering for stochastic time delay systems with missing measurements. IEEE Trans Signal Process 54:2579–2587
4. Sahebsara M, Chen T, Shah SL (2007) Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations. Int J Control 80:292–301
5. Sun SL, Xie LH, Xiao WD, Soh YC (2008) Optimal linear estimation for systems with multiple packet dropouts. Automatica 44:1333–1342
6. Sun SL, Xie LH, Xiao WD, Soh YC (2008) Optimal full-order and reduced-order estimators for discrete-time systems with multiple packet dropouts. IEEE Trans Signal Process 56:4031–4038
7. Liang Y, Chen T, Pan Q (2010) Optimal linear state estimator with multiple packet dropouts. IEEE Trans Autom Control 55:1428–1433
8. Sun SL (2009) Optimal estimator for systems with finite consecutive packet dropouts. IEEE Signal Process Lett 16:557–560
9. Ma J, Sun SL (2011) Information fusion estimators for systems with multiple sensors of different packet dropout rates. Inf Fusion 12:213–222
10. Sun SL (2004) Multi-sensor optimal information fusion kalman filter with application. Aerosp Sci Technol 8:57–62

Chapter 150
Method of Obtaining Dynamic Stress Intensity Factor by Measuring Crack Mouth Opening Displacement on a 3-Point Bending Specimen

Yayu Huang, Xiangping Hu, Yujie Shen and Taohong Liao

Abstract This paper focuses on how to obtain the Dynamic Stress Intensity Factor by measuring Crack Mouth Opening Displacement on a pre-cracked 3-Point Bending Specimen. During the experiments, one uses impactor to hit on a pre-crack 3-Point Bending Specimen and at the same time uses a high-speed camera to take a set of photos of the crack. Crack Mouth Opening Displacement can then be measured from these photos. According to the relationship of J integral, Crack Open Mouth Displacement and the Dynamic Stress Intensity Factor, the Dynamic Stress Intensity Factor could be obtained. This result could be useful for guiding the practical engineering design.

Keywords 3-point bending specimen · Crack mouth opening displacement · J integral · Dynamic stress intensity

Y. Huang · X. Hu · Y. Shen · T. Liao
Mechanical and Electrical Engineering,
Kunming University of Science and Technology,
Kunming, China
e-mail: kmhuangyayu@yahoo.com.cn

X. Hu (✉) · Y. Shen
Department of Mechanical Engineering,
Blekinge Institute of Technology,
Karlskrona, Sweden
e-mail: kmhuxp@gmail.com

Y. Shen
e-mail: fpspaul@hotmail.com

T. Liao
e-mail: liao.taohong@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_150,

150.1 Introduction

It is known that there are always some defects in the mechanical parts, such as different kinds of cracks. There are typically three ways of applying a force to enable a crack to propagate and so the crack propagation is divided to three types, i.e. mode I, II and III, or the so called opening mode, sliding mode and tearing mode [1, 2]. Cracks are quite common in reality and crucial for structure fracture, fatigue and failure which will influence structure's life. Dynamic Stress Intensity Factor (DSIF) K is an important fracture mechanics factor to predict the stress state close to the crack tip and it could be theoretically used to predict a criterion for failure of homogeneous elastic materials [1, 2]. The DSIF K_I, for the opening mode, is common studied and used in many literatures [3–5]. Huang Y. and Hu X. did both theoretical analysis and finite element analysis for the DSIF based on the pre-cracked 3-Point Bending Specimen (3PBS) [3]. Huang Y. et al. then extended the theoretical analysis for the DSIF based on the more complicated equivalent system for the pre-cracked 3PBS [4]. For a standard experiment, there is one recommended structure by Dahlberg T. and Ekberg A. [1]. Jiang F. et al. have chosen this recommended structure to do both theoretical study and experiments analysis for the DSIF [5].

It is difficult to obtain the DSIF K_I experimentally. However, the Crack Mouth Opening Displacement (CMOD) of the crack is easier to be obtained through the experiments, so one could obtain the DSIF K_I by measuring the CMOD through the relationships between them and then use K_I to determine the stress state of the crack and predict the life time of the mechanical parts. In fact, in the relationship between CMOD and DSIF, another important factor is also needed, i.e. J integral. The J integral is a direct strict fracture mechanics factor for elastic–plastic stress and strain fields of the crack tip. It can be used to describe the strength of crack tip stress–strain field. Fortunately, it could be also obtained from the experiments. For more information about the J integral, see for example [1, 2].

This paper emphasizes on obtaining the DSIF by measuring CMOD on a pre-cracked 3-Point Bending Specimen, by using high-speed camera to take a set of photos of the crack to measure the CMOD. The structure of paper for the following sections is organized as follows: Sect. 150.2 is used to introduce method of measuring the CMOD by using the high-speed camera. Section 150.3 explains how to use the CMOD and J integral to obtain the crucial fracture mechanics factor DSIF. Section 150.4 is the general discussion and future research to end this paper.

150.2 Method of Measuring the CMOD

CMOD is a parameter for measuring the displacement at the crack tip when the specimen suffering load. There are some methods for measuring the CMOD, the first one is by using the microscope to observe and measure, or using the

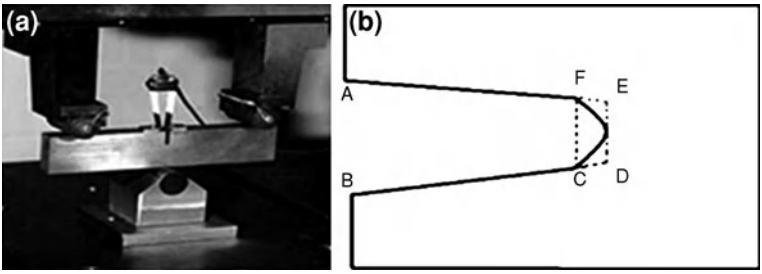


Fig. 150.1 CMOD measurement (**a**) [6] and crack tip of the 3PBS (**b**)

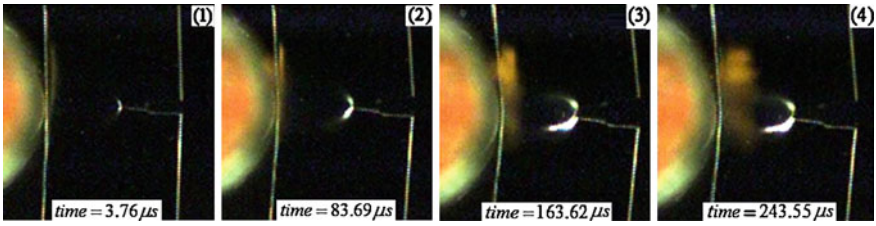


Fig. 150.2 A series of experiment photos collection

high-speed camera to capture the photos. The other one is to use a clip gauge either attached to knife edges mounted at the crack mouth or integral knife edges machined into the notch [6], as shown in Fig. 150.1a. These gauges comprise two cantilever beams on which placed four stain gauges. By measuring the elastic strains and then calibration, the CMOD could be obtained experimentally. From the analysis of the relationship between the applied force and the CMOD form the experiments, it enables fracture toughness to be determined in terms of the DSIF K_I, CMOD and J integral. There are also some similar factors, such as Crack Opening Displacement (COD) and Crack Tip Opening Displacement (CTOD) but will not be included in this paper.

One should notice that there is no scientifically standard definition for how to measure the CMOD and CTOD, i.e., some researchers suggest that the distance CF in the crack tip, as shown in Fig. 150.1b, could be used as CMOD, but others insist that the distance DE is exactly the CMOD.

The high-speed camera is used in this paper to capture a series of photos of the crack tip of the 3PBS. Some of the photos are shown in Fig. 150.2. These photos give the information which could be used for measuring the CMOD since the dimensions of the pre-cracked specimen are measured before the experiments.

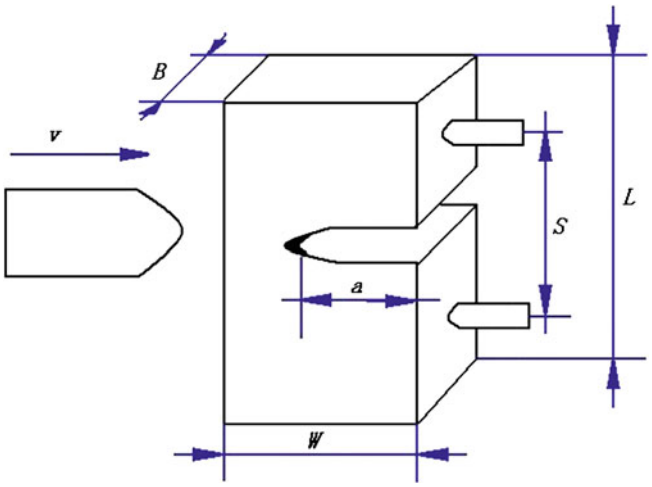


Fig. 150.3 Testing model of the pre-cracked 3PBS

Table 150.1 Parameter values of the testing model [mm]

Parameters	L	S	W	B	a
Values	40.52	26	6.51	3.99	W/2

The geometry structure of the pre-cracked 3PBS is given in Fig. 150.3 and the parameters are given in Table 150.1. L is the total length, S is the length between the two support points, W is the width, B is the thickness and a the length of the crack.

150.3 DSIF Using CMOD and J Integral

According to literature [7], there is one relationship between the J integral and the CMOD, and this relationship is defined as a factor called $\beta(v)$. So it is possible to obtain this factor $\beta(v)$ from the model and then the relationship between CMOD and DSIF K_I could be obtained. The definition of the factor $\beta(v)$ is given as

$$\beta(v) = \frac{\sum_{i=1}^{k} J_i(v)\delta_M(v)}{\sigma_Y \sum_{i=1}^{k} \delta_M(v)^2}. \tag{150.1}$$

From the series of the photos from the high-speed camera, one could obtain the CMOD in each picture by pix as unit with the software Phantom and the result is shown in Fig. 150.4.

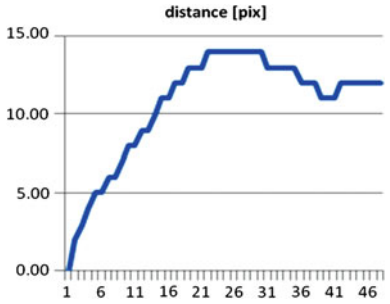


Fig. 150.4 Displacement measured from picture collection

On the other hand, since the width of the 3PBS is known, it is possible to obtain the length of one pix in the photo, i.e., $W = 6.51[\text{mm}] = 82[\text{pix}]$ results that $1[\text{pix}] = 0.0794[\text{mm}]$.

From the simple relations (150.2) and (150.3),

$$J = \frac{K_I^2}{E}, \tag{150.2}$$

$$J = \beta(v) \times \sigma_Y \times \text{CMOD}, \tag{150.3}$$

one could get the DSIF as

$$K_I = \sqrt{E \times \beta(v) \times \sigma_Y \times \text{CMOD}}, \tag{150.4}$$

where, β is determined by the structures and the properties of the materials, E is the Young's modulus, and σ_Y is the yielding stress. From the known data from the photos, it is easy to obtain the factor $\beta(v)$ from Matlab as

$$\beta(v) = 4.3863. \tag{150.5}$$

So from (150.4), it will be possible to obtain the DSIF K_I by just measuring the CMOD of the given model. The values of CMODs in the experiments are shown in Fig. 150.5a. The yielding stress of the material in the experiment is $\sigma_Y = 373[\text{MPa}]$, and substituting it, together with (150.5), into (150.4) gives the value of DSIF K_I

$$K_I \doteq 130\left[\text{MPam}^{1/2}\right].$$

In order to verify the result, the values of the J integral directly from the ABAQUS and from (150.3) are compared. The results are shown in Fig. 150.5b. One could notice that these two curves fit each rather well and the error between these two may be from the measurement errors for CMOD and the error from converting the pix to CMOD. So this indicates that it will be possible to get the DSIF K_I by measuring the CMOD, when the J integral is difficult to obtain through other approaches.

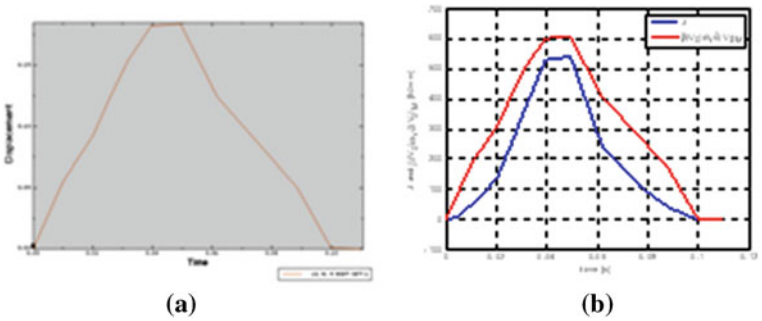


Fig. 150.5 The values of CMOD (**a**) and the *J* integral for comparison (**b**)

150.4 Discussion and Future Research

This paper has discussed one approach to get the DSIF K_I through the CMOD on a pre-cracked 3PBS experimentally. A high-speed camera is used to take a set of photos for measuring the CMOD. The formula for DSIF K_I is also simply derived and used for calculation. The result shows that it is trustful to get the DSIF K_I by measuring the CMOD when the other approaches are out of reach. This result is useful for guiding the future research of the DSIF and also useful for practical engineering design of the ballast tamping machine in the China Railway Large Maintenance Machinery CO. LTD., Kunming.

Future research is under processing, such more materials are under testing and different types of cracks are chosen for both theoretical study and experiments.

Acknowledgments Author Hu X. expresses his gratitude to Dr. Sharon Kao-Walter from the Blekinge Institute of Technology, Sweden, for her useful support during the whole research. Author Hu X., Shen Y. and Liao T. appreciate to PhD student Leon A. for the ABAQUS simulations. Author Huang Y. wants to expresses his gratitude to the China Railway Large Maintenance Machinery CO. LTD., Kunming, for supporting the research.

References

1. Dahlberg T, Ekberg A (2002) Failure fracture fatigue: an introduction. Student litteratur, Sweden
2. Anderson TL (2005) Fracture mechanics, fundamentals and applications. CRC press, Taylor and Francis Group, USA
3. Huang Y, Hu X (2011) Fracture mechanics theoretical modelling and fem analysis of 3-point bending specimen suffering cyclic loads. Proceedings of the 9th International Conference on Reliability, Maintainability and Safety pp 308–312
4. Huang Y, Hu X, Liao T (2011) Fracture mechanics mathematical modelling for dynamic stress intensity factor of 3-point bending specimen, accept for Proceedings of the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce

5. Jiang F, Rohatgi A, Vecchio KS, Cheney JL (2004) Analysis of the dynamic responses for a pre-cracked three-point bend specimen. Int J Fract 127:147–165
6. Pisarski H Fracture toughness testing, http://www.twi.co.uk/content/kscsw011.html
7. Kao HR (1991) Correlation between J-integral and CMOD of an impact loaded 3-point bend specimen, technical report. Lund Institute of Technology, Sweden

Chapter 151
Fault Diagnosis of ZPW-2000A System Based on PSO-GA-BP Algorithm

Bo Fan, Shiwu Yang, Xingmin Wang and Yuhang Wang

Abstract ZPW-2000A joint-less frequency shift automatic block system has a crucial and realistic significance in maintaining the safety of railway transportation. The traditional fault diagnosis for ZPW-2000A is inefficient and troublesome. A new fault diagnosis method was proposed based on BP neural network, Particle Swarm Optimization Algorithm and Genetic Algorithm (PSO-GA-BP). In order to compare with the diagnostic performance of PSO-GA-BP, some other hybrid algorithms were also simulated. It came to the conclusion that PSO-GA-BP algorithm was more effective in convergent rate, convergent accuracy and diagnostic accuracy according to the simulation results. The PSO-GA-BP neural network is efficient, convenient and provides a new approach to fault diagnosis.

Keywords Fault diagnosis · ZPW-2000A · Neural network · PSO · GA · BP

B. Fan (✉) · S. Yang (✉) · X. Wang (✉)
Beijing Jiaotong University,
100044 Beijing, China
e-mail: 09120282@bjtu.edu.cn

S. Yang
e-mail: ysw@bjtu.edu.cn

X. Wang
e-mail: 09120308@bjtu.edu.cn

Y. Wang (✉)
China University of petroleum Beijing,
102200 Beijing, China
e-mail: Wangyuhang.11@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_151,

151.1 Introduction

The signal equipment in a railway station is the main facility used to ensure railway safety which is an important goal that rail transport pursues; therefore to ensure the trouble-free working of the signal equipment is of great significance.

The ZPW-2000A system is becoming more and more popular as an occlusion device in many railway stations because of its high reliability meeting the requirements of high security.

How to diagnose the faults of ZPW-2000A system quickly and effectively is an important problem before technical staff. The matching conflict, combinatorial explosion and infinite recursion of traditional expert systems make the expert system work slowly and inefficiently. What is more, as to the new faults, the traditional expert systems can do nothing. BP neural network is a new network which has self-learning ability and memory functions [1, 2]. It is the most maturely and widely used in fault diagnosis. But the inherent problem of BPs easy to fall into local minimum makes the network have low accuracy [3, 4]. If particle swarm optimization algorithm and BP neural network work together so that the weights and thresholds of BP neural network are optimized by PSO algorithm [5], the defects in neural networks can be better overcome. In addition, the neural network not only improves the convergence rate, but also enhances the generalization ability. However, PSO algorithm [6] is easy to fall into local minima in the search. Then if we combine genetic algorithm which has selection operator, crossover operator, and mutation operator with PSO algorithm, the PSO can easily jump out of local optimal value to the global optimum with the help of GA.

Based on the idea above, a new fault diagnosis network will be designed based on PSO-GA-BP algorithm in this paper. In order to compare with the diagnostic performance of PSO-GA-BP, GA-BP (algorithm mixed GA algorithm with BP algorithm), BP algorithm and PSO-BP (algorithm mixed PSO algorithm with BP algorithm) are also simulated. Simulation result shows that the network has a better efficiency and accuracy in fault diagnosis.

151.2 Introduction to PSO-GA-BP Algorithm

151.2.1 The Principle of BP Neural Network

BP network is a multilayer feed-forward network. BP network learning process consists of two parts: forward propagation of information and back propagation of error. Input information is processed by the hidden layer and the output layer which will output the results to the outside world. If the actual output does not match the expected output, then transfer to the back-propagation. The error signal goes back along the original path. The weights and thresholds of every layer of neurons will be modified one by one in this process. This iterative process

continues until the network output error is reduced to an acceptable level, or the preset number of times comes to the end.

151.2.2 The Particle Swarm Optimization Algorithm

Suppose that the state of particle j at time t in n-dimensional search space is expressed in formula (151.1) and formula (151.2).

$$X_j(t) = \left[x_{j1}(t), x_{j2}(t), \ldots, x_{jn}(t)\right] \tag{151.1}$$

$$V_j(t) = \left[v_{j1}(t), v_{j2}(t), \ldots, v_{jn}(t)\right] \tag{151.2}$$

In the expression above, $X_j(t)$ and $V_j(t)$ indicate the position and velocity vectors of particle$j$at time$t$. The status of particle$j$in the n-dimensional search space is updated according to formula (151.3) and formula (151.4).

$$x_{jk}(t+1) = x_{jk}(t) + v_{jk}(t+1) \tag{151.3}$$

$$v_{jk}(t+1) = \omega * v_{jk}(t) + c_1 * r_1\left(p_{jk} - x_{jk}(t)\right) + c_2 * r_2\left(p_{gk} - x_{jk}(t)\right) \tag{151.4}$$

In the expression above, p_{jk} is the best position of particle$j$ in the history in the k-dimensional search space. And p_{gk} is the best position of all particles in the k-dimensional search space.

151.2.3 The Design of PSO-GA-BP Algorithm

The number of each PSO particle dimension is equal to the number of weights and thresholds of BP neural network. Each dimension corresponds to a weight. The fitness function of PSO is the network's total error. In the process of calculation, if the average fitness value of the particle swarm keep unchanging during certain generations at the same time the required precision is not achieved, Then the diversity of particle populations is lost in other words PSO algorithm falls into a local optimum. At this time Genetic algorithm can be operated in order to help PSO algorithm jump out of local optimal value, so that the PSO algorithm can continue. Finally the optimal location of particle swarm which is our need will be the output.

151.2.4 The Steps of PSO-GA-BP Algorithm

(1) Decide the network err intended to reach, maximum permissible number of iterative steps, reach range, maximum speed and the number of particles according to the network size. And pbest is the best position of particle in the history. Then gbest is the best position of particle swarm.

(2) Initialize particle swarm and set the initial position and velocity of the particle.
(3) Calculate the fitness of each particle and update pbest and gbest according to the particle's fitness value
(4) Determine whether the diversity of species is lost. If the answer is yes, turn to step 7.
(5) Each particle updates its position and velocity according to its own pbest and gbest of the particle swarm.
(6) Check whether the end conditions of the maximum iterative generation is arrived or whether the solution is in the range of error allowed are satisfied. If the answer is yes, then the algorithm is end. If the answer is no, then turn to step 3.
(7) We can see each particle in the particle swarm as a single genetic individual. Then GA is operated for several generations. Calculate each individual's fitness value. If gbest is better than that in PSO, assign each genetic individual in GA to each individual particle in PSO. Then turn to step 3. Otherwise, check whether the number of rounds is equal to the given value. If no, the next round of genetic algorithm will be operated. If yes, output the best particle. PSO-GA-BP algorithm is end.

The flowcharts are shown in Fig. 151.1.

151.3 The Construction of ZPW-2000A Fault Diagnosis Network

151.3.1 The Selection of Variables

There are two criteria that must be given attention. First, the selected inputs which have to be detected and extracted must have a great impact on the outputs. The correlation of selected inputs is unrelated or very small. In this paper the inputs of the network are as follows.

The voltage of main track, the voltage of Small rail, the voltage of track, the empty-voltage of attenuator "XGJ" and the voltage of analog disk.

151.3.2 Data Normalized Processing

The selected inputs often have different physical meaning and different dimension, it is necessary to normalize the data samples. The normalized equation is shown in formula (151.5).

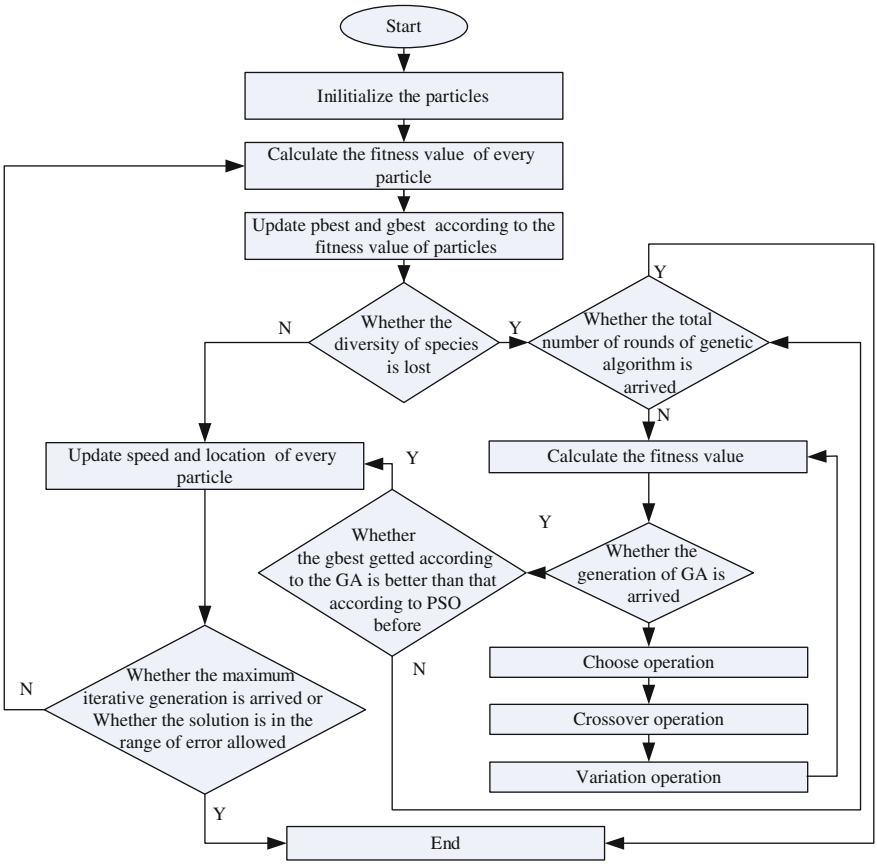


Fig. 151.1 PSO-GA-BP algorithm flow chart

$$T = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{151.5}$$

In the expression above, X is the sample data. And $X_{\max}$ is the max value among the sample data. Then $X_{\min}$ is the min value among the sample data. In this paper, we selected historical data in a railway station in Anhui Province as the sample data which will be used to study and train the PSO-GA-BP network. A total of 3000 groups of sample data are selected in this paper. Seven groups of normalized sample data are shown in Table 151.1.

151.3.3 The Number of Neurons Nodes in Hidden-Layer

The BP neural network which is selected in this paper has three layers with one single hidden layer. The number of neurons nodes in hidden-layer determines the

Table 151.1 Seven groups of normalized sample data

Serial number	Main track voltage	Small rail voltage	Track voltage	Empty-voltage of attenuator "XGJ"	Analog disk voltage	Fault type
1	0.932383	0.847063	0.756017	0.825358	0.505142	Shared channel failure
2	0.824982	0.481765	0.69498	0.713428	0.541679	Main rail fault
3	0.690057	0.828538	0.679592	0.86765	0.582315	Small rail fault
4	0.570296	0.488839	0.770244	0.208252	0.536518	Small rail fault
5	0.550893	0.432521	0.774261	0.85941	0.566057	Indoor circuit fault
6	0.550978	0.45201	0.191171	0.879296	0.588794	Attenuation box failure
7	0.564119	0.479943	0.718021	0.763961	0.552113	Indoor circuit fault

performance of the neural network. We can determine the number in this way. First, we get the approximate range of the number through the experiential rule $m = \sqrt{n + l} + a$. Second, try each number in the range to obtain the best one. After repeated experiments the number of neurons nodes in hidden-layer is 20. Therefore the construction of the BP neural network is 5-20-5.

151.3.4 PSO-GA-BP Algorithm Parameters Setting

The population size is 60 in PSO-GA-BP. The value of particle's position is random value ranging from 0 to 1. The value of particle's velocity is random value ranging from –1 to 1. After many tests, the Maximum velocity is 0.55. Inertia weight factor w decreases linearly from 0.9 to 0.4 in the iteration process. The coefficients c_1=c_2=2. The iterative generation is 3000 in PSO-GA-BP. The GA crossover probability p_c is 0.88. Mutation probability p_m is 0.08. The number of rounds is 5. The number of generation of every round is 100.

151.4 Simulation Results and Analysis

While the first 2700 groups of sample data are used as training samples, the remaining 300 groups are used as test samples. In order to compare with PSO-GA-BP algorithm, GA-BP algorithm, BP algorithm and PSO-BP algorithm are also simulated.

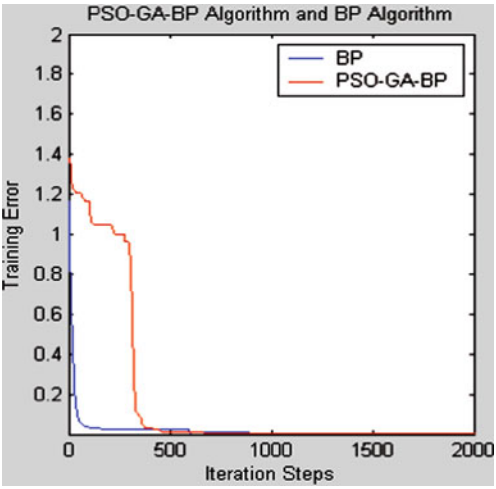


Fig. 151.2 Comparison of convergence rate of PSO-GA-BP algorithm and BP algorithm

151.4.1 Comparison of Convergence Rate of PSO-GA-BP Algorithm and BP Algorithm

Error evolution curves of PSO-GA-BP algorithm and BP algorithm are shown in Fig. 151.2 at the condition that the training error precision of PSO-GA-BP algorithm is 0.00001, while that of BP algorithm is 0.0001.

From Fig. 151.2 we can see that the convergence rate of PSO-GA-BP algorithm is slower than that of BP algorithm at the very beginning. But with the Iteration steps increasing, convergence rate of PSO-GA-BP algorithm is much faster. In order to reach the training error precision, the PSO-GA-BP algorithm needs only 440 step iterations. Its error precision is 0.0000001 at 770th step iterations. As to BP algorithm, it needs 2000 step iterations to reach its training error precision. So it comes to the conclusion that PSO-GA-BP algorithm is better than BP algorithm in fault diagnosis in this paper.

151.4.2 Comparison of Convergence Rate of PSO-GA-BP Algorithm and PSO-BP Algorithm

Error evolution curves of PSO-GA-BP algorithm and PSO-BP algorithm are shown in Fig. 151.3 at the condition that the training error precision of the two algorithms is 0.00001.

From Fig. 151.3 we can see that the convergence rate of PSO-BP algorithm is faster at the very beginning. But with the Iteration steps increasing more and more flat areas appear that PSO often falls into local minima so that the convergence rate is slowing and the network accuracy is lower. On the contrary, such a problem

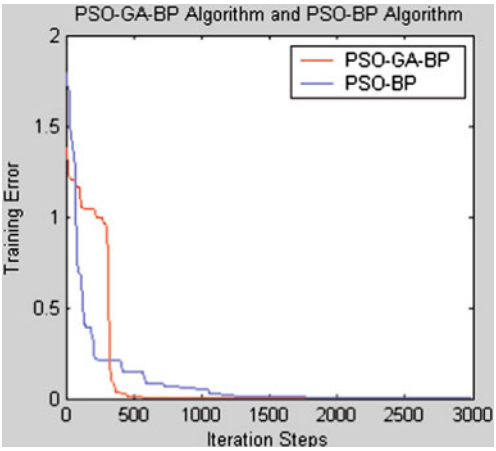


Fig. 151.3 Comparison of convergence rate of PSO-GA-BP algorithm and PSO-BP algorithm

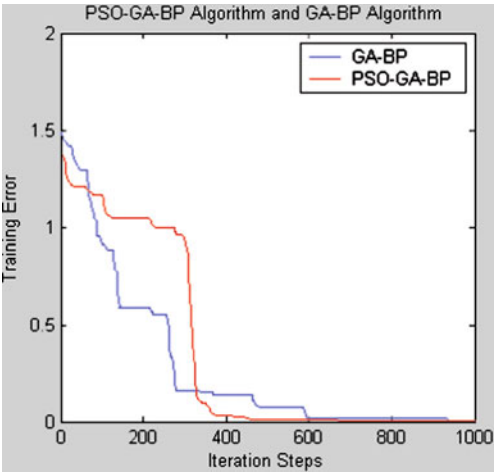


Fig. 151.4 Comparison of convergence rate of PSO-GA-BP algorithm and GA-BP algorithm

does not come to PSO-GA-BP algorithm. So it comes to the conclusion that PSO-GA-BP algorithm is better than PSO-BP algorithm in fault diagnosis in this paper.

151.4.3 Comparison of Convergence Rate of PSO-GA-BP Algorithm and GA-BP Algorithm

Error evolution curves of PSO-GA-BP algorithm and GA-BP algorithm are shown in Fig. 151.4 at the condition that the training error precision of the two algorithms is 0.00001.

Because operations in GA-BP are more complex operations in PSO-GA-BP. Convergence rate and network efficiency of PSO-GA-BP algorithm are much

Table 151.2 Comparison of performance of the four algorithms

Algorithm	Time (s)	Algorithm accuracy	Testing accuracy (%)
BP	2523.45	0.0001	75
GA-BP	536.56	0.00001	90.3
PSO-BP	745.70	0.00001	87.6
PSO-GA-BP	350.2	0.00001	94.9

better. So it comes to the conclusion that PSO-GA-BP algorithm is better than GA-BP algorithm in fault diagnosis in this paper.

151.4.4 Comparison of Performance of the Four Algorithms

As for the four algorithms, time used to complete the training samples, the algorithm accuracy and the accuracy of testing the remaining 300 groups of samples are shown in Table 152.2.

151.5 Conclusions

Using this model, the actual diagnosis is processed to some typical faults. Results showed that PSO-GA-BP network which is more effective in convergent rate, convergent accuracy and diagnostic accuracy is efficient, convenient and provides a new approach to fault diagnosis of ZPW-2000A system.

References

1. Shi Yong, Zhao Yao (2010) The forecast of logistics cost of coal enterprises based on PSO-BP. J Liaoning Tech Univ 29(5):945–948
2. Wang Han, Liu Yu (2007) Back propagation neural network training based on particle swarm optimization with best influential particle. Comput Eng App 43(1):69–71
3. Zhao Xiao, Zhou Yin (2009) Hybrid optimization algorithm maintaining independence of PSO and GA. Comput Eng App 45(1):53–55
4. Yi Zhao HC, Zhang (2005) An improved rapid algorithm for BP network. J Beijing Inst Mach 1(3):31–34
5. Wang Na, Chen De (2004) An Improved BP Neural Network and its Application in Fault Diagnosis. Comput Eng App 5(9):20–22
6. Fan Long, Wang Yu (2011) On Applying an Improved PSO Algorithm in Grid Intrusion Detection System. Comput App Softw 3(9):274–278

Chapter 152
Mine Workface Gas Emission Time-Series Fractal Properties

Meiying Qiao, Jianjun Qiao, Hui Tao and Xiaoping Ma

Abstract Mine workface gas emission time series has a complex nonlinear characteristics, R/S method is presented to analyze their fractal characteristics in this paper. First, R/S basic principle is introduced and long-range correlation properties are analyzed; at the same time the effective correlation length is calculated. Second, the algorithm program is implemented under MATLAB200b environment, and two noise sequences are used to verify it. Finally, three gas outburst time-series data and two non-outburst time-series data of Hebi 10th mine 1,113 workface are analyzed using above program. The results show that: they Hurst index values are greater than 0.5, whether exist gas outbursts or not, which show workface gas emission time series has positive correlation feature, and this feature of non-outburst's is obviously stronger than outburst's. Furthermore the effective correlation length of non-outburst gas time series is also significantly longer than outburst occurs.

Keywords Mine face · Gas emission · Time series

M. Qiao (✉) · H. Tao
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: qiaomy@foxmail.com

H. Tao
e-mail: taohui@hpu.edu.cn

M. Qiao · X. Ma
School of Information and Electrical Engineering, China University Mining and Technology, Xuzhou 221008, China
e-mail: xpma@cumt.edu.cn

J. Qiao
JiaoZuo WanFang Power Manufacturing Co. Ltd, Jiaozuo 454003, China
e-mail: 149178640@qq.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_152,

152.1 Introduction

The coal and gas outburst is a complex nonlinear dynamics action, and its formation conditions, predisposing factors and evolution possess diversity, complexity and randomness, etc., so the dynamic behavior of the time-series data of gas emission in the working faces shows unity of uncertainty and randomness. Chaotic dynamics which range between deterministic and stochastic is introduced into the face of gas emission prediction, and it has become a hot research [1].

Shi Shiliang first verified gas emission in the working face possessing chaotic characteristics by using the power spectrum and Poincare section, however, the methods used are classified qualitative research and it possessed certain limitations [2]. Wang Kai introduced the quantitative method such as G-P algorithm and Laypunov Index to analyze the time series of gas emission (abbreviated Q series), but he studied little about the time series before the gas outburst [3]. He Liwen analyzed more in-depth on the time series on the basis of the previous studies [4]. However, studying the chaotic dynamics conduct of gas emission in the working face can provide a theoretical basis for its short-term predictability, and the existing prediction methods cannot meet the long-term forecast. R/S analysis (Rescaled Range Analysis) is a method to study time series which is basis of fractal theory. The method is particularly suitable to analyze nonlinear time series with long-range correlation properties, but also it need not be made any assumptions. It is classified as non-parameter analysis with good stability [5]. Currently, R/S analysis method has been used in many areas to do long-term trend analysis: Huang Cunhan, forecasted the mine discharge using it [6]; Selvaratnam adopted the method to predict the time series for the stock market [7]. Li Yexue used it to analyze rock deformation trend [8]; Li Yuanyao predicted landslides deformation using the method [5].Li Ming, adopted the method to study the trend of traffic [9]; Wong Wenjing had studied long memory properties of financial time use method [10].

The R/S analysis method has been introduced [4, 11], to coal working face for analyzing the gas emission of the fractional Brownian motion in long-term trend based on the existing achievement in this article. Algorithm step has been given and been achieved by MALTAB 2009b, while it has been verified through using two noise sequences based on R/S analysis algorithms principle in detail. Gas emission time series is analyzed such as 1113 working face of the Hebi10th mine. The Hurst value of time series which had taken place three times outburst and the value Hurst exponent of two times not outburst time series are respectively calculated, and the Hurst value of several time series been studied and time limits associated with long-range are comparatively analyzed.

152.2 R/S Analysis Basic Principle

152.2.1 Definition and Solving of Hurst Exponent

R/S analysis method is used to determine the timing signal of fractal characteristics and long-range correlation, then to distinguish between time series of random and non-random by calculating the Hurst exponent H value of nonlinear time series in order to determine the continuing trend and strength. Its basic principle is as follows

Set the length N time series data $\{x(1), x(2), \ldots, x(N)\}$. It is divided into intervals with adjacent and non-overlapping $M_a(a = 1, 2, \ldots, A)$, their length be $n(2 \leq n < L)$. In here, L is the largest sub-interval length. From the above analysis, their relationship can be gained: $A \cdot n = N$. If each element of M_a is expressed $m_{k,a}(k = 1, 2, \ldots, n)$, mathematical expectation of each subinterval will be get:

$$X_a = \frac{1}{n}\sum_{k=1}^{n} m_{k,a} \tag{152.1}$$

Deviation from the mean difference of each sub-sequence of sequence can be calculated by formula: $\Delta X_{k,a} = m_{k,a} - X_a$. Clearly, the mean of $\Delta X_{k,a}$ is zero, this step is re-calibrated or normalized (standardized).

Therefore, the cumulative mean deviation $Y_{k,a}$, standard deviation S_{Ma} and the range R_{Ma} can be expressed as follows:

$$Y_{k,a} = \sum_{i=1}^{k} \Delta X_{k,a}, (k = 1, 2, \ldots, n) \tag{152.2}$$

$$S_{Ma} = \sqrt{\frac{1}{n}\sum_{k=1}^{n} \Delta X_{k,a}} \tag{152.3}$$

$$R_{Ma} = \max\left(Y_{k,a}\right) - \min\left(Y_{k,a}\right), (k = 1, 2, \ldots, n) \tag{152.4}$$

Calculating rescaling range value of each sub-interval R_{Mz}/S_{Ma} and the average range rescaling R_n/S_n of a interval, it can get the formula as follows

$$R_n/S_n = \frac{1}{A}\sum_{a=1}^{A} (R_{Ma}/S_{Ma}) \tag{152.5}$$

$R_n/S_n \propto n^H$ are proved as power-law relationship by Mandelbrot. So $\lg(R_n/S_n)$and $\lg n$ can be expressed with linear relationship.

$$\lg(R_n/S_n) = \lg C + H \lg n \tag{152.6}$$

Therefore, in the course of the actual solving, point$(n, R_n/S_n)$ is drawn in double logarithmic coordinates, and regression analysis is carried out. If the coordinate points showed a good linear relationship in the double logarithmic coordinates, the slope rate of the line is the value H that is, Hurst exponent [12].

152.2.2 Hurst Exponent Analysis to Determine the Long-Range Correlation

The results show that if the value H is close to 0.5, this time series can be described with the random walk, such as using computer to generate random time-series data, or to disrupt the time series of the Nile flow, then do R/S analysis, obtained Hurst exponent values is also close to 0.5. $0.5 < H \leq 1$indicates state is continuous, and it implies that the time series is a persistent or trend-enhanced sequence. If the sequence of the previous period is up (down), then the next period will be more likely to be positive (negative). H more close to 1, the correlation is stronger, at same time, it is more close to 0.5, the noise is greater, and the trend more uncertain. On the contrary, $0 \leq H < 0.5$ is a state against the continuing, the time series is anti-persistent. If the sequence of the previous period is up (down), then the next period will be more likely to be down (up). Anti-persistence will be gradually strengthened as H close to zero. In short, as long as $H \neq 0.5$, the time-series data can be described with the fractal Brownian motion [13].

Many time series are non-normal probability distribution in the real world, and their distribution follows the inverse power law, that the mean peak at the higher ministry and more fat tail is typical fractal distribution. Fractal dimension α is used to express the peak level and the amount of fat tail, its relationship with the Hurst exponent is $\alpha = 1/H$. Therefore, according to the above characteristics, it is very important to solve a time-series Hurst exponent value. It can understand time-series gas emission trends through researching Hurst exponent [13].

152.2.3 Time-Series Long-Range Correlation Limits Analysis

Most long-range correlation of complex nonlinear time-series possesses time-scale boundaries, if excess of this time scale, and it will show an unrelated random behavior, then self-similar characteristics will be lost. Therefore, the study of this time scale is also a very meaningful work, T_m is expressed in this time scale. According to the above analysis in double logarithmic coordinates to map out the plot$(n, R_n/S_n)$, if there is a significant turning point, it indicates that logarithm of the critical time scale appear in here. Then long-range relevant limits of effective time series can be obtained through the corresponding conversion.

Table 152.1 1,113 face three gas emission instances

Time	Place	Elevation (m)	Vertical Depth (m)	Coal emission Quantity (t)	Gas emission Quantity (m3)
2003.03.19	1,113 bottom panel	−552.9	713	22	2523
2003.03.30	1,113 main gate	−530	678	2	240
2003.04.10	1,113 main gate	−530	690	57	1,1174

152.3 Algorithm Steps and the Case Study

152.3.1 Algorithm Steps of Hurst Exponent

First, according to the basic principles of Hurst exponent, Hurst value is calculated using programming under the MATLAB2009b environment, the specific algorithm steps are the following.

The time-series data are imported into the workspace, the average value is gained by using mean function, and the variance sequence are calculated, accumulated sum vector is obtained by using cumsum function. Finally, the average of the ratio of sample range and standard deviation are obtained. In this step the entire time-series sequence is sub-sequence.

The sequence is divided into two sub-sequences; the sample range R_n/S_n is again obtained.

Whole time series are divided into three, four,…, and the above calculation steps are repeated. The calculation is stopped when the dimensions of sub-sequences are less than a given threshold. In this article this threshold is 15.

The sub-sequence dimensions and the corresponding sample range R_n/S_n are taken logarithm, respectively. In double logarithmic coordinates, the least-squares fitting function ployfit is used to fit these points. The fitted slope value is Hurst value.

Draw the fitting curve, and find a clear turning point at the abscissa value. Convert this value to long-range correlation limits.

152.3.2 Mining Face Gas Outburst Instance

In this paper, the Hebi 10th mine 1,113 blasting mine work face time-series data is used as the object. By the end of April 2003, there have occurred five times mine and gas outburst in Hebi 10th mine, and three instances have occurred in the 1,113 face, so it is very representative that this work face is selected as object. The time-series data are the gas emission concentration percentage. The face gas emission data are measured by using KJ93 environmental monitoring system. The system

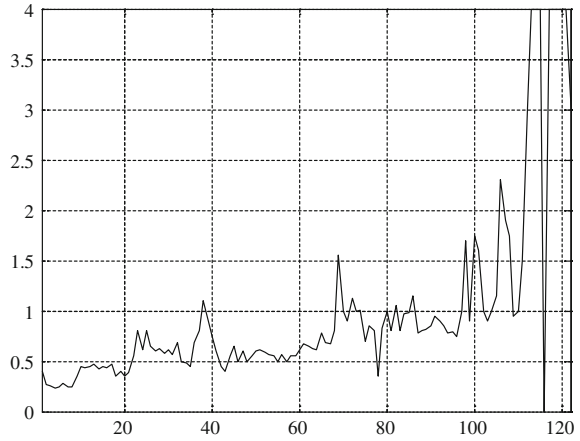


Fig. 152.1 2003.3.19 gas emission quantity

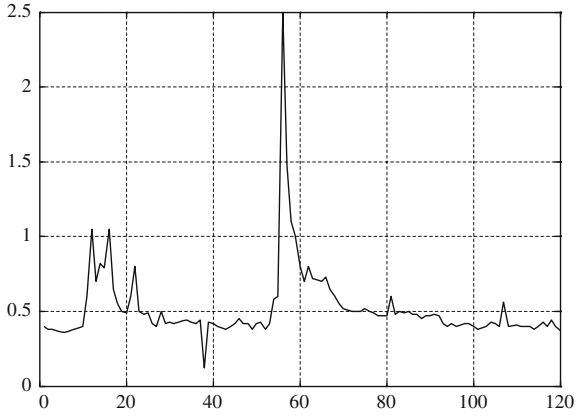


Fig. 152.2 2003.3.30 gas emission quantity

can display the amplitude of the gas concentration range of 0–4%. [11]. There are 120 time-series data. Each interval between two points is twelve minutes. Table 152.1 shows three mine and gas emissions statistical table.

Three times gas emission time-series curve is shown in Figs. 152.1, 152.2, 152.3. Starting point is zero o 'clock, end point is 24 o 'clock.

152.3.3 Hurst Exponent Analysis of Time Series

Based on the above algorithm steps, the programming is achieved by using MATLAB2009b. According to Hurst exponent and long-term time-series correlation analysis, the Hurst exponent value of white noise time-series close to 0.5, while the colored noise Hurst index less than 0.5. First, the program is verified by

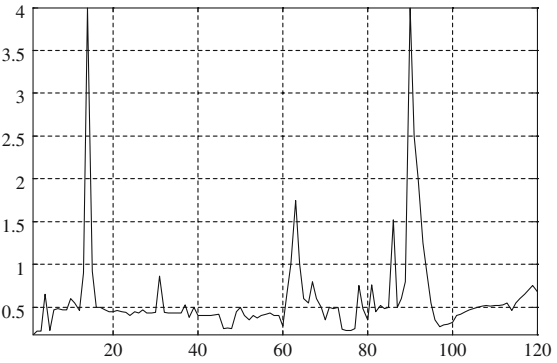


Fig. 152.3 2003.4.10 gas emission quantity

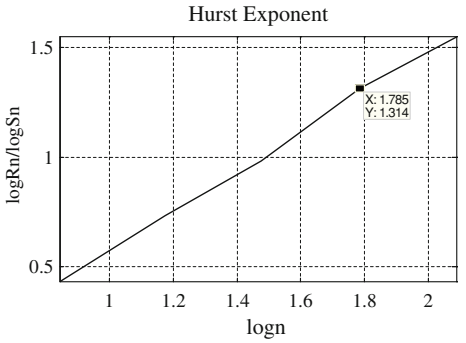


Fig. 152.4 2003.3.19 gas hurst exponent curve

using white noise and colored noise data. As the actual time-series data length are short, selecting a shorter time-series data makes sense. Data length, respectively 100,200,400 random colored noise sequences are generated by using MALTAB. For each the length data respectively run 10 Hurst index procedures. Eight Hurst index less than 0.5 when the data length 100; nine Hurst index less than 0.5 when length 200 and 400. The same method used to verify the white noise sequence data, the average of 10 times Hurst exponent is 0.521446 when data length is 100, a length of 200 the average Hurst exponent of 0.517392, a length of 400 mean Hurst exponent is 0.515577, the mean is 0.513631 when length 1000. So the programmer is right, and has certain robustness.

Figures 152.4, 152.5, 152.6 respectively are the fitted Hurst curve figure of three gas outburst time-series data in the 1,113 workface. Figs. 152.7 and 152.8, respectively, the time-series Hurst exponent curve of no gas emission. Several cases of the Hurst index characteristics are summarized in Table 152.2.

Figures 152.4, 152.5, 152.6 and Table 152.2 show that the three emission time-series Hurst exponent values are greater than 0.5, which indicate that the three time-series data obey fractal Brownian motion, and then is long-term positive correlation characteristics, within a certain time period of the performance will be characterized by continuity, that trend is sustainable.

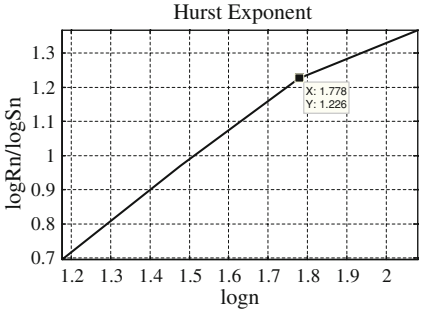


Fig. 152.5 2003.3.30 gas hurst exponent curve

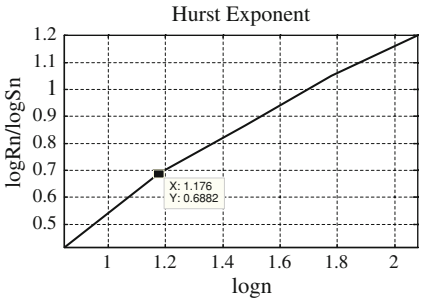


Fig. 152.6 2003.4.10 gas hurst exponent curve

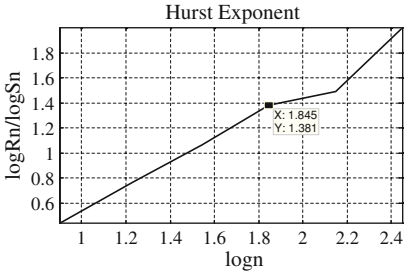


Fig. 152.7 No outburst gas data 1 hurst exponent

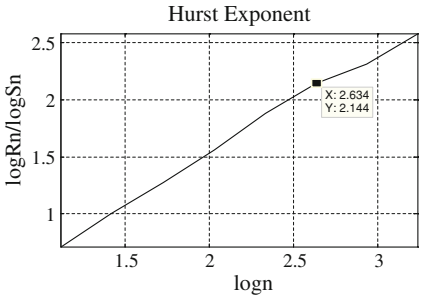


Fig. 152.8 No outburst gas data 2 hurst exponent

Table 152.2 Features table of 1,113 workface gas time-series hurst exponent

	Hurst exponent	Turning point abscissa value	effective correlation length	Fractal dimension
2003.03.19	0.8841	1.785	101.785	1.1309
2003.03.30	0.7888	1.778	101.778	1.2677
2003.04.10	0.6019	1.176	101.176	1.5825
No outburst gas data 1	0.9502	1.845	101.845	1.0524
No outburst gas data 2	0.8859	2.634	102.634	1.1288

Table 152.2 shows that gas emission time-series Hurst exponent maximum value is March 19 data, and its positive trend is the strongest, so the effective correlation length of the time series is the longest and its predictable time is also the longest. Figure 152.4 also shows that the overall gas emission is a growing trend. March 30 of the Hurst gas emission time series is 0.7888, as can be seen from Fig. 152.5 twice small anomalies phenomenon occur before the gas emission, and however, its Hurst value indicates the positive trend is still very strong, and its effective correlation time is shorter than March 19. The minimum Hurst exponent is April 10 gas emission time series, the Hurst exponent value 0.6019 has been closer to 0.5, indicating the time-series trends fluctuations, and enduring relevance is not obvious. Nevertheless, the Hurst exponent is still larger than 0.5, indicating that this time series still has a positive correlation, from its time-series emission curves can also be seen gas emission beyond the limiting value more, and there were two times gas emission.

In addition, in the article Hurst exponent two groups time-series data of 1,113 work face not emission has been also studied, their curve is shown in Figs. 152.7 and 152.8. From the Figs. 152.7 and 152.8 and Table 152.2 can be seen Hurst exponent values of not emission time series are higher than emission time series.

152.4 Conclusions

In this paper, long-range correlation properties of gas emission time series in the workface have been studied by R/S methods. This method can be used as an effective way to predict workface gas outburst, and it not only effectively analyzed gas emission time series when gas is out bursting, and there be significant forecast gas emission of daily prediction, but can also be used as a conventional drainage effects test. It will be a new face gas outburst prediction method that combined the long-range correlation with the short-term predictability of chaos characteristics. While, it will be continued in-depth study as the direction in this paper. Dynamic fractal characteristics of the gas emission time series wants to be further studied, so it must be linked with the gas outburst examples, which would be inconsistent with the safety in coal mines.

References

1. Liwen H (2010) Nonlinear characteristic analysis and study of the theory for prediction and simulation about gas emission in the working faces of coal mine, Central South University. pp 34–35
2. Shiliang S, Yi S, Liwen H, Chuanqu Z (2006) Research on determination of chaotic characteristics of gas gush based on time series in excavation working face of coal mine. J China Coal Soc 31:701–705
3. Kai W, Yibo W, Jie L (2007) Study on nonlinear characteristics of dynamic precursor of coal and gas outburst. J Min Saf Eng 24:22–26
4. Liwen H, Shiliang S, Yi S, Ying L (2008) Complexity and measurement of complex degree of gas gush in heading faces of coal mine. J China Coal Soc 33:547–550
5. Yuanyao L, Kunlong Y, Wenming C (2010) Application of R/S method in forecast of landslide deformation trend. Chin J Geotech Eng 8:1291–1296
6. Cunhan H, Tao F, Weijun W, Hui L (2010) Mine water inrush prediction based on fractal and support vector machines. J China Coal Soc. pp 806–810
7. Selvaratnam S, KirleyM (2006) Predicting stock market time series using evolutionary artificial neural networks with hurst exponent input windows, in AI 2006: Advances in Artificial Intelligence, Springer Berlin. pp 617–626
8. Yexue L, Jianfeng L (2010) Study on deformation characteristic of surrounding rock by R/S method and fractal theory. J Sichuan Univ (Engineering Science edition) 34:43–48
9. Li M, Lim SC, Hu BJ et al. (2007) Towards describing multi-fractality of traffic using local hurst function, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Beijing. pp 1012–1020
10. Wenjin W (2009) Long memory analysis and forecasting of financial time series, Tianjin University. pp 15–16
11. Yan-wei L, Mingju L, Gangsheng W (2005) Hebi 10th coal mine gas emission characteristics and predictor selection and application. Coal Mine Saf 36:18–21
12. Chen Y, Sun R, Zhou A (2010) An improved hurst parameter estimator based on fractional fourier transform. Telecommun Syst 43:3–4
13. Benassi A, Bertrand P, Cohen S (2000) Identification of the hurst index of a step fractional brownian motion. Stat Infer Stoch Process 2:121–125

Chapter 153
Researches on the Laser-Induced Sound as the Sound Source of Imaging Sonar

Yan-Hui Wang, Jian-Zhou Mao and Zhi-Hong Xiu

Abstract To determine whether Laser Induced Sound (LIS) could be used as sound source of imaging sonar for minesweeper. Experimental environment and some important parameters are described. The structure and acoustic parameters of LIS transducer are articulated with a series of diagrams and formulae. Width and the directivity index of the beam of LIS transducer are achieved. The analysis of the feasibility of LIS as sound source of imaging sonar play an important role in this paper, which consist of four parts: analysis of distance resolution of sonar with LIS, the detection range of LIS under background noise, reverberation, thorough analysis of the feasibility of LIS as imaging sonar's sound source. The results show that sonar with LIS as sound source could detect a metal ball with radius of 3 cm or the target with target strength of −50 dB, its range resolution and the maximum detection range are 5–8 times higher than the existing imaging sonar's.

Keywords Laser induced sound · Imaging sonar · Sound source · Range resolution

153.1 Introduction

Laser pulse focused in distilled water, if the power density of the focus beyond the breakdown threshold of water, would cause optical breakdown [1]. The sound caused by optical breakdown called Laser Induced Sound(LIS) [2]. This paper

Y.-H. Wang (✉) · J.-Z. Mao · Z.-H. Xiu
Department of Operation and Command, Dalian Naval Academy,
Dalian 116018, China
e-mail: bagum_wyh163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_153,

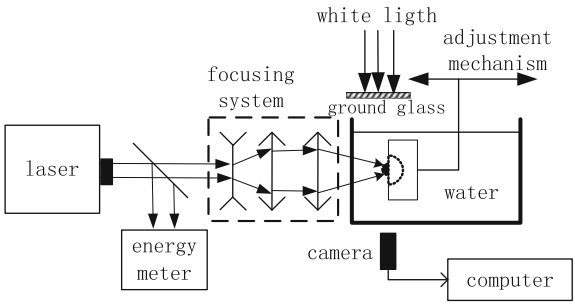


Fig. 153.1 The setup of experiment [6]

conducted a study to determine whether LIS could be used as sound source of imaging sonar. Range resolution and detect distance of LIS sonar were analyzed under typical oceans environment [3–6].

153.2 Experimental Setup

153.2.1 Experimental Setup Description

The schematic of the experimental setup is showed in Fig. 153.1 [5], a beam of laser pulse shot on a set of focusing system and focused in distilled water, which cause optical breakdown and produce LIS. Hydrophone received LIS and store the data in a computer for further analysis.

The sound level is tested under condition that the laser pulse's energy is 600 mJ and the distance of hydrophone to the center of breakdown is 15 cm, the mean value of 100 times is showed in Fig. 153.2. The first peak is sound of plasma swelling, the second peak is the sound of vacuolated collapsing.

153.2.2 The Principle of LIS's Transducer

Figure 153.3 shows the LIS transducer described in reference 5, laser beam come from point 1 directed to a set of lens and focus on focal point O in distilled water, which cause optical breakdown and produce LIS. There are 2 ways LIS can transmit to surface 7, one way is from point O to paraboloid then to surface 7 without hemispherical surface 5 (such as from point O to point A then to point N), the other way is from point 5 to hemispherical surface 5 to paraboloid surface 4 then to surface 7 (such as from point O to point B to point C then to point M). Based on acoustic theory of sound-hard-surface [7] and reflection,the sound-ray by

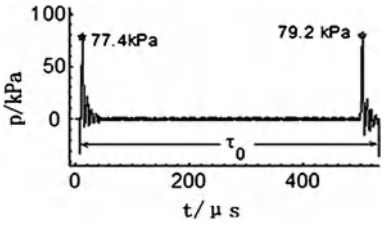


Fig. 153.2 Laser-induced sound $p(t)$

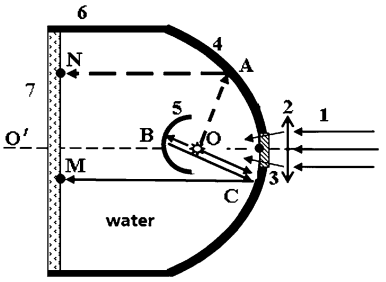


Fig. 153.3 The sketch of laser-induced sound transducer

each way at surface 7 is parallel. With the theory of point-source combination [7], LIS beyond surface 7 is a beam width certain width but not plane wave.

153.2.3 Acoustic Parameter of LIS Transducer

The directivity function of LIS transducer can be described with formula (153.1) [7].

$$b(\theta) = \frac{p(\theta)}{p(0)} = \left| \frac{2J_1(kr\sin\theta)}{kr\sin\theta} \right| \tag{153.1}$$

where P is the acoustic pressure, θ is the deviation angle from axis OO', J_1 is Bessel function, r is the radius of the cylinder, k is wave number.

Based on the definition of directivity index DI_t, can deduce that:

$$DI_t = 10\lg\left(\frac{b^2(0)}{\int_0^{\frac{\pi}{2}} b^2(\theta)\sin(\theta)d\theta}\right) = 10\lg\left(\frac{2kr}{1 - J_0^2(kr) - J_1^2(kr)}\right) \tag{153.2}$$

The beam width of LIS, namely, θ_t can be described with formula (153.3). When the radius of hemispherical surface in Fig. 153.3 is 25 cm, the output wave of LIS transducer can be showed as in Fig. 153.4.

$$b(\theta_t/2) = 1/\sqrt{2} \tag{153.3}$$

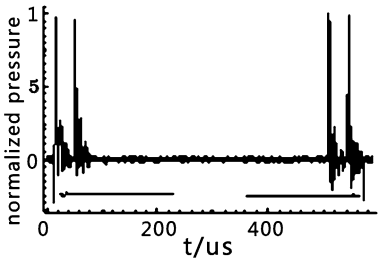


Fig. 153.4 The signal $P_0(t)$ of transducer

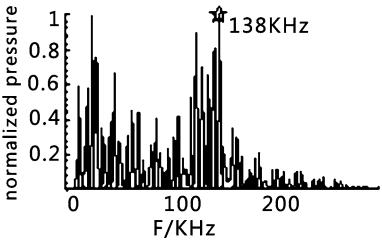


Fig. 153.5 The power spectrum of signal $P_0(t)$

153.3 Results

153.3.1 Range Resolution of the System

Take main lobe and side lobe into account considerately, the constant delay to distinguish (A_τ) and range resolution(ΔR) of LIS signal, namely, $P_0(t)$ could be described with formula (153.4) and (153.5)

$$A_\tau = \int_{-\infty}^{\infty} |P_0(f)|^4 df \Big/ \left[\int_{-\infty}^{\infty} |P_0(f)|^2 df \right]^2 \tag{153.4}$$

$$\Delta R = cA_\tau/2 \tag{153.5}$$

where $P(f)$ is the fourier transform of LIS singnal $P_0(t)$, f is the frequency, c is the velocity of sound in water, take as 1,500 m/s.

Power spectrum of signal $P_0(t)$ is illustrated in Fig. 153.5, its total spectrum is wider than 150 kHz, the main frequency is 138 kHz, the wave number is $k = 2\pi c/f_0 = 578$, with formula (153.4) and (153.5) can work out: $A_r = 7$ μs, $\Delta R = 5.3$ mm, it means that sonar with LIS as the sound source can recognize target with demission large than ΔR. It proves that as far as distance resolution is consider LIS could be used as the sound source of imaging sonar.

153.3.2 Detect Range Under Noise Background

Take minehunter sonar as example, the detect formula under background noise [8] is:

$$2\mathrm{PL} = \mathrm{SL} + \mathrm{TS} - \mathrm{N} + DI_r + 10\lg(\tau_1) - 5\lg(d) + 5\lg(n) \tag{153.6}$$

where PL is the lost of propagation, TS is target strength, N is noise of environment, $\mathrm{DI_r}$ is directivity index of receive array, τ_1 is the width of LIS in Fig. 153.4, d is detection index, n is pulse number when making up decision. The lost of propagation PL can be calculated with formula (153.7).

$$\mathrm{PL} = 20\lg(\mathrm{L}) + \alpha L \times 10^{-3} \tag{153.7}$$

In formula (153.7), α is absorptance coefficient(dB/km), L is the max detect distance(m). SL could be deduced with the definition of sound source as formula (153.8).

$$SL = 10\lg\left(\frac{1}{\tau_0}\int_0^{\tau_0}\left(\frac{l}{l_r}\cdot p(t)\right)^2 dt \Big/ p_r^2\right) + \mathrm{DI_t} \tag{153.8}$$

where $P(t)$ is the LIS signal of Fig. 153.2, τ_0 is the LIS pulse time of Fig. 153.2, P_r is reference sound pressure taken as 10^{-6}Pa, l is the distance from sound receiver to center point of laser, take 0.15 m. l_r is the reference distance, take 1 m. Supposing that LIS is used to detect a ball with radius of α, the target strength is:

$$TS = 10\lg(\mathrm{a}^2/4) \tag{153.9}$$

If a surface with sound insulation plane [8] is used as the receive array, then the directivity index can be work out with formula (153.10). Where θ_r is the width of the received beam.

$$\mathrm{DI_r} = 3 + 10\lg(100/\theta_r) \tag{153.10}$$

153.3.3 Detect Range Under Reverberate Background

For minehunter sonar the detect formula under background of reverberation can be explained with formula [8] (153.11):

$$10lg(L) = 10lg(B/\theta_r) - S_b - 11 + TS - 5\lg(d) + 5\lg(n) \tag{153.11}$$

where $\mathrm{S_b}$ is the strength of backscattering, B is the bandwidth of $P(t)$ which can be estimated with formula [8, 9] (153.12).

Table 153.1 The maximum detection distance

Para background	TS(dB)	θ_r(°)	L(m)
Noise	−36.5	1.0	371.9
		0.2	470.7
	−50.0	1.0	283.7
		0.2	375.0
Reverberant	−36.5	1.0	548.6
		0.2	2743.0
	−50.0	1.0	24.3
		0.2	121.9

$$\mathrm{B} = 1/A_\tau \tag{153.12}$$

where A_τ is the constant delay to distinguish signal $P_0(t)$ in Fig. 153.4. If the value is 7 μs then we can work out B to be 143 kHz.

153.3.4 Possibility as Sound Source of Imaging Sonar

In this paper a ball with radius of 0.6 cm with target strength of −50 dB, namely, a target similar to a fish is calculated, and the results are shown in Table 153.1. From the table we can reach the conclusion that the maximum detection range of LIS sonar is 5–8 times higher than the existing imaging sonar's 10 m ∼ 50 m [10].

153.4 Conclusion

LIS is a kind of repeatable and steady and high distance resolution sound source, especially it has the advantage over traditional sound source in detecting distance and distance resolution. Main frequency, bandwidth and sound source level of LIS can be achieved by changing laser's parameter and liquid. Calculation and derivation both proved that LIS could be used as sound source of imaging sonar to detect distance and distance resolution are considered.

References

1. Kennedy PK (1995) A first-order model for computation of laser induced breakdown thresholds in ocular and aqueous media. Part I-theory. IEEE J Quantum Electron QE-31:2241–2249
2. Vogel A (1999) Energy balance of optical breakdown in water at nanosecond to femto second time scales. Appl Phys B 68:271–280

3. Rong-fu L, Guei-hua C, Zuo-xi T, Guo-ming S (2003) Laser-sound remote sensing technique[M]. National defence industry press, Beijing, pp 3–64
4. Blackmon F, Antonelli L, Kalinowski A (2005) A remote optical system for port and harbor defense. Proc SPIE 5780:99–106
5. Yuhong W, Jiang'an W, Siguang Z et al (2008) Design of sonar transducer based on the laser-induced sound [J]. Chin J Acoust 33(6):562–565
6. Yuhong W, Jiang'an W, Ronghua W (2009) Influences of liquids characteristics and the absorbed energy to the laser induced expanding sound [J]. High Power Laser Part Beams 21(7):998–1002
7. Gong-huan D, Zhe-ming Z, Xiu-fen G (2001) Sounds basic theory[M]. Nanjing University Publishing House, Nanjing, pp 199–354
8. De-shi W (2004) Sonar for practising engineers third edition[M], vol 9. (translator). Publishing House of Electronics Industry, Beijing, pp 13–228
9. Mao-yong L, You-an K (1984) Radar signal theory[M]. National Defense Industry Press, Beijing, pp 93–127
10. Meng-an L (2002) Underwater sound engineering[M]. Zhejian Science and Technology Press, Hangzhou, pp 341–371

Chapter 154
Analysis and Design of Sensors Based on MEMS Process

De En, Nana Wei, Huanghe Wei, Yangyang Liu and Xiaoguang Chen

Abstract Due to the features of light weight, small size, high intelligence, MEMS is widely used in the field of nature and engineering science, such as communication, aerospace, biomedical and materials science and so on. In this paper, MEMS manufacturing process is introduced, the structure and working principle of MEMS' acceleration seismic sensors are analyzed and designed.

Keywords MEMS · Acceleration · Sensor · Manufacturing technology · Seism

154.1 Introduction

MEMS is short for Micro Electro Mechanical Systems, which was raised in Europe and other developed countries in the late 1970s. It is a micron-scale's device which is similar to integrated circuit, mainly includes miniature institutions, miniature sensors, micro actuators, corresponding processing circuit, and other few parts.

MEMS technology has been widely used in various fields of science. In the geophysical field, MEMS technology has been gradually applied. Using it to

D. En (✉) · N. Wei · H. Wei · Y. Liu · X. Chen
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: ende@hpu.edu.cn

N. Wei
e-mail: amyweinana@126.com

H. Wei
e-mail: weihuanghe@yeah.net

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_154,

develop a new generation of digital seismic acceleration sensor, its output shows the characteristics of the acceleration detector range from DC to AC, very low intrinsic noise and distortion, direct digital output, high vector fidelity and good low frequency response [1].

154.2 Manufacturing Technology of MEMS

MEMS use conventional silicon micromachining techniques, which can only processing silicon, including surface micromachining and bulk micromachining. Using of the technology has been successfully developed part of the MEMS devices, such as acceleration sensors, pressure sensors and micro-gyroscope. As Fig. 154.1 is shown that mask film of corrosion-resistant on the silicon is formed by using deposition, ion implantation, oxidation, or electroforming, then etching the film by photolithography, the last further corrosion of silicon body, then the desired microstructure is obtained [2].

154.2.1 Surface Micromachining

Surface micromachining proposed in the 1980s which is one of the popular MEMS processing methods. Usually thin layer can be divided into structure layer and the sacrificial layer, it is suitable for processing thin film components which suspense on the surface, the size is relatively small and there is a certain freedom of movement. The material of structure layer is polysilicon, silicon nitride, aluminum, etc. which is used to make movable micro-mechanical structure; Sacrificial layer is used as an intermediate cushion layer, it is mostly oxides and organic materials, layers of these materials will be repeated deposition and etching processes [3].

154.2.2 Bulk Micromachining

Bulk micromachining is a method that entity structure processing can be obtained through directly remove silicon substrate, including wet etching and dry etching. The process of wet etching includes materials oxidation and oxide dissolved by chemical reactions, its material is silicon, quartz, including wet chemical etching and wet electrochemical etching. Dry etching, including Reactive ion sputter etching based on physical effects, plasma etching based on chemical and both physical and chemical reaction sputtering etching role.

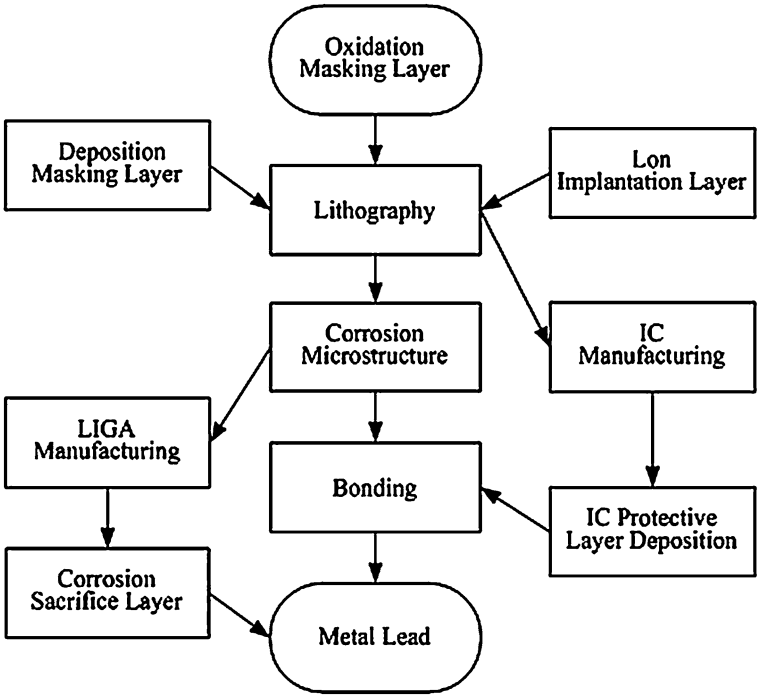


Fig. 154.1 Micromachining process flow chart

154.3 Structure and Principle Analysis of MEMS Sensor

154.3.1 Sensing Technology of MEMS

MEMS sensor technology is signal devices and systems of converting the electrical by using microelectronics micro-mechanical processing out and using sensitive components such as capacitor, piezoelectric, pressure resistance, thermal power decoupling, resonant, tunnel current, and so on. Sensor's development direction is the array, integration, and intelligence at present [4].

MEMS acceleration sensor includes Micro Electro Mechanical System and Application Specific Integrated Circuit.

1. Basic circuit principle

MEMS sensors are divided into open and closed loop. As Fig. 154.2 shown is the open loop MEMS sensor, the series of accelerometers can measure both dynamic acceleration and the static acceleration. Because of its big noise, it cannot be used as geophones.

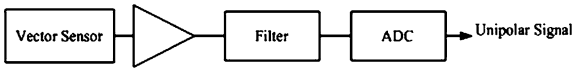


Fig. 154.2 Open-loop acceleration sensor schematic

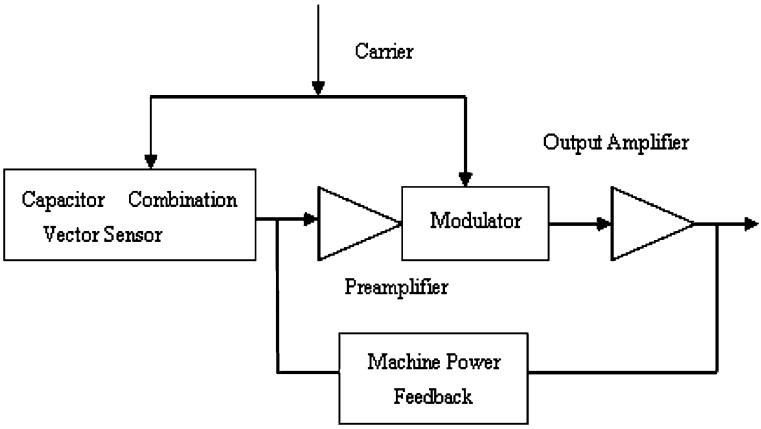
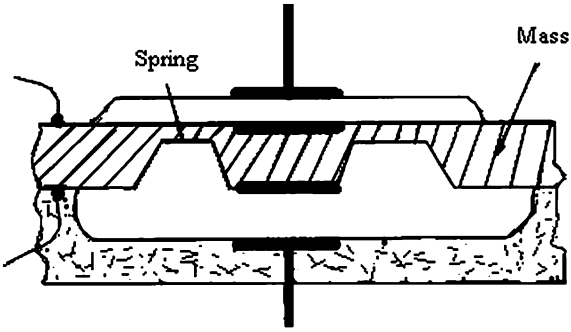


Fig. 154.3 Closed-loop accelerometer schematic

Fig. 154.4 MEMS structure section



Closed loop acceleration sensor based on principle of capacitance change, acceleration amplification and converted into the displacement of the capacitor plates by inertia mass. Differential capacitor change into electrical signals through the detection circuit, after force balance feedback loop, encouraging the movable capacitor plate is always in equilibrium position. At the same time take the feedback signal as output, it shows the size of the input acceleration, as Fig. 154.3 is shown [5, 6].

2. Structural analysis of mechanical systems MEMS

MEMS mechanical vibration system consists of mass, spring, cover, the framework (as Fig. 154.4 is shown) upper and lower surfaces of mass distribute metal conductive layer, below the roof top and bottom of the metal distribute metal

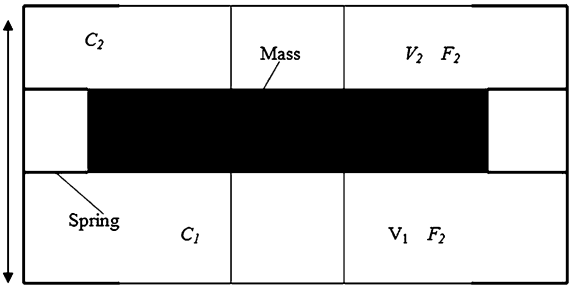


Fig. 154.5 MEMS sensor operating principle

conductive layer, so formed a differential capacitor, with the corresponding external circuit can become capacitive accelerometer. From the structure to see, inertia mass and spring which hanging in the structure around constitute the MEMS sensor. Changing the coefficient of elasticity can make the resonance frequency to achieve kHz and beyond the frequency band of seismic signal. When designed it to work under the resonant frequency, the sensor can show its acceleration characteristics.

Geophone needs very low noise. MEMS output noise is white noise, which is the main source of noise in MEMS. In an absolute temperature, output noise power spectral density N and inertia plastid M, resonance damping B, and other parameters has the following relationship:

$$N^2 = \frac{4KTB}{M^2 g^2} \tag{154.1}$$

where K is the Boltzmann's constant.

154.3.2 Principle Analysis of MEMS Sensor

As Fig. 154.5 is shown that each MEMS consists of four independent silicon chips, upper and lower surfaces of mass is the negative capacitor, plate of upper and lower is the capacitor's positive, the distance between two plates is about 10 μm, the largest displacement of the mass is from −5 μm to +5 μm.

Suppose two capacitors are C_1 and C_2, the plate spacing are X_1 and X_2, When the value of acceleration of moving inertia mass, capacitance ratio is

$$\frac{C_1}{C_2} = \frac{X_2}{X_1} \approx \frac{1 + 2\Delta X}{X_1} \tag{154.2}$$

Schematic diagram of the MEMS sensor, capacitance of between under cover and mass is C_1, applied voltage is V_1; Capacitance of between upper cover and mass is C_2, applied voltage is V_2. The sensor is in a dormant state when it is unpowered. Gravity pulls down the quality of the body, right now $C_1 > C_2$; The loop began to work after power up, adjust the size of V_1 and V_2 to produce a

force to overcome gravity, sensor achieve a balance until C_1 is equal to C_2 and F1 plus F_2 is equal to g, then ready to record the signal. When receiving seismic signal, continuous sampling the value of C_1 and C_2, the proportion changed as the mass moved, at the same time the value of V_1 and V_2 changed by negative feedback loop circuit to produce compensation, make the quality body move to the center. Sensor base on the mass back to the center to maintain the required output signal correction obtained, and the vertical axis of the sensor work was supported by the spring to prevent the seismic signal, no signal output.

154.4 Conclusions

Capacitive acceleration sensor based on MEMS has some features, such as small size, efficient immunity to interference, fitting to work in the powerful electromagnetic and high or low temperature environment, its dynamic response time is short, suitable for dynamic measurements, and the high-frequency system can be used to measure changes in the parameters of high-speed, so it has better applicable foreground.

Acknowledgements This paper is supported by the followings: (1) National Natural Science Foundation of China (41074090), (2) Key Technologies R and D Program of Henan Province (022102210360), (3) China Postdoctoral Science Foundation (2005038468), (4) Ph.D. Programs Foundation of Henan Polytechnic University (B2008-10).

References

1. Guizhang L, Xin Z (2010) MEMS design [M]. Scientific Publisherst, Beijing
2. Desheng L, Donghong W, Jinwei S, Peng J (2002) MEMS technology and its application[M]. Harbin Institute of Technology Press, Ha er bin
3. Linggang K (2005) MEMS sensor principle and structure analysis[J]. China Academic Journal Electronic Publishing House, p 21
4. Guihui C, Wanming Z, Cheng R (2010) Xi Jianzhong: research of new seismic geophone based on MEMS [J]. Transducer and Microsystem Technologies, p 29
5. Jonathan Bernstein et al (2002) MEMS inertial sensing technology
6. Guang lin liu Tai sheng liu et al (2003) On the development trend of seismic geophones [J]. Progress in exploration geophysics, p 26

Chapter 155
The Study of Displacement Measurement Instrument of Displacement Sensor Based on Polarized Light

De EN, Xiaoguang Chen, Yangyang Liu, Nana Wei and Huanghe Wei

Abstract A kind of displacement measurement instruments based on polarized light displacement sensor is presented in this chapter. The signals from two polarized light displacement sensors that respectively indicate the value of coordinate X and coordinate Y are subdivided by a special circuit. After the subdivided signals are identified and counted, a microprocessor processes the identified and counted signals to obtain the values of the displacement and implements the error correction algorithm. The characteristics of the instrument are reliable, easy to extend and convenient to debug.

Keywords Displacement measurement · Measuring precision · Polarized light sensor

155.1 Introduction

In recent years, displacement measuring instrument used for X–Y table is widely used in precision measuring machine and MEMS, and the measuring precision and system extensibility of this instrument was put forward higher request.

D. EN (✉) · X. Chen · Y. Liu · N. Wei · H. Wei
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: ende@hpu.edu.cn

X. Chen
e-mail: chzhm2010@yeah.net

Y. Liu
e-mail: lvshucn@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_155,

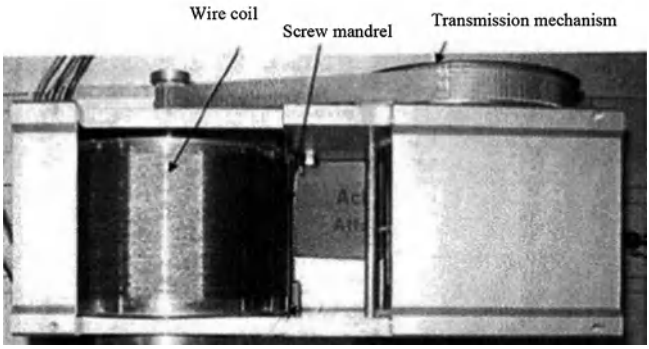


Fig. 155.1 Real figure of conversion institutions

This chapter introduces the study of displacement measurement instrument of displacement sensor based on polarized light, it subdivided the output signal which come from the polarized light sensor, and the subdivided signal was identified direction and counted respectively. After this signal was processed by the microprocessor, we can get the displacement measured value of the coordinate X and Y direction. Finally, according to the error correction algorithm confirmed displacement measured value to improve the measurement accuracy of the measuring instrument.

155.2 Measurement Principle and Structure

155.2.1 Machinery Conversion Institutions of Displacement Sensor

Machinery conversion institutions use the traction positioning line to cause rotation of the wire coil, and produce the angular displacement by the transmission mechanism of conveyor belt. Figure 155.1 is a kind of real figure of conversion institutions which transform linear displacement into angular displacement [1].

155.2.2 Polarized Light Angular Displacement Sensor

Figure 155.2 gives the principle of polarized light angular displacement sensor. Aperture is in the front of the LED, Polarizer P1 is fixed, Analyzer P2 run with angular displacement. If luminous intensity of the light through P1 is I_0, and P2 has no additional absorption except the polarization property, the luminous intensity through P2 can be written as:

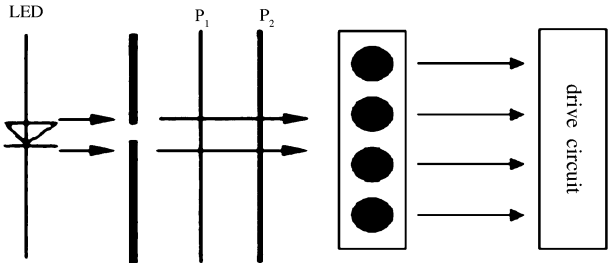


Fig. 155.2 Polarized light angular displacement sensor

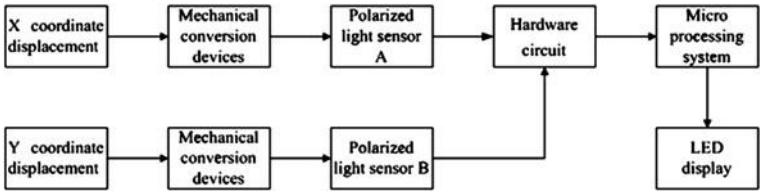


Fig. 155.3 System structure flow diagram

$$I = I_0 \cos \alpha \tag{155.1}$$

where α is the included angle of P1 and P2.

The output luminous intensity signal is converted to sinusoidal voltage signal by four light sensors proportionately. At last, the power and amplitude of output signal is amplified by drive circuit. The output luminous intensity signal is converted to sinusoidal voltage signal by four light sensors proportionately. At last, the power and amplitude of output signal is amplified by drive circuit [2].

155.2.3 System Structure

Displacement Measurement Instrument of Displacement Sensor Based on Polarized Light is made up of machinery conversion device, polarized light angular displacement sensors, hardware circuit, Micro Processor Systems and LED display. System structure is shown in Fig. 155.3.

The signals from two polarized light displacement sensors that respectively indicate the value of coordinate X and coordinate Y are subdivided by a special circuit. After the subdivided signals are identified and counted, a microprocessor processes the identified and counted signals to obtain the values of the displacement, and implements the error correction algorithm and sends the signals to the LED display at the same time.

155.3 Hardware Circuit

155.3.1 Signal Subdivision Circuit

Figure 155.3 gives the flow of signal processing. Signal subdivision circuit is made up of fivefold frequency circuit and fourfold frequency circuit. The signals from two polarized light sensors are subdivided at the subdivision circuit. In order to simplify process of the output four-way signals, use QA740204 which is fivefold frequency integrated circuit to process signals. After filtered by QA740204 and the signals are phase shifted and wave shaped, we can get two-way orthogonal square wave signals. Fourfold frequency circuit adopts integrated circuit QA740210 which do fourfold frequency to orthogonal square wave signals from five frequency circuit, and at last, we get 20-fold pulse signals (XCP+ and XCP−,YCP+ and YCP−) (Fig. 155.4).

155.3.2 Direction Discrimination and Signal Count Circuit

In order to count correctly, the direction discrimination circuit is needed. This circuit includes two units to distinguish the direction of X coordinate and Y coordinate. Microprocessor gets 3-byte data by the data bus, and compares this data with setting value. When the measurement data is greater than the setting value, movement direction is positive, otherwise is negative [3].

As shown in Fig. 155.5, signal count circuit is made of two 18-byte reversible counter. UT1 ∼ UT10 use 4-byte binary count chips 74LS193. The UT5 and UT10 reversible count chips are regarded as two binary counter, so five pieces of reversible count chips UT1 ∼ UT5 are connected to be 18-byte reversible counter. When the table moves positive, the reversible counter accumulates the pulse count. When the table moves negative, the reversible counter will accumulate negative value of the pulse to get the correct measuring results.

155.3.3 Micro Processing System

Microprocessor system is mainly made up of the buffer, memory and microprocessor, as shown in Fig. 155.6. The buffer unit is made up of two buffers including six piece of 74LS244, three pieces of them constitute one buffer, and they can temporarily store value of the reversible counter.

In order to improve the accuracy of the measurement instrument, the linear interpolation error correction algorithm is used to amend the value [4]. The memory is used for storing the system setting parameters, the data collected and error correction data tables. P89C58 is used for microprocessor; it gets binary data

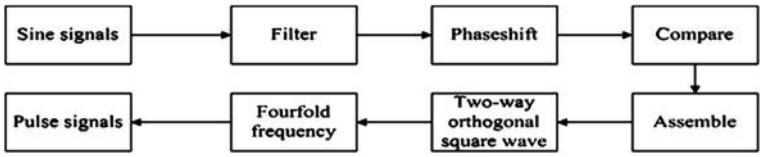


Fig. 155.4 System structure flow chart

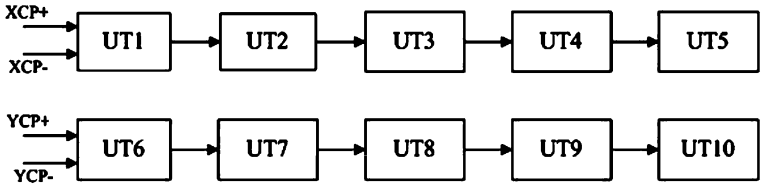
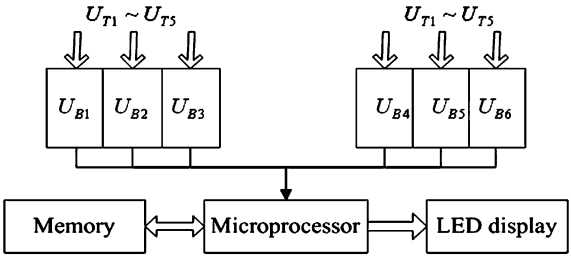


Fig. 155.5 Signal count circuit

Fig. 155.6 Micro processing system structural diagram



according to a certain time sequence from data bus. After calculating and processing, the processor gets the decimal data of measurement. Then the processor takes error correction according to the error table and sends the revised data to LED display.

155.4 The Software System

The software system uses the software development platform of KEIL C, using C language and assembly language-mixed programming. Assembly language application procedure and interrupt subroutine, written in C language are the main program. The model which C language program called the assembly procedure ensures the microprocessor to data collection and processing efficiency and strengthens the readability of the program. Application procedure mainly includes coordinates of the reset, divide, absolute value and the relative value of the conversion, the transformation of the measurement model, and computer communications, printing, etc.

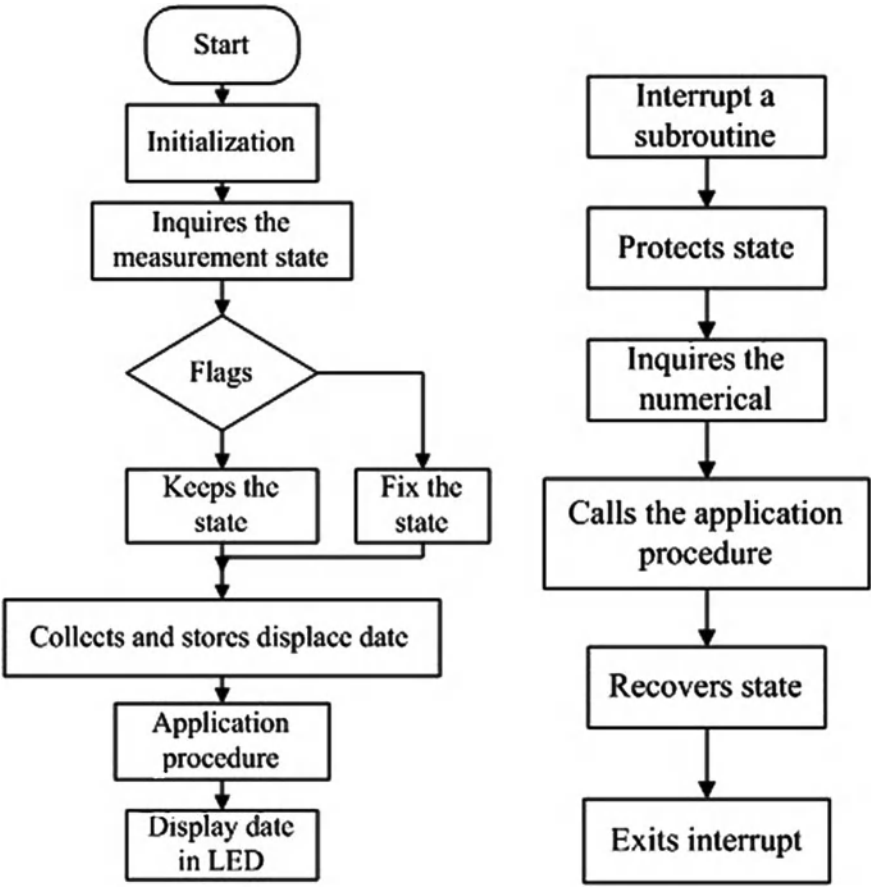


Fig. 155.7 The main program flow diagram and interruption subroutine flow diagram

Interrupt procedure including keyboard interrupt and the stage of the mobile zero interrupt function, can accomplish in real-time measurement value of the processing and other application functions. What the Fig. 155.7 shows is the main program flow diagram and interruption subroutine flow diagram.

155.5 The Experimental Results and Analysis

The normal line in the range of 200 mm, scribed line for 1 mm, interval time line for the 0.12 um is used as the standard; the whole measuring range is measured by every 10 mm. For coordinate X: the measuring range is $-100 \sim 100$ mm. There is 21 point to be measured; for coordinate Y: the measuring range is $-90 \sim 80$ mm; there is 18 point to be measured [5]. The measuring result is shown in Table 155.1.

Table 155.1 The measuring result of coordinate X and coordinate Y

Coordinate X true value/mm	Coordinate X measured value/mm	Coordinate Y true value/mm	Coordinate Y measured value/mm
−100.000	−99.999		
−90.000	−89.998	−90.000	−90.003
−80.000	−79.998	−80.000	−80.001
−70.000	−69.999	−70.000	−70.002
−60.000	−59.997	−60.000	−60.001
−50.000	−49.998	−50.000	−50.002
−40.000	−39.998	−40.000	−40.001
−30.000	−29.999	−30.000	−30.001
−20.000	−19.998	−20.000	−20.002
−10.000	−9.999	−10.000	−10.001
0.000	0.000	0.000	0.000
10.000	10.000	10.000	9.999
20.000	20.001	20.000	19.998
30.000	30.002	30.000	29.999
40.000	40.001	40.000	39.999
50.000	50.001	50.000	49.999
60.000	60.000	60.000	59.998
70.000	70.001	70.000	69.999
80.000	80.002	80.000	79.998
90.000	90.002	90.000	89.997
100.000	100.001		

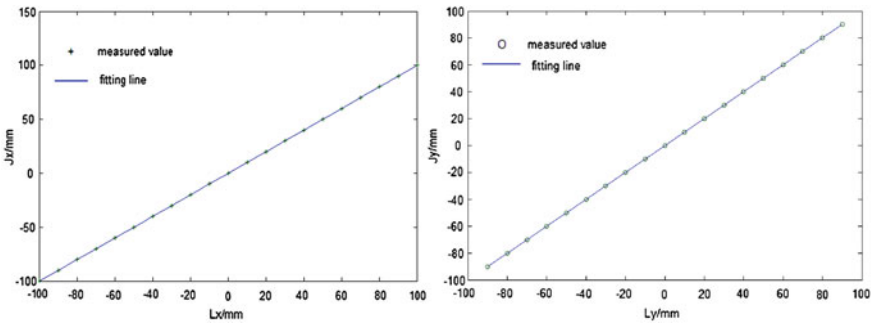


Fig. 155.8 The measured value and fitting line

Using least square method to fit the value of Table 155.1, we get the function relation between the true value and the measured value.

$$Jx = 0.00113 + Lx \tag{155.2}$$

$$Jy = 0.00109 + Ly \tag{155.3}$$

where *Lx* is the coordinate x true value;
Jx is the coordinate x measured value;
Ly is the coordinate y true value;
Jy is the coordinate y measured value
Measurement data and the fitting line are shown in Fig. 155.8.

Acknowledgments This chapter is supported by the followings: (1) National Natural Science Foundation of China (41074090), (2) Key Technologies R and D Program of Henan Province (022102210360), (3) China Postdoctoral Science Foundation (2005038468), (4) Ph.D. Programs Foundation of Henan Polytechnic University (B2008-10).

References

1. Guanglin L, Taisheng L et al (2003) On the development trend of seismic geophones [J]. Progress in exploration geophysics. p 26
2. Qiushu L, Wanbo Z et al (2005) The Principle and implementation of single chip computer 160 subdivision [J]. Instrum J, p 26
3. Wei L (1999) Research on the principle and method of polarized light displacement sensor [J]. Instrum J 3:221–224
4. Weimin Z (1997) A novel high sensitive photo electric micro-pressure sensor. Chin J Sci Instrum vol 18
5. Guolei C (2006) Based on the high speed to the single chip microcomputer control system design displacement measurement [J]. Ind Autom 4:44–46

Chapter 156
Optimal Filter for Stochastic Uncertain Systems with Multiplicative Noise and Sensor Failure Rates

Jing Ma and Shuli Sun

Abstract The current paper is concerned with the optimal linear filter problem for linear uncertain discrete-time stochastic systems with multiplicative noise and sensor failures. Based on the innovation analysis approach, the optimal linear filter is developed in least mean square sense. Compared with the robust filter in the literature, the proposed filter avoids the artificial filtering gain that brings the additional computation. The steady-state filter is also investigated. A sufficient condition for the existence of the steady-state filter is given. The asymptotic stability of the optimal filter is analyzed. Illustrative example shows the effectiveness of the proposed algorithm.

Keywords Optimal linear filter · Multiplicative noise · Sensor failures · Innovation analysis

156.1 Introduction

In recent years, the research on networked control and estimation has attracted much attention. Compared with traditional systems, networked systems have many advantages, such as low cost, inherent robustness and power requirements, as well as

J. Ma
School of Mathematics Science, Heilongjiang University,
Harbin 150080, China
e-mail: majing@student.hlju.edu.cn

S. Sun (✉)
School of Electrical Engineering, Heilongjiang University,
Harbin 150080, China
e-mail: sunsl@hlju.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_156,

high reliability [1–3]. However, several new challenging issues have been posed due to limited communication capacity and unreliable communication networks, which are the results of network-induced delay and uncertain observation (or sensor failures). For systems with random delayed measurements, many algorithms have been proposed, such as a modification of the conventional minimum variance state estimator [2] and the least mean square filter [3]. For the system of uncertain observation, the linear minimum variance filters are designed in [4–7], where the uncertain observation is described by a Bernoulli distributed stochastic variable. In [8], a recursive estimator is presented based on covariance information approach.

In many practical estimation problems, such as the tracking of a maneuvering target subject to disturbance, the estimation performance requirements are robust. The robust H_2/H_∞ filters independent of the parameter perturbations are designed for linear perturbed systems with steady-state error variance constraints [9]. A robust finite-horizon Kalman filter is designed for uncertain systems with both additive and multiplicative noises [10]. In [11, 12], the filter is proposed for systems subjected to parameter uncertainty in both the state and output matrices, respectively. But the results above are obtained based on deterministic perturbations or norm-bounded uncertainty in the system matrix. In a recent research effort [13], a robust and resilient minimum variance unbiased filter is designed for systems with stochastic uncertain parameters in the system matrix and multiple sensors of different sensor failure rates. In order to guarantee resilient operation, [13] introduces stochastic perturbations in the filtering gain. But how to select the variances and coefficients of stochastic perturbations is of subjective.

In this paper, we explore the optimal filter problem based on the system of [13] where the parameter uncertainties are allowed to be stochastic white noises with known variances perturbations and appear in the system state matrix. The system measurements are subject to sensor failures, which often occur in networked systems. Differently from [13] where a robust filter is designed by completing square, here the optimal linear filter is studied in least mean square sense by the innovation analysis approach. The form of our linear filter is simple and the computation burden is reduced. Furthermore, we remove the assumption of independent noises. The steady-state filter is also investigated. A sufficient condition for the convergence of the optimal linear filter is given.

156.2 Problem Formulation

Consider the following discrete-time stochastic uncertain systems with multiplicative noise and multiple sensors of different sensors failure rates

$$x(t+1) = \left(A_0(t) + \sum_{k=1}^{n_\alpha} \alpha_k(t) A_k(t)\right) x(t) + F(t)w(t) \tag{156.1}$$

$$y_i(t) = \gamma_i(t) C_i(t) x(t) + D_i(t) v_i(t), \quad i = 1, 2, \ldots, L \tag{156.2}$$

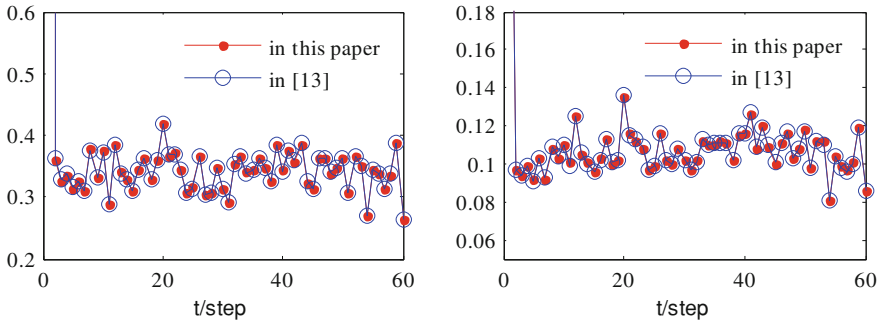


Fig. 156.1 Comparison of mean square error of filter in this paper and [13]

where $x(t) \in R^n$ is the state, $y_i(t) \in R^{m_i}$, $i = 1, 2, \ldots, L$, are sensor outputs, multiplicative noise $\alpha_k(t)$, $k = 1, 2, \ldots, n_\alpha$ are scalar white noise sequences that are uncorrelated with other random variables and are introduced for the structured perturbations in the system matrix. They are of mean zero and variance Q_{α_k} and are mutually uncorrelated. The matrices $A_k(t)$, $k = 1, 2, \ldots, n_\alpha$ are known and signify the direction of parameter perturbations. Also zero mean of the $\alpha_k(t)$ means that parameter perturbations in both directions are equally likely. The family $\gamma_i(t)$, $i = 1, 2, \ldots, L$ are scalar binary Bernoulli distributed random variables with the probabilities $\text{Prob}\{\gamma_i(t) = 1\} = \pi_i$ and $\text{Prob}\{\gamma_i(t) = 0\} = 1 - \pi_i$, where $0 \le \pi_i \le 1$, $i = 1, 2, \ldots, L$, and are uncorrelated with other random variables. $A_0(t)$, $F(t)$, $C_i(t)$, $D_i(t)$ are known matrices with suitable dimensions. Subscript i denotes the ith sensor, and L is the number of sensors.

In this paper, the mathematical expectation E operates on $\gamma_i(t)$ as well as $w(t), v_i(t), i = 1, 2, \ldots, L$, and $\alpha_k(t), k = 1, 2, \ldots, n_\alpha$. We make the following assumptions.

Assumption 1 $w(t)$ and $v_i(t)$ are correlated white noises with zero means and variances Q_w, Q_{v_i} and cross-covariance matrices S_i, $i = 1, 2, \ldots, L$.

Assumption 2 The initial state $x(0)$ is uncorrelated with $\alpha_k(t)$, $\gamma_i(t)$, $w(t)$ and $v_i(t), i = 1, 2, \ldots, L$, $k = 1, 2, \ldots, n_\alpha$ and satisfies that $\text{E}[x(0)] = \mu_0, \text{E}[(x(0) - \mu_0)(x(0) - \mu_0)^{\text{T}}] = p_0$.

The model in (156.1) and (156.2) can be rewritten in a more compact form as

$$x(t+1) = (A_0(t) + \Sigma_\alpha(t))x(t) + F(t)w(t) \tag{156.3}$$

$$y(t) = \Gamma(t)C(t)x(t) + D(t)v(t) \tag{156.4}$$

where

$$\Sigma_\alpha(t) = \sum_{k=1}^{n_\alpha} \alpha_k(t)A_k(t),$$

$$y(t) = \begin{bmatrix} y_1^{\text{T}}(t) & \ldots & y_L^{\text{T}}(t) \end{bmatrix}^{\text{T}}, v(t) = \begin{bmatrix} v_1^{\text{T}}(t) & \ldots & v_1^{\text{T}}(t) \end{bmatrix}^{\text{T}},$$

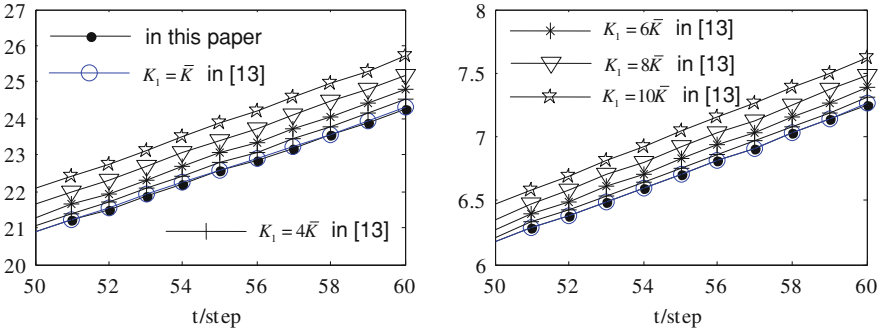


Fig. 156.2 Comparison of accumulated error squares of filter in this paper and [13] with different K_1, where $\bar{K}$ is the K_1 in [13]

$$C(t) = \left[C_1^{\mathrm{T}}(t) \quad \cdots \quad C_L^{\mathrm{T}}(t) \right]^{\mathrm{T}}, \Gamma(t) = \mathrm{diag}\left(\gamma_1(t) I_{m_1}, \ldots, \gamma_L(t) I_{m_L}\right), \\ D(t) = \mathrm{diag}\left(D_1(t), \ldots, D_L(t)\right) \tag{156.5}$$

where diag(•) denotes a diagonal matrix which diagonal elements are consist of •.

Further, we have the following noise statistical information

$$Q_v = \mathrm{E}[v(t)v^{\mathrm{T}}(t)] = (Q_{v_{ij}}),\ S = \mathrm{E}[w(t)v^{\mathrm{T}}(t)] = [S_1 \quad \ldots \quad S_L], \quad i,j = 1,2,\ldots,L \tag{156.6}$$

where $Q_{v_{ii}} = Q_{v_i}$.

We first introduce the following linear robust and resilient minimum variance unbiased filter in [13] for comparing the performance of the two filters.

$$\hat{x}(t+1) = A_0(t)\hat{x}(t) + \left(K_0(t) \sum_{j=1}^{n_\beta} \beta_j(t) K_j(t) \right) \left(y(t) - \bar{\Gamma} C(t) \hat{x}(t) \right) \tag{156.7}$$

where $\bar{\Gamma} = \mathrm{E}[\Gamma(t)]$ and $K_0(t)$ is the gain matrix. The initial value is $\hat{x}(0) = \mu_0$.

Remark 1 Stochastic parameters $\beta_j(t), j = 1, \ldots, n_\beta$ are introduced in the filtering gain to guarantee resilient operation. The additional gain term $\sum_{j=1}^{n_\beta} \beta_j(t) K_j(t)$ models the computation or implementation errors associated with the filtering gain. The family $\beta_j(t)$ are scalar white noise sequences of mean zero (meaning that errors are equally likely to result in smaller or larger values for the components of the gain matrix) and variance Q_{β_j} and are mutually uncorrelated.

Remark 2 In order to guarantee resilient operation, the $\beta_j(t)$ is introduced. But how to select the variance Q_{β_j} and coefficient $K_j(t)$ of stochastic perturbations $\beta_j(t)$ is not given. At the same time, different estimation performances are obtained by selecting different variances and coefficients which can be verified via the latter simulation research. Further, uncorrelated noises are assumed in the filter in [13].

All the above motivated us to find new filter $\hat{x}(t)$ for the state $x(t)$ based on the received measurements $(y_i(t), \ldots, y_i(0))$, $i = 1, 2, \ldots, L$. They are optimal in linear minimum variance sense for system with stochastic perturbations in the system matrix. Also, it is not necessary to introduce any artificially additional gain. Next, we will investigate the optimal linear filter based on the innovation analysis approach.

156.3 Optimal Linear Filter

Theorem 1 *For the system* (156.3) *and* (156.4) *satisfying Assumptions* 1 *and* 2, *the recursive filter is given by*

$$\hat{x}(t+1) = A_0(t)\hat{x}(t) + K(t)\varepsilon(t) \tag{156.8}$$

$$\varepsilon(t) = y(t) - \bar{\Gamma}C(t)\hat{x}(t) \tag{156.9}$$

$$K(t) = \left(A_0(t)P(t)C^{\mathrm{T}}(t)\bar{\Gamma} + F(t)SD^{\mathrm{T}}(t)\right)Q_{\varepsilon}^{-1}(t) \tag{156.10}$$

$$Q_{\varepsilon}(t) = \bar{\Gamma}C(t)P(t)C^{\mathrm{T}}(t)\bar{\Gamma}^{\mathrm{T}} + \Xi \odot \left(C(t)X(t)C^{\mathrm{T}}(t)\right) + D(t)Q_vD^{\mathrm{T}}(t) \tag{156.11}$$

$$X(t+1) = A_0(t)X(t)A_0^{\mathrm{T}}(t) + \sum_{k=1}^{n_\alpha} Q_{\alpha_k}A_k(t)X(t)A_k^{\mathrm{T}}(t) + F(t)Q_wF^{\mathrm{T}}(t) \tag{156.12}$$

$$\begin{aligned} P(t+1) &= [A_0(t) - K(t)\bar{\Gamma}C(t)]P(t)[A_0(t) - K(t)\bar{\Gamma}C(t)]^{\mathrm{T}} \\ &+ \sum_{k=1}^{n_\alpha} Q_{\alpha_k}A_k(t)X(t)A_k^{\mathrm{T}}(t) + K(t)\left[\Xi \odot \left(C(t)X(t)C^{\mathrm{T}}(t)\right)\right]K^{\mathrm{T}}(t) \\ &+ [F(t) \quad -K(t)D(t)]Q[F(t) \quad -K(t)D(t)]^{\mathrm{T}} \end{aligned} \tag{156.13}$$

where $\odot$ *is the Hadamard product defined by* $B \odot C = (b_{ij} \times c_{ij})_{r\times s}$, *where* $B = (b_{ij})_{r\times s}$ *and* $C = (c_{ij})_{r\times s}$. $\Xi = diag(\pi_1(1-\pi_1)I_{m_1}, \ldots, \pi_L(1-\pi_L)I_{m_L})$, $Q = \begin{bmatrix} Q_w & S \\ S^{\mathrm{T}} & Q_v \end{bmatrix}$, $\varepsilon(t)$ *is the innovation sequence with covariance* $Q_\varepsilon(t)$, $K(t)$ *is the filtering gain matrix*, $P(t+1)$ *is the filtering error covariance matrix, and* $X(t)$*is the state covariance matrix of system* (156.3) and (156.4). *The initial values are* $\hat{x}(0) = \mu_0$, $X(0) = \mu_0\mu_0^{\mathrm{T}} + p_0$ *and* $P(0) = p_0$.

Proof From projection theory [14], we can easily obtain (156.8) and (156.9), where the filtering gain matrix is defined by

$$K(t) = \mathrm{E}\left[x(t+1)\varepsilon^{\mathrm{T}}(t)\right]Q_{\varepsilon}^{-1}(t) \tag{156.14}$$

Substituting (156.4) into (156.9), the innovation $\varepsilon(t)$ can be rewritten as

$$\varepsilon(t) = [\Gamma(t) - \bar{\Gamma}]C(t)x(t) + \bar{\Gamma}C(t)\tilde{x}(t) + D(t)v(t) \tag{156.15}$$

Substituting (156.3) and (156.15) into (156.14), and using $\mathrm{E}[\Gamma(t) - \bar{\Gamma}] = 0$, $\mathrm{E}[\sum_{\alpha}(t)] = 0$, $\tilde{x}(t) \perp w(t)$ and $x(t) \perp v(t)$ where symbol $\perp$ denotes orthogonality, we can obtain (156.10). Further, the innovation covariance matrix (156.11) can be obtained from $Q_{\varepsilon}(t) = \mathrm{E}[\varepsilon(t)\varepsilon^{\mathrm{T}}(t)]$ by using the uncorrelation of $\tilde{x}(t)$, $x(t)$ and $w(t)$.

By substituting (156.3) into $X(t) = \mathrm{E}[x(t)x^{\mathrm{T}}(t)]$, we can obtain the state covariance matrix $X(t)$ is computed by (156.12).

Subtracting (156.8) from (156.3) and using (156.15), we have the one-step predictor error equation

$$\begin{aligned}\tilde{x}(t+1) = x(t+1) - \hat{x}(t+1) = [A_0(t) - K(t)\bar{\Gamma}C(t)]\tilde{x}(t) - K(t)(\Gamma(t) - \bar{\Gamma})C(t)x(t) \\ + \Sigma_{\alpha}(t)x(t) + F(t)w(t) - K(t)D(t)v(t)\end{aligned} \tag{156.16}$$

Noting $\tilde{x}(t) \perp w(t)$, $\tilde{x}(t) \perp v(t)$, $\mathrm{E}[\Gamma(t) - \bar{\Gamma}] = 0$, and $\mathrm{E}[\sum_{\alpha}(t)] = 0$, (156.13) is obtained.

Corollary 1 *Theorem* 1 *is reduced to the case that* $w(t)$ *and* $v_i(t)$ *is independent of each other when the cross-covariance matrices take the value of* 0, i.e., $S_i = 0$, $i = 1, 2, \ldots, L$.

Remark 3 In fact, the filter in [13] is equal to the corollary 1 when $Q_{\beta_j} = 0$, $K_j(t) = 0$, $j = 1, \ldots, n_{\beta}$. In other words, our filter has the simple form and smaller computation load than the filter in [13]. Further, the accuracy of the former is very close to that of the latter via simulation research.

Remark 4 From Theorem 1, it can be easily observed that our filter (156.8) is reduced to the standard Kalman filter when $\alpha_k(t) = 0, k = 1, 2, \ldots, n_{\alpha}$, and $\gamma_i(t) = 1, i = 1, 2, \ldots, L$.

156.4 Optimal Linear Steady-State Filter

In the preceding section, we have obtained the optimal linear filter. In this section, we will investigate the steady-state filtering problem. We shall consider the corresponding time-invariant system (156.3) and (156.4), where $A_0, A_k, k = 1, 2, \ldots, n_{\alpha}$, F, C and D are constant matrices.

Theorem 2 *For the corresponding time-invariant system* (156.3) *and* (156.4), if $\rho(A_0 \otimes A_0^{\mathrm{T}} + \sum_{k=1}^{n_{\alpha}} Q_{\alpha_k} A_k \otimes A_k^{\mathrm{T}}) < 1$ *and* A_0 *is stable, the solutions* $X(t)$ *and* $P(t+1)$*with any initial values* $X(0)$ and $P(0) \geq 0$ *converge, respectively, to the unique positive semi-definite solutions* X and Σ *of the following Lyapunov and algebraic Riccati equations*:

$$X = A_0 X A_0^{\mathrm{T}} + \sum_{k=1}^{n_{\alpha}} Q_{\alpha_k} A_k X A_k^{\mathrm{T}} + F Q_w F^{\mathrm{T}} \tag{156.17}$$

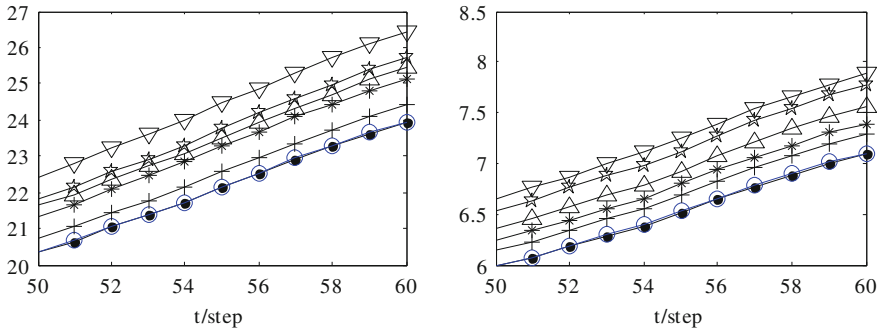


Fig. 156.3 Comparison of accumulated error squares of filter in this paper and [13] with different Q_{β_1}:$Q_{\beta_1} = 0.01$—○— [13], —•— this paper; $Q_{\beta_1} = 0.04$—+— [13], $Q_{\beta_1} = 0.08$—*— [13], $Q_{\beta_1} = 1.2$—△— [13], $Q_{\beta_1} = 1.6$—☆— [13], $Q_{\beta_1} = 2$—▽—, [13]

$$\begin{aligned}\Sigma &= [A_0 - K\bar{\Gamma}C]\Sigma[A_0 - K\bar{\Gamma}C]^{\mathrm{T}} + \sum_{k=1}^{n_\alpha} Q_{\alpha_k} A_k X A_k^{\mathrm{T}} \\ &+[F \quad -KD]Q[F \quad -KD]^{\mathrm{T}} + K[\Xi \odot (CXC^{\mathrm{T}})]K^{\mathrm{T}}\end{aligned} \tag{156.18}$$

where $\rho(\bullet)$ *is the spectrum radius of matrix* $\bullet$ *and* $\otimes$ *is the Kronecker product.*

Moreover, we have that $K = \lim_{t\to\infty} K(t)$, $\Sigma = \lim_{t\to\infty} P(t)$, $Q_\varepsilon = \lim_{t\to\infty} Q_\varepsilon(t)$. Furthermore, the steady-state filter

$$\hat{x}(t+1) = (A_0 - K\bar{\Gamma}C)\hat{x}(t) + Ky(t) \tag{156.19}$$

is asymptotically stable.

Proof Form $\rho(\bullet) < 1$, we have $X = \lim_{x\to\infty} X(t)$ [15]. Furthermore, the stability of A_0 means that corresponding time-invariant system is detectable and stabilizable. Then, a similar proof as in [16] can be applied to show $X(t)$ and $P(t)$ of equations (156.12) and (156.13) with any initial conditions $X(0)$ and $P(0) \geq 0$ converge to the unique positive semi-definite solutions X and Σ of (156.17) and (156.18), and $(A_0 - K\bar{\Gamma}C)$ is a stable matrix, which implies the stability of the steady-state filter (156.19).

156.5 Simulation Research

Consider the stochastic system as in [13] with two sensors, where $v_i(t)$ and $w(t)$ is independent of each other and all the parameters are the same as in [13], i.e., $A_0 = \begin{bmatrix} 0.3 & 0.7 \\ 0.2 & 0.6 \end{bmatrix}$, $A_1 = \begin{bmatrix} 0.1 & 0.05 \\ 0.2 & 0.1 \end{bmatrix}$, $F = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$, $C_1 = [0.5 \quad 1]$, $C_2 = [1 \quad 1]$, $D_1 = 0.5$, $D_2 = 0.2$, $Q_w = 0.3$, $Q_{v_1} = Q_{v_2} = 0.3$, $Q_\alpha = 0.1$, $Q_{\beta_1} = 0.01, K_1 = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & -0.1 \end{bmatrix}$ and $n_\beta = 1$. Our aim is (a) to compare the accuracy of the two filters

$\hat{x}(t+1)$ and (b) to compare the accuracy of the filter in [13] with different Q_{β_1} and $K_1(t)$.

Figures 156.1, 156.2 and 156.3 show the comparison of estimation accuracy of our filter and the filter in [13] over an average of 200 runs of Monte Carlo method. From Fig. 156.1, we can see that the accuracy of our filter and the filter in [13] is almost coincident. From Figs. 156.2, 156.3, we can see that the performance of the filter in [13] becomes worst as Q_{β_1} and $K_1(t)$ increase.

156.6 Conclusion

For the discrete-time stochastic system subject to multiple sensors of different failure rates and multiplicative noise, the optimal linear filter has been developed via an innovation analysis approach. Differently from [13] where the stochastic perturbations are allowed in the filtering gain to guarantee resilient operation, here the optimal filter is designed by applying the innovation analysis method. Simulation comparison shows that the accuracy of our optimal filter is very close to that of the robust filter in [13]. Further, the performance of the filter in [13] becomes worse as Q_{β_1} and $K_1(t)$ increase. The steady-state filter has also been investigated. A sufficient condition for the convergence of the optimal linear filter has been obtained.

Acknowledgments This work was supported by Natural Science Foundation of China under Grant NSFC-60874062, Key Project of Chinese Ministry of Education under Grant 209038, Program for New Century Excellent Talents in University under Grant NCET-10-0133, 1154-NCET-01, Program for High-qualified Talents under Grant Hdtd2010-03, Province Key Laboratory, and Foundation of Heilongjiang Education Committee under Grant No. 12511431.

References

1. Matveev AS, Savkin AV (2003) The problem of state estimation via asynchronous communication channels with irregular transmission times. IEEE Trans Autom Control 48:670–676
2. Ray A, Liou LW, Shen JH (1993) State estimation using randomly delayed measurements. J Dyn Syst Measurement Contr 115:19–26
3. Yaz E, Ray A (1998) Linear unbiased state estimation under randomly varying bounded sensor delay. Appl Math Lett 11:27–32
4. Nahi N (1969) Optimal recursive estimation with uncertain observation. IEEE Trans Inf Theory 15:457–462
5. Hadidi MT, Schwartz SC (1979) Linear recursive state estimators under uncertain observations. IEEE Trans Autom Control 24:944–948
6. Sahebsara M, Chen TW, Shah SL (2007) Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations. Int J Contr 80:292–301
7. Sun SL, Xie LH, Xiao WD, Soh YC (2008) Optimal linear estimation for systems with multiple packet dropouts. Automatica 44:1333–1342

8. Nakamori S, Caballero-Aguila R, Hermoso-Carazo A, Linares-Perez J (2003) Linear recursive discrete-time estimators using covariance information under uncertain observations. Sign proces 83:1553–1559
9. Wang ZD, Huang B (2000) Robust $H_2/H\infty$ filtering for linear systems with error variance constraints. IEEE Trans Autom Control 48:2463–2467
10. Yang FW, Wang ZD, Hung YS (2002) Robust Kalman filtering for discrete time-varying uncertain systems with multiplicative noises. IEEE Trans Autom Control 47:1179–1183
11. Xie LH, Soh YC, De Souza CE (1994) Robust Kalman filtering for uncertain discrete-time systems. IEEE Trans Autom Control 39:1310–1314
12. Shaked U, De Souza CE (1995) Robust minimum variance filtering. IEEE Trans Sign Proces 43:2474–2483
13. Hounkpevi FO, Yaz EE (2007) Robust minimum variance linear state estimators for multiple sensors with different failure rates. Automatica 43:1274–1280
14. Wang ZD, Ho DWC, Liu XH (2004) Robust filtering under randomly varying sensor delay with variance constraints. IEEE Trans Circuits Syst II 51:320–326
15. De Koning WL (1984) Optimal estimation of linear discrete-time systems with stochastic parameters. Automatica 20:113–115
16. Anderson BDO, Moore JB (1979) Optimal filtering. Prentice-Hall, Englewood Cliffs

Chapter 157
Study on Timing Sequences of Interface Program of Digital Temperature Sensor DS18B20

Huang Wenli and He Linlin

Abstract DS18B20 belongs to a kind of single-bus digital temperature sensor with corresponding simple hardware circuit connection. As a result it is necessary to make use of software to provide complicated and accurate timing logical sequences to realize data transfer correctly between the sensor and host computer, and finally to ensure the accuracy of data the host receives. It is the most foundational element to provide accurate timing logical sequences making use of some kind of relatively complicated software structure for data transfer of the single-bus digital sensors.

Keywords DS18B20 · Temperature sensor · Timing logical sequences · Software structure

157.1 Introduction

DS18B20 is an improved version of single-bus digital temperature sensor DS1820 released by the American Dallas Semiconductor Corporation. Due to the single-bus structure, the corresponding hardware circuit of DS18B20 becomes pretty simple, as a result only one wire is needed to transmit temperature data, simultaneously the working power source may also be achieved through this wire

H. Wenli (✉) · H. Linlin
School of Mechanical and Electrical Engineering, Zhengzhou Institute of Aeronautical Industry Management, Daxue Middle Road 2, Zhengzhou 450010, China
e-mail: hwl@zzia.edu.cn

H. Linlin
e-mail: helinlin@zzia.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_157,

instead of more other wires [1, 2]. Based on the simple hardware circuit, the relatively complex software structure should be designed to provide reasonable logical timing sequences for data transmission of DS18B20 so as to guarantee the accuracy and reliability that the sensor of DS18B20 works [3, 4]. This chapter thoroughly analyzed the logical timing sequences in the process of data transmission of single-bus digital temperature sensor DS18B20, in order that the users of this sensor can refer the fruits to promote the settlement of temperature measurements under different circumstances and to enhance the corresponding precision of temperature measured.

157.2 Basic Features and Makeup of DS18B20

The single-bus structure is the primary feature of DS18B20. As a result the device has characteristics of simple hardware circuit connection, strong anti-jamming ability of digital data signals and wide temperature measurement range of −55 °C to +125 °C and so on. The measuring precision, up to 0.0625 °C, can be changed by software to deal with different actual circumstances. Due to sole serial number each sensor occupies, multiple sensors can be articulated on one bus to adapt for distributed temperature measurements. There are two kinds of power supply for the sensor: one is from external power supply, the other is parasitic power supply through single-bus wire which is able to save connecting wires as well as expand operating voltage range to 3 or 5.5 V. The time executing one 12-bit temperature signal conversion is very short, only 750 ms. The sensor has negative voltage property, namely for reversed voltage polarity the sensor would not be destroyed owing to high heat, but does not work as usual.

For DS18B20, the external wire connection is very simple and the A/D transformation of data, the production of working logical timing sequences, etc. are completed within the sensor. The internal structure of the sensor mainly consisted of four parts: (1) a 64-bit lithographic ROM used to store the serial number of the sensor; (2) a temperature sensor used to measure temperature; (3) the non-volatile temperature alarm triggers TH and TL into which the upper and lower thresholds of temperature alarm can be input; (4) the high-speed scratchpad storage which works with setting the resolution that temperature transforms, temporarily storing temperature value to wait for the host computer to read, estimating the validity of transformed temperature value and resetting the temperature alarm sign inside sensor, etc. [4, 5].

The host computer operates the sensor through the ROM and the scratchpad commands which are used commonly are showed in Table 157.1. Relevant commands present in the form of a hexadecimal number with 8-bit word length.The host computer that controls the sensor must begin with reading ROM, then operating the scratchpad storage to accomplish the delivery of relevant codes and the measurement of temperature.

Table 157.1 The common operating command codes used by DS18B20

Operations	Command codes	Functional interpretations of codes
ROM operation	Read ROM [33h]	Read serial number. Errors would be caused if there are more than one sensor on a single bus
	Match ROM [55h]	Locate the sensor. The command is followed by a 64-bit data sequence to match some sensor from many sensors tied on a bus
	Skip ROM [CCh]	Skip the ROM operation. The command permits the host computer to perform the scratchpad operation straightly to save time, which is only suitable for single point bus system, instead, errors would occur for multi-point bus system
Scratchpad operation	Write scratchpad [4Eh]	Write scratchpad. This command permits to input two bytes of data into the sensor. A reset pulse can be sent out to stop writing at any time
	Read scratchpad [BEh]	Read scratchpad. The command reads the temperature value from the sensor. A reset pulse would be used to stop reading at any time
	Convert T [44h]	Convert the temperature. This command lets the sensor launch temperature A/D transformation once, hereafter wait for the host to read

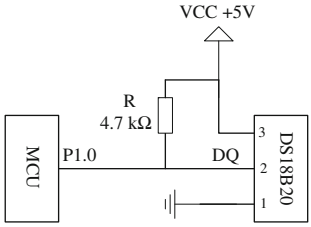


Fig. 157.1 The electric schematic connecting the microcontroller with DS18B20

157.3 Digital Signal Transmitting Timing Sequences of DS18B20

It is very prominent to execute temperature measurement based on single-bus structure for DS18B20; however, the strict timing sequences must be complied with because the correct timing sequences are basic premise of achievement of correct data. The general steps of data transmission between the host and the sensor include: the host starts a reset pulse to initialize the sensor, then ROM command, followed by scratchpad command and finally the sensor receives data to finish data transmission one time. Except for the correct operating steps, the correct time gaps, viz. time slots, must also be guaranteed between the steps. There are three kinds of communications with the host for DS18B20: initialization, reading and writing, all of which must comply with the strict time slots. So the following, respectively, discusses the time slots of the above three in depth. Only one sensor on a connecting wire, as shown in Fig. 157.1, is the case expounded here in order to clarify logical relationships of time slots as a main problem.

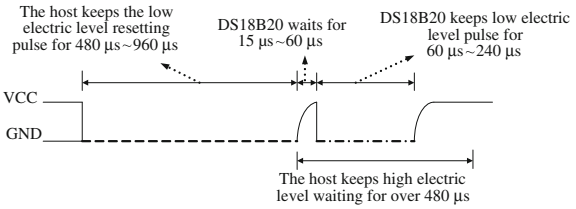


Fig. 157.2 The initializing timing slots of DS18B20

157.3.1 The Initializing Timing Sequences

The initializing operation, as one of bottom basic operations of the sensor, launches the sensor. The initiating pulse is equal to a responding password by which a communicating bridge is established between the host and the sensor to be ready for the next operation. The initiating pulse is made up of a reset pulse sent out by the host and then an existing pulse coming from the sensor. The reset pulse can be regarded as a request for the sensor from the host and the subsequent existing pulse as a response to the host from the sensor. By the request and the response the communicating channel between the host and the sensor is established to prepare for next data transmission, viz. data reading or writing.

The logical timing slots of the initializing pulse are shown in Fig. 157.2. The initiating reset pulse consists of the following: first, the host sets data line DQ to zero with time length of 480–960 μs as the dashed line shows in Fig. 157.2, which is the pulse signal that the host requests the sensor DS18B20; then, the host sets one to induce DQ line one set by pull-up resistor with time length of at least 480 μs. Due to only one line DQ between the host and the sensor it is evident as long as either of both ends of the line DQ is zero, then the line DQ would be zero, viz. the device connected with one line would be AND relationship in logic. Consequently, it may be seen, only after the host is set one, the host would be at the situation of receiving data. Hereafter, the decisive right of electric potential of line DQ is given to the sensor, and the host waits for the response of the sensor; finally, after receiving the requesting pulse, the host set zero with time length of over 480 μs, and the sensor waits for a time lag of 15–60 μs after the host converses from zero to one, viz. during the duration line DQ, resting with the pull-up resistor, keeps one for 15–60 μs. After this, the sensor knows the host is ready and correspondingly sends out the response signal of a low electric potential pulse of time lag 60–240 μs, as the dotted line shows in Fig. 157.2. After the response pulse keeps 60–240 μs, the sensor sets one to induce the line DQ one set by pull-up resistor again. So far, a communicating bridge between the host and the sensor is established to prepare for next data transmission. Easy to know from the above, while the initializing process has been completed, both the host and the sensor are set one and the line DQ is set one by pull-up resistor, as a result the host and the sensor are in conditions which can be read, and also be written.

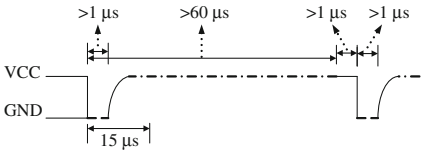


Fig. 157.3 The reading timing slots of DS18B20

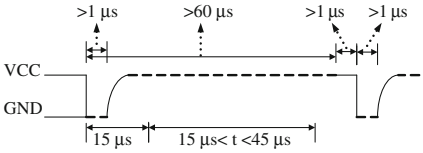


Fig. 157.4 The writing timing slots of DS18B20

157.3.2 The Reading Timing Sequences

The data-read operation is one of bottom basic operations of the sensor. All data are read in a way that the least significant bit is first. The common procedures for reading data include: first, the host sends out a zero electric potential pulse signal of time lag of over 1 μs to launch data transmission when the line DQ is set one by pull-up resistor, after the sensor is initialized. Hereafter, the host sets one with time length of over 60 μs to be able to receive data; then, within 15 μs after receiving the launching pulse signal, the sensor sends temperature data to the line DQ for the host to read; finally, as close to 15 μs the host collects data on the line DQ to get 1-bit binary numbers of 8-bit binary representation of temperature. One-by-one by analogy, the 8-bit binary numbers, viz. the temperature values at some time, are achieved in the way that the least significant bit is the first. Figure. 157.3 shows the data-read timing slots for DS18B20, in which the dashed line is the role of the host, that is, the electric potential of line DQ is up to the host, and the arc line is transition period, and the dotted line is the role of the sensor. It can be seen although the action of reading data is completed once in only 15 μs, the duration of reading data is beyond 60 μs. Moreover there must be a time gap of at least 1 μs between both consecutive reading periods, within which both the host and the sensor set one and the line DQ is set one by the pull-up resistor.

157.3.3 The Writing Timing Sequences

The data-write operation is also one of the underlying basic operations. All data are written in the way that the least significant bit is the first. The basic steps of writing data consist of: first, when the line DQ is set one by pull-up resistor after the sensor has been initialized, the host sends out a zero electric potential pulse with time length of over 1 μs as launching signal of data transmission; then, the

host sends due-out signal of zero or one to the line DQ within following 15 μs and waits for the sensor to read; finally, after the sensor has received the launching signal for 15 μs, the sensor would finish data collecting within next 45 μs, that is, the due-out signal of the host would be kept at least 60 μs on the line DQ. And so on, the commands consisted of binary numbers are input into the sensor bit by the way that the least significant bit is first. The data-write timing slots for DS18B20 are shown in Fig. 157.4 in which the dashed line is a period of the host and the arc line for a period of transition. It is evident from Fig. 157.4 that the number value is one to have been input into the sensor. It may be seen that a writing cycle is at least 60 μs and there must be a time gap of not less than 1 μs when both the host and the sensor set one and consequently the line DQ is set one through the pull-up resistor.

157.5 Conclusions

Based on the single-bus structure, the sensor DS18B20 only needs one data wire to finish relevant hardware circuit connection with the host and then is enough to finish corresponding signal acquisition, A/D conversion, temperature measurement and so on functions. A strict communicating agreement must be complied with to guarantee accurate data transmission for electric devices characterized by the single-bus structure due to their limited communicating channels with the host. For DS18B20, the correct communicating agreement is constituted by strict operating steps and correct timing slots between two consecutive steps, which are presented and realized by corresponding software program structure.

References

1. Jin WZ (2000) The principle and application of the single wire digital temperature sensor. Appl Electron Tech 6:66–68
2. Ni TL (2010) Application on the temperature and humidity measurement and control system of the single-bus sensor DHT11. Microcontrollers Embeded Syst 6:60–62
3. Gu ZY, Liu LY, Du ZH (2002) DS18B20 C language programming interface. Microcontrollers Embeded Syst 7:22–24
4. Liu JT, Mao SK (2005) Principle of DS18B20 and its interface design based on C. Instrum Meters User 12(6):138–140
5. Zhao ZH, Zhao XW (2003) Principle and application of DS18B20 intellect thermometer. Mod Electron Tech 7:32–34

Chapter 158
A Novel Purely Digital Temperature and Humidity Monitoring Control System

Wenli Huang and Jianfeng Cui

Abstract A kind of novel purely digital temperature and humidity controller is introduced in-depth. The temperature signal is received from digital sensor DS18B20 of single-bus structure made by Dallas Corporation in America, and the humidity signal from the sensor DHT11 similar to DS18B20 in structure made by Aosong Corporation in Guangzhou, China. The hardware electric schematics, including single-chip microcomputer (SCM) circuit, temperature and humidity gauging circuit, nixietube displaying circuits, alarm circuit, and the corresponding software design, including data transfer of DS18B20 and data transfer of DHT11, are also introduced.

Keywords DS18B20 · DHT11 · Single-bus · Digital signal

158.1 Introduction

Both temperature and humidity parameters need to be supervised under many circumstances such as the vegetable cropping in greenhouse, the file storage, the chemical fiber processing, and the highly accurate experiments. The strict temperature and humidity control is very important condition and basic safeguard for product qualification, production safety, test accuracy, and operation success [1, 2].

W. Huang (✉) · J. Cui
School of Mechanical and Electrical Engineering,
Zhengzhou Institute of Aeronautical Industry Management,
Daxue Middle Road 2, Zhengzhou 450015, China
e-mail: hwl@zzia.edu.cn

J. Cui
e-mail: cuijianfeng@zzia.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_158,

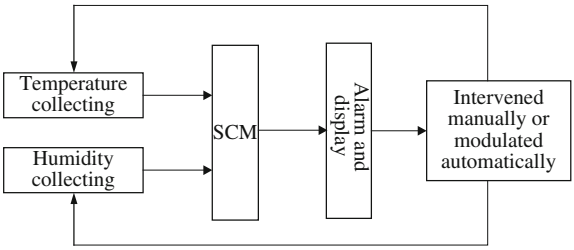


Fig. 158.1 Structure frame chart of temperature and humidity controller

The temperature and humidity control system can be implemented based on many precepts such as CAN bus and integrated circuit, single chip microcomputer (SCM) and microcomputer system, and analog sensor and digital sensor [3–5]. Each temperature and humidity monitoring control system has its own characteristic to satisfy different users and actual applications in different conditions. This paper introduced a novel purely digital temperature and humidity controller based on the SCM STC89C52, the single-bus digital temperature sensor DS18B20 made in American Dallas Corporation and the humidity sensor DHT11 similar to DS18B20 in structure made in Chinese Aosong Corporation. The temperature and humidity controller has higher application and promotion worth due to its simple structure, low cost, high accuracy, strong antijamming ability, etc.

158.2 Hardware Circuit Design

158.2.1 SCM Circuit Design

This control system, centered around the SCM STC89C52, has constructed the serial port junction circuit, alarm circuit, temperature and humidity monitoring circuit, nixietube displaying circuit, and so on modules to compose a whole temperature and humidity monitoring controller, as structure frame chart shown in Fig. 158.1. The logical processes that temperature and humidity are measured and monitored are as follows: the digital temperature and humidity signals output from temperature and humidity sensors are input into SCM separately and analyzed; then if the temperature and humidity values exceed a set threshold one, the alarm circuit starts a buzzer to output an alarm voice to remind human's intervention, or initiate a relevant circuit expanded easily based on this controller to realize self-regulated temperature and humidity. The functional connections of all pins of the SCM STC89C52 of this controller are shown in Table 158.1.

158.2.2 Temperature and Humidity Gauging Circuit Design

Temperature is measured making use of the temperature sensor DS18B20 of American Dallas Corporation, which is characteristic of two kinds of power supply

Table 158.1 The functional electric connections of all pins of the SCM STC89C52

Pins of STC89C52	Functions or external electric components connected	Notes
P1.0	DS18B20 DATA	Data wire of DS18B20
P1.1	DHT11 DATA	Data wire of DHT11
P1.2	Buzzer	
XTAL1, XTAL2	A external crystal oscillator 12 MHz	
GND	Grounding	
VCC	A power supply +5 V	
RST	SCM reset	Reset with a high electric potential
P0.0–P0.7	Segment-chosen signals of every nixietube	A parallel connection of 8 nixietubes
P2.0–P2.7	Bit-chosen signals of each nixietube	
RXD,TXD	The input–output wires in serial connection	

modes adapted to a wide voltage range; the single-bus data wire interface fashion which does not require any peripheral components in use to achieve highly accurate temperature measurement of programable 9–12 bits; temperature measured is directly output by digital signals with fast measuring speed as well as strong anti-jamming error-correction ability; a wide temperature measuring range of −55 to 125°C with resolution of 0.5°C; etc.

Humidity is measured by temperature and humidity sensor DHT11 of Aosong Corporation in Guangzhou, which is characteristic of as follows: a single-bus technology, small size, low power consumption, fast speed-responding, strong anti-jamming ability, simple operation and control, and high performance-to-price ratio; signal transmitting distance of up to 20 m; relatively wide humidity measuring range of from 20 to 90% with an accuracy of ±5%, and temperature measuring range of from 0 to 50°C with an accuracy of ±1°C and a response time of less than 5 s. It can be seen, due to the narrower range of temperature output of DHT11, the controller takes advantage of the temperature sensor DS18B20 to expand temperature measuring range stead of the temperature output of DHT11, but the humidity output signal taken from DHT11.

Both temperature and humidity signals are achieved from digital sensors, thus greatly simplifying the hardware circuit design as well as maintaining a good monitoring performance with low cost, which reflects high application value of the control system. An electric schematic of temperature and humidity sensors of the controller is shown in Fig. 158.2.

158.2.3 Nixietube Displaying Circuits

Temperature and humidity are displayed, respectively, by use of four eight-segment nixietubes. The connection of common cathode is adopted for every

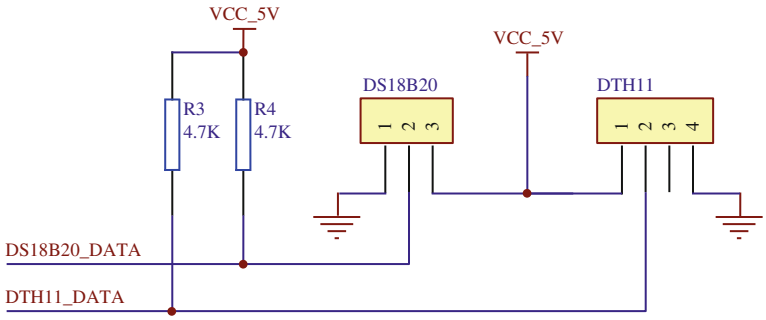


Fig. 158.2 An electric schematic of temperature and humidity sensors

nixietube, and a PNP triode is connected with each nixietube to enhance relevant driving power for the connection of common cathode of several nixietubes. Some kind of dynamic display of temperature and humidity: the same segment-code signal for every nixietube at one time, and only one bit-code signal is valid for some nixietube, that is, at any moment only one nixietube can be allowed to display. Each temperature and humidity values become extinguished immediately after displayed, however, displaying for many times in a very short time results in continuous existence of temperature and humidity values based on human persistence of vision. Eight segment codes of each of eight nixietubes are connected in parallel with pins of P0.0–P0.7 of microcontroller, and the bit code of each of eight nixietubes is connected respectively with P2.0–P2.7 of microcontroller to realize time-sharing display of eight nixietubes. In order to save book space and indicate clearly, Fig. 158.3 only shows electric connection of one of the eight nixietubes with similar electric structure.

158.2.4 Alarm Circuit

The buzzer alarming circuit connected with P1.2 of the microcontroller is connected in series with a PNP triode to enhance corresponding driving power, as shown in Fig. 158.4. The buzzer launches alarm when the pin of P1.2 is low electric level.

158.3 Software Development

Because of very simple hardware electric connection of the single-bus digital temperature and humidity sensors the measurement, control, and display of their temperature and humidity signals have to be implemented by relatively

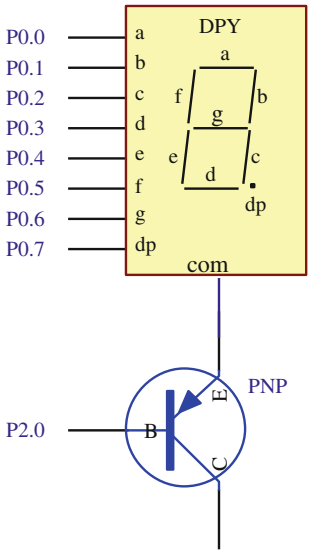


Fig. 158.3 Electric schematic of a nixietube

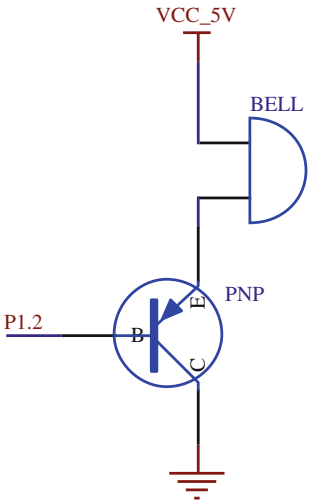


Fig. 158.4 Electric schematic of the buzzer

sophisticate software structure. Through mere a data line DATA, the communication between sensors and the host must comply with strict protocols to complete data transfer and information exchange, in accordance with established procedures and timing sequences. On the premise of stable operating sequences or procedures the software design is very flexible and complicated, because the timing slot between two procedures is a number value range. The corresponding interface program can include many flexible forms which need be probed into deeply, in practice.

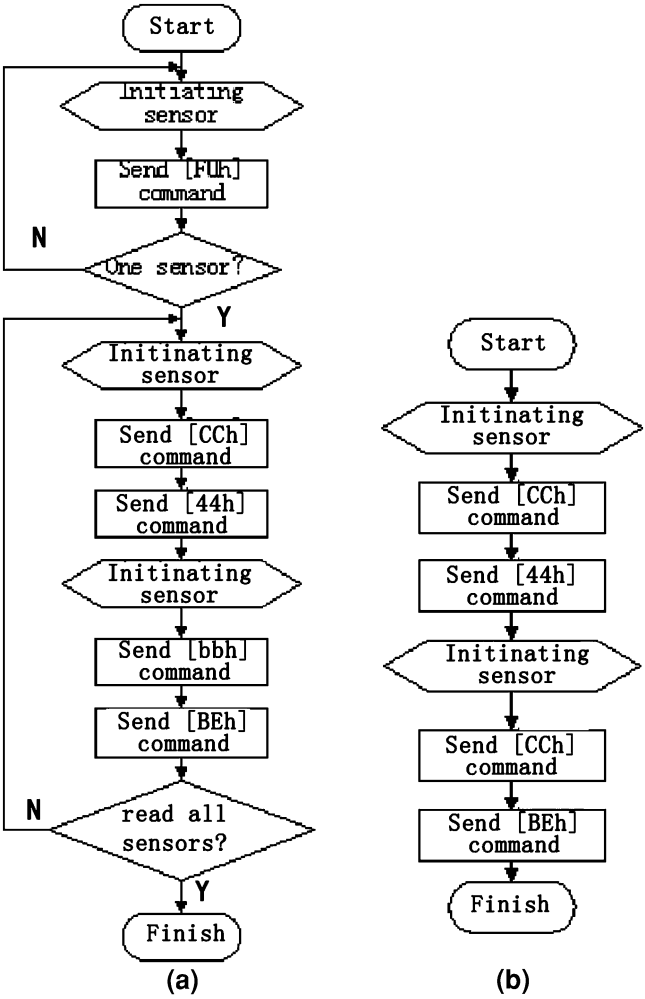


Fig. 158.5 The flow chart operating DS18B20

158.3.1 Data Transfer of DS18B20

The host is connected with DS18B20 through single data line DATA which presents high electric level at free time, so that the sensor may either read data or write data. To Operate DS18B20 mainly includes two kinds: read data and write data, which appear in the form of ROM operating commands or memory operating commands. The commands used commonly include: [F0h] which chooses some sensor from many sensors connected in one data line; [33h] reads the serial number of some sensor based on the command of [F0h]; [55h] locates some sensor; [CCh] skips ROM operation; [4Eh] writes the memory; [BEh] reads the memory; [44h]

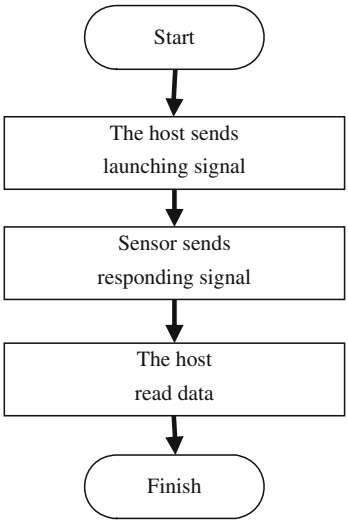


Fig. 158.6 The flow chart operating DHT11

completes temperature A/D conversion, etc. (see [6] in detail). The flow chart of operating the sensor is shown in Fig. 158.5 in which the chart (a) is one instance that there are many sensors to be connected with one data line; the chart (b) is the other instance that there is only one sensor for one data line.

For Fig. 158.5a, it is necessary to determine operating object, that is, which sensor should be executed, prior to dominating any sensor on account of many sensors on single bus. The identification of sensors is resorting to their own sole series number. Compared with Fig. 158.5a, b is that there is only one sensor on single bus. Due to only one sensor the procedure of identifying sensors may be omitted, consequently the operating procedures would be simplified greatly. Figure 158.5 merely shows the basic operating flow procedures. What needs to be paid attention to in particular is, besides essential timing sequences among all steps, the strict timing slot between two consecutive steps, too, must be obeyed so as to accomplish valid performance. That is, the function of each operating step is carried out without exception based on a group of TTL electric level pulses with strict timing sequences and timing slots. The interface program of temperature sensor DS18B20 and microcontroller may refer to [6, 7].

158.3.2 Data Transfer of DHT11

Similar to DS18B20, there is only single bus, too, between the host and DHT11, and the single bus is at high electric level to ensure that the sensor can either read data or write data. If no start signal, the sensor cannot carry on temperature and humidity measurement on its own initiative. After receiving start signal, the sensor converts to high power loss mode from low one to accomplish acquisition and measurement of temperature and humidity once. Immediately following the start

signal of the host, the sensor sends out responding signal and then delivers temperature and humidity values to the host, next resumes the low power loss mode again to prepare for next temperature and humidity measurement. The flow chart of temperature and humidity measurement making use of DHT11 is shown in Fig. 158.6. Similarly each step must obey strict timing sequences and timing slot to ensure proper function. The relevant program can be seen in [8].

Both DS18B20 and DHT11 are digital sensors of single-bus structure; however, compared with DS18B20, the information delivery between the host and DHT11 is through TTL electric level pulses with different time length instead of operating commands identified by different hexadecimal notation groups. In addition, DHT11 is only suitable for temperature and humidity measurement of single bus and single point conditions, correspondingly the measurement is simple.

158.4 Conclusions

A kind of new purely digital temperature and humidity monitoring system of integrating comprehensive excellences of temperature sensor DS18B20 and humidity sensor DHT11 is introduced in this paper. The system achieves real-time monitoring environmental temperature and humidity, has characteristics of simple structure, high performance-to-price ratio, strong anti-interference ability, and so on, and manifests higher practical worth. In addition, the microcontroller used in the system supports online programing named ISP/IAP function; therefore, it is very simple to revise temperature and humidity threshold value, which lets the system flexibly apply with a variety of practical applications with higher scalability.

References

1. Lin YF (2003) An inspecting system of temperature and humidity measurement fireproofing and guard against theft in grain depots. China Agricultural University Library, China
2. Fu K, Peng CW, Yuan XF (1997) Design and application of a kind of microcontroller of automatically monitoring temperature, humidity and fire, theft alarm. Appl Electron Tech 2:23–24, 28
3. Liu Y, Jin TD (2010) Design of intelligent system for humidity and temperature acquisition based on CAN-bus technology. Mod Electron Tech 1:126–128
4. Xia FL (1999) The design of a temperature humidity controller based on the AT89C51 monolithic microcontroller. Ind Instrum Autom 6:32–34
5. Feng XY, Ge RY (2006) The temperature/humidity measurement and control system based on digital temperature/humidity sensor SHT11. Process Autom Instrum 1:62–64
6. Liu JT, Mao SK (2005) Principle of DS18B20 and its interface design based on C. Instrum Meters User 12(6):138–140
7. Jin WZ (2000) The principle and application of the single wire digital temperature sensor. Appl Electron Tech 6:66–68
8. Ni TL (2010) Application on the temperature and humidity measurement and control system of the single-bus sensor DHT11. Microcontrollers Embeded Syst 6:60–62

Chapter 159
Design of Counting System Based on Pyroelectric Infrared Sensor

Yanhua Zhang, Lu Yang and Ni Jing

Abstract In order to facilitate to get real-time statistic number of people in and out of the building, based on pyroelectric infrared sensor, peripheral circuit and its corresponding control circuit were designed and real-time statistics on the number of people in the building were realized. General structure of the system was given,and hardware and software design of counting system were mainly stated. Practice shows that the counting system of building has higher stability, response speed and precision, and it has low price, so it has a broad application prospect.

Keywords Pyroelectric infrared sensors · Statistics of people number · Direction recognition

159.1 Introduction

A real-time and accurate statistics system for the number of people can effectively provide passenger flow volume for residential area, library, subway, supermarkets and other places, and can provide protection for public safety and arrangement work. Now for no crowed place, the main statistic methods for the number of people are feature analysis method with digital video and physical sensor method. Feature analysis method with digital video finds human body and makes statistics through inspection and analysis of the people characteristics in the scene sequence images, it has advantages of less investment and easy to integrate [1–4]. Physical

Y. Zhang (✉) · L. Yang · N. Jing
National Key Laboratory for Electronic Measurement Technology,
North University of China, Taiyuan, Shanxi, China
e-mail: tyzhzyh@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_159,

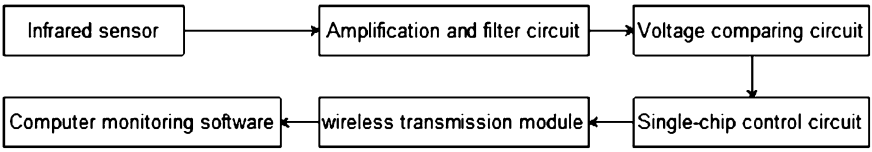


Fig. 159.1 Structure diagram of the system

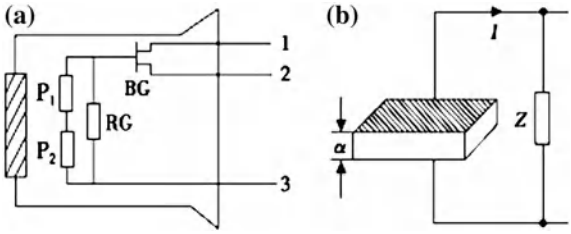


Fig. 159.2 Infrared receiving devices. **a** Internal structure of pyroelectric infrared. **b** Circuit diagram of single hotsensor detection unit

sensor method realizes counting through perceiving people passing in and out of the monitored area with physical sensors, and it has the advantages of accuracy and rapidity [5, 6]. But these methods did not mention whether they can identify directions of stream of people, and in many occasions, such as library and market, an intelligent detection display system is needed which can detect the number passing in and out of the place and detect the real-time number of remaining people. So it can be known when is the peak of the people number, which helps to arrange work scientifically and reasonably.

In the paper, based on the pyroelectric infrared sensor, peripheral circuits are designed and the movement direction of coming in (left) or going out (right) can be identified. Then different channel signals trigger the corresponding control circuits, real-time and accurate monitoring the number of people can be realized through the programing with single chip computer and connecting computer.

159.2 The Design of the System

Structure diagram of the system is shown in Fig. 159.1. When someone goes through the sensor detection range, the signal induced by the sensor will be transmitted to the next amplification and filter circuit, the signal after amplification and filter circuit will be transmitted to voltage comparing circuit, which can well recognize two direction of passing in and out. Then through the microcontroller programing the signal of different direction will be converted to the signal of the number of people, and in the meantime the signal of the number of people will be sent to computer by serial port programing through the wireless module, and then

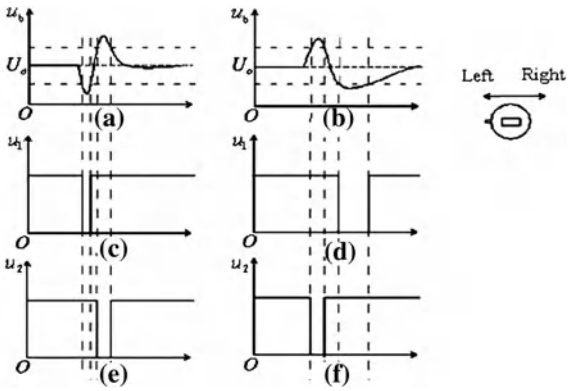


Fig. 159.3 Output signal of single person entering

statistical information of the number of people will be in real-time display with the graphical interface through the computer monitoring software.

159.2.1 Pyroelectric Infrared Sensor

Pyroelectric infrared sensor RE200B is used as the core component of receiving infrared radiation from human body in the system. It has the function of distinguishing direction, it can detect the change of infrared ray energy radiated by human body in non-contact form and converts it into voltage signal output. After being amplified, the voltage signal can drive all kinds of control circuit. As shown in Fig. 159.2(a), RE200B has two sensitive units P1and P2. When the two thermoelectric units with consistent feature are connected in series reversely or into differential balance circuit, compared with the structure of single thermoelectric unit [Fig. 159.2(b)], output signal can not only distinguish direction, but can also keep down interference due to temperature variation. When human body passes detection view from two different directions, there will be two alternating voltage signal output with opposite phase and they can reflect the human body walking direction in or out.

159.2.2 Signal Processing Circuits

The output of pyroelectric infrared sensor RE200B is very weak, and it contains noise caused by environment, so it must be filtered and amplified. In the circuit, amplifying circuit is made up by LM324 and the output voltage is made smooth through regulation resistance adjusting bias voltage. But it cannot guarantee signal stability because of environmental temperature influence, so differential

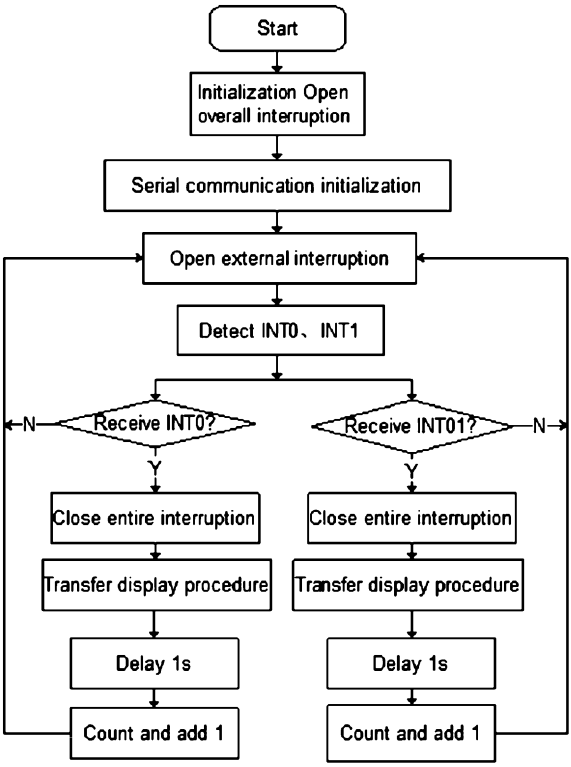


Fig. 159.4 Process of main program

coefficient circuit is entered. Regulation resistance adjustment protects stable static output of differential coefficient circuit from the influence of ambient temperature and stable at about U_O. When someone or infrared source passes sensor window from right or left, it is can be seen that the output of differential coefficient circuit would be different according to different passing direction, as shown in Fig. 159.3(a) and (b).

Although sensor RE200B can distinguish human movement direction, direction output signal cannot be recognized by single-chip, so voltage comparing circuits are added after amplifying circuits to distinguish two signals going in and out and to be recognized by single-chip microcomputer. Suppose the output of last level circuit is u_b and the outputs after two voltage comparing circuit are u_1 and u_2. When someone passes sensor window from left to right, u_1 firstly appears low voltage and soon after u_2 appears low voltage, as shown in Fig. 159.3(c) and (e). When someone passes sensor window from right to left, u_2 firstly appears low voltage and soon after $u_1$appears low voltage, as shown in Fig. 159.3(d) and (f). Direction can be judged when single-chip uses interruption to obtain the order of two low voltages.

159.2.3 Design of Single-Chip Control Circuit

In the design, AT89C2051 is chosen as the core of control circuit of single-chip microcomputer. External interruption INT0 and INT1 of single-chip microcomputer are connected, respectively, to two groups high- and low-level signals from voltage comparator. An interrupt request signal can be produced only when external interruption pins of single-chip receive negative jump signal. When human body infrared signal is detected, corresponding external interruption jumps into a low voltage level and makes an application for interruption to CPU. The direction of people in and out is determined according to judging interruption order produced at INT0 and INT1 pins.

The program design mainly considers signal acquisition, the different speed of people in and out, whether they are in or out, the environment interference and real-time display output and so on, whose accuracy directly relates to the reliability and practicability of counting device. Signal acquisition is completed by scanning periodic signal, every 2 $\sim$ 8 ms external interruption INT0 and INT1 pins are acquired. External interruption, timing counter T0 and T1 cooperate. The process of main program is shown in Fig. 159.4.

Signals are processed by single-chip microcomputer and then output to eight display module of a, b, c, d, e, g, f and DP to drive nixietube to display the total number of people in room, and at the same time the data of room are transmitted to master control system in the way of wireless serial communication to show results, so managers can look over the number of people in the room quickly from data terminal.

159.2.4 Design of Monitoring Software

System software is compiled based on VC++ of Microsoft company, including the system parameters Settings module, receiving data real-time display module and the data storage module.

159.3 Conclusion

The monitoring system for the number of people in building has been tried out in some enterprise building. Practice proves that the system is stable, real-time and for non-high population places the statistic number of people is accurate. Besides, the system has characteristics of low cost, small volume, light weight, easy to set and easy to use, so it has a broad application prospect.

References

1. Valle JD, Oliveira LES (2007) People counting in low density video sequences. LNCS 4872: PSIVT 2007:737–748
2. Chen L, Tao J (2007) People counting using iterative mean-shift fitting with symmetry measure. International conference on information and communication security, pp 890–895
3. Jun H, Yan C (2009) Video surveillance system for pedestrian counting. Video Appl Project 33(2):63–65
4. Lei Q, Guijun T, Zhou S (2010) Adaptive multi-feature-fused tracking algorithm for people counting application. Sci Technol Eng 10(4):910–913
5. Hashimoto K, Morinaka K (1997) People count system using multi-sensing application. International conference on solid-state sensors and actuators, pp 1291–1294
6. Ping H, Kaitao Y, Xiaoling M, Jiye W (2009) Design and applications of counting system for floating population. J MUC 18(3):64–72

Chapter 160
Temperature Compensation Method of Acceleration Sensor SD1221

Yanhua Zhang, Lu Yang and Zeyu Yang

Abstract Temperature error is in a large proportion of the total error in the micro-capacitive acceleration sensor, so the quality of temperature compensation has a direct impact on the measurement accuracy of the sensor. Based on the analysis of changes of acceleration sensor SD1221 affected by temperature, a linear forecast amendment method was put forward. Actual tests show that the method can reduce the impact of temperature changes effectively, simplify the amendment process and meet the requirements in a wide temperature range.

Keywords SD1221 · Temperature compensation · Forecast amendments

160.1 Introduction

In micro electro-mechanical systems, the temperature is an important question. Temperature error has a large proportion in its total errors for micro-capacitive acceleration sensor, so temperature compensation effect has a direct impact on total errors of the sensor. There are a lot of methods of temperature compensation, which are basically divided into hardware compensation and software compensation. Hardware compensation is to add resistance network in sensor circuit to achieve temperature compensation, while software compensation is to use software programing to modify sensor output data and achieve temperature compensation in digital instrument [1–3]. Different kinds of sensors have different

Y. Zhang (✉) · L. Yang · Z. Yang
National Key Laboratory for Electronic Measurement Technology,
North University of China, Taiyuan, Shanxi, China
e-mail: tyzyhzyh@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_160,

temperature characteristic and need to adopt different compensation methods. How to obtain their similarities and simplify amendment process are necessarily solved problems in batch application of sensors. According to the temperature variation characteristics of capacitive accelerometer sensor SD1221, a linear forecast compensation method is put forward in this chapter which can simplify amendment process and can be used in wide temperature range.

160.2 Characteristics of SD1221 Acceleration Sensor

Acceleration sensor SD1221 of SD Company is a single integrated chip based on MEMS technology. Under the action of acceleration, the gap between the inertial quality block in chip and detected electrode will change and result in capacitance change. Capacitance change results in equivalent voltage change at both ends of capacitance and vibration signal is changed into electrical signal [4, 5]. There are two analog outputs Vp and Vn with the opposite polarity in SD1221. When reference voltage is +2.5 V, both differential output and one-way output can be achieved. Its output is current. Since nitrogen is filled inside the chips, optimal damping is offered for the tiny gaps between inertia quality and external electrode, thus acceleration sensor has flat frequency response in a wider temperature range. Meanwhile, a high precision temperature sensor is built in the chip, which makes it to be used on highly required temperature occasions.

Typical application circuit of SD1221 is shown in Fig. 160.1, R_1 and R_2 separate 5 V to produce 2.5 V reference voltage, and TL084A constitutes differential amplifier circuit. When $R_3 = R_5$, $R_4 = R_6$, output $-V_g = (V_n - V_p)R_6/R_5$. C_3, R_4 and C_4, R_6 constitute a low-pass filter. The temperature output current is changed into temperature voltage V_T by R_T, voltage V_T and temperature T meets $V_T = R_T[(500\,\mu A) + (1.5\,\mu A)(T - 25)]$, and the slope is $k_T = \Delta V_T/\Delta T = (1.5\,\mu A)R_T$.

160.3 Temperature Variation leftacteristics of Acceleration Sensor SD1221

For 1221L-050, measuring range is ±50 g, frequency response is 2 kHz and sensitivity is 0.08 V/g. Choose $R_3 = R_5 = 8.2\,k\Omega$, $R_4 = R_6 = 8.2\,k\Omega$, then the output of differential amplifier circuit is $-V_g = V_n - V_p$.

Make the sensor in +1 g state, place platinum resistance beside the sensor to monitor temperature changes, adjust temperature of test box to −55,−45,−20, 0, 20, 40, 60 and 70°C, respectively, keep warm 1.5 h at each temperature point, track and measure V_p, V_n, $-V_g$, measurement results are in Table 160.1. Measured value, theoretical value and deviation between them are in Table 160.2.

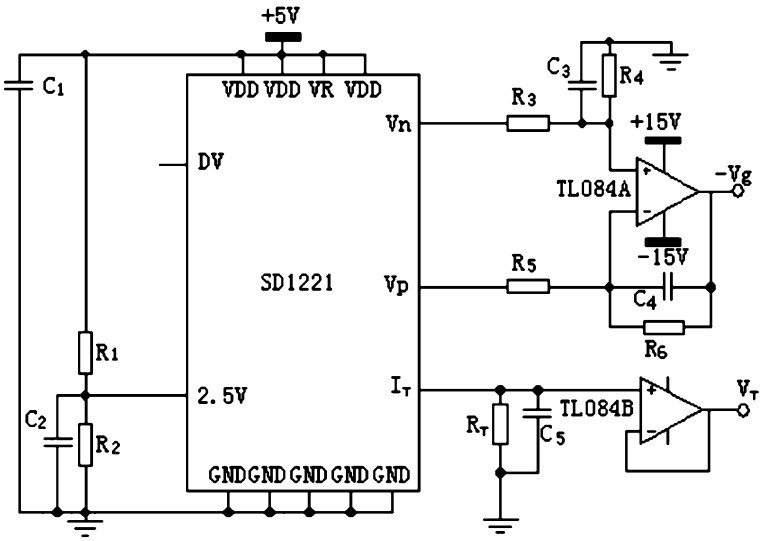


Fig. 160.1 Application circuit of SD122l

Table 160.1 Measurement data of V_P, V_n, and $-V_g$ changing with temperature T

T(°C)	−55	−45	−20	0	20	40	60	70
V_p(V)	2.459	2.457	2.451	2.446	2.441	2.434	2.427	2.423
V_n(V)	2.348	2.350	2.356	2.361	2.365	2.369	2.374	2.376
$-V_g$(V)	−0.111	−0.107	−0.095	−0.085	−0.076	−0.065	−0.053	−0.047

It can be seen from Tables 160.1 and 160.2 that sensor produces great output changes as changes of environmental temperature, output deviation of the sensor is smaller at small temperature range, while deviation is very big at bigger temperature range. Therefore, when it is used in the wide temperature range, temperature compensation is needed.

160.4 Temperature Compensation Method

Suppose that the output of acceleration sensor V_g consists of two components: one is the component under the action of acceleration, recorded as V'_g, and the other is the component under the influence of temperature, recorded as V'_T. Considering the two kinds of function dividedly can independently study respective influence: when acceleration is unchanged,V'_g is constant, the change of V_g is mainly dependent on V'_T with the change of temperature. It can be known from

Table 160.2 Measured value, theoretical value of V_g and deviation between them

T(°C)	−55	−45	−20	0	20	40	60	70
Measured (V)	0.111	0.107	0.095	0.085	0.076	0.065	0.053	0.047
Theoretical (V)	0.080	0.080	0.080	0.080	0.080	0.080	0.080	0.080
Deviation (V)	0.031	0.027	0.015	0.005	−0.004	−0.015	−0.027	−0.033
Relative Error (%)	38.75	33.75	18.75	6.25	−5.00	−18.75	−33.75	−41.25

Table 160.2 that change of V'_T with temperature is approximately linear, so suppose $V'_T = k_g T, k_g$ is the slope, T is temperature variable, therefore:

$$V_g = V'_T + V'_g = k_g T + V'_g$$

The main purpose of temperature compensation is to reduce the influence of temperature on the output that is to say reducing V'_T to minimum. So we can find a variable V_T whose change with temperature is a linear relation and whose slope has just opposite symbol with k_g, Namely it is met:

$$V_T = -k_g T + V_b$$

In the above formulaV_b is offset. Use V_T to compensate V_g, suppose the output after compensation is V_{bc}, then:

$$V_{bc} = V_T + V_g = (-k_g T + V_b) + (k_g T + V'_g) = V_b + V'_g$$

The output after compensation V_{bc} has nothing to do with temperature. Suppose V_{out} is V_{bc} minusV_b, then:

$$V_{out} = V_{bc} - V_b = V'_g$$

From above formula it can be seen that output V_{out} only relates to acceleration.

Using a linear model to compensate is easy to implement with hardware circuits and easy to adjust parameter. The compensation mode can be realized by the circuit as shown in Fig. 160.2.

The circuit consists of TL084C, TL084D and AD620. TL084C realizes temperature compensation, TL084D realizes bias output voltage through parting 2.5 V reference source, AD620 realizes unbiased adjustment and sensitivity adjustment. Suppose the gain of AD620 is G, then circuit output can be expressed as:

$$V_{out} = G\left[(V_p - V_n) + V_T - V_b\right]$$

In this formula, the required output can be obtained only through adjusting three parameters G, V_T and V_b. Specific adjustment process is as follows:

1. Firstly determine two parameters at 1 g $R_T$and V_b, the method is as follows:

A: Measure and obtain change values $-V_g = V_n - V_p$ at temperature from −55 to 70°C.

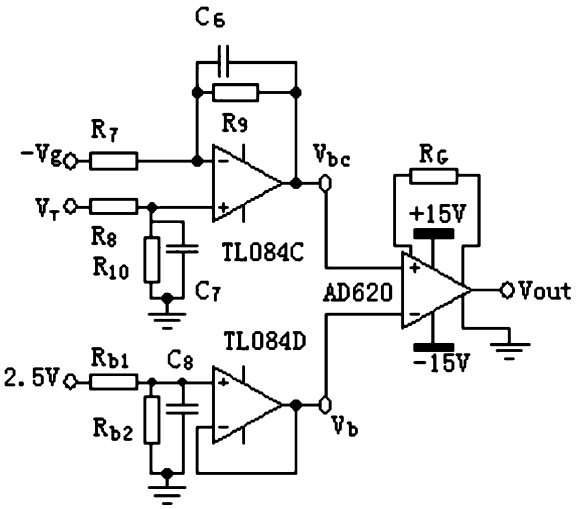


Fig. 160.2 Linear compensation and sensitivity adjustment circuit

Table 160.3 Output of $R_T = 0.34\,\text{k}\Omega$ $V_b = 0.1572\,\text{V}$ $G = 1$

$T(^\circ)$	−55	−45	−20	0	20	40	60	70
V_{out} (V)	0.083	0.084	0.085	0.085	0.086	0.085	0.084	0.083

B: Using function polyfit (T, V_g, 1) in Matlab to calculate the slope k_g and the offset V'_g. From measured values V_g in the Table 160.1, $k_g = -0.5099$, $V'_g = 0.0843(\text{V})$ can be calculated, then the slope of the temperature sensor is $k_T = -k_g = 0.5099$.

C: From the output leftacteristic slope of temperature sensor $k_T = (1.5\,\mu\text{A})R_T$, it can be got:

$$R_T = k_T/1.5 = 0.34\,(\text{k}\Omega)$$

Substitute the value of R_T into output leftacteristic expression of temperature sensor, it can be got:

$$V_b = R_T\,[(500\,\mu\text{A}) + (1.5\,\mu\text{A})(-25)] = 0.1572\,(\text{V})$$

2. Calculate the gain G when output is 0.5 V at 1 g.

A: Suppose $G = 1$, Calculate the outputs at different temperature at 1 g, $V_{out} = V_{bc} - V_b = V_g + V_T - V_b$, as shown in Table 160.3.

B: Calculate the mean $V_m = [\max(V_{out}) + \min(V_{out})]/2 = 0.0843\,(\text{V})$.

C: If the output at +1 g is $V_{out} = 0.5\,\text{V}$, the gain is $G = V_{out}/V_m = 5.9286$.

D: Calculate gain controlling resistance $R_G = 49.4/(G-1) = 10.023\,\text{k}\Omega$.

Table 160.4 Output of $R_T = 340\,\Omega$ $V_b = 0.1563\,V$ $G = 5.94$

$T(°)$	−55	−45	−20	0	20	40	60	70
V_{out}(V)	0.498	0.505	0.509	0.511	0.518	0.513	0.502	0.497

Table 160.5 Measured value, theoretical value and deviation

T (°)	−55	−45	−20	0	20	40	60	70
Measured (V)	0.498	0.505	0.509	0.511	0.518	0.513	0.502	0.497
Theoretical (V)	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
Deviation (V)	−0.002	0.005	0.009	0.011	0.018	0.013	0.002	−0.003
Relative error (%)	−0.326	0.980	1.871	2.108	3.534	2.583	0.445	−0.623

Table 160.6 Measured value, theoretical value and deviation

T(°C)	−55	−45	−20	0	20	40	60	70
Before compensation (%)	38.8	33.85	18.8	6.25	−5.00	−18.8	−33.8	−41.2
After compensation (%)	−0.33	0.98	1.87	2.11	3.53	2.58	0.44	−0.62

160.5 Error Analysis before and after Compensation

According to above analysis, $R_T = 0.34\,k\Omega$, $R_{b1} = 15\,k\Omega$, $R_{b2} = 1\,k\Omega$ and $R_G = 10\,k\Omega$ are determined. Insert corresponding resistances in compensation circuit and test again according to former test method, the change of measured V_{out} with temperature is shown in Table 160.4. Measured value, theoretical value and deviation between them are shown in Table 160.5. Relative error before and after compensation are shown in Table 160.6.

From Table 160.6 we can know that: before compensation, output deviation at low temperature area(−55 ~ −20°C)and at high temperature (40–70°C) are over 15%, but after compensation, output deviation at whole temperature area(−55 ~ 70°C) is less than 4%. So the method can preferably reduce the influence of temperature change.

160.6 Conclusion

Using a linear forecast to simulate temperature influence of acceleration sensor is completely consistent with change rule of measured data, and isolated V_g' can completely reflect the true value of sensor output without temperature influence. The proposed compensation plan only requires measuring response curves of sensor at different temperature at 1 g state, the precise values of adjustment parameters can be predicted, so it facilitates batch production.

References

1. Guo T (2003) Study on the temperature compensation method of piezoresistance silicon film sensor. Micronanoelectron Technol 40(8):489–491
2. Wang X (2007) Intellectual correction and temperature compensation of transducer based on LS-SVM. Transducer Microsyst Technol 26(3):76–79
3. Gang ZT, Gang DM, Long YZ, Yang Y (2010) An accurate thermal errors compensating method for embedded pressure measuring system. Appl Electron Tech 36(6):101–103
4. Xiaogang T (2005) MEMS differential capacitor acceleration sensor. J Instrum Technol Sens 42(12):8–15
5. ShaoHua W (2007) Research on micro capacitive wireless accelerometer based on MEMS. Transducer Microsyst Technol 26(6):94–96

Chapter 161
Acquisition of Weak GPS Signals Based on Cross Correlated Interference Elimination

Chao Deng and Dandan Wang

Abstract The cross-correlated interference between strong and weak signals is a problem pressed for solution during acquisition of weak GPS signals. This paper adopts an algorithm combined successive interference cancellation and ingenious differentially coherent integration to solve the problem. The cross correlated interference is eliminated by successive interference cancellation, and then the weak GPS signals are captured by ingenious differentially coherent integration in a short time. The experimental results indicate that this algorithm can eliminate the cross correlated interference between strong and weak signals effectively and capture the weak GPS signals whose C/N are low down to 15 dB/Hz. It can improve the sensitivity of GPS software receiver.

Keywords Acquisition of weak GPS signals · Cross correlated interference · Successive interference cancellation · Ingenious differentially coherent integration

161.1 Introduction

In global positioning system (GPS) system, when the difference between strong and weak signals is up to 20 dB, the interference of the strong signals will affect the normal work of receiver. The elimination of cross-correlation is to remove the

C. Deng (✉) · D. Wang
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: superdeng5@163.com

D. Wang
e-mail: wangdan8696@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_161,

cross-correlation between weak and strong signals from the correlation during correlative operation of weak signals [1].

At present, there are a lot of articles present the solutions to the issue of cross correlated interference between GPS signals [2]. In the literature [3] the successive interference cancellation was used to remove the cross correlated interference, then the differentially coherent integration was used to capture the weak signals and the algorithm can capture the weak GPS signals whose C/N are low down to 23 dB/Hz. But it is complicated and low sensitive. In the literature [4], an ingenious differentially coherent integration was used to capture the weak signals directly; this algorithm is simple and effective. But there is not a part to remove the cross correlated interference. In this paper, the advantages of the two algorithms in literature [3] and literature [4] were combined. The successive interference cancellation was used to remove the cross correlated interference and the ingenious differentially coherent integration was used to capture the weak GPS signals. This algorithm can eliminate the cross correlated interference effectively and capture the weak GPS signals whose C/N are low down to 15 dB/Hz.

161.2 Acquisition of Weak GPS Signals

161.2.1 Elimination of Cross Correlated Interference

Suppose there are M satellites in the GPS base band signals, as shown in formula (161.1), there are strong signals from S satellites and weak signals from M satellites, and s_m is the signal from number m satellite.

$$r(k) = \sum_{m=1}^{M} s_m \tag{161.1}$$

The in-phase component and quadrature component of the weak signal's correlation value from m satellites during acquisition is shown in formula (161.2) [5].

$$\begin{aligned} \begin{pmatrix} I_w(k) \\ Q_w(k) \end{pmatrix} &= A_w d_w(k) \begin{pmatrix} \cos\phi_w \\ \sin\phi_w \end{pmatrix} + v(k) + \sum_{s=1, s\neq w}^{M} \frac{A_s}{t_c} \int_{kt_c}^{(k+N)t_c} g_{\mathrm{ws}}(t)dt \\ g_{\mathrm{ws}}(t) &= d_s(t-\tau_{\mathrm{wsk}}^{D}) c_s(t-\tau_{\mathrm{wsk}}) c_w(t) \cdot \begin{pmatrix} \cos(2\pi\Delta f_{D,\mathrm{ws}} t + \varphi_{\mathrm{ws}}) \\ \sin(2\pi\Delta f_{D,\mathrm{ws}} t) + \varphi_{\mathrm{ws}} \end{pmatrix} \end{aligned} \tag{161.2}$$

On the right side of this formula, the first part is auto-correlation of ideal signals, the third part is cross-correlation interference. V is noise. $\Delta f_{D,\mathrm{ws}}$ and ϕ_{ws} is

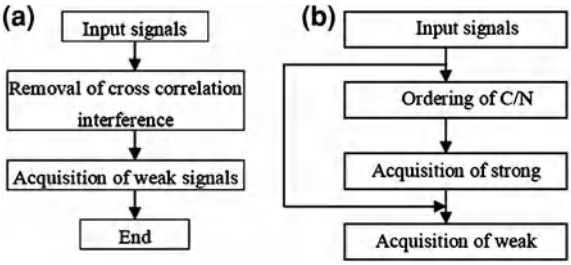


Fig. 161.1 The flow chart. **a** The block diagram of synthesis algorithm. **b** The flowchart of successive interference cancellation

the relative Doppler frequency offset and carrier phase offset. The elimination of cross correlated is to remove the part three in the formula (161.2).

161.2.2 Acquisition of Weak GPS Signals

Acquisition of GPS signals is a searching process on both C/A code phase according to correlation of C/A code and carrier frequency. The most typical algorithm is based on FFT; it is operated in a PRN cycle which is 1 ms. When the GPS signals are weaker, it is not effective to capture the signals by only 1 ms data. At this time, adding the length of the data and the accumulating times can be used to capture the signals. Usually acquisition algorithms of weak GPS signals are coherent combining, non-coherent combining and differentially-coherent combining.

The coherent combining only accumulates the signal energy according to the correlation of signals during accumulation time; the noise energy does not increase with the accumulation time. But the jumping of navigation data bit will cancel part or even the whole signals. The non-coherent combining accumulates the sum of the squared correlation; this algorithm will accumulate both the signals energy and noise energy and produce S/N loss which is called squared loss. The differentially-coherent combining accumulates the correlation after a multiplication to reduce the squared loss and the jumping influence of the navigation data bit according to the non-related of the correlation between two adjacent epochs. This is the most efficient one among those acquisition algorithms of weak GPS signals.

161.3 Realization of the Algorithm

In this paper, the successive interference cancellation was used to remove the cross correlated interference; then the ingenious differentially coherent integration was used to capture the weak GPS signals. Removing the interference while capturing the signals. The block diagram can be seen in Fig. 161.1a.

161.3.1 Removal of Cross Correlation Interference: The Successive Interference Cancellation

In this paper, the successive interference cancellation was used to remove the cross correlated part which is also the third part of formula (161.1) [3]. The block diagram of the successive interference cancellation is shown in Fig. 161.1b.

As shown in Fig. 161.1b, order the C/N of all satellites signals from input signals firstly; then capture the strongest signals and adaptively adjust the accumulation time. The strongest signals will be reconstructed by the received parameters which are amplitude of the signals, Doppler frequency shift of the carrier and C/A code phase. At last, cancel the strongest signals by the input signals to remove the interference. Remove all of the strongest signals by repeating the steps described above; and then capture the weak signals.

161.3.2 Acquisition of Weak Signals: The Ingenious Differentially Coherent Integration

This paper adopted a high sensitive algorithm based on the ingenious differentially coherent integration mentioned in the literature [4] to capture the signals. This algorithm is improved on the base of differentially-coherent combining which is shown in formula (161.3).

$$z = \sum_{k=2}^{K} \sqrt[4]{(I_k I_{k-1} + Q_k Q_{k-1})^2 + (Q_k I_{k-1} - I_k Q_{k-1})^2} \tag{161.3}$$

where k is accumulation times, I and Q is respectively the correlations of I branch signal and Q branch signal.

Firstly, replace the amplitude envelope by the energy envelope. Then suppose the code phase, Doppler frequency and phase is in accordance with those of the reference signal. After coherent combining, the signals exist in the I branch and the signals mean value remain unchanged between the corresponding points. And there is only noise in Q branch; the mean value is zero and the amplitude remain unchanged between the corresponding points. The consequence can be got as follow formula shown.

$$(Q_k I_{k-1} - I_k Q_{k-1})^2 \approx 0 \tag{161.4}$$

And the consequence can be simplified as follow.

$$z = \sum_{k=2}^{K} (I_k I_{k-1} + Q_k Q_{k-1}) \tag{161.5}$$

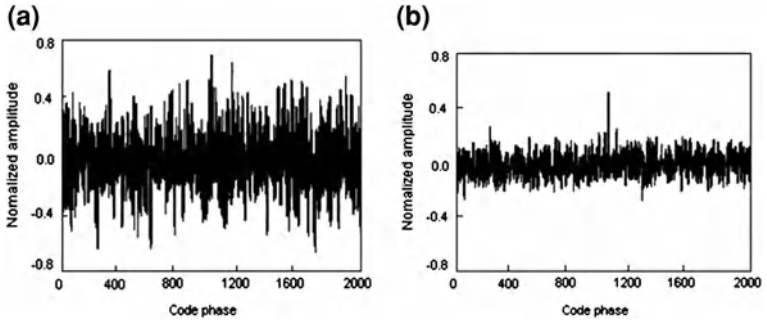


Fig. 161.2 Autocorrelation results. **a** Autocorrelation result without cross correlation mitigation. **b** Autocorrelation result with cross correlation mitigation

The formula (161.5) is the realization of the ingenious differentially coherent integration and this proves it is simpler to realize. The experiments as follows will prove the performance of this algorithm rise or not.

161.4 Experiments and Analysis

There are there strong signals and one weak signal in the simulation and the weak signal is 20 dB lower than the satellite signal. The simulation is shown in Fig. 161.3. It is traditional acquisition algorithm in Fig. 161.2a; the signal can not be captured properly due to the interference. Which is shown in Fig. 161.2b is the correlation of the successive interference cancellation. It can be seen that the autocorrelation peak is obvious and the signal is captured successfully. This algorithm can achieve a better effect on removing the cross correlated interference.

The ingenious differentially coherent integration is used to capture the GPS signals when there is no cross correlated interference. As shown in Fig. 161.3a, K is the accumulation times of non-coherent combining and M is accumulation time of coherent combining before non-coherent combining; and the unit of M is ms. The relationship of acquisition probability changed with C/N is got when K and M is selected different value by simulation. When accumulation times are 10, the signals whose C/N are low down to 18 dB/Hz can be captured successfully and the acquisition probability is up to 100%. It proves the upper hand of this algorithm in high sensitive acquisition.

Select the C/N respectively and combine the successive interference cancellation and the ingenious differentially coherent integration to cancel the strong signals and capture the weak signals. As shown in Fig. 161.3b, when cancel two strong signals, the weak signals whose C/N are above 20 dB/Hz can be captured successfully; when cancel all of the strong signals, the weak signals whose C/N are above 15 dB/Hz can be captured successfully.

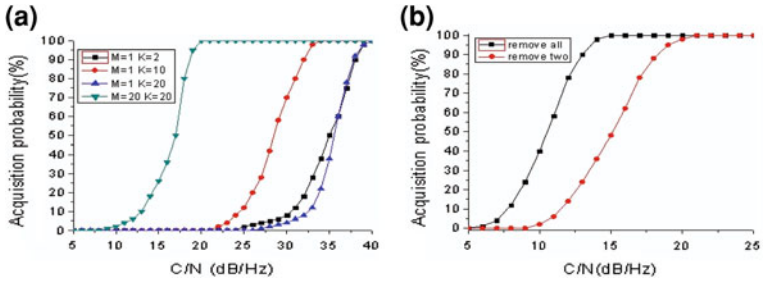


Fig. 161.3 Acquisition probability with C/N. **a** Before removing of strong signals. **b** After removing of strong signals

161.5 Conclusion

In this paper, a new algorithm combined successive interference cancellation and ingenious differentially coherent integration is presented. It can remove the cross correlated interference of the strong signals efficiently and adaptively adjust the accumulation time according to the C/N of the satellite signals to capture the weak signals in a short time. It can improve the sensitivity of GPS software receiver and is simple to realize. The experiments shows that this algorithm can eliminate the cross correlated interference between strong and weak signals effectively and capture the weak GPS signals whose C/N are low down to 15 dB/Hz.

Acknowledgments This paper is supported by Henan Polytechnic University Postgraduate Dissertations Foundation. The author would like to thank the Editor for helpful comments that greatly improved this paper.

References

1. Hao J (2009) Implementation of cross: correlation interference mitigation algorithm for GNSS receiver [J]. Microelectronics 39(6):857–860 Dec
2. Aekermann J (1980) Parameter space design of robust control systems [J]. IEEE Tmm Auto Cont 1(25):1058
3. Wang C, He D (2009) Acquisition of weak GPS signals based on successive interference cancellation [J]. High Technol Lett 19(12):228–1232
4. Tang B, Dong X (2009) A new high-sensitivity acquisition method of software GPS receiver based on Ingenious Differentially Coherent Integration [J]. Signal Process 25(5):832–836 May
5. Eamonn PG, Andrew GD (2007) Cross correlation mitigation techniques for software GPS C/A code receivers. In: Proceedings of the international global navigation satellite systems society IGNSS symposium, Sydney, Australia

Chapter 162
An Analysis on Improved Wireless Channel Allocation Strategy

Lei Shi and Caihong Zhao

Abstract With the wide application of GPRS/GSM mobile access network, the efficient channel allocation strategy has shown importance in supporting the quality of network service. In view of GSM/GPRS network limitations, which has a low utilization about wireless resource and channel, and based on the existing resource allocation strategy, an improved channel allocation strategy is proposed. Numerical results indicate that new strategy can decrease the voice call blocking probability, reducing the average transmission time of GPRS packet, improving the channel utilization, etc.

Keywords GSM/GPRS · Channel allocation · Blocking probability

162.1 Introduction

In the Global system for mobile communications/General packet radio service (GSM/GPRS) wireless communications, wireless channel distribution strategy determines the allocation and use of radio resource. With the application of GPRS service and the increasing number of mobile users, limited radio resources becomes more and more precious, and urgently needs to improve the capacity of

L. Shi (✉)
Department of Computer Engineering, Henan Polytechnic Institute, Nanyang 473009, Henan, China
e-mail: shileiftp@163.com

C. Zhao
Department of Electrical Engineering, Jiyuan Vocational and Technical College, Jiyuan 459000, Henan, China
e-mail: zhaocaihong1974@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_162,

the system—the rational allocation and utilization of wireless channel resources [1, 2]. GPRS network is a GSM network as the bearer network, and they share the same base stations and spectrum resources. At the same time, GPRS and GSM network channel is also shared. Channel sharing has two implications, one is the channel sharing of voice and data services, another is that between the data users shares same data channel [3]. Due to the limitations of network equipment, provided channel is limited for GPRS, it needs to find a balance between the two. An unreasonable allocation strategy not only makes access to the mobile station not just a good quality of service(QOS), but will also decline the quality of the original communication(COS). Therefore, a reasonable channel allocation strategy is an extremely important factor for the whole access network QOS and COS, not only to remain a good QOS network, but also to improve resource utilization and network capacity [4, 5].

162.2 An Existing Channel Allocation Scheme and Its Performance Comparisons

1. Fixed Channel Allocation Scheme(FCAS)

Suppose that the current system remaining number of channels is *N*, when *N* is not less than *M* for a need data packets of *M*-channel transmission, the system will assign the *M* channel for this packet. Otherwise, GPRS packet transfer request will be rejected, and when there is no channel voice calls only will be blocked [6].

2. Dynamical Channel Allocation Scheme(DCAS)

According to the current available channel resources, DRA will allocate up to M channels for this call for the data call of up to the M channels transfer. If the remaining number of channel is $N \geq M$, then assign M channels. If $0 \leq N < M$, it will be assigned the N channels; if $N = 0$, this data calls will be blocked. Actually, DRA strategy is that the default used channel number is at least 1 or 2. When there is no available channel, voice calls will be blocked.

3. Hybrid Channel Allocation Scheme(HCAS)

This strategy combines the fixed channel allocation and dynamic channel allocation strategy, in which is each system is assigned a fixed set of channels and a flexible dynamic channel reserves, preallocation method depends on the known business change, these flexible channels are allocated those scheduled system to resolve the known business predictable changes. Using the predictable allocation, business volume of each system is continuous or periodically detected, then the channel based on these measured values assigns to different business. Therefore, HCAS is the combination of the product of FCAS and DCAS, so it is said HCAS [7].

In the current case of radio resources shortage, above strategy is either fixed GPRS packet to use number of channels (FCAS), or the default GPRS packet to

use at least channel number which is 1(DCAS), or the actual channel allocation produces frequency interference(HCAS), which both will result in the decline of radio resources efficiency and channel utilization.

Therefore, based on the existing channel allocation scheme, proposed an improved dynamic channel allocation strategy, which is channel re-allocation and de-allocation scheme(CRA-DAS). Numerical results indicate that the new strategy could well reduce the voice call blocking rate, reduce the GPRS packet transmission time, improve the channel utilization etc.

162.3 Improved Wireless Channel Distribution Strategy

When a voice call arrives, it needs a channel. If remaining channel $N > 0$, the channel will be assigned to this call; if $N = 0$, but there is $j_{(\mathrm{m})}$ ($\mathrm{m} > 1$) which will split a channel to voice call, at the same time it becomes $j_{(\mathrm{m}-1)}$, otherwise call blocking, that is, CDAS. When a voice call is completed or switched, and not new voice calls or cut, a channel will be released. If there is GPRS data packet transmission of a lower QOS, then the released channel will be allocated to the GPRS data packet, which is the CRAS. Noted here, if remaining channel $N > 0$, and all of the GPRS data packets have reached their own QOS ($j_{(\mathrm{m})}$), the newly released channel will remain idle. Whether there is the cache, the analysis M value changing for the impact of the system performance is not inevitable link. Therefore, the analysis does not consider caching strategies [8].

Then, in the CRA-DAS strategy the system channel state can be expressed as follows: $z = (i, j_{(1)}, j_{(2)}, \ldots, j_{(M)})$ is that the voice calls quantity is i, the GPRS calls quantity of occupied channel number m is $j_{(\mathrm{m})}$, the GPRS calls quantity of occupied channel number $m + 1$ is $j_{(\mathrm{m}+1)}$. If the total channels number is C, under the CRA-DAS, all the possible channel state set is in the following form:

$$Z \equiv \left\{ z \middle| 0 < i + \sum_{m=1}^{M} m j_m \le C;\ 0 \le j_m \le \left\lfloor \frac{C}{m} \right\rfloor,\ m = 1, 2, \ldots\ldots, M \right\} \tag{162.1}$$

Based $\pi(z)(z \in Z)$ is a steady-state probability in $z = (i, j_{(1)}, j_{(2)}, \ldots, j_{(M)})$ state, all the states of $\pi(z) = \{\pi_1, \pi_2, \ldots, \pi_{S_{\max}}\}$ can get the various state transition matrix $R_{S_{\max} \times S_{\max}}$, where $\sum_{z \in Z} \pi(z)$ is equal to 1, the $\pi(z)R = \pi(z)$ and the $\sum_{z \in Z} \pi(z) = 1$ can attain such a $\pi(z)$. The probability transfer equation among the various states can have a continuous multi-dimensional Markov chain to get [9, 10].

Suppose that P_{bv} is voice blocking probability, P_{bg} is GPRS data blocking probability

$$P_{bv} = \sum_{i+j=C,\, j_2=0,\, j_3=0, \cdots,\, j_M=0} \pi(z) \tag{162.2}$$

$$P_{bg} = \sum_{i+\sum_{q=1}^{M} qj_q = C} \pi(z) \tag{162.3}$$

Suppose that $E(T_g)$ is the average transmission time of GPRS data packet, λ_g is GPRS packet arrival rate, γ is busy channel utilization, there are:

$$E(T_g) = \frac{\sum_{z \in Z} (\sum_{m=1}^{M} j_m) \pi(z)}{\lambda(1 - P_{bg})} \tag{162.4}$$

$$\gamma = \frac{\sum_{z \in Z} (\sum_{m=1}^{M} j_m) \pi(z)}{C} \tag{162.5}$$

162.4 Numerical Analysis

Suppose that a district containing the total channels are 15, the average arrival rate of voice calls is to obey the Poisson distribution which is $\lambda_{vn} = 0.04$, the average service time is to obey the exponential distribution which is $\mu_C = 1/180.0$ $\lambda_g = (0.3 \sim 3)$ packet/sec should separately be considered $M = 2$, $M = 3$, $M = 4$. The simulation results separately show the value, as shown in Fig. 162.1.

As shown from Fig. 162.1a and d, when using the improved CRA-DAS with the increase of λ_g, under the different value of M, the voice blocking probability P_{bv} and the GPRS packet average transmission time has dropped significantly have a significant decrease. With the increase of M, a packet will get several channels, the date being transmitted probability is the greater, and to make the average transmission time of GPRS packet is reduced. And then, voice calls can use more channels at a certain time range.

As shown from Fig. 162.1c, when using the improved CRA-DAS, channel utilization has also been a little improvement, this is because the CRA-DAS can effectively use all the free channels.

As is shown from the Fig. 162.1b, when using CRA-DAS, the GPRS packet blocking probability has slight increases, but the average transmission time of GPRS packet is significantly reduced, thereby the deterioration of compensating GPRS blocking probability.

The simulation results can be analyzed, when M is greater, the overall optimization effect will be more obvious. And using CRA-DAS and under the conditions of $M = 3$ and $M = 4$, voice call blocking is nearly the same; at the same time, it is not in order to ensure the data packet transmission by increasing M, which can make the larger packet blocking rate and the lower channel utilization.

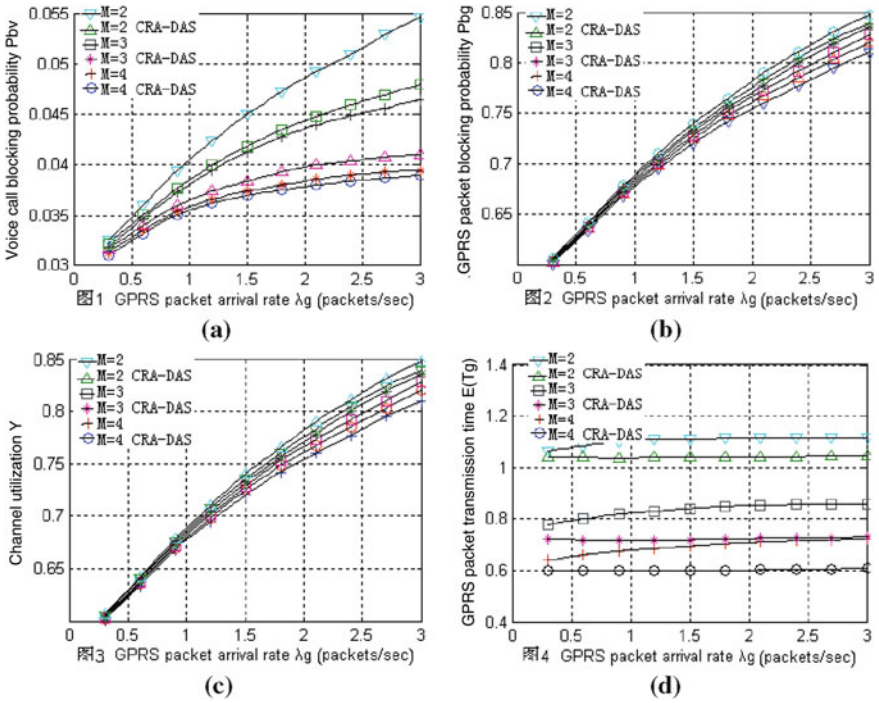


Fig. 162.1 The simulation results

Therefore, a GPRS package for multi-channel demand depends on the different QOS and the system operating environment.

So, an ideal channel allocation strategy should at least satisfy the following conditions:

1. High utilization rate of channel resources
2. Call blocking probability as low as possible
3. Ensure business continuity

The above three conditions are contradictory to each other, it is not possible to find a so-called ideal channel allocation scheme. Only in the certain circumstances, the optimal channel allocation strategy meets certain requirements. So, when to determine a system should be adopted channel allocation strategy, it must be integrated into the above three conditions, using different focusing for different system [11].

162.5 Conclusions

Through the analysis of several current GSM/GPRS system channel allocation strategy, an improved channel allocation strategy CRA-DAS is proposed which is analyzed in theory and performance. The numerical results show that the new

strategy can lower the voice call blocking probability, reduce the average transmission time of GPRS packet, improve the channel utilization, etc. At present, the third-generation communication systems dynamic channel allocation scheme is a real-time requirement, which is further optimized for analysis and comparison of its programs. At the same time, it can further improve the utilization efficiency of radio resources.

References

1. He X (2003) Digital mobile communication technology and applications. Machinery Industry Press, Anyang
2. Zhang Q, Zhang D (2004) Mobile communications infrastructure. China Water Power Press, Beijing
3. Han B (2004) GPRS principle and network optimization. Machinery Industry Press, Anyang
4. Ma L, Cheng L (2008) PDCH channel configuration method and research. Mobile 19:63–68
5. Q He, Z Fang (2003) Wireless capacity planning of data services. Communications market, Issue 9, pp 78–8
6. Ghribi B, Logrippo L (2000) Understanding GPRS: the GSM packet radio service. Comput Networks 34(5):763–779
7. Ermel M, Müller T, Schuler J, Schweigel M, Begain K (2000) Performance of GSM networks with general packet radio services. Perform Eval 48(1–4):285–310
8. Lindemann C, Thümmler A (2003) Performance analysis of the general packet radio service. Comput Networks 41(1):1–17
9. Meo M, Ajmone Marsan M (2004) Resource management policies in GPRS systems. Perform Eval 56(1–4):73–92 March
10. Zhang Y, Soong B, Ma M (2006) A dynamic channel assignment scheme for voice/data integration in GPRS networks. Comput Commun 29(8):1163–1173
11. Kang J, Nath B (2006) A study on switching voice traffic seamlessly between GSM and GPRS cellular networks. Comput Commun 29(8):1079–1093

Chapter 163
The Waveform Generator Based on Quartus II Software

Lili Zhao, Jian Liu, Weiwei Zhang, Caijuan Shi and Xiaodong Yan

Abstract With the development of technology in computer software, hardware and Very Large Scale Integrated Circuits (VLSI), EDA is gradually taking the place of the traditional method of designing and has become the basic method of modern electronic designing. Many EDA software functions are more comprehensive. The Quartus II software offers an intuitive and user-friendly interface based on customer feedback, and the popular Maxplus II interface. This paper introduces the operational method of Quartus II software with waveform generator design as an example, and gives the simulation result with Quartus II software.

Keywords EDA · Quartus II · Waveform generator · Simulation

163.1 Introduction

Waveform generator is a very common device, which is used in many designs. The most common waveform is the sin wave, the square wave and the triangle wave. Although these three kinds of waveforms in analog circuit are easy to get, it will make hardware circuit more trivial, the circuit is easily subjected to interference. Using FPGA/CPLD to design then cannot appear in this circumstance. With FPGA/CPLD to implement it allows us to become more familiar with algorithms about signal waveform and principles of signal sampling and waveform generation.

L. Zhao (✉) · J. Liu · W. Zhang · C. Shi · X. Yan
College of Information, Hebei United University,
Tangshan 063009, China
e-mail: zhaolili@heut.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_163,

Quartus II software is the fourth generation of Altera programmable device software development, in addition to its inherited Maxplus II software features and components of all objects, but it also includes many new features and supports for new FPGA/CPLD device families [1, 2]. Quartus II software provides a complete, multiplatform design environment that easily adapts to your specific design needs. The design, synthesis, place and route, timing analysis, simulation, are integrated into a seamless environment [3]. This paper introduces the operational method of Quartus II software with waveform generator design as an example, and gives the simulation result with Quartus II software.

163.2 The Functional Requirements of the Waveform Generator

Waveform generator can produce sine, square and triangle wave, and adjustable frequency. It consists of three modules: decoding, frequency and three wave generation. Decoding module is to control output waveform. It is compiled by the input signal, respectively. In the three outputs it will output "100", "010", "001" cases, corresponding to the square wave, triangle wave, and sine wave. Frequency module is controlled by adjusting the external clock signal frequency to meet the equipment or the user's needs.

For the sine wave, make a sine wave table, which is calculated to form, then loop through the table to achieve the generation of sine wave. First cycle of the waveform is divided into a number of points, such as 50 points, 50 points will be calculated which correspond to the voltage and then compared with the standard voltage, and analog to digital conversion into the digital and then to assign. In this paper, the sample is 64 points per cycle, and waveform output value is between 0 and 255, $256/32 = 8$, the average difference between adjacent samples is 8. As the sine wave characteristics, although the sampling interval is the same, but the difference between adjacent samples is different. Difference between adjacent samples is less where the wave steepness is lower, and vice versa. When the difference between sampling points is 1, 2, 3, 4, 6, 6, 8, 8, 10, 10, 11, 12, 12, 12, 13, 13 incremental change, and then reverse decline, resulting in the most accurate waveform, showing the best results.

Square wave in the digital signal is easier to get, as long as the alternating output all zeroes or all ones, 32 clock cycles delay to flip once.

Triangle design is easier than the sine wave, because its algorithm is simpler. Because it is linear, we can use simple addition and subtraction algorithm. Triangle wave is achieved by constant plus the same number by "1,000"–"11,111,111" and then continue to reduce the same number. As every 32 clock cycles to change the time, $256/32 = 8$, so every time plus or minus 8–64 clock cycles to achieve a triangular clock cycle.

163.3 Design of the Waveform Generator Based on Quartus II Software

163.3.1 Creating a Project and a Design in the Quartus II Software

When we design digital circuit system based on the Quartus II software, any one design is a project (Project). A Quartus II project includes all of the design files, software source files, and other related files necessary for the eventual implementation of a design in a programmable logic device [4]. The design folder of the waveform generator named Wave_a.

Open the Quartus II software interface, on the File menu, click New Project Wizard to create a project for the waveform generator, specify project directory, name, and top-level entity [5]. The design file of the project called Wave. In the Quartus II software, input design files about waveform generator, and save them to the folder Wave_a. On the Project menu, click Add/Remove Files in Project to open the folder Wave_a, then click…, from the folder select the VHDL files related the design into the project folder Wave_a.

163.3.2 Compile the Design

You can start a full compilation in the Quartus II software, which includes the Analysis & Synthesis module, or you can start Analysis & Synthesis separately. You can perform an Analysis and Elaboration to check a design for syntax and semantic errors without performing a complete Analysis & Synthesis or use the Analyze Current File command on the Processing menu to check a signal design file for syntax errors. The design of the project is fitted into the FPGA/CPLD target device, and produced a variety of uses of the output file, such as functional and timing simulation files, the programming files [6].

Specify a target device or device family before compile the design. Open the Quartus II software interface, on the File menu, click Open Project, select project Wave, then on the Assignments menu, click Settings. Select device and use filters to narrow the devices displayed. Next, on the Processing menu, click Start Compilation, perform full compilation. Compilation stops after file is analyzed for syntax errors, which are displayed on the compiler messages. Double-clicking the text of the error that pops up with the wrong file, the flashing cursor can be found at or near the error, modify and save it, again compilation. After full compilation is successful, you can select entity Wave on the Project Navigator window, click the right mouse button, select Locate in RTL Viewer. The Quartus II RTL Viewer provides graphical representations of the waveform generator (Fig. 163.1). In addition to the schematic view, the RTL Viewer has a hierarchy list that lists the

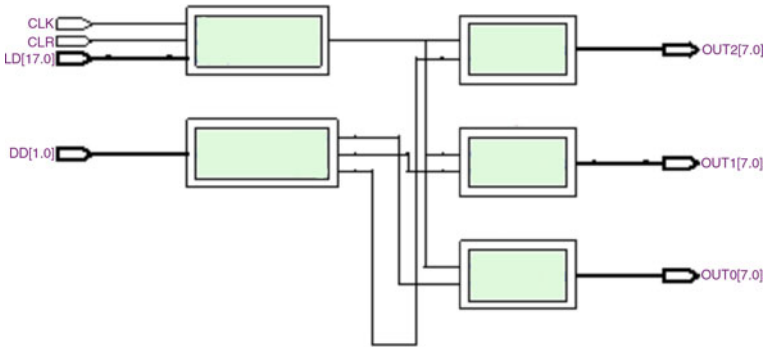


Fig. 163.1 The RTL viewer of waveform generator

instances, primitives, pins, and nets for the entire design netlist. The RTL Viewer of waveform generator is composed of the five modules Divide, Decode, Square, Sine and Triangle. From the RTL Viewer is relatively easy to see how it works.

163.3.3 Simulation

Run simulation after compilation. The Quartus II Simulator tests and debugs the logical operation and internal timing of the design entities in a project. The simulator can perform two types of simulations: functional simulation and timing simulation. Quartus II Analysis and Synthesis must be run successfully before running a functional simulation. Functional simulation is only to test the logical operation of the design. The Timing Analyzer must be run successfully before running a timing simulation. A timing simulation includes timing delays of each device atom block and the routing delays.

After full compilation is successful, to create a Vector Waveform File, perform the following steps: (1) On the File menu, click New. The New dialog box appears. (2) Click the other Files tab, and select Vector Waveform File. (3) Click OK. A blank Waveform Editor window appears. (4) Add nodes and buses. To add a node or bus, on the Edit menu, click Insert Node or Bus. The Insert Node or Bus dialog box appears. All nodes and buses, as well as the internal signals, are listed under Name in the Waveform Editor window. (5) Select any nodes or buses from the Nodes Found list and click > to include it in the waveform, or you can click >> to include all nodes and buses displayed in the Nodes Found list. (6) Click OK. Create a waveform for a signal. After you edit your waveform, save the waveform file wave (.vwf). (7) Run the Simulation [7].

When DD = 00, the output is a square wave. The simulation waveform is shown (Fig. 163.2).

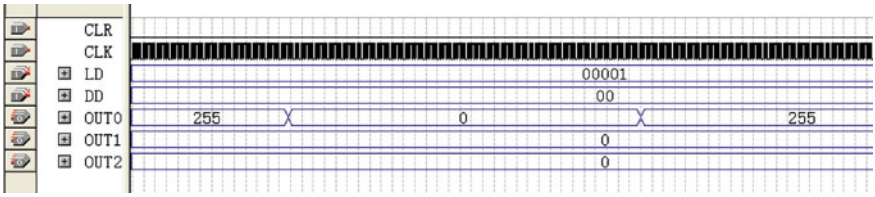


Fig. 163.2 Simulation of square wave

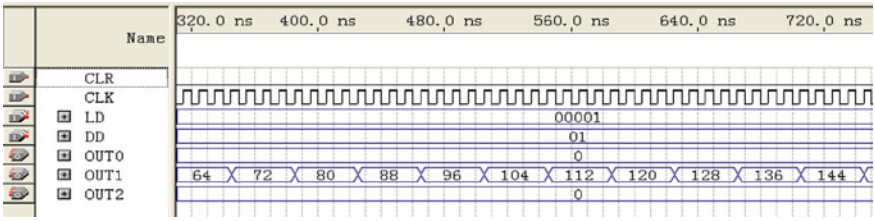


Fig. 163.3 Simulation of triangle wave

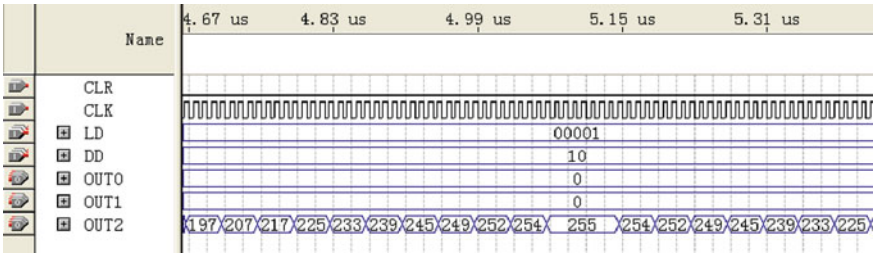


Fig. 163.4 Simulation of sine wave

When DD = 01, the output is a triangular wave. The simulation waveform is shown (Fig. 163.3).

When DD = 10, the output is a sine wave. The simulation waveform is shown (Fig. 163.4).

163.4 Programming

After compilation and simulation are successful, then program the device with the Programmer and Altera programming hardware. The Quartus II Programmer has the same functionality as the Maxplus II Programmer, including programming, verifying, examining, and blank checking operations. Additionally, the Quartus II Programmer now supports the erase capability for CPLD. To improve usability,

the Quartus II Programmer displays all programming-related information in one window. The Assembler automatically converts the Fitter's device, logic cell, and pin assignments into a programming image for the device, in the form of one or more Programmer Object Files (.pof) or SRAM Object Files (.sof) for the target device. The Programmer uses the Programmer Object Files and SRAM Object Files generated by the Assembler to program or configure all Altera devices supported by the Quartus II software.

Open the Tools menu, click Programmer to open the Quartus II Programmer. At first, on the Mode drop-down box, select JTAG. Then, click Hardware Setup, select ByteBlasterMV on the dialog box. Finally, save the programmer settings as a Chain Description File (wave.cdf) and click Start to program the target device.

163.5 Conclusion

In this paper, a waveform signal generator about sine, triangle and square wave, frequency and type of adjustable output waveform is designed, based on EDA platform, the use of VHDL and Quartus II software. As a result of top-down design ideas, simple design, where we did not consider the use of discrete components and the resulting debug, modify the trouble, but also for functional testing and timing simulation and hardware testing, and easy to change and the corresponding into the design file.

As the microelectronics and computer technology rapidly develop, the scale of programmable device is increasing. Many EDA software functions are more comprehensive. The Quartus II software offers an intuitive and user-friendly interface based on customer feedback, and the popular Maxplus II interface. Altera Quartus II design software has been upgraded to 11.0 version. The Quartus II software version 11.0 delivers the production release of Altera's new system-level integration tool known as Qsys [8]. The Qsys system integration tool saves time and effort in the FPGA design process by enabling faster system development and design reuse.

References

1. Altera Corp (2003) Stratix device handbook. Altera Corporation, San Jose
2. Goslin GR, (2010) A guide to using field programmable gate arrays for application-specific digital signal processing performance [D]. www.xinlinx.com.dsp
3. Altera Corp (2003) DSP builder user guide. Altera Corporation, San Jose
4. Idris MYI, Yaacob M, Razak Z (2006) A vhdl implementation of uart design with bist capability[J]. Malaysian J Comput Sci 19(1):73–86
5. Pan S, Huang J, Wang G (2003) Modern DSP techniques. Xidian University Press, Xian
6. Jiang G (2009) VHDL design of digital systems. China Machine Press, Beijing
7. Li X, Zhang H (2009) Digital system design. Science Press, Beijing
8. Altera Corp (2011) Quartus II version 11.0 handbook. Altera Corporation, San Jose

Chapter 164
An Improved Probabilistic Data Association Method for Multi-Target Tracking in Clutter

Yazhao Wang, Yingmin Jia, Junping Du and Fashan Yu

Abstract Improved data association methods for multi-target tracking are discussed in this paper. We introduce an exquisite gating technique based on the well-known probabilistic data association (PDA) filter, which calculates the association probabilities to the targets being tracked for each validated measurement. An important feature of the new algorithm is that it can divide the joint validation region into several parts in which the association probabilities can be calculated with more accuracy. The algorithm is tested and compared with both the generic PDA and joint probabilistic data association (JPDA) filter. Simulations are also presented to illustrate the effectiveness of the algorithm.

Keywords Multi-target tracking · Probabilistic data association · Gating technique · Validation region

Y. Wang (✉)
The Seventh Research Division and the Department of Systems and Control,
Beihang University (BUAA), Beijing 100191, China
e-mail: yzwang@ss.buaa.edu.cn

Y. Jia
Key Laboratory of Mathematics, Informatics and Behavioral Semantics (LMIB),
Ministry of Education, SMSS, Beihang University (BUAA), Beijing 100191, China
e-mail: ymjia@buaa.edu.cn

J. Du
Beijing Key Laboratory of Intelligent Telecommunications
Software and Multimedia, School of Computer Science and Technology,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
e-mail: junpingdu@126.com

F. Yu
School of Electrical Engineering and Automation, Henan Polytechnic University,
Jiaozuo 454000, Henan, China
e-mail: yufs@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_164,

164.1 Introduction

Data association plays an important role in multi-target tracking systems and a number of strategies are available to solve this problem [1–3]. One famous solution is the probabilistic data association (PDA) filter, which associates the validated measurements with the current tracks using a probabilistic measure [4, 5]. Thus, the tracks can be updated with a weighted sum of the measurements. This was proved to be a very effective method for tracking a single target in dense clutter with low computational time cost. However, for multi-target tracking, the PDA filter does not consider the measurements originated from other targets in the validation region of a particular target, and tracking will probably be failed [6]. To account for this problem, an extension was derived called joint probabilistic data association (JPDA) [5, 7]. In JPDA, the joint association probabilities are calculated over all targets and measurements, however, coming up with high time cost in calculating the joint probabilities. Meanwhile, many types of gating technique have been studied to improve the capability of tracking algorithms [8–10]. Since the measurement uncertainty has close relations with gating selections, we aim to develop a more efficient method by using the gating techniques. In this paper, we present an exquisite gating technique which can divide the joint validation region into several parts so that the gate weight in each sub gating region can be obtained for each target. This is followed by a new suboptimal algorithm termed exquisite PDA (EPDA), and the new method is compared with the generic PDA and JPDA.

The remainder of this paper is structured as follows. The state space model for multi-target tracking is described in Sect. 164.2. In Sect. 164.3, an exquisite gating technique is introduced and incorporated into the framework of PDA filter. Simulation results and conclusions are given in Sects. 164.4 and 164.5, respectively.

164.2 Tracking Model

In multi-target tracking, there are two stages of processing, data association and filtering. Data association is to work out which measurement is originated from which target, that is to say, we first require a partition of measurements to know the sequence of correct measurements associated with each target. Filtering is used to estimate the tracks with the assigned measurements. Consider the following linear dynamic and measurement state space model of target j,

$$x_k^{(j)} = Fx_{k-1}^{(j)} + w_{k-1}^{(j)} \tag{164.1}$$

$$z_k^{(j)} = Hx_k^{(j)} + v_k^{(j)} \tag{164.2}$$

where $x_k^{(j)}$ are state vector, $w_{k-1}^{(j)}$ and $v_k^{(j)}$ are white Gaussian noise with zero mean and covariance $Q_{k-1}^{(j)}$ and $R_k^{(j)}$, respectively. F and H are matrices with compatible sizes. At time step k, tracks are modeled as a Markov process and depend on the previous results at time step $k-1$. The state vector, the measurement, and the state covariance matrix are predicted as in Kalman filter [11, 12],

$$\hat{x}_{k|k-1}^{(j)} = F\hat{x}_{k-1|k-1}^{(j)}, \quad \hat{z}_{k|k-1}^{(j)} = H\hat{x}_{k|k-1}^{(j)}, \quad P_{k|k-1}^{(j)} = FP_{k-1|k-1}^{(j)}F^T + Q_k^{(j)}, \quad (164.3)$$

where the covariance $P_{k-1|k-1}^{(j)}$ is available from time step $k-1$. The innovation covariance matrix is

$$S_k^{(j)} = HP_{k|k-1}^{(j)}H^T + R_k^{(j)}. \quad (164.4)$$

After matching the measurements and tracks by data association, each track is updated with the associated measurements $z_k^{(j)}$ as

$$\hat{x}_{k|k}^{(j)} = \hat{x}_{k|k-1}^{(j)} + W_k^{(j)}(z_k^{(j)} - \hat{z}_{k|k-1}^{(j)}), \quad P_{k|k}^{(j)} = P_{k|k-1}^{(j)} - W_k^{(j)}P_{k-1|k-1}^{(j)}W_k^{(j)T}, \quad (164.5)$$

where the gain $W_k^{(j)} = P_{k|k-1}^{(j)}H^T(S_k^{(j)})^{-1}$.

164.3 Improved Probabilistic Data Association Filter

The PDA assumes a measurement is validated for target j if it falls inside the elliptical region centered at $\hat{z}_j$ (the predicted measurement of target j) as shown in Fig. 164.1, and all the validated measurements in this particular validation region are candidates to be association with target j. However, if the measurements from other targets are also in this validation region, the obtained association probabilities will be incorrect and tracking may possibly fail. For example, PDA filter is correct for Fig. 164.1 (left) but wrong for Fig. 164.1 (right). To solve this problem, we introduce a gating technique to divide the validation region into parts by using the information as shown in Fig. 164.1, so that the data association can be restricted to the sub-validation-regions. Thus, the sub gate weights $P_r^{(j)}$ (r is the identifier of the sub validation) can be calculated and the data association probabilities will be estimated with more accuracy.

164.3.1 Exquisite Gating Technique

Now, we will show how to calculate $P_r^{(j)}$ by using an exquisite gating technique (a similar scheme has been incorporated into the RBPF data association [13]).

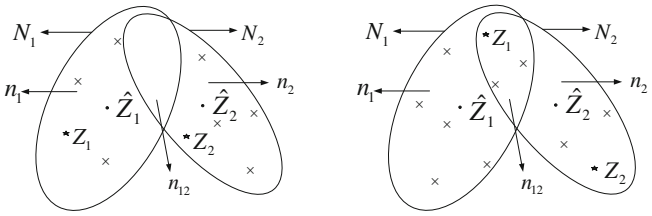


Fig. 164.1 Two possible examples of measurement validation. The target originated measurements $z_j(j = 1, 2)$ and clutter originated measurements "×" are depicted. In this example, the number of validated measurements in two elliptical regions is N_j, and the number of validated measurements in sub region $r(r = 1, 2, 3)$ is n_r

At time step k and for all targets $j(j = 1, \cdots, T)$, the computational formula is as follows:

$$N_j = n_j + \sum_{\mathrm{Con}(1)} n_{jk_1} + \sum_{\mathrm{Con}(2)} n_{jk_1k_2} +, \cdots, + \sum_{\mathrm{Con}(T-1)} n_{jk_1k_2\cdots k_{T-1}} \tag{164.6}$$

$$\begin{aligned} P_j^{(j)} &= n_j/N_j;\ P_{jk_1}^{(j)} = n_{jk_1}/N_j, \quad k_1 \in \mathrm{Con}(1); \\ &\vdots \end{aligned} \tag{164.7}$$

$$P_{jk_1k_2\cdots k_{T-1}}^{(j)} = n_{jk_1k_2\cdots k_{T-1}}/N_j, \quad \{k_1, k_2, \cdots, k_{T-1}\} \in \mathrm{Con}(T-1), \tag{164.8}$$

where,

$$\begin{aligned} \mathrm{Con}(1) &= \{k_1 \ :\ 1 \le k_1 \le T \,\mathrm{and}\, k_1 \ne j\}, \\ &\vdots \end{aligned} \tag{164.9}$$

$$\begin{aligned} \mathrm{Con}(T-1) = \{&(k_1, k_2, \cdots, k_{T-1}) \ :\ k_1, k_2, \cdots, k_{T-1} \\ &\text{are } T-1 \text{ different numbers in } \{1, 2, \cdots, T\}/\{j\}\}. \end{aligned} \tag{164.10}$$

N_j is the number of validated measurements in the elliptical region of target j, and $n_{jk_1k_2\cdots k_t}\,(t = 1, 2, \cdots, T-1)$ is the number of validated measurements in the intersection (if present) of validation regions which are related to a targets with the identifier $\{j, k_1, k_2, \cdots, k_t\}$. In addition, $n_{jk_1k_2\cdots k_t}$ is independent of the order of $j, k_1, k_2, \cdots, k_t$. That is, for any permutations $\{j, i_1, i_2, \cdots, i_t\}$ of $\{j, k_1, k_2, \cdots, k_t\}$, $n_{jk_1k_2\cdots k_t}$ does not change. As a result, $P_{jk_1k_2\cdots k_t}^{(j)}$ is the estimated gate weight of target j in the sub validation region $jk_1k_2\cdots k_t$.

164.3.2 Probabilistic Data Association Filter Using Exquisite Gating Technique

The exquisite gating technique based probabilistic data association algorithm (EPDA) is presented in Algorithm 1. The algorithm includes four major parts: prediction, measurement validation, data association and finally state estimation. In this algorithm, dividing the joint validation region and calculating the gate weights are quite different from the conventional PDA filter, while the association calculation steps just need some small changes.

Algorithm 1 *EPDA for Multi-Target Tracking in Clutter*

1. For $k = 1$ to n, and $j = 1$ to T do
2. Do prediction step, the procedures are given as (164.3) and (164.4).
3. Select validated measurements and calculate gate weight $P_r^{(j)}$ for target j in each sub-validation-region r according to (164.6–164.10).
4. Compute the association probability for $z_{r,k}^{(j)}$ being the correct measurement, as

$$\beta_{i,r,k}^{(j)} = \begin{cases} \dfrac{P_r^{(j)} E_{i,k}}{1 - P_D^{(j)} P_G^{(j)} + \sum_{j=1}^{m_{r,k}^{(j)}} E_{i,k}}, & i = 1, \cdots, m_{r,k}^{(j)} \\ \dfrac{P_r^{(j)} E_{i,k}}{1 - P_D^{(j)} P_G^{(j)} + \sum_{j=1}^{m_{r,k}^{(j)}} E_{i,k}}, & i = 0 \end{cases} \tag{164.11}$$

where $P_D^{(j)}$ and $P_G^{(j)}$ is the detection probability and gate probability of target j, respectively. And $i = 0$ means that none is correct, $r = 1, \cdots, G_k^{(j)}$ is the identifier of sub-validation-region, ($G_k^{(j)}$ is the number of the sub-validation-regions which contain the validated measurements from target j), and

$$E_{i,k} = N\left[z_{r,k}^{(j)}; \hat{z}_{k|k-1}^{(j)}, S_k^{(j)}\right] P_D^{(j)} \Big/ \lambda \tag{164.12}$$

is the likelihood ratio of the measurement $z_{r,k}^{j}$ originating from the target rather than from clutter. Note that, the set of validated measurements of target j is $z_k^{(j)} = \left\{z_{r,k}^{(j)}\right\}_{r=1}^{G_k^{(j)}}$, where $z_{r,k}^{(j)} = \left\{z_{i,r,k}^{(j)}\right\}_{r=1}^{m_{r,k}^{(j)}}$, with $m_{r,k}^{(j)}$ the number of the validated measurements in sub-validation-region r at time step k. And the number of clutter measurements obeys a Poisson distribution with the Poisson number λ [5].

5. Do state estimation step (filtering).

a. Calculate the gain $W_k^{(j)} = P_{k|k-1}^{(j)} H^T (S_k^{(j)})^{-1}$ in the same way as Kalman filter.

b. Calculate the updated state as

$$\hat{x}_{k|k}^{(j)} = \hat{x}_{k|k-1}^{(j)} + W_k^{(j)} v_k^{(j)} \tag{164.13}$$

where the combined innovation is

$$v_k^{(j)} = \sum\nolimits_{r=1}^{G_k^{(j)}} \sum\nolimits_{i=1}^{m_{r,k}^{(j)}} \beta_{i,r,k}^{(j)} v_{i,r,k}^{(j)}, \tag{164.14}$$

with $v_{i,r,k}^{(j)} = z_{i,r,k}^{(j)} - \hat{z}_{k|k-1}^{(j)}$.

c. Calculate the covariance of the updated state as

$$P_{k|k}^{(j)} = \sum\nolimits_{r=1}^{G_k^{(j)}} \beta_{0,r,k}^{(j)} P_{k|k-1}^{(j)} + \sum\nolimits_{r=1}^{G_k^{(j)}} (1 - \beta_{0,r,k}^{(j)}) \bar{P}_{k|k}^{(j)} + \tilde{P}_k^{(j)}, \tag{164.15}$$

where the covariance of state updated with the correct measurement is

$$\bar{P}_{k|k}^{(j)} = P_{k|k-1}^{(j)} - W_k^{(j)} S_k^{(j)} (W_k^{(j)})^T, \tag{164.16}$$

and the spread innovations term is

$$\tilde{P}_k^{(j)} \triangleq W_k^{(j)} \left[\sum\nolimits_{r=1}^{G_k^{(j)}} \sum\nolimits_{i=1}^{m_{r,k}^{(j)}} \beta_{i,r,k}^{(j)} v_{i,r,k}^{(j)} (v_{i,r,k}^{(j)})^T - v_k^{(j)} (v_k^{(j)})^T \right] (W_k^{(j)})^T. \tag{164.17}$$

6. End for.

164.4 Simulation Results

We model each target with constant velocity model in 2-dimensional Cartesian coordinates. The dynamic and measurement models are with the same form of (164.1) and (164.2), where

$$F = \begin{pmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad Q_{k-1}^{(j)} = q \begin{pmatrix} T^3/3 & 0 & T^2/2 & 0 \\ 0 & T^3/3 & 0 & T^2/2 \\ T^2/2 & 0 & T & 0 \\ 0 & T^2/2 & 0 & T \end{pmatrix},$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

with the sample interval $T = 0.1$ and the spectral density of the noise $q = 0.1$. To save simulations time, we set the local rectangular detection region centered at (x_k, y_k), where $x_k = 1/2 \sum_{j=1}^{2} x_k^{(j)}$ and $y_k = 1/2 \sum_{j=1}^{2} y_k^{(j)}$, with the area $s = 96$. We set the other parameters as: $P_D^{(j)} = 0.8$, $P_G^{(j)} = 0.9997$ with the gate threshold $\gamma = 16$, and the total time steps $n = 180$. All simulations are performed on a PC with a 2.8 GHz Intel processor.

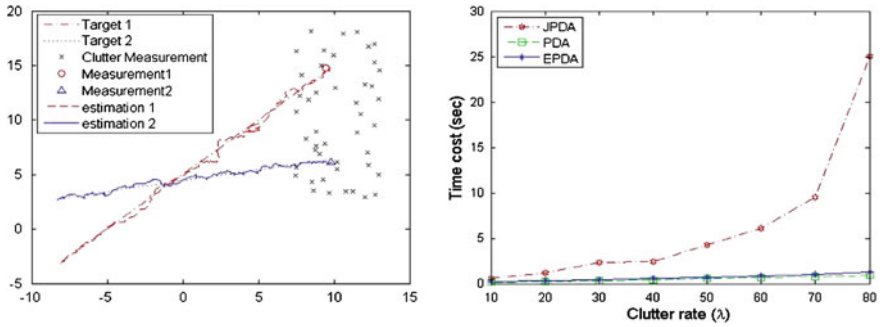


Fig. 164.2 Example trajectories and comparison of the average time cost

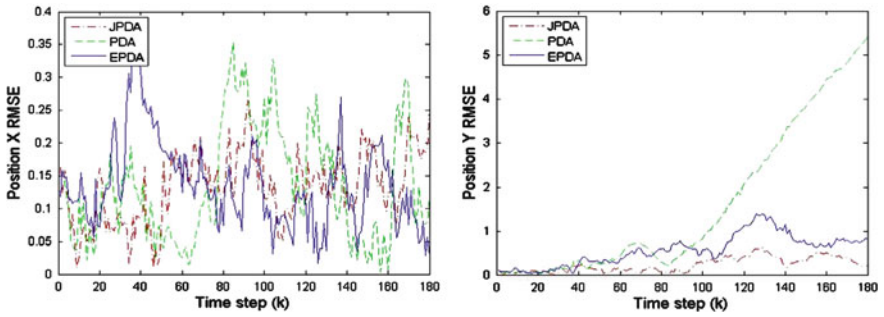


Fig. 164.3 Comparisons of the average position estimation errors

In the simulation, we consider a typical two-target-crossing example as shown in Fig. 164.2 (left). The time cost of the three algorithms with different λ is shown in Fig. 164.2 (right), it can be seen that the computational time of EPDA is nearly the same as that of the generic PDA but far less than JPDA, especially in dense clutter environment. Set $\lambda = 40/\text{s}$, Fig. 164.3 shows the root mean square error (RMSE) of positions and velocities of the three algorithms from 50 different runs. We can see that the EPDA has slight larger RMSE values than JPDA but much lower than PDA.

164.5 Conclusions

In this paper, we have shown that the PDA filter can be significantly improved by using the exquisite gating technique. This approach (EPDA) goes some way to compensate for the weakness of the PDA and needs far less time than the JPDA filter. Future work should be done to study the tracking abilities of EPDA and apply this method to other gate based multi-target tracking algorithms.

Acknowledgments This work was supported by the NSFC (60727002, 60774003, 60921001, 90916024), the MOE (20030006003), the COSTIND (A2120061303), the National 973 Program (2005CB321902).

References

1. Bar-Shalom Y, Fortmann TE (1988) Tracking and data association. Academic, San Diego
2. Blackman SS (2004) Multiple hypothesis tracking for multiple target tracking. IEEE Aerosp Electron Syst Mag 19(1):5–18 Jan
3. Oh S, Russell S, Sastry S (2009) Markov chain Monte Carlo data association for multi-target tracking. IEEE Trans Autom Control 54(3):481–497
4. Kirubarajan T, Bar-Shalom Y (2004) Probabilistic data association for target tracking in clutter. Proc IEEE 92(3):536–557
5. Bar-Shalom Y, Daum F, Huang J (2009) The probabilistic data association filter: estimation in the presence of measurement origin uncertainty. IEEE Control Syst Mag 29:82–100
6. Fitzgerald RJ (1985) Track biases and coalescence with probabilistic data association. IEEE Trans Aerosp Electron Syst 21(6):822–825
7. Roecker JA (1994) A class JPDA of near optimal algorithms. IEEE Trans Aerosp Electron Syst 30(2):504–510
8. Collins J, Uhlmann J (1992) Efficient gating in data association with multivariate distributed states. IEEE Trans Aerosp Electron Syst 28(3):909–916
9. Wang X, Challa S, Evans R (2002) Gating techniques for maneuvering target tracking in clutter. IEEE Trans Aerosp Electron Syst 38(3):1087–1097
10. Musicki D, Morelande MR (2004) Gate volume estimation for target tracking. In: International conference on information fusion
11. Ristic B, Arulampalam S, Gordon N (2004) Beyond the Kalman filter. Artech House, Boston
12. Zhongzhi L, Xuegang W (2010) Reverse prediction adaptive Kalman filtering algorithm for maneuvering target tracking. J Comput Inf Syst 6(10):3257–3266 Oct
13. Wang Y, Jia Y (2011) Adaptive measurement validation in multi-target tracking using local information

Chapter 165
Improved Wavelet Threshold Denoising Method for Railway Signal

Juan Zou, Shijie Jia and Shaohua Chen

Abstract Railway signals are inevitably mixed with a variety of noise. Therefore noise reduction becomes crucial in railway operations. In this paper, the wavelet threshold denoising method was applied to railway signals. A new wavelet threshold function was proposed to overcome the disadvantages of soft threshold and hard threshold. The new method is more smoothing. More details of the signals are reserved. The obtained simulation results show that the new thresholding algorithm has better denoising effect in signal to noise ratio (SNR), minimum mean square error (MSE), and results in closer approximation to the real signals, compared with hard thresholding and soft thresholding methods.

Keywords Wavelet analysis · Threshold function · Railway signal de-noise

165.1 Introduction

The railway signal is the most important factor in train control systems. It carries instructions and information to control the train line, time interval and speed [1]. Unfortunately, as railway signal transmission channels are very bad, signals are

J. Zou (✉) · S. Jia · S. Chen
School of Electrical and Information Engineering,
Dalian Jiaotong University, Dalian 116028, China
e-mail: zouliangman@sina.com

S. Jia
e-mail: jiashijie@dl.cn

S. Chen
e-mail: chengshineng@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_165,

usually contaminated by a variety of noise. In addition to traditional white gauss noises, the noise occurs mainly due to interference of unbalanced traction current at 50 Hz power frequency and its odd harmonics frequency (especially 150, 250 Hz). Therefore, the noise reduction becomes crucial in the railway operation based on analysis of railway signals.

There are mainly three railway FSK signals [2] used in China: the AC counting signal, the 18 information FSK signal and the UM-71 FSK signal. The carrier frequencies of these signals cover widely from a few tens Hertz to about 2600 Hz. Traditional linear time-invariant filtering is not suitable for the railway signal denoising, because this linear filtering approach cannot separate noise from signal where their Fourier spectra overlap. However, wavelet analysis is entirely different. The chief advantage of wavelet transform is the ability of partial and detailed analysis in desired scales. Because the main noisy energy is centralized in detailed component of wavelet decomposition, the wavelet threshold function can be applied to separated signals to remove noise. The method of wavelet denoising has been extensive applied in image compression, feature extraction, edge detection, and so on [3]. In this paper, the wavelet threshold denoising method is employed, and a new thresholding method is proposed based on hard and soft thresholding methods.

The remainder of this paper is organized as follows. In Sect. 165.2, we review the traditional wavelet threshold denoising methods and introduce the wavelet threshold denoising methods. In Sect. 165.3, we simulate these three threshold denoising methods to study their performance. Finally, In Sect. 165.4, we draw some conclusions and present the directions of future work.

165.2 Wavelet Denoising

The general wavelet denoising procedure is as follows [4–6]:

- Apply wavelet transform to the noisy signal to produce the noisy wavelet coefficients.
- Select appropriate threshold limit at each level and threshold method (hard or soft thresholding) to best remove the noises.
- Reconstruct the signal (apply the inverse wavelet transform) to obtain a reconstructed signal.

165.2.1 Hard Thresholding

Hard thresholding sets zeros for all wavelet coefficients whose absolute value is less than the specified threshold limit. For threshold λ, the hard thresholding algorithm is expressed as follows:

$$\tilde{d}_{j,k} = \begin{cases} \mathrm{d}_{j,k}, & |\mathrm{d}_{j,k}| \geq \lambda \\ 0, & |\mathrm{d}_{j,k}| < \lambda \end{cases} \tag{165.1}$$

where $d_{j,k}$ is the wavelet coefficient of noisy signal, $\tilde{\mathrm{d}}_{j,k}$ is the wavelet coefficient after thresholding, λ is the threshold, which is the universal threshold, given by [7]:

$$\lambda = \sigma\sqrt{2\lg N} \tag{165.2}$$

Hard thresholding maintains the scale of the signal but introduces ringing and artifacts after reconstruction due to a discontinuity in the wavelet coefficients.

Where N is the length of the signal, σ is the noise standard deviation, its standard deviation can be estimated from the median of its detail coefficients $\mathrm{d_j}$, with j = 1..M, and is computed as follows:

$$\sigma = \frac{\text{median}(|\mathrm{d_j}|)}{0.6745} \tag{165.3}$$

Hard thresholding maintains the scale of the signal but introduces ringing and artifacts after reconstruction due to a discontinuity in the wavelet coefficients.

165.2.2 Soft Thresholding

Soft thresholding sets all coefficients below the threshold to zero and also reduces the magnitude of remaining coefficients by the threshold value. The soft thresholding algorithm is given by:

$$\tilde{d}_{j,k} = \begin{cases} d_{j,k} - \lambda, & d_{j,k} \geq \lambda \\ 0, & |d_{j,k}| < \lambda \\ d_{j,k} + \lambda, & d_{j,k\leq} - \lambda \end{cases} \tag{165.4}$$

Where$\lambda$is defined by equation (165.2), (165.3).

Soft thresholding eliminates this discontinuity resulting in smoother signals but slightly decreases the magnitude of the reconstructed signal.

165.2.3 New Thresholding

This algorithm is combined hard thresholding with soft thresholding. Taking into account the larger threshold λ, the threshold is reduced from λ to $\lambda/2$. When $d_{j,k} \geq \lambda$, its procession is the same as hard thresholding, retaining the wavelet coefficient. When $\lambda/2 < |d_{j,k}| < \lambda$, the wavelet coefficient is performed according to the new thresholding method. When $d_{j,k} \geq \lambda$, it is performed as soft thresholding. The expression is as follows:

$$\tilde{d}_{j,k} = \begin{cases} d_{j,k}, & |d_{j,k}| \geq \lambda \\ -\frac{(2d_{j,k}+\lambda)^2}{\lambda} - \frac{1}{4}\left(\frac{d_{j,k}+\lambda}{\lambda}\right)^3, & -\lambda < d_{j,k} \leq -\frac{\lambda}{2} \\ 0, & |d_{j,k}| < \frac{\lambda}{2} \\ \frac{(2d_{j,k}-\lambda)^2}{\lambda} - \frac{1}{4}\left(\frac{d_{j,k}-\lambda}{\lambda}\right)^3, & \frac{\lambda}{2} < d_{j,k} \leq \lambda \end{cases} \tag{165.5}$$

This new method not only has the advantage of hard thresholding and soft thresholding method but also overcomes their shortcoming. It is more smoothing. And more details of the signals are reserved, which results in closer approximation to the real signals.

In this method, we choose a different thresholding function. It is given by:

$$\lambda = \sigma\sqrt{2\ln L_j}/j \quad j = 1, \ldots, J \tag{165.6}$$

Where σ is defined by Eq. (165.2), L_j is the length of wavelet decomposition coefficients on the jth level. J is the maximal number of the wavelet decomposition levels.

This threshold function is not universal, but individual on every level, which is different from the threshold function used in hard and soft methods.

165.3 Experiments and Results

There are mainly two railway signals in China: AC counting signal and railway FSK signal. In this paper, we choice the railway FSK signal to be de-noised.

165.3.1 Railway FSK Signal

The base frequency signal of the FSK is a square wave, which is given by:

$$f(t) = \begin{cases} A & -T/4 \leq t < T/4 \\ -A & T/4 \leq t < 3T/4 \end{cases} \tag{165.7}$$

where A is the amplitude of the square wave, T is its period. The carrier signal with the frequency f_0 is modulated by $f(t)$. The offset angular frequency $\Delta\omega$ and the offset frequency Δf of the FSK signal can be expressed by:

$$\Delta\omega = kA, \Delta f = kA/2\pi \tag{165.8}$$

where, k is a coefficient, which denotes the sensitivity of the frequency shifter, the unit is Hz/V. The Instantaneous angular frequency of the FSK signal is:

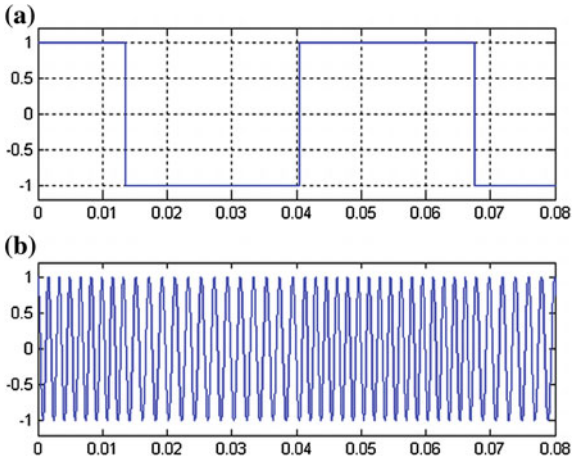


Fig. 165.1 Railway frequency-shift Signal; **a** the square wave with period T = 1/18.5 s, base frequency f = 18.5 Hz and amplitude A = 1; **b** the railway FSK signal modulated by the square wave with 550 Hz base frequency and 55 Hz offset frequency

$$\omega(t) = \begin{cases} \omega 0 + \Delta\omega & -T/4 \le t < T/4 \\ \omega 0 - \Delta\omega & T/4 \le t < 3T/4 \end{cases} \tag{165.9}$$

The Instantaneous phase of the frequency-shift signal is:

$$\theta(t) = \int \omega(t)\mathrm{d}t = \omega_0 t + \int kf(t)\mathrm{d}t = \omega_0 t + g(t) \tag{165.10}$$

Where,

$$g(t) = \int kf(t)\mathrm{d}t = \begin{cases} \Delta\omega \times t & -T/4 \le t < T/4 \\ \Delta\omega \times (T/2 - t) & T/4 \le t < 3T/4 \end{cases} \tag{165.11}$$

Time domain expression of FSK signal is defined as follows:

$$S(t) = A\cos\theta(t) = A\cos\theta[\omega_0 t + g(t)] \tag{165.12}$$

165.3.2 Results

Here, we adopt the 18 information FSK system. Figure 165.1a shows the square wave signal with an 18.5 Hz base frequency. And Fig. 165.1b shows the FSK signal with a 550 Hz carrier frequency and a 55 Hz offset frequency.

In addition to the white Gaussian noise, the noise corrupted to the railway signals mainly is 50 Hz power frequency and its odd harmonics (especially 150, 250 Hz) interference, so in the simulation trial the noise added to original signal shown in Fig. 165.1 consists of white Gaussian noise, 50 Hz power frequency,150 and 250 Hz noise. And the 50 Hz power frequency and its odd harmonics noise

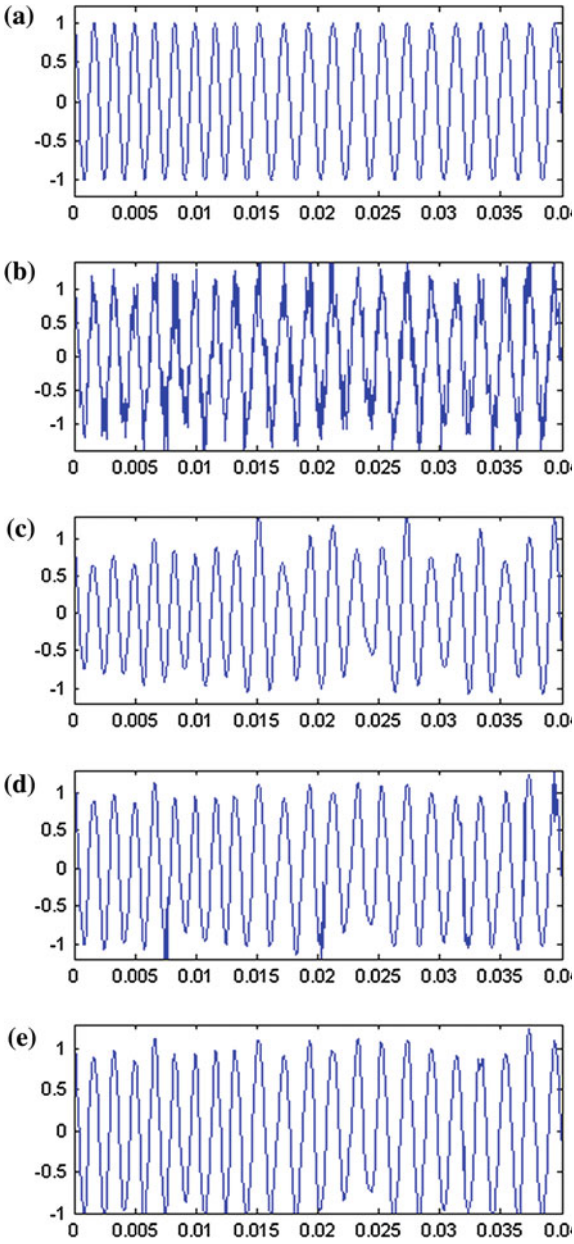


Fig. 165.2 Result of the wavelet denoising algorithm; **a** the original railway frequency-shift signal; **b** the noisy signal with SNR being 8; **c** the result of soft thresholding algorithm; **d** the result of hard thresholding algorithm; **e** the result of the new thresholding algorithm

Table 165.1 SNR and MSE of the reconstructed signals using three different algorithms

Algorithm	SNR	MSE
Noisy signal	7.9460	0.2832
Soft thresholding	10.5833	0.2090
Hard thresholding	13.9923	0.1412
Proposed thresholding	15.0430	0.1251

consists chiefly of 50 Hz frequency counting for 70%, 150 Hz noise counting for 20% and 250 Hz noise counting for 10%. The noisy railway FSK signal with SNR of 8 dB is shown in Fig. 165.2b.

Figure 165.2c–e illustrate the simulation results of soft thresholding method, hard thresholding method and new thresholding method respectively. We should note that, along this work, we choose 4-layer wavelet decomposition, and the wavelet function 'db7' is used in the wavelet denoising process. Figure 165.2 shows that this new method is more smoothing than hard thresholding method and reserves more details of the signals than the soft thresholding method.

Besides SNR being commonly used, the measure is used to compare the relative performance of the denoising methods is the mean squared error (MSE) defined as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{n}(s(n) - y(n))^2 \tag{165.13}$$

The symbol s(n) represents the original noiseless signal and y(n) represents the reconstructed signal.

The analysis of Table 165.1 reveals the new thresholding method outperforming traditional hard thresholding and soft thresholding methods in both SNR and MSE.

165.4 Conclusions

The wavelet threshold denoising methods for noisy railway signal is discussed. A new wavelet threshold denoising method is proposed to overcome the disadvantages of the soft threshold and hard threshold. The obtained simulation results show the new thresholding method outperforms traditional hard thresholding and soft thresholding methods in both SNR and MSE and is closer approximation to the real signals. Our future research will focus on finding more accommodating wavelet and employing the new method for other kinds of noise.

References

1. Yu D (2008) Section signal and train control syitem. Railway Publishing House, Beijing, pp 38–64
2. Hu YH (2002) The present situation and development trend of chinese railway signal. Chinese Railway pp 39–41
3. Yang J (2005) Wavelet analysis and its engineering applications. China Machine Press, China, pp 108–110
4. Donoho DL, Johnstone LM (1995) Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 12:1200–1224
5. Zhu YQ, Yang XL (2008) Several new methods based on wavelet thresholding denoising. Electron Test 2(2):18–22

6. Liu HB, Han SQ (2007) Wavelet de-noising based on novel thresholding function and best decomposition scale. Comput Eng Appl 43(24):72–74
7. Donoho DL (1995) Denoising by soft thresholding. IEEE Trans Inf 3:613–627

Chapter 166
Performance of 100 Gbit/s Transmission System Based on Non-Return to Zero and Return to Zero in Single Mode Fiber High-dispersion Fiber

Chunbo Jiang

Abstract Transmission performance of 100 Gbit/s signal with non-return to zero (NRZ) and return to Zero (RZ) in single mode fiber (SMF) high-dispersion fiber is investigated. The bit error rate (BER) against transmission distance, laser power and chromatic dispersion is simulated and analyzed. The maximum transmission distance is obtained for different modulation format. Comparison of NRZ and RZ transmission over SMF high-dispersion fiber, it is shown clearly the advantage of NRZ modulation scheme for 100 Gbit/s transmission over long SMF link. The scheme of chromatic dispersion compensation is also studied.

Keywords Bit error rate · Non-return to zero (NRZ) · Return to zero (RZ)

166.1 Introduction

To date, high speed transmission has been investigated as a means for upgrading the transmission capacity of the currently installed fiber network. The currently installed fiber network comprises predominantly standard non dispersion-shifted fiber or SMF, exhibiting chromatic dispersion values of D = 16 ps/km/nm at a wavelength of 1550 nm. For applications of 100 Gbit/s transmission in this network, an important issue to investigate is whether NRZ (non return-to-zero) or RZ (return-to-zero) is the more optimal modulation scheme for obtaining large transmission spans [1–4]. Moreover, next-generation high-speed transmission

C. Jiang (✉)
Potevio institute of technology, No.6 Haidian North, Second Street, Haidian, Beijing, 100080, People's Republic of China
e-mail: bocoj@cpit.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_166,

systems will require techniques to compensate for signal distortion due to dispersion and fiber nonlinearities.

In this paper, the effect of modulation formats in 100 bit/s system based SMF high-dispersion fiber is investigated. The transmitter and receiver scheme of RZ and NRZ formats is depicted firstly. The BER against transmission distance, laser power and chromatic dispersion is simulated and analyzed. More, the performance of individual dispersion compensation by utilizing erbium-doped optical fiber amplifier (EDFA) is studied. Comparison of NRZ and RZ transmission over SMF high-dispersion fiber, it is shown clearly the advantage of NRZ modulation scheme for 100 Gbit/s transmission over long SMF link.

166.2 Transmitter and Receiver Scheme of 100 Gbit/s NRZ and RZ Using Mach–Zehnder Interferometer

The block diagram of the typical optical transmission system over SMF high-dispersion fiber is shown in Fig. 166.1, which comprises a 100 Gbit/s transmitter, a transmission line and a receiver. The transmitter is different in the RZ and NRZ transmission. In both cases, a 100 Gbit/s pseudorandom bit sequence (PRBS) of length $2^{15} - 1$ is generated at a wavelength of 1550 nm. In the NRZ scheme, the output is an optical NRZ data signal obtained by modulation of the CW light of a distributed feedback (DFB) laser in a LiNbO3 Mach–Zehnder modulator. In the RZ scheme, the transmitter comprises a tunable modelocked semiconductor laser, an external modulator and a fiber delay-line multiplexer. The transmission line consists of SMF, dispersion compensation and EDFA. In optical receiver, the optical signal emerging from the transmission link is filtered by an optical filter and detected by a photoelectric detector (PD). The output signal of PD is processed by electronic low pass Bessel filter. Then the data is received.

166.3 Performance Simulation

In this section, the transmission performance of 100 Gbit/s RZ and NRZ is investigated with respect to chromatic dispersion (CD), transmission distance and laser power. Moreover, compensation of chromatic dispersion is also studied.

166.3.1 BER Performance Without Compensation of Chromatic Dispersion

Figure 166.2 shows the BER as a function of transmission distance for NRZ modulation format. From Fig. 166.2, it can be seen that the maximum SMF length

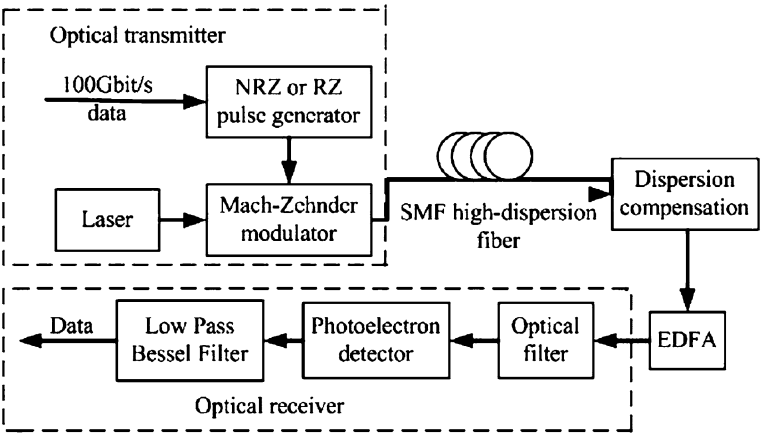


Fig. 166.1 Block diagram of the typical optical transmission system over SMF high-dispersion fiber

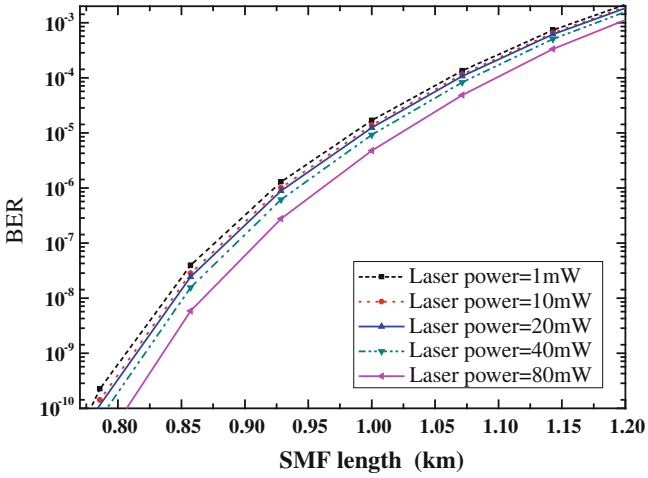


Fig. 166.2 BER against transmission distance and laser power for NRZ modulation format

of laser power 1, 10, 20, 40 and 80 mw for BER $= 10^{-9}$ is found to be 0.81, 0.82, 0.821, 0.824 and 0.84 km, respectively. The BER degrades with the increase of SMF length when the laser powers are the same. The BER is improved with the increase of the laser power when the transmission distances are the same. At the condition of BER less than 10^{-9}, the maximum transmission distance is less than

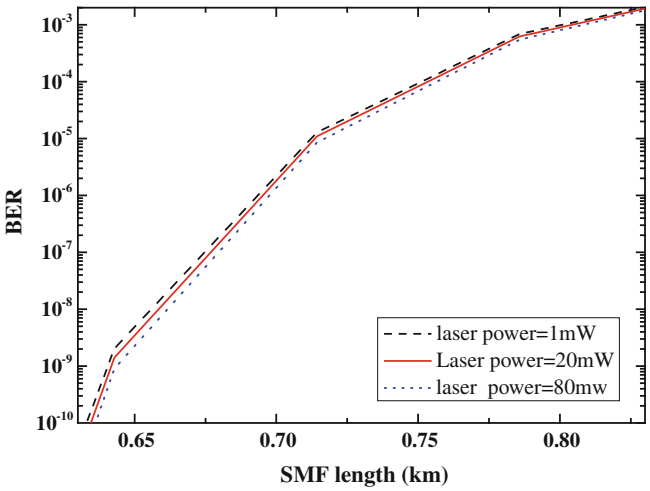


Fig. 166.3 BER against transmission distance and laser power for RZ modulation format

0.85 km when the laser power is less than 80 mw. At the condition of BER less than 10^{-3}, the maximum transmission distance is less than 1.2 km when the laser power is less than 80 mw. Note that such predetermined BER threshold (10^{-3}) can achieve the corrected BER, providing the forward-error correction (FEC) overhead with enhanced Reed–Solomon (RS) with concatenation is applied [5].

Figure 166.3 shows the BER as a function of transmission distance for RZ modulation format. From Fig. 166.3, it can be seen that the maximum SMF length of laser power 1, 20 and 80 mw for BER $= 10^{-9}$ is found to be 0.635, 0.64, and 0.642 km, respectively. The BER also degrades with the increase of SMF length when the laser powers are the same. The BER is also improved with the increase of the laser power when transmission distances are the same. At the condition of BER less than 10^{-9}, the maximum transmission distance is less than 0.65 km when laser power is less than 80 mw. If the FEC technology is applied in system at the condition of BER less than 10^{-3}, the maximum transmission distance is less than 1 km when the laser power is less than 80 mw.

Comparison of Figs. 166.2 and 166.3 shows that the maximum transmission distance of NRZ system is larger than that of RZ system when the laser powers are the same. Moreover, both maximum transmission distances are less than 1.2 km. This is caused by the chromatic dispersion of SMF fiber. In order to realize long distance transmission, the chromatic dispersion must be compensated.

Assuming 80 mw of laser power and 0.7 km of SMF length, the curve of BER against chromatic dispersion coefficient of SMF is shown in Fig. 166.4. It can be shown that the BER degrades with the chromatic dispersion coefficient of SMF for NRZ and RZ modulation format. Moreover, the variance of BER in RZ modulation format system is larger than that of NRZ modulation format.

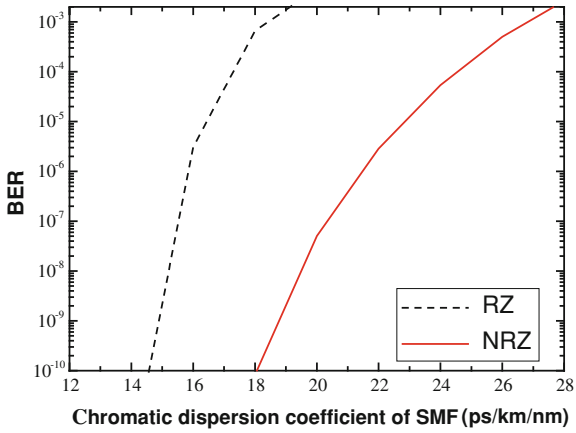


Fig. 166.4 BER against Chromatic dispersion coefficient of SMF with different modulation formats

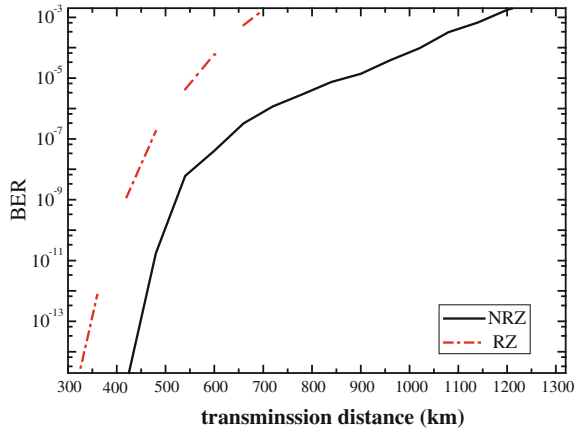


Fig. 166.5 BER performance after compensation of chromatic dispersion when the laser power is 1 mw with different modulation formats

166.3.2 Compensation of Chromatic Dispersion

In order to realize long distance transmission, some schemes for compensation of chromatic dispersion such as EDFA, chromatic dispersion fiber gratings, dispersion compensation fiber were proposed. In this paper, EDFA is used as the dispersion compensation device. The transmission line consists of SMF loop, which consists of a SMF length of 50 km with an EDFA gain of 10 dB and a SMF length of 10 km with an EDFA gain of 5 dB.

Figure 166.5 shows the BER performance after compensation of chromatic dispersion when laser power is 1 mw. From Fig. 166.5, it can be seen that the maximum transmission distance for NRZ and RZ is about 540 and 420 km at BER $= 10^{-9}$, respectively. For FEC system, the maximum transmission distance for NRZ and RZ is about 1,140 and 720 km, respectively.

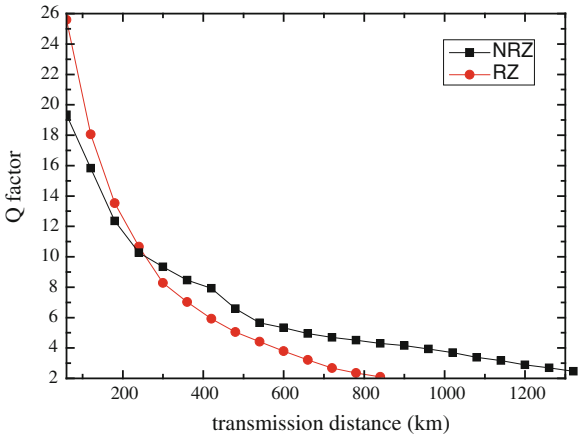


Fig. 166.6 Q factor after compensation of chromatic dispersion when laser power is 1 mw with different modulation formats

Q factor after compensation of chromatic dispersion when the laser power is 1 mw with different modulation formats is shown in Fig. 166.6. It can be seen that the Q factor of RZ modulation format is more than that of NRZ at transmission distance less than 360 km. When transmission distance is less than 360 km, the performance of RZ modulation scheme is better than that of NRZ. When transmission distance span is between 360 and 720 km, performance of NRZ modulation scheme is better than that of RZ. When transmission distance is more than 720 km, NRZ modulation scheme should be used.

166.4 Conclusion

Transmission performance of 100 Gbit/s signal with NRZ and RZ in SMF high-dispersion fiber is investigated. The BER against transmission distance, laser power and chromatic dispersion is simulated and analyzed. When the laser power is less than 80 mw, the maximum transmission distance of SMF with D = 16 ps/km/nm is less than 1.2 km for both RZ and NRZ system without compensation of chromatic dispersion. Comparison of NRZ and RZ transmission over SMF high-dispersion fiber clearly shows the advantage of NRZ modulation scheme for 100Gbit/s transmission over long SMF link.

Compensation of chromatic dispersion is also studied. After chromatic dispersion compensation, the performance of the system is improved. When the transmission distance is less than 360 km and the laser power is 1 mw, the performance of RZ modulation scheme is better than that of NRZ. When the transmission distance span is between 360 and 720 km and the laser power is 1 mw, the performance of NRZ modulation scheme is better than that of RZ. When the transmission distance is more than 720 km and the laser power is 1 mw, the NRZ modulation scheme should be used.

References

1. Tanizawa K, Kurosu T, Namiki S (2010) 1.8-ps RZ-pulse 43-Gbps transmissions over 126-km DSF with parametric tunable dispersion compensation. OFC/NFOEC, pp 1–3
2. Yu J (2010) 1.2 Tbit/s orthogonal PDM-RZ-QPSK DWDM signal transmission over 1040 km SMF-28. Electron Lett 46(11):775–777
3. Li Z, Cheng L, Yang Y, Lu C, Lau APT, Tam H, Wai PKA, Wang C, Xu X, Deng J, Xiong Q (2009) 1, 500-km SSMF transmission of mixed 40-Gb/s CS-RZ duobinary and 100-Gb/s CS-RZ DQPSK signals. Photonics Technol Lett 21(16):1148–1150
4. Nguyen Tan H, Matsuura M, Kishi N (2011) Parallel WDM signal processing in mixed NRZ and RZ transmission networks using a single optical gate with multiple switching windows. IEEE J Sel Top Quantum Electron 99:1–9
5. Liu X, Chandrasekhar S, Gnauck A and Wei X (2003) OSNR margin assessment for optical transmissions in the nonlinear regime with forward error correction. In: Proceedings OFC, Atlanta, GA, ThN2

Chapter 167
Wide-Area Backup Protection Algorithm Based on Multisource Information Fusion

Zhiqin He, Zhe Zhang, Xianggen Yin, Zhenxing Li and Xing Deng

Abstract Wide-area backup protection is a novel technique presented in recent years, which uses multisource information to make fault identification decision and availably prevents catastrophic blackouts. However, the property of algorithm will be deteriorated when information error exists. For the sake of improving the property of wide-area backup protection, this paper proposes an information fusion strategy using genetic algorithm to optimize the decision-making function. The operation information from local directional pilot relay and impedance relay is applied to establish the fitness function. Simulation tests based on IEEE 3-generator and 9-bus system validates the presented algorithm under as many as 3 bits information error.

Keywords Backup · Directional pilot relay · Genetic algorithm · Impedance relay · Maloperation · Wide-area protection

167.1 Introduction

Protective relaying is important for power grid to clear faults and sustain system stability. However, traditional backup protection is easy to maloperate under stressed conditions and induces catastrophic blackouts [1]. In order to

Z. He (✉) · Z. Zhang · X. Yin · Z. Li · X. Deng
State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Electrical Power Security and High Efficiency Lab, Huazhong University of Science and technology, Wuhan 430074, People's Republic of China
e-mail: 99211421@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_167,

overcome above problems, wide-area backup protection (WABP) is presented in recent years, which adopts multisource information from each substation within the power grid to make comprehensive decisions [2]. Various protection principles are presented to search fault elements correctly and avoid improper tripping [3, 4]. In order to overcome the difficulty about information deficiency and error, relative improvements are also introduced. Reference [5] presents a method for bad data identification using Dempster-Shafer evidence theory. It is helpful to filter electrical parameters from current or voltage transformers in substations. On the other hand, Ref. [6] utilizes expert system to optimize the decision-making procedure of WABP by fusing logical information. But expert system depends on a large amount of priori knowledge and is hard to realize on-line computation.

A new information fusion method based on genetic algorithm (GA) is proposed. The operation logic of wide-area directional pilot protection is adapted, and a new fitness function combining with the operation data from pilot relay and impedance relay is established. This information fusion mechanism has simple principle and is to be realized in practical grids.

This paper is organized as follows. Section 167.2 introduces the operation logic of wide-area directional pilot protection. The information fusion algorithm is demonstrated in Sect. 167.3. A detailed simulation tests using MATLAB 7 is presented in Sect. 167.4 for proving the effectiveness of this method. Finally, the paper ends with conclusions in Sect. 167.5.

167.2 Wide-Area Directional Pilot Protection Principle

Directional pilot relay is used widely in practical grid so as to realize faulted elements isolation quickly. For its excellent performance in identifying different types of faults, protection engineers absorb it into wide-area backup protections and make relevant improvements. The IEEE 3-generator and 9-bus system [7] is applied to briefly illustrate the principle of wide-area directional pilot protection.

As shown in Fig. 167.1, all the branches (L_1–L_6) within the dotted boundaries are protected by the WABP system. When a fault occurs on L_6 at F_1, the intelligent electronic devices (IEDs) at both end of L_6 will operate to clear this fault. An incidence matrix shown in Eq. 167.1 illustrates the relationship between buses (substations) and branches (IEDs). Thereafter, a weighted incidence matrix representing the operation results can be listed as Eq. 167.2.

The number '1' in W represents the forward direction; number '-1' represents the backward direction. Calculate the column vector L_1–L_6, respectively, then the fault element identification result can be represented by an array: $R = [0,\ 0,\ 0,\ 0,\ 0,\ 2]$. Thus, the actual faulty line L_6 can be found and necessary tripping decision will be sent to IED_3 and IED_4 finally. In short, the operation principle of wide-area

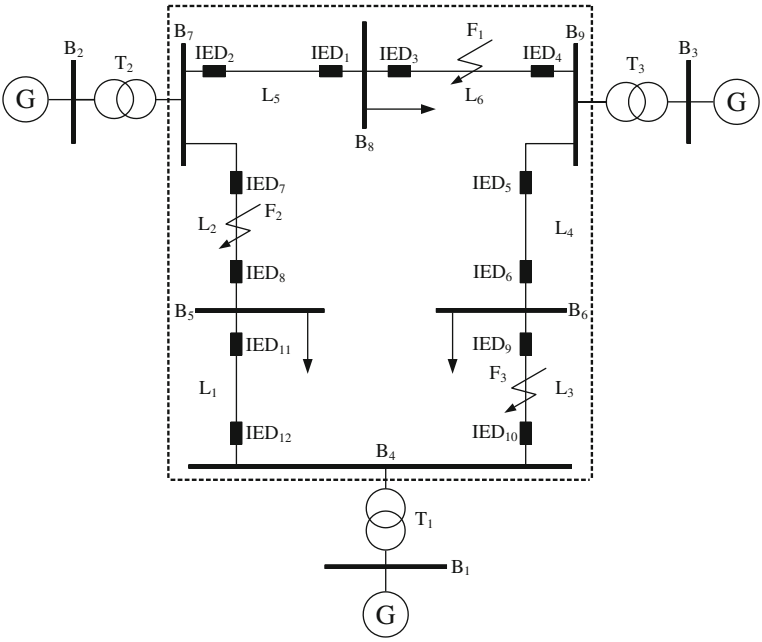


Fig. 167.1 IEEE 3-generator and 9-bus test system

directional pilot protection is simple to realize, however, it relies on sampled data with high reliability. If information loss or error occurs, the faulty devices will not be found exactly. Moreover, the WABP system may misoperate and the power grid may collapse.

$$I = \begin{array}{c} \\ B_4 \\ B_5 \\ B_6 \\ B_7 \\ B_8 \\ B_9 \end{array} \begin{array}{c} \begin{array}{cccccc} L_1 & L_2 & L_3 & L_4 & L_5 & L_6 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array} \tag{167.1}$$

$$W = \begin{array}{c} \\ B_4 \\ B_5 \\ B_6 \\ B_7 \\ B_8 \\ B_9 \end{array} \begin{array}{c} \begin{array}{cccccc} L_1 & L_2 & L_3 & L_4 & L_5 & L_6 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix} \end{array} \tag{167.2}$$

167.3 Multisource Information Fusion Using Genetic Algorithm

It is clear that Eq. (167.2) is vulnerable to subject to the scenario of information loss or error. Therefore, an information fusion scheme using GA is proposed. GA is an optimization algorithm and has been used widely in power system research, such as: reactive power optimization [8], load dispatch [9], etc. Consider the operation speed of backup protection is slower than primary protection because of communication delay, the operation information of primary protection including directional pilot relay, zone-1 and -2 of impedance relay is utilized comprehensively to realize multisource information fusion.

Let the forward fault be '1' and the backward fault be '0', the fault direction set of protected lines can be represented by a chromosome using binary encoding mode. For the IEEE 9-bus system, the chromosome contains 12 bits neglecting the impact of load: (Fig. 167.2).

The evolutionary population is constituted by a certain number of chromosomes. In general, the number of chromosome in one population is among 20–200 so that the diversity of population can be kept. In this method, most of the chromosomes are generated randomly. On the other hand, one of the chromosomes in population should be adopted by practical measured information. Assume that the number of initial chromosomes within the population is N, it could comprise two parts: $\mathrm{N}-1$ generated chromosomes (GC) and 1 measured chromosome (MC).

Afterwards, the fitness function is also established to evaluate the property of evolutionary solution:

$$\min F = \sum_{i=1}^{M} \left| P_i - P_i^*(D, Z1, Z2) \right| \tag{167.3}$$

where M represents the number of protection IEDs; P_i is the operation status of traditional line protection and corresponds to $IED_i\,(i = 1, \ldots, 12)$; D represents the operation results of directional pilot relay; $Z1$ and $Z2$ are the operation results of zone-1 and -2 of impedance relay, respectively. Consider the impedance relay is used commonly in practical high voltage (HV) grid, this paper combine it with directional pilot relay to realize information fusion.

The expectation function of P_i^*can be further demonstrated as follows:

$$\begin{cases} p_i^* = (Z1_i \oplus Z2_i) \otimes F_i \otimes F_{i+1} & \text{if i is an odd number} \\ p_i^* = (Z1_i \oplus Z2_i) \otimes F_{i-1} \otimes F_i & \text{if i is an even number} \end{cases} \tag{167.4}$$

where symbol '$\oplus$' represents logic OR; symbol '$\otimes$' represents logic AND.

Figure 167.3 shows the procedure of GA which contains population initialization and evolutionary operation of chromosome. The latter includes selection, crossover and mutation operation. The introduction of these operations is not

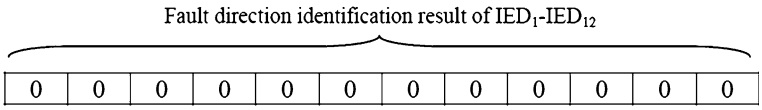


Fig. 167.2 Chromosome structure

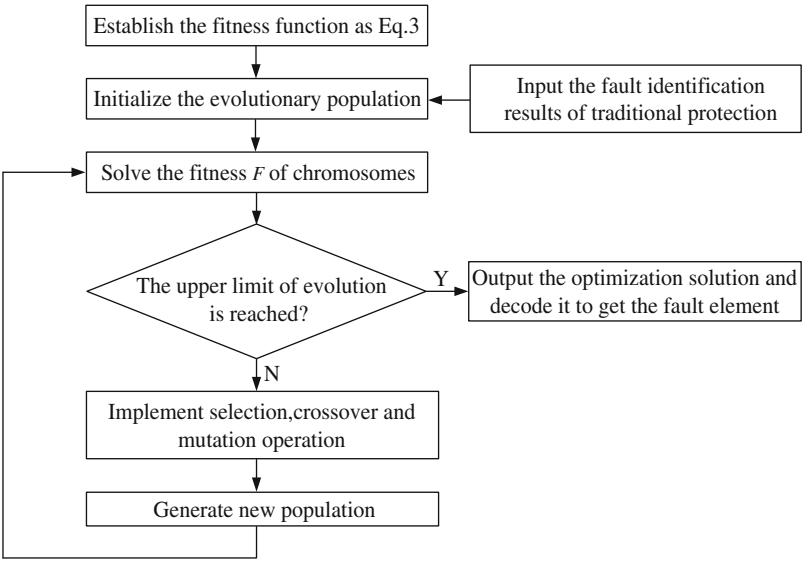


Fig. 167.3 Flow chart of information fusion

Table 167.1 Multisource information fusion results with 1 bit error

Fault position	Measured chromosome	Evolution result of GA	Decision result
F_1	000100101111000100 100000000100000100	011101100101	L_6
F_2	100000000100100000 111111000101000100	100110110101	L_2
F_3	010100000100100000 100000111111100000	010110101110	L_3

illustrated here due to space limitations. The evolutionary computation will not terminate until the condition of convergence is satisfied.

167.4 Simulation Test

The IEEE 3-generator and 9-bus system shown in Fig. 167.1 is provided to testify the property of GA, which is realized by MATLAB 7.0. There are 3 faults set in the test system: F1, F2 and F3 are located in the middle of L6, L2 and L3,

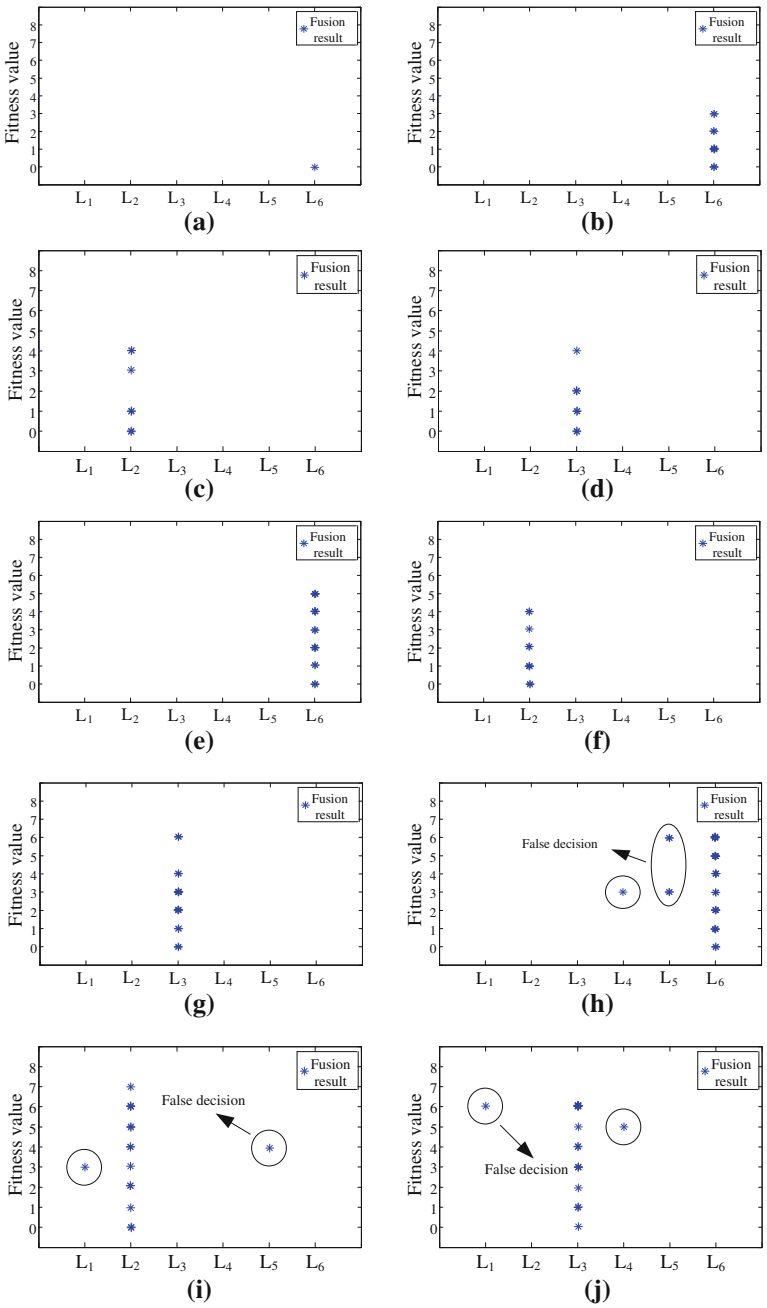


Fig. 167.4 Multisource information fusion results for other conditions. **a** F1 without false data. **b** F1 with 2 bits false data. **c** F2 with 2 bits false data. **d** F3 with 2 bits false data. **e** F1 with 3 bits false data. **f** F2 with 3 bits false data. **g** F3 with 3 bits false data. **h** F1 with 4 bits false data. **i** F2 with 4 bits false data. **j** F3 with 4 bits false data

respectively. Each IED should input 3 bits data about the operation result of relative directional pilot relay and impedance relay (zone-1, -2). In this paper, the scale of evolutionary population is set to 50, the championship method is applied to implement selection operation, the crossover and mutation probability are set to 0.8 and 0.03, respectively. The decision result is represented by binary code, and then the actual faulty line can be found. When the evolution number reaches 500, the iteration procedure will terminate. Then, the solution with the highest fitness value will be chosen as the final information fusion result.

Table 167.1 shows the fusion test results with 1 bit false data. The false data is marked by shadow. Simulation results clearly prove the correctness of this method.

Meanwhile, test results for 2–4 bits random error are shown in Fig. 167.4. Compared to Fig. 167.4a, the optimized fitness values of other conditions increase gradually when the degree of information error increases. When information error reaches 4 bits, the property of GA will degrade and false decision would be made. Therefore, the presented algorithm could make correct decision under at most 3 bits error. Yet, the probability of 4 relays making incorrect decisions at the same time is very low in practical power grid, and the excellent performance of proposed algorithm can be validated ultimately.

167.5 Conclusion

This paper proposes a novel wide-area backup protection algorithm based on multisource information fusion. GA is applied to realize fault elements identification and misoperation correction. The operation information comprising directional pilot relay and impedance relay from each IED at the ends of protected transmission lines is collected to construct chromosome in the evolutionary population. Simulation tests based on IEEE 3-generator and 9-bus system validates the presented algorithm under less than 4 bits data error. Further developments will focus on improving the content of fused information and advancing the efficiency of evolution.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (50837002, 50877031), National High Technology Research and Development of China (863 Programme) (2009AA05Z208) and the Specialized Research Fund for the Doctoral Program of Higher Education (20090142110055).

References

1. Damir N, George B, Gene H (2010) IEEE PSRC report on performance of relaying during wide-area stressed conditions. IEEE Trans Power Delivery 25(1):3–16
2. Phadke A, Thorp J (2008) Synchronized phasor measurements and their applications. Springer, New York, pp 197–221
3. Lin X, Li Z, Wu K (2009) Principles and implementations of hierarchical region defensive systems of power grid. IEEE Trans Power Delivery 24(1):30–37

4. Eissa M, Masoud M, Elanwar M (2010) A novel back up wide area protection technique for power transmission grids using phasor measurement unit. IEEE Trans Power Delivery 25(1):270–278
5. Zhu L, Shi D, Duan X (2011) Evidence theory-based fake measurement identification and fault-tolerant protection in digital substations. IET Gener Transm Distrib 5(1):119–126
6. Tan J, Crossley P, Kirschen D (2000) An expert system for the back-up protection of a transmission network. IEEE Trans Power Delivery 15(2):508–514
7. Church C, Morsi W, Hawary M (2011) Voltage collapse detection using ant colony optimization for smart grid. Electr Power Syst Res 81:1723–1730
8. Masoum M, Ladjevardi M, Jafarian A (2004) Optimal placement, replacement and sizing of capacitor banks in distorted distribution networks by genetic algorithms. IEEE Trans Power Delivery 19(4):1794–1801
9. Park J, Lee K, Shin J (2005) A particle swarm optimization for economic dispatch with non-smooth cost functions. IEEE Trans Power Syst 20(1):34–42

Chapter 168
P2P-Based Self-Organized Streaming Media CDN Content Delivery Strategy

Qi-xing Xu and Zi-ao Zhan

Abstract In order to cope with the shortcomings of the existing streaming media CDN, and accomplish large-scale deployment streaming media delivery system, in this chapter, we introduce peer to peer(P2P) idea into streaming media CDN. Then we put forward a framework of PBSMCDN(P2P-Based Self-organized Streaming Media Content Delivery Network)and study streaming media content delivery strategies, and present a Random push algorithm based on partial access feature of steaming media. The experimental results show that the algorithm is effective.

Keywords P2P · CDN · Streaming media · Architecture · Delivery strategy

168.1 Introduction

The main reason for large-scale deploying streaming media delivery service difficulty on Internet is that the resource of servers such as output bandwidth and CPU computing power is too restricted to bear plentiful parallel media flow.

Although some technologies, such as streaming media caching technology [1, 2], streaming media CDN(SMCDN)technology [3], and P2P streaming media technology [4] have partly solved the problem of overspending network resources when streaming media are distributed, but they are still faulty. Therefore, large-scale application of streaming media delivery hasn't been universalized on Internet so far.

Q. Xu (✉) · Z. Zhan
Henan Institute of Engineering, Zhengzhou 451191, China
e-mail: xuqix@sina.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_168,

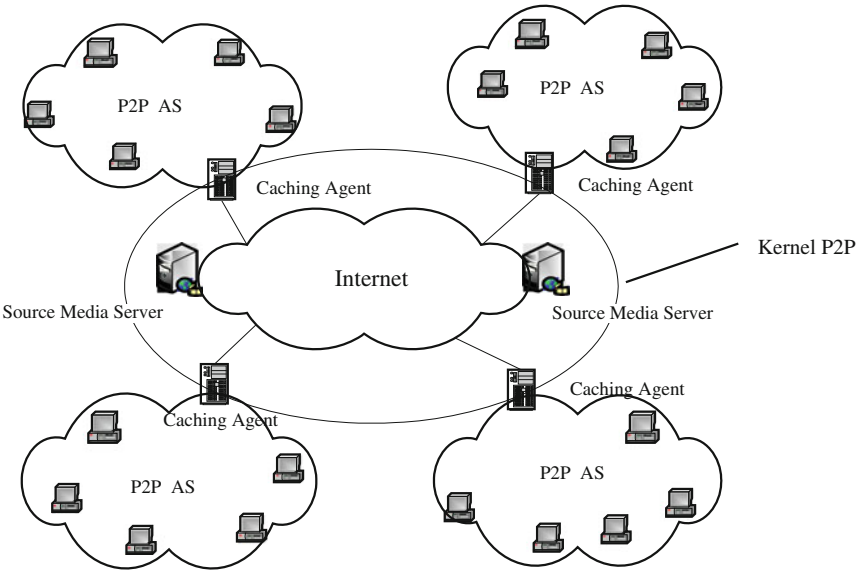


Fig. 168.1 The framework of PBSMCDN

The hybrid CDN + P2P mode of streaming media on demand [5] is by now the most large-scale and expansible delivery mode of the streaming media service. However, its Kernel CDN is also based on the existing streaming media CDN. As a result, it is still difficult to obviate the inherent shortcomings of the existing streaming media CDN.

In order to cope with the shortcomings of the existing streaming media CDN system, and accomplish large-scale deploying streaming media delivery system, in this paper, we introduce P2P idea into streaming media CDN. Then, we put forward a framework of PBSMCDN, and study streaming media content delivery strategies.

168.2 The Framework of PBSMCDN

168.2.1 The Framework of PBSMCDN

In recent years, network application based on P2P is becoming a popular research topic and it provides a new way to deploy and accomplish network service.

Utilizing the constructing technology of P2P networks and taking CDN + P2P mode [5] as example, an architecture of two-level PBSMCDN is proposed as shown in Fig. 168.1.

The basic ideas are as follows:

Primarily, lots of streaming media caching servers in Wide Area Networks should be constituted to the first-level P2P networks, by which the original content at the streaming media servers may be quickly distributed to every caching proxy collaborating with each another to enhance the whole performance of the system. The first-level P2P networks is the core of system, so it is also called kernel P2P.

Secondly, P2P AS of streaming media is brought into the edge networks. The AS is a admixing P2P streaming media system made up with one or more caching proxy servers and their clients, called Second-level P2P Networks or edge P2P. The purpose of bringing in P2P AS is to reduce caching proxy servers and original media server bandwidth consumption.

168.2.2 Three Layers Architecture of PBSMCDN

There are three types of nodes in PBSMCDN, named source media server, caching proxy server and client respectively.

A source media server is a streaming media server or a group of streaming media servers, and its main function is responding to the access requests for streaming media content from caching proxy server and client to provide the content. The caching proxy server, as a proxy for media transmit, serves the end users by receiving median content from the source server or the proxy server.

A lot of caching proxy servers ranged on Internet are organized together by structured P2P network technologies to form the Kernel P2P network. They deliver streaming media content to caching proxy servers and coordinate them and sequentially achieve sharing the cache balancing load. Their main function is to deliver streaming media content from source media servers to the edge of networks.

The clients belonging to caching proxy servers are organized Edge P2P Streaming Media ASs by P2P technologies, which mainly function as extending the service power of caching proxy servers and enhancing system concurrency.

Thus, the PBSMCDN consists of three layers: source media server layer, kernel P2P network layer and edge P2P AS layer, as shown in Fig. 168.2.

168.3 Random Push Strategy Based on the Partial Access Feature of Steaming Media

168.3.1 Content Delivery Methods in Steaming Media CDN

Currently, there are two main methods of steaming media content delivery from the source media server to the caching agent server. One is called “pull” way. In this way, the new content will be sent to the caching agent server from the source media server only when the client requests and a copy will be cached.

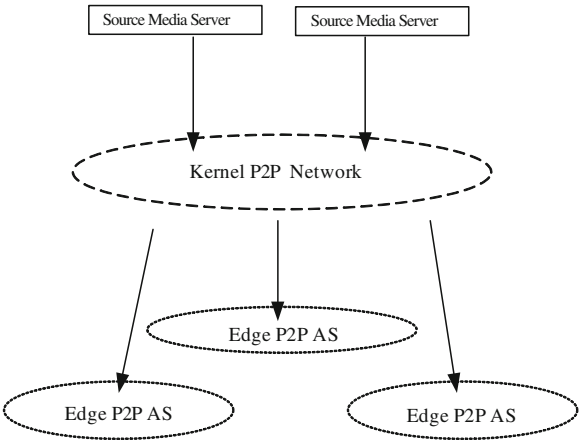


Fig. 168.2 Three layers architecture of PBSMCDN

Another is called "push" way. In this way, the new streaming media content is sent actively to each caching agent server close to the customer by the streaming media CDN content delivery system.

In the "pull" approach, the content will be distributed only when needed by customers, so the cache server's cache space utilization is high, but for the first client approaching the new content, a longer startup delay is needed; in the "push" approach, the source media server decide the delivery destiny simply base on its known information. Often, customers will send the unneeded contents back to the caching agent server, resulting in a waste of network resources and caching agent server stored resources. However, new access of customers will have a shorter startup time, as a result of getting direct contents from the caching agent server.

In self-organized streaming media CDN based on P2P, the kernel P2P network is a streaming media CDN with caching proxy server as node, so the content delivery method is mainly "pull", but in order to overcome the limitations of the "pull" mode, in this paper, we proposed to introduce the "push" method into the system, and a random push strategy based on the partial access feature of steaming media using a combination of "push" and "pull".

168.3.2 The Feature of Partial Access of Steaming Media and the Segmentation Method

Studies have shown that, in the same class within the media file, the popularity of different parts (access probability) is often not the same [6]. Reference [6] Studies have shown that different users on different parts of the object with the same media access show bimodal characteristics, only about 55% of users can play a stream object completely, the majority of the rest users can only play 5% of the streaming media objects before it stopped.

Due to the huge volume of streaming media files and the great popularity differences among separate parts of one steaming media file, in order to improve the cache efficiency, streaming media caching systems generally use segment caching strategy, i.e. one streaming media file will be divided into multiple segments and take these segments as the unit of streaming media data for caching. Theoretically, the shorter each steaming media object segment is, the smaller the possibility is for parts with larger popularity differences be assigned in the same block. Therefore, it is more likely to improve the cache space utilization and optimize the proxy cache performance. However, the smaller the segment is, the more the amount of segment will be, resulting in complex cache management and more additional expenditure.

In order to solve this problem, according to the basic characteristics of the partial access to streaming media files, namely, the access frequency of former part of the file decreases rapidly while that of the latter part tends to decrease gently, we propose a segmentation strategy in which the segment length gradually increased, i.e. the segmentation is shorter in parts whose popularity vary greatly, and vice versa. The specific segmentation methods are as follows.

$$l_i = \begin{cases} \left(\frac{1}{\beta}\right)^{i-1} \times \delta \; i < m \\ \left(\frac{1}{\beta}\right)^{m-1} \times \delta \; i < m \end{cases} \tag{168.1}$$

Parameters: i is the serial number of segment, δ is the length of first segment, $\beta < 1$ is increasing factor of segment, $m > 1$ is the increasing the length of the segment.

168.3.3 Research of the Active "Push" of Steaming Media Contents

Generally, Streaming media caching proxy server uses the "pull" method to obtain contents, but when the source media server adds new contents, certain impact will be bring about to the source media server load, since the popularity of these new programs are generally higher, and all agents have yet to cache the content. To solve this problem, the method is to introduce active "push" technology, i.e. send these new contents to interested caching proxy servers before they are demanded. By doing this, not only the load fluctuations of source media servers can be effectively solved, but the playing quality of initial customers can also be improved.

For a source media streaming media server that has been introduced to a "push" technology, when a new group streaming media file is added, it needs to push these contents to the caching proxy servers that demand them. Therefore, it must be aware of these files' push destinations. According to the feature of streaming media's partial access, each streaming media file is segment based on

the formula (1), and the source media streaming media server must know which few caching proxy servers each stream segment should be pushed to. Therefore, the first problem of streaming media content delivery is the destination choice.

168.3.3.1 The Choice of Push-Purposed Proxy Server

There are two cases to determine the destination of the push: one is scheduled in advance, the other is not, or the server does not support it. For the case of pre-scheduled, the pre-caching proxy server which the customer belongs to can be directly uses as the push destination. So we mainly aimed at studying the situation that is no scheduled.

As customers of caching proxy servers that did not schedule these new contents might also be interested in them, these caching proxy servers can also serve as 'push' targets. But this does not go for every caching proxy server, since there is great possibility that all or a majority of customers of one caching proxy server have completely no interest in some contents of one source media server. To solve this problem, this paper takes the level of interest for new content of the customer base of caching proxy server as basis for streaming content "push" destinations.

However, it is difficult to get the customers' interested in those new streaming media file. Therefore, statistics of historical information should be added to the source media server in order to predict customer interest.

Definition 1 The caching proxy server i belongs to the customer base for certain types of streaming media files degree of interest α_i from formula (168.2) gives.

$$\alpha_i = \frac{D_i}{D} \tag{168.2}$$

Parameters: D is the total amount of certain types of streaming media files, D_i is caching proxy server i request amount of the file.

When α_i is greater than a threshold θ, the caching proxy server i is the "push" potential destination of certain types of streaming media files. The number of potential destinations such file recorded as Nc.

168.3.3.2 Random Push Algorithm Based on the Partial Access Feature of Steaming Media Files

Once the potential destinations are determined, generally it is not able to push all the contents to potential destinations, as different files and different segments of the file have different popularities. If too many copies are pushed to segments with lower popularities, the cache space utilization of caching proxy server will be reduced, thus affecting the performance of the system. Therefore, when choosing

destination for each segment, the number of its destination should be decided by their popularity, i.e. the pushing number of copies.

Suppose a new streaming media file is divided into K segments, one to access the streaming media file requests access to the probability of the *i*-th segment is P_i, then the mean number of copies of M needed to "push" for the *i*-th segment:

$$M = N_c \times P_i$$

For a new streaming media file, Pi is difficult to estimate, we use formula (3) to estimate the value.

$$p_i = \begin{cases} \gamma\beta^{i-1} & i < m \\ \gamma\beta^{m-1} & i \geq m \end{cases} \tag{168.3}$$

Parameters: $0 < \gamma \leq 1$ Based on the expected popularity of streaming media files to determine, β and m are the parameters of streaming media file's segment.

Identify the mean number of copies of M needed to "push", and then consider which proxy cache server it should be pushed to. A simple selection method is, basing on the agency's interest rankings in the streaming media files, simply choose the first M streaming media proxy cache server from Nc to push. This method is simple, but the selection method is too absolute, and it is not in accordance with the random access of customers when the system runs, for a caching agent server with smaller degree of interest may also be qualified to act as a destination, and there is possibility that their customers may have a high demand on the file, though the possibility might be small. Therefore, absolutely not choosing these agents can have an impact on the integrity, so this paper gives a random selection algorithm according to probability.

A is set for the selected caching proxy server, M is the number of needed to "push", the proxy i is selected by probability P_i, then M can be calculated by algorithm 1.

algorithm 1:

For i = 1 to N_c
the proxy i is selected by P_i;
If you select will be added to A;

END

P_i is calculated as follows:

$$P_i = \frac{\alpha_i}{\sum_{j=1}^{N_c} \alpha_j}$$

Parameters: α_1 , α_2 , ..., α_{Nc} is the degree of interest of caching proxy server to certain type streaming media file.

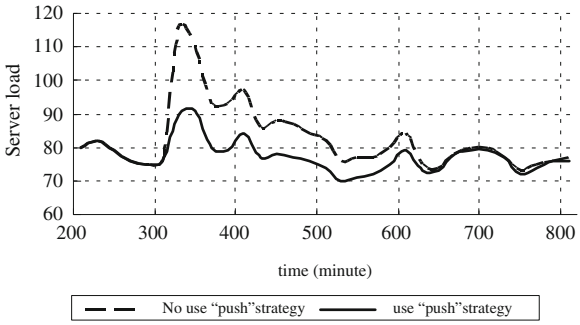


Fig. 168.3 Compare of the server load using different distributing method

168.3.3.3 Simulating Experiment

In order to verify the effectiveness of the push strategy and related algorithms, we use NS2 simulation environment. Assume that the system has a streaming media server, which caches 300 60 min streaming media files, and 10 caching proxy server using LRU cache replacement algorithm, each with a cache space which can accommodate 3,600 min streaming data.

The popularity of streaming media files obeys the Zipf's delivery of parameter $\theta = 0.271$; assume that the interval of customer request's arrival time obeys exponentially delivery with an arrival rate 3/min. For simplicity, each file has a equal length of 2 min, that is to take $\delta = 2$, $\alpha = 1$. To generate the access frequency, use formula (168.3), take γ as 1, and select β randomly among 0–0.5 for different files.

In the experiment, we assume that each streaming media file coding is CBR, ignores the actual network factors (bandwidth, congestion, connection topology, etc.), and not taking the impact of streaming media VCR operation into account. Timing starts when the system turns into a stable state (all the proxy cache spaces of proxy caches are covered). When the system has been running for 300 min, add 20 60 min streaming media files to the server. Assume that the popularity of these documents were all higher than the source media streaming files. Figure 168.3 shows comparison of changes in the source media server load, from which we can see the effectiveness of the pushing strategy that this paper proposed.

168.4 Summary

Streaming media content on-demand is an important growing point of Internet application development, but the supports for it of the existing systems have their flaws, leaving streaming media content on-demand applications far from being large-scale deployed and applied. In order to support open and autonomous on-demand streaming services' large-scale deployment on the Internet, in this paper,

we first utilize the efficient routing mechanism in P2P and its inherent self-organizing feature, proposed the self-organization streaming media CDN architecture based on P2P. Then, aiming on the content delivery from the source media server to the proxy server, random-content "push" algorithm based on the partially access feature of streaming media files is presented. The effectiveness of the algorithm has been proved by experiments.

References

1. Paknikar S, Kankanhalli M, Ramakrishman KR et al (2000) A caching and streaming framework for multimedia[J]. In: Proceedings of ACM multimedia. Los Angeles, pp 13–20
2. Tran DA, HUA KA, Sheu S (2003) A new caching architecture for efficient video services on the internet[J]. In: Proceedings of IEEE symposium on applications and the internet. Orlando, pp 172–181
3. Zhi-guo L (2004) The study of streaming media content distributing networks[D]. Computing Technology Research Institution, Science Academy of China, Beijing
4. Xiang Z, Zhang Q, Zhu W, Zhang Z, Zhang Y (2004) Peer-to-Peer based muhimedia delivery service[J]. IEEE Trans Multimedia 6(4):343–355
5. YI C, SUN JZ (2005) Mohamed H et al An analytical study of peer-to-peer media streaming systems [J]. ACM Trans Multimedia Comput, Commun App (4):354–376
6. Acharya S, Smith B, Parnes P (2010) Characterizing user access to videos on the world wide web. Multimedia Comput Networking

Chapter 169
Improved Sparse Multipath Channel Estimation via Modified Orthogonal Matching Pursuit

Jing Lu and Rui Wang

Abstract Impulse response of sparse multipath channel (SMPC) can be recovered from a short training sequence since most entries of SMPC are zeros. Though the ordinary orthogonal matching pursuit (OMP) algorithm provides a very fast implementation of SMPC estimation, it suffers from the coherence of atoms in dictionary, especially in the case of SMPC with a large delay spread and short training sequence. In this chapter, a modified OMP method is proposed and a sensing dictionary is designed adaptively to improve the performance of the OMP algorithm. Numeral experiments illustrate that the proposed algorithm based on adaptive sensing dictionary outperforms the ordinary OMP algorithm.

Keywords Terms—sparse multipath channel · Generalized orthogonal matching pursuit · Sensing dictionary · Sparse approximation

169.1 Introduction

The problem of channel estimation has been studied extensively and some methods have been proposed in the literature. Conventional methods for channel estimation were based on least-square (LS) algorithm. Unfortunately, all entries of the solution obtained by these methods were nonzeros and it was wrong in the case of sparse multipath channel (SMPC). SMPC is frequently encountered in wireless communication applications and has only a small portion of entries significantly

J. Lu (✉) · R. Wang
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: lujing@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_169,

different from zero. Taking advantage of the sparsity, impulse response of SMPC can be recovered from relatively small number of received data and training data. However, finding the sparsest solution is an NP-Hard combinatorial problem and massive works have been down to develop suboptimal methods for this problem.

Relax methods [1] and greedy algorithms [2, 3] are the most popular methods for finding the sparse solution. In particular, greedy algorithms, such as matching pursuit (MP) [2] and orthogonal matching pursuit (OMP) [3] can provide a very fast implementation of sparse approximation [4]. Some methods for sparse channel estimation have been proposed based on MP [5, 6]. However, according to the sufficient condition developed by Tropp [7], both MP and OMP suffer from highly coherence of redundant dictionary, especially in the case of SMPC with either large time delay spread or relatively small number of training data and received data. Recently, a modified OMP algorithm was developed to improve the performance of the ordinary OMP algorithm in the case of highly coherent dictionary by introducing a sensing dictionary [8]. However, this algorithm only considered the noiseless situation and the sensing dictionary is non-adaptively designed, which is independent of the received data. In this chapter, a novel OMP algorithm is proposed to improve the performance of the ordinary OMP algorithm. An adaptively designed sensing dictionary is constructed and posterior information is utilized efficiently to prevent false atoms from being selected due to highly coherence between atoms in the ordinary dictionary. Numeral experiments illustrate that the performance of the proposed algorithm based on adaptive sensing dictionary is much better than that of the ordinary OMP algorithm.

The remaining sections are organized as follows. In Sect. 169.2, the sparse multipath channel model is presented and the coherence between atoms is formulated. The adaptive approach to design sensing dictionary is given in Sect. 169.3. Finally, we compare the performance of the proposed algorithm with other algorithms via simulations over wireless Gauss channel in Sect. 169.4 and conclusions are given in Sect. 169.5.

169.2 Problem Formulation

Let us transmit the training sequences $s(n)$, $n = 0, 1, \ldots, N-1$, through a stationary multipath sparse channel. The training sequence symbols $s(n)$ for $n < 0$ can be obtained from the previous estimates or for the first arriving frame they are assumed to be zero [5]. The received base-band signal samples can be modeled as

$$r_t = \sum_{i=0}^{L-1} s(t-i)h_i + e_t, \tag{169.1}$$

where $t = 0, 1, \ldots, N-1$, h_i is the channel impulse response with length L and e_t is additive white Gaussian noise with zero mean and variance σ_e^2. Denote the

power of training sequence and the received signal by σ_s^2 and σ_r^2, respectively. In the vector form, we have

$$\mathbf{r} = \mathbf{S}\mathbf{h} + \mathbf{e}, \tag{169.2}$$

where $\mathbf{h} = [\, h_0\, h_1 ! \cdots h_{L-1}\,]^T$, $\mathbf{r} = [\, r_0\, r_1\, \cdots\, r_{N-1}\,]^T$, $\mathbf{e} = [\, e_0\, e_1\, \cdots\, e_{N-1}\,]^T$, T denotes transposition and $\mathbf{S}$ is the known training matrix given by

$$\begin{aligned}\mathbf{S} &= \begin{bmatrix} s(0) & s(-1) & \cdots & s(-L+1) \\ s(1) & s(0) & \cdots & s(1) \\ \vdots & \vdots & \ddots & \vdots \\ s(N-1) & s(N-2) & \cdots & s(N-L) \end{bmatrix} \\ &= [\, \mathbf{s}_0 \quad \mathbf{s}_1 \quad \cdots \quad \mathbf{s}_{L-1}\,]\end{aligned}$$

Denote the number of nonzero entries of $\mathbf{h}$ as K. The channel $\mathbf{h}$ is sparse if $K \ll L$ holds. In the context of sparse analysis, $\mathbf{S}$ is called as dictionary and the column vector $\mathbf{s}_i\ (i = 0, 1, \ldots, L-1)$ as atom. As a result of short training sequence, which improves throughput efficiency for the systems where transmitted packet length is short, the dictionary is highly redundant. In other word, the dimension of the received base-band signal vector $\mathbf{r}$ is much smaller than the number of atoms in the dictionary, i.e., $N \ll L$.

The ordinary OMP algorithm iteratively selects the atom that correlates most strongly with the residual signal. At each step k, the best atom $\mathbf{s}_{m_k}$ is selected by solving the simple optimization

$$m_k = \arg \max_{0 \le i \le L-1} \hat{h}_i^{(k)} \tag{169.4}$$

where

$$\hat{\mathbf{h}}^{(k)} = \left[\, \hat{h}_0^{(k)} \quad \hat{h}_1^{(k)} \quad \cdots \quad \hat{h}_{L-1}^{(k)} \,\right]^T = \left|\mathbf{S}^T \mathbf{g}_{k-1}\right|, \tag{169.5}$$

where $k = 1, 2, \ldots, K$. We have $\mathbf{g}_0 = \mathbf{r}$ for initialization and $\mathbf{g}_k = \mathbf{P}_k \mathbf{r}$ for $k = 1, 2, \ldots, K-1$, where

$$\mathbf{P}_k = \mathbf{I}_M - \hat{\mathbf{A}}^{(k)} \left(\hat{\mathbf{A}}^{(k)T} \hat{\mathbf{A}}^{(k)} \right)^{-1} \hat{\mathbf{A}}^{(k)T} \tag{169.6}$$

$$\hat{\mathbf{A}}^{(k)} = [\, \mathbf{s}_{m_1} \quad \mathbf{s}_{m_2} \quad \cdots \quad \mathbf{s}_{m_k}\,] \tag{169.7}$$

and $\mathbf{I}_M$ is an identity matrix.

To illustrate the effect of coherence between atoms, e.g., at the initialization step, we express the sparse channel estimation as

$$\hat{\mathbf{h}}^{(1)} = \left|\mathbf{S}^T \mathbf{g}_0\right| = \left|\mathbf{S}^T \mathbf{r}\right| = \left|\mathbf{S}^T (\mathbf{S}\mathbf{h} + \mathbf{e})\right|, \tag{169.8}$$

or

$$\hat{h}_i^{(1)} = \left| \sum_{l=0}^{L-1} \mathbf{s}_i^T \mathbf{s}_l h_l + \mathbf{s}_i^T \mathbf{e} \right|. \tag{169.9}$$

for $i = 0, 1, \ldots, L-1$. If $h_l = 0$, the coherence $\mathbf{s}_i^T \mathbf{s}_l$ can not affect the estimated value of $\hat{h}_i^{(1)}$. However, if $h_l \neq 0$, $\mathbf{s}_i^T \mathbf{s}_l$ will draw the estimated value of $\hat{h}_i^{(1)}$ away from its correct value h_i. As a result, we may either choose a false atom when $h_i = 0$ or omit a correct atom when $h_i \neq 0$ at this step if the coherence is large enough. Here, the problem is how to mitigate the effect of the coherence on the performance of OMP.

169.3 The Proposed Algorithm

In order to identify the correct atoms in the case of high coherence, we resort to the modified OMP based on a sensing dictionary $\mathbf{W}$, and use $\hat{\mathbf{h}}^{(k)} = \left| \mathbf{W}^T \mathbf{g}_{k-1} \right|$ rather than $\hat{\mathbf{h}}^{(k)} = \left| \mathbf{S}^T \mathbf{g}_{k-1} \right|$ in (169.5). Obviously, the ordinary OMP is a special case of the general OMP with $\mathbf{W} = \mathbf{S}$. Given the received signal $\mathbf{r}$, the probability of appearance in the reconstruction of $\mathbf{r}$ is different for different atom [9]. Therefore, the adaptive sensing vector is taken as the solution to the following optimization

$$\min_{\mathbf{w}_i} \mathbf{w}_i^T \mathbf{S} \mathbf{U}^{(k)} \mathbf{S}^T \mathbf{w}_i \tag{169.10}$$

$$s.t. \;\; \mathbf{s}_i^H \mathbf{w}_i = 1 \tag{169.11}$$

where $i = 0, 1, \ldots, L-1$, $\mathbf{U}^{(k)} = \mathrm{diag}\left(\left| \hat{\mathbf{h}}^{(k)} \right|^{\rho} \right)$, $\hat{\mathbf{h}}^{(k)} = \left| (\mathbf{W}^{(k)})^T \mathbf{r} \right|$ and $\rho > 0$. Similarly, the closed-form solution can be given by

$$\mathbf{w}_i = \mathbf{D}_i \mathbf{s}_i \tag{169.12}$$

where

$$\mathbf{D}_i = \frac{1}{\mathbf{s}_i^T \left(\mathbf{S} \mathbf{U}^{(k)} \mathbf{S}^T + \beta \mathrm{I}_M \right)^{-1} \mathbf{s}_i} \left(\mathbf{S} \mathbf{U}^{(k)} \mathbf{S}^T + \beta \mathbf{I}_M \right)^{-1} \tag{169.13}$$

for $i = 0, 1, \ldots, L-1$, and β is a positive regularization parameter. We here take the correlation between the received vector (or the residual vector) and each atom in the dictionary as an approximate measure of this probability. Because $\mathbf{U}^{(k)}$ is calculated from the sensing dictionary itself, we must set an initial sensing dictionary such as $\mathbf{W} = \mathbf{S}$.

The sensing dictionary given by (169.13) is the adaptive function as a result of the adaptive minimum coherence optimization with the distortionless response constraint. It is easy to see that the sensing dictionary given by the non-adaptive design method [8], which is completely determined by the dictionary $\mathbf{S}$ and independent of the received signal, corresponds to a special case of (169.13) with $\mathbf{U}^{(k)} = \mathbf{I}_L$ ($\mathbf{I}_L$ is an identity matrix) at each step.

Accordingly, the modified OMP algorithm based on adaptive sensing dictionary can be summarized as follows.

1. Initialization: $\mathbf{g}_0 = \mathbf{r}$, $\mathbf{W} = \mathbf{S}$, $k = 1$, $\beta > 0$ and $\rho > 0$;
2. For $i = 0, 1, \ldots, L-1$, repeat the following process for J times:

$$\hat{\mathbf{h}}^{(k)} = \left|\mathbf{W}^T \mathbf{g}_{k-1}\right|,\ \mathbf{U}^{(k)} = \mathrm{diag}\left(\left|\hat{\mathbf{h}}^{(k)}\right|^{\rho}\right)$$

$$\mathbf{D}_i = \frac{1}{\mathbf{s}_i^T \left(\mathbf{S}\mathbf{U}^{(k)}\mathbf{S}^T + \beta \mathbf{I}_M\right)^{-1} \mathbf{s}_i} \left(\mathbf{S}\mathbf{U}^{(k)}\mathbf{S}^T + \beta \mathbf{I}_M\right)^{-1},$$

$$\mathbf{w}_i = \mathbf{D}_i \mathbf{s}_i,\ \mathbf{W} = \left[\mathbf{w}_0 \quad \mathbf{w}_1 \quad \cdots \quad \mathbf{w}_{L-1}\right];$$

3. $\hat{\mathbf{h}}^{(k)} = \left[\hat{h}_0^{(k)} \quad \hat{h}_1^{(k)} \quad \cdots \quad \hat{h}_{L-1}^{(k)}\right]^T = \left|\mathbf{W}^T \mathbf{g}_{k-1}\right|$
$m_k = \arg \max_{0 \le i \le L-1} \hat{h}_i^{(k)}$,

$$\hat{\mathbf{A}}^{(k)} = \left[\mathbf{s}_{m_1} \quad \mathbf{s}_{m_2} \quad \cdots \quad \mathbf{s}_{m_k}\right],$$

$$\mathbf{P}_k = \mathbf{I}_M - \hat{\mathbf{A}}^{(k)} \left(\hat{\mathbf{A}}^{(k)T} \hat{\mathbf{A}}^{(k)}\right)^{-1} \hat{\mathbf{A}}^{(k)T},\ \mathbf{g}_{k+1} = \mathbf{P}_k \mathbf{r}$$

4. $k = k + 1$, go to (169.2) and repeat until $k = K$.

Finally, the position of the nonzero entries of SMPC is detected by $\left[m_1 \quad m_2 \quad \cdots \quad m_K\right]$, and the corresponding nonzero values are estimated as $\left(\hat{\mathbf{A}}^{(k)T} \hat{\mathbf{A}}^{(k)}\right)^{-1} \hat{\mathbf{A}}^{(k)T} \mathbf{r}$. To reduce the computation cost, the sensing dictionary can be calculated only for $k = 1$ and used at the subsequent steps.

169.4 Simulation Results

To gain some insights into the performance of the proposed algorithm, we carry out some experiments of SMPC estimation. The nonzero entries of SMPC are drawn randomly from a uniform distribution on $[-1, -0.2] \cup [0.2, 1]$ and the number of nonzero entries is K = 5. The position of nonzero entry of $\mathbf{h}$ is generated randomly. The channel length is set as L = 50 or 100, the length of training

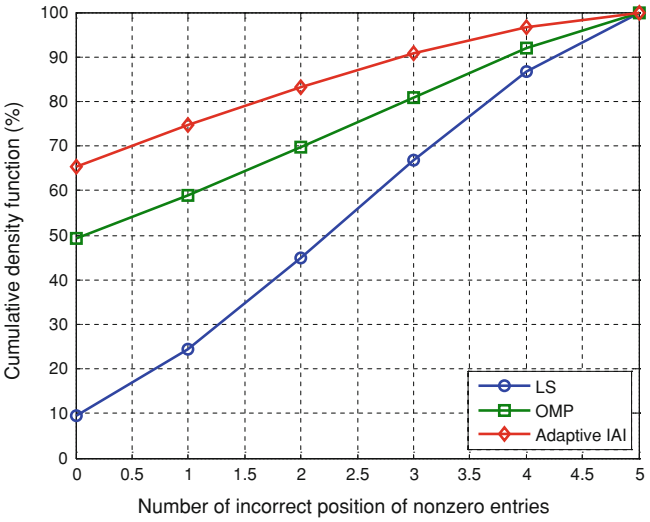


Fig. 169.1 Cumulative density function of the number of incorrect position of nonzero entries with N = 30, L = 50

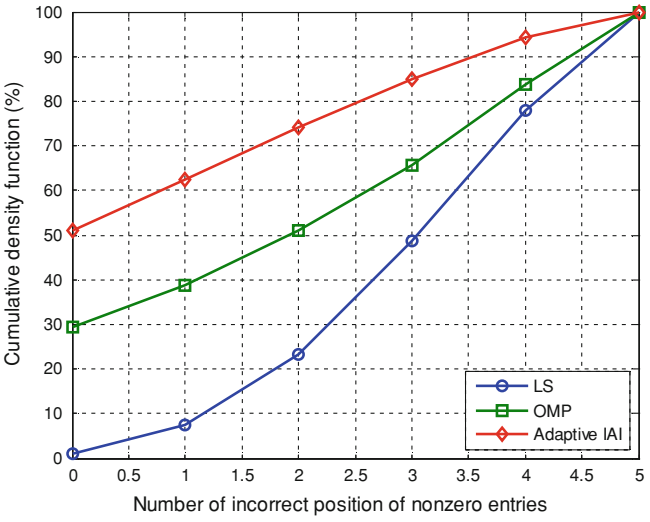


Fig. 169.2 Cumulative density function of the number of incorrect position of nonzero entries with N = 30, L = 100

sequence is N = 30, and the signal to noise ratio (SNR) is 10 dB. The other involved parameters used in the algorithms are set to $\rho = 3$, $\beta = 0.1$ and $J = 10$, which may be further optimized to obtain better performance. Simulation results are obtained over 10,000 independent Monte Carlo trials.

We compare performance of the modified OMP algorithm based on adaptive sensing dictionary with that of the least squares method (LS) and the ordinary OMP algorithm. We compare the ability of these algorithms to detect the nonzero entries of SMPC. The cumulative density functions (CDF) of the number of incorrectly detected nonzero components for the channel length values of 50 and 100 are shown in Figs. 169.1 and 169.2, respectively. From these CDF functions, we see that the proposed algorithm based on adaptive sensing dictionary gives more accurate detection of nonzero entries of SMPC than other algorithms. These results show that the modified OMP algorithm based on adaptive sensing dictionary significantly outperforms the other methods, especially in the case of more redundant dictionary of the latter ($L = 100$).

169.5 Conclusions

In this chapter, we propose a modified OMP algorithm by introducing an adaptive sensing dictionary. This algorithm significantly improves the performance of sparse multipath channels estimation especially in the case of sparser and longer SMPC. The numeral experiments indicate that the proposed algorithm outperforms both the ordinary OMP algorithm and the modified OMP algorithm based on non-adaptive sensing dictionary.

References

1. Chen SS, Donoho DL, Saunders MA (1999) Atomic decomposition by basis pursuit. SIAM J Sci Comput 20(1):33–61
2. Mallat SG, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Tran ASSP 41(12):3397–3415
3. Pati YC, Rezaiifar R, Krishnaprasad PS (1993) Orthogonal matching pursuit: recursive function approximations to wavelet decomposition. In: Proceeding of the 27th Annual Asilomar Conference on Signals Systems and Computers, pp 40–44
4. Gribonval R, Krstulovic S (2006) MPTK: matching pursuit made tractable. In: Proceedings of international conference on acoustics, speech, and signal processing (ICASSP'06), pp 496–499
5. Cotter SF, Rao BD (2002) Sparse channel estimation via matching pursuit with application to equalization. IEEE Tran Commun 50(3):374–377
6. Kim S, Iltis RA (2004) A matching-pursuit/GSIC-based algorithm for DS-CDMA sparse-channel estimation. IEEE Signal Process Lett 11(1):12–15
7. Tropp J (2004) Greed is good: algorithmic results for sparse approximation. IEEE Trans Inf Theor 50:2231–2242
8. Schnass K, Vandergheynst P (2008) Dictionary preconditioning for greedy algorithms. IEEE Trans Inf Theor 56(5):1994–2002
9. Divorra EO, Granai L, Vandergheynst P (2006) On the use of a priori information for sparse signal approximations. IEEE Trans Inf Theor 54(9):3468–3452

Chapter 170
The Study on Distribution Optimization Method Between Supply Chain Strategic Partners

Yang Jianhua and Mu Dandan

Abstract The distribution problem in supply chain is studied in this paper. To minimize the logistics cost in upstream and downstream enterprises in supply chain, bi-level programming method is innovatively used in this paper to establish the optimization model of the supply chain network distribution. Solving methods for this model are discussed and then an efficient arithmetic combined with layered iterative method is proposed. The model is calculated by Lindo software to verify the feasibility, efficiency and superiority of this method.

Keywords Bi-level programming · Layered iterative · Supply chain · Distribution optimization

170.1 Introduction

In traditional model, the object of the logistics planning for individual enterprise is to reduce the logistics cost for itself. Economic order quantity optimization and the network transportation cost minimization theory are usually used to minimize the inventory and transportation cost. However, in supply chain strategic partnership

Y. Jianhua (✉)
Shandong Institute of Business and Technology, Yantai
264005, People's Republic of China
e-mail: billyang@eyou.com

M. Dandan
Shanghai Maritime University, Shanghai 200135,
People's Republic of China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_170,

background, upstream and downstream enterprises with mutual cooperation relationship have their own constraints, control the decision variables respectively and have different profit goals. But the two enterprises are not independent with each other. The decision-making level in enterprises coalition could effectively coordinate with other enterprises in logistics decision-making, causing individual enterprise to be restricted to achieve the logistics goal of individual enterprise.

The final plan might not be the optimal one to individual enterprise but could minimize the logistics cost of the whole supply chain and consolidate the supply chain partnership relationship. This paper is proposed in the above background [1].

170.2 Problems Description

In supply chain strategic partnership background, upstream and downstream enterprises have their own constraints, control the decision variables respectively and have different benefit goals. But the two enterprises are not independent with each other. One the constraint condition and objective function of one enterprise include the decision variables controlled by the other enterprise in achieving the goal of win–win by cooperation. The final result might be overall optimization, but both enterprises could not realize their own optimization goals [1].

With the globalization of market, downstream enterprises must provide products and related service to geographically separate customers. To accelerate the reaction velocity, downstream enterprises establish many production lines in many places to meet the customer requirement. The customer requirement of the majority enterprises changes for the impact of the market competition. But in the industry significantly influenced by seasonal effect, the management decision makers have to consider the changing trend in logistics planning. Therefore, the logistics planning cannot be designed according to the demand in busy season to avoid the idle resources in low season. The material suppliers have to make their own adjustment (establishing new production lines or adjusting the distribution of the existing production lines) for certain decision of manufacturers (the adjustment of assembly lines or the increase of supply) to achieve the profit goal. Here the first-level network represents providing materials from upstream enterprises to downstream enterprises and the second-level network represents delivering goods from downstream enterprises to customers. The supply chain network structure is shown as Fig. 170.1. The factors we should determine include the distribution quantity in the lines of the second-level network (the output of production lines in downstream enterprises) and the distribution quantity in the lines of the first-level network.

The Bi-level Programming model can be used to describe the upstream and downstream enterprises distribution optimization problem under the supply chain strategic partnership background, for two different economic entities with different objectives and constraints are involved in this problem.

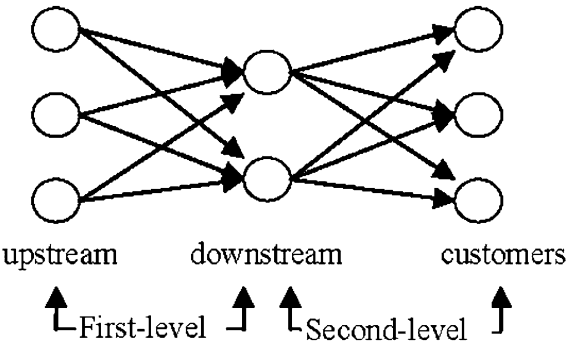


Fig. 170.1 Supply chain network structure

170.3 Bi-Level Programming Description

The optimization of supply chain distribution could be a Leader–Follower problem of which downstream manufactures are Leaders and upstream suppliers are Followers. The change of plans in downstream enterprises could cause the change of constraints and objective profits in upstream enterprises. By the same token, the decision variables controlled by downstream enterprises could cause the adjustment of the goal in upstream enterprises. This relationship could be described as Bi-level Programming.

The Bi-level Programming can be described as the following form:

$$(U)\ \min F(x,y) = a^T x + b^T y \tag{170.1}$$

$x \geq 0$, $y = y(x)$ obtained from the following planning

$$(L)\ \min f(x,y) = c^T x + d^T y \tag{170.2}$$

$$\text{s.t.}\quad Ax + By \leq r,\ y \geq 0$$

Wherein F, f—the objective function determined by Upper Planning and Lower Planning.

x, y—the decision variables of Upper Planning and Lower Planning.

Bi-level Programming model is consisted of two submodels including Upper Planning (U) and Lower Planning (L).

The upper layer decision maker influences the lower layer decision maker by setting up the value of x. Therefore, the feasible constraints set of the lower layer decision maker is restricted, while the action of the lower layer decision maker will influence the decision making of upper layer by the value of y. So the lower layer decision variable y is the function of upper layer decision variable x, namely y = y(x). This function is generally called reaction function [2].

Compared with Single-level Programming, the merits of Bi-level Programming are:

a. Analyze simultaneously two different and conflicting entities in the decision making process;
b. Multi-value decision-making method of Bi-level Programming more corresponds with the actual situation;
c. One complicated problem can be separated and considered as two layers with mutual effect.

Therefore, Bi-level Programming is widely used in the decomposition technique of large scale system, the hierarchical planning method in production control field, object and multi-objet programming, balanced programming, game theory, cybernetics and transportation planning and so on [3].

170.4 The Construction and Solution of Bi-Level Programming Model

170.4.1 The Construction of Bi-Level Programming Model

Suppose: only upstream and downstream enterprises exists in supply chain; the core enterprise is the downstream enterprise facing the final customers directly; the transportation cost is related with the volume in direct proportion; distribution outsourcing and inventory situation are not considered; the problem is about single product distribution; goods required by assemble line and demand points are transported once; the transportation cost of the first-level network is shared by the two enterprises in proportion; the idle cost should be calculated by idle productivity of production lines.

$$(U): \min F = \sum_i \sum_j C_{ij} X_{ij} + q \sum_i \sum_j C'_{ij} Y_{ij} + \sum_i a_i \left(\max P_i - \sum_j X_{ij} \right) \tag{170.3}$$

$$(L): \min f = (1-q) \sum_i \sum_j C'_{ij} Y_{ij} + \sum_i a'_i \left(\max P'_i - \sum_j Y_{ij} \right) \tag{170.4}$$

$$\text{s.t.} \quad \sum_j X_{ij} \le \max P_i, \ \ i = 1, 2, \ldots, n$$

$$\sum_j y_{ij} \le \max P'_i, \ \ i = 1, 2, \ldots, n$$

$$\sum_j X_{ij} \le \sum_j Y_{ji}, \ \ i = 1, 2, \ldots, n$$

C_{ij} represents the transportation fee from manufacture point *i* to demand point j; X_{ij} represents the transportation volume from manufacture point *i* to demand point j; P_i represents the maximum distribution of supplier *i*, P'_i represents the maximum distribution of manufacture i; C'_{ij} represents the transportation fee from supplier *i* to manufacture j; Y_{ij} represents the transportation volume from supplier *i* to manufacture j; a_i and a'_i represent the idle cost of the manufacture and the supplier respectively; q represents the proportion ratio.

170.4.2 The Solution of Bi-Level Programming Model

Generally, the solution of Bi-level Programming problem is complicated. One of reasons is that Bi-level Programming problem is an NP-hard problem. Ben-Ayed and Blair studied this problem thoroughly in 1988 [4] and proposed that there is no multinomial solving arithmetic for Bi-level Programming problem even if it is a simple one. Nonconvexity of Bi-level Programming is another reason for exceptional complicatedness of the solution of Bi-level Programming problem [5].

The solution of Bi-level Programming model is generally to find the expression of reaction function. But this function is nonlinear function and the expression is unknown. In the Bi-level Programming model proposed by this article, the decision variable X_{ij} of Upper Planning (U) and the decision variable Y_{ij} of Lower Planning (L) have relative internal relations, however, it is hard to directly find the relational expression between the two, namely the expression of the reaction function is unknown. Therefore, the Bi-level Programming model can be solved by layered iterative method. The basic line of thinking is as follows:

Firstly, progress the initialization preparation work including setting up the initial solution of downstream distribution volume X_{ij}, substitute it in Lower Planning (L), solving the Lower Planning problem by existing optimization algorithm and obtain the distribution volume of upstream suppliers Y_{ij}; Substitute the above Y_{ij} in Upper Planning (U) and obtain the new distribution volume X_{ij}; Secondly, solve the lower layer problem again and obtain the new Y_{ij}; Repeat the calculation process in this way and finally have a chance of obtaining the solution converged to the optimum solution of the primary Bi-level Programming model.

The specific calculation procedure is as follows:

Step 0. Set up the initial solution of downstream enterprise distribution X_{ij}^0, and make iterate $N = 0$;

Step 1. Solve the lower layer problem by the given X_{ij}^0 and obtain the upstream distribution volume Y_{ij}^{N+1} and the objective function value of lower layer programming L_{N+1};

Step 2. Solve the upper layer problem by given Y_{ij}^{N+1} and obtain the downstream enterprises distribution volume X_{ij}^{N+1} and the objective function value of upper layer programming U_{N+1};

Step 3. Make $\Phi_U = \left|\frac{U_{N+1} - U_N}{U_{N+1}}\right|$, $\Phi_L = \left|\frac{L_{N+1} - L_N}{L_{N+1}}\right|$

If $\max(\Phi_U, \Phi_L) \leq \mu$, then stop, μ is iterate accuracy; otherwise, make $N = N+1$ and turn to step 1;

Step 4. Output the optimum solution and the optimum value.

170.5 Case Analysis

Beverage Company A has two can packing companies in Xiamen and Dongguan. The maximum output of the two companies is 25500, 51100, (ten thousand cans) respectively, and idle cost is 0.1, 0.2 (ten thousand yuan/ten thousand cans). The demand places are Shaoxing and Sanshui with the demand of 14,479, and the transportation cost is (0.6 0.5 0.65 0.25). The beverage can supplier B has canners in Zhangjiagang, Hangzhou and Fanyu to provide cans to Xiamen and Dongguan. The maximum output of the three companies is 14600,51100,29200. The transportation cost is (0.6 0.7 0.6 0.65 0.5 0.35) and idle cost is 0.1,0.2,0.3. The two companies share the transportation cost of the fist-level network. Company A share 20% and Company B share 80% [6].

The objective function of upper and lower planning is obtained:

$$\min F = 0.6x_{11} + 0.5x_{12} + 0.65x_{21} + 0.25x_{22} + 0.12y_{11} + 0.14y_{12} + 0.12y_{21} + 0.13y_{22} + 0.1y_{31} + 0.07y_{32} + 0.1(25550 - x_{11} - x_{12}) + 0.2(51100 - x_{21} - x_{22}) \tag{170.5}$$

$$\min f = 0.48y_{11} + 0.56y_{12} + 0.48y_{21} + 0.52y_{22} + 0.4y_{31} + 0.28y_{32} + 0.1(14600 - y_{11} - y_{12}) + 0.2(51100 - y_{21} - y_{22}) + 0.3(29200 - y_{31} - y_{32}) \tag{170.6}$$

Therefore

$$(U)\ \min F = 0.5x_{11} + 0.4x_{12} + 0.45x_{21} + 0.05x_{22} + 0.12y_{11} + 0.14y_{12} + 0.12y_{21} + 0.13y_{22} + 0.1y_{31} + 0.07y_{32} + 12775 \tag{170.7}$$

$$(L)\ \min f = 0.38y_{11} + 0.46y_{12} + 0.28y_{21} + 0.32y_{22} + 0.1y_{31} - 0.02y_{32} + 20440 \tag{170.8}$$

$$\text{s.t.}\quad x_{11} + x_{21} = 14479$$

$$x_{12} + x_{22} = 14479$$

$$x_{11} + x_{12} \leq 25550$$

$$x_{21} + x_{22} \leq 51100$$

$$-x_{11} - x_{12} \leq -2000$$

$$-x_{21} - x_{22} \leq -2000$$

$$x_{11} + x_{12} - y_{11} - y_{21} - y_{31} \leq 0$$

$$x_{21} + x_{22} - y_{12} - y_{22} - y_{32} \leq 0$$

$$y_{11} + y_{12} \leq 14660$$

$$y_{21} + y_{22} \leq 51100$$

$$y_{31} + y_{32} \leq 29200$$

$$-y_{11} - y_{12} \leq -2000$$

$$-y_{21} - y_{22} \leq -2000$$

$$-y_{31} - y_{32} \leq -2000$$

$$x_{11}, x_{12}, x_{21}, x_{22}, y_{11}, y_{12}, y_{21}, y_{22}, y_{31}, y_{32} \geq 0$$

Set up the iterate accuracy $\mu = 0.001$

Step 0. Set up the initial solution

Sale/Pro.	Shaoxing	Sanshui	V	Max.Output	Min.Output
Xiamen	0.6	0.5	0.1	25,500	2000
Dongguan	0.65	0.25	0.2	51,100	2000
Demand	14,479	14,479	47,642		

The solution is obtained by table dispatching method

Sale/Pro.	Shaoxing	Sanshui	V
Xiamen	2000	0	23500
Dongguan	12,479	12,479	24,142

$X^0 = [2000\ 0\ 12479\ 14479]$

Step 1. Solve the Lower Planning (L) by given

$$L: \min f = 0.38y_{11} + 0.46y_{12} + 0.28y_{21} + 0.32y_{22} + 0.1y_{31} - 0.02y_{32} + 20440 \tag{170.9}$$

$$\text{s.t.} \quad y_{11} + y_{21} + y_{31} \geq 2000$$

$$y_{12} + y_{22} + y_{32} \geq 26958$$

$$2000 \leq y_{11} + y_{12} \leq 14600$$

$$2000 \leq y_{21} + y_{22} \leq 51100$$

$$2000 \leq y_{31} + y_{32} \leq 29200$$

The solution is obtained by Lindo software:
$Y^1 = [2000\ 0\ 2000\ 0\ 0\ 29200]$
$L_1 = 21176$

Step 2. Substitute Y^1 in Upper Planning (U)

$$(U): \min F = 0.5x_{11} + 0.4x_{12} + 0.45x_{21} + 0.05x_{22} + 15299 \qquad (170.10)$$

s.t. $x_{11} + x_{21} = 14479$

$$x_{12} + x_{22} = 14479$$

$$2000 \leq x_{11} + x_{12} \leq 4000$$

$$2000 \leq x_{21} + x_{22} \leq 29200$$

$X^1 = [2000\ 0\ 12479\ 14479]$
$U_1 = 22638.5$

Step 3. Iterate and obtain $L_2 = 21176$, $U_2 = 22638.5$

$$\Phi_U = \left|\frac{U_2 - U_1}{U_2}\right| = 0,\ \Phi_L = \left|\frac{L_2 - L_1}{L_2}\right| = 0$$

$$\max(\Phi_U, \Phi_L) \leq 0.001$$

Therefore,Upper Planning (U)
Optimum solution is
$X^* = [2000\ 0\ 12479\ 14479]$
Optimum value is $U^* = 22638.5$
Lower Planning (L)
Optimum solution is
$Y^* = [2000\ 0\ 2000\ 0\ 0\ 29200]$
Optimum value is $L^* = 21176$

170.6 Conclusion

The optimization problem of supply chain distribution is described by Bi-level Programming model in this article combined with the calculation of specific case. The initial solution is obtained by table dispatching method to reduce the iterations and improve the efficiency of calculation. The feasibility of applying Bi-level Programming model to supply chain strategic partnership distribution optimization problem is approved.

Acknowledgments Natural Social Science Foundation of P.R.China, No: 10BGL011; National Postdoctoral Foundation, No. 20100471548;Humanity & Social Science Foundation of Education Ministry, No: 10YJC630327 & 09YJA630085; Shandong Province Nature Sciences Foundation, No: ZR2009HL005.

References

1. Huijun S, Ziyou G (2003) Bi-level optimization model for distribution system of supply chain. J Manage Sci 6(3):66–70
2. Zhigang Z, Xinyi G (2006) Bi-level programming method for distribution network model in supply chain. Univ Shanghai Sci Technol 28:300–302
3. Lixin S, Yan C, Zhaogang S (2005) A model and solution algorithm based on bi-level programming for the partner selecting of virtual logistic enterprise alliance. Sci Technol Eng 5(4):245
4. Huijun S, Ziyou G (2004) An optimization model for two-echelon distribution network design in supply chain based on bi-level programming. J Ind Eng/Eng Manage 1:68–71
5. Ziyou G, Yifan S, Bingfeng S (2000) Network design of urban transportation consecutive balance—theory and method. China Railway Press, Beijing
6. Peng G (2006) The study on logistics planning of supply chain strategic partnership based on bi-level programming model. Zhejing University, Zhejing

Part XII
Computer Application and Communication Engineering

Chapter 171
Advanced Route Design Based on Properties of Nodes in Opportunistic Networks

Qilie Liu, Xiang Pang, Yun Li and Yingjun Pan

Abstract Opportunistic network is a kind of highly partitioned wireless ad hoc networks. Message ferry (MF) as an availably solution exploit controlled mobility to forward message in a specific route. In this paper, we study the problem of efficient data delivery in opportunistic networks, and propose an Advanced Route Design Based on Properties of nodes (ARDBP) to find the appropriate nodes which should be visited on the next route to improve the data delivery performance. The enhanced route algorithm designs the choice of ferry to deliver messages according the different properties of nodes. We evaluate the performance via ONE simulation to confirm the enhanced approach is efficient in data delivery performance, compared to the Nearest Neighbor Algorithm of Traveling Salesman Problem (TSP).

Keywords Opportunistic networks · Message ferry · Delivery performance · Routing design

171.1 Introduction

In opportunistic networks, nodes are deployed sparely so that the networks are often vulnerable and links between nodes may break down for a long time at any time. One existing challenge is the question of how to improve the data delivery

Q. Liu (✉) · X. Pang · Y. Li
Key Lab of Mobile Communication technology,
Chongqing University of Posts and Telecommunications,
Chongqing 400065, China
e-mail: liuql@cqupt.edu.cn

Q. Liu · Y. Pan
The Key Lab of OptoElectronics Technology and System,
Ministry of Education, Chongqing University,
Chongqing 400044, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_171,

performance. In recent years routing in ad hoc networks is an active research field, and a number of routing algorithms have been proposed [1–3]. To enhance the connectivity, a Message Ferrying (MF) mechanism has been proposed [4, 5]. The mechanism has shown that MF can provide communication service and delivery performance for nodes in opportunistic networks. Therefore, the key challenge is how to find the ferry routes to enhance the connectivity and delivery performance. The ME problem has been proved to be an NP-complete problem in wireless sensor networks [6]. In recent years, most of the methods are using a single ferry to visit all nodes on one route in the shortest time. In addition, PBS [7] algorithm studied the data loss rate for a given mobile element speed and the minimum speed to avoid buffer overflow.

In this paper, we concentrate on the route design for a single ferry and study the efficient data delivery in opportunistic networks. We propose an advanced route design based on different properties of nodes (ADRBP) to design the path selection strategies for ferry to deliver messages.

This paper is organized as follows. Section 171.2 proposes the network model. In Sect. 171.3, we describe the ferry route problem and present a new scheduling algorithm, Sect. 171.4 discusses the performance analyses. Section 171.5 discusses the simulation results and Sect. 171.6 outlines some directions for future work.

171.2 Network Model and Discussion

171.2.1 Network Model

We consider such a distribution of nodes in a network topology, where nodes with higher data generation rate are generally distributed closer to the center of the topology, nodes with lower data generation rate are generally distributed more remote locations, and nodes with higher generation rate will receive more data. If ferry visits all the nodes in the topology every route, it will generate a very long path for ferry and cause large data delay for the entire topology. In this route method, ferry takes the same path access strategy to visit all nodes. The ferry cannot be timely accessed to nodes with high data generation rate and cost a lot of time to access to nodes with low data generation rate. The method impacts the data transmission delay and delivery rate and other properties.

According to these analyses, we proposed an advanced route design algorithm for ferry to take different access frequencies policy according to the different nature of nodes, so that the nodes with higher data generation are visited more frequently and the nodes with lower data generation are visited timely.

171.2.2 Parameter Notation and Problem Formulation

In this paper, we assume that: regular nodes are stationary and ferry moves at a constant speed v. The ferry and nodes have limited wireless communication range, the size of each packet is normalized to 1, the ferry has no buffer limitation, all nodes have the same limited buffer and transmit rate between ferry and a regular node is r. Let $N = \{1, 2, \ldots, n\}$ denote a set of n regular nodes. A ferry route can be mathematically described as an optimized (O) permutation of the regular nodes.

Let b_{ij} denote the number of packets destined to node j from node i and d_{ij} is the average delay for traffic from node i to j a given ferry path. In this paper, we no longer consider the data transmission time between ferry and nodes. We have the average delay d_{ij}: for messages from node i to node j consists the waiting time d_{ij}^{w} at node i before ferry receive and the carrying time d_{ij}^{T} at ferry before delivered to node j, as $d_{ij} = d_{ij}^{w} + d_{ij}^{T}$.

For any given path O, we define the objective function C_O as the average delivery time for all traffic, the average message delay over all the packets can be defined as:

$$C_o = E\left[\frac{\sum b_{ij} d_{ij}}{\sum b_{ij}}\right] \tag{171.1}$$

The problem of ferry route design is finding an optimal path to minimize the objective function C_O (171.1).

171.3 Related Strategies

In this paper, with the ARDBP algorithm, we first put nodes into different sets according to data generation rate of nodes. Then, we divide nodes in each set according to their location distribution topology into different subsets. Finally we choose the nodes which ferry should visit on the next route and compute the shortest path of the nodes in these subsets.

171.3.1 Nodes Segmentation

We have the following assumptions: Supposing ferry and node's physical size can be negligible, ferry can change its direction anytime and without considering delay.

Let r_i denote the data generation rate of node $i = 1, 2, \ldots, n$, all nodes have the same limited buffer b. We group the nodes into different sets according to their data production rate, and denote the set as S_j. Assuming the number of sets is M, we denote $r_{\min}$ and $r_{\max}$ as the minimum and maximum data generation rate of all nodes, the node i in set S_j is satisfied with the following equation:

$$\begin{cases} 2^{-j} r_{\max} \le r_i \le 2^{-(j-1)} r_{\max}, & j = 1, \ldots, M-1 \\ r_{\min} \le r_i \le 2^{j} r_{\max}, & j = M \end{cases} \tag{171.2}$$

We can get different sets of nodes according to the formula (171.2). The range of data generation rate of nodes in set B_i is twice the data generation rate range of nodes in set $S_{i+1.}$ Therefore, we should visit the nodes data with generation rate more frequently than others.

171.3.2 Division Region of Nodes

According to the location distribution of nodes in the network topology, we can divide the region of node sets into different subregions. The KD-tree method had been presented in the literature [8], which is utilized to partition the topology. We use the 2D-tree method in our algorithm, and the two dimensions are the length and width of the topology.

Figure 171.1 shows an example of 2D-tree algorithm to partition nodes in the topology. In the first step, we choose nodes from set S_1 in the whole region A_1, then divide A_1 into two subregions A_{21} and A_{21} by their average x-coordinate value: $\frac{1}{N}\sum_{i=1}^{N} x_i$ (N denotes the number of nodes which have been chosen, x_i is their x-coordinates), the A_{21} is the area which x-coordinates smaller than the average horizontal coordinate value and the rest to the A_{21}. In the next step, we respectively choose nodes from the set S_2 in the sub-regions A_{21} and A_{21}, and divide them into two subregions A_{31}, A_{32} and A_{33}, A_{34} according to the node's average y-coordinate value: $\frac{1}{N}\sum_{i=1}^{N} y_i$ (N denotes the number of nodes which have been chosen, y_i is their y-coordinates), repeat until the subregions of nodes in the last set S_M have been divided.

We can obtain the following relationship between the subregions A_{ij} as shown in Fig. 171.2, the low region belongs to the top region when connected through the line.

The number of subregions in the set S_i is less than 2^{i-1}. The total number is less than $2^{M}-1$. The subregion A_{ij} can partition into $A^{j}_{i+1,2\text{-}1}$ and $A^{j}_{i+1,2}$ at most. It shows that the combination of two lower subregions connected the top region through the line is equivalent to the upper subregion.

171.3.3 Route Design

In this section, we choose the appropriate nodes to design the route for the ferry. We have to solve two problems: how to make the ferry to access the nodes with higher data generation rate more frequently and how to access the nodes with very low data production rate timely.

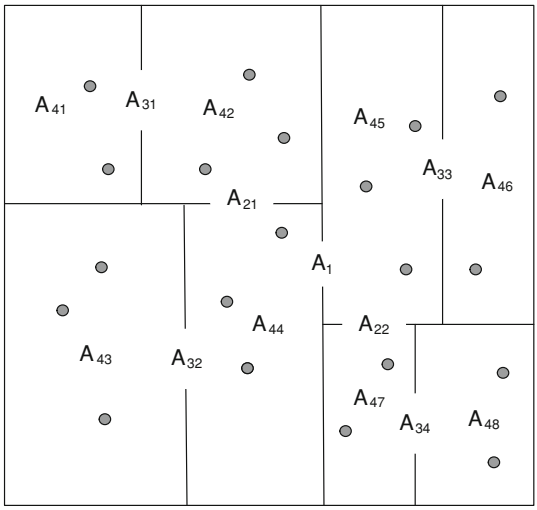


Fig. 171.1 Nodes in a 2 dimension space cut by 2D-tree

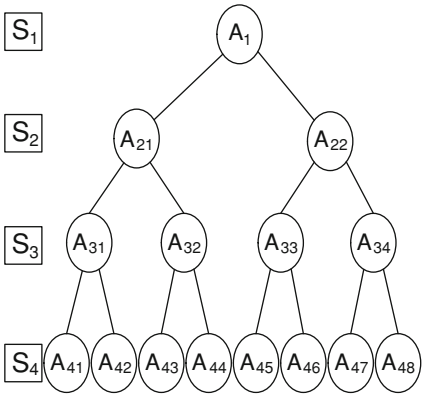


Fig. 171.2 The relationship between subregions

When the ferry has known the nodes should be visited on the route, then the route problem is reduced to the Traveling Salesman Problem (TSP) for ferry. There are a large of algorithms to calculate the TSP paths. We adopt the Nearest Neighbor algorithm.

We choose the front m sets of nodes satisfied with the number of nodes in the front m sets no less than half of the total number of nodes. First of all, we choose the nodes in the set S_1 from the whole region A_1, and then choose the nodes in the set S_2 from subregion A_{21} which belongs to A_1. After that we choose the nodes in B_3 from the subregion S_3 of the area A_{31} which belongs to A_{21}. Repeated until choose nods in the set S_m. The chosen nodes are the ferry should access on the first route, and then the problem is reduced to the TSP for ferry. In the second selection, we choose the nodes in the set S_1 from the whole region A_1, and then choose the nodes in the set S_2 from sub-region A_{22} which belongs to A_1. After that we choose

the nodes in B_3 from the subregion S_3 of the area A_{33} which belongs to A_{22}. Repeat until choose nods in the set S_m. So we ensure the nodes and calculate a shortest path as the second path. We can get the third and forth path for ferry. Finally ferry visits nodes according to this order. This route design solves the problem of how to visit the nodes with higher data generation rate more frequently.

The second problem is how to access nodes with very low data production rate timely. Let d_i denote the accumulation of message delay of the node i from the last visit time of ferry to the time t:

$$d_i = \int_{t_i}^{t} r_i(t-x)dx \tag{171.3}$$

Let $\frac{1}{2n}\sum r_i T^2$ denote the average accumulation of message delay when ferry visit all n nodes on every route, and the period of time is T. We compared to the two values to decide when to visit the nodes that have not been chosen. When the node's accumulation of message delay is satisfied with $d_i = \int_{t_i}^{t} r_i(t-x)dx \geq \frac{1}{2}\sum r_i T^2$, we should put the node into the next route.

From the above discussion, we first choose the front m sets of nodes satisfied with the number of nodes in the front m sets no less than half of the total: $\Sigma_{i=1}^{m} S_i \geq \frac{n}{2}$; then we choose the other nodes according to their accumulation of message delay.

171.4 Performance Analysis

171.4.1 The Shortest Path

Without considering the data loss, for a selected path π that ferry visited all nodes in every path, we denote i is the ID which node is visited at order i, and its data generation rate is r_i. The ferry will periodic visit the fixed path, the interval time is one period T. Now we consider the average data delay for the selected path π. According to the formula (171.1), we can get the objective function:

$$[C_\pi]_{\min} = E\left[\frac{\sum b_{ij} d_{ij}}{\sum b_{ij}}\right]_{\min} \Rightarrow E\left[\sum b_{ij} d_{ij}\right]_{\min} \tag{171.4}$$

Since ferry visit the all nodes periodic, $E\left[\sum b_{ij}\right]$ is a constant. Let $T_{\Pi(i),\Pi(i=1)}$ denote ferry's travel time from i to $i+1$ in the path π. Then Eq. 171.4 minimizes the total travel time of ferry.

$$[C_\pi]_{\min} \Rightarrow \sum_{k=1}^{n} T_{\pi(k),\pi(k+1)} \tag{171.5}$$

The path selection of visiting all nodes for ferry is turned to the shortest route of TSP. Assume that we use the Nearest Neighbor algorithm to find the optimal route O. Supposing the cycle time of the route is T, according to (171.1) we have the average data delay C_O:

$$C_O = \frac{\sum b_{ij} d_{ij}}{\sum b_{ij}} = \frac{\frac{T}{2}\sum_i r_i T}{\sum_i r_i T} = \frac{T}{2} \tag{171.6}$$

From Eq. 171.6, the average data delivery delay is only relating to the route's cycle time T.

171.4.2 The Route Design

When the number of nodes in the ith set is 2^{i-1}, and uniform distribution on the circumference with diameter $2^{i-1}d$, the nodes in the ith set have the same data generation rate $2^{1-i}r_1$. We denote this distribution as the standard model. n and M can be linked with $M = log_2(N+1)$. The data generation rate of nodes in set i are all r_i, and the value of r_i is calculated as: $r = 2^{1-i}r_1$, $i = 1,\ldots,m$.

We use the route design algorithm for ferry to visit the noses with standard model in every set. On every path, ferry has the same cycle time T. The average delay of set i is:

$$\bar{d}_i = \bar{d}_i^w + \bar{d}_i^T = 2^{i-2}T + \frac{1}{m}\left(2^{-1}T + T + \cdots + 2^{m-2}T\right) \tag{171.7}$$

According to (171.1) and (171.7) we have the average data delay C_O as following:

$$C_O = \frac{\sum_{ij} b_{ij} d_{ij}}{\sum_{ij} b_{ij}} = \frac{2^m - 1}{m} T \tag{171.8}$$

From Eq. 171.8, the average data delay is related to the number of sets M and the cycle time T.

171.4.3 Complexity Analyses

Supposing there are n nodes with the standard distribution model in the network and divided into M sets. Then each step in the complexity for ADRBP is:

(1) Nodes segmentation according to contrast node's data generation rate, the complexity is $O(n)$.

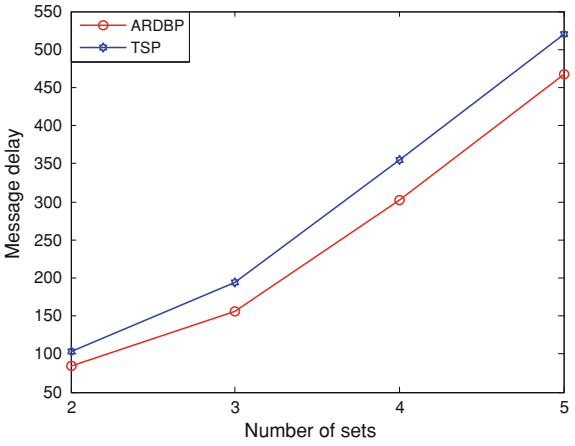


Fig. 171.3 Message delay for number of sets

(2) In the division region of nodes step, we partition nodes in every set into 2^{i-1} subregion; the computation order of complexity is $O(2^M)$.
(3) In the route design step, we first confirm the sets of nodes should be selected with the complexity of $O(n)$. Then the complexity of calculating the shortest path for $m-1$ nodes is $O((m-1)^2)$. The complexity of confirming accumulation of data delay and distance of other nodes is $O(n^2 2^{M-1})$. At last to connect all paths to a whole path, the complexity is $O(2^{M-2})$.

Therefore, the computation order of complexity of ADRBP in the completely calculation is $O(n^2 2^{M-1}+2^M)$.

171.5 Simulations

In this section, we evaluate the performance through ONE simulation, and we use the following default settings. Nodes are distributed in 100 * 100 unit^2 area. All nodes have the same limited buffer. The size of a message is one unit. The single ferry has the speed of 1 unit in a unit time. The transmission range of both sensor node and ferry is 2 units, and the transmission rate is 100 messages in a unit time. We evaluate the performance for 10000 time units on each simulation.

At first, we consider the nodes distribution with standard model, as the number of sets m is equal to 2, 3, 4 and 5. We respectively compared the performance of average message transmission delay between TSP and ADRBP, the result is shown in Fig. 171.3.

Then we evaluate the performance of ADRBP and TSP algorithm with the number of nodes n equal to 10, 20, 30, 40 and 50. The nodes are randomly distributed in the topology and divided into 3 sets. From Figs. 171.4 and 171.5, the

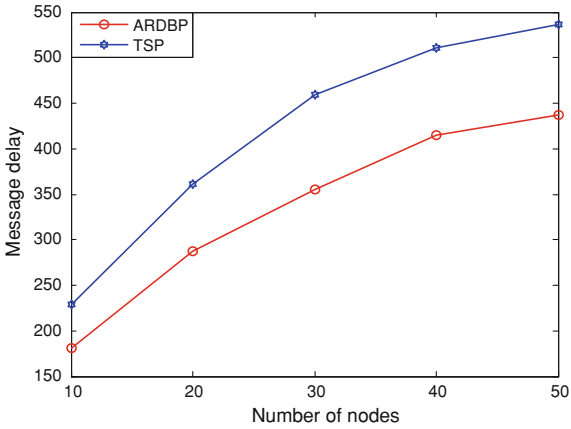


Fig. 171.4 Message delay for number of nodes

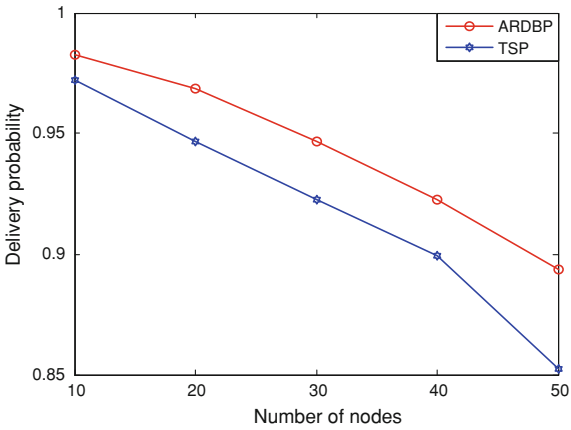


Fig. 171.5 Delivery probability for number of nodes

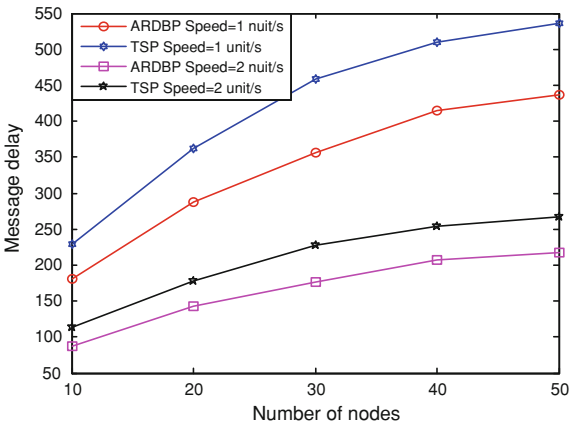


Fig. 171.6 Message delay for speed of ferry

ADRBP has a better performance of average data delay and delivery probability with the increasing number of nodes.

Additionally, we consider the performance of speed of ferry. The message delay decreases with the speed of ferry increases in both algorithms, because ferry can move quickly between nodes. The performance about average message delay is shown in Fig. 171.6.

171.6 Conclusion and Future Work

The nodes in the network have different properties such as data generation rates, location distribution, we propose the ARDBP algorithm such that the ferry can visit the nodes with higher data generation more frequently and visit the nodes with lower data generation timely. We evaluate the performance of ADRBP algorithm via *ONE* simulations, and the result showed that the ADRBP algorithm can increase the performance of average message delay. Our future work will consider communicating nodes with multiple ferries in opportunistic networks.

Acknowledgments This paper is Supported by the National Science Foundation of China (61071118), the Science and Technology Research Project of Chongqing Municipal Education Commission of China (KJ110524), the Natural Science Foundation Project of CQ CSTC(CSTC2010BA4010), Program for New Century Excellent Talents in University (NCET-07-0914), the Grant by the Science & Tech. Commission of Chongqing (CSTC2009BB2279), and China Postdoctoral Science Foundation (20090451158), and the National Science Foundation of Chongqing University of Posts and Telecommunications (A2009-28, A2009-41).

References

1. Beaufour A, Leopold M, Bonnet P (2002) Smart-tag based data dissemination, First ACM WSAN Workshop, Sept 2002
2. Davis J, Fagg A, Levine B (2001) Wearable computer as packet transport mechanisms in highly-partitioned ad-hoc networks, International Symposium on Wearable Computing, Oct 2001
3. Vahdat A, Becker D (2000) Epidemic routing for partially-connected an hoc networks, Technical report, Duke University
4. Zhao W, Ammar M (2003) Message ferrying: proactive routing in highly-partitioned wireless ad hoc network, Proceedings of the 9th IEEE Workshop on Future Trends in Distributed Computing System, Puerto Rico, May
5. Zhao W, Ammar M, Zegura E (2004) A message ferrying approach for data delivery in sparse mobile ad hoc networks, ACM MobiHoc, pp 187–198
6. Somasundara AA, Ramamoorthy A, Srivastava MB (2004) Mobile element scheduling for efficient data collection in wireless sensor networks with dynamic deadlines, In 25th IEEE International RealTime Systems Symposium(RTSS'04), pp 296–305
7. Elissa K (2005) Partitioning based mobile element scheduling in wireless sensor networks in Scend Annual IEEE Communications Society Conference
8. Bentley JL Multidimensional binary search trees used for associative searching, Communications of the ACM, 18(9):509–517

Chapter 172
Study of Two Adaptive Beam-Forming Algorithms

Duan Li and Wei Li

Abstract The technology of adaptive beam-forming is the key technology in smart antenna. The performance of the whole smart antenna system is directly affected by the convergency and computation complexity of the adaptive algorithm. Based on the principle of smart antenna adaptive beam-forming, this chapter analyzes the Least-Mean Squares (LMS) and Recursive Least-Squares (RLS) algorithms, and suggests an improvement of LMS algorithm. The convergence result and performance indicators were analyzed through simulation. In the same environment, the two algorithms were compared. Finally, a DSP + FPGA beam forming system is designed based on adaptive algorithm.

Keywords Smart antenna · Adaptive beam forming · LMS · RLS

172.1 Introduction

Smart antenna technology is attracting more and more attention which is one of the key technologies for the future wireless communication. Smart antenna beam-forming algorithm can be divided into two classifications: blind algorithm and non-blind algorithm in terms of whether it needs reference signal [1].

D. Li (✉)
School of Electrical Engineering and Automation,
Henna Polytechnic University, Jiaozuo 454000, China
e-mail: liduan@hpu.edu.cn

W. Li
School of Information and Engineering, Huanghe Science
and Technology University, Zhengzhou 450000, China
e-mail: nysq216@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_172,

Two non-blind beam-forming algorithms which are Least-Mean Squares (LMS) algorithm and Recursive Least-Squares (RLS) algorithm are analyzed and compared, and the simulation results are also given.

172.2 Smart Antenna Beam Forming Technology

The beam-forming system can get a higher SNR and super target resolution [2]. There is an antenna array model, as the Fig. 172.1.

Suppose there are $J + 1$ signals, among them 1 expected signal and J interference signals. The Direction of Arrive (DOA) of expected signal is θ_d. J interference signals' DOA are $\theta_1, \theta_2, \ldots\ldots, \theta_J$. The input M*1−D vector signal x(t) can be shown as:

$$x(t) = As(t) + n(t) \tag{172.1}$$

Among them,

$$x(t) = [x_0(t), x_1(t), \ldots, x_{M-1}(t)]^T \tag{172.2}$$

$s(t)$ is a $(J + 1) \times 1$−D vector that represent 1 desired signal and J interference signals.

$$s(t) = [s_d(t), s_1(t), \ldots, s_J(t)]^T \tag{172.3}$$

$n(t)$ is M*1−D noise vector that represent an additive white noise with σ_n^2 variance.

$$n(t) = [n_0(t), n_1(t), \ldots, n_{M-1}(t)]^T \tag{172.4}$$

Array steering matrix A is:

$$A = [a(\theta_d), a(\theta_1), \ldots, a(\theta_J)] \tag{172.5}$$

Among them, column vector $a(\theta_d)$ is the steering matrix of desired signal, $a(\theta_1), \ldots, a(\theta_j)$ are J interference signals' steering matrix.

There,

$$a(\theta_k) = [1, e^{-j\phi_k}, \ldots, e^{-j(M-1)\phi_k}]^T \tag{172.6}$$

$$\phi_k = \frac{2\pi d}{\lambda} \sin\theta_k (k = d,\ 1,\ 2, \ldots,\ J) \tag{172.7}$$

Then we can get output signal:

$$y(t) = w^H x(t) = \sum_{m=0}^{M-1} w_m^* x_m(t) \tag{172.8}$$

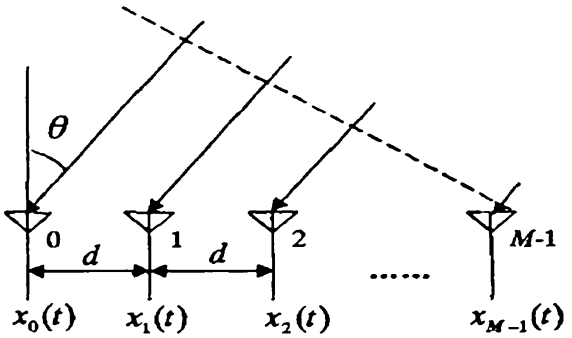


Fig. 172.1 Uniform linear array model

There, weight vector $w = [w_0, w_1, \ldots, w_{M-1}]$

Amplitude beam response is:

$$F(\theta) = |y(t)| = \left|w^H a(\theta)\right| \tag{172.9}$$

The weights of beam pointing in the θ_0 direction:

$$w(\theta_0) = a(\theta_0) = [1, e^{-j\phi_0}, \ldots, e^{-j(M-1)\phi_0}]^T \tag{172.10}$$

There, $\phi_0 = \dfrac{2\pi d}{\lambda}\sin\theta_0$.

The beam pattern is:

$$\begin{aligned} F(\theta) &= \left|w^H(\theta_0)a(\theta)\right| = \left|\sum_{m=1}^{M} e^{j(m-1)(\phi_0-\phi)}\right| \\ &= \left|\frac{\sin[(M\pi d/\lambda)(\sin\theta_0 - \sin\theta)]}{\sin[(\pi d/\lambda)(\sin\theta_0 - \sin\theta)]}\right| \end{aligned} \tag{172.11}$$

When $\theta = \theta_0$, the maximum gain of amplitude beam response can be got. By changing the numerics of θ_0, the beam-forming can point anywhere.

So the technique of adaptive beam forming is obtaining desired signal and avoiding all the interferences and noises by changing the complex weights through adaptive algorithm.

In the article, E[•] show mathematic expectation, $(\bullet)^{-1}$ show inversion, $[\bullet]^{\mathrm{T}}$ show transpose, (•)* show complex conjugate.

172.3 Adaptive Non-Blind Beam Forming Algorithm

Adaptive beam forming algorithm which can be grouped into blind beam forming algorithms and non-blind beam forming algorithms is the key of adaptive beam forming technology. Non-blind algorithm needs reference signal. LMS algorithm

and RLS algorithm are non-blind beam forming algorithms. They are based on the LMS error rule [3].

172.3.1 Least-Mean Squares Algorithm

LMS algorithm does not rely on correlation matrix, but uses the nonlinear optimization, known as steepest descent method, to calculate the metrics [4]. Then, the metrics can be shown as:

$$w_N(k+1) = w_N(k) - \frac{1}{2}\mu\nabla J(w_N(k)) \tag{172.12}$$

LMS algorithm has an important position in the field of adaptive technology, because of its simple principle and less amount of calculation. Finally, we can get the formula to update weight vector as follows.

$$e_N(k) = w_N^H(k)X_N - d_N(k) \tag{172.13}$$

$$w_N(k+1) = w_N(k) - \mu X_N e_N^*(k) \quad 0\langle\mu\langle\text{Trace}(R) \tag{172.14}$$

The algorithm would have fast convergence speed, if weight vector has minimal change in the process of each iteration. The modified algorithm is normalized least-mean square (NLMS) algorithm. NLMS is actually a normalized varying step-size LMS, whose speed of constringency and effect of track is better than traditional LMS [5]. The formula to update weight vector is:

$$w_N(k+1) = w_N(k) - \frac{\mu}{\|X_N\|^2}X_N e_N^*(k) \tag{172.15}$$

172.3.2 Recursive Least-Squares Algorithm

Based on the principle of Least-squares algorithm, RLS algorithm minimizes the squares sum of array output error in every snapshot RLS needs all the array data acquired after initialization and achieves matrix inversion using iterative method Therefore, it has fast convergence and lower arithmetic complexity [5]. Recursive formula of RLS algorithm is as follows.

$$v(n) = p(n-1)x(n) \tag{172.16}$$

Filter gain vector:

$$z(n) = \frac{\lambda^{-1}v(n)}{1+\lambda^{-1}x^H(n)v(n)} \tag{172.17}$$

$$\varepsilon(n) = d(n) - w^H(n-1)x(n) \tag{172.18}$$

Updated weight formula:

$$w(n) = w(n-1) + z(n)\varepsilon^*(n) \tag{172.19}$$

The covariance matrice inverse:

$$p(n) = \lambda^{-1}(1 - z(n)x^H(n))p(n-1) \tag{172.20}$$

There, $w_0 = 0$, $p_0 = \delta^{-1}I$, δ are some smaller constant, λ is forgetting factor. $0 < \lambda < 1$.

172.4 The Analysis of Simulation Result

Supposing the simulated conditions are: The space of ULA is equal to half wavelength of the transmit signal. The number of antenna elements is eight. Supposing antenna arrays receive two user signals from -45 to $0°$. And the coming direction of desired signal is $-45°$, the other is interference signal. Each element of input signal is 1 or -1. The average power of signals is one; Noise vector is Gaussian noise, noise power is 0.01 and signal to noise ratio is 20 dB.

According to the convergence principle of LMS algorithm, step factor $0 < \mu < 2/\text{trace}$. Make $\mu = 0.0031$ and $\mu = 0.0123$, the direction diagram and convergence curve are gained as Fig. 172.2.

The diagram shows that step factor μ makes little effect on direction diagram. But from the convergence curve of Fig. 172.2b, it can be seen that increasing μ makes convergence speed much faster. If $\mu = 0.0123$, convergence will be achieved when $n = 40$. But if $\mu = 0.0031$, convergence can be achieved just when $n = 130$. So the step-length needs to be carefully calibrated. If the step factor μ is too small, convergence speed will be too slow; but if step factor is too large, the performance will be easily in disorders.

According to the convergence principle of RLS algorithm, forgetting factor λ must be $0 < \lambda < 1$. Make $\lambda = 0.99$ and $\lambda = 0.4$, the direction diagram and convergence curve can be gained as Fig. 172.3.

The Fig. 172.3 shows that the algorithm has good convergence performance and numerical stability if λ is bigger. If λ is too small, it causes bad convergence performance and numerical instability.

172.5 Design for Digital Beam Forming System

Digital beam forming module is usually made of two parts: one is optimal weights forming network with the key of digital signal processor (DSP) and adaptive algorithm; another is adaptive beam forming network constituted of dynamic

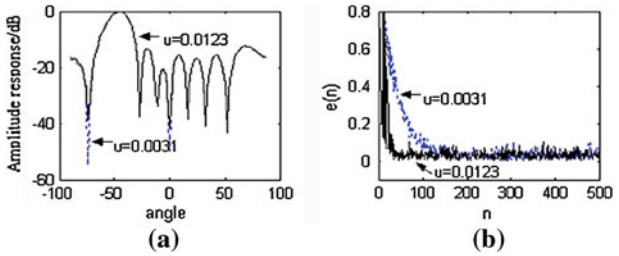


Fig. 172.2 Pattern and convergent curve of LMS algorithm. **a** Pattern, **b** convergent curve

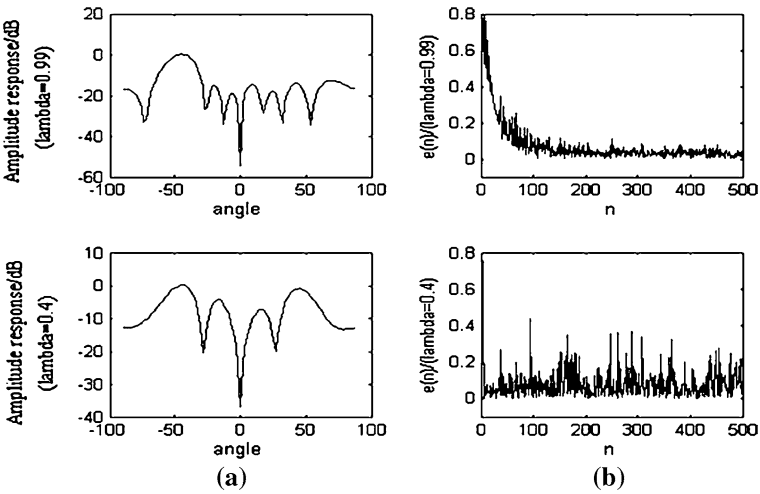


Fig. 172.3 Pattern and convergent curve of LMS algorithm. **a** Pattern, **b** convergent curve

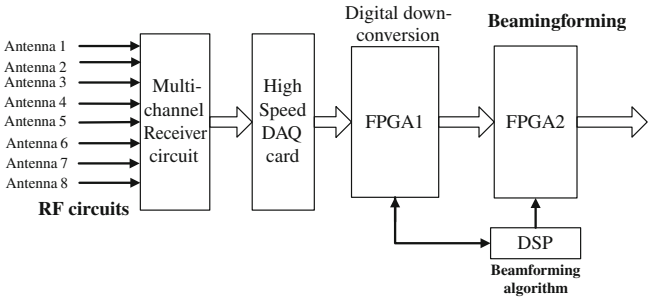


Fig. 172.4 Beam forming system diagram

adaptive weighting network. The circuit is as shown in Fig. 172.4: Use the FPGA to finish the work that is time-consuming but proper for parallel computing and use the DSP to finish the comparative complicated calculation [6].

A radio wave from antenna can't directly enter Data Acquisition Card (DAQ Card) because of the interference and noise it brings. So it needs the first handle from receiver.

Receiver usually adopts the way of super heterodyne to achieve down-conversion and filtering of received signal of arrays and enlarge signals to the standards that A/D circuits needs. Then the radio wave is sent to the DAQ Card to acquire. The signal bandwidth is wide. So the high-speed DAQ Card PCI-9820 with 30 MHz bandwidth made by ADLINK Technology is adopted [7].

This design adopts FPGA1 to devise digital down-conversion. FPGA uses EP2C35 which is the series of Cyclone from Altera Company. EP2C35 can totally provide the RAM with a large memory of 483,840 bit, logical unit is 33,216 and more than 300 pins that is available for users. DSP adopts TMS320C6701 from TI Company. TMS320C6701 is a high performance float-point DSP with 32 bits. It adopts the structure of Ve loci TI VLIW. One instruction word composed 8 commands with 32 bits. TMS320C6701 can concurrently execute 8 commands in a clock cycle. Clock frequency is 167 MHz. The capacity of peak operation achieves 1,336 MIPS. For single precision arithmetic it can reach 1G FLOPS and for double precision arithmetic it can attain 250 M FLOPS. ROM with 64 kbit and RAM with 64 kbit are integrated on chip. DSP sends the processed data to the final beam forming unit, and FPGA2 takes the task of full array beam forming.

172.6 Conclusion

Comparing the interference canceling ability of two smart antenna non-blind beam forming algorithms with different step and forgetting factors. LMS algorithm is easy to realize and has stable convergence, but convergence speed is slow. If the step factor μ is too small, convergence speed will be too slow; if step factor is too large, the performance will be easily in disorder. In order to obtain fast convergence, initial step is bigger. Then diminish μ to reduce noise and misadjustment when approaching optimal solution. RLS algorithm has fast convergence speed, but convergence performance is bad and has numerical instability. It would have better convergence performance and numerical stability if λ is bigger. The paper also puts forward the implementation method of digital beam forming based on FPGA + DSP, then gives the design of hardware circuit.

Acknowledgments This work is supported by the foundation of The National High-Tech Research and Development Plan of China (Grand No. 2008AA01Z103) Doctor foundation of henna polytechnic university(No: B2009-63)

References

1. Yang S, Ma H et al (2006) Advanced antenna technique [M]. Press of Harbin engineering university, Harbin, pp 258–272
2. Chryssomallis M (2000) Smart Antennas [J]. IEEE Antennas Propag Mag 42(3):132–134
3. Zhu Z, Lv J, et al. (2006) Study and application of digital beam forming in radar [J]. J Chin Acad Electron Inf Technol (3)
4. Gu Y, Tang K (2003) Optimal variable step-size least-mean-square model and practical algorithm [J]. J Tsinghua Univ (Nat Sci Edn) 43(1):9–15
5. Hajian M, van de Kasteele N, Ligthart LP (2002–2004) Minimum-span constant modulus array for a smart antenna testbed [J]. Vision, Image Signal Process, IEEE Proc 149(2):120–127
6. Todd WN, James S (2002) Digital beam forming and calibration for smart antennas using real time FPGA processing [J]. IEEE MTT-S Digest 56(5):307–310
7. Zhang X, Zheng B (2000) Signal processing for communications [M]. China National Defence Industry Press, Beijing

Chapter 173
Performance Analysis of 10 Gbit/s Optical Orthogonal Frequency Division Multiplexing (OFDM) Signal Transmission System Over Multimode Fiber

Chunbo Jiang

Abstract Performance of 10 Gbit/s optical orthogonal frequency division multiplexing (OFDM) signal transmission system over multimode fiber (MMF) is researched. It is simulated that the constellation diagram change with transmission, chromatic dispersion, different Multi-quadrature amplitude modulation (MQAM) modulation mode. It is shown that dispersion is the main factor to limit the transmission distance. As a number of subcarriers are related to dispersion and bandwidth utilization, a chromatic dispersion compromise needs to be considered.

Keywords Optical orthogonal frequency division multiplexing (OFDM) · Multimode fiber (MMF) · Multi-quadrature amplitude modulation (MQAM)

173.1 Introduction

Multimode fiber (MMF) is widely used in short distance optical communications for its advantages, for example, high efficiency on coupling with laser source, low system cost convenience for fiber connecting operation, and convenience for fiber connecting operation [1]. It is becoming increasingly attractive for high-speed short-reach application scenarios such as storage networks. However, the mode dispersion in MMF will cause the inter-symbol interference (ISI), which restricts the bandwidth and transmission capacity. In order to solve this problem, optical

C. Jiang (✉)
Potevio Institute of Technology, No. 6 Haidian North,
Second Street Haidian District, Beijing 100080,
People's Republic of China
e-mail: bocoj@cpit.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_173,

orthogonal frequency division multiplexing (OFDM) technology is applied in transmission system with MMF [2].

Optical OFDM has attracted extensive interests worldwide for its advantages such as providing a cost-effective, high-speed, "future-proof" technical solution due to full utilization of the rapid advances in modern digital signal processing (DSP) technology, inherent resistance to linear system impairments, tolerance to imperfect components and system. When optical OFDM is used in optical system over MMF, it is an effective solution to ISI caused by a dispersive channel.

Recently, many technologies for optical OFDM over MMF has been proposed and researched such as coherent optical OFDM (CO-OFDM) [3] and direct-detection optical OFDM (DDO-OFDM) over MMF [4, 5], coded-OFDM [6], adaptively modulated optical OFDM [7], and so on.

In this paper, performance of optical OFDM signal transmission system over MMF is investigated. It is simulated and analyzed that the constellation diagram of the optical OFDM system change with transmission, chromatic dispersion and different Multi-quadrature amplitude modulation (MQAM) modulation mode.

173.2 Scheme of Optical OFDM Transmission System Over MMF

The block diagram of the typical optical OFDM transmission system over MMF is shown in Fig. 173.1, which consists of an OFDM signal modulation, an optical signal sender, an optical transmitter over MMF, an optical signal receiver and an OFDM signal demodulation.

At the OFDM signal modulation, an incoming binary data sequence is divided into N parallel sub-carrier data streams. Each sub-carrier is then converted into complex values according to a MQAM or multiphase shift keying (MPSK) mapping. The MQAM or MPSK mapping is chosen according to the frequency response and noise of the transmission link at this particular frequency. In order to obtain the real value time-domain waveform for N sub-carriers at the output of the Inverse Fast Fourier Transform (IFFT), the length of transform has to be $2N$ points with the condition that the data in the second half of the $2N$ IFFT points has to be the complex conjugate of the data in the first half. Then, the output of IFFT is serialized using a parallel to serial converter (P/S). After added cyclic prefix (CP), the signal is converted to an analogue signal for transmission using a digital to analogue converter (DAC). The DAC output is, in general, complex and is applied to an IQ modulator for transmission over the channel, which is named as RF up-convert.

In the receiver, the optical signal emerging from the MMF link is detected by a photoelectric detector (PD). After RF down-convert, the inverse of the sender with samples of the incoming waveform is captured using an analogue to digital converter (ADC).The CP is removed by RCP. The samples are aligned into symbols after the serial to parallel converter (S/P) and then processed using a fast Fourier

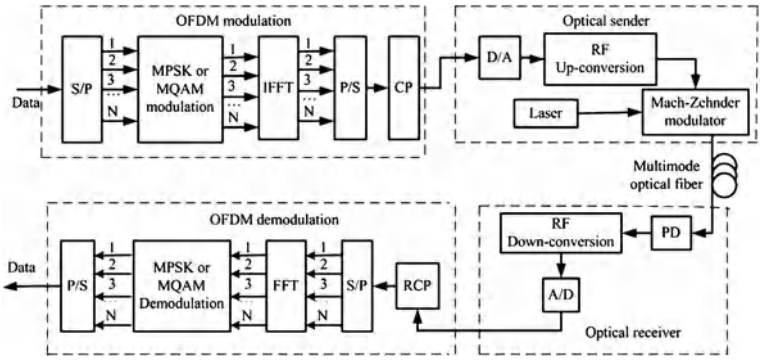


Fig. 173.1 Optical OFDM transmission system based MMF

transform (FFT). The output of FFT is demodulated by corresponding MQAM or MPSK demodulation. The original data sequence is obtained after P/S converter.

173.3 Performance Analysis of 10 Gbit/s Optical OFDM System Over MMF

In order to analyze the performance of 10 Gbit/s optical OFDM system over MMF, the simulation system is set up. The length of pseudo-random bit sequence generator is $2^{16}-1$. Frequency of IQ modulation is 7.5 GHz. Output power of optical laser power with wavelength 1,550 nm is 0 dB m, which is modulated by $LiNbO_3$ Mach–Zehnder Modulator. The MMF attenuation is 2.61 dB/km.

173.3.1 Performance Against Different Modulation

In this section, the performance changing with the different QAM modulation is described.

On the condition of 90 dB of SNR, 10 km of transmission distance, 2.61 dB/km of MMF attenuation and 20 ps/(nm km) of dispersion, the constellation diagram with 4 and 16QAM is simulated.

Figure 173.2 shows the constellation diagram with different QAM modulation at the same SNR, transmission distance and MMF. From Fig. 173.2, it can be seen that the performance of 4QAM is better than that of 16QAM. This is because that the spacing of constellation dot of 4QAM is bigger than that of 16QAM. The noise tolerance of lower step QAM is better than that of higher step QAM. So the higher step QAM is affected easily by noise signal at the receiver. But better symbol coding efficiency ratio and better bandwidth utilization ratio are obtained in higher

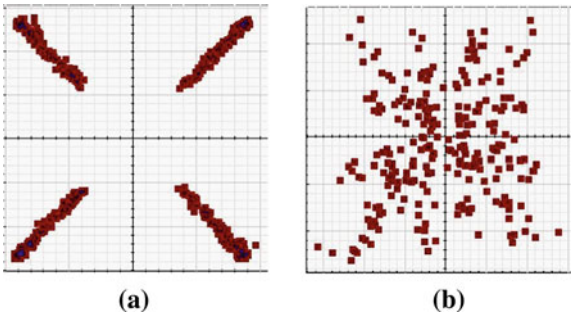


Fig. 173.2 Constellation diagram with 4 and 16QAM at the same SNR, transmission distance and MMF. **a** 4QAM **b** 16QAM

QAM system. Thus, higher QAM modulation should be chosen if this modulation mode can meet the system demand at the same SNR.

173.3.2 Chromatic Dispersion Tolerance

Chromatic dispersion emerges when optical pulse with different frequencies transmits in optical fiber. This is because different frequency optical wavelength has different transmission rate. Chromatic dispersion in MMF is more than that of signal mode fiber. In this section, the relation between chromatic dispersion coefficient and transmission distance is studied through simulation in 10 Gbit/s 4QAM optical OFDM system with 256 sub-carriers.

Figure 173.3 shows the constellation diagram with different chromatic dispersion coefficient in 4QAM optical OFDM system after 10 km MMF transmission. When chromatic dispersion coefficient is 5 ps/(nm km), constellation diagram is best. When chromatic dispersion coefficient increases, the constellation diagram degrades. When chromatic dispersion coefficient is 60 ps/(nm km), four part cannot be separate from constellation diagram. It can be seen that constellation diagram is degraded with the increase of chromatic dispersion coefficient. This is because orthogonal character between sub-carrier is destroyed at receiver for receiving signal wave shape distortion due to chromatic dispersion of optical OFDM in MMF. Then the ISI and inter-carrier interference (ICI) at receiver increase and bit error ratio degrades.

Figure 173.4 shows the constellation diagram with different chromatic dispersion coefficient in 16QAM optical OFDM system after 5 km MMF transmission. When chromatic dispersion coefficient is 20 ps/(nm km), constellation diagram is best. When chromatic dispersion coefficient increases, the constellation diagram degrades. When chromatic dispersion coefficient is 60 ps/(nm km), the constellation diagram is distorted completely. It can be seen that constellation diagram is degraded with the increase of chromatic dispersion coefficient. The bit error rate is degraded due to phase distortion coming from the phase noise, which is produced by phase wide increase for chromatic dispersion.

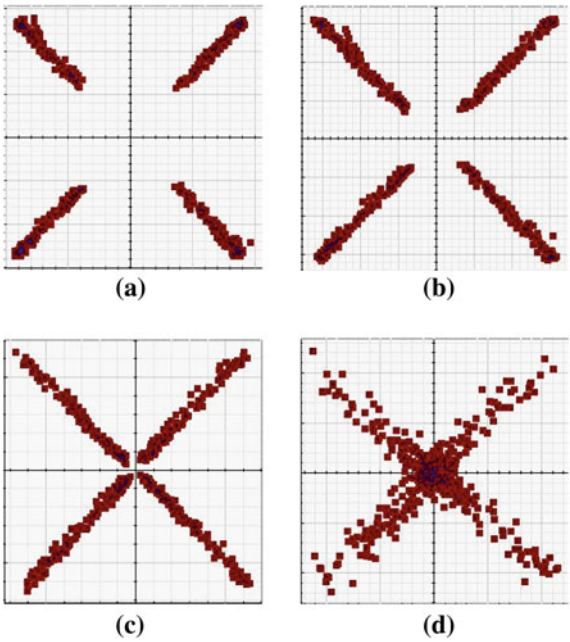


Fig. 173.3 Constellation diagram with different chromatic dispersion coefficient in 4QAM optical OFDM system after 10 km MMF transmission. **a** 5 ps/(nm km), **b** 20 ps/(nm km), **c** 40 ps/(nm km), **d** 60 ps/(nm km)

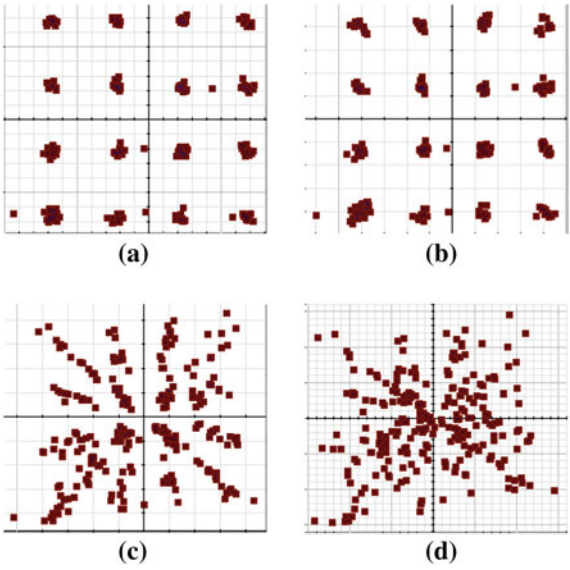


Fig. 173.4 Constellation diagram with different chromatic dispersion coefficient in 4QAM optical OFDM system after 5 km MMF transmission. **a** 20 ps/(nm km), **b** 40 ps/(nm km), **c** 60 ps/(nm km), **d** 70 ps/(nm km)

Comparing Fig. 173.3 with Fig. 173.4, it can be seen that lower step QAM has longer transmission distance than higher step QAM at the same chromatic dispersion coefficient.

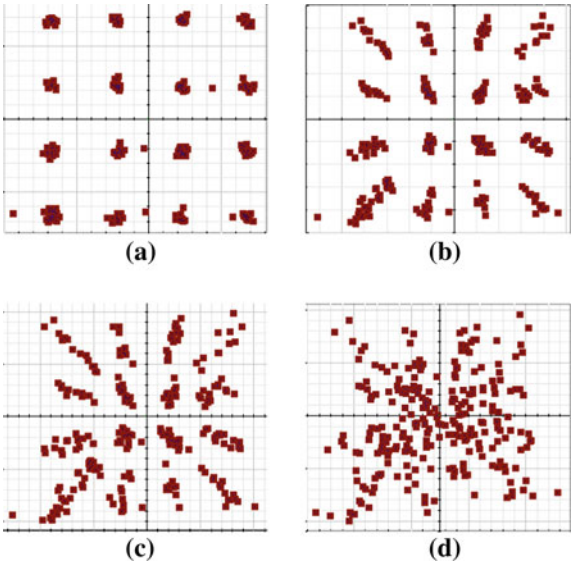


Fig. 173.5 Constellation diagram with different transmission distance in 16QAM optical OFDM system. **a** 5 km, **b** 7 km, **c** 8 km, **d** 9 km

Moreover, in order to improve the performance of system, chromatic dispersion must be compensated during optical transmission or electronic transmission.

173.3.3 Performance Changing with Transmission Distance

In this section, performance of optical OFDM system over MMF changing with transmission distance is simulated and analyzed. Figure 173.5 shows the constellation diagram with different transmission distance in 16QAM optical OFDM system based on 20 ps/(nm km) of chromatic dispersion coefficient and 2.61 dB/km of MMF attenuation. From Fig. 173.5, it can be seen that constellation diagram is degraded rapidly with 1 km increase of transmission distance. This is because that effect of chromatic dispersion for performance is more than that of other parameter such as number of sub-carrier. Thus, chromatic dispersion is the main factor of limiting transmission distance for optical OFDM system over MMF with larger chromatic dispersion and no chromatic dispersion compensate.

173.4 Conclusion

Performance of 10 Gbit/s optical OFDM signal transmission system over MMF is simulated and analyzed. It can be seen that the performance of 4QAM is better than that of 16QAM. The noise tolerance of lower step QAM is better than that of

higher step QAM. Constellation diagram is degraded with the increase of chromatic dispersion coefficient. Lower step QAM has longer transmission distance than higher step QAM at the same chromatic dispersion coefficient. Constellation diagram is degraded rapidly with 1 km increase of transmission distance. Moreover, dispersion is the main factor to limit the transmission distance. For number of subcarriers is related to dispersion and bandwidth utilization, chromatic dispersion must be compensated during optical transmission or electronic transmission in order to improve the performance of system.

References

1. Cunningham D (2008) Multimode fiber data communication. OFC/NFOEC, San Diego, CA, OWR1
2. Jolley NE, Kee H, Rickard R, Tang J, Cordina K (2005) Generation and propagation of a 1550 nm 10 Gbit/s optical orthogonal frequency division multiplexed signal over 1000 m of multimode fibre using a directly modulated DFB. OFC/NFOEC, Anaheim, CA, OFP3
3. Tang JM, Shore KA (2007) Maximizing the transmission performance of adaptively modulated optical OFDM signals in multimode-fiber links by optimizing analog-to-digital converters. J Lightwave Technol 25(3):787–798
4. Tong Z, Yang Q, Ma Y, Shieh W (2008) 21.4 Gbit\s transmission over 200 km multimode fiber using coherent optical OFDM. Electron Lett 44(23):1373–1374
5. Schich W (2009) High spectral efficiency coherent optical OFDM for 1 Tb/s ethernet transport. OSA/NFOEC, OWW1
6. Djordjevic IB (2010) Coded-orthogonal frequency division multiplexing in hybrid optical networks. IET Optoelectron 4(1):17–28
7. Giacoumidis E, Jin XQ, Tsokanos A, Tang JM (2010) Statistical performance comparisons of optical OFDM adaptive loading algorithms in multimode fiber-based transmission system. J IEEE Photonics 2(6):1051–1058

Chapter 174
Research of Botnet Detection Based on Multi-Stage Classifier

Xin Liang Wang, Nan Lu and CuiCui Wang

Abstract The botnet represents a growing threat to the network security and the variants that emerge endlessly. Although the known anomaly detection algorithms can detect the unknown botnet to some degree, the generalization and the accuracy of the models remain to be improved. Whereas in this paper we proposed a botnet anomaly detection structure based on multi-level classifier, the first level classifier proposes a new periodical communication detection method. Compared with spectrum analysis, the algorithm has a lower complexity and can make online and real-time detection of the botnet; the second level classifier establishes the decision tree using the statistical characteristics of the IP pairs of the periodical communication. The result of the experiment indicates that compared with the algorithms only adopting the periodical communication or that based on the flow statistical characteristics, the multi-level classifier model not only has the higher generalization but also the higher accuracy in detecting the unknown botnet.

Keywords Multi-level classifier · Periodicit · Statistical characteristics of flow · Botnet

X. L. Wang (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Kaifeng, China
e-mail: wangxinliang@bupt.edu.cn

N. Lu · C. Wang
School of Information and Communication Engineering,
Beijing University of Posts and Telecommunications,
Beijing, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_174,

174.1 Introduction

Recently, botnet is the major reason for the security lacks such as the leakage of the privacy and financial information, the pilferage of the network bank passwords, the overflowing of the spam as well as the DDOS attack and so on. The botnet not only caused losses to the internet users, but also harmed the interest of the corporations and even did harm to the information security of the states and the military. At present, the known botnet can be detected purely depending on the deep packet inspection algorithms, but it cannot detect the unknown botnet. As for the anomaly detection of the botnet based on the flow statistical characteristics [1, 2], although it can detect the unknown botnet, the model designed only can detect a nominal sum of botnets from a huge mass of the network data and the detection accuracy and the generalization ability can hardly be assured. Meanwhile, the experiment results of the document [2] indicates that simply choosing the characteristics of some flow data and using the machine learning to take the anomaly detection of the botnet cannot achieve good results.

As for the anomaly detection [3, 4] of the botnet based on the periodical communication behaviors, although the method using the discrete-time sequence analysis is effective to some degree in detecting the botnet with periodical communication behaviors, the botnet's behaviors of keeping up a correspondence may not be exactly periodical. When the communication intervals of the botnet randomly changes, the spectrum components will become complex and the detection accuracy of the algorithms will be affected. At the same time, the periodical communication behaviors exist in the normal network flows, such as QQ, Skype as well as 360 Anti-virus and so on, the communication interval of which could be stable or could randomly change. The known detection methods cannot solve this problem well. Moreover, the temporal complication and the special complication of the detection algorithm based on the spectrum analysis is too high and is not suited for the online and real-time botnet detection.

The anomaly detection of the botnet based on the attack behaviors [5, 6], detecting the botnet on the basis of the behaviors such as scanning, spam as well as the DDOS attack and so on, has achieved good results. But this method takes effect only when the attacks occur and the controller of the botnet can avoid the detection of this method through reduction of the frequency of the scanning and transmitting the spam.

In view of the above methods having both advantages and disadvantages, in this paper, a new algorithm of detecting periodical communication is proposed. Compared with the spectrum analysis algorithm, it has the advantage in low complexity and real-time detection. On that basis, it detects the botnet combined the detection method which is based on the periodical communication with the detection method that based on the statistical characteristics of IP, which improves the accuracy of he detection and can detect the unknown botnet.

174.2 Data Acquisition

174.2.1 Offline Sample Flow Acquisition

Some kinds of botnet programs were downloaded from some hacker forums. The downloaded samples run under the condition of offline network and the flows were captured. The offline network is a small-scale LAN consisting of three hosts. Host 1 serves as the control server of the botnet, while hosts 2 and 3 install the virtual machine with three kinds of operation systems including windows XP, windows server and windows server 2000. So hosts 2 and 3 are both equipped with three operation systems. So the botnet experiment environment with one controlling host and 6 controlled hosts is simulated.

174.2.2 Capture the Online Sample Flows

Obtain the latest malicious IP addresses from the noted abroad web sites and then take advantage of the deployed network flows monitoring equipments to acquisition the raw flows.

174.2.3 The Data Set

The data is divided into 10 groups: Trace 1 ~ Trace 10. The Trace 1 ~ Trace 4 are training model sets which belong to data sets 1 and Trace 5 ~ Trace 10 are served as testing model sets, belonging to the data set 2. Trace 1 ~ Trace 3 and Trace 5 ~ Trace 7 are normal network flows captured by 6 personal machines through long term. While Trace 4, Trace 8 and Trace 9 are anomaly flows obtained through long-term data acquisition of 12 kinds of botnets (including win32.Hack.Drwolf.nd, Win32.Hack.Ceckno and so on) in the offline network. Trace 4 is the flow of six kinds of botnets and Traces 8 and 9 contain the flows of another six kinds. While Trace 10 is the anomaly flows of the 13th kind of botnet captured in the online network.

174.3 The Detection Frame of the Multilevel Classifier

174.3.1 Detection Frame

Definition 1 Strictly periodical communication behavior is IP pairs communicating with fixed intervals (Fig. 174.1).

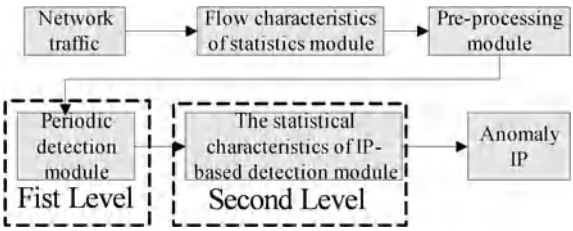


Fig. 174.1 The detection frame of the multi-level classifier

Definition 2 Periodical communication is IP pairs communicating with the interval T and T randomly changing in a certain range.

The characteristics of the periodical communication are based on the botnet. This paper proposes the detection frame of the multilevel classifier, which consists of statistical module of the flow characteristics, preprocessing module, periodical communication detection module as well as the botnet detection module based on the statistical characteristics of the IP pairs. The flow characteristics statistical module and preprocessing module process the network flows and the output is the valid flows. The periodical communication module obtains the IP pairs with periodical communication based on the valid flows and then it takes advantage of the statistical characteristics of the IP pairs to build the decision tree to detecting the anomaly traffic of the botnet, which takes good effect.

174.3.2 Flow Characteristics Statistical Module

This module can make statistical analysis of the network flows and it can draw several characteristics of the data flows. The details of the characteristics are shown in Table 174.1.

174.3.3 Preprocessing Module

In the normal network traffic, there are some packets only with SYN flag. Through the analysis of the captured packets, these kinds of packets are produced by repeatedly connecting requests when users of the QQ music players choose songs that are not available. When the botnet servers are shut down by the network security administrator, the bots in the network still send numerous SYN connecting packets, producing amount of data flows of the SYN packets. Because of the sophisticated causes of the SYN packets this paper focuses on the botnet detection based on the periodical communication, where the SYN packets were filtered by the preprocess module.

This paper proposes the detection based on the periodical traffic of the botnet, so the captured anomaly flows mainly contain the keep-alive packets while the

Table 174.1 Statistical characteristics

Index	Meaning	Index	Meaning
Sumpayloadlen	The sum of the payload length	sumpktlen	The sum of the packet length
Sumpayloadlenup	The sum of the payload length in the upstream packets	sumpktlenup	The sum of upstream packet lengths
sumpayloadlendown	The sum of the payload length of the downstream packets	sumpktlendown	The sum of downstream packet length
avepktlen	The average length of the packet	packetnum	The number of packets
maxpayloadlen	The maximum length of the payload of the packet	pktnumup	The number of upstream packets
minpayloadlen	The minimum length of the payload of the packet	pktnumdown	The number of downstream packets

flows of the other functions (such as reporting the system information, issuing the controlling command etc.) are mostly neglected.

174.3.4 Periodical Communication Detection Module

The detection of the characteristics of the periodical communication behavior is proposed because communication intervals between the bot hosts and the controlling servers are not the same and they can change randomly. The periodical communication detection purely depending on the spectrum analysis cannot meet the command of the real network. Moreover, the spectrum analysis has high time and space complexity, which is not suited for the high-speed link network environment.

This paper is on the basis of the sequential hypothesis testing algorithm [7] and through improving this algorithm. It implemented the real-time detection of the periodical communication behavior according to the real network. The detailed procedure of the algorithm is shown as follows: the flows of the IP pairs are equivalent as connecting sequences and then the sequences are tested to implement the detection of periodicity.

Definition 3 Connection. Suppose the connection is confirmed by the number of flows in a fixed interval and the connection is represented as Y, and then Y_i represents whether the communication of the IP pairs happened. Besides,

$$Y_i = \begin{cases} 0 & \text{if the connection of flow numbers } > 0, \text{ exist communication} \\ 1 & \text{if the connection of flow numbers } = 0, \text{ not exist communication} \end{cases}$$

Definition 4 The judge unity is the IP pairs remaining to be judged. The judge unity is expressed as H_0 and H_1. H_0 represents the set of IP pairs with periodical communication behaviors and H_1 represents the set of IP pairs with no periodical communication.

The random variable Y_i has four potential values shown as follows:

$$P_r[Y_i = 0/H_0] = \theta_0 \, P_r[Y_i = 1/H_0] = 1 - \theta_0$$

$$P_r[Y_i = 0/H_1] = \theta_1 \, P_r[Y_i = 1/H_1] = 1 - \theta_1$$

To ensure the accuracy of the sequential hypothesis testing algorithm, it is of vital importance to define the statistical behavior of H_0 and H_1 accurately. If an IP pair with periodical communicating behaviors is fixed, the possibility of communicating during a certain interval will be far greater than the possibility of no communication. So the value of θ_0 is considerably large.

As for the IP pairs with no periodical communicating behaviors, the possibility of there occurrence during a certain interval is comparatively much lower and the value of θ_1 is smaller.

If the connection Y_i happens, the sequential hypothesis testing algorithm will update the likelihood ratio $\Lambda(Y) = \prod_{i=1}^{n} \frac{P_r[Y_i/H_1]}{P_r[Y_i/H_0]}$.

As the continuation the IP communicating behaviors, $\Lambda(Y)$ will update continuously. And comparing the updated $\Lambda(Y)$ with the threshold η_0 and η_1. If $\Lambda(Y) > \eta_1$, and then it can be decided whether this IP pairs belongs to the H_1 set. If $\Lambda(Y) < \eta_0$, this IP pair belongs to H_0. In this paper we assumed that $\theta_0 = 0.75, \theta_1 = 0.25, \eta_0 = 0.004, \eta_1 = 250$.

As for the bots with fixed communicating period, this algorithm can make fast detection and it can detect bots with random communicating period. Assuming that the intervals of bots sending packets change in the range of 5 s and 5 min, the spectrum analysis algorithm cannot detect the periodical communicating behavior of the botnet. The sequential hypothesis testing algorithm will assume the connecting interval as 5 min. After limited fixed intervals, the hosts with strictly periodical communicating behaviors corresponds the inequality $\Lambda(Y) < \eta_0$. While the hosts have no strict periodical communication, if the random interval is in the assumed range, it can be decided as bots with periodical communication.

Definition 5 False positive = the number of the normal IP that is falsely judged as anomaly IP/the number of the anomaly IP.

Definition 6 False negative = the number the anomaly IP that is falsely judged as the normal IP/the number of the anomaly IP.

The result of periodical communication module is shown in Table 174.2. The experiment result indicates that the periodical communicating detection module has high accuracy in anomaly detection. But the false positive is 70.6% and the false negative is 169%. Therefore, although the improved periodical communicating

Table 174.2 The experiment result of the detection algorithm of the periodical communicating detection module

Data set	False positive (%)	False negative
Data set 1	70.6	0
Data set 2	169	0

detection module has low complexity and it implemented the real-time and online botnet detection, the detection accuracy remains to be improved.

174.3.5 The Botnet Detection Module Based on the Statistical Characteristics of the IP Pairs

The detection of the characteristics of the periodical communication behavior is proposed because communication intervals between the bot hosts and the controlling servers are not the same and they can change randomly. The periodical communication detection purely depends on the spectrum analysis which cannot meet the command of the real network. Moreover, the spectrum analysis has high time and space complexity, which is not suit for the high-speed link network environment.

Through the first level periodical communicating module, the IP pairs with periodical communicating behaviors can be obtained, which include the IP pairs of the botnet with periodical communication and normal IP pairs with periodical communication. The later IP pairs are almost produced by the messaging software (such as QQ, Skype and 360 Anti-virus etc.). These kinds of IP pairs will affect the accuracy of the botnet detection. So it is necessary to make further classification with the second level classifier.

Definition 7 The statistical characteristic of the connection. In this paper, the connection is during the interval i, and the corresponding flows are expressed as $f_1, f_2, \ldots, f_n$.

$$\text{sumpayloadlen}(Y_i) = \sum_{k=1}^{n} \text{sumpayloadlen}(f_k), \quad \text{avepktlen}(Y_i) = (\sum_{k=1}^{n} \text{avepktlen}(f_k))/n$$

$$\text{maxpayloadlen}(Y_i) = \max(\text{maxpayloadlen}(f_1), \ldots, \text{maxpayloadlen}(f_n)), \quad \text{minpayloadlen}(Y_i) = \min(\text{minpayloadlen}(f_1), \ldots, \text{minpayloadlen}(f_n))$$

The computing method of the other corresponding statistical characteristics of the connection is the same as $\text{sumpayloadlen}(Y_i)$.

Definition 8 The statistical characteristics of the IP pairs. The statistical characteristics of the IP pairs are calculated based on the connection statistical

Table 174.3 The experiment result of the botnet detection module is based on the IP pairs of statistical characteristics

Data set	Algorithms	False positive	False negative
data set1	J48	0	0
data set1	BayesNet	0	0
data set1	Naïve Bayes	23.5%	0
data set2	J48	6.5%	0
data set2	BayesNet	6.5%	35.5%
data set2	Naïve Bayes	41.9%	0

characteristics. Assuming that during the judgment of the first level classifier, the corresponding connections of the IP pairs with periodical communications are expressed as IP pairs are expressed as:

$$\text{sumpayloadlen}(I) = (\sum_{i=1}^{n} \text{sumpayloadlen}(Y_i))/n \qquad \text{maxpayloadlen}(I) = \max(\text{maxpayloadlen}(Y_1), \ldots, \text{maxpayloadlen}(Y_n))$$

$$\text{minpayloadlen}(I) = \min(\text{minpayloadlen}(Y_1), \ldots, \text{minpayloadlen}(Y_n))$$

The computing method of other statistical characteristics of the IP pairs is the same as sumpayloadlen(I).

This paper takes advantage of the C45 decision tree algorithm (J48) provided by the weak platform, bytes network, the naïve Bayes algorithm to implement the detection model and compares and analyzes the performance of multiple models. The detection results of the model is shown in Table 174.3. Data set 1 serves as the training model and data set 2 serves as the testing set to detect the accuracy and generalization ability of the training model. The results showed that the accuracy of J48 is larger than the ByesNet and NaiveByes. As for the seven kinds of unknown botnets, for achieving a comparatively high detection accuracy, the second level classifier J48 is chosen to implement the model.

174.3.6 The Botnet Detection Module Based on the Statistical Characteristics of the IP Pairs

The detecting result of the multilevel classifier is shown in Table 174.4. It can be known that the false positive and the false negative are 6.5% and 0. Compared with the detection purely on the periodical communicating detection module, the index of false positive has been improved by a large scale. If only using the flow statistical characteristics based on the J48 decision tree algorithm to implement the anomaly detection of the botnet, although the detection result of the training model is good, when detecting the training model with the data set2, 40407 of the 75059

Table 174.4 The result of the multilevel classifier detection algorithm

Data set	False positive	False negative
Data set1 (train)	0	0
Data set2 (testing)	6.5%	0

anomaly flows are false negative alarms, the case of false negative alarms is severe. It can be concluded that the accuracy of detecting the unknown the botnet using the model implemented with flow statistical characteristics remains to be improved. As for the multilevel classifier model using six kinds of botnets, although there remains a few wrongly judged, it can detect seven unknown botnets accurately, which indicates that this model is equipped with the ability to detect the unknown botnet and has fine generalization ability.

174.4 Conclusions and the Future Study Trends

174.4.1 Conclusions

This paper implemented the multilevel classifier model based on the periodical communicating behavior to detect the botnet.

The first level completed the detection of the botnet with the strictly periodical and periodical to be communicating behaviors, providing foundations for the second level to detect the statistical characteristics of the IP pairs. It can be known from Table 174.2 that, despite the periodical communicating detection module having high accuracy in anomaly detection, which is as high as 100%, the amount of falsely judging the normal IP as anomaly IP is large, which affects the performance of the periodical communicating detection module. Based on the IP pairs with periodical communication, the second level module of detecting IP pair statistical characteristics takes advantage of multiple machine learning algorithms to implement the IP pairs statistical model on the basis of the data of six kinds of botnets and the normal traffic of several personal computers. Besides, it takes the anomaly data of another seven kinds of botnets and the data of several personal computers as the testing data. The experiment result shows that the generalization ability and the detecting performance of the model implemented with decision tree algorithm are significantly better than other machine learning algorithms. Therefore, the second level classifier completed the botnet detection using the statistical model implemented with the decision tree model. It can be known from Table 174.4, compared with the algorithm based on the periodical communicating detection, the detecting structure based on the multilevel classifiers not only has low time and space complexity and high generalization ability, but also has high accuracy as to the unknown botnet detection.

174.4.2 Future Study Trends

Although the model is equipped with the ability of detecting the unknown botnet, the amount of botnet samples is small. Therefore, more botnet samples need to be obtained to improve the statistical model of the botnet IP pairs further. Furthermore, the detection result of the multilevel classifier model remains in some cases that a few normal IP pairs are wrongly judged as anomaly traffic. Despite the amount being small, how to improve the detection accuracy further needs to be studied.

References

1. Wang T, Yu S-Z (2009) Centralized botnet detection by traffic aggregation [C]. 2009 IEEE International Symposium on Parallel and Distributed Processing with Applications, United States, pp 86–93
2. Livadas C, Walsh R, Lapsley D, Strayer WT (2006) Using machine learning techniques to identify botnet traffic [C]. In 2nd IEEE LCN Workshop on Network Security (WoNS'2006), pp 967–974
3. AsSadhan B, Moura JMF, Lapsley D, Jones C, Strayer WT (2009) Detecting botnets using command and control traffic [C]. Eighth IEEE International Symposium on Network Computing and Applications, pp 156–162
4. Chen L, Wang X, Zhao X, Li W (2010) Research of botnet anomaly detection alogrithm based on private protocol [C]. Proceedings of 2010 3rd IEEE International Conference on Broadband Network and Multimedia Technology, pp 55–59
5. Gu G, Zhang J, Lee W (2008) BotSniffer: detecting botnet command and control channels in network traffic [C]. In Proceedings of the 15th Annual Network and Distributed System Security Symposium (NDSS'08)
6. Gu G, Perdisci R, Zhang J, Lee W (2008) BotMiner: clustering analysis of network traffic for protocol- and structure-independent botnet detection [C]. In Proceedings of the USENIX Security Symposium (Security), pp 139–154
7. Wald A (1947) Sequential analysis[M]. Wiley, New York

Chapter 175
Network Topology Model and Fault Analysis for Electrical Control Systems

Yuan Haibin

Abstract The safety analysis and fault diagnosis is an important issue for the maintenance process of large complex industrial system such as electrical control system, and the integrated system is required to be taken into account when all the effect factors are included. Based on the above consideration, the topology description using complex network is introduced in this paper, the nodes represent the components in the electrical control system and the edges represent the relationship between components. The characteristics of the model are analyzed from the complex networks viewpoint, and fault analysis and propagation inference process is given based on the proposed network topology. Application study is given to validate the feasibility of the proposed method.

Keywords Electrical system · Topology · Fault diagnosis · Complex networks

175.1 Introduction

Electrical devices such as relays, circuit breakers and connecters are used widely in many fields. The safety and reliability is important to make sure that the system is working under proper operation. One useful way of improving system safety and reliability is to rely on fault prognosis [1], as a result, fault knowledge base is important in processing and understanding the condition of the system.

Y. Haibin (✉)
School of Automation Science and Electrical Engineering,
Beihang University, Beijing 100191, China
e-mail: yuanhb@buaa.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_175,

Fault modeling is a useful way in evaluating working condition for a specific electrical control system, which can provide reference for activity of prediction and maintenance. As a result, the cost will decrease greatly and long working life expectation will be achieved. Theory of fault modeling is studied and applications of fault modeling are employed extensively in many fields such as chemical industry, power system, traffic control and networks [2]. In chemical industry, SDG is first used for fault analysis, and fault tree is derived from SDG modeling, with software package is developed for visualization of fault analysis [3, 4], at the same time, Petri net based fault modeling is studied in power system, moreover, interactive computer aided GPNT tool is developed as well [5, 6]. Recently, fault modeling platform for electrical control system is studied by research in order to meet the requirement of the system level prognosis health management. Electrical components fault diagnosis model is proposed based on RBFNN [7], while fuzzy comprehensive evaluation model is proposed in [8]. Furthermore, AHP method is applied to evaluate fault condition in a qualitative way [9].

In view of the research mentioned above, it can be seen that the risk analysis technology for electrical utilities of electrical control system is mainly based on computational purpose, most researches are based on the one part of the entire system and little is done for practical purpose, on the other hand, the method of fault modeling are mostly dependant on single electrical component activity rather than analysis process of system architecture. Due to the complexity in electric control system, integrated health monitoring and diagnostic system must has the ability to provide an autonomous, timely, and accurate assessment in order to dynamically identify any degradation in functional performance that may affect safety or successful field operations as well as to identify the specific subsystems that require maintenance to restore full operational capability. The difficulty in analyzing electrical control system character lies in the fact that there are many components involved with each component can interact with the others, changes in one component can affect the others within the system, the components represent sources in the system much in the same way that nodes represent sources in the complex networks.

In this paper, the topology model and fault analysis for electrical control system is proposed and studied based on complex networks. A chief advantage that a complex network is employed to represent electrical control system is that integrated system analysis result can be obtained. Also, some new inference method can be achieved to deal with fault diagnosis and evaluate system performance or operation condition. Using the state information of the components, the computational formulations based on node event probability are derived. The rest of the paper is organized as follows. The characters of the network property for electrical control system is analyzed and discussed followed by the introduction, after that, fault analysis and propagation process is discussed and algorithm is proposed, finally, result of practical fault analysis is shown and discussed in order to validate the feasibility of the proposed method, and conclusion and research direction are given as summary.

175.2 Network Topology and Parameters Description

The study of complex network is mainly based on graph theory, recently, the structure and dynamics of complex is studied extensively, and research article can be found in many literatures. Small-world network and scale-free network are two models that are studied extensively. The property of mall-network is associated with the presence of localized clusters and has been observed in many different real networks such as biological network and technological network. Watts and Strogatz [10] defined small-world networks as those networks having both a small value of average shortest path and a high clustering coefficient. Scale-free network presences the property of heterogeneous connection, in which certain fractions of vertices are highly connected. Such network displays power law shaped degree distribution $P(k) = Ak^{-\gamma}$ with exponents varying in the range between 2 and 3. Weighted network and unweighted network are two different, for unweighted network, the linkage between two nodes are present or not, for example, when the linkage is noted by 1 when two node are connected and 0 when there is no linkage between two nodes, which mean the strength is either 1 or 0, while the strength of the linkage between two nodes is assigned value for weighted network. Weighted network is preferable in dealing with the real world.

For electrical control system, the components are regarded as nodes or vertices, which represent the relays, circuit breakers, switches, generators and so on. The linkage between two components is regarded as edge or arch. Then the complex network model for the electrical control system is defined as $N = (V, E)$, such that $N \neq \phi$, where $V = \{v_1, v_2, \ldots, v_k, \ldots, v_n | k \in [1, n]\}$ represents the set of nodes, and E is ordered pairs of element from V, $E = \left\{(v_i, v_j) | v_i \in V, v_j \in V\right\}$ represents the set of edges from node i to node j. Degree of node i is noted as k_i, which means the number of edges connected to node i, and distance d_{ij} between node i and node j is defined as the shortest number of edges from node i to node j, and network distance is defined as $D = \max_{i,j} d_{ij}$, for undirected network, the mean geodesic distance between node pair in a network is:

$$l = \frac{2}{n(n+1)} \sum_{i,j \in n, i \neq j} d_{ij} \tag{175.1}$$

where n denotes the number of nodes in the network. The clustering coefficient for node v_i is defined as:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{175.2}$$

where C_i means the ratio between number of triangles connected to node i and number of triples centered on node i.

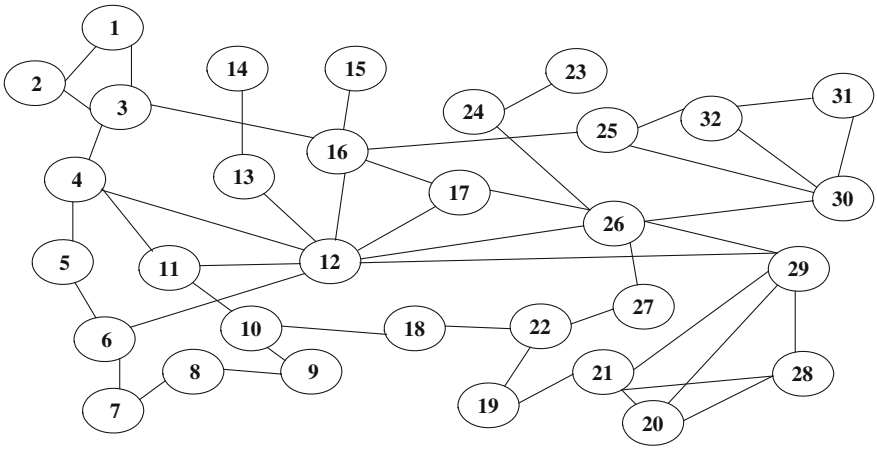
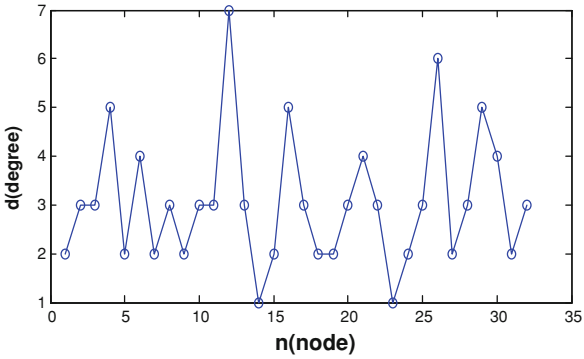


Fig. 175.1 Network topology of electrical control system

Fig. 175.2 Node degree distribution (for node n = 32)



175.3 Network-Based Fault Model and its Character

In this section, the electrical control system that is used for aerospace power supply is shown to exploit the fault model construction and fault analysis method [11], the operation function for the system is divided into three separate subsystems, with various components included in the system, based on the definition mentioned above, there are 32 nodes in the system, the network topology is shown in Fig. 175.1, the statistical character for the system is shown in Fig. 175.2, it can be seen from Fig. 175.2 that degree for node 12 is 7, followed by node 26 with 6 and node 4, node 16 and node 29 with degree of 5, which means these nodes are more likely affected by the fault influence from the other nodes, also, these node are more subjected to possible fault propagation paths than those nodes that have smaller degree, as a result, these nodes are more important in the process of the system operation. By computation from Eq. (175.2), the network average distance

Table 1 Node description in the network topology

Node	Description	Node	Description	Node	Description
1	Left generator	12	BTB	23	Right generator
2	LG control unit	13	Left EPC	24	RG bus
3	LG breaker	14	GS BUS	25	RG control unit
4	LG bus	15	EXR POWER	26	RG bus
5	LG LED	16	Bus bar power unit	27	RG RED
6	Left circuit breaker	17	Right EPC	28	SEBUS
7	Left battery	18	GB 100	29	RBC
8	Left BEC	19	Right ERBD	30	APB
9	Left ELBD	20	Right battery	31	APUG
10	Left BUS	21	Right BEC	32	APU GCU
11	Left GLED	22	Right EBUS		

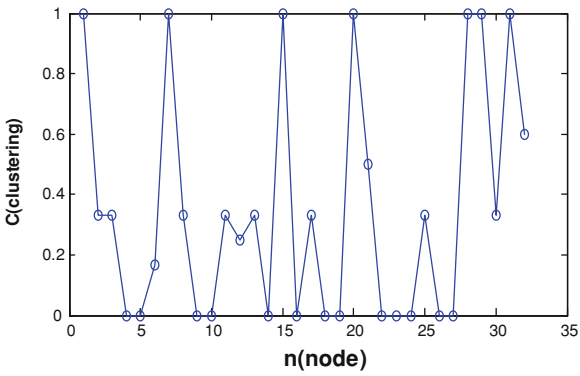


Fig. 175.3 Node clustering (for node n = 32)

is $L = 3.821$, and average clustering coefficient is $\bar{C} = 0.3125$. It can be seen from the computation result that the network for the electrical control system has small average distance and larger clustering coefficient. The topology construction in Fig. 175.1 is based on the physical component connection relationship of the aerospace power supply system, each component can be further divided into some sub nodes, and hence the network can be regarded as small-world network. (Table 175.1, Fig. 175.3)

175.4 Fault Propagation Analysis

Inference process of fault propagation is one of the important parts based on the network topology. In order to represent fault propagation path in the form of complex networks model, the relationship between two components has to be considered. Here, adjacency matrix is given to indicate the causal relationship from one node to another. Assuming $J = \{w_{ij}\}$, the element w_{ij} in the matrix

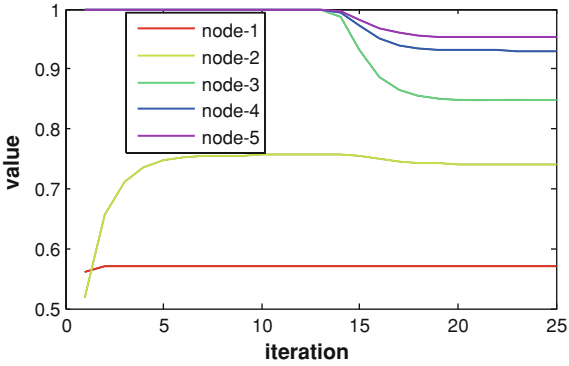


Fig. 175.4 Node values update

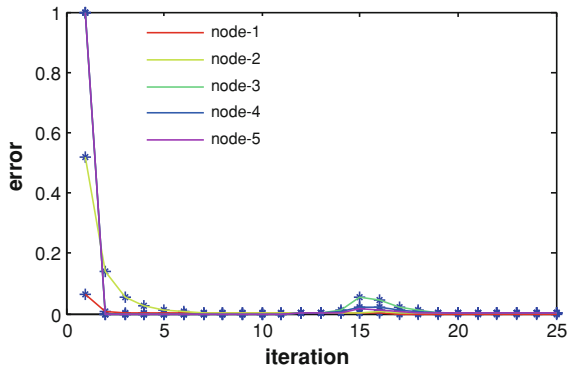


Fig. 175.5 Error versus iteration

indicate the fault propagation probability from node i to node j, if node i and node j has direct connection, then $w_{ij} = \mu_{ij}$, $\mu_{ij} \in [0, 1]$, else$w_{ij} = \mu_{ij} = 0$. The inference process for node i is denoted by:

$$v_i(k+1) = f\left(\sum_{j=1}^{c} w_{ij} v_j(k) + w_{0i}\right) \tag{175.3}$$

where $f(\bullet)$ denotes the shareholding function mapping the result of the sum to one of the valid values for the effect node, which means the fault from v_j propagates to v_i. The inference process can be realized by optimization algorithm, which we have studied in Ref. [12]. Figures 175.4 and 175.5 show failure cause effect for five nodes labeled from 1 to 5, it is assumed that initial node 1 is the cause node, while the other nodes are in normal state, and after sometimes, the fault propagates starts and begin disperse, the other nodes are affected simultaneously, and finally we can see that node 5 is the mostly affected by the fault from node 1, while node 2 is the lest affected by the fault from node 1 because of the two propagation path from node 1. Hebb learning algorithm is used to calculate the state value for node at every time step.

175.5 Summary

In this paper, we proposed a novel method of fault model and propagation analysis using the complex network in order to achieve integrated characteristics of electrical control systems. The aerospace power supply system is studied as an example to exploit the construction of the network topology and the node degree and clustering coefficient of the network is studied. The proposed method can be used to exploit fault analysis and maintenance management. Further study will be carried on the fault propagation dynamics in the network model.

Acknowledgments This research is supported by National Natural Science Foundation of China under Grant No. 60974058 and Ph.D. Programs Foundation of Ministry of Education of China under Grant No. 20070006047.

References

1. Liu L, Logan KP, Cartes DA (2007) Fault detection, diagnostics, and prognostics: software agent solutions. IEEE Trans Veh Technol 56:1613–1622
2. Doraiswami R, Diduch CP, Tang Jiong (2010) A new diagnostic model for identifying parametric faults. IEEE Trans Control Syst Technol 18:533–544
3. Shiozaki J, Shibata B, Matsuyama H, O'shima E (1989) Fault diagnosis of chemical processes utilizing signed directed graphs-improvement by using temporal information. IEEE Trans Ind Electron 36:469–474
4. Yuhong Z, Yiming Gu (2006) Visual method of fault diagnosis for a complicated process. J Chem Ind Eng (China) 57:2140–2145
5. Xiaowei Z, Hua X, Jihong C, Fei H (2002) Design of graphic Petri net tool and its implementation. J Huazhong Univ Sci Technol (China) 30:43–46
6. Luo X, Kezunovic M (2008) Implementing fuzzy reasoning Petri-Nets for fault section estimation. IEEE Trans Power Deliv 23:676–685
7. Haiwen Y, Hong L, Haibin Y (2006) Fault diagnosis of control electric component based on RBF neural network. J Beijing Univ Aeronaut Astronaut 32:544–547
8. Haiwen Y, Haibin Y, Hong L (2006) Investigation of comprehensive fault decision for control electrical apparatus element based on fuzzy computation. Chin J Sci Instrum 8:860–862
9. Haibin Y, Haiwen Y, Hong L (2006) Comprehensive decision of electrical apparatus control system failure based on AHP. Electrotech Appl 25:42–44
10. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:409–410
11. Songhua S (2005) Air vehicle and space vehicle power supply system. Beihang Publication, Beijing
12. Haibin Y, Haiwen Y, Yingyi L (2010) IEICE Technical report, vol 110, pp 251–254

Chapter 176
CDS: A Multi-Channel MAC Protocol with Collision Detection for Wireless Sensor Networks

Yahong Guo and Desheng Zhang

Abstract This paper proposes a multi-channel MAC protocol, CDS, for WSNs to tackle the collision incurred by channel selection. By using collision detection, when the collision occurs, a receiver can timely detect it and notify the sender to abort this transmission. Performance improves because the remainder of this message is not sent on this channel unnecessarily. Instead, this channel is released for other more productive transmissions. Therefore, CDS can efficiently handle the collision, and conserve more energy to extend the lifetime of WSNs. More importantly, CDS is fully distributed with no requirements of time synchronization or multi-radio, so it is practical to realize CDS in sensor nodes. Moreover, CDS is evaluated in testbed experiments, and the results indicate that when the number of channels, loads and number of neighbors increase, CDS significantly improves throughput.

Keywords WSN · MAC protocols · Collision detection

176.1 Introduction

Wireless sensor networks (WSNs) have a wide range of potential applications [1], but current off-the-shelf sensor devices only provide very limited bandwidth in a single channel, which undermines the performance of these applications [2].

Y. Guo (✉)
School of Infomation Science and Technology,
Heilongjiang University, Harbin Heilongjiang, China
e-mail: jbli@hlju.edu.cn

D. Zhang
School of Computer Science and Technology,
Heilongjiang University, Harbin Heilongjiang, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_176,

A potential solution is to utilize multi-channel MAC (mcMAC) with parallel communication to improve WSNs performance. Unfortunately, in distributed multi-channel WSNs with low cost sensor nodes, it is very challenging for the nodes to choose an idle channel for transmission. Many efforts had been made to let nodes choose an idle channel for transmission, but few works aim to handle the collision if the selected channel is not idle.

The most famous scheme to handle collision is CSMA/CD (carrier sense multiple access with collision detection) in WLAN. Under CSMA/CD, the sender simultaneously sends and listens on the wired channel. If the collision is detected, the sender immediately aborts this packet, and thus the remainder of the packet is not sent unnecessarily. However, this efficient scheme cannot be used in wireless networks due to the following constraints. (1) The sender in wireless network cannot send and listen on the same channel simultaneously [3]. (2) Although the sender can send and listen simultaneously, e.g., with another radio, the self signal is too strong to let the senders detect a collision [4]. (3) Although the sender can detect a collision, the multi-hop characteristic of wireless network makes this detection in vain, because the collision at the sender does not necessarily indicate the collision at the receiver and vice versa [5]. Limited by these constraints, a node in wireless network has to handles collision with CSMA/CA (collision avoidance) using backoff scheme, i.e., binary exponential backoff. However, when the collision occurs, the corrupted remainder of packet is still unnecessarily sent and received. More importantly, the channel is occupied by corrupted packets, and cannot be efficiently utilized. Thus, compared with CSMA/CD, CSMA/CA is suboptimal. It would be desirable to emulate CSMA/CD like method even in wireless network.

In order to gain the benefit of CSMA/CD in the WSNs, in this paper, we propose a collision detection-based mcMAC CDS (collision detection for WSNs), which handles the collision incurred by channel selection in WSNs scenarios. The high-level idea behind this scheme is simple. When detecting the collision, the receiver employs a new kind of packets, NTF, with highest priority (i.e., with the shortest interframe spacing) to notify the sender that the packet received by the receiver is corrupted and this channel cannot be used for further transmission. This NTF can be sent on either the Data Channel (DC) with the data packet, or the control channel (CC) using another radio. The former involves delay, but does not involve any new hardware cost; whereas the latter almost simultaneously notifies the sender when the collision occurs, but involves multi-radio cost. In this paper, we focus on the former, i.e., we utilize the single radio with acceptable delay in CDS to emulate multi-radio scheme.

Even with the collision detection from the receiver, another challenge remains. The packet length of WSNs (e.g., 30–50 bytes) is significantly shorter than the packet length of general wired and wireless networks (e.g., 512 + bytes) [1] and thus the remainder of corrupted packet is also short. Therefore, the benefit of CSMA/CD is diminished. To fully gain the benefit of CSMA/CD, in this paper we propose CDS where multiple short data packets (i.e., a message) are sent

continuously without ACK interruption to emulate a long data packet in wired or wireless networks. This paper hopes to contribute as follows:

(1) This paper makes an early attempt to apply the idea of CSMA/CD in single radio multi-channel WSNs. No previous work does such investigation. Although some works provide a solution with collision detection in wireless network [4, 6], they assume that the node has two radios or antennas, and did not consider the packet length issue in WSNs.
(2) This paper proposes a mcMAC CDS with practical constraints of WSNs in mind. By taking the packet length issue into account, CDS is capable of fully taking advantage of CSMA/CD, and does not involve additional energy consumption, e.g., multi-radio scheme or time synchronization, in order to extend the lifetime of WSNs.
(3) This paper implements and evaluates CDS on a small scale prototype of 10 nodes. The testbed results indicate that ARM achieves 67% more throughput ratios at most in a small-scale network. More importantly, the testbed results demonstrate that further progress in collision detection is feasible and worth pursuing.

176.2 Related Work

Synchronous mcMACs for WSMs Zhou et al. [1] proposed MMSN which is the first mcMAC that takes into account the restrictions imposed by WSNs. The senders in MMSN switch their current channels to channels of the intended receivers at the beginning of every slot when they have packets to send [7]. Kim et al. [8] proposed Y-MAC for WSNs where time is divided into several fixed-length frames. The frames are composed of a broadcast period and a unicast period. The difference between Y-MAC and other mcMACs is that Y-MAC schedules the receivers rather than the senders to achieve low energy consumption to extend the lifetime of WSNs.

Salajegheh et al. [9] proposed HyMAC for WSNs where the communication period consists of a number of frames, which are divided into scheduled slots and contention slots. The base station allocates specific time slots and channels to all the nodes for communication. Jovanovic et al. [10] proposed TFMAC for WSNs, which works similar to HyMAC except that the schedules are made by all the nodes rather than the base station.

Asynchronous mcMACs for WSNs Luo et al. [11] exploited Distributed Information SHaring mechanism (DISH) and proposed CAM-MAC for ad hoc networks. In CAM-MAC, when a node-pair performs a channel reservation on the CC, all the neighbors may send cooperative packets to invalidate the reservation if they aware of the fact that the selected DC or the intended receiver is unavailable. Luo et al. [12] proposed ALTU based on altruistic cooperation, which introduces some specialized nodes called altruists whose only role is to acquire and share

channel usage information. Le et al. [13] proposed PMC which utilizes a control theory approach to dynamically add available channel in a distributed method. In PMC, nodes work on current available channels by CSMA, and decide whether to switch to the next available channel based on certain parameters, which vary with channel utilization from time to time. However, computing methods of these parameters need further discussion.

Wu et al. [14] proposed TMCP which is a multi-channel protocol that does not require time synchronization among the nodes. However, this protocol is more like a topology control protocol than a MAC protocol. Zhou et al. [15] proposed CUMAC using cooperation for underwater WSNs, but it requires an extra hardware cost, i.e., tone device, on every node in WSNs, which increases the cost of WSNs.

Summary Whether synchronous or asynchronous mcMACs for WSNs, they usually focus on the choice of the actually idle channel for transmission, and this objective will involves too much cost, e.g., time synchronization, cooperation communication, complicated parameter computation, extra hardware, etc. On the contrary, in this paper, we admit the flaw of channel selection and aim at handling the possible collision caused by it. Via avoiding the above-mentioned cost, CDS is expected to significantly improve the network performance.

176.3 The Collision Detection of CDS

Since this paper aims at proposing an attempt to rethink collision resolution in WSNs, this section only provides a prototype design of CDS and sophisticated optimal scheme leaves for future work. Three assumptions are made as follows. (1) Wireless bandwidth is divided into one CC for control information and K DCs for data packet and NTF packets, and every channel is orthogonal to the others. (2) CDS can be supplemented by the most of channel section schemes, and focuses on the collision detection and resolution. (3) All the nodes are assumed to be equipped with the same half-duplex radio.

When the node in CDS actually has a message containing multiply packets to send, five steps must be accomplished as in Fig. 176.1. (1) When the sender S has a message to send to the receiver R, S selects a DC based on its channel selection. Then, under CSMA scheme, S sends an RTS with channel reservation information to R on the CC. (2) After R receives this RTS, R send a CTS to confirm this reservation on the CC. Next, R switches its channel to the DC that S had selected before and waits for data packets. (3) After S receives this CTS, S also switches its current channel to the same DC and sends all the data packets belonging to a message to R. (4) After R correctly receives all these data packets, R sends an ACK to S to confirm this message transmission. (5) After S receives this ACK, S and R switch back to the CC together.

Two questions arise with respect to the design and applicability of CDS. First in step (4), what if R cannot correctly receive all the packets, i.e., collision occurs? If

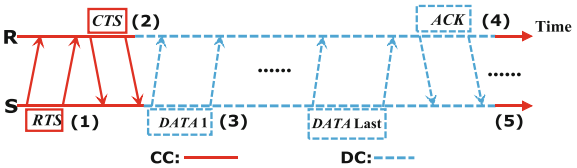


Fig. 176.1 Communication scheme

collision occurs, R first drops the corrupted packet immediately, and then chooses another DC according to its channel selection, and sends ID of this DC to S in an NTF, and finally switches to this DC. As S receives this NTF in a short interframe spacing between every two data packets, S aborts this transmission on this DC, and switches to the DC specified in NTF to retransmit.

Second, in step (4), why R sends an ACK after continuously receiving all the data packets belonging to one message instead of one data packet? There are two main reasons why we utilize this scheme in CDS. (1) Since the packet in WSNs is shorter than in other networks, the probability that a data packet experiences a collision is also smaller. If every data packet requires an ACK, many ACKs will be sent in vain. This may cause the excessive ACK overhead problem [4]. Therefore, we employ only one ACK to confirm a whole message containing multiply data packets. (2) In CDS, if the receiver does not correctly receive a data packet due to the collision, the receiver can utilize NTF to supplement this one ACK scheme by timely notifying the sender to stop sending the rest of the data packets, and to retransmit both the failed packet and the rest of packets in another DC. This one ACK and possible NTF combined scheme seems to be naïve, but how much of the gain we can obtain is still unclear. We will investigate this issue in Sect. 176.4.

176.4 Performance Evaluation

We built a sensor node platform, Hawk, for our testbed experiments in this section. Several experiments were conducted to evaluate CDS's performance. Hawk utilizes μC/OS [16], where each node is equipped with an nRF905 radio and an MSP430 processor. A hawk node is shown in Fig. 176.2a. Three LEDs are used, i.e., red, green, and yellow, on each node to indicate specific events. For instance, when the red LED is on, it indicates that the node is communicating on a DC. Meanwhile, the green LED and the yellow LED jointly indicate the specific number of DCs, e.g., in Fig. 176.2b the red and yellow LEDs of node 9 are on, which indicates it is communicating on DC3. When the red LED is off and the yellow and green LEDs are on, it indicates the node is on the CC. When all the three LEDs are off, it indicates the node is inactive.

The testbed consists of 10 hawk nodes which are completely connected as shown in Fig. 176.2b. In such a case, all the nodes are within the communication range of each other, similar to the work in [11]. The size of data packet is 32 Byte as in the simulation, and data transmission rate is 100 Kbps. All the nodes

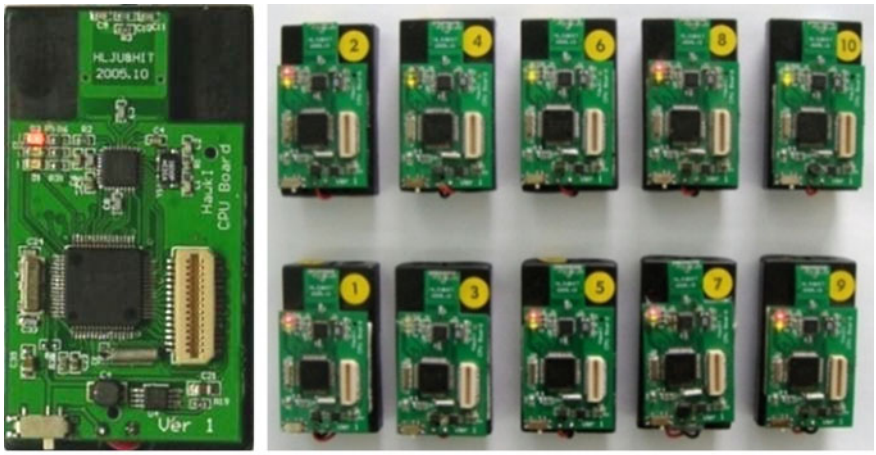


Fig. 176.2 Testbed experiments **a** Hawk node. **b** A snapshot in the testbed experiment

randomly choose a neighbor to initiate a transmission. Every experiment was repeated 10 times. When an experiment is finished, all the nodes send their total amount of bytes received during the experiment to a Sink node one by one. Sink is connected to a desktop computer, and thus throughput can be successfully obtained. The throughput is computed with the same method in the simulation. Due to the time synchronization of MMSN and the complexity of PMC for parameter computations, only CAM-MAC, CDS, CDS-R and CDS-O were implemented for throughput comparisons.

The throughput is explored as NCBR is fixed to 5 and the different TNCs are used. From Fig. 176.3a, it is observed that CAM-MAC has higher throughput than CDS and its two varieties, when TNC is less than or equal to 3. CAM-MAC does not have to detect the collision incurred by channel selection, when the nodes have packets to send. This collision detection compromises the throughput of CDS, when collision is less serious. However, as observed from Fig. 176.3a, CDS indeed achieves better throughput when TNC is larger than or equal to 4. A possible explanation for this throughput improvement is that when more DCs are available, the collision incurred by the channel selection becomes more serious. And the collision detection scheme utilized by CDS is more efficient than ordinary retransmission employed by CAM-MAC. In addition, CDS has similar throughput compared with CDS-R when TNC is small; whereas CDS outperforms CDS-R, when TNC is larger than 2. These results indicate that the collision detection utilized by CDS is more beneficial in the network with more channels. In fact, current off-the-shelf WSNs radios such as CC2420 already offer multiple channels more than 15 for parallel communications. Therefore, the collision detection scheme is both feasible and applicable in real WSNs application. When compared

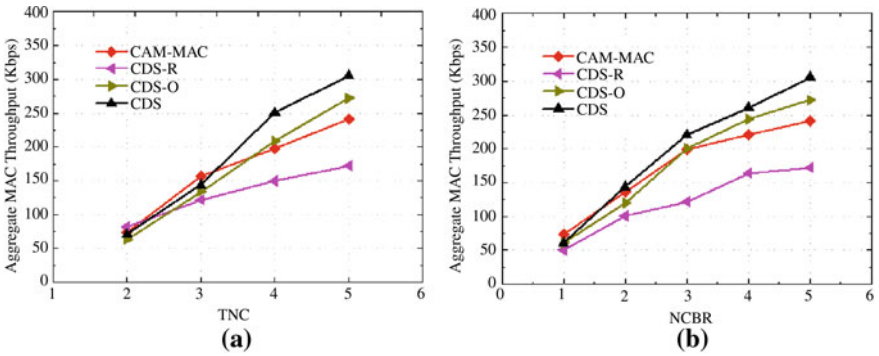


Fig. 176.3 Testbed experiment results. **a** Throughtput versus TNC. **b** Throughtput vs. NCBR

with CDS-O, CDS shows a roughly 20% better throughput, which indicates that the excessive ACK overhead indeed affects the performance of protocols.

When TNC is fixed at 5 and the loads are varying, the throughput is observed in Fig. 176.3b. It shows that CDS and its two varieties have lower throughput than CAM-MAC, as loads are small, i.e., NCBR is equal to or less than 2. This is due to the fact that when fewer nodes are communicating, cooperative scheme of CAM-MAC works better in avoiding possible collision than collision detection scheme of CDS. However, when loads are heavy, fewer nodes are left for cooperation, so collision becomes more serious. Therefore, CDS outperforms CAM-MAC as NCBR is equal to or larger than 3. The scheme used by CAM-MAC becomes inefficient, because in the network under heavy loads, it is very challenging to completely eliminate the collision, no matter which channel selection is used. CAM-MAC aims to avoid the collision before it occurs, while CDS aims to handle the collision when it occurs. Note that CDS performs better than CDS-R under all the loads. This indicates that collision detection works well not only in the network with heavy loads but also in the network with light loads. In addition, compared with CDS-O, CDS also has 10–20% higher throughput. This is because without excessive ACK overhead, the entire transmission time is shorter, but the reliability does not become lower, and thus the throughput becomes higher.

176.5 Conclusion

To handle the collision incurred by channel selection, this paper proposes a new multi-channel MAC protocol CDS with collision detection scheme for WSNs. CDS is compatible with the most of channel selection. Instead of involving more cost to avoid the collision, CDS aims to detect and to handle the collision, when the collision occurs. The extensive simulations are conducted to evaluate the performance of CDS in large scale WSNs. The simulation results show that with

collision detection scheme, CDS is capable of efficiently handling the collision. Furthermore, CDS is implemented in a real testbed, and the experimental results show that the collision detection scheme actually enables CDS to achieve better throughput. More importantly, the real testbed results demonstrate that further progress in collision detection scheme under WSNs is both feasible and worth pursuing.

References

1. Zhou G, Wu Y, Yan T et al (2010) A multifrequency Mac specially designed for wireless sensor network applications, In ACM TECS, 9(4):1–41
2. Culler D, Estrin D, Srivastava M (2004) Overview of sensor networks, In IEEE Computer, Special Issue on Sensor Networks
3. So J, Vaidya N (2004) Multi-Channel MAC for ad hoc networks: handling multi-channel hidden terminals using a single transceiver, In MOBIHOC
4. Sen S, Choudhury RR, Nelakuditi S (2010) CSMA/CN: carrier sense multiple access with collision notification, In ACM MobiCom
5. Wu SL, Lin CY, Tseng YC, Sheu JP (2008) A new multi-channel MAC protocol with on-demand channel assignment for multi-hop mobile ad hoc networks, In ISPAN
6. Peng J, Cheng L, Sikdar B (2007) A wireless MAC protocol with collision detection, In IEEE TMC, 6(12):1357
7. Zhou G, Huang C, Yan T et al (2006) MMSN: multi-frequency media access control for wireless sensor networks, In INFOCOM
8. Kim Y, Shin H et al (2008) Y-MAC: an energy-efficient multi-channel MAC protocol for dense wireless sensor networks, In IPSN
9. Salajegheh M, Soroush H, Kalis A (2007) HyMAC:hybrid TDMA/FDMA medium access control protocol for wireless sensor networks, In PIMRC
10. Jovanovic M, Djordjevic G (2007) TFMAC: multi-channel MAC protocol for wireless sensor networks, In TELSIKS
11. Luo T, Motani M, Srinivasan V (2009) Cooperative asynchronous multichannel MAC: design, analysis, and implementation, In IEEE TMC, 8(3):338–352
12. Luo T, Motani M, Srinivasan V (2007) Altruistic cooperation for energy-efficient multi-channel MAC protocols, In MOBICOM
13. Le HK, Henriksson D et al (2008) A practical multi-channel media access control protocol for wireless sensor networks, In IPSN
14. Y Wu, Stankovic JA et al (2008) Realistic and efficient multi-channel communications in wireless sensor networks, In INFOCOM
15. Zhou Z et al (2010) Handling triple hidden terminal problems for multi-channel mac in long-delay underwater sensor networks, In INFOCOM
16. Labrosse J MicroC/OS-II, the real-time kernel, 2nd edn. ISBN:1-57820-103-9

Chapter 177
Dynamic Positioning Control System of the Dredger

Zhang Yu-hua and Jiang Jian-guo

Abstract A Dredger's dynamic positioning system (DPS) is a real-time control system composed of several complex algorithms. A dredger's dynamic positioning real-time control system is designed with Siemens Industrial-Personal-Computer (IPC)-based controller using WinAC RTX software. A real-time industrial Ethernet (Profinet IO) structure is used in this system. Traditional Hardware-Programmable-Logic-Controller-based platform is replaced with the WinAC software-based one in the Industrial Personal Computer. Software and hardware resources in Industrial Personal Computer are well utilized. This design not only improves the operation speed and saves the cost, but also improves the control real-time performance and enhances the network' robustness.

Keywords WinAC RTX · Profinet IO · Dynamic positioning system · Dredger

Z. Yu-hua (✉) · J. Jian-guo
Key Lab of Control of Power Transmission and Transformation,
Ministry of Education, Department of Electrical Engineering,
Shanghai Jiaotong University, Shanghai 200240, China
e-mail: azhyh@sjtu.edu.cn

J. Jian-guo
e-mail: bjiang@sjtu.edu.cn

Z. Yu-hua
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jaozuo 454000, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_177,

177.1 Introduction

Dredgers' Dynamic positioning system (DPS) is necessary for the dredger to keep in a specific position or pre-defined path through the action of propellers when the dredger is dredging or backfilling of offshore deep groove, or pumping ashore. Since the 1990s, dynamic positioning systems of dredgers have been rapid developed as the dredging market increasing rapidly. The corresponding technology has been developed also from analog to digital, up to network mode. And the study about the DPS is also popular in recently years [1–3].

In order to remain at a certain position or orientation, The DPS must overcome interferences by the different propellers distributed around the ship with its inherent power. The interferences include wind, waves, currents, and dredging action and so on. Therefore, there are three steps in the DPS's control processes. First, information is collected; including interferences, given reference values, the vessel's position and heading, and so on. Second, control algorithms are carried out, including complex algorithms of filter, control, and thrust distribution. Then the information of thrust is transmitted to the propellers, the propellers produce thrust according to calculation results by the computer. So, it is hard to achieve those complex algorithms in Hardware-Programmable-Logic-Controller-based platform, and large amounts of data transmission requires rapid real-time network.

One real-time industrial Ethernet network of IPC-based Profinet IO is used in this DPS instead of the traditional Programmable-Logic-Controller (PLC) hardware-based one. The IPC's rich soft-hardware resources can be well utilized, which not only improves the velocity of control and saves the cost, but also enhances the system's real-time ability and the network robustness.

177.2 System Design

Profinet IO mode is adopted in this DPS, which is Real-time industrial Ethernet of IPC-based distributed IO structure. The system structure is shown in Fig. 177.1. All the Measuring equipments and propellers are connected to network through distributed IO modules ET200s.

WinAC RTX as the core control software for the IPC-based industrial Ethernet, Which Performs complex calculations of filtering, control, and thrust distribution programs. Ethernet card CP1616 is used as the controller in the network. ET200ses get information of position, heading, attitude, dredging reaction, and wind force from the sensors and connected to the IPC via Profinet network. Filtering procedures isolate the ship's the low-frequency components position information which reflects the true position of the ship. Thrust that each propeller should yield is calculated by the thrust allocation procedures according to the principle of the minimizing power consumption. Then the Dredger is moved to the specified position or maintained at the given position with the thrust action. Filters, control,

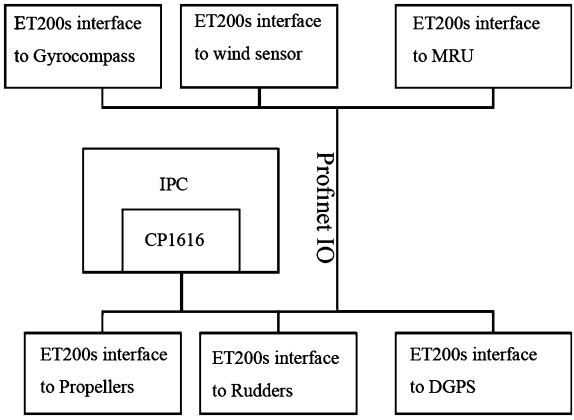


Fig. 177.1 Structure of the DPS' control system

and thrust distribution programs are written by Microsoft C/C++ language, and are compiled to dynamic link libraries by WinAC ODK. These dynamic link libraries can be called in STEP7 Program to achieve the task of the DPS.

177.3 Hardwares

177.3.1 IPC

Siemens compact and high performing IPC477C is selected in this system. The HMI IPC477C is available as an embedded PC-System, preinstalled visualization software WinCC flexible and Software PLC WinAC RTX, and the CP1616 Ethernet card is onboard. IPC477C equipped with highly efficient energy-saving Intel ® Core ™ 2 Duo processor and SSD solid state drive. So IPC477C can meet the requirements of the dredger's DPS.

177.3.2 ET200s

ET200s modules are used to connect field devices to IPC in Profinet IO systems. ET200s is widely used in various industrial control field for its Bit-modular and highly scalable advantages. It is an important part of Siemens ET200 distributed IO system. IM151 interface module of ET200s can be inserted up to 64 IO modules (Including digital input and output, analog input and output, technology module, SSI encoder modules, stepper motor control module, pulse output module, serial communication module, and so on). About 128 bytes of input and output data can be handled and there are no slot restrictions in this modular. So it is very convenient for the system expansion and renovation.

177.3.3 Position Measurement Device

The PEM's choice depends on the ship's missions and the level of the DPS. Differential global positioning system (DGPS) is an ideal choice for the dredge because of its constantly changing location and not too high positioning accuracy requirements. DGPS can fully meet the positioning measuring requirements of the dredger' DPS because DGPS with characteristics such as free of range and depth restrictions, global scope, ±1 m measurement accuracy.

177.3.4 Gyrocompass

In order to measure bow roll angle, ANSCH ¨ UTZ standard 20 digital Gyrocompass is used in this DPS with static accuracy of ±0.1° and the dynamic accuracy of ±0.4°.

177.3.5 Move Reference Unit

Dynamic motion sensor is a high-accuracy hydrographic surveying instrument in dynamic condition. It is widely used as auxiliary equipment on vessel. It can enhance the measuring accuracy by correcting offset from heave, roll, and pitch of vessel. Move reference unit measure the ship's pitch, roll and heave for compensation of the GPS antenna. And Kongsberg Seatex\MRU is selected.

177.3.6 Wind Sensor

In order to compensate for the wind disturbance, It is important to accurately measure wind speed and wind direction, for the wind is a major disturbance in the DPS. Usually the wind sensor' accuracy required is below: wind speed ±0.3 m/s, measurement threshold 1 m/s, wind direction 1°, accuracy ±3°. The wind sensor should be installed at the top of the mast and the wind shadow should be avoided.

177.4 Soft Wares

177.4.1 WinAC RTX 2009

WinAC RTX is PC-based automation solutions for the control system of high-speed and precise time required, which has "hard time" and "anti-death" features. WinAC has broken the traditional template-style PLC's limitations in resources,

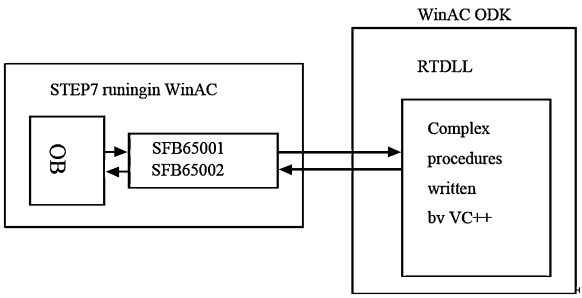


Fig. 177.2 Relationship of WinAC, WinAC ODK and VC++ procedures

interface, and function. The IPC' high computing ability and large storage space can adapt to variety of automation applications [3]. WinAC can complete the demanding of real-time control tasks with ODK. In which a high-speed interface to program is provided by ODK for users. C/C++ Control algorithms code can be integrated as a standard PLC program function block in the software PLC ladder [4, 5]. It is a high-performance system for real-time control programs only occupying a small part of CPU resources of IPC in the control cycle. So there are enough resources of CPU that can be used by other parallel complex tasks.

177.4.2 Control Programs

There are many complex control programs in the DPS, such as low-frequency signal filtering procedures, control procedures, thrust allocation procedures, data collection procedures, and so on [1]. Taking into account the complexity of the algorithms, those complex programs are developed to many functions in dynamic link library (DLL) by the WinAC ODK and VC++ development toolkit, and then these functions can be called by WinAC with the system function blocks SFB65001 and SFB65002 in STEP7. Figure 177.2 shows the relationship between the WinAC and the control programs.

177.4.3 WinAC Control Programs

There are some procedures in WinAC control programs, such as data exchange procedures between IPC and WinAC, Fault handler, and so on. Flow chart is shown in Fig. 177.3.

WinAC programs were compiled by the programming tool Step 7, and then were download to WinAC. Modular design was adopted and control functions were compiled into different submodules. These submodules were called in OB1 according to the circumstances. Fault handler is called automatically when the system fails to run.

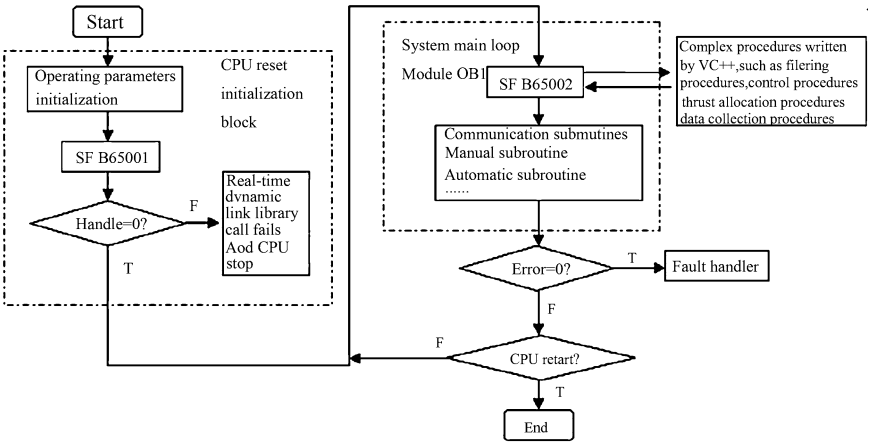


Fig. 177.3 Main flow chart of control program

177.5 Conclusions

A DPS is designed based on real-time Profinet IO with the following characteristics compared with the conventional system

1. IPC-based integrated automation DPS increased running efficiency. IPC-based and software PLC control structure simplified the hardware structure.
2. Complex control programs can be integrated to real-time control program easily.
3. Rapid real-time requirements are met and programing effort is reduced, which enhance the systems real-time and increase the system's developing speed.
4. HMI and control programs can be integrated in one computer, which saves installation space and cost.

Acknowledgments This work was supported by 863 project in China (2011AA050403), and Control Engineering key subject open-laboratory open-fund of Henan Province in China (KG2009-02).

References

1. Tannuri EA, Morishita HM (2006) Experimental and numerical evaluation of a typical dynamic positioning system. Appl Ocean Res 28:133–146
2. Tannuri EA, Agostinho AC (2010) Dynamic positioning systems: an experimental analysis of sliding mode control. Control Eng Pract 18:1121–1132
3. Leavitt JA (2009) Optimal thrust allocation in a dynamic positioning system. Trans soc Naval Architects Mar Eng 116:153–165

4. Neumann P, Pöschmann A (2005) Ethernet-based real-time communications with PROFINET IO. WSEAS Trans Commun 4:235–245
5. Hanzálek Zdenêk (2010) Profinet IO IRT message scheduling with temporal constraints. IEEE Trans Ind Inf 6:369–380
6. Kleines, Harald, Detert, Sebastian (2008) Performance aspects of PROFINET IO. IEEE Trans Nucl Sci 55:290–294

Chapter 178
Research on Technology for Context Awareness in Pervasive Computing Environment

Xianquan Zeng and Yudong Feng

Abstract To facilitate the development and deployment of context-awareness application in pervasive computing, a context-awareness framework needs to be used to adapt to different execution contexts and hide heterogeneity from applications. Therefore, after examining the concept of context and context awareness computing and some key technologies for context-awareness in pervasive computing, including context acquisition methods, context modeling and reasoning, and programming context etc., a dynamic context-awareness framework is put forward to effectively manage the context information and services mobility which has not been supported in the previous models. The model provides the standard interface to realize the interaction among the heterogeneous devices and hide heterogeneity from applications.

Keywords Context · Context awareness · Pervasive computing · Framework

178.1 Introduction

Ubiquitous computing environments within Weiser's vision encompass a spectrum of computation and communication devices that need to weave themselves into the fabric of our everyday existence so that they can be accessed by anyone, anywhere,

X. Zeng (✉)
School of Computer Science and Technology, Xuchang University, Xuchang, China
e-mail: xianquanzeng@hotmail.com

Y. Feng
Information Technology Department, Zhengzhou Electric Power College, Zhengzhou, China
e-mail: fengyudong@zepc.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_178,

anytime or anything [1]. This vision places a number of requirements on pervasive computing applications to gracefully integrate themselves with human users and to operate in high dynamic environments placing minimal overhead on users. These requirements are important as users are continually exposed to distractions in what is a highly mobile environment, where walking, driving, and other real world activities preoccupy the user. To minimize such negative effects, pervasive systems applications must be aware of and adapt to the situation in which they are running, including the state of the physical space, the users and the available resources. So the applications in pervasive computing need to be context-awareness in order to respond quickly to their dynamic computing environments and provide smarter services.

To enable designers to easily build context-aware applications, there needs to be the architectural support to provide the general mechanisms required by context. General context-aware systems and architectures will require these mechanisms to support context sensing and context reasoning. In context sensing, mechanisms are provided to deal with the acquisition of information from the physical environment while context reasoning deals with the interpretation of the acquired information. Here, on the fundament of analyzing the main approaches presented in the literature for creating context based architectures and the methods used for context acquisition and reasoning, we propose a generic framework for context-awareness computing to supports the context information management and service mobility.

The remainder of this paper is organized as follows. Sections 178.2 and 178.3 introduce the concepts and technologies on context awareness computing respectively. Section 178.4 presents a framework for context-awareness computing by focusing on the management of context information and service mobility. Conclusions and future works are presented in Sect. 178.5.

178.2 Concept of Context and Context-Awareness

Context is one of the core concepts in the context awareness computing. With the many implementations of applications and services that make use of context information, context has been interpreted from different aspects. Here, two representative definitions used in the pervasive computing field are presented.

One is from Schilit and Adams [2], they give the definition of context focusing on extension of context. They think that the context consists of three parts: (1) user environment, including user position, marking, social relationships, etc. (2) computing environment, including computing resources, computing power of devices, network connectivity, etc. (3) physical environment, such as light, noise level, etc. The other definition is given by Dey from the connotation of context. Dey [3] defined context as “any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered

relevant to the interaction between a user and an application, including the user and the application themselves".

From the various refinements in the definition of the term context, Dey's definition have become widely accepted and adopted to provide a consistent understanding of the subjects studied by researchers.

According to a system is context awareness if it uses context to provide relevant information and/or services to the user, here context awareness is defined as a system that can respond intelligently in either a reactive or proactive fashion to the combined information acquired by the applications, either explicitly through the physical environment (i.e. through sensors) or implicitly (i.e. by reasoning about this environment). This context awareness in turn enhances services provided to users by taking into account their personal tastes; it makes the environment more knowledgeable and it promotes autonomy and adaptability in applications themselves, liberating users from being concerned with distractions.

178.3 Key Technologies for Context-Awareness

178.3.1 Context Acquisition Methods

The method of context-data acquisition is very important when designing context aware systems because it predefines the architectural style of the system at least to some extent. Chen defines the following are three categories of context acquisition methods [4]:

Direct access to hardware sensors. The sensors are integrated into the devices. The client software gathers the context information directly from these sensors rather than by a specialized infrastructure. In this approach, the high-level applications can have great controls over the operations of the low-level sensors, thus can have better knowledge about how different data is collected and computed. However, this method presents a number of problems with regard to extensibility and reusability. It is difficult to extend the application that is gathering the data and because the sensor data acquisition is tightly coupled with the application and the reusability of code is limited.

Facilitated by a middleware infrastructure. The idea is that instead of letting the applications to manage the low-level sensing details, middleware infrastructures are provided to facilitate sensing. Context acquisition middleware typically built into the hosting devices or platform on which the context-aware applications operate. Compared to direct sensor access, this method hide low level sensing details, which eases the extensibility and simplifies the issue of the reusability of hardware dependent sensing code.

Acquire context from a context server. This approach shifts the context acquisition procedures into the implementation of a server entity that runs on a resource-rich device. The server entity provides context information to different

context-aware applications in a distributed environment. This approach overcomes the drawbacks noted with hosting device of a context-aware application that has limited computing resource. Besides the reuse of sensors, the usage of a context server has the advantage of relieving clients of resource intensive operations.

178.3.2 Context Modeling and Reasoning

Context modeling and reasoning relates to the task of using context data in an intelligent way and is perhaps the most challenging of contemporary research tasks in creating context-awareness. A more precise definition for context reasoning is deducing new information relevant to the use of application(s) and user(s) from the various sources of context-data. Service behavior adaptation is an output from context reasoning where services can adapt their functionality and behavior to a context by applying either rule-based or learning-based logic.

With regard to rule-based logic, there are several approaches, including First Order Logic (FOL), fuzzy logic, description logic and temporal logic. In these approaches, actions are triggered by a set of rules whenever the current context changes. However, these approaches do not allow actions to be performed when the user's context has not been envisioned beforehand [5]. Within learning-based reasoning, several approaches have been proposed: case-based reasoning, neural networks and Bayesian approaches [6]. In case-based reasoning, problem solving is based on remembering specific experiences that might be useful for the problem being solved. Neural networks have been used in context reasoning as a method of learning and adapting to context. Bayesian networks provide a probability-based reasoning method that is very useful to model the uncertain nature and to reason about the probabilistic occurrence of a context. The learning methods essentially provide a more sophisticated approach compared to rule-based approaches, which lack robustness and flexibility, are confined to narrow problem domains and are difficult to maintain and update throughout their life-cycle.

Context modeling is a topic that is very closely related to context reasoning. As there will be a wide range of heterogeneous context information in pervasive systems, modeling of contextual information is highly necessary. Strang et al. [7] presented a survey of six context modeling approaches: Key-value modeling, Markup scheme modeling, Graphical modeling (UML, ER, etc.), Object-oriented modeling, Logic-based modeling and Ontology-based modeling approaches. Key-value modeling techniques uses key value pairs to model context by providing the value of context information to an application as an environmental variable. Markup Scheme modeling techniques is a hierarchical data structure consisting of markup tags with attributes and content. Within graphical models, the Unified Modeling Language (UML) is appropriate to model context. Object-oriented modeling allows modeling of context by using object oriented techniques, offering the full power of object orientation, including encapsulation, reusability and inheritance. Logic-based models have a high degree of formality where facts,

Table 178.1 Comparison of context modeling method

	Key-value modeling	Markup scheme modeling	Graphical modeling	Object-oriented modeling	Logic-based modeling	Ontology-based modeling
Ability of expressiveness	Low	Medium	High	High	Higher	Higher
Complexity	Low	Low	Medium	Medium	Higher	High
Degree of formality	Low	Low	Low	Medium	Higher	High
Degree of reasoning	Low	Low	Low	Low	High	Higher

expressions and rules are used to define the context model. In ontology-based models, ontology present a description of the concepts and relationships and provide a very promising instrument for modeling contextual information due to their high and formal expressiveness and due to the possibilities for applying ontology reasoning techniques.

Based-on analysis of modeling techniques, it is evident that the most promising context modeling approaches for pervasive computing systems to date are ontology. The comparison results of six context modeling method are shown in Table 178.1.

178.3.3 Programming Context

In the past few years, a large number of diverse context information frameworks have been developed. However, the proliferation of context-aware applications is inhibited by the absence of a generic approach to such framework development and by a lack of programming support to rapidly develop them. To build a context-aware application, developers need to either design or implement their own applications from scratch, requiring them to write code that directly interacts with devices, or to use a toolkit that hides many of the device details from them. The problem with this approach lies in the fact that program behaviors in pervasive systems are difficult to predict and understand. To reduce this complexity and to enable programmers to easily develop and deploy context aware applications, a programming framework needs to be used.

This issue of programming for pervasive systems is not well understood. Existing approaches, such as PIMA [8] and Aura [9], utilize the idea that the application model for pervasive computing must decouple the application logic from details that are specific to the run-time environment. The Aura project created a programming model for task based pervasive computing. In addition programming languages, such as Telescript, Obliq and Java, have been investigated to address the practical concerns of building distributed pervasive systems, but none of these approaches have specifically considered context as a central modeling

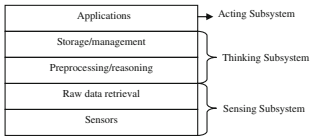


Fig. 178.1 Abstract layered architecture for context awareness system

concept. As a result a number of programming methodologies have been devised, including Object-oriented (OOP), Aspec-oriented (AOP), Feature-oriented (FOP) and Context-oriented programming (COP). Using these programming methodologies and frameworks, programmers can focus on modeling and using context information and functionality specific to their application while relying on a basic infrastructure to handle the actual management and distribution of this information.

178.4 Framework for Context-Awareness System

The framework of context awareness system is primarily responsible for managing, coordinating and scheduling the devices, contexts, physical environment etc., which constitute the computational environment of system, and realize the interoperability among objects. On the other hand, the framework can also shield the complex, diverse and dynamic computational environment, and provide a unified framework and application programming interface (API) for application development. Figure 178.1 [10] categorized the components of a context-awareness pervasive system as sensing, thinking and acting. Sensing forms the initial part of the equation where it looks at how raw environmental information could be acquired, represented and interpreted in a cohesive manner.

Currently, there has been a large number of the system framework for context awareness computing such as ContextTookit, Gaia, PACE, SOCAM and CoBrA. etc. These frameworks have already carried out developing appropriate high-level programming abstractions and toolkits to simplify the process of sensory data acquisition, however, these frameworks have been only used in specific application, and cannot effectively support services mobility which should be supported in pervasive computing [11]. After examining previous work on system framework for context-awareness computing, we propose a system framework for context-awareness computing, which can support the management of service mobility and context information. The overall structure of this framework is shown in Fig. 178.2. The organization is logical and doesn't describe each component's actual location or distribution. The framework of context awareness consists of context manager and service mobility manager. Context manager deduces the context information using the raw data from the sensors and delivers the information to the desired context awareness service. Service mobility manager manages the service mobility to realize the transfer of computing tasks without interruption.

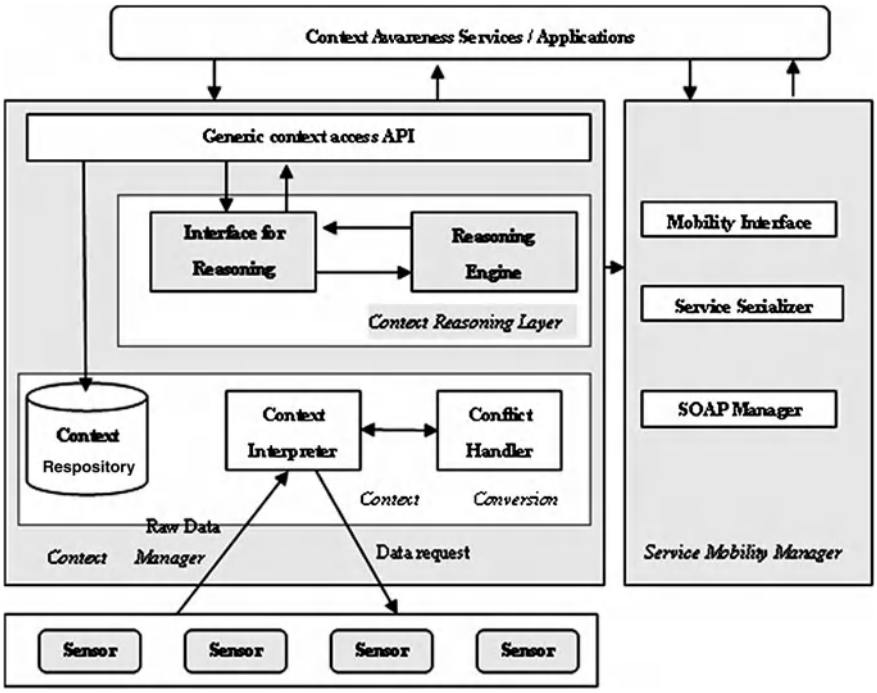


Fig. 178.2 An overview of context awareness framework

178.4.1 Context Manager

The context manager aims to collect, manage, combine the context information and forward it to the context awareness service/application and is constructed into the context interpreter, conflict hander, context repository, reasoning engine, interface for reasoning engine and generic context access API. Context interpreter and conflict handler implement the context interpreters and conversion and generate the basic context. The context repository maintains a database of context data. Interface for reasoning engine and generic context access API provide a mechanism for interaction. Reasoning engine can deducing the complex context according to the rule defined by the context-awareness application. Generally, the raw data collecting from the sensors can be inaccurate and inconsistent, therefore, it is necessary for conflict handler to detect and resolve the context conflict. The context interpreter creates an appropriate basic context using the data from the sensors as well as the context information stored in the database and provided by conflict handler. The basic context is delivered the reasoning engine through the interface for reasoning engine, and converted into a complex context using the rules. The generated complex context is sent to the context awareness service that requires the corresponding context information through the interface for reasoning engine and generic context access API.

178.4.2 Service Mobility Manager

Service mobility manager is designed to support user and service mobility and consists of mobility interface, serializer and SOAP manager. Mobility interface manages the user and service mobility. Seralizer is responsible for transforming the service status information into XML format file. SOAP manager sends the SOAP message to the destination or receive the SOAP message from a remote location. When the Mobility Interface receives a mobility request from the application, the service state switched into the mobility state. The status information prior to mobility is serialized into the XML format using the Serializer, and service mobility is converted into SOAP message and delivered to the destination. If mobility is successful, the service currently being executed is deleted from registry.

178.5 Conclusion and Future Works

In the pervasive computing environment, applications use various kinds of contexts to adapt to the evolving environments and provide smarter service. So applications must be aware of incessant context changes in their environment and adjust their behaviors accordingly. In this study, we have examined the key technologies for context awareness computing, including context acquisition, context modeling and reasoning, programming context. We also designed a framework of context awareness computing which supports the management of context information and service mobility. The framework is composed by context manager and service mobility manager. Context manager takes charge of the management of context information from the sensors, and service mobility manager manages the user and service mobility to realize the seamless mobile computing.

Currently, the work on context awareness framework delivered in this paper is designed based on OSGi and in its final development stage. Although it has proved useful in its current state, the framework still has room for improvements. In the future work, we will reduce the complexity and increase the reusability of the system architecture.

Acknowledgments This research is supported by Natural Science Research Project under Grant No.2009C520007, Education Department of Henan Province.

References

1. Abowd GA, Dey AK, Brown PJ, Davies N, Smith M, Steggles P (1999) Towards a better understanding of context and context-awareness. Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing, Karlsruhe

2. Want R, Hopper A, Falcao V, Gibbons J (1992) The active badge location system. ACM Trans Inf Syst 10(1):91–102
3. Dey AK (2001) Providing architectural support for building context-aware applications, PhD thesis, Georgia Institute of Technology
4. Devaraju A, Hoh S, Hartley M (2007) A context gathering framework for context-aware mobile solutions. Proceedings of the 4th International Conference on Mobile Technology, Applications and Systems (Mobility 2007), pp 39–46, Sept 10–12, Singapore
5. Jie S, ZhaoHui W (2006) Context reasoning technologies in ubiquitous computing environment, Lecture Notes in Computer Science. In: Embedded and Ubiquitous Computing
6. Binh AT, Lee Y-K, Sung-Young L (2005) Modeling and reasoning about uncertainty in context-aware systems, IEEE International Conference on e-Business Engineering
7. Strang T, Linnhoff-Popien C (2004) A context modeling survey. In UbiComp 1st International Workshop on Advanced Context Modeling, Reasoning and Management, Nottingham
8. Banavar G, Beck J, Gluzberg E, Munson J, Sussman J, Zukowski D (2000) Challenges: an application model for pervasive computing. In Proceedings of the 6th Annual international Conference on Mobile Computing and Networking, Boston, Aug 06–11
9. Garlan D, Siewiorek DP, Smailagic A, Steenkiste P (2002) Project aura: toward distraction free pervasive computing, Pervasive Computing, IEEE, 1(2):22–31, April–June, 2002
10. Loke S (2006) Context-aware pervasive systems: architectures for a new breed of applications. AUERBACH, Dec 7, 2006, ISBN-10: 0849372550
11. Lee S, Lee J (2006) Dynamic context aware system for ubiquitous computing environment, PRIMA, 2006, LNAI 4088:409–419

Chapter 179
Optimization of Iterations Based on Turbo Code Decoding

Yanping Xu

Abstract To reduce the iterations of Turbo code decoding without degrading the quality of image, which was used as signal source of a new scheme, is presented. Because each bit of binary system takes on varying degrees of content of other systems such as decimal system in the course of data format conversion, ensuring the iterations of high-bit and giving up the low-bits can maintain the image definition as well as decrease the sum of iterations obviously. Simulations of comparing this method with some widely used normalization techniques show that the proposed approach can achieve reducing 50% sum of iterations of decoding with higher overall BER and cannot debase the figure sharpness.

Keywords Turbo code · Iterative decoding · Optimization · Image quality

179.1 Introduction

Owing to the Fine decode performance approaching to Shannon theory limit, Turbo code has been confirmed one of the channel coding standard of 3G mobile communication system. TURBO code can improve communication system performance by using iterative decode method, but too much iterations will increase

Y. Xu (✉)
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo, China
e-mail: xubaby@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_179,

the complexity and delay of decoding and power attenuation. For this reason, Turbo code applications were limited in real-time communications. Through optimization the stopping criterion of iteration decoding, it can reduce unnecessary iterations and increase the calculation complexity at the same time [1–4]. From another perspective, each bit of the binary data converted from non-binary data in actual communication system contains different proportion of the information. With bigger proportion bit creating greater influence on transmission quality in certain SNR, iterations should be operated differently depending on proportion of the information [5].

179.2 Information Proportion of Binary Code Bit

In digital communication, each bit of the binary data converted from non-binary data contains different proportion of the information. As an example of decimal to binary conversion, we can gain 8 bits binary data after one decimal data converted:

$$D = \sum_{i=1}^{8} 2^{i-1} \times b_i$$

where D is the decimal data, $b_i (i = 1, 2, \ldots, 8)$ is bit position of the binary data.

In this equation, the bit corresponds to relevant proportion of the information as expressed below:

$$\frac{2^{i-1}}{2^8} \times 100\%$$

Each bit binary data in communication process occurs error probability is similar, but its influence on the quality of communications is different.

After the gray image that contains 200 × 256 pixels is converted to binary data, there are about 6,656 random errors on $i(i = 1, 2, \ldots, 8)$ bit in sequence. That is to say, one of the bit's BER is about 0.13 orderly and the others are equal to 0. At last, the binary data loaded noise is converted to restored image.

As shown in Fig. 179.1, the proportion of the information of bit which is from the 1st to 4th is about 3% on condition that BER is similar in each bit. The influence of the image quality is very small and the image is still clarity. The proportion of 5th and 6th bit is 6.25 and 12.5%, respectively. Similar BER will make the image clarity to be getting worse. The proportion of 7th and 8th bit increase to 25 and 50% separately, Similar BER has already made the image terrible.

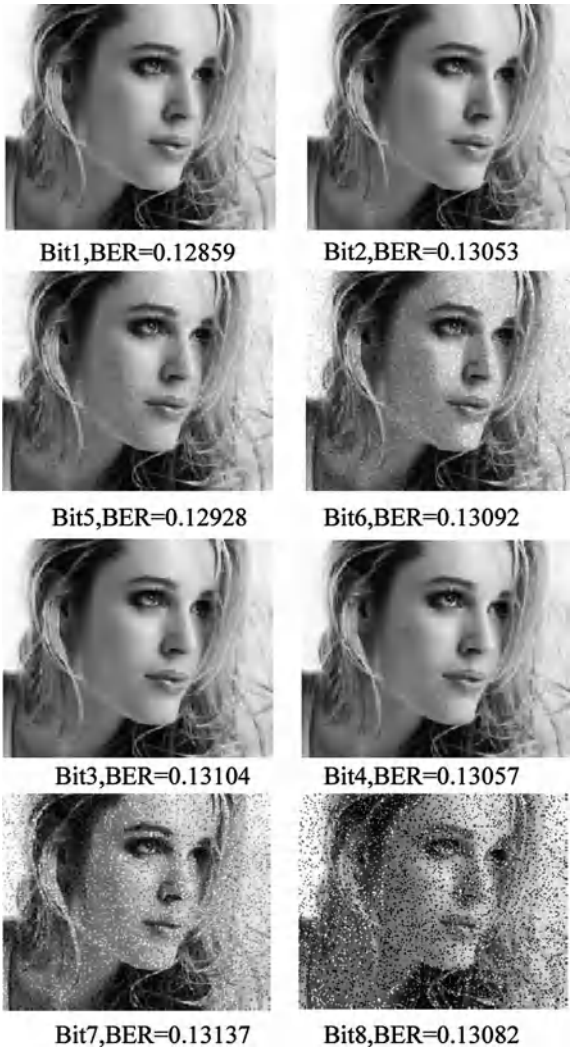


Fig. 179.1 Image clarity comparison of each bit with similar BER

179.3 Parameter Optimization and Simulation of Bit's Iterative Decoding Based on Turbo Code

According to above analysis, a method is proposed bitwise iterative decoding. Because each bit carry different proportion of the information, they should be treated differently. The more proportion of the information bit carries, the more iterations should be used, and vice versa.

Table 179.1 Comparison of different iteration schemes

Scheme	n1, n2, n3	Total iterations	Decoding image clarity
1	8, 8, 8	64	Best (Fig. 179.2b)
2	4, 4, 4	32	Worst (Fig. 179.2c)
3	3, 4, 6	32	Worse (Fig. 179.2d)
4	2, 4, 8	32	Better (Fig. 179.2e)
5	1, 6, 8	32	Better (Fig. 179.2f)
6	0, 8, 8	32	Best (Fig. 179.2g)
7	0, 4, 6	20	Worse (Fig. 179.2h)
8	0, 4, 8	24	Better (Fig. 179.2i)
9	0, 6, 8	28	Best (Fig. 179.2j)

The proportion of the information the bits from first to fourth carried is small and approximately same. So it can use the same iterations that noted as n1. The proportion of the information the fifth and sixth bits carried is medium, its iterations is noted as n2. The proportion of the information the seventh and eighth bits carried is large, its iterations is noted as n3. Obviously, $n3 \geq n2 \geq n1$.

Utilize the image data as Fig. 179.1 to carry on simulation and optimize n1, n2, n3. Parameters in this model are adopted as follow: AWGN channel, s/n = 1.5, LOG-MAP algorithm, module 2 symmetry random interleaver, 1,024 interleaver length.

In order to compare the iterations and image clarity, firstly a scheme that all bits iterate 4 is regarded as reference. Secondly adjust n1, n2, n3, at the same time total iterations remains constant. Thirdly compare all schemes and choice one, which total iterations is minimum and image clarity is best. Total iterations form is: $(4 \times n1 + 2 \times n2 + 2 \times n3)$.

Table 179.1 shows the specific scheme of iterations. Figure 179.2 indicates the image clarity after decoding. Compared with the following parameters, including bits BER, total BER, 7 $\sim$ 8th bits BER, four analysis figures are obtained, shown in Figs. 179.3 and 179.4 separately.

The image clarity after decoding is not only determined by the total BER, but also the BER of bits from fifth to eighth, especially the BER of 7th and 8th bits.

Therefore, ensure n3 is enough big, properly reduce n2 and n1 = 0, which can ensure image quality after decoding and greatly reduce the iterations and decoding time. The image quality (or image clarity) of 6th scheme and 9th scheme is the same as that of 1st scheme basically. But the iterations of 6th scheme and 9th scheme are only 50% of 1st scheme's iterations, and even less.

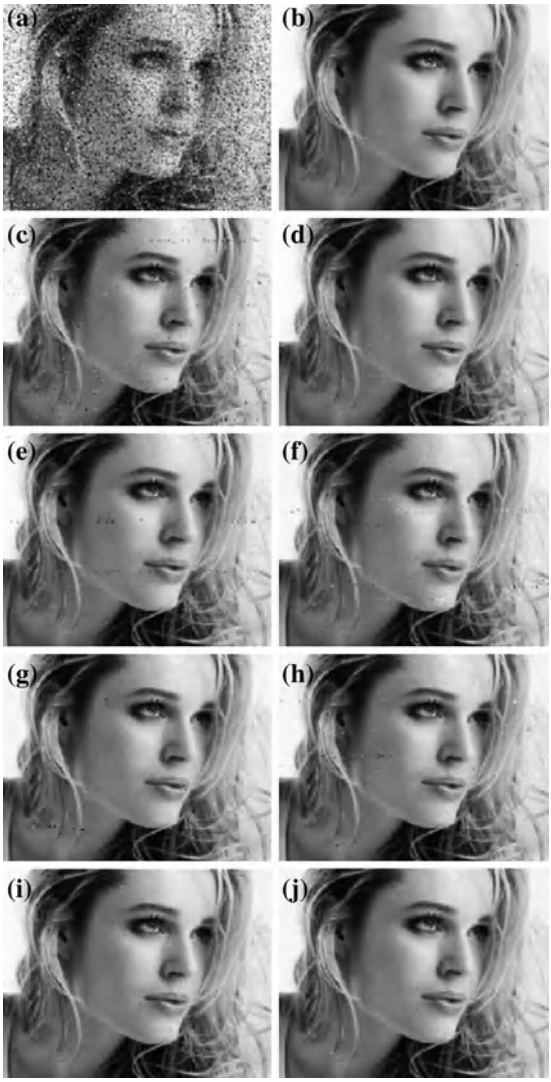


Fig. 179.2 Image clarity comparison of different iterations schemes. **a** 0, 0, 0. **b** 8, 8, 8. **c** 4, 4, 4. **d** 3, 4, 6. **e** 2, 4, 8. **f** 1, 6, 8. **g** 0, 8, 8. **h** 0, 4, 6. **i** 0, 4, 8. **j** 0, 6, 8

179.4 Conclusion

Because of the conversion of binary data and other forms of data, it makes different binary bits to carry different proportions of the information in digital communication. So during Turbo decoding in the image, more iterations in high-bits are used to guarantee lower BER and image clarity, less iteration in low-bits or data in low-bits can be decoded without iteration directly to reduce decoding iterations. Image clarity almost has no effect after decoding.

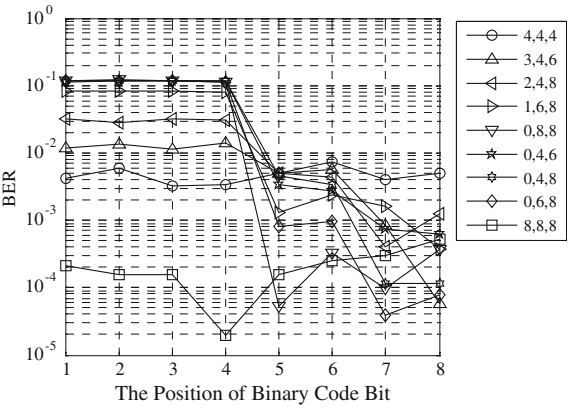


Fig. 179.3 Each bit BER of different schemes

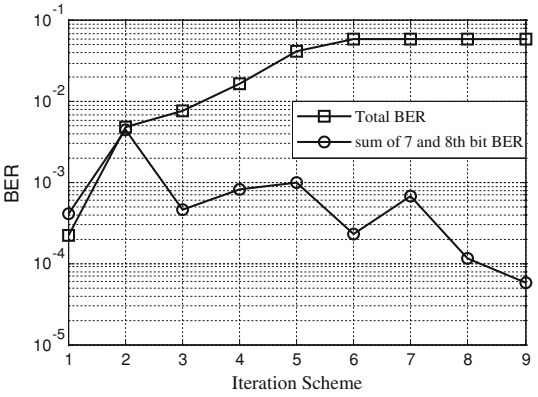


Fig. 179.4 Total BER and the sum of 7and 8th bit BER of different schemes

Acknowledgments The authors gratefully acknowledge the support provided by the Henan Polytechnic University Research Grants in 2010.

References

1. Shim BS, Jeong DH, Lim SJ, Kim HY (2006) A new stopping criterion for turbo codes. Proceedings of the International Conference on Advanced Communication Technology (ICACT), Feb, 2006: 1107–1111
2. Pang Y, Zhang C, Wang Y (2006) A simple iteration stopping criterion for Turbo code. Radio Commun Technol 32(6):25–26
3. Wang H, Chen S, Zhu C (2007) Optimal design for turbo code's iterative decoding. Comput Eng App 43(3):99–101
4. Pan P, Dong-jian H, Tian C (2009) New iterative stopping criteria for Turbo code decoding. Comput Eng App 45(27):66–68
5. Xu YP (2009) Research and analysis for the performance of Turbo-code in atmospheric laser communication system. Jiaozuo: Henan Polytechnic university

Chapter 180
Error Resilience Technique for Network Transmission in Wireless Video Surveillance System

Yu Jianguo and Shang Jianzhen

Abstract To solve the problem of error resilience in wireless network transmission, the method of adopting slice coding was proposed based on Flexible Macroblock Ordering (FMO) mode and intra-block refreshing. It combined the above three factors: slice coding might reduce the length of each packet, so it was easy for error resilience because the probability of packet loss was reduced accordingly; FMO mode made full use of spatial information for error correction and reduced bit rate; intra-block refreshing might eliminate the incorrect transmission by the loss of inter prediction macroblocks. The results of simulative experiments have proved that the method is of low time delay, less complex and high reliability, which can effectively solve the errors in wireless transmission and improve video quality.

Keywords Video surveillance · Wireless transmission · Error resilience

180.1 Introduction

Wireless video surveillance system is a small and remote digital monitoring system, which is simple to operate. It is used with network cameras and may be connected to network with or without wire. It is easy to be installed and positioned

Y. Jianguo (✉)
Zhengzhou Institute of Aeronautical Industry Management,
ZhengZhou 450015, China
e-mail: yjg@zzia.edu.cn

S. Jianzhen
Henan Business College, Zhengzhou 450044, Henan, China
e-mail: shangjianzhen@foxmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_180,

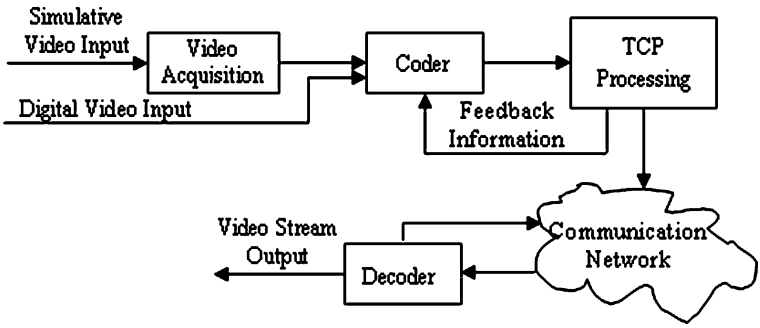


Fig. 180.1 Framework of video network transmission system

and no extra equipment is required, such as a dedicated computer or a video recorder. Either a mobile phone or a computer can be used as the monitoring terminal, and the user is able to receive alarm information and review video surveillance whenever and wherever possible. The system is stable, reliable, economical and practical and is available for fire and theft protection, security service, personnel monitoring, remote management, etc.

180.2 Framework of Wireless Network Transmission System

To study the problem of controlling errors during transmission in wireless video surveillance system, the framework of wireless video transmission system was provided for the analysis of processing video stream in the system, and the error-resilient techniques for all links in wireless network transmission were proposed.

A complete wireless video network transmission system included video acquisition, video coding, transmission control protocol processing, communication network and video decoding [1], which provided video application service in communication network based on TCP/IP with the characteristics of random time delay and packet loss. The principle was shown in Fig. 180.1.

In the system shown in Fig. 180.1, processing and transmission of the whole video stream was described as follows: at the sending end, simulative video was sampled to get digital video and then it was coded or the digital video input was coded directly to generate video bit stream for network communication that was adapted to the transmission; according to the feedback information, the available transmission bandwidth for network was estimated and the output rate of the coder was regulated adaptively (including the regulation of source and channel rates), so that the video bit stream could satisfy the limit of available bandwidth for current network transmission; at the receiving end, the received video stream was decoded, the video signals were reconstructed, the parameters of current network

transmission were calculated (e.g. the packet loss rate in transmission, etc.) and the feedback control information was sent out.

In accordance with the processing flow of video stream in the system, the error-resistance ability of video data could be improved with the help of a series of error control tools in the transmission environment of high bit-error and packet-loss rates, especially during the processes of coding and decoding. There were many error-control methods at the coding end, which were generally classified into two categories, i.e. coding algorithm and transmission control. Among them, image segmentation, selection of reference images, intra coding, data partition (applied in extended profile) and specific parameter sets in H.264. Flexible Macroblock Ordering (FMO), redundant slices, etc. belonged to the error-control in coding algorithm, while Forward Error Correction (FEC) in transport layer, RTP package, package interleaving, etc. belonged to the category of transmission control. The error-concealment technique was applied more in the video error-control at the decoding end, whose aim was to minimize the influence by transmission error as much as possible. The most simple and basic method was to make full use of time and space correlation between video data and adopt methods such as prediction, and interpolation. to recover incorrect data as much as possible [2].

180.3 Problems and Factors to be Considered During Transmission in Wireless Network

The network for wireless video surveillance system was error-prone, in which data were easy to be transferred incorrectly, such as packet loss and, bit inversion.

180.3.1 Problems During Network Transmission in Wireless Video Surveillance System

Since the error resilience technique in some video surveillance systems was very sensitive to bit error in channels, even a single primary error might result in rapid deterioration of video quality; and it was more intolerable to lose packet in IP network, which might result in the collapse of decoding terminal if severely. In particular, when the packet that was transmitting error was regarded as the reference frame, the image based on it would also go wrong, and it might attribute to the wrong transmission. In the transmission of wireless network, TCP protocol could ensure a reliable transmission, but its retransmission mechanism brought about long time delay, which was unacceptable in real-time video application. The UDP protocol could reduce time delay, but it was of high packet-loss rate without the guarantee of QOS; especially when it was applied in wireless network environment, the high rate of packet loss had exerted great influence on the quality of

video. Based on the overall consideration of above factors, UDP protocol might be used in transport layer, while in the top layer, a proper error-resilient strategy was necessary to reduce the rate of packet loss and effective measures were adopted to recover lost data, so as to minimize the visual influence by transmission error [3].

180.3.2 Factors to be Considered During Real-time Transmission in Wireless Video Surveillance System

Firstly, in wireless network, the transmission was error-prone and the packet-loss rate was high, so it was necessary to reduce the rate of packet loss in transmission as soon as possible [4].

Secondly, terminals were generally portable mobile equipment with weak computing and restricted memory abilities; therefore, it was important to notice the complexity of various usage methods in consideration of efficiency.

Thirdly, it required a high real-time capability in the video surveillance, so low time delay was also a factor to be considered in transmission.

180.4 Optimizing Error-Resilient Strategy in Wireless Video Surveillance System

For the unreliable channels in video transmission, the key was to make full use of limited bandwidth resources to achieve reliable transmission, i.e. the error resilience control technique.

180.4.1 Transmission of Parameter Sets

Both of sequence and picture parameter sets were the most important to the whole visual bit stream. Therefore, it was necessary to ensure the reliable transmission of parameter sets, which was usually achieved by out-band transmission or repeated retransmission [5]. In the application of most videos, the whole video stream would not change its properties from the beginning to the end that was to say, the sequence and picture parameters never changed in the process of connection. Consequently, all necessary parameter sets (usually one parameter set only) might be sent out as the affiliated parts of capability negotiation/inform process [6]. Only after the development of all necessary parameter sets, the RTP session should start. In the surveillance system, out-band transmission might be adopted to transfer the parameter sets to clients by the application of TCP protocol before the session.

180.4.2 Optimizing Error-Resilient Scheme During Network Transmission in Wireless Video Surveillance System

In the video surveillance system with short time delay, it was an important and difficult task to achieve the real-time transmission of video stream. Confined by the computing and storing capabilities of mobile terminals, we could only adopt error-resilient algorithm with low computing complexity and low requirement for memory.

For the requirement of low time delay, the mechanisms of feedback and RTP retransmission could not be used. Since H.264 used in the surveillance system was baseline profile, which did not support data partition, so data partition and the unequal error protection based on it could not be used, either [7]. In order to reduce data traffic in wireless environment, the method of redundant slices was not applicable, either.

In consideration of the above factors, block coding could be used during the network transmission in wireless video surveillance system, together with the error-resilient strategy such as FMO structure and intra block refreshing.

Slice coding might reduce the length of each packet, so that the probability of packet loss was reduced accordingly; and even in case of packet loss, it would be easier for error resilience [8]. There were 3 types of data block in H.264: head information, intradata block and inter data block. Head information included the type of macro block, quantitative parameters, motion vector, etc. which was the most important and should be provided with maximum protection, so these blocks were regarded as Class A in H.264 [9]. Intra data block included intra CBPs and intra coefficient, which could prevent spread of error bits effectively, and they belonged to Class B. Interdata block included inter CBPs and inter coefficient, which was the least important, so they were regarded as Class C. In this way, it was not so easy to lose important information and the capability of error resilience was enhanced.

FMO mode (interleaving mode) could reduce bit rate, as well as make full use of spatial information for error correction. FMO classified macro blocks in the same frame into different slice groups, so the error was concealed effectively by means of acquiring the information of macro blocks adjacent to the lost one from other slice groups that were receiving correctly, when the error occurred in a macro block or several macro blocks [10]. In this way, some adjacent macro blocks might be distributed in different slices, so it helped to restore the wrong image better, and avoided spreading the error as well.

Intra-block refreshing was used to eliminate the incorrect transmission by the loss of inter prediction macro blocks. By means of scattering adjacent macro blocks into different slices (groups), FMO helped the receiving end improve its property of error concealment, utilizing the spatial correlation. In spite of that, it was still difficult to avoid the network error, such as packet loss and its result of the mismatch of reference frames at the encoding and decoding ends, and the influence by accumulated spatial errors was severe, so it could only be restored rapidly by

Table 180.1 Testing environment in wireless network

No.	Bit rate (kbps)	Video length (s)	Error rate (%)	Application environment
1	64	180	0.051	PCS
2	64	180	0.017	PCS
3	128	180	0.05	PCS
4	128	180	0.02	PCS
5	128	180	0.41	PCS
6	64	60	0.93	PSS
7	64	60	0.29	PSS
8	64	60	0.32	PSS

Table 180.2 Comparison of results before and after the error-resilient strategy

Sequence	Format	Frame rate (fps)	Frames	Error- resilient strategy	RTP packets	Packets lost	Min Y	Avg Y
Carphone	QCIF	15	380	Yes	2526	32	32.08	25.01
Carphone	QCIF	15	380	No	2325	29	30.67	22.53
Forenlan	QCIF	15	380	Yes	2511	29	30.57	23.41
Forenlan	QCIF	15	380	No	2384	31	26.42	21.22

encoding the image region as intra blocks [11]. However, full Frame I would not be inserted in the application of video phone, for the instant bit rate and its introduction would increase time delay dramatically. Therefore, the coding tool FMO for error resilience, which could improve the property of error concealment at the decoding end effectively, should be combined with intra block refreshing to get a more effective scheme of mixed error resilience.

Next, the error-resilient strategy would be tested.

Table 180.1 showed several testing conditions in wireless network that was proposed in Reference, and the first condition was adopted in this paper, i.e. 64 kbps bit rate, 0.051% error rate; as the reference software of JVT's H.264/AVC, JM 9.2 was adopted for encoding and decoding, which had patched up some bugs. IPPP···IPPP (one Frame I from every other 150 frames) was adopted for testing video sequence; RTP packet was output from the encoder, the packet damage was simulated with general testing software in offline environment and the file after damage should be decoded by the decoder directly. In the decoder, if some frame could not be restored, then it would copy the previous frame directly to ensure the integrity of video sequence. The testing results could be evaluated by peak-value single-to-noise ratio of intensity.

Other parameters for the encoder and decoder included: ProflleDC = 66, LevelDC = l0, COnstralnedSetlflag = 1, NUmberReferenceFrames = 3, the bit rate was within 64 kbps, Loop Filter, Symbolmode = UVLC, out-Fi1eMode = RTP were not adopted; decoder: considering that there were many slices in Frame I, the MAX-NUM-SLICES in the program should be given a high value (200 or so).

The testing results were shown as Table 180.2.

It could be seen from the experimental results that the application of slice coding, together with the error-resilient strategy of FMO and intra block refreshing during the network transmission in wireless video surveillance system would improve the quality of video and the average intensity PSNR could be improved 2 dB or so.

180.5 Conclusion

In this paper, it proposed the error-resilient strategy of combining FMO and intra-block refreshing based on slice coding. The adoption of complex macro blocks reduced distortion that was related to error concealment, so the quantity of intra blocks in interencoding frames decreased, thus effectively improved the efficiency of encoding as well as the quality of video.

References

1. Liqing WAN, Xueming LI (2007) Error control and error concealment techniques used in H. 264. Beijing: Comput Eng Appl 16:53–57
2. Yalin SANG, Zhenghua ZHANG, Dawei LU (2008) Intra refresh technology for wireless video transmission over low rate channels. J Yang Zhou Univ (Nat Sci Ed) 11(3):35–39
3. Lisha XIE, Yuan ZHANG (2007) Joint FMO and rate-distortion optimized mode selection for H. 264 error-resilient video coding. Beijing: J Commun 28:128–132
4. Hong-bin Y, Songyu Y, Ci W (2010) Highly efficient, low delay architecture for transporting H. 264 video over wireless channel. Sign Process: Image Commun 19:369–385
5. Wenger S (2010) H. 264/AVC over IP. IEEE Trans Circuits Syst Video Technol 13(7): 645–656
6. Hammer TS, Hannuksela MM, Wiegand T (2010) H. 264/AVC in wireless environments. IEEE Trans Circuits Systems Video Technol 7(13):657–673
7. Zhiping XIE, Gengsheng Z, Guiming HE (2006) H. 264 error resilient coding scheme based on flexible macroblock ordering. Beijing: Comput Eng Appl 31:200–202
8. Zhang Y, Gao W, Lu Y et al (2007) Joint source channel rate distortion optimization for H. 264 video coding over error-prone networks. IEEE Trans Multimedia 9(3):445–454
9. Seong HJ, Jayant N (2004) An efficient error resilient technique for applications of one way video using transcoding and analysis by synthesis. In: Proceedings of the IEEE global telecommunications conference workshops, Texas, USA, pp 428–432
10. Chiou H-J, Lee Y-R, Lin C-W (2005) Content aware error-resilient transcoding using prioritized intra-refresh for video streaming. J Vis Commun Image Represent 16(4):563–588
11. Zhen HAN, Ruimin HU, Hao CHEN, Ming LI (2009) Error resilient video transcoding based on layered rate distortion optimization. Chin J Comput 32(6):1152–1160

Chapter 181
Research on Spread Spectrum Communication Anti-Multipath Jamming Based on MATLAB

Xuejun Li, Zhicheng He and Jigang Wu

Abstract Aiming at the problem that the multi-path interference is always the bottleneck which prevents the enhancement of the narrowband communication quality and the uncertainty of radio channel and channel decline are the negative factors which prevent the enhancement of the wireless communication quality, the mechanism of the spread spectrum communication multi-path interference is investigated in this paper by the research on spread spectrum communication. The spread spectrum communication of anti-multipath interference model is built and the relevant MATLAB program is compiled, the prospective simulation result is obtained, all of these will directly verify that the spread spectrum communication system multi-path interference performance is very good.

Keywords Multi-path interference · MATLAB · Spread spectrum communication

181.1 Introduction

As is known to all, in satellite communications, mobile communications, microwave communication and other wireless communications, because of radio transmission waves encounter all sorts of reflector causing reflection and

X. Li (✉) · Z. He · J. Wu
Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China
e-mail: hnkjdxlxj@163.com

Z. He
e-mail: zhichenghe861205@126.com

J. Wu
e-mail: jgwuhust@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_181,

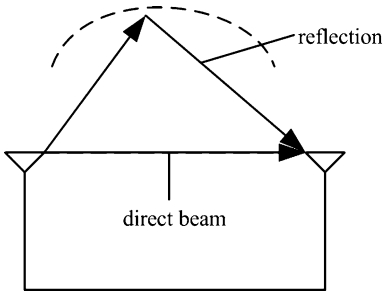


Fig. 181.1 Multi-path effect schemes

refraction, direct to the receiver's signal may take shape multi-path interference, cause communication system modem output appear error and seriously result in communication failure. But it is considered that existing conventional communication is difficult to deal with the multi-path interference, the spread spectrum communication has good inhibitory ability, which depends largely on the spread spectrum communication the pseudo random sequence related characteristics of the cycle. To spread spectrum communication multi-path interference good inhibitory ability, but no one make a holistic research for spread spectrum communication multi-path interference ability. So Therefore this chapter will studies spread spectrum communication system multi-path interference by MATLAB, established its multi-path interference model, write related procedures, in particular the simulation conditions, operation simulation program, obtained the expected related laboratory data and results, which provide a basis for others studying spread spectrum communication multi-path interference [1].

181.2 Principle of Spread Spectrum Communication Multi-Path Interference

Multi-path interference is a interference of communication, especially in the field of mobile communication common and the influence very serious interference, it belongs to a multiplier interference. Multi-path interference is due to meet in the process of spreading radio waves all kinds of reflectors (such as the ionosphere, the troposphere, high mountains and building etc.) caused reflection and scattering. In the receiver, direct path signals and reflection path signals cause group of reflection that signals of the random interference formed, as shown in Fig. 181.1.

Figure 181.2 is multi-path transmission base-band synthesis waveform. The formation of the multi-path is related of the environment of radio, terrain, the ground and so on. Due to multi-path interference signal frequency selective fading and path poorer cause delay τ, which make signal serious distortion and waveform broadening and lead to information waveform overlap. This not only can cause noise increasing and bit error rate rising, making communication quality reduced. And it may make some communication system not working.

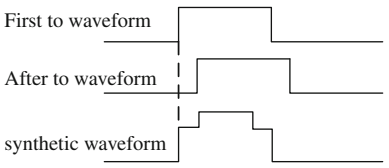


Fig. 181.2 Multi-path transmission base-band synthetic waveform

Suppose the received signal of the receiver is $r(t)$, emission signal is$A\cos\omega_0 t$, while

$$\begin{aligned} r(t) &= \sum_{i=1}^{n} \mu_i(t)\cos\omega_0[t-\tau_i(t)] \\ &= \sum_{i=1}^{n} \mu_i(t)\cos[\omega_0 t-\varphi_i(t)] \end{aligned} \tag{181.1}$$

In the Formula $\mu_i(t), \tau_i(t), \varphi_i(t)$ separately is receipt signal of first I path amplitude, propagation delay, additional phase, $\varphi_i(t) = -\omega_0\tau_i(t)$, a large of observation showed, $\mu_i(t)$ and $\varphi_i(t)$ with the change of time compared to launch the cycle of carrier is much slower, therefore upper type can be written as:

$$\begin{aligned} r(t) &= \sum_{i=1}^{n} \mu_i(t)\cos\varphi_i(t)\cos\omega_0 t \\ &\quad - \sum_{i=1}^{n} \mu_i(t)\sin\varphi_i(t)\sin\omega_0 t \\ &= X_C(t)\cos\omega_0 t + X_S(t)\sin\omega_0 t \\ &= V(t)\cos[\omega_0 t+\varphi(t)] \end{aligned} \tag{181.2}$$

$$\text{which}: V(t) = \sqrt{X_C^2(t)+X_S^2(t)}$$

$$(t) = \arctan\frac{X_S(t)}{X_C(t)}$$

$$X_C(t) = \sum_{i=1}^{n} \mu_i(t)\cos_i(t)\cos\omega_0 t$$

$$X_S(t) = \sum_{i=1}^{n} \mu_i(t)\sin_i(t)\sin\omega_0 t$$

Because $\mu_i(t)$ and $\varphi_i(t)$ holds that is slow change random process. Therefore, $V(t)$ and $\varphi(t)$ as well as $X_c(t)$ and $X_s(t)$ both for slow change random process, $r(t)$ is narrow band process. That is known by the type: the first, from the waveform looking, multi-path spread the results in the single frequency unascertained signal to modulated signals of, such signal calling declined signal. The second, Looking from spectrum, multi-path caused the frequency

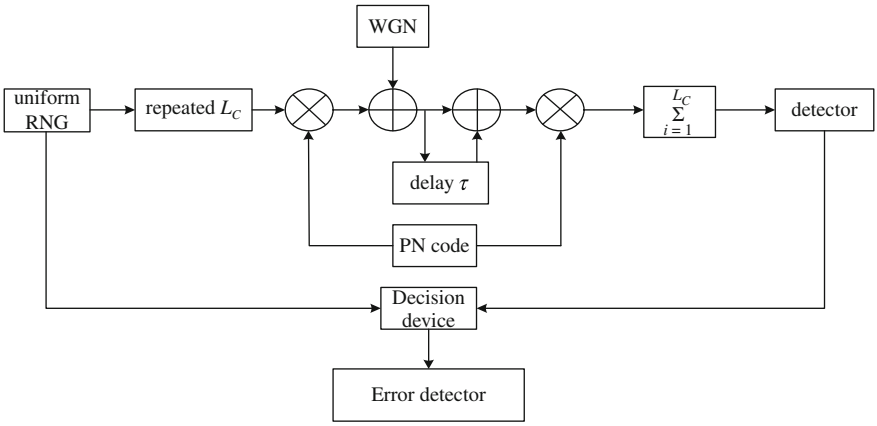


Fig. 181.3 Model of anti-multipath jamming simulation

dispersion, namely the single frequency into a narrow band spectrum. In general case, $V(t)$ obeys Rayleigh distribution, $\varphi(t)$ obeys Uniform distribution, while $r(t)$ is Narrowband Gaussian process. Multi-path spread caused fading and frequency dispersion, also may occur the frequency selective decline, namely signal spectrum some of the components a kind of attenuation phenomenon, and it related of multi-path spread delay difference. Supposing the largest delay difference is τ_m, while definition $\Delta f = 1/\tau_m$ is multi-path spread media related bandwidth [2, 3]. If the spectrum of transmission waveform is wider than Δf, while the waveform will produce clear frequency selective declining. So in order to cause no obvious frequency selective declining, Transmission wave band must be smaller than multi-path transmission medium related bandwidth Δf.

181.3 Multi-Path Interference Resistance Simulation Model

Because in white gaussian noise (WGN) channel, the bit error rate of straight expansion of BPSK way and the traditional not expansion of BPSK way modulation is the same, at the same time because this paper is studying performance of the system that spread spectrum bring, so in order to simplify the simulation model, I omit BPSK modulation [4, 5]. At the same time I assume PN code of system synchronization problems is the ideal state, that is, the PN code is completely in sync.

By the above theory foundation set up multi-path interference resistance simulation model as shown in Fig. 181.3 below. In the model, random number generator model produce a series of binary (+1) information data, each information bit repeated L_c time, L_c represents for each information bit containing code number

of PN code. Contains every bit L_c repeat sequence multiplied by the other a random number generator produces PN sequence $c(n)$, quite to information were spread spectrum modulation. As we discussed channel of additive WGN environment, so using WGN produce White Gaussian Noise. And then adding WGN that variance is $\sigma^2 = N_0/2$ and delay unit τ, let stay transmission signal after it for delay. In the modem, it makes a cross-correlation operation with the PN sequence, and sample of forming all kinds of information bit sum (integral operation). The output of the adder sent to the decision device, which make signal and threshold 0 for comparison and certain data transmitted for +1 or −1, error counter record wrongly convicted numbers by decision device [6]. The simulation model is shown in Fig. 181.3.

Parts of the simulation codes as below:

```
SNRindB1 = 0;
SNRindB2 = 5;
SNRindB3 = 10;
SNRindB4 = 12;
SNRindB5 = 15;
SNRindB6 = 17;
for i = 1:length(Tc)

  smld_err_prb1(i) = wubitlv(SNRindB1,Lc,Tc(i));
  smld_err_prb2(i) = wubitlv(SNRindB2,Lc,Tc(i));
  smld_err_prb3(i) = wubitlv(SNRindB3,Lc,Tc(i));
  smld_err_prb4(i) = wubitlv(SNRindB4,Lc,Tc(i));
  smld_err_prb5(i) = wubitlv(SNRindB5,Lc,Tc(i));
  smld_err_prb6(i) = wubitlv(SNRindB6,Lc,Tc(i));
  echo off
.....
trans_sig = sqrt(E_chip)*repeat_data.*pn_seq;

  noise = sgma*randn(1,Lc);
  a = zeros(1,Tc);
  delay_sig = [a,trans_sig];
  rec_sig = [trans_sig,a] + [noise,a] + delay_sig;
  temp = rec_sig.*[pn_seq,a];
  decision_variable = sum(temp);
.....
```

181.4 Analysis of Simulation Result

Operation simulation program, get the simulation results is shown in Fig. 181.4.

The simulation results such as epigraph above, we can see that the bigger of the signal-to-noise ratio and the better of the anti multi-path effect. From the

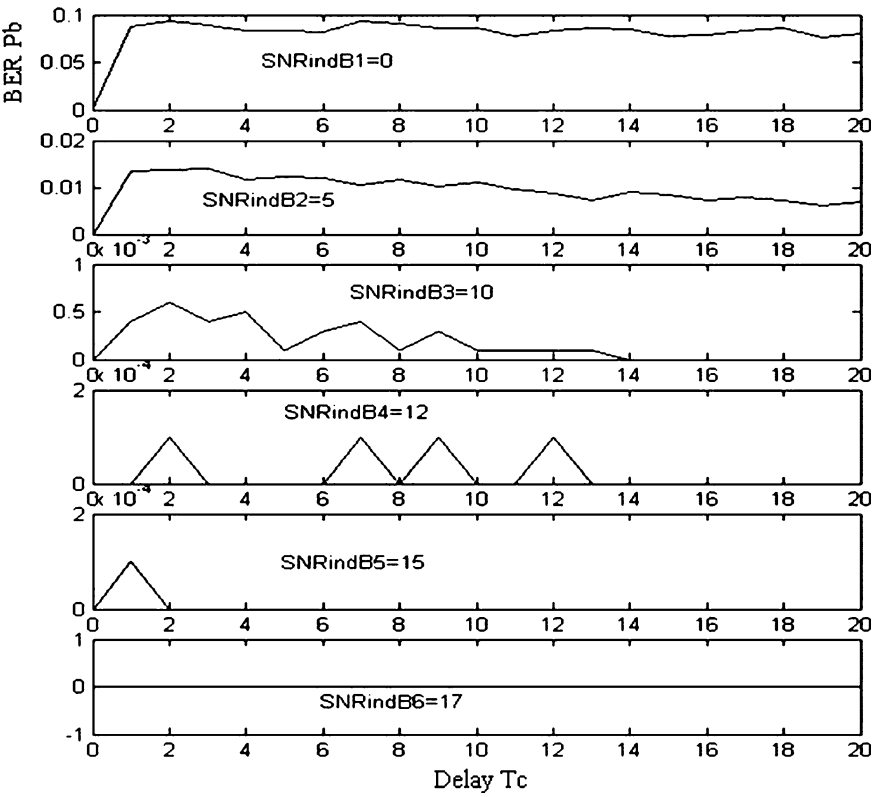


Fig. 181.4 Bit error rate curve of DS system in different multi-path interference

simulation chart known, when S/N = 0 dB, bit error rate is too big (>0.02), the system can't work normally. At the same time can see, when S/N = 5 dB, bit error rate is too litter (<0.02), the system can work normally. And we can see, From the six curves, when after the signal-to-noise ratio is 5 dB, S/N increase a litter that will be able to make at least one order of magnitude of the bit error rate decreased, and theory described in part that delay time to grow more, the lower bit error rate. Also we can see, from the curve, although bit error rate is grossly trending to down with the growth of the delay, but bit error rate is broken line in local, this should is due to that delay is corresponding to the frequency selective decline, when delay is a certain value, just the decline frequency aligned with the signal center frequency, so the bit error rate is big, also, some value of decline frequency is not aligned with signal center frequency, bit error rate is small. Therefore known that the system have good inhibitory effect for the multi-path effect.

181.5 Conclusion

Spread Spectrum Communication becomes the core of the third generation mobile communication technology because of its strong anti-interference, anti-fading and anti-multipath jamming performance. This paper mainly studies the spread spectrum communication ability of anti-multipath jamming. Using MATLAB establish spread spectrum communication anti-interference model and writing related programs, and by given simulation conditions, operation simulation program. The results show that the bigger of the signal-to-noise ratio and the lower of the bit error rate. Bit error rate is grossly trending to down with the growth of the delay. Although in communications, a part of the spectrum may be fading, but it won't course much deterioration.

Acknowledgments Financial support from Aid program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province, the CEEUSRO special plan of Hunan province (2010XK6066), Hunan University Science and Technology Achievement Industrialization Cultivation Project (10CY008), New Century Excellent Talents (NCET-08-0677), Natural Science Foundation of Hunan Province Key Project (09JJ8005), is gratefully acknowledged.

References

1. Zhao G (2009) A high rate underwater,acoustic data communications transceiver. National Defence Industrial press, Beijing, pp 24–159
2. Xu MY, Shao YB (2005) Research on anti-multi-path frequency-hopping communication system in shallow water acoustic channel. Xi'an university of electronic science and technology press, Xi'an, pp 126–253
3. Li J, Wang M, Zhao LG (2009) Anti-jamming and anti-multi path performances of generalized FH/BFSK. Comput Dev Appl 22(9):54–59
4. Zhang L, Zhen SQ (2007) Performance criteria for spread spectrum communications. Electr Transm Autom 29(3):23–27
5. Luke HD, Schoccen HD (1995) Odd-perfect, almost binary correlation sequences. IEEE Trans Aerosp Electron Syst 31(1):495–498
6. Yang ZT, Zhang KZ, Huang XF (1995) Anti-jam, anti-multipath spread spectrum OFDM system. Univ Electron Sci Technol J 24(4):360–364

Chapter 182
Application of Signal Sparse Decomposition Based on DEPSO

Mu Fusheng, Liu Chao, Deng Ling and Huang Sheng

Abstract To reduce the calculation of signal sparse decomposition, put forward to use the hybrid algorithm, DEPSO which mixed the characteristics of particle swarm optimization (PSO) and difference optimization algorithm (DE), to complete the searching of the best atoms of signal sparse decomposition based on matching pursuit (MP) algorithm. This method not only helps to get the best matching parameters with high precision, but also avoids falling into the partial optimal solution which is common to single optimization algorithm. The result of experiment shows that, compared with the single algorithm, using DEPSO greatly improved the searching speed, efficiency and accuracy of decomposition. Besides, the calculation has also dropped down dramatically.

Keywords Particle swarm optimization · Difference optimization algorithm · Matching pursuit · Sparse decomposition

182.1 Introduction

Generally, the signal sparse decomposition is very important to signal processing. In view of the complex signal's approximation problem, Mallat et al. summarized the predecessor's research results and in 1993 put forward that Matching Pursuit

M. Fusheng (✉) · L. Chao · D. Ling · H. Sheng
College of Mechanical and Electricity Engineering, Central South University, Changsha 410083, China
e-mail: csuztg@163.com

L. Chao
e-mail: liuyachao1987@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_182,

(MP) algorithm used for the general decomposition of non-quadrature signal. At present, in signal compression, denoising, coding, identifying and time–frequency distribution and extracting weak signal MP is widely used [1].

In practical application, researchers usually find the best matching atom of each MP with Newton method. However, this method is difficult to obtain the global optimal solution but also requires a lot of calculations. Thus, in recent years, many scholars did research how to reduce the calculation of decompose signal based MP. Fan Hong and others used the hybrid coding genetic algorithm to find the best matching atom and achieved certain results; based on the characteristics of atoms and FFT algorithm, Yin Zhongke et al. put forward fast effective algorithms; Yin Ming et al. used such as ant colony algorithm to search the best atoms decompose based on MP and Shu Weijie artificial fish-swarm algorithm (AFSA) [2–6].These research results are very helpful to reduce the calculation of MP signal decomposition. However, because each algorithm has its limitations, it helps to drop down the calculation of signal decomposition drastically and at the same time, the accuracy of reconstructed signal could not be guaranteed.

Learning from these findings, this chapter proposes using the hybrid algorithm, DEPSO, to complete the searching of the best atoms of signal sparse decomposition based on MP algorithm. Using DEPSO could avoid falling into the partial optimal solution which is common to single optimization algorithm. After the algorithm imports the chiasma or variation operations, the adaptability of the algorithm has made a lot of improvement. The result of experiment shows that, using DEPSO algorithm to find the best atom of signal decomposition based on MP, the searching speed, efficiency and accuracy of decomposition has been greatly improved.

182.2 MP Algorithm

Set to decompose signals $f \in H$ and H is finite dimensional Hilbert-space [7]. D is over-complete atom dictionary ($D \in H$). Firstly, select the atoms, g_{γ_0}, which matched with decomposition signal greatly from over-complete atom dictionary, D. g_{γ_0} should meet that $\left| < f, g_{\gamma_0} > \right|$ is the largest one among inner product between atoms of D dictionary and f. Therefore, f can be expressed as:

$$f = < f, g_{\gamma_0} > g_{\gamma_0} + R^1 f \tag{182.1}$$

In Eq. 182.1, $R^1 f$ is the residual amount of the first decomposition of original signal. Using the above methods, after decomposing n steps, the signal will be expressed as:

$$f = \sum_{k=0}^{n-1} < R^k f, g_{\gamma_k} > + R^n f \tag{182.2}$$

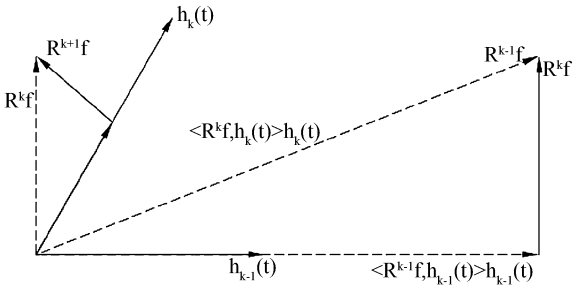


Fig. 182.1 The process of signal decomposition based on MP

Besides, g_{γ_k} meets:

$$\left| < R^k f, g_{\gamma_k} > \right| = \left| \sup_{\gamma \in \Gamma} < \mathrm{R^k f}, \mathrm{g}_\gamma > \right| \tag{182.3}$$

If the length of signal is limited and n increases infinitely, $\|R^n f\|$ will exponentially decay to zero. And then, the signal will be expressed as:

$$f = \sum_{k=0}^{n-1} < R^k f, g_{\gamma_k} > \tag{182.4}$$

From the process of decomposition based on MP, it is known that decomposed every times signal or residual signal must select the atom form over-completed dictionary, calculate and obtain the maximum value. In other words, every step of the calculation should be based on Eq. 182.3 and changed to the multiple inner product calculation in a high dimension. So this is the main reason why the amount of calculation of signal decomposition is large [8]. If take Eq. 182.3 as objective optimization function and use DEPSO algorithm, the amount of calculation will be dropped down obviously (Fig. 182.1).

182.3 Hybrid Algorithm: DEPSO

182.3.1 Particle Swarm Optimization (PSO) Algorithm

PSO algorithm is put forward by American scholars, Kennedy and Eberhart, in 1995 [9]. The basic ideas of PSO are: randomly initialize a group of no-volume and no-quality particles and take each particle as a feasible solution of optimization problems, the quality of which is measured by fitness function. Each particle will be moving in the feasible solution space, the direction and distance of which is determined by variable velocity. Usually the particles will follow the current best particle and get the optimal solution through generationally searching many times in the end. Suppose a group consisted of M particles move in a certain speed in the

D-dimensional space. At t, the position of particle i is x_{id}^t and the best position of particle moving is the individual extreme p_{id}^t. The best position of the group moving is the global extreme p_{gd}^t and the velocity is v_{id}^t. Therefore, through Eqs. 182.5 and 182.6 the velocity and position of particle at $t+1$ is updated:

$$v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id}^t - x_{id}^t) + c_2 r_2 (p_{gd}^t - x_{gd}^t) \tag{182.5}$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \tag{182.6}$$

ω is inertia weight factor and shows the inheritance of the previous velocity of particle. Smaller ω is benefit to searching globally and larger ω is good at doing locally. $c_1,\ c_2$ is acceleration constant and shows how the best position of particle or group moving take effect on the velocity of particle or group. $r_1,\ r_2$ is random numbers uniformly distributed in (0,1).

182.3.2 Difference Optimization Algorithm (DE)

According to differential evolution algorithm, at the beginning, it will get a random initial group and then across choosing, hybridizing and mutating continuously calculate iteratively. According to individual's fitness value, keep good individual, wash out inferior and lead the searching process to the optimal solution [10].

For a minimization problem max $F(x)$, DE starts from initial group x_i^t consisting of N candidate solutions, $i = 1, 2, \ldots, N$. i is group number and t is iterative number.

Among mutating, the random variation v_i^t is born by Eq. 182.7. In Eq. 182.7, $r_1, r_2, r_3 \in \{1, 2 \ldots N\}$ is random number. $F \in [0, 2]$ is weighted factor.

$$v_i^t = x_{r_1}^t + F(x_{r_2}^t - x_{r_3}^t) \tag{182.7}$$

Among hybridizing, new group $x_i^{t'} = [x_{t1}', x_{t2}', \ldots, x_{td}']$ is measured by random vector $v_i^t = [v_{i1}, v_{i2}, \ldots, v_{id}]$ and target vector $x_i^t = [x_{i_1}, x_{i_2}, \ldots, x_{i_d}]$.

$$x_{ij}' = \begin{cases} v_{ij} & \text{if } r \text{ and } b(j) \le CR \quad \text{or} \quad j = r \text{ and } r(i) \\ x_{ij} & \text{if } r \text{ and } b(j) \le CR \quad \text{and} \quad j = r \text{ and } r(j) \end{cases} \tag{182.8}$$

$j \in [1, D]$; r and $b(j) \in [0, 1]$ is the value of j time in same random number generator; $CR \in [0, 1]$ is mutation probability; r and $r(i) \in [1, 2, \ldots, D]$ is the index of randomly choosing and get a parameter at least. Among choosing, DE uses greedy strategy.

182.3.3 Signal Sparse Decomposition Based on DEPSO and MP

PSO and DE are both the heuristic algorithm based on the group. The major difference between them is the way to produce. PSO algorithm converges faster in

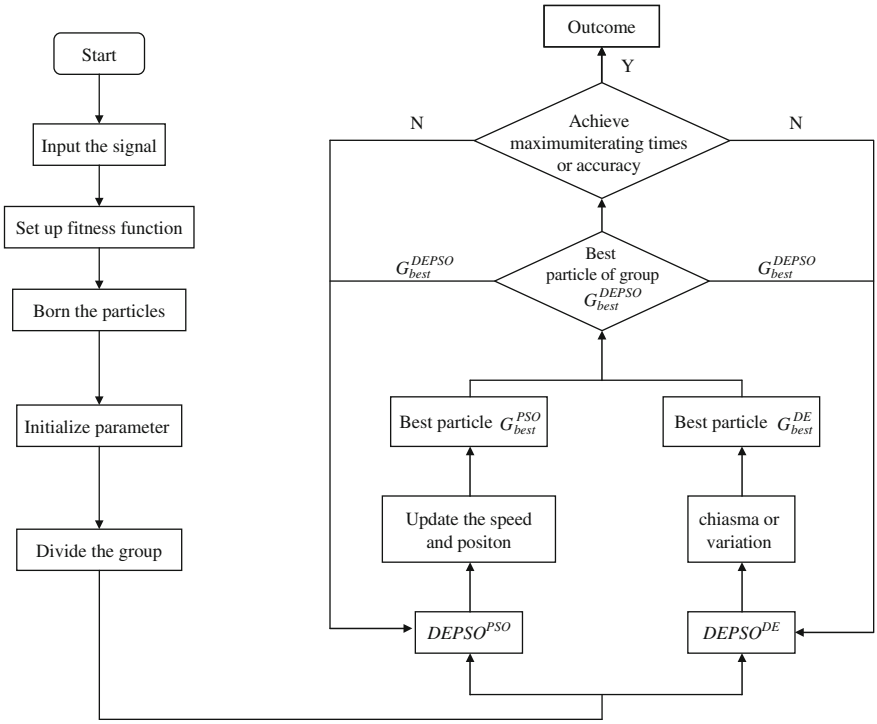


Fig. 182.2 Flow chart of sparse decomposition algorithm Based on DEPSO and MP

the early optimization but in practice application the diversity of group lost too fast and it is easy to born premature stagnation later. So it will fall into local optimum. If only use PSO to make signal sparse decomposition, the accuracy and speed couldn't be guaranteed. Mutation operator of DE is benefit to searching globally and make sure the diversity of group; crossover operator can improve the ability of locally searching and speed up the convergence rate; selection operators have certain memory ability and can keep excellent individuals. However, as the loss of diversity of group the speed of convergence will be dropped and the algorithm is easy to fall into local optimum in the later evolution [11].

In order to effectively use the advantages of PSO and DE, advise using DEPSO which max with DE and PSO to solve the problem that the calculation of signal sparse decomposition is large.

Figure 182.2 shows the process of signal sparse decomposition based on DEPSO and MP.

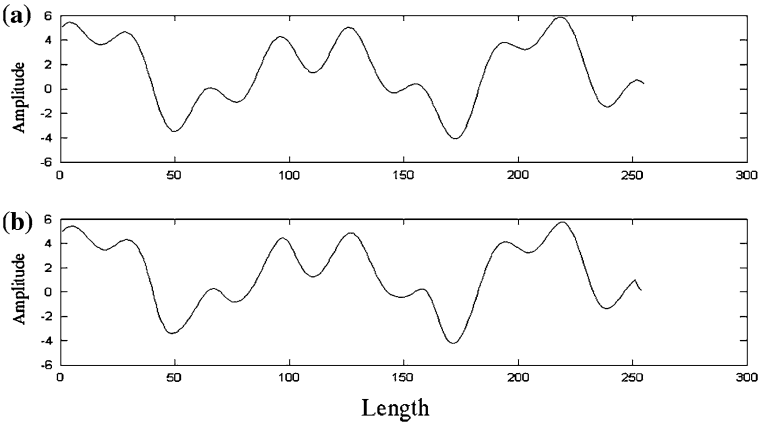


Fig. 182.3 Waveform of the original signal and reconstruction-signal

Table 182.1 Contrast of speed of signal decomposition

Algorithm	Rate
General	1
DEPSO	35

182.4 Simulation and Analysis

The length of signal simulated is 256 and its signal function is:

$$f(t) = 2\cos 10x + e^{\sin 5x} + 3\sin(\pi x + \frac{\pi}{4}) \tag{182.9}$$

The signal contains periodic attenuation vibration and cosine components and can effectively show the normal and the mechanical fault operation. Figure 182.3a shows the waveform of signal. According to the literature [7], the time–frequency atoms which is used to signal sparse decomposition based on MP is Gabor:

$$g_{(s,u,v,\Phi)}(t) = e^{\left(\frac{-sv(t-u)}{\sqrt{1-s^2}}\right)} \sin(v(t-u) + \Phi) \tag{182.10}$$

When the number of iteration is maximum or the error between the best particle and signal characteristics parameters is within the allowable range, the operation is successful. Figure 182.3b shows the reconstruction-signal after iterating across the hybrid algorithm. From the figure, see that the reconstruction-signal is nearly same with the original signal. Because DE and PSO is the intelligence algorithm based on the group and their outcome is random, sometimes the signal is not stable. However, this does not affect the actual application of the result of signal sparse decomposition.

Because the speed of signal decomposition depends on the calculation conditions (hardware and software), it is not much significance to give the actual time of sparse decomposition. The rate of signal sparse decomposition based on general algorithm and MP is defined 1. Table 182.1 shows how many times of the rate of signal sparse decomposition based on the MP and DEPSO to the general algorithm and MP.

182.5 Conclusion

This chapter proposed using the hybrid algorithm, DEPSO based on PSO and DE, for the signal sparse decomposition based on MP. And it avoids falling into the partial optimal solution which is common to single optimization algorithm. After the algorithm imports the chiasma or variation operations, the adaptability of the algorithm has been improved greatly. In addition, it is easy to avoid the premature stagnation later. It is significant to solve the problem as the speed of DE algorithm is convergence. Therefore, the calculation of signal sparse decomposition is dropped down greatly. What is more is it takes a good effect on reconstruction and extraction of signal.

References

1. Jianjun Z, Zhongsheng W (2009) Improving matching pursuit (MP) decomposition of signals using improved particle swarm optimization(IPSO) algorithm. J Northwestern Polytech Univ, pp 157–161
2. Hong F, Qingfeng M, Youyun Z (2005) Matching pursuit via genetic algorithm based on hybrid coding. J Xi' Jiao Tong Univ, pp 295–299
3. Zhongke Y, Jianying W, Jun S (2005) Sparse decomposition based on structural properties of atom dictionary. J Southwest Jiao Tong Univ, pp 173–178
4. Ming Y, Zhongke Y, Jianying W (2006) Signal sparse decomposition based on MP with AA. Comput Eng Appl, pp 47–48
5. Weijie S, Zhigang Y, Zhongke Y (2009) Signal MP-based sparse decomposition with artificial fish-swarm algorithm. Appl Res Comput, pp 66–67
6. Yu Y (2008) Signal sparse decomposition based on OMP and PSO. Microcomput Inf, pp 178–179
7. Zhang MS (1993) Matching pursuit with time-frequency dictionaries. IEEE Trans Signal Process, pp 3397–3415
8. Rui G, Huanan X, Gang H (2008) MP algorithm for signal sparse decomposition based on over-complete. Sci Technol Eng, pp 914–916
9. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of IEEE international conference on neutral network, pp 1942–1948
10. Price KV (1996) Differential evolution: a fast and simple numerical optimizer. In: Proceedings of the 1996 biennial conference of the North American fuzzy information processing society. Piscataway NJ, USA: IEEE pp 524–527
11. Lijun L, Lijing T, Ben N (2007) A novel hybrid global optimization algorithm based on particle swarm optimization and differential evolution. Inf Control, pp 708–714

Chapter 183
Transmission of Quantum Information and Experimental Measurement of Remote Operation of Quantum State

Z. G. Wang

Abstract Remote operation of quantum state is an effective way to realize the quantum communication. This paper introduces the basic principles of quantum communication and quantum entanglement. And the operation of the quantum information is essentially dealing with the quantum state. Due to the existence of quantum entanglement, the measurement of quantum state will result in the collapse of the associated quantum state. Therefore, this paper puts forward implementation scheme for remote operation and measurement of quantum information.

Keywords Quantum state · Quantum entanglement · Quantum communication · Remote operation

183.1 Introduction

The quantum information science is a new interdisciplinary subject combining quantum mechanics and information science, and it involves mathematics, physics, computer and many other fields, which opens up a new way [1] for the application of quantum mechanics and the development of information science. The transmission and remote operation of the quantum information is the important content of quantum information and also a kind of technology that may be put into practical use at present. In the quantum communication, it is required to

Z. G. Wang (✉)
School of Physics and Electronic Electrical Engineering,
Huaiyin Normal University, Huai'an 223300, Jiangsu, China
e-mail: wangzg0216@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_183,

establish not only the classic channel but also the quantum channel, in other words, the quantum entanglement among all communication sides should be realized. Quantum entanglement is a kind of important information resource. When two or more particles are under the entanglement state, an operation is made on one particle, and then other particles will have their quantum state changed immediately and correspondingly despite their position. Therefore, the quantum entanglement plays a key role in the quantum communication and quantum calculation.

183.2 Basic Theory of Quantum Communication

Classic information system uses Bit as information unit. In physics, Bit is a two-state system, which can be prepared as one among two identifiable states, just as yes or no, true or false, 0 or 1. Quantum information science takes quantum state as information unit, and this is quantum bit. And its physical carrier is any two-state quantum system, such as the two-level atom, the particle with spin of 1/2, and the photon with two polarization directions, etc. [2]. Transmission of quantum information is the transmission of quantum state in the channel. Information processing is to carry out corresponding unitary conversion for quantum state, and the information extraction, to carry out quantum measurement for quantum information system. The quantum information processing is actually the operation on quantum state. Due to the existence of quantum entanglement, the measurement of quantum state will result in the collapse of the associated quantum state. Therefore, there is an essential difference between the extraction of quantum bit information and the acquisition of classic bit information. Generally, the quantum bit is in a superposition state of two logic states:

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad |\alpha|^2 + |\beta|^2 = 1 \tag{183.1}$$

wherein, α and β are complex numbers. The possibility for $|\Psi\rangle$ to be $|0\rangle$ is $|\alpha|^2$, and the possibility for $|\Psi\rangle$ to be $|1\rangle$ is $|\beta|^2$. The quantum bit state can be on any quantum state between $|0>$ and $|1>$, and it has no definitive value.

From the classic bit, a special case of quantum bit ($\alpha = 0$ 或 $\beta = 0$) can be realized. As for a atom which only has two possible quantum states $|0\rangle$ and $|1\rangle$ including one ground state and one excitation state, it may be only under the state $|0\rangle$ or $|1\rangle$, and at this time it is corresponding to the classic bit, and also it can be under the superposition state of $|0\rangle$ and $|1\rangle$, such as the quantum state $|\Psi\rangle$ in formula (183.1), and at this time it is corresponding to the quantum bit. If we use $|0\rangle$ and $|1\rangle$ these two independent states as basic vectors, we can then stretch a two-dimension complex vector space, which is a two-dimension Hilbert space. Generally, n times qubits will form a 2^n-dimension Hilbert space, and there are 2^n states in mutual orthometric state. Generally, 2^n ground state is $|i\rangle$, and i is a n-bit binary number. The general state of n quantum bits can be expressed as the linear superposition of 2^n ground state:

$$|\Psi\rangle = \sum_{i=000\ldots0}^{111\ldots1} \Psi_i |i\rangle \tag{183.2}$$

wherein, i is respectively n times of 0 and 1, and Ψ_i is a complex number which satisfies $\sum_i |\Psi_i|^2 = 1$.

183.3 Quantum Entanglement

Quantum entanglement is a peculiar phenomenon of quantum mechanics in the quantum system with multiple subsystems, namely the measuring result of one subsystem cannot be independent of the measurement parameters of other subsystems [3], which differs from classic physics. Since the creation of quantum mechanics, there have been ceaseless controversies about the interpretation of its fundamental principles and the understanding of its basic concepts, which are mainly between the classic physicists represented by Einstein and the Copenhagen school represented by Bohr. The core controversies virtually involve the "entangled state" and its non-local interaction. In 1935, Einstein, Podolsky and Rosen put forward the well-known EPR Paradox [4], indicating the future development direction of the essential questions of quantum mechanics, and the concept of quantum entangled state was created exactly on the basis of that direction.

"Entangled state" refers to a special form of quantum state in the multiplexed system. In any representation, it cannot be written to the direct product form of two-subsystem quantum state. The four Bell bases of the two 1/2 spin particle systems are the most common two-body entangled states:

$$|\Phi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 \pm |1\rangle_1|1\rangle_2) \tag{183.3}$$
$$|\Psi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_2 \pm |1\rangle_1|1\rangle_2)$$

It is assumed that two atoms correspond to two quantum states $|0\rangle$ and $|1\rangle$, which can be in the superposition state of formula (183.3):

$$|\Psi^{-}\rangle_{12}\frac{1}{\sqrt{2}}(|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2)$$

There into, $|0\rangle_1|1\rangle_2$ mean the first atom stays in state $|0\rangle$, and the second atom, in state $|1\rangle$; $|1\rangle_1|0\rangle_2$ mean the first atom stays in state $|1\rangle$, and the second atom, in state $|0\rangle$. When the above two atoms are in superposition state $|\Psi^{-}\rangle$, we can conclude they are in the entangled state. However, we only know one atom is in state $|0\rangle$ and the other in state $|1\rangle$ and we have no idea which one is in state $|0\rangle$ or $|1\rangle$. Therefore, the two atoms are entangled.

There is a intriguing trait for the two entangled atoms. Once we conduct measurement for one particle and define its state, we can immediately know the state of the other one, no matter how far they are from each other. For example, for the two-atom system in state $|\Psi-\rangle$, if Atom 1 is measured and found in state $|0\rangle$, then, we will immediately know the state of Atom 2. Quantum entanglement is a useful information resource, which plays a key role in quantum teleportation, quantum dense coding, quantum key distribution and acceleration of quantum computing, quantum error correction and error protection, etc. At present, the most perfect entangled state prepared experimentally is the entangled photon pairs produced by means of parametric down-conversion as well as the four-particle entangled state prepared in the ion trap.

183.4 Realization of Remote Operation on Quantum State

In 1993, Bennett and other five scientists from four countries put forward a solution of quantum teleportation [5]. Its basic idea: to realize the transfer of the unknown quantum state of some object, the information of the original can be divided into the classic information and quantum information, which will be transferred to the receiver through the classic channel and quantum channel respectively. The classic information is obtained through the sender's some measurement of the original; the quantum information is the remaining information that the sender doesn't extract in measurement. After getting the above two kinds of information, the receiver can make a perfect replica of the original. In that process, the original is not transferred to the receiver and always stays with the sender. Only the original's quantum state is transferred. The sender even has no idea about the quantum state; while the receiver transforms other urstoff to be the same quantum state as the original. The quantum state of the original is damaged when the sender conducts measurement and extracts classic information. Therefore, that is a teleportation of quantum state. Since the classic information is essential to the teleportation of quantum state and the information transfer rate of classic information cannot be higher than the light velocity, quantum teleportation doesn't go against the principle of maximum light velocity of the relativity theory.

In Fig. 183.1, it is assumed that Particle 1, in the state of $|\Phi\rangle 1 = \alpha|0\rangle 1 + \beta|1\rangle\ 1$, is located at Alice; Particle pair (2, 3) is the prepared EPR particle pair; Particle 2 is located at Alice and Particle 3, located at Bob. The particle teleportation is to transfer the quantum state $|\Phi\rangle 1$ of Particle 1 at Alice to Particle 3 at Bob and to keep Particle 3 staying in $|\Phi\rangle 3$, while Particle 1 itself won't be transferred. Through the joint measurement of Bell base, the relevance of Particle 3 and 2 collapses to the corresponding quantum states. When the measuring result of particle pair (1, 2) is $|\Phi_{12}^{(-)}\rangle = \alpha|0\rangle 3 + \beta|1\rangle\ 3$, Alice informs Bob via the classic channel; while Particle 3 is in the quantum state $|\Phi\rangle 3$ to be transferred, so that the unknown quantum state of Particle 1 is transferred to Particle 3 through

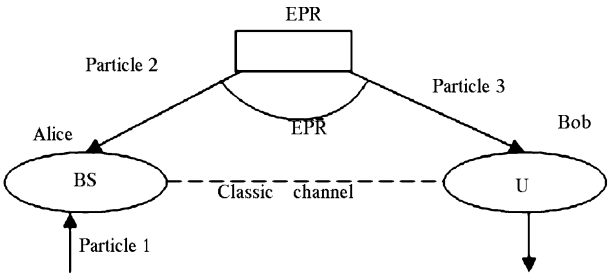


Fig. 183.1 Schematic diagram of quantum remote operation

teleportation. After Bell base measurement of Particle Pair (1, 2), the probability of $|\Phi_{12}^{(-)}\rangle$ is 1/4.

The traits of quantum teleportation: not Particle 1 itself but the quantum state will be transferred. Moreover, after the measurement of Alice, the initial state is damaged; therefore the process doesn't belong to quantum cloning [6].

183.5 Experimental Progress of Remote Operation

We can obtain the specific solution to quantum communication by means of non-orthogonal quantum state identifier. See Fig. 183.2.

When Alice, the sender, wants to send the information code "0", she will control the Optical Switch C_1 to prevent Photon v_1 from passing through the Analyzer π_1 but allow it to transmit freely along another line, so that when Photon v_2 passes through the Analyzer π_2, the space superposition state similar to $|\Phi_{12}^{(+)}\rangle$ or $|\Phi_{12}^{(-)}\rangle$ will be formed. Through observing the output of non-orthogonal quantum state identifier, Bob, the receiver, will know that Photon v_2 is in the space superposition state and decode the information to be "0". In addition, when Alice wants to send the Information Code "1", she will control the Optical Switch C_1 to let Photon v_1 pass through the Analyzer π_1 to be measured, so that the overall quantum state of incident photon pair will collapse. Therefore, when Photon v_2 passes through the Analyzer π_2, no space superposition state will be formed and it will collapse in the state on single path similar to $|\Phi\rangle_1$ or $|\Phi\rangle_2$. At that time, through observing the output of non-orthogonal quantum state identifier, Bob will know that Photon v_2 is not in the space superposition state and decode the information as "1". As a result, both sides of communication can carry out quantum communication in accordance with the above communication rules and process.

At present, lots of groups have realized the remote operation of quantum information through experiments. (1) In December, 1997, Austrian Innsbruck Zeilinger Group initially succeeded in demonstrating quantum teleportation by

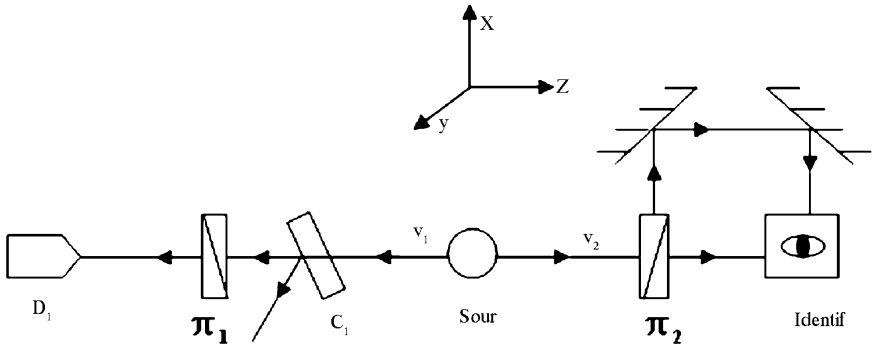


Fig. 183.2 Solution to quantum communication

means of experiment [7] and published a thesis in Nature, which was the first success in the world to transfer one quantum state from the photon in Place A to the photon in Place B through experiment. In the experiment, only the "state" indicating quantum information is transferred; while the photon, the information carrier, is not transferred. (2) At the beginning of 1998, Italian Rome Group published another experimental result of quantum teleportation in Phys. Rev. Let [8]. In the experiment, a more simple method was adopted to transfer the quantum state from one photon in the entangled photon pair to the other photon. (3) At the end of 1998, CIT Group of Professor Kimble at American California Institute of Technology realized the quantum teleportation of continuous variables. (4) In late 1998, the research fellow of Los Alamos, USA realized the teleportation of nuclear spin quantum state by means of nuclear magnetic resonance (NMR) [9]. (5) In 2001, American Shih Group successfully realized the quantum teleportation experiment by implementing Bell base measurement by means of nonlinear method in pulse parametric down-conversion. (6) In 2002, Martini Group in Rome, Italy, was reported to have realized the solution to the teleportation of vacuum and single photon entangled quantum bit in two different field modes [10]. (7) On June 17, 2002, Ping Koy Lam Research Group at ANU (Australian National University) announced that they had succeeded in carrying out the "teleportation" of information coded laser beams.

There are also outstanding achievements in the research of this domain in China. In 2003, Gao Jiangrui Group from Institute of Photoelectricity, Shanxi University, completed the experiment of invisible quantum teleportation of continuous variables. In February, 2003, Professor Pan published a thesis titled "Teleportation of Free Quantum State" in Nature [11], proposing that if the luminance of transferred quantum state is properly reduced, the quantum state can be transferred successfully without damaging the transferred state, so that the solution avoiding destroying the quantum state during the invisible teleportation is found. University of Science and Technology of China have achieved breakthroughs in the realization of the two-quantum bit composite system quantum state

teleportation experiment [12] and the quantum teleportation experiment by means of the embedded storage of photon and atom quantum bit [13], and thus realized the quantum key distribution and encryption picture transmission of 125 km commercial optical fiber from Beijing to Tianjin.

183.6 Conclusion

The remote operation of quantum state is a simple means in quantum communication. The operation is of great significance for people to understand and explore the mysterious laws of the nature. Moreover, it uses the quantum state as the information carrier, and realizes the secret quantum communication that cannot be decoded in principle by completing the transmission of high-capacity information through the transferring of quantum states. This method features a high reliability and a strong security. Besides, it can save resources and reduce the complexity of communication. If the experimental technology is promoted, it will play an important role in its application in quantum computing and communication, etc.

In addition, there are numerous technical problems to be solved from the conversion from theory to practical application, such as how to effectively produce multiparticle entangled state, how to avoid the environmental influence on the entangled state to prevent the transferred quantum information from disclosing to the environment, and how to improve the fidelity of reproduction. It is believed that, with the efforts of scientific and technical workers, human beings will certainly realize the stride leap from the classic communication times to the quantum communication era.

Acknowledgments This work is supported by the "eleventh five-year plan" planning subject of modern education technology in Jiangsu Province (2009—R—12161); and important subject of basis education research in Huaiyin Normal University (H.S.B. [2008] No. 156).

References

1. Nielsen MA, Chuang IL (2000) Quantum computation and quantum information. Cambridge University Press, Cambridge
2. Zonghai C, Daoyi D, Chenbin Z (2005) Introduction to quantum control, vol 12. Press of University of Science and Technology of China, Hefei
3. Zheng-Wei Z, Guang-Can G (2000) Quantum entangled states. Physics 29:695–699
4. Einstein A, Podolsky B, Rosen N (1935) Can quantum-mechanical description of physical reality be considered complete? Phys Rev 47:777
5. Bennett CH, Brassard G, Crepeau C et al (1993) Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. Phys Rev Lett 70:1895–1899
6. Peng X, Guang-Can G (2002) Quantum communication. Physics 31:385–391
7. Bouwmeester D, Pan JW, Mattle K et al (1997) Experimental quantum teleportation. Nature 390:575–579

8. Boschi D, Branca S, De Martini F et al (1998) Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein–Podolsky–Rosen channels. Phys Rev Lett 80:1121–1125
9. Nielsen MA, Knill E, Laflamme R (1998) Complete quantum teleportation using nuclear magnetic resonance. Nature 396:52–55
10. Lombardi E, Sciarrino F, Popescu S et al (2002) Teleportation of a vacuum-one-photon qubit. Phys Rev Lett 88:070402
11. Pan JW, Gasparoni S, Aspelmeyer M et al (2003) Experimental realization of freely propagating teleported qubits [J]. Nature 421:721–725
12. Zhang Q, Goebel A, Wagenknecht C et al (2006) Experimental quantum teleportation of a two-qubit composite system. Nat Phys 2:678
13. Chen YA, Chen S, Yuan ZS et al (2008) Memory-built-in quantum teleportation with photonic and atomic qubits. Nat Phys 4:103

Chapter 184
Thyristor Ports Voltage in Double Anti-Star-Controlled Rectifier Circuit with Balancing Reactor

Feng Zhao and Bingjiao Wu

Abstract High-power controlled rectifier circuit includes a balanced reactor with dual reverse star-controlled rectifier, and multiple rectifier circuit. In this paper, we analyze and calculate the voltage of the thyristor for the working process of a dual reverse star-controlled rectifier with balanced reactor, and MATLAB simulation, making a basis for the study of the application in the occasions which requires low voltage, high current, and provides a basis for the exact choice of power electronic devices in order to save energy and reduce costs.

Keywords Dual reverse star controlled rectifier · Thyristor · MATLAB simulation

184.1 Introduction

With the development of computer, communications technology, and the automobile industry, the high current with low voltage circuit applies more and more widely [1, 2]. In the applications of strong magnetic field devices, electroplating, electrolysis, and large-scale nuclear fusion devices and so on, often need a few to a dozen voltage V, and current of several to tens of thousands of security of the adjustable DC power supply [3, 4]. In the design of this low voltage with high-current adjustable DC power supply, double anti-star-controlled rectifier circuit with balancing reactor is always applied. Many research topics are how to improve

F. Zhao · B. Wu (✉)
Institute of Automation and Electrical Engineering,
Lan Zhou Jiao Tong University, Lan zhou, Gan su, China
e-mail: wu-bingjiao@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_184,

the circuit structure and the structure of series and parallel circuits, and very few of the calculation of the voltage across the thyristors and simulation analysis [5, 6]. The voltage across the thyristors in the double anti-star-controlled rectifier circuit with balancing reactor is analyzed and calculated in detail in this paper, which lay the foundation for selecting the appropriate thyristors and providing a theoretical basis for developing new type electronic devices which are used in the double anti-star-controlled rectifier circuit with balancing reactor specially.

184.2 Introduction of the SCR Working Conditions

Thyristor is a high-power rectifier components, its rectified voltage can be controlled. When the AC voltage supply for the rectifier circuit is constant, the output voltage can be uniformly regulated. In the rectifier circuit, in the time of the thyristor enduring the passive voltage, the input trigger pulse time is changed, the size of the control angle is changed in fact, the different available DC voltage values can be received in the load, and thus the output voltage is controlled [7].

The working conditions of SCR are as follows:

1. When the thyristor withstand voltage against passive, no matter whatever the gate withstand voltage, the thyristor is in the state of off.
2. When the thyristor withstand voltage is passive, only in the case of the thyristor gate bear passive voltage, thyristor can turned on.
3. When the thyristor is in conduction case, as long as there are some positive voltage, regardless of how is the gate voltage, the thyristor remains on, the role of the gate loss after the thyristor turn on.
4. When the Thyristor is in conduction case, when the main circuit voltage (or current) is reduced to near zero, the thyristor is off.

184.3 Double Anti-Star Controlled Rectifier Circuit with Balancing Reactor

Double anti-star-controlled rectifier circuit with balancing reactor because of the role of the balancing reactor, it can product the magnetic filed, and then it can product the reactive power, but the reactive power can only be occupied but cannot be consumed [8]. Because some of the energy is occupied by the reactive power which is not participating in the exchanging of the other energy. But it also can affect the efficiency of the system. But for this circuit, the balancing reactor is the important situation for its normal working. The balancing reactor is a large inductance, it play an important role in enabling the two rows Thyristors to turn on normally, then it can play a role in balancing the energy in the occasions which require low-voltage, high current.

In three phases double anti-star controlled rectifier circuit with balancing reactor, the accurate triggering of the Thyristors is the important base of the circuit which can work normally. In this kind of circuit, the controlled silicon triggered controlled board which is used widely have high reliability. It is power controlled device moving phase, the core of the equipment is the particularly integrated circuit which have high characteristic and high reliability armed controlled silicon triggered product using the foreign technology. The output triggered pulse has high stability and symmetry, and also cannot change with the air condition of the environment. It do not need adjust itself to have different symmetry and limited condition in using process. It can not need a scope when adjust it in practice. Because the two rows Thyristors must be turn on in turn in practice, then the circuit have high requiring of triggered pulse in the triggered circuit, so this kind of circuit need to solve the practical problem that is the triggered pulse of the two rows Thyristors in practice. Thus it is necessary to analyze the ports voltage of the Thyristors in this circuit, which is the foundation and of analyzing this circuit and the application of this circuit in practice.

184.4 Thyristor Ports Voltage in Double Anti-Star Controlled Rectifier Circuit With Balancing Reactor

Thyristor is semi-controlled type devices, although many places have to be replaced by full-controlled device [9], but its current and voltage capacity withstanding voltage is still the highest in power electronic devices, and its is reliable, so its is still more important position in large capacity applications. Double anti-star controlled rectifier circuit with balancing reactor is usually used in the occasion of large capacity, so the use of the thyristor is common in this situation. In double anti-star controlled rectifier circuit with balancing reactor, the voltage across the thyristors is analyzed using the withstanding the voltage of VT1 as an example. When the firing angle $\alpha = 0$, the voltage of VT1 is calculated and analyzed theoretically. $\alpha = 0$ is in the natural commutation point, that is, the thyristor is trigged and turn on when the voltage across the thyristors is positive. It is proposed in the ideal case, the voltage across the thyristor is zero when the thyristors is on, the thyristor bear two terminal voltage when the thyristors is off. The phase voltage can be expressed by the following:

$$U_a = \sqrt{2}U\sin(\omega t) \tag{184.1}$$

$$U_b = \sqrt{2}U\sin(\omega t - 120^\circ) \tag{184.2}$$

$$U_c = \sqrt{2}U\sin(\omega t + 120^\circ) \tag{184.3}$$

$$U_a^{'} = \sqrt{2}U\sin(\omega t - 180^\circ) \tag{184.4}$$

$$U_b^{'} = \sqrt{2}U\sin(\omega t - 120^\circ) \tag{184.5}$$

$$U_c^{'} = \sqrt{2}U\sin(\omega t - 60^\circ) \tag{184.6}$$

When ωt is between the $30^\circ \sim 150^\circ$, U_a is the maximum forward voltage, VT1 is turning on, in the conduction period, the voltage across the thyristors VT1 is 0.

When ωt is between the $150^\circ \sim 270^\circ$, because U_b has a maximum forward voltage, the tube VT3 is turning on during this period, in a dual reverse star controlled rectifier circuit with balanced reactor, because of the role of the balanced reactor, the on or off situation of VT4, VT6, VT2 which is connected to U_a', U_b', U_c' tube can be classified.

During ωt in $150^\circ \sim 210^\circ$, $U_c' > 0, U_b > 0$, the tubes connected to U_b' and U_c' are triggered and turned on, that is to say the VT3, VT2 is turning on. VT1, VT2, VT3, VT4, VT5, VT6 are common cathode connection, the withstanding voltage of VT1 is that U_a minus the smallest voltage value connected to the tube turning on, that is to say the voltage of VT1 is the value of U_a minus the smaller of U_b and U_c. When $\omega t = 180^\circ$, $U_c' = U_b'$, when $\omega t < 180^\circ$ for example, $\omega t = 150^\circ$ then:

$$U_c' = \sqrt{2}U\,sin(\omega t 60^\circ) = \sqrt{2}U \tag{184.7}$$

$$U_b' = \sqrt{2}U\,sin(\omega t 120^\circ) = \frac{\sqrt{2}}{2}U \tag{184.8}$$

During $150^{\circ} \sim 180^{\circ}$, $U_b < U_c'$, the voltage crossing VT1 is U_{ab}.

$$U_{ab} = U_a - U_b = \sqrt{6}U\,sin(\omega t + 30^\circ) \tag{184.9}$$

When $\omega t = 150^\circ$, $U_{ab} = 0$, when $\omega t = 180^\circ$, $U_{ab} = -\frac{\sqrt{6}}{2}U$ When $\omega t > 180^\circ$, if $\omega t = 210^\circ, U_b = \sqrt{2}\,U\sin(\omega t - 120^\circ) = \sqrt{2}U$ Because $U_c' = \sqrt{2}\,U\sin(\omega t - 60^\circ) = \frac{\sqrt{2}}{2}\,U$, then we can know $U_c' < U_b$, the voltage crossing VT1 is $U_{ac}' = U_a - U_c' = \sqrt{2}\,U\,\sin(\omega t + 60^\circ)$

When $\omega t = 180^\circ$, $U_{ac}' = -\frac{\sqrt{6}}{2}U$
When $\omega t = 210^\circ$, $U_{ac}' = -\sqrt{2}U$
During $210^\circ \sim 270^\circ$, VT3 connected to U_b and VT4 connected to U_a' are turning on
When $\omega t = 240^\circ$, $U_b = U_a'$,
During $210^\circ \sim 240^\circ$, when $\omega t < 240^\circ$, if $\omega t = 210^\circ$
$U_b = \sqrt{2}\,U\sin(\omega t - 120^\circ) = \sqrt{2}U$
$U_a' = \sqrt{2}U\sin(\omega t - 180^\circ) = \frac{\sqrt{2}}{2}U$,
$U_a' < U_b$ then the voltage crossing VT1 is $U_{aa}^{'}$,
$U_{aa}' = \sqrt{2}U\sin\omega t - \sqrt{2}U\sin(\omega t - 180^\circ) = 2\sqrt{2}U\sin\omega t$
When $\omega t = 210^\circ$, $U_{aa}' = -\sqrt{2}U$,
When $\omega t = 240^\circ$, $U_{aa}' = -\sqrt{6}U$,

During 240° ~ 270°, when $\omega t > 240°$, if $\omega t = 270°$,

$$U_b = \sqrt{2}\,U\,\sin(\omega t - 120°) = \frac{\sqrt{2}}{2}U$$

$U'_a = \sqrt{2}\,U\sin(\omega t - 180°) = \sqrt{2}U$,
$U_b < U'_a$ then the votage crossing VT1 is U_{ab}, $U_{ab} = U_a - U_b = \sqrt{6}\,U\sin(\omega t + 30°)$
When $\omega t = 240°$, $U_{ab} = -\sqrt{6}U$,
When $\omega t = 270°$, $U_{ab} = -\frac{\sqrt{18}}{2}U$

Analysis process is proceeding according to this rule, we can receive the voltage of the thyristor VT1 in theoretical analysis. When we calculate the voltage of VT1, wu must know the voltage value of $U_a, U_b, U_c, U_c, U'_a, U'_b, U'_c$, comparing which is bigger or smaller, then we can know the line voltage which is calculated using the two of the phase voltage. In this paper we only calculate the value of the turning points, because in the process of the middle value, the ports voltage value is content.

When we simulate the wave of the thyristor VT1 ports value we also using the same method to analyze the voltage value.

184.5 The Simulation and Analysis of the Voltage Crossing Thyristor VT1 in MATLAB Electrical Power Electronics Module

Based on the narrative above, the corresponding simulation model can be created, in which the phase difference is 120° between the two-phase voltage source, and we use the common thyristor the of factor of the thyristor's default data. The trigger pulse and the supply voltage have the same frequency −50 Hz. We can set the trigger pulse interval according to the above content. Control section contains the control parameters and pulse logic timing module; the main circulation circuit load current, while the weak signals flows in the control circuit; the control strategy with the good portability comes from the pulse generator, providing theoretical support and practical method specific for the using of the double anti-star-controlled rectifier circuit with balancing reactor.

The system simulation graph in MATLAB is in Fig. 184.1, the simulation goal is to observe that the voltage across the thyristor is consistent or inconsistent with the theoretical analysis, so the dates in this module are the default data based on the original module, the inductor voltage is setting large based on the role played by the large inductive in the circuit, which is consistent with its use conditions in practice. Figure 184.2 is the voltage simulation waveforms of thyristor VT1. It can

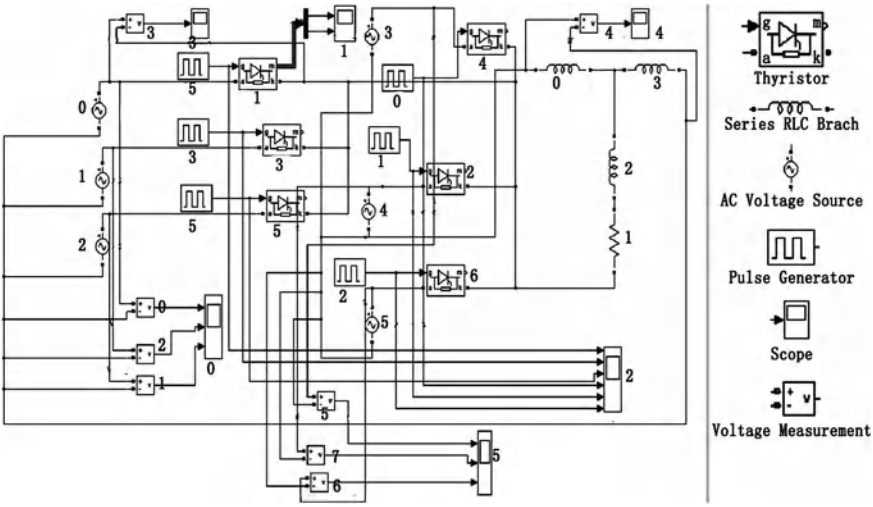


Fig. 184.1 The voltage of thyristor VT1 simulation chart

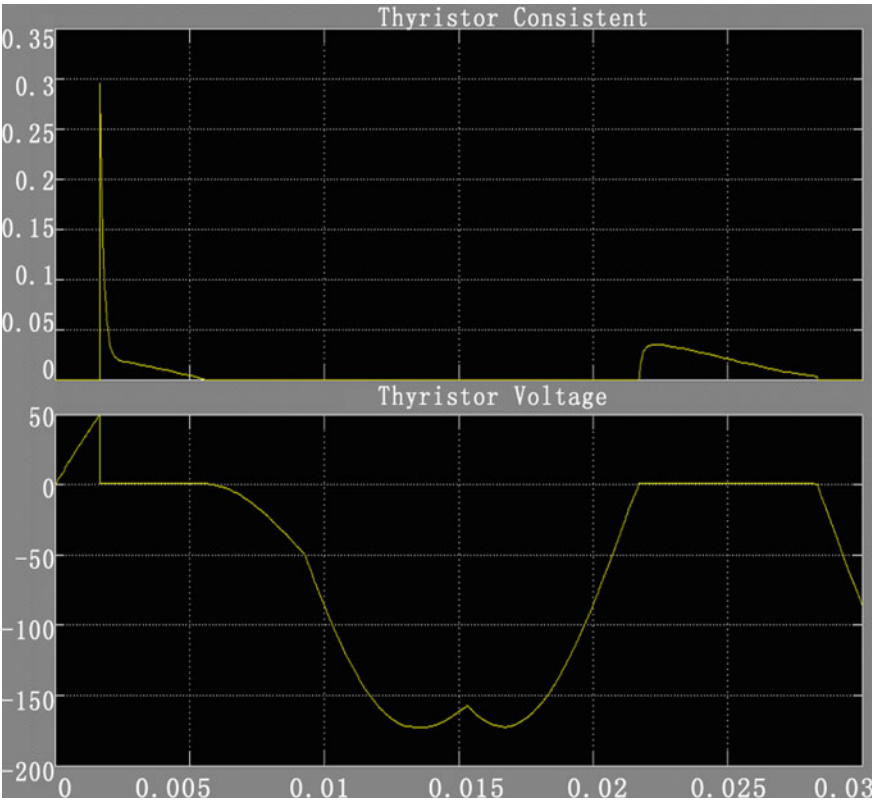


Fig. 184.2 Simulation wave chart

be seen from the figure that it is consistent with the theoretical analysis results above.

184.6 Summary

In this paper, in the double anti-star-controlled rectifier circuit with balancing reactor, the voltage on the thyristor VT1 is analyzed in theory and it is also simulated correspondingly in MATLAB, the simulation result can also illustrate the correction of theory. Thyristor as a half-controlled type device has broad application prospects in the double anti-star-controlled rectifier circuit with balancing reactor. Through theoretical analysis and simulation we can use all kinds of thyristors in the double anti-star-controlled rectifier circuit with balancing reactor, and we can select appropriate capacity and type of thyristor on the basis of theoretical analysis.

We can also see from the simulation graph that the current wave form is consistent with the voltage wave form in the circuit. When the thyristors is turning on, the voltage value of the thyristors is zero, the current value is the value of the circuit current value. Because of the characteristics of the thyristor itself, it produces a current pulse in changing current; then there are two pulses in the simulation graph. It can not only produces in this circuit, but it can also produce in other types of circuit. So we cannot analyze this appearance in detail.

References

1. Chen J (2004) Power electronics [M]. Higher Education Press set, Beijing
2. Wang Z, Huang J (2000) Power electronic technology [M]. China Machine Press set, combined with the corresponding simulation graph, Beijing
3. Xu L, X Liu (2003) Dual rectifier circuit star without the balance of parallel reactor technology [J]. The Middle Low-voltage Electrical Appliances 2(37):24–28
4. Nitsche A, Sapper J (1994) Power supply and quench protection for the WENDELSTEIN 7-x magnet. Fusion Technol pp 755–758
5. Wang Yi, Chen Xi (2003) Closed-loop space vector modulation control of matrix converter [J]. Proceedings of the CSEE 23(6):165–169
6. Burany N (1989) Safe control of four-quadrant switches [J]. IEEE IAS, 190.1194
7. Shin DH, Ch QH (1989) Improved PWM method of forced commutated cycloconverters. J IEE Proc B 136(3):121–126
8. Wheeler PW, Clare JC (2002) A vector controlled MCT matrix converter induction motor drive with minimized commutation times and enhanced waveform quality [J]. IEEE IAS[C] 12(1):466–472
9. Huber L, Borojevic D (2005) 130 roievic.space vector modulated three-phases to three-phase matrix convener with input power factor correction [J]. IEEE Trans Ind App, 31(6): 1234–1246

Chapter 185
Research for Management System of Mining Mechanical and Electrical Equipment Based on Wireless Network

Hui He and Jizheng Sun

Abstract With sustained and rapid development of China's economy since the beginning of reform and opening up, the demand for all kinds of energy resources has been increasing greatly, especially coal, as one of the basic energy, the demand for which is most obvious. In 2009 our county's coal output reached 3.05 million tons, accounting for 77.2% of primary energy. However, the management system of electromechanical equipment is still not perfect, which severely restricts safe and highly effective coal production. It is very tedious to get information about electromechanical equipment, for in nowadays specialized management personals check, maintain, and register the equipment in the fixed place and at the fixed time, which results in heavy workload and low efficiency. To improve the supervision and efficiency of electromechanical equipment maintenance, to prevent electromechanical equipment omissions caused by human, to guarantee the electromechanical equipment to run normally, efficiently and safely, to reduce electromechanical equipment's influence upon production, to improve the production efficiency and the modern management level and to realize electromechanical equipment's maintenance management modernize and informationize, we design and launch an information dynamic management system of mine electromechanical equipment that based on wireless network.

H. He (✉)
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: herunhaocili@163.com

J. Sun
School of Computer and Information Engineering,
Jiaozuo Teachers College Profile, Jiaozuo 454003, China
e-mail: sjzjzsz@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_185,

Keywords Internet of things · Mine electromechanical equipment · Dynamic management system · Information network

185.1 Introduction

185.1.1 The Basic Logic

Research and develop a set of equipment management information system consisting of website system and database, handheld device, wireless Internet module and information acquisition module.

Website system and database is used to storage equipment information and to acquire equipment and personnel information by visiting Websites according to different permissions. Handheld device is used to identify equipment and accept datum transmitted from node and also modify and upload datum. Wireless Internet module is to establish various datum acquisitions and send them to handheld devices. The last one, information acquisition module is for equipment information collection.

185.1.2 Overall Design Idea

According to equipment requirements, various sensor detections are installed. The sensor detection transmit datum to wireless nodes, and then via wireless protocol wireless nodes transmit them to handheld devices, which can display datum, that is, to display various states of equipments. Handheld devices can also upload datum to upper computer database, while upper computer can check equipment information through the database. And at the same time handheld device achieve sign-in function by using radio-frequency technique.

185.2 Website System and Database

185.2.1 System Composition Structure

As shown in Fig. 185.2, taking database configuration files as pivot, database configuration file is composed of left side user interface application (or applications) and right side database program (or database). Application program adopt techniques of ASP.NET 2.0 to write program under the program design environment of Visual Studio 2005. The main programming languages include HTML

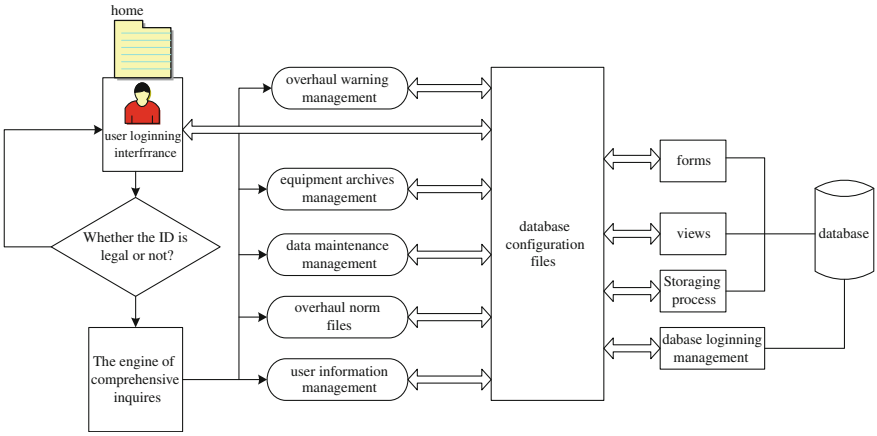


Fig. 185.1 The diagram of the system principle

language, VB 2005 language, JAVASCRIPT language and XML language, etc. Database uses SQL SERVER 2000 for constructing, mainly including constructing forms, writing views and storaging process for application program to call, and so on. At the same time, in order to be able to make application program log on to data system successfully, the users' access rights of APS.NET 2.0 need to be set in the database landing settings, as the database landing management setting shown in Fig. 185.1.

185.2.2 The Server Application Design

What is shown in Fig. 185.2 is the whole application structure. The system consists of seven big function modules. They are user login, comprehensive query engine, overhaul warning management, equipment archives management, data maintenance management, maintenance norm file and user's information management.

185.3 Handheld Devices

It uses VB 6.0 to write message server. Server program is automatically activated at startup and search the database content. It sends corresponding early warning information to relevant responsible persons according to management personnel's setting information. All these will improve the transmission of early warning information to a great extent. It will improve the management efficiency by using

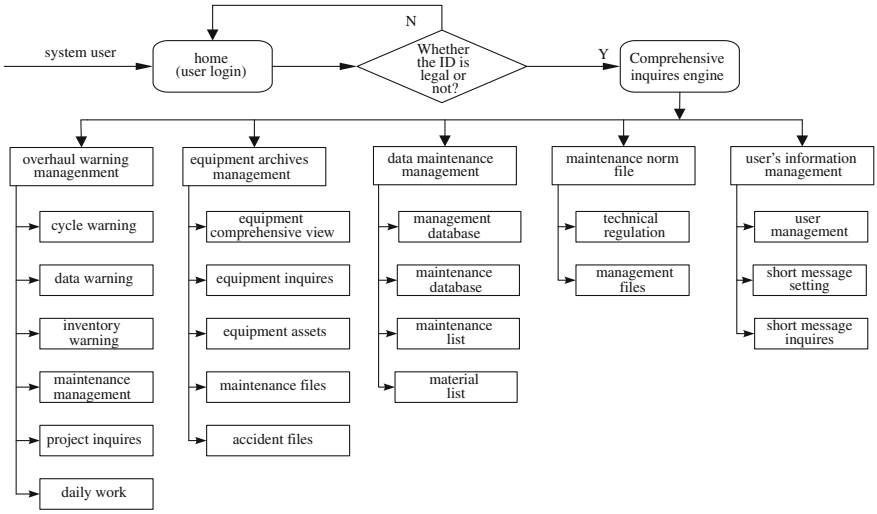


Fig. 185.2 System function structure chart

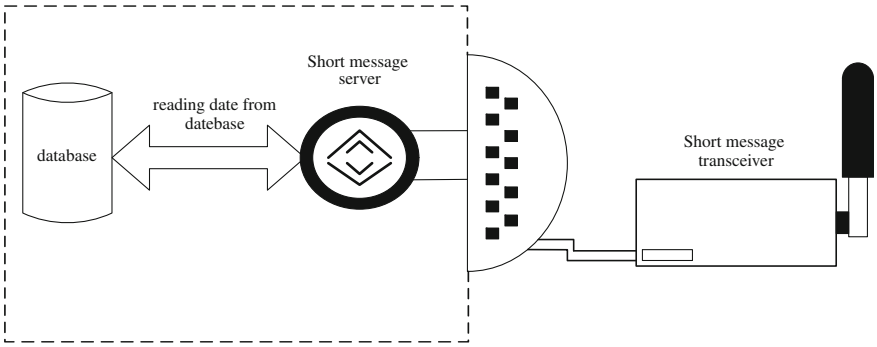


Fig. 185.3 Principle of short message server

handheld device to manage equipment information. Besides, the use of radio-frequency card, which as equipment identification method, will makes the information more accurate and convenient.

What is shown in Fig. 185.3 is the message server's working principle and design framework. The one in dash–dotted frame is the software part, showing how message server and database exchange datum. Meanwhile, message server controls message transceiver through computer's serial communication port, so as to achieve message transmission.

185.4 Wireless Network Modules and Information Collecting Modules

The handheld equipment can identify the mechanical and electrical equipment, and send commands to the microcontroller on the device. Through the sender, the microcontrollers can read the information on the mechanical and electrical equipment. The data signal can be transmitted to the handheld device through the wireless transmission. And through WIFI, the data signal can be further transmitted to the Internet. People with different permissions can have access to the website. It can also send messages to administrators to remind them.

185.4.1 Zigbee Topology Analysis

Zigbee is a kind of wireless network technology with the characteristics of short, low power consumption, and based on the basis of the IEEE. 802.15.4. The complete protocol stack (PC) of ZigBee consists of Application layer (APL), network layer (NWK), Media Access Control (MAC) and physical layer (PHY).

Zigbee network supports two kinds of physical equipment, namely Full Function Device (FFD) and Reduced Function Device (RFD). Generally speaking, FFD can act as network coordinator, can communicate with any equipment communication; RFD can't act as network coordinator, and only communicate with FFD. In any Zigbee network, there are network coordinating function equipments and (RFD) existing. The former communicates with any equipment, and compatibility sets up network communication. The latter, because of its simple structure, some just contains the IEEE standard agreement, has litter application in upper layer.

The above two kinds of physical equipment constitute the Zigbee network, two styles common network topology structure of which are star network topology structure and peer to peer network topological structure. In the star network topology structure, a FFD communicate with some other RFD. And in the peer to peer network topological structure, as long as the network equipment is within the communication scope of any of the other equipment, can directly communicate with the others, and doesn't through the coordinator for transshipment. A peer to peer network topology structure can use very complex network forms to be realized. The messages sent by any equipment in the network can reach any other routing equipment through Multi-hop routing.

185.4.2 Network Structure Design

According to the above analysis, the star network is constructed relatively simply, and is easy to realize up higher stability. Therefore, we design Multi-star network topology structure, which is shown as Fig. 185.4.

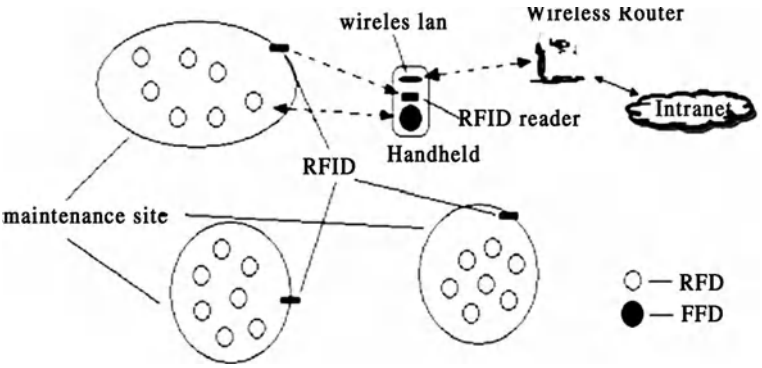


Fig. 185.4 The structural drawing of Zigbee network topology

For access to the site for the unit division, we set several Zigbee RFD nodes in each network area. Each network FFD is placed in handheld mobile network terminal (namely handheld devices). At each maintenance site RF electronic label is set up to mark this position. In the hand-held devices, RFID read module is configured. And when the personnel arrive to a maintenance spot with hand-held devices, RFID read module reads scene local information, and at the same time other related information is recorded. After the hand-held devices confirm the corresponding location, FFD is derived to send up RF signal, waking up corresponding RFD to set up star network. Then the handheld devices read data information coming back from every RFD, and complete the collection of scene data information. Then again, each RFD gets into sleeping state to save battery power. When the handheld device is bringing to the coal well's mouth, it will inspect the wireless network signal, and transfer data to the local area network. The network is design simply and flexibly configured, which could solve the wireless signal transmission path and solve the problem of the wireless signal in coal well.

185.5 Conclusion

Developing the dynamic information management system of mining mechanical and electrical equipment is of significant importance. This system adopts APS NET 2.0 and database technologies, through the network management, it uses electrical equipment to maintain and manage the database technologies, which has optimized the maintenance and management system of the coal mining electrical and mechanical equipment, and has achieved the network construction of the equipment. What's more, it has also improved the coal mining electrical and mechanical maintenance and management system.

It has improved the working efficiency of the management stuff and has provided a guarantee for the production of the mining district. The system uses the method of real-time calculation and updating to achieve an early warning mechanism to effectively promote the supervision and management of administrators. Thus, it ensures the implementation of maintenance.

Using the handheld devices to manage the information on the device has effectively improved the management efficiency. Using the network of wireless sender technology as a means of information collecting can computerize the information on the low-level device.

Bibliography

1. He LS, Zhang DP, Lai H (2005) Reconfigurable virtual instrument system. Chin J Mech Eng 41:78–81 September
2. Wang X, Ding EJ, Dai J (2008) Coal mine strao skid protecting and monitoring system based on WSN. Instrum Tech Sens 8(12):32–34
3. Yang Z, Li J (2008) Application of wireless sensors networks in a mine monitoring system. Min Process Equip 14(4):24–26
4. Liu P, Zhen L, Hu B (2009) Wireless sensor network system used in underground gas monitoring. Ind Mine Autom 8(6):8–10
5. Zhang ZY, Guo TL, Li ZJ (2009) Tracing system for underpit operators based on chip CC2431. Commun Technol 21(3):111–113

Chapter 186
Stripes Modulating Photograph Technology Researches to Obtain the Shape Information of Fast Moving Objects

Honge Luo, Jinliang Gu, Ping Chen, Yan Xia and Baoming Li

Abstract The stripe modulating photograph technology for shape information of fast moving objects was supposed. This technology resolves the problem of take stripes modulating images of fast moving objects with high quality because of the moving blur. This system consists of short-pulsed laser grating modulate illumination source, CCD cameras with high resolution and the timing control system. The moving blur of the fast moving objects in the images was decreased efficiently for the laser stripe illumination source with about 12 ns half-wavelength. Then the stripe modulated image of fast moving objects with high quality was obtained. This technology makes it possible to get shape information of the fast moving objects with little distort.

Keywords Stripe modulate · Short pulsed laser · CCD cameras with high resolution

186.1 Introduction

In order to research the deconstruction of the projectile with high velocity, and to measure the three-dimensional (3D) information this method is use. There are contact and non-contact ways in 3D measurement system. The contact 3D measurement ways have the advantage of high precision and can measure the objects with complex shapes, and its disadvantage is that it can only measure the static

H. Luo (✉) · J. Gu · P. Chen · Y. Xia · B. Li
Science and Technology on Transient Physics Laboratory,
Nanjing University of Science and Technology,
Nanjing 210094, Jiangsu, China
e-mail: xiaowufengyun@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_186,

targets. Obviously it is not suitable for the moving projectile's 3D measurement. The non-contact 3D measurement ways are mainly based on optical methods such as Fourier transform profilometry (FTP) [1–3], phase measuring profilometry (PMP) [4], spatial phase detection (SPD) [5] and other 3D reconstruction method etc. 3D reconstruction based on FTP has quick data acquisition and can get the 3D information of the object by unwrapping the phase from only one stripe image. The 3D reconstruction based on PMP has high precision, however, it needs three or more digital phase shift which is can't be used in the fast moving objects' transient 3D reconstruction. The 3D reconstruction based on SPD has fast process speed, but when the objects are defective, the phase of the defect changes too large which will result in a lot of errors. Therefore, using FTP technique to get the 3D information of the fast moving object is more feasible. Currently the FTP are mainly used in the 3D reconstruction of the static target or target with slow velocity such as people's face, plaster casts, insect wings etc. [3, 6]. The researches on 3D reconstruction of moving objects with high velocity based on FTP was rarely reported for the difficulties in obtain of clear stripe image of the projectile because of the moving blur. Till now there was no photograph system which can capture the stripe image of the fast moving objects with high resolution which can satisfy the demand of the FTP process in 3D reconstruction. The shadow photograph systems [7] can capture the images of the fast moving objects with very high solution, however, the stripe images of the objects can't be captured by these systems and no 3D information can be processed. The high speed cameras of the ultra fast speed cameras can capture the images of high speed moving objects too, but the solution of always was low which was not high enough for the demand of 3D reconstruction. In order to solve this problem, the stripes modulating photograph technology based on the CCD photograph technology with high solution and the grating stripe modulation technology with laser illumination was suggested. When the projectile flies into the field of view of the photograph system, the grating was illuminated by the laser then the stripes are generated in the field of view. The stripes will be modulated by the shape of the projectile and the modulation stripe information will be captured by the CCD camera with high resolution. For the laser has very short pulse width which will frozen the stripe image of the fast speed moving objects and generate very little moving blur. The stripe images captured by this system have high resolution and solving the moving blur problem which can satisfy the demand of the 3D reconstruction based on FTP.

186.2 The Principal and the Structure

The structure of the stripes modulating photograph is shown in Fig. 186.1.

In Fig. 186.1, the optical axis of the CCD camera and the optical axis of the grating projection set with laser illumination intersect at the center of the field of view O with an angle. The stripes generated by the grating projection set with laser

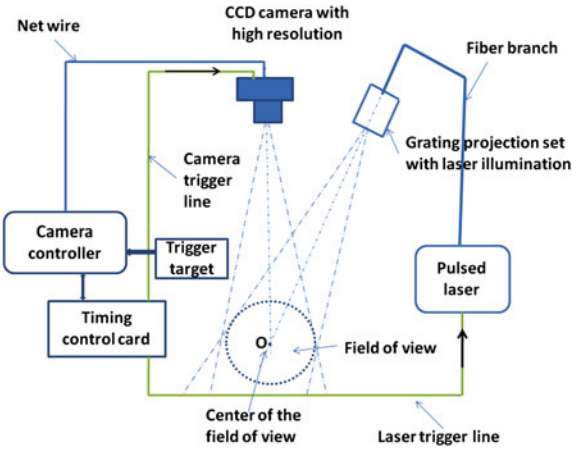


Fig. 186.1 The structure of the stripes modulating photograph system

illumination illuminate all the field of view equably which supplies the light source for the flying object's photograph.

A trigger signal was generated to trigger the schedule controlling circuit when the moving object touches the trigger target. Delay time of the schedule controlling circuit was set by user interface in camera controller to control the laser source to spark and the second time and the CCD cameras to start exposure, respectively. At the end of the delay of the CCD camera, the CCD camera starts to exposure. During the exposure of the CCD camera, at the end of the delay of the laser trigger signal, the laser pulse was generated and was translated into the grating projection set by the fiber branch. The grating was illuminated by the laser pulse and grating stripes was projected to the shape of the flying object with high speed. The even stripes were modulated by the shape of the moving object. Meanwhile the modulated stripes including the 3D information of the shape of the moving object with high speed will be captured by the CCD camera. Finally the image captured will be translated to the camera controller by high speed cable to be processed.

186.2.1 The Calculation of the Laser Pulse

If the Laser pulse energy is W (μJ) and the laser spot diameter is d (cm), then the surface energy density of radiation E (μJ/cm^2) is:

$$E = \frac{4W}{\pi d^2} \tag{186.1}$$

If the reflection coefficient of the reflective screen is ρ, and the CCD lens numerical aperture is D/f', according to the formula of surface illumination, the energy density of CCD E_0 can be calculated as:

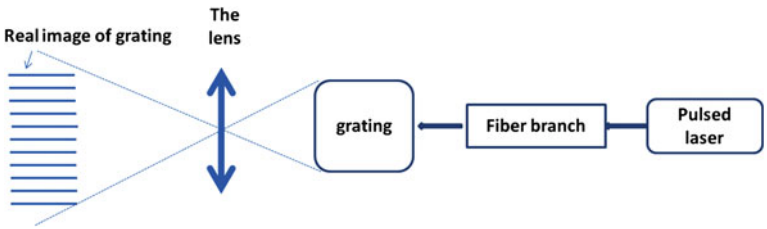


Fig. 186.2 The diagram of the grating projection set with laser illumination

$$E_0 = \frac{\pi}{4} \times \left(\frac{D}{f'}\right)^2 \rho E \tag{186.2}$$

From (186.1, 186.2) laser light energy is:

$$W = \frac{E_0 d^2}{\rho} \times \left(\frac{f'}{D}\right)^2 \tag{186.3}$$

If the field of view of the photograph system is $300 \times 300\ \text{mm}^2$, then the diameter of the illumination round area should be

$$d = 300\sqrt{2} = 424\,\text{mm} = 42.4\,\text{cm}$$

When $E_0 = 0.05\ \text{uJ/cm}^2$, satisfactory image quality of CCD can be obtained.

If the diffuse reflection coefficient ρ of the object is 0.4, and D/f' =1/8, then from Eq. 186.3 the energy of the laser pulse is about:

$$W = 14,382\,\mu\text{J} \approx 14.4\,\text{mJ}$$

Therefore, the pulsed laser with the pulse energy from 10 to 80 mJ adjustable was chosen. The during of the laser pulse was 12 ns which guarantee the small moving blur of the flying objects with high speed and the wavelength of the laser was 532 nm.

186.2.2 Grating Projection Set with Laser Illumination

In order to get clear periodic stripes of the entire field of view, the photograph principle of the lens was used. The diagram of the grating projection set with laser illumination was shown in Fig. 186.2. In the projection set, the laser pulse was translated to illuminate the grating through the branch of fiber. Then the real image of the illuminated grating was formed in field of view. Thus the clear periodic stripes were obtained.

The grating projection set with laser illumination was shown in Fig. 186.3.

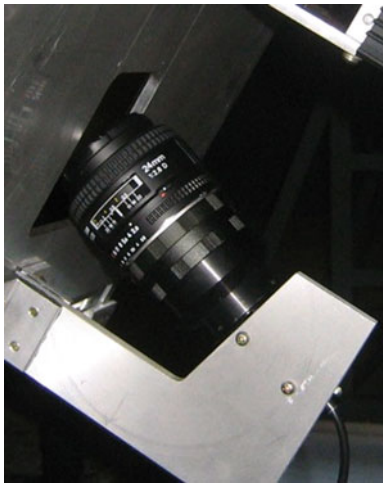


Fig. 186.3 The grating projection set with laser illumination

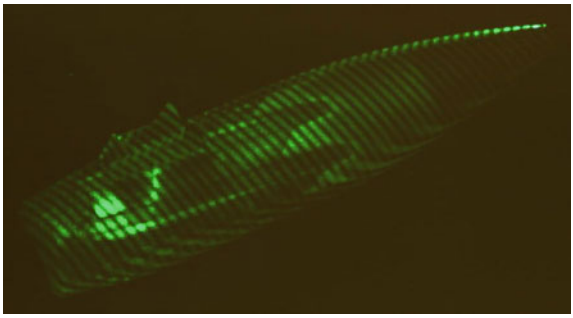


Fig. 186.4 Image captured

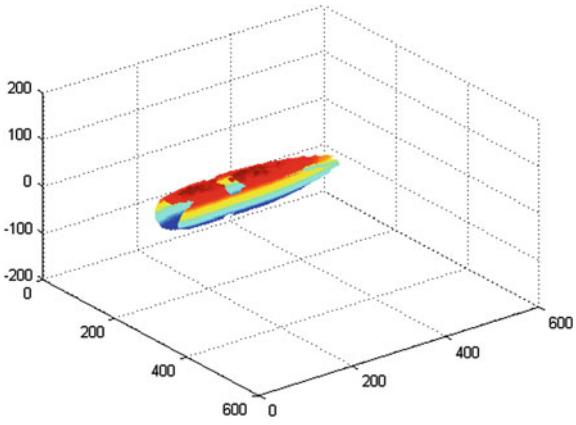


Fig. 186.5 3D information if the 14.5 projectile

186.3 Results

The system was used to capture the modulation stripe image of projectile fired by the 14.5 trajectory gun in flying with the speed about 1,000 m/s in experiments. The image captured was shown in Fig. 186.4.

The captured image has clear modulation stripes and can satisfy the command of the 3D reconstruction based on FTP. Three-Dimension information of the surface of the flying projectile processed from the modulation stripes images in Fig. 186.4 was shown in Fig. 186.5.

186.4 Conclusions

Stripes modulating photograph system applied to the 3D reconstruction of high speed objects was suggested and accomplished based on grating projection set with laser illumination and CCD technology with high resolution. Experiments were done to capture the modulation stripes of the 14.5 projectile with the speed of 1,000 m/s and the stripe image can satisfy the command of the 3D reconstruction based on FTP because of the little moving blur. The software work will continue to improve the precision of the 3D reconstruction of flying objects with high speed.

Acknowledgments This project was supported by the science and research plan of Nanjing university of science and technology (no. 2010ZYTS093), and supported by National laboratory funds.

References

1. Takeda M, Mutoh K (1983) Fourier transform profilometry for the automatic measurement 3-D object shapes. Appl Opt 22(24):3977–3982
2. Lin J-F, Su X-Y (1995) Two-dimensional Fourier transform profilometry for the automatic measurement of three-dimensional object shapes. Opt Eng 34(11):3297–3302
3. Guo H, Huang PS (2007) Absolute phase retrieval for 3D shape measurement by the Fourier transform method. Proc of SPIE 6762:676204-1–676204-10
4. Srinivasan V, Li HC, Halioua M (1984) Automated phase measuring profilometry of 3-D diffuse object. App Opt 23(18):3105–3108
5. Toyooka S, Tominaga M (1984) Spatial fringe scanning for optical phase measurement. Opt Comm 2(51):68
6. Spagnolo GS, Guattari G, Sapia C et al (2000) Contouring of artwork surface by fringe projection and FFT analysis. Opt Lasers Eng 33:141–156
7. Honge L, Jinliang G, Ping C et al (2010) A velocity measurement method based on digital orthogonal shadow photograph system in ballistic range. The J Chin Univ Posts Telecommun 17(1):127–130

Chapter 187
Study on the Construction of Distributed Collaborative Product Design Platform Based on the Internet of Things

Jiang-bo Zheng and Hong-Xue Yan

Abstract To enhance the response to customers' demand and the efficiency of designing product in networked collaborated environment, a new concept for product design, which emphasizes the supply chain view is put forward based on the Internet of Things (IOT). In accordance with the requirements of collaborated design, the relative application technologies of IOT are deeply analyzed. To optimize and carry out the product design flow through the network, a framework of Distributed Collaborative Product Design Platform established on the IOT is given. The fundamental structure, feasibilities, functions and key technologies of this platform are described in detail. All of these researches provide significant practical references for enterprises to build the related product design platform for improving the design quality.

Keywords Distributed collaborative product design platform · Internet of things · Supply chain view

187.1 Introduction

The development of information and network technologies has greatly promoted the networked collaborative product design through the network, which makes designers from various places to communicate with each other via information networks so as to contribute their skills and knowledge conveniently and realize off-site design. With the development of the Internet of Things (IOT) technologies,

J. Zheng (✉) · H.-X. Yan
School of Management, Jinan University, Guangzhou, China
e-mail: zhengjbjnu@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_187,

the collaboration on product design is now not limited within Internet. IOT introduces new element for product design. For example, some new technologies like RFID, sensors and so on can not only provide purified, real time, total information of a product for designer, but also play a great role in managing the design process, supervising and guiding designers' behaviors.

As to networked collaboration product design, the prevailing researches pay much attention on the basis of Computer Supported Cooperate Work (CSCW) [1, 2]. Generally, CSCW can integrate various resources of different designers and enterprises so that they can share data quickly and improve the designing efficiency. Some symbolic researches in this field are: Arguing that any changes in shape can be described by the related change of model parameter, the literature [3] suggests designers parameterizing the visual characters and sharing three-dimension model by using CSCW to realize cooperative design. The literature [4] emphasizes the role of product database to support the large amount of information in networked cooperative design, and it also studies how to manage database to keep the consistency and concurrency of the product model. The literature [5] describes the construction of cooperative design platform, introducing the function of each part and the necessary technologies to achieve the goal. The literatures [6–9] advance a cooperative design platform based on grid mainly, solving the problem of different data co-existence by using grids.

These achievements have built a solid foundation for the further research. However, there are still some limitations. For example, some literatures focus on the way to unify the product's figures format; some discuss the management and classification of database, while some tend to apply the existing technologies. But none of these give enough attention to get information before the designing, nor discussing on the supervising and controlling of the design process. Further more, there is no expectation for probable questions-solving during the process. Therefore, the whole process will have considerable risks. Fortunately, the rapid development of IOT may provide some solutions for these problems. Hoping to make some contributions to the relative research, this paper deeply studies the new concept of designing products, puts forward a distributed collaborative product design platform based on IOT, and discusses the technologies needed.

187.2 Designing Product in the IOT Environment

Today, the rapid changes of consumers' needs make product design more and more complex. Actually, besides sophisticated products, even some common products can not be designed by a single designer (even an enterprise). On the contrary, the design needs the collaboration of designers and enterprises from different places. Now it is recognized that such collaboration can be carried out through Distributed Collaborative Products Design Platform (DCPDP), which focuses on how to use existing communication technologies to support designers from different places, including the establishment of dynamic alliances, management of design tasks,

supervision of development process and the management of relationships among designers. In this paper, a DCPDP based on the Internet of Things is defined as: A platform through which the designers located in different places communicate and cooperate for the same tasks. What makes it outstanding is that the application of IOT, which can help people to manage and supervise the designers and the designing process. This platform can also provide all the information and parameters anytime from anyplace, including the first-hand data about the environment. All these are beneficial to designing products more efficiently and successfully.

187.2.1 New View on Products Design

Product design is an intelligent activity, which begins from the analysis of customers' demand, going through concept design, structure design, detail design and ending up with product descriptions in the form of documents. Changes in competition make enterprises to be aware of the importance of building strategic partnerships. Under the pattern of over-corporate cooperation, during the process of design, designers should establish a concept based on the supply chain that consists of suppliers, manufacturers, retailers and final customers.

The goal of concept design is to get the basic scheme of the product, including function design, layout design and preliminary structure design. Simply speaking, this phase begins with the market analysis, and designers should analyze data from the view of customers in order to determine the original design proposal. Then, with the help of internal and external experts, the designer can subdivide and specify the abstract design. For instance, to reflect customers' demand more accurately, the designer can get some valuable information as customers' impression and preference on the product from the channel dealers such as distributors and retailers during the process of market research. Then the designer can forecast and decide the core questions combining with the data from market research. In the process of developing, a designer may invite experts from such institutions to give advice for the product's configuration, layout and structure, etc. At the same time, the designer may also need advice from suppliers for the parts choosing. All these have great influence on the form of final product.

The structure design is to specify the concept into actual product structure. Namely, the designer should decide the key parameters for each part in this phase. Cooperating with suppliers to get data of current parts on size, specifications and other attributes, the designer may ask them to provide the usage of this product. Further more, the designer may also let suppliers participate in some of the designing and give suggestions and supports, as well as other material suppliers may provide professional instructions on the performance and features of the new materials.

The detail design is to determine all the concrete parameters for each component and part. Using engineering drawing or 3D geometric model, the designer can determine the shape, size, tolerance and other parameters of each part. Taking into

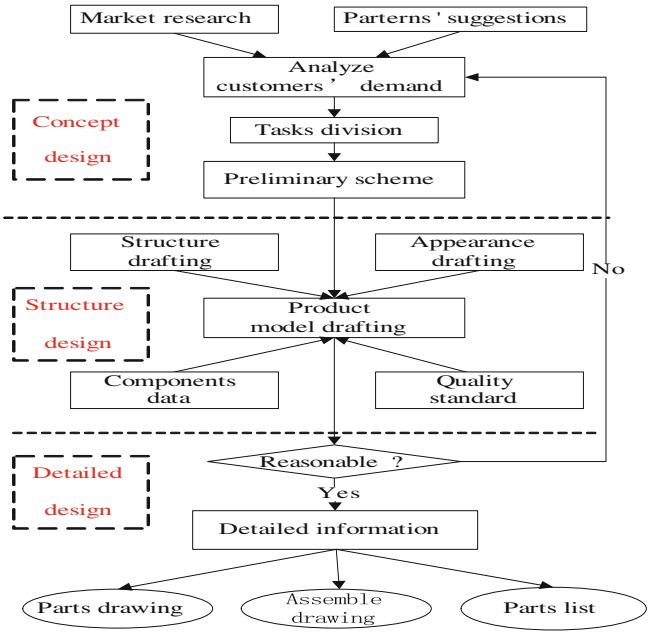


Fig. 187.1 Design flow chart based on SCM

account the importance of designing flow, the designer may invite outside experts again to participate in order that the designer can gain more practical suggestions.

Based on the descriptions above, we can get the design flow chart as Fig. 187.1.

187.2.2 Significance of IOT for Product Design

187.2.2.1 Providing Overall Information

By means of tracking products and analyzing the sale data, the designer could get detailed information about the advantages and disadvantages of products to improve the concept design. Besides, the application of IOT makes CRM more convenient and fast, and the system on this foundation makes information-getting, customer-accessing and after-sell-survey much easier.

In the structure design, collaborating with suppliers through IOT, the designer can get environment and product data directly. Then, the designer can refer suppliers to reducing defectiveness and costs.

In the detail design, by using Electronic Product Code (EPC) to get the specific number of size and other parameters, the designer can provide correct data for changing drawing into instructive process flow figure.

187.2.2.2 Controlling Design Process

By setting a deadline for the whole task and dividing it into several subtasks according to the Work Breakdown Structure (WBS), the design task can be assigned to the designers. At the same time, the manager can accelerate the subtasks according to the standard to control the design progress. To guarantee the task to be carried out on time, the manager can give some implications to designers at the right time.

187.2.2.3 Managing Designers

During the whole process, the relation between designers is collaborative and designers can communicate with each other anytime via IOT because the platform is of powerful communicating ability. An enterprise adopting this platform will no longer use the fixed working time, because designers can arrange their own time reasonably and the only thing they have to do is submitting the drafting or drawing on time. Besides, designers can ask for coordination meetings via the platform to make necessary changes according to the practical environment only if the applicants are more than a certain number.

187.3 Basic Structure and Functions of the Platform

We put forward a DCPDP based on the circumstance of IOT through which companies and designers from different places cooperate with each other and they are managed and coordinated in real-time. Therefore, this platform is designed according to the functions that IOT can provide.

187.3.1 Basic Structure

The basic structure is as Fig. 187.2. Applying interface is to help designers' work based on internet and computers, and designers in different areas have access to the interface. They can also browse different kinds of information about the product from the interface and design the part that belongs to them and send the results to be dealt and saved. It contains two kinds of style: 2-D drafting and 3-D drafting. Designers can lessen the complexity by using 2-D design in early stage. In the final stage, designers have to finish 3-D design and work out the final result. Make some micro modulate on its size and shape through using design data offered by the platform, and finally designers can submit the product. Cloud computing platform mainly provides powerful calculation to accumulate, choose, process the gathered information, form constrain formulation and solve them, and then save the final data in databases.

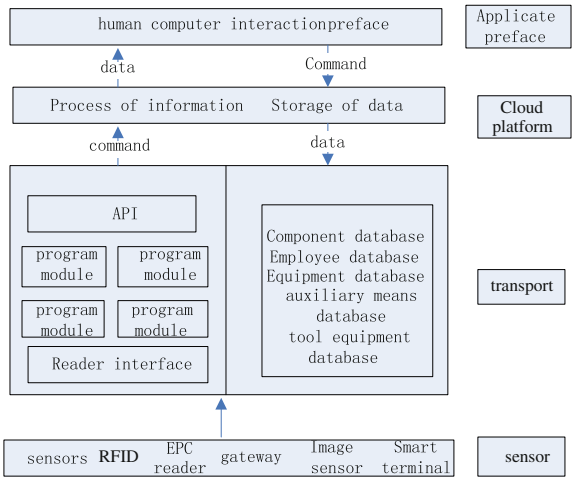


Fig. 187.2 Basic structure of DCPDP

187.3.2 Main Functions

EPC (Electronic Product Code) and RFID (Radio Frequent Identification) can help designers to get data of the parts under any circumstances. In the process of design, designers can browse these data at any time they want, and they can foresee the possible risks and problems in a more convenient way. Therefore, they can improve the design quality and reduce the failure rate.

Based on storing data collaboratively, cloud computing can process these data rapidly so that design results can be received, processed and classified timely. It doesn't only reduce the complexity, but also guarantee the quality.

The technologies of video, audio and electronic conference make designers communicate with each other freely and monitor others' progress at any time. With appropriate process-controlling, designers can reduce the conflicts. Besides, this platform has the common functions as data and user management. The main functions of this platform are as Fig. 187.3.

187.4 Key Technologies

187.4.1 Getting Information

The data of products and environment from which computers get information via IOT are in form of codes, which needs to be decoded into languages by the middleware to be understood by normal designers. Middleware is a series of 'programming module that has certain characters. It is located between operating system software and utility software, mainly providing developing circumstance for upper level. It can process all the information and event flow, also has the

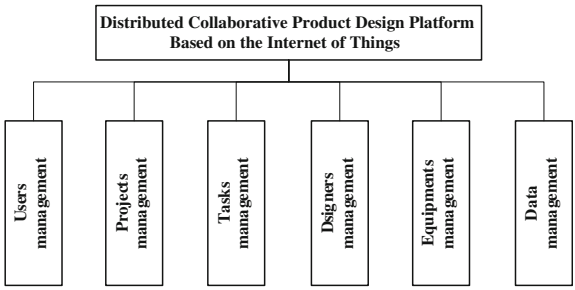


Fig. 187.3 Main functions of the platform

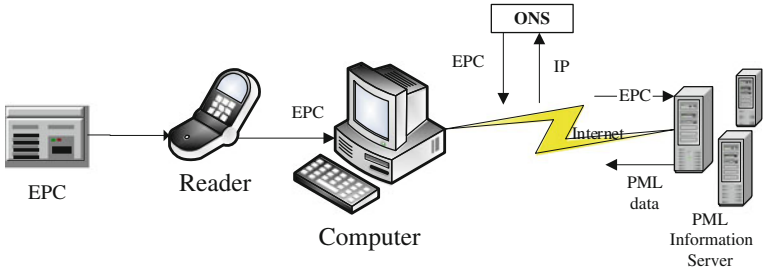


Fig. 187.4 Working principle of information-getting

function of filtering, grouping and counting. So designers can reduce the data that are sent to information system and prevent information from being read incorrectly. The principle is as Fig. 187.4.

When RFID readers identify the EPC, the information in it will be sent to the middleware. Then, the middleware system will find the location where information are stored by ONS (Object Name Service) on internet or intranet and then indicate the server which has these data including PML (Physical Markup Language) to analyze and send information to cloud platform to be stored. PML, which is simple and universal and is applied to control physical things and environment far away, is a new kind of computer language. All the documents are stored in a specific server to provide documents for other computers. PML server will be maintained by manufacturer, and it also stores the other products' information.

187.4.2 Technologies to Process Information

The development of IOT builds up a substantial foundation for distributed collaborative design, and the broadness of information it collects will provide great support for designers to fulfill the design comprehensively and correctively. However, how to analyze and apply the relative technologies correctively is

another key problem. By dividing the whole task into small ones by computer process programs, the cloud computing platform can process millionaire information within a second. Virtual technologies change physical resources into virtual ones that have the same functions and connections. This can reduce the complexity and present a logical condition for designers. Besides, it also makes single server support many virtual machines' works. Collaboratively storing data, this can express data abstractly in computer languages so that designers can manage the unified data and guarantee the safety and reliability.

187.5 Conclusions

In information era, communication will no longer be limited by time and areas. Collaboration becomes the main measure of designing and producing. Distributed collaborative products design platform based on the Internet of Things will provide great convenience to enhance design efficiency, which is to be of great significance for designers and managers to make full use of resources and to promote the management by integrating dominant resources that are located in different places. All of these studies in this paper provide practical references for enterprises to build distributed collaborative products design platform based on the Internet of Things.

Acknowledgments This work is sponsored by (1) the Philosophy and Social Science Fund of Chinese Ministry of Education, foundation item 09YJC630103; (2) the Fundamental Research Funds for the Central Universities, foundation item 10JYB2041.

References

1. Edmond D, Hofstede AHM (2000) A reflective infrastructure for work-flow adaptability [J]. Data Knowl Eng 34(3):271–304
2. Lei W, Wei G (2004) Systematic research on network-based cooperation design and manufacture platform system towards regional multi-sectors Industry [J]. China Mech Eng 15(19):1778–1781
3. ZHOU Y-b, LUO T-h (2008) 3D modeling based on internet collaborative design [J]. J Chongqing Univ Nat Sci Ed 31(9):1038–1043
4. TANG SH, MAY S, Cheng G (2008) Controlling of concurrency and consistency in the feature oriented database of product [J]. J Mach Des 25(10):43–45
5. WANG H-f, LIU T-t, ZHANG Y-l (2007) Research and realization on network cooperated technical design system [J]. J Mach Des 24(12):1–7
6. LUO Z-m, CHEN X-h, SUN Y-n (2009) Design of collaborative platform based on grid services [J]. Comput Eng Des 30(22):66–67
7. FAN s-q, YU ze, GUO h-j (2009) The influence of the Internet of Thing on Supply Chain Management [J].China Econ Trade Herald 19: 66
8. NING X-f, ZHANG C-y, WAN W, LI R-b, LI J-q (2010) Research and design of Internet of things architecture based on LTE system [J]. J Comput App 1:6–9
9. Gu X-j, Qi G-n, Tan J-r (2007) Introductions of modern manufacturing system engineering [M]. pp 22–30 Zhejiang University Press, Zhejiang

Chapter 188
A Cultural Particle Swarm Optimization Algorithm for Flow Shop Scheduling Problem

Yu-lin Zhang and Xia Zhu

Abstract To describe the minimization of the makespan in permutation flow shop scheduling problem, Cultural Particle Swarm Optimization (CPSO) is proposed. Two kinds of spaces, population space and belief space, are set in the algorithm. The population space is evolved with adaptive PSO strategy and the belief space is with update function. The interaction between two spaces is by the acceptance operation and impact operation using synchronous transmission mode. The proposed algorithm is tested on different scale benchmarks and compared with the other representative algorithms. The result shows that CPSO has faster convergence speed and is better than other algorithms in not only the solution quality but also the stability.

Keywords Particle swarm optimization · Cultural algorithm · Flow shop scheduling · Non-polynomial hard · Population space, the belief space

188.1 Introduction

The job-shop scheduling problem (JSP) consists in scheduling a set of jobs on a set of machines with the objective to minimize a certain criterion, subject to the constraint that each job has a specified processing order through all machines which are fixed and known in advance. It is well known that the job-shop scheduling problem (JSP) is in the class of NP-hard, non-polynomial (NP)-complete, and has proved to be a difficult task for human planners and schedulers,

Y. Zhang (✉) · X. Zhu
Electronic and Electrical Engineering, Huaiyin Institute of Technology,
Huaian, Jiangsu, China
e-mail: zhangyulin@hyit.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_188,

particularly if optimal solutions are required. Predecessors have proposed many different approaches to JSP and have obtained rich harvest, but current algorithms even moderately size problems cannot be solved to guarantee optimality.

Consider minimizing the maximum completion time of the classic replacement pipeline [1] scheduling problem. Denoted $p_{i,j}$ stands processing time, while the workpiece i is working on the machine j, $\pi = (j_1, j_2, \ldots, j_n)$stands all parts of a sort, Π is for the collection of all sorts, $C(j_i, k)$is for the completion of processing time while the workpiece j_i is on the machine k. Assuming the workpiece to the machine by machine 1 to m for processing the order, then the workpiece completion time on each machine mathematical formula can be described as follows.

$$C(j_1, 1) = p_{j_1,1} \tag{188.1}$$

$$C(j_i, 1) = C(j_{i-1}, 1) + p_{j_i,1} \quad i = 2, \ldots, n \tag{188.2}$$

$$C(j_1, k) = C(j_1, k-1) + p_{j_1,k} \quad k = 2, \ldots, m \tag{188.3}$$

$$C(j_i, k) = \max\{C(j_{i-1}, k), C(j_i, k-1)\} + p_{j_i,k} \quad i = 2, \ldots, n; \; k = 2, \ldots, m \tag{188.4}$$

$$C_{\max}(\pi) = C(j_n, m) \tag{188.5}$$

$$\forall \pi \in \Pi \quad \pi^* = \arg\{C_{\max}(\pi) = C(j_n, m)\} \rightarrow \min \tag{188.6}$$

Among them, the Eq. (188.5) is the maximum completion time, Eq. (188.6) indicates that minimizing the maximum completion time scheduling sorting scheme.

Particle swarm optimization is an evolutionary computation technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling. A significant number of successful applications of PSO to job-shop scheduling problems have been appearing [2–6]. Reynolds proposed a cultural algorithm, at different levels from the micro-level and macro-level simulation of biological evolution and cultural aspects of evolution, the evolution of the two-tier structure of the algorithm framework is conducive to global convergence of the algorithm.

This paper will introduce the cultural PSO algorithm (CPSO), proposes a cultural particle swarm algorithm to solve job shop scheduling problem of minimizing the maximum completion time problem. Two kinds of spaces, population space and belief space, are set in the algorithm. The population space is evolved with adaptive PSO strategy and the belief space is with update function. The interaction between two spaces is by the acceptance operation and impact operation using synchronous transmission mode. The algorithm uses the information contained in outstanding individuals, enhance the convergence of the algorithm. Under the same parameters set by the algorithm, respectively, with the literature on the algorithm, the results showed that CPSO not only improves the convergence of particle swarm speed, avoids local optimization problem, and avoids the occurrence of premature groups.

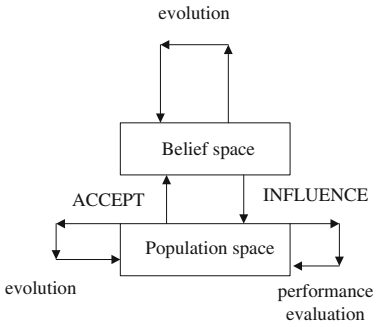


Fig. 188.1 The evolution framework

188.2 Cultural Particle Swarm Optimization Algorithm

In this paper, CPSO uses double-cultural evolution of the structure. The evolution of the framework is shown in Fig. 188.1. CPSO generally include three elements: population space, the belief space and communication channel which includes: acceptance function, update function, and influence function.

188.2.1 The Design of Population Space

PSO is basically developed through simulation of bird flocking in two-dimension space. Velocity of each agent can be modified by the following equation:

$$\begin{aligned} v_i(t+1) = \omega * v_i(t) + c_1 * rand() * (p_i\text{best} - x_i(t)) \\ + c_2 * rand() * (g\text{best} - x_i(t)) \end{aligned} \tag{188.7}$$

The current position (searching point in the solution space) can be modified by the following equation:

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{188.8}$$

where, $v(t)$: Velocity of agent; ω : weighting function; [0.1, 0.9];c_1, c_2: weighting factor; $rand()$: Random number between 0 and 1; $x(t)$: Current position of agent; $p$best: Best of agent; $g$best: Best of the group.

188.2.2 The Design of Belief Space

Belief space is defined as $< S, N[n] >$. Where, S is a collection of the best individual of the environmental knowledge. Standard knowledge N saves the changes of n main variable parameter. Each section has been described as $< Ij, Lj, Uj >$,

$j = 1, 2, \ldots, n$ where $I = [l, u] = \{l \le x \le u, x \in R\}$ expresses the value of the variable$x$domain boundary. Upper and lower bounds (u, l) of the range are given by the initialization. The fitness value Lj of the objective function corresponds to lower range lj of parameters j. The fitness value Uj of the objective function corresponds to Lower range uj of parameters j.

188.3 Accept Function

Population space transfer individual experiences to belief space by accept function, in fact; transfer a set of optimal subset. In the optimization problem, in general take a certain first few; usually take 20–25%, but also can change the rule according to environment

Accept (): In the group space evolution of PSO, when running time is Acc, replace the poor individuals of belief space with the global best.

$$Acc = B_{\text{num}} + \frac{\text{iter}}{\text{iter}_{\max}} \times D_{\text{num}} \tag{188.9}$$

Among them, $B_{\text{num}}$and D_{num} are two parameters used to adjust the number of acceptance and influence.

188.3.1 Update Belief Space

Different types of knowledge have their own corresponding update functions. Here a brief knowledge of the corresponding update functions of environment and standard knowledge of are introduced. The current and previous best individual are stored in Environmental knowledge, $S = \left\{\vec{S^t}, \vec{S^{t-1}}\right\}$ respectively stands the best individual of t generation and $t - 1$ generation. The update rules are as follows.

$\vec{S^{t+1}} = \vec{x^t_{\text{best}}}$, if $f\left(\vec{x^t_{\text{best}}}\right) < f\left(\vec{S^t}\right)$; $\vec{S^{t+1}} = \vec{S^t}$, Otherwise

Standard knowledge N update rules are as follows.

$l_j^{t+1} = x_{i,j}^t$, if $x_{i,j}^t \le l_j^t$ or $f\left(x_i^t\right) < L_j^t$; $l_j^{t+1} = l_j^t$, Otherwise $L_j^{t+1} = f\left(x_i^t\right)$, if $x_{i,j}^t \le l_j^t$ or $f\left(x_i^t\right) < L_j^t$; $L_j^{t+1} = L_j^t$, Otherwise

Here, individual i affects lower boundary of the parameters j.l_j^t: The lower boundary of parameter j in t iteration. The fitness value L_j^t corresponds to l_j^t. Similarly, individual k affects lower boundary of the parameters j. u_j^t: The lower boundary of parameter j in t iteration. The fitness value U_j^t corresponds to u_j^t.

$u_j^{t+1} = x_{k,j}^t$, if $x_{k,j}^t \ge u_j^t$ or $f\left(x_k^t < U_j^t\right)$; $u_j^{t+1} = u_j^t$, Otherwise

Table 188.1 Machine data and process constraints

Workpiece	Operation number		
	1	2	3
Operation time			
J1	3	2	5
J2	3	5	1
J3	2	5	3
Machine order			
J1	M1	M2	M3
J2	M1	M3	M2
J3	M2	M1	M3

$U_j^{t+1} = f\left(x_k^t\right)$, if $x_{k,j}^t \geq u_j^t$ or$f\left(x_k^t < U_j^t\right)$; $U_j^{t+1} = U_j^t$, Otherwise

188.3.2 Influence Function

Belief space form space experience, modify the individual behavior by affect function to get a higher evolution efficiency.

Influence (): In the group space evolution of PSO, when running time is Inf, replace the poor individuals of belief space with the global best.

$$\text{Inf} = B_{\text{num}} + \frac{\text{iter}_{\max} - \text{iter}}{\text{iter}_{\max}} \times D_{\text{num}} \tag{188.10}$$

Among them, $\text{iter}_{\max}$ is maximum of pre-set particle swarm evolution, iter is the current time of evolution.

188.4 Experiments

188.4.1 Experiment Analysis

Take a 3/3/G/$C_{\max}$ job shop scheduling problem for example, Table 188.1 shows the problem of processing time and process constraints. According to the combination features and JSP technology constraints, the coding for the CPSO must be specially designed. First, convert continuous to discrete values by using ROV rule, use disjunctive graph as a priority to decide on the same machine conflict sequence of operations, generate scheduling by using the critical path method.

1. ROV rule

Each particle position vector is produced by random key encoding; Position vector of particles k is a matrix X_k:

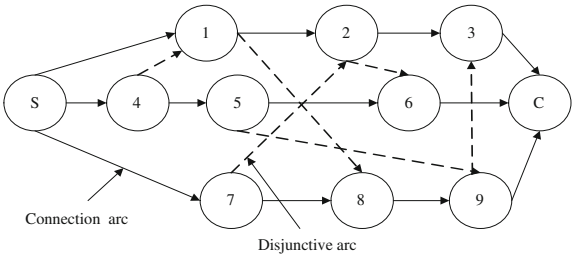


Fig. 188.2 Disjunctive graph

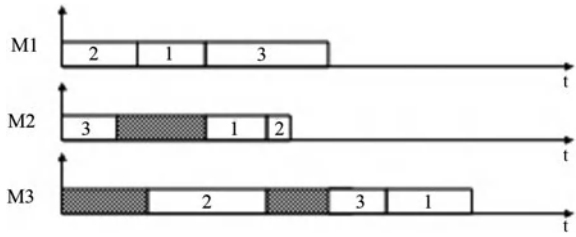


Fig. 188.3 Gantt chart

$$X_k = \begin{bmatrix} x_{11}^k & x_{12}^k & x_{1n}^k \\ x_{21}^k & x_{22}^k & x_{2n}^k \\ x_{m1}^k & x_{m2}^k & x_{nm}^k \end{bmatrix}$$

Each element of X_k is a real value, the i line of X_k corresponds to machine i. Use ROV rule, respectively, code each line of X_k, and get all the machine operations on the priority order. Suppose the k particle's position is as follows:

$$X_k = \begin{bmatrix} 1.34 & 1.16 & 3.50 \\ 3.28 & 0.01 & 0.67 \\ 3.30 & 3.41 & 2.42 \end{bmatrix}$$

Then for the machine 1, the sequence is 2-1-3 using ROV rules, for machine 2, the sequence is 3-1-2, for machine 3, the sequence is 2-3-1.

2. Disjunctive graph.

When get an orderly operation table. First set up the direction of the arc map according to table, get an acyclic graph.

According to Fig. 188.2, we can get the critical path (4-1-8-9-3).Corresponding Gantt chart is shown in Fig. 188.3

3. The CPSO algorithm to solve JSP

Based on the previous part of the introduction, the proposed CPSO algorithm to solve the JSP is as follows.

Table 188.2 CPSO algorithm performance

Problem	n	m	Best value	Average value	Poor value	Average time (S)
LA01	10	5	666	666	666	17.8
LA06	15	5	930	930	930	61.1
LA11	20	5	1222	1222	1222	145.6
LA16	10	10	945	945.5	946	148.3
LA21	15	10	1024	1024.4	1078	733.7
LA26	20	10	1218	1218	1218	2449.5
LA31	30	10	1789	1789	1789	3883.6
LA36	15	15	1270	1270	1780	3550.2

Table 188.3 Optimal performance comparison between CPSO and other PSO algorithm

Problem	n	m	CPSO	Literature [4]	Literature [5]	Literature [6]
LA01	10	5	666	666	666	666
LA06	15	5	930	930	930	930
LA11	20	5	1222	1222	1222	1222
LA16	10	10	945	945	946	946
LA21	15	10	1024	1056	1078	1078
LA26	20	10	1218	1218	1218	1218
LA31	30	10	1789	1789	1789	1789
LA36	15	15	1270	1280	1287	1271

(1) Initialization: Set the population space and belief space the size of the population: $n = 30$:population size, $\omega = 0.9$:the inertia weight factor, $c_1 = c_2 = 2$:speed factor, $x_{\min} = 0$:the smallest particles position value, $x_{\max} = 4.0$:maximum position, $v_{\min} = -4.0$:particles minimum speed, $v_{\max} = 4.0$:maximum speed value;
(2) Initialize the particle's positions and velocities based on random key code. Evaluate the target value according to the disjunctive graph;
(3) Calculate the particle's actual input and the fitness function;
(4) Update the belief space of particle positions and the training samples, and calculate the particle's actual input and the fitness function;
(5) If iter%Inf == 0, do acceptable operation;
(6) Update the group's position and velocity of population space, according to evaluation of the disjunctive graph, and test whether the particle's velocity and position are out of range, and if so, adjust the speed and position;
(7) Calculate the actual input of each particle and the fitness function;
(8) Update the particle position and velocity of belief space, according to the evaluation of the disjunctive graph, and test whether the particle's velocity and position are out of range, and if so, adjust the speed and position;
(9) Calculate the particle's actual input and the fitness function of the belief space;
(10) If iter% Inf == 0, do affect the operation;

(11) When the error come to the initial set value or the maximum number of iterations come, output the scheduling program corresponding to *g*best and target; otherwise iter = iter + 1, return to Step 3.

188.4.2 Experiment Simulation

In order to verify the performance of JSP CPSO algorithm, take a typical JSP example (LA); these tests are widely in the literature for testing. Given class of problems by Lawrence LA, select one of eight in this issue of different sizes: LA01, LA06, LA11, LA16, LA21, LA26, LA31, and LA36.

Hardware simulation environment: Intel Core (TM) 2 DUO CPU T5250 @1.50 GHz/2046 MB; Software Platform: VISTA operating systems, MATLAB 7.0 compile software

Tables 188.2 and 188.3 show that CPSO algorithm for the numerical examples are able to obtain the optimal solution, which indicates that CPSO algorithm has a good search quality. As a result of cultural evolution of double-layer structure of the algorithm, this paper presents not only the culture of PSO has a higher operating efficiency, but also has good global search capability. Tracking individual extreme and global extreme to updated themselves and achieve the evolution of groups, then through continuous interaction space with best individual, speed up the convergence rate.

188.5 Summary

This paper introduces the basic principles of cultural algorithms to PSO, proposes a cultural particle swarm algorithm (CPSO), and typically takes examples of JSP (LA) for the simulation examples, and verifies the effectiveness of the algorithm. Simulation results show that: CPSO improves the convergence rate, avoids local optimum and premature convergence problem, and has certain advantages.

References

1. Reynolds RG (1994) An introduction to cultural algorithms [C]. In: Proceedings of the 3rd annual conference on evolutionary programming. World Scientific, River Edge, pp 131–139
2. Tasgetiren MF, Liang YC, Sevkli M et al (2007) A particle swarm optimization algorithm for makespan and total flow time minimization in the permutation flow shop sequencing problem [J]. Eur J Oper Res 177(3):1930–1947
3. Lian ZG, Gu XS, Jiao B (2008) A novel particle swarm optimization algorithm for permutation flow shop scheduling to minimize makespan [J]. Chaos Solitons Fractals 35(5):851–861
4. Sha DY, Hsu CY (2006) A hybrid particle swarm optimization for job shop scheduling problem. Comput Ind Eng (51):791–808

5. Tasgetiren MF, Sevkli M, Liang C, Gencyilmaz G (2004) Particle warm optimization algorithm for permutation flow s hop sequencing problem [A]. Lecture Notes in Computer Science [C]. Springer, Berlin 3172:382–389
6. Chang Z (2009) A hybrid algorithm for flow-shop scheduling problem. Acta Autom Sin 35(3):332–336. (In Chinese)

Chapter 189
A Scheduling Method Based on Deadlines in MapReduce

Zhang Xiaohong, Ju Shui and Jiao Zhibin

Abstract When Map tasks are scheduled to the nodes without the input data, they will cause the remote data access delay which degrades the performance of MapReduce. The delay can be hidden by prefetching the input data to the nodes predicted to run the tasks. However, when the actual nodes running the tasks are not identical with the predicted nodes, the default method of Hadoop MapReduce cannot schedule the tasks to the nodes with the prefetched data. This makes hiding the delay by prefetching difficultly. In this paper, we propose a new scheduling method to handle the problem. The method sets scheduling deadlines for the tasks of which the input data have been prefetched. It schedules the tasks to the nodes with the prefetched data before the deadlines or to other nodes after the deadlines. The experiments have shown that the new method scheduled more than 90% of tasks to the nodes with the input data or the prefetched data, and improved performance of MapReduce by 6.5 ~ 12%.

Keywords Prefetching technique · Scheduling · MapReduce

Z. Xiaohong (✉) · J. Shui · J. Zhibin
Henan Polytechnic University, School of Computer Science and Technology, Jiaozuo 454000, People's Republic of China
e-mail: xh.zhang@hpu.edu.cn

J. Shui
e-mail: jushui@hpu.edu.cn

J. Zhibin
e-mail: jiaozhb@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_189,

189.1 Introduction

MapReduce [1] is a parallel processing model for large scale and data intensive applications. It provides convenient APIs which make the development of parallel programs efficiently. Due to the loose decoupling architecture, MapReduce can provide great scalability and fault tolerant, Hadoop MapReduce [2], as an open source of MapReduce, has been adopted widely [3–7].

Due to the computing resource competition and the limitation of the default scheduling policy, almost in each job, there exist some Map tasks which cannot be scheduled to the nodes with the input data. These tasks will cause the remote data access delay which degrades the performance of MapReduce. Zhang et al. [8] proposed a prefetching technique to hide the delay. The technique first predictes the nodes to execute a job according to the order that nodes request new tasks, and then selects the Map tasks to cause the delay. Finally, it prefetches the input data before the selected tasks were scheduled.

However, the actual order that nodes request tasks is not identical with the predicted order because of the uncertainty in MapReduce and the limitation of the predicting model. In this situation, the default method Hadoop MapReduce cannot schedule all selected Map tasks to the nodes with the prefetched data, and hence some of the delay cannot be hidden by the technique.

Zaharia et al. [9] proposed a delay scheduling method to assign Map tasks to the nodes with the input data. However, the method takes fairness as the cost. Zhang et al. [10] proposed a method to schedule Map tasks to the nodes with the input data according to future task requests. Tian et al. [11], Isard et al. [12], Sandholm et al. [13], Polo et al. [14] and Dhok et al. [15] also proposed new scheduling methods in Hadoop MapReduce. However, the methods were not used to schedule tasks to the nodes with the input data.

In this paper, we propose a scheduling method based on deadlines. The method sets scheduling deadlines for the Map tasks of which the input data have been prefetched. It assigns the tasks to the nodes with the prefetched data before the deadlines or to other nodes after the deadlines. We have implemented the technique in Hadoop-0.20.1. The experiment results have shown that the technique scheduled more tasks to the nodes with the prefetched data, and improved the performance of MapReduce.

This paper is organized as follows. Section 189.2 describes the scheduling method. Section 189.3 introduces the experiment environment and discusses the results. Section 189.4 concludes this paper.

189.2 Scheduling Tasks Based on Deadlines

Due to the uncertainty in a MapReduce cluster and the limitation of the model predicting the order that nodes request tasks, the actual order that nodes request tasks is not identical with the predicted order. In this situation, the default

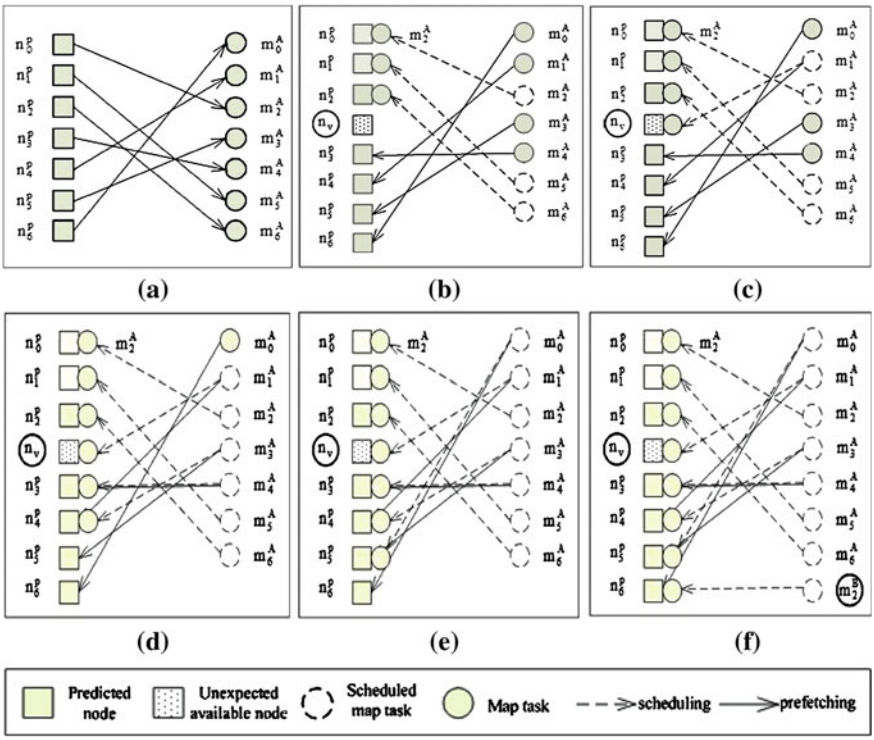


Fig. 189.1 Difficulty of scheduling tasks to prefetched input. **a** Prefetching input of tasks to different predicted nodes; **b** An unpredicted node requesting a task; **c** scheduling m_1^A to n_v; **d** scheduling m_3^A to n_4^p; **e** scheduling m_5^A to n_5^p; **f** scheduling m_2^B to n_6^p

scheduling method of Hadoop MapReduce cannot assign Map tasks to the nodes with the prefetched data, which makes the prefetching technique work inefficiently.

For simple presentation, we use the same method in [8] to describe a MapReduce job and the nodes predicted to run the job. Supposing A includes 7 Map tasks, and n_3^p is the set of the Map tasks in A, $M^A = \{m_0^A, m_1^A, \ldots, m_6^A\}$. N^P is the set of the nodes predicted to run A, $N^P = \{n_0^p, n_1^p, \ldots, n_6^p\}$.

Supposing $\forall m_i^A \in M^A$ and $i \in [0, 6]$, $\forall n_k^p \in N^P$ and $k \in [0, 6]$, the input data of m_i^A is not stored on n_k^p. When m_i^A runs on n_k^p, it will cause the remote data access delay. To hide the delay, the input data of m_i^A is prefetched to n_k^p before m_i^A is scheduled. Figure 189.1a shows the prefetching relationship according to the technique in [8]. In the figure, the input data of a task is prefetched to the node connecting to the task by a line.

Supposing n_v requests a task more early than n_4^p as Fig. 189.1b shows. The default method of Hadoop MapReduce schedules tasks according to the distance

between the input data of the tasks and request nodes. The input data of m_1^A is nearest to n_v, so m_1^A is scheduled to n_v, just as Fig. 189.1c shows.

When n_4^p requests a task, the default method can only select the task (m_3^A) of which the input data is not prefetched to n_4^p but nearest to n_4^p, just as Fig. 189.1d shows. According to the same policy, the default method schedules m_0^A to n_5^p and m_2^B to n_6^p, just as Fig. 189.1e and f show. m_3^A and m_0^A are not scheduled to the nodes with the prefetched data or the nodes with the input data, the remote data access delay caused by these tasks cannot be hidden by the prefetching technique.

To improve the efficiency of the prefetching technique, we propose a scheduling method based on deadlines. The method sets a scheduling deadline for each task of which the input data has been prefetched. It will preferentially schedule the tasks to the nodes with the prefetched input data before the deadlines or to other nodes after the deadlines.

If an unpredicted node request a task, the new method first selects a task according to the distance between the input data of tasks and the node, and then check the scheduling deadline of the selected task. If the deadline is larger than the current system time, the new method will skip the task, and select another task for the request node.

For each node, the new method first computes the difference between the predicted time to request a task and the actual time, and then set (the difference + the predicted time) as the scheduling deadline for the task of which the input data has been prefetched by the node. If an unpredicted node requests a task before scheduling deadlines, no task can be scheduled to the node. To avoid such occurrence, the new method sets two constraints when selecting a task for a request node.

Constraint 1: if the input data of tasks have been prefetched, schedules a task according to the distance between the input data of tasks and the request node, regardless of scheduling deadlines.

Constraint 2: if the node which prefetched the input data of a task becomes unavailable, sets the deadline of the task with 0.

To facilitate scheduling, we use a special list to store the prefetching information and name this list as Pscheduling. In the list, each element includes a prefetching record. Each record store a task, the node prefetching the input data of the task. After receiving a request from a node, the new method schedules a task as follows:

(1) Set the head job of the waiting queue as the current job
(2) If there exists the task of which the input data has been prefetched by the node, execute (6); otherwise, execute (3).
(3) If there exists the task of which the input data has been stored on the node, execute (6); otherwise,execute (7).
(4) If the input data of the task has been prefetched, execute (9); otherwise, execute (5).

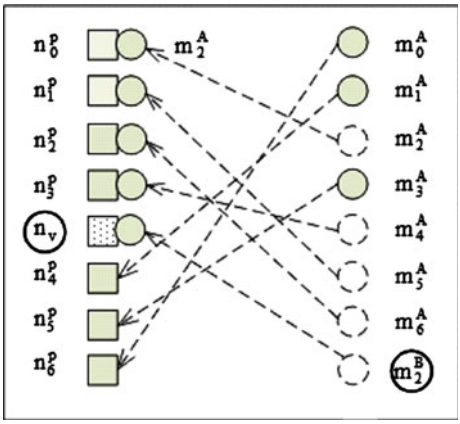


Fig. 189.2 Scheduling Map tasks to the node with the prefetched data

(5) Delete the task from the waiting task, and return the task.
(6) Delete the task from Pscheduling, and execute (5).
(7) If there exists the task of which the input data is nearest to the node, execute (4); otherwise, execute (13).
(8) Return null.
(9) If the current system time is smaller than the deadline of the task, execute (10); otherwise, execute (6);
(10) If the task is the last waiting task of the current job, execute (11); otherwise, execute (12).
(11) If the current job is the last waiting job, execute (6); otherwise, set the next waiting job as the current job, execute (2).
(12) Skip the task and execute (7).
(13) If the current job is not the last one in the waiting queue, set the next job as the current job, and execute (2); otherwise, execute (8).

When n_v requests a new task, the new method first finds the task (m_1^A) of which the input data is nearest to n_v, and then checks the scheduling deadline of m_1^A. If the current system time is smaller than the deadline, the new method will skip m_1^A and find another task for n_v according to the distance between the input data of tasks and n_v.

Supposing m_3^A and m_0^A are the two tasks found successively, these tasks cannot be scheduled to n_v because the current system time is smaller than their deadlines. The new method skips the tasks, and selects another task for n_v. At last, m_2^B is scheduled to n_v, as Fig. 189.2 shows. Comparing with the default scheduling method of Hadoop MapReduce, the new method can schedule m_3^A and m_0^A to the nodes with the prefetched data, and the remote data access delay caused by the tasks can be hidden by the prefetching technique in [8].

Table 189.1 Information of jobs running in our experiment

Job type	Input size	Map tasks	Reduce tasks	Total number
Big job	3.74 GB	30	2	7
Small job	383.03 MB	3	2	15

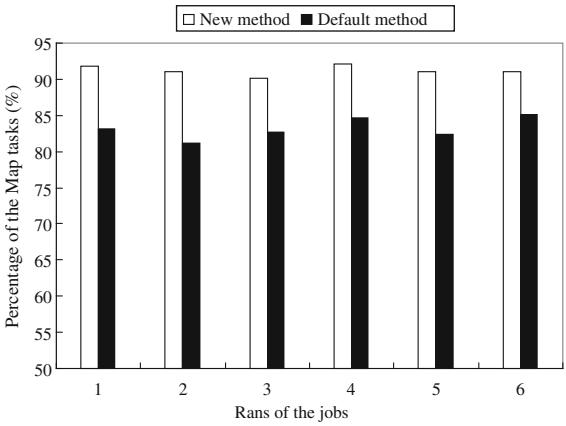


Fig. 189.3 Percentage of the tasks scheduled to the nodes with the input data or the prefetched data

189.3 Experiments

We have implemented the new scheduling method in Hadoop-0.20.1. We deployed the Hadoop with the prefetching technique and the default scheduling method and the Hadoop with the prefetching technique and the new scheduling method on a Linux cluster. The cluster includes 15 nodes connected by Gigabit Ethernet. We ran the same jobs in the Hadoop environments, and Table 189.1 shows the details of these jobs.

In each Hadoop, we ran a set of jobs comprised of 7 big jobs and 15 small jobs. We compared the normalized executing time and the percentage of the Map tasks scheduled to the nodes with the input data or the prefetched data when the jobs ran in the Hadoop environments, respectively.

We counted the Map tasks scheduled to the nodes with the input data or the prefetched data. To compute the percentage, the total number was divided by the total number of all Map tasks in the jobs. The results were described in Fig. 189.3. According to the figure, the new method scheduled more Map tasks to the nodes with the input data or the prefetched data when comparing with the default scheduling method.

We recorded the execution time of each ran of the jobs in each Hadoop environment. To calculate the normalized execution time, the two execution time of each ran were divided by the longer of the times. We described the results in Fig. 189.4. According to the figure, the performance of Hadoop can be improved by 6.5 ~ 12%.

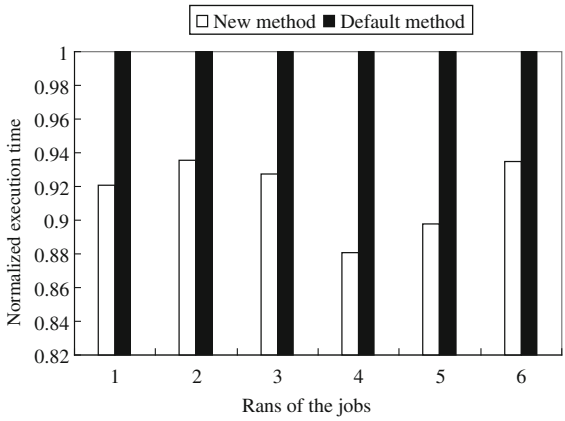


Fig. 189.4 Normalized execution time

189.4 Conclusion

We have proposed a prefetching technique to hide the remote data access delay caused by the Map tasks processed on the nodes without the input data. However, if the actual nodes running a job are not identical with the nodes predicted to run the job, the default method of Hadoop cannot schedule Map tasks to the nodes with the prefetched data, and the prefetching technique cannot hide the delay efficiently.

To avoid such occurrence, we proposed a scheduling method based on deadline. We have implemented the method in Hadoop-0.20.1 and evaluated the method on a Linux cluster. The experiment results have shown that the method scheduled more Map tasks to the nodes with the input data or the prefetched data, and improved the performance of MapReduce by 6.5 $\sim$ 12% when comparing with the default scheduling method of Hadoop.

References

1. Dean J, Ghemawat S (2004) MapReduce: simplified data processing on large clusters. In: Proceedings of 6th Symposium on Operating System Design and Implementation (OSDI), pp 137–150. USENIX
2. http://hadoop.apache.org, 2010
3. Huang X (2009) Research. Development and practice of cloud computing in China mobile. http://labs.chinamobile.com/mblog/47_40345
4. Leo S, Santoni F, Zanetti G (2009) Biodoop: bioinformatics on hadoop. In: Proceedings of International Conference on Parallel Processing Workshops.Vienna, pp.415–422. IEEE
5. Liu X, Han J, He J et al (2009) An approach to parallel processing of spatial data on clusters. Chinese High Technol Lett 19(10):991–997
6. Wegener D, Mock M, Adranale D et al. (2009) Toolkit-based high-performance data mining of large data on mapreduce clusters. In: Proceedings of the IEEE International Conference on Data Mining Workshops, pp 296–301. IEEE

7. Zhang S, Han J, Liu Z et al. (2009) Spatial queries evaluation with mapreduce. In: Proceedings of the Eighth International Conference on Grid and Cooperative Computing (GCC), pp 287–292. IEEE
8. Zhang X, Feng S, Fan J et al. (2012) Improving mapreduce performance by prefetching before scheduling. J High Technol Lett, vol 2
9. Zaharia M, Borthakur D, Sarma JS et al. (2009) Job scheduling for multi-user mapreduce clusters, Technical Report No.UCB/EECS-2009-55[R]. University of California at Berkeley
10. Zhang, Zhong Z, Feng S et al. (2011) Improving data locality of mapreduce by scheduling in homogeneous computing environments. In: Proceedings of IEEE 9th International Symposium on Parallel and Distributed Processing with Applications (ISPA), pp 120–126. IEEE
11. Tian C, Zhou H, He Y, Zha L (2009) A dynamic mapreduce scheduler for heterogeneous workloads. In: Proceedings of the Eighth International Conference on Grid and Cooperative Computing(GCC), pp 218–224. IEEE
12. Isard M, Prabhakaran V, Currey J, Wieder U, Talwar K, Goldberg A (2009) Quincy: fair scheduling for distributed computing clusters. In: Proceedings of the ACM SIGOPS symposium on Operating systems principles(SOSP), pp 261–276. ACM
13. Sandholm T, Lai K (2010) Dynamic proportional share scheduling in hadoop. In: Proceedings of Proceedings of the 15th international conference on Job scheduling strategies for parallel processing, pp 110–131. Springer
14. Polo J, Carrera D, Becerra Y, Steinder M, Whalley I (2010) Performance driven ask co-scheduling for mapreduce environments. In: Proceedings of IEEE/IFIP Network Operations and Management Symposium(NOMS), pp 373–380. IEEE
15. Dhok J, Varma V (2010) Using pattern classification for task assignment in mapreduce. http://researchweb.iiit.ac.in/∼jaideep/

Chapter 190
Product Data Sharing and Integration Technology in Network Supported Collaborative Design

Wang Qi, Ren Zhong-wei and Yu Lian-yong

Abstract In network supported collaborative design, the sharing and integration of multi-site and heterogeneous data is one of the key problems. The flexible data model and open access characteristics of XML can provide powerful support for the product data sharing and integration. This paper presents the XML-based solution of product data sharing and integration. On the basis of summarizing STEP and the XML technology characteristic respectively, the technology of the integration of XML language and STEP standard to express product is analyzed. Aiming at the related question in data transmission and exchange using XML, such as data parse, data storage and data query, this paper all gives the detailed explanation.

Keywords Product data sharing and integration · Network supported collaborative design · XML

190.1 Introduction

With the development of internet technology and manufacturing, the network supported collaborative design becomes a general trend of the product design. The network supported collaborative design stressed that the designers related with

W. Qi (✉) · R. Zhong-wei
School of Mechanical Engineering, Shenyang University of Technology, Shenyang, China
e-mail: angelwangqi@sina.com

Y. Lian-yong
Yantai Wellsys Electronic Co. Ltd, Yantai, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_190,

various stages of product life cycle working in groups complete the product development with the network supporting [1]. Possibly, not only the software and hardware environment of the designers are different, but also the geographical position is not the same place. Therefore, to enable the participants working in heterogeneous platform and different places carrying on the joint operation smoothly, one of key questions is to realize the information sharing and integration.

Among the traditional data exchange methods, the STEP neutral file is quite mature, also obtained the widespread application. But the STEP neutral file mainly faces the enterprise interior to carry on the product data exchange and integration, is unable to satisfy the need of product data exchange in the network environment. The XML-based technology provided the enormous convenience for data description, data exchange and transmission through the network.

The purpose of this paper is to realize the product data sharing and integration in network supported collaborative design based on XML.

190.2 Product Data Expression

190.2.1 STEP-Based Product Data Expression

STEP (Standard for the Exchange of Product Model Data) developed by the International Standards Organization (ISO), is a comprehensive standard that describes how to represent and exchange digital product information. STEP standards provide a neutral mechanism for the description of product information covered the entire product life cycle [2]. STEP is one of the important tools for solving the information integration under the traditional manufacture pattern. At present, the commonly used CAD application software, such as CAM, CAE, CAPP and so on, supports the STEP standard application agreement and takes the STEP neutral document as a foundation to realize the integration with the CAD system [3]. EXPRESS is information modeling language and the core of the STEP standard.

However, the question still exists in the course of the STEP-based collaborative design. The data exchange way based on STEP lacks the enough flexibility and the extension. The character-based EXPRESS entity can not realizing the transmission online with the general network agreement (for example FTP, HTTP) and can not be supported by the browser and the database. All the documents generated during the design must be re-transferred, which includes a large amount of redundant information. This method is not suitable for product data sharing and integration between heterogeneous systems in network supported collaborative design.

190.2.2 XML-Based Product Data Expression

XML (eXtensible Markup Language) is group of standards which W3C founded in the late-1990s. XML may enable the interactive cooperation which carries on through the network to have the good reliability and the interoperability [4].

XML has the following main characteristics:

- The good data storage form. The semantic marking in XML document, on the one hand defined the element hierarchical structure, on the other hand explained the meaning of the element.
- The independence of content. XML may describe various data stored in heterogeneous environment ignoring the concrete application.
- The diversity of display format. An extremely bright characteristic of XML is the data display format separated form the data expression. This character may realize seamless connection of data from different sources and cause the data exchange and transmission online very easily.
- The scalability. We can increase the new data field in XML document and does not destroy the existing structure. One thing we needed to do for changing the data model is changing the data model definition, like DTD (document type definition) and so on. We do not need to recode the existing object.
- The capacity of local update. When the data is partly changed, we can only transmit the data being changed but the whole document from the server to the client side.
- The high flexibility. XML is an open standard and does not have the pre-defined label. The user may define the construction of data by himself.

These characteristics enable XML to become most popular and practical tool meeting the need of data sharing and transmission on network. But, we cannot establish the unified standards and the specification for the product model structure with the XML definition. Thus, we cannot realize the product information sharing for entire life cycle. Therefore, the basis of the XML-based product data sharing is the establishment of the unified profession standard and the product modeling standard.

190.2.3 Product Data Expression Based on STEP and XML

We realize the product data sharing and integration by using XML to complete the STEP-based product data recognition online. This method not only retains the advantages of STEP file in product information description but also displays the superiority of XML in the data online transmission.

To realize the representation of STEP with XML, the core mission is mapping EXPRESS expression to XML schema structure. In November, 1999, ISO promoted “XML expression model for EXPRESS driver data” (ISO 10303-28, was called part28) [5]. Part28 stipulated the related standard for model transformed and

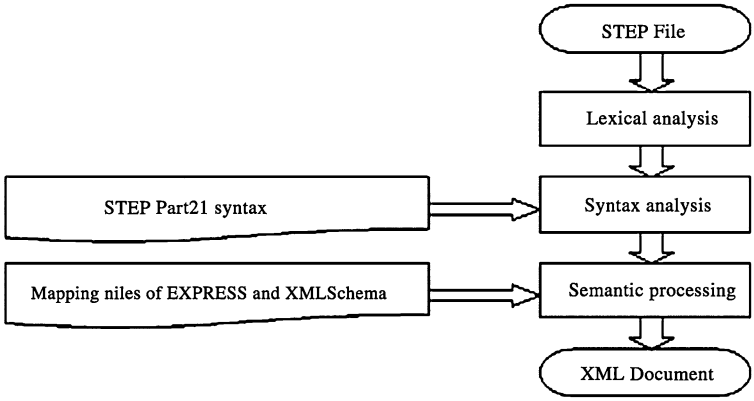


Fig. 190.1 STEP file conversion

provided the theoretical foundation of establishing the mapping model between the XML document and the STEP file. The process of STEP file conversion is shown in Fig. 190.1.

STEP data analysis plays a very important role in this process. STEP data analysis is to go through three steps: lexical analysis, syntax analysis and semantic processing. Lexical analyzer is supposed to read the characters from the STEP file and to complete word recognition according to STEP Part21 syntax. Syntax analyzer is used to get a hierarchical structure of the STEP physical file and call the corresponding function to analyze the syntax for each STEP syntax unit. The function of semantic processor contains a syntax-directed translation and structural transformation processing. The content of syntax-directed translation includes three parts: first, searching the STEP keywords (obtained from the lexical analyzer) in STEP Application Protocol library and getting the corresponding semantic information; second, matching the semantic information to the hierarchical structure information (obtained from the syntax analyzer) and inserting the information into the corresponding location of the parsing program and to form the grammar translation; last, realizing the structure transform according to the documents logical organization ruled by XML Schema, thus the translation grammar will be transferred to the corresponding XML documents.

190.3 XML-Based Product Data Sharing

To realize the XML-based data sharing, three aspects of the work need to be completed: parse XML document, the XML data storage and the XML data query.

190.3.1 Parse XML Document

The XML data parse is the basis of the XML-based data processing. Extracting the data and structural information from the XML documents in text form, we can generate the tree structure of the XML document. It provides an interface for the XML application to realize the information processing by operating data objects (rather than a text file).

Among the XML analysis methods, DOM (Document Object Model) and SAX (Simple API for XML) analysis are commonly used [6]. DOM is an object interface which has nothing to do with the language. The application procedure may visit the XML internal data through this interface. The DOM model may realize the comprehensive and dynamic access to the entire XML documents. But the increasing size of the XML documents will cause the memory consumption and the extension of the process time. SAX is an event driven standard interface of XML analysis. The user provides the processor conforms to the interface definition. SAX will create the related event when the special mark is discovered in XML analysis and call the corresponding processing function from the processor. But, the SAX cannot access XML documents directly and do not support revising XML in situ.

190.3.2 XML Data Storage

XML data processing based on relational database is a feasible and widely used way among all ways of storing and managing XML data.

There are three methods to store the XML documents in relational database: the fine grain method, the coarse grain method and the granularity method [7]. The fine grain storage method is simple and also obtains the widespread support of commercial relational DBMS. So we may take the fine grain storage method as the preferred selection. The basic philosophy of the fine grain storage method is each structure in the XML documents being assigned the only identifier in a relational database, namely each structure has the respective table. Each element, attribute and character data, as structure constituent, is stored separately in the independent field of the database table.

The transformation algorithm between product data XML pattern and relational database physical model will be explained concretely as followed:

1. Subtract the naming space of the XML pattern. The main application of naming space in the XML pattern is to provide the only name for some specific personnel or the organization. The naming space has nothing to do with the relational scheme.
2. Establish the relational scheme by founding a physical schema for each complex type in XML pattern. Thus, each element in complex type is transformed to a corresponding attribute of physical schema.

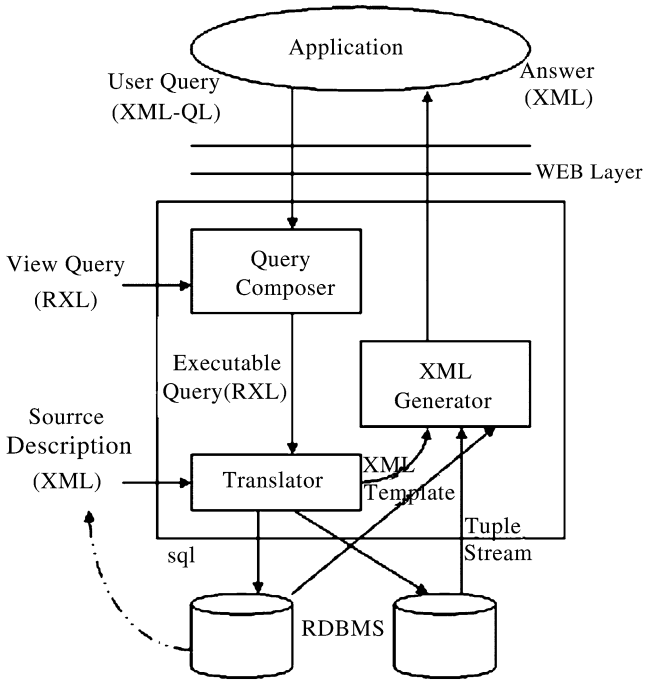


Fig. 190.2 XML data query

3. Confirm the primary key in the physical schema. The element with the unique restraint in the XML pattern just is the primary key of the physical schema.
4. Compose a relational database by all complex types in the XML pattern.
5. Confirm the foreign key in relational scheme. The nesting complex type will be transformed to the foreign key in the physical schema.
6. The simple enumeration type in the XML pattern will be restrained with CHECK in the physical model.

190.3.3 XML Data Query

XML is emerging as a standard to define such semi-structured data which is self-describing. Also, to retrieve data from such semi-structured databases, special semi-structured query languages have been proposed; examples are XQL, Lorel, XML-QL, or XML-GL. In our work we take as an example the XML-QL query language. The process of XML data query is showed in Fig. 190.2.

XML-QL enables data extraction from XML documents and allows to expressing mappings between different ontologies. To query unknown or

unpredictable data structures, XML-QL and other query languages for semi-structured data have two features in common: regular path expressions and the ability to query the schema (i.e., attributes and references). In addition, XML-QL has a powerful restructuring mechanism; that is, the result of an XML-QL query could be a complex XML document. We will mostly ignore the restructuring part and concentrate on the data extraction part of the language because data extraction is always the costly part of query evaluation. The overall approach is to translate an XML query into SQL, let the RDBMS execute the query, and do some post-processing in order to get the right query result (e.g., generate XML if this is requested by the query). All three of these steps can be carried out in a straightforward way, but of course the most interesting step is the translation into SQL.

190.4 Conclusion

Product data sharing and integration is especially important in network supported collaborative design. In view of many strong points of XML data on description and transfer, XML has been used more and more widely for data representation and exchange. This paper gives the XML-based solution for data sharing and integration in network supported collaborative design. For some key technology, just like data conversion, data storage and data query, this paper gives the choice of methods and strategies very detail drawing on the results of the current study. The research content of this paper has some referential significance in network supported collaborative design area.

References

1. Maher ML, Rutherford JH (1997) A model for synchronous collaborative design using CAD and database management. Res Eng Des 9(7):95–98
2. Gui-Chuan T (1998) STEP-based model mapping and implementation of product data exchange and knowledge modeling. Zhejiang University, Hangzhou
3. Chin KS, Zhao Y, Mok CK (2002) STEP-based multiview intergrated product modelling for concurrent engineering. Adv Manuf Technol 20:896–906
4. Renner A (2001) XML data and object databases: a perfect couple. In: Proceedings of the 17th international conference on data engineering, Heidelberg, Germany pp 143–148
5. Cheng-Feng JIAN, Jian-Rong TAN (2001) Description and identification of STEP product data with XML. J Comput Aided Des Comput Graphics 13:984–990
6. Fiebig T, Helmer S, Kanne C (2002) Anatomy of a native XML base management system. VLDB J 11(4):292–314
7. Florescu D, Kossmann D (2002) Storing and querying XML data using an RDBMS. IEEE Comput Soc 22(3):27–34

Chapter 191
A Solution for Management Information Systems Integration Based on WebService

Guowei Wang, Zhibin Zhang and Manjun Xue

Abstract By enveloping the data source into WebServices and constructing Universal Description Discovery and Integration (UDDI), the solution gives a method of data integration based on Service Oriented Architecture (SOA), then presents WebService and Extensible Markup Language (XML) mapping model to realize data interchange, and the process of registration, query and invocation of WebServices to realize data integration. The solution this paper presented has practical value for universities and large enterprises to integrate information systems.

Keywords SOA · WebService · XML · Data synchronization · Integration

191.1 Introduction

With the development of computer and network technology, many enterprises developed management information systems (MISs) to improve work efficiency, however, since having no unified plan and design, these MISs were developed by

G. Wang (✉)
School of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: wangguowei@hpu.edu.cn

Z. Zhang
Morden Educational Technology Center,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zhangzhibin@hpu.edu.cn

M. Xue
School of Architectural and Artistic Design,
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: xuemanjun@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_191,

different technology and different platform at different period, also, the database, operation system and program language are all different from each other, which brought the difficulty for data interchange and data synchronization. For this reason, these MISs worked in a semi-automatic state. Usually, we defined these MISs as "information islands" [1], and the solution of MIS integration became the prime way to resolve the problem of data synchronization in heterogeneous computing environment.

191.2 Corresponding Technologies Analysis

SOA. SOA is emerging as the premier integration and architecture framework in heterogeneous computing environment. The basic idea of SOA is service cored and integrates information resources base on standard service, which enables information resources reconfigurable and reusable. SOA is coarse-grained, loose coupled service architecture. Services communicate with simply and precisely defined interface, not involving bottom layer programming interface and communication models. Therefore, it is suitable to resolve some practical problems in enterprise business applications, such as distributed business services communication, information resource integration and so on.

WebService. WebService is a good realize technology of SOA at present. It has a series criteria and protocol which ensure the functions to be carried out, such as web services description Language (WSDL), UDDI, simple object access protocol (SOAP) and so on. WebService is loose coupled, the service requestor binds provider dynamically only in need, and it is platform independent. Because of communication with SOAP and XML message format, it requires application system supporting SOAP and XML only. WebService is very suitable for heterogeneous environment.

By encapsulating information, data object and business into WebServices, using SOA and WebService and XML can present an effective solution to resolve the problem of MISs integration based on the extensible, structural advantage of XML.

191.3 The Solution for MIS Integration Based on WebService

191.3.1 Architecture of the Solution

The solution can be divided into three tier architectures: application tier, service tier and resource tier.

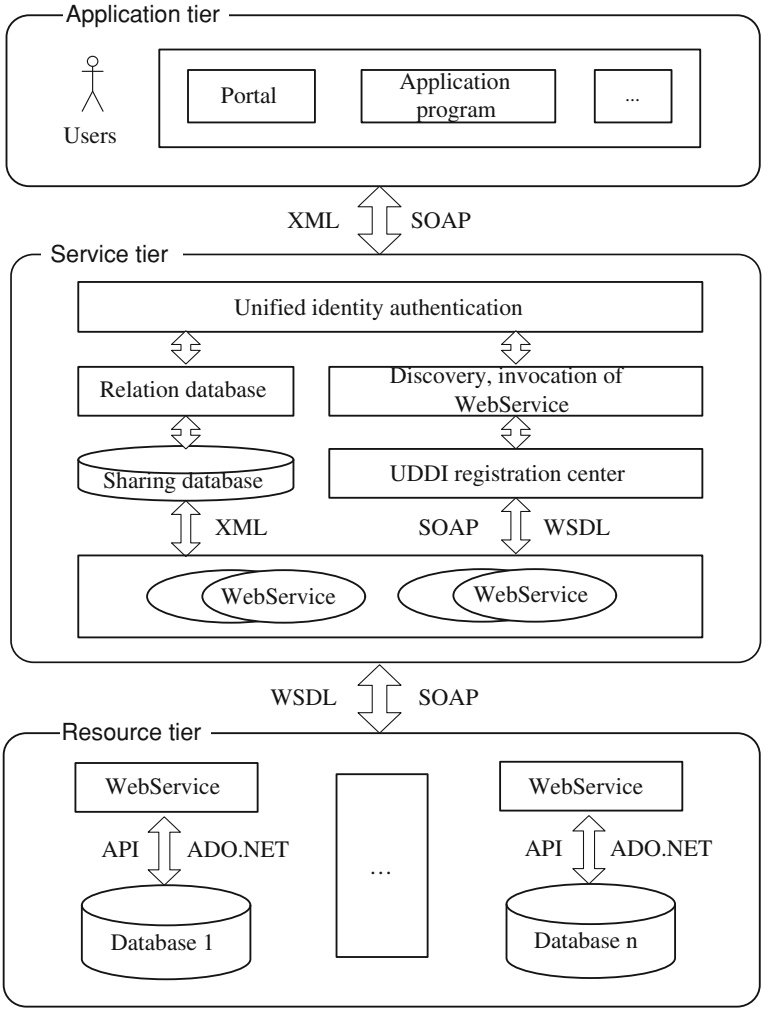


Fig. 191.1 The architecture of the solution

The architecture of the solution as shown in the Fig. 191.1. Resource tier is composed of all sorts of database of the MIS which need to be integrated. Service tier is the core of the solution, including WebService encapsulation, composition, registration, query, invocation and database mapping module. Application tier consists of the application client of MISs and information portal.

191.3.2 Function of Each Tier

Application tier. User or clients sent SOAP services request through application program or information portal.

Service tier. The service tier is the combination of SOA, XML and WebService, act as the core of the solution. Service tier accomplish these functions: data transmission and data interchange between service provider and service consumer, function of WebService registration, discovery and invocation, dynamic data interchange between XML and XML, and between XML and relation database.

Resource tier. Resource tier is the beginning of the SOA-based architecture. According to the demand of MIS integration, resource tier connect the database by API function or ADO.NET, then encapsulate the data into WebService. WebService is a SOA standard function unit use XML as transmission media to perform special mission, it consists of WebServices description (WSDL) and services data. The WSDL document defined the location, transmission protocol, interface of the WebService by XML. Unit WebService is the prime entity and the minimum service granularity can't be divided, and didn't include other WebServices.

191.4 Implementation of Key Parts

191.4.1 Table Structure of Database and XML Model Mapping Rules

The data format and data dictionary of MISS are all various, but data sharing and synchronization needs to work in a unified data format and rules, so a unified data model is needed. The structure of sharing database was designed according to "education management informatization standard one part: school management information standard".

The mapping rules between structure of database model and XML model as follows [2]:

- **<Table name="">…</Table>.** Each table in database corresponds a XML structure document, the property value of "name" equal to the table name of database.
- **<col name="" flag="" ref="">…</col>.** Sub-item of the XML document's root-item. The property value of the "name" equal to the column name of database table, and the property value of the "flag" show the column in database table is primary key or foreign key or is null and the property value of the "ref" is the table name of the foreign key related primary table.

191.4.2 WebServices Encapsulation of MIS

191.4.2.1 Encapsulation of API-Supported MIS

According to the WebServices standard, by encapsulating the function component into WebServices component, then describe the service by WSDL and write the data that has been extracted from the MIS into XML document. The whole procedure can be divided into the following stages: quotation of API function library, declaration of WebServices object, declaration of WebServices method, calling API function to operate database and release the quotation of local application object [3].

191.4.2.2 Encapsulation of ADO-Accessed MIS

Encapsulation of ADO-access MIS whose database is SQLServer or ORACL can extract data by SQL mainly depend on WriteXml() method of DataSet component in.Net framework to acquire XML document.

191.4.3 Integration and Transformation of WebServices

By using open XML standard and extensible stylesheet language transformations (XSLT) to analyze and extract XML document then integrate different XML documents and map XML document and relation database to realize data exchange and data sharing based on the XMLSchema of unit fine-grained WebServices.

191.4.3.1 Integration of Different XML Document

Based on the mentioned mapping rules, XPath and XQuery [4] the query language of XML transforms the XML document of unit WebService into unified structure XML document model by making the query progress an executable function series to meet the requirement of complicated query [5]. The whole progress as follow.

First step. Analyze the grammar of the XQuery request, then, read the transformed XML document.
Second step. Process the query according to XPath expressions of XQuery request.
Third step. Generate the goal XML document by unified mapping rules after the needed data has been retrieved.

The goal XML document is a union dataset of several databases of MIS that can be mapped and transformed to relation database and encapsulated into coarse-grained services and registered in UDDI registration center.

191.4.3.2 Mapping and Transformation from XML to Relation Database

The method of transformation from XML to relation database can be classified into structure mapping and model mapping. Compared to structure mapping method, model mapping has these advantages: support any static or dynamic XML data memory, support XML data memory without database model extension [6]. So this paper use table based mapping method of model mapping to map and transform XML into sharing relation database [7]. View structure defined XML document model as table or tables to map table or view in relation database, items or properties of XML document map the names of the columns of relation database, contents of the XML document map the values of the columns of relation database.

191.4.3.3 Mapping and Transformation from Relation Database to XML Document

The mapping and transformation include structure mapping and data transformation, at last generate XMLschema document and XML data document, the progress is similar with the encapsulation of WebServices.

191.4.4 Operation of WebServices in UDDI

191.4.4.1 Webservice Model Establishment

Before the establishment of WebService model, it is necessary to confirm the existence of the tModel file which is indispensable to the implementation of WebService. If the WebService based on the existed WSDL, record its name and tModelKey, otherwise, create a new tModel to replace the interface, the new tModel should has a unified resource name and the index to locate the WSDL file [8].

191.4.4.2 Webservice Registration

The registration can be realized by programming or user interfaces. The following codes is the registration progress of using Microsoft UDDI.NET SDK.

```
saveTModel stm = new SaveTModel();
stm.TModels.Add();
stm.TModels[0].Name = ''TestStuInfo'';
stm.TModels[0].OverviewDoc.OverviewURL = ''http://
192.168.11.67/TestWebService/TestStuService.asmx?WSDL'';
try
{   TModelDetail tmd = stm.Send();
```

```
   sTModelKey = tmd.TModels[0].TModelKey;
}
catch (UddiException ue)
{   Console.WriteLine (ue.Message);
   return;}
catch (Exception e)
{   Console.WriteLine (e.Message);
   return;
}
```

191.4.5 Webservice Discovery

The whole progress of inquiry program can be divided into the following steps: create inquiry object and set inquiry universal resource locator (URL) parameter; declare FindBusiness object and set the name and properties of the inquiry WebService, acquire the BusinessList object then search BusinessInfo, ServiceInfo, BusinessService Binding Template in turn, at last, acquire Access Point object and invoke Get Text() method to get the access URL of the service.

191.4.6 WebService Invocation

After the inquiry, services request got URL of the given WebService, then invoke it by using SOAP message through firewall, at last bind it into the application flow to achieve data exchange. In practical application, it is necessary to establish a proxy class that can communicate the given WebService. In order to using the proxy class, a CS file must be generated and added to the application project.

191.5 Conclusions

Data exchange and sharing between heterogeneous MIS is the key point of MIS integration. The solution this paper presented based on SOA, WebServices and XML under.net framework has practical value to universities and large enterprises. Because of little data security factor and access control has been concerned, next work will be focused on the research and implementation of Single Sign On (SSO) and Unified Identity Authentication (UIA).

Acknowledgments The research of this paper was supported by soft science research program of Henan province and youth fund program of Henan Polytechnic University. The authors thank the research team for the valuable discussions and suggestions.

References

1. Yujian Z, Yuelin Z (2009) Acta Scifntiarum Naturalium Universitatis Sunyaatseni. Build Online Trading Syst Relying Packaged Campus Card 48:108–110
2. Tao Z (2009) Design and implementation of heterogeneous data exchange based on XML and web services. Qingdao, China
3. Tianjing (2007) Research and application of enterprises integration based on SOA. Wuhan, China
4. Hong Z, Xuhui LI, Mengchi LIU (2009) Application research of computers. XML Query Approach Based Log XML Tree 26:609
5. Pengfei Z, Naiqian L (2010) Computer Engineering. XML-Based Distrib Database Integr Syst 36:71
6. Song Yuqing G, Chengquan L (2010) Application research of computers. Res Mapp Method XML Doc Relational Database 27:951
7. Liming W, Jun Z, Ming C (2007) Mapping mechanism based on table model between XML and database. J Military Commun Technol 28:46–47
8. Januszewski K (2001) Web service description and discovery using UDDI, Part I, http://msdn.microsoft.com/en-us/library/aa480517.aspx

Chapter 192
Application of Association Rule Mining in College Teaching Evaluation

Sanjing Geng and Zhenghui Guo

Abstract Association rule mining is the most important and most mature one of the research methods and techniques that can find the services a meaningful connection between the hidden and rules. This paper introduces the classical association rule mining algorithm Apriori algorithm and association rules combined with the teaching evaluation, evaluation of data from a large extract useful information hidden, so as to educational administration departments to provide decision support basis for improving teaching quality.

Keywords Association rule mining · Apriori algorithm · Teaching evaluation

192.1 Introduction

The quality of instruction is fundamental of the survival and development. To improve the quality of education, apart from strengthening the teaching administration, the University must introduce the necessary means of inspecting the quality of education, which is evaluating the quality of teaching. The teaching evaluation play the role of inspection and inspire in the in the process of teaching and is important measures and assurance.

S. Geng (✉)
School of Computer Science and Technology, Henan Polytechinc University, Jiaozuo, China
e-mail: gengsj@hpu.edu.cn

Z. Guo
Modern Education and Technology Center, Henan Polytechnic University, Jiaozuo, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_192,

At present, University uses the main method which is network evaluation. The method is based on internet teaching evaluation system and collects data easily and completes statistics and analysis functions and shortens appraisal period. To some extent, the method guarantees the quality evaluation. But the data collected have the very rate of utilization, generally teaching evaluation system is simply calculate and statistics and then gets the evaluation of all the result which belong to all teachers in the terms. The university assesses the different teachers by the different levels which are divided into great and good and qualified and unqualified, etc. The end of the evaluations and promotions is decided by your level. They don't analyze profoundly data collected. Don't mine and think inside information. Long a time, there is no better way or proposals how to boost the teacher's teaching quality. Just as every teaches the term of the teaching quality inspection of an instrument, had not fundamentally enhance the teacher's teaching and the quality of instruction.

Therefore, problems and weaknesses in the system of institutions of higher education are changed from the technology and further use the cumulative number of data, the university changes the existing data to available knowledge. Easy to provide support for the school management decisions and to improve the management level and quality of teaching methods, which has become one of the problem for solution.

192.2 Association Rule Mining Algorithms

192.2.1 The Theory of Associated Rules

Association rule mining is the most important and one of the most active research methods, which is brought forward by Agrawal in 1993 [1].It finds interesting association or correlation relationships among a large set of data items.

A typical example of association rule mining is market basket analysis [2]. This process analyzes customer buying habits by finding associations between the different items that customers place in their "shopping baskets". The discovery of such associations can help retailers develop marketing strategies by gaining insight into which items are frequently purchased together by customers.

There are two important concept, support and confidence [3] when people description the associated rules. The former is used to measure the associated rules on the importance of he whole data, the latter measure is the its credibility of rules. Generally speaking, the user is interested in the rules when the pertinent rules have the support and credibility.

In general, given a transactional database, association rule mining is the procedure of the associated rules when the user set the little support and the credibility.

Association rule mining is a two-step process:

1. Find all frequent itemsets: By definition, each of these itemsets will occur at least as frequently as a pre-determined minimum support count.
2. Generate strong association rules from the frequent itemsets: By definition these rules must satisfy minimum support and minimum confidence.

192.2.2 The Apriori Algorithm: Finding Frequent Itemsets

The Apriori Algorithm is put forward by Agrawal in 1994 [2]. It is applied to find the frequent itemsets of database. So far as the algorithm is still the classic association rule mining algorithms have been widely discussed and applied.

Apriori is an influential algorithm for mining frequent itemsets for Boolean association rules. The name of the algorithm is based on the fact that the algorithm uses prior knowledge of frequent itemset properties, as we shall see below.

Algorithm 1 Apriori[4]
Input:Database, D, of transacions, minsup_count.
Output: L, frequent itemsets in D.
L1 = find_frequent_1-itemsets(D);
FOR (k = 2; Lk-1 $\neq$ Φ; k ++){

Ck = apriori-gen(Lk − 1, min_sup);
FOR each transactions t∈D

Ct = subset (Ck, t);

FOR each candidates c∈Ct {

c.count ++;
}
Lk = { c∈Ck | c.count ≥ min_sup }
}
Return L = ∪k Lk

192.2.3 The Generation Algorithm of Association Rule

All frequent itemsets is obtained through apriori algorithms, associated rule are generated through algorithms as following:

Algorithm 2 From a given frequent itemsets create strong association rule [4].
Input: frequent itemsets: minimum confidence minconf
Output: strong association rule.
Rule-generate (L, minconf)

FOR each frequent itemset lk in L
Genrules(lk,lk);

192.3 Application of Association Rule Mining in Teaching Evaluation

192.3.1 Data Collection and Preprocessing

The paper takes the evaluation data in the term for example, which digs the associated rules.

Data collection uses data in the two tables. One is form from educational evaluation of the database which makes use of main field such as teacher number teacher names teaching attitude and teaching ability, another is the table from the personnel management system such as another is from the personnel management system of export interest to the table.

The export value of data from educational evaluation system is 139388, Evaluation of teachers was 590. The results of the evaluation of each teacher are calculated according to a certain approach. The outstanding teachers is 82, accounting for 14% of teachers being evaluated, a good teacher is 167, accounting for evaluation of teachers 28%, medium teachers is 201, accounting for 34%, qualified teachers is 141, accounting for 24% of teachers being evaluated.

Given database will exist in general incomplete, noisy and inconsistent data, such as records of individual property missing values or incorrect property values [5]. The same property value has repeated and may the result of data or data type inconsistencies. Therefore, to address the various situations that may arise pre-analysis to ensure the accuracy of data, improve data quality to ensure mining results.

The raw data pre-processing collected has three steps which are cleaning data, eliminating duplicate records, converting data type. The Apriori algorithm is a Boolean association rules mining frequent itemsets algorithm. So the quantitative data collected has certain treatments, such as a table of teacher age, education, job title, teaching attitude, teaching ability and other indicator, these quantitative data into discrete Boolean data processing.

192.3.2 Mining and Analyzing Results

Apriori algorithm used in the evaluation of the level of good data handling good and good for the 248 records for mining. Suppose the minimum support association rule is 10%, minimum confidence is 50%, the association rule mining produces 25 strong association rules. 10 of them restore the code according to the rules defined in each rule to calculate the rules support and confidence, and the results shown in Table 192.1.

Among them, Rule 2 is between the ages of 22–35 years old and the title of instructor evaluation rating of good teachers, support is 33%, confidence level is 93%; rule 16 that the title of associate professor and the evaluation rating of good teachers support is 18%, confidence level is 66%; other rules can be used for a similar interpretation.

Table 192.1 Part of the association rules

Rule	Sex	Age	Title	Education	Grade	Support/%	Confidence/%
1		22–35	Lecturer	Master		40	96
2		22–35	Lecturer		Good	33	93
5		22–35	Lecturer	Master	Good	28	87
11			Professor		Good	10	71
16			Associate professor		Good	18	66
12				Master	Good	49	70
25				Bachelor	Good	10	61
17		22–35			Good	39	66
23		36–50			Good	21	64
15	Male				Good	39	67

The following analysis:

1. Teachers who is between the ages of 22–35 years old and have the title of lecturer and the degree of Master accounted for the master's relatively large proportion of teachers and students' evaluation is good.
2. Teacher in the title of professor or associate professor has relatively small ratios and good evaluations.
3. Between the ages of 36–50 years old, middle-aged teacher evaluation is good.

By the above rules can be found: At present, young teachers are the backbone of school teaching, high academic qualifications, knowledge breadth, teaching quality. High-grade, highly educated older teachers, teaching experience, teaching and research levels are higher, and thus the quality of teaching is better, the title of high school teachers should be encouraged for the class, teachers improve high-grade classroom teachers in the ratio. Timetable for the class, the classroom teachers should pay attention to age, education and titles allocation. Schools should engage in teaching and research activities for young teachers and high-grade, highly educated teachers should learn more exchanges each other and continue to progress and common development

The record contains 590 data sheet library as a data mining excavation. Suppose the minimum support association rule is 20%, minimum confidence of 80%, the association rule mining produced 20 strong association rules. To which the rule 9 of the code reduction calculated according to the definition of each rule support and confidence degrees, and the results shown in Table 192.2.

Among them, Rule 5 that good teaching organization and teaching methods, teaching on the good support is 86, 98% confidence level; rule 17 that the effect on the teaching ability of good teaching is good support 85% confidence level is 96%; other rules can be used for a similar interpretation.

The following analysis:

1. Rules 2, 3, 5 and 18 show that the teaching attitude of teacher focuses on the organization of teaching, class high utilization rate, flexible teaching methods, classroom atmosphere, good teachers who obtain teaching achievements.

Table 192.2 Part of the association rules

Rule	Teaching attitude	Teaching ability	Teaching content	Teaching organization	Teaching methods	Teaching effect	Support/ %	Confidence/ %
2	Good			Good			86	99
3				Good	Good		88	99
5				Good	Good	Good	86	98
8		Good		Good			87	98
9					Good	Good	87	98
10		Good		Good		Good	85	98
12		Good				Good	86	97
14		Good	Good				85	97
18				Good		Good	89	96

2. Rules 8, 10, 12 the a firmer foundation of the good teacher have well-organized lectures, teaching ability, and focus on the organization, teaching good effect.
3. Rule 9 that the use of good teaching methods can achieve good teaching results.
4. Rule 14 shows that solid teaching basic skills, teaching ability, teaching content focused, moderate depth, and integrate theory with practice and extensive knowledge.

By the above rules can be found: the teaching attitude of teachers to obtain the necessary prerequisite for good teaching effectiveness, teachers with good teaching ability grasp the teaching content and teaching organization based in the teaching process should focus on flexible use of different teaching method, so as to achieve good teaching.

192.4 Conclusion

The main purpose of the paper is to apply association rule mined to the current teaching evaluation system and uses the accumulation of a large number of schools for each semester of teaching evaluation data, combined with the basic information database of teachers and student achievement database, the use of association rule mining, large amounts of data from in data mining, the entire teaching and learning activities for teachers to evaluate and analyze, explore the teacher's gender, age, title, education, teaching attitude, teaching ability, teaching content, teaching organization, teaching methods, teaching effectiveness indicators and their relationship and education, job title, teaching attitude, teaching ability, teaching content, teaching organization, teaching methods, the association between relationship. It can provide a reasonable allocation of the class teacher to stimulate student motivation for learning, encourage teachers to better carry out teaching activities, educational administration departments to provide decision support based on improving the quality of teaching.

References

1. Han J, Kamber M (2001) DATA MINING concepts and techniques. Higer Education Press, BeiJing
2. Shao F (2009) Principle and algorithm of data mining. Science Press, BeiJing
3. Yao J (2009) The principle and application of data warehouse and data mining technology. Plublishing House of Eletronics Industy, BeiJing
4. Mao G (2005) Principle and algorithm of data mining. TsingHua University Press, BeiJing
5. Zheng J (2009) Application of association rules in teaching appraisal. J Yangtze Uni Nat Sci Edit 6(4):149–150 (Sci&Eng)

Chapter 193
Support Vector Machine Selective Ensemble Learning on Feature Clustering and Genetic Algorithm

Hui Tao, Xiao-ping Ma and Mei-ying Qiao

Abstract A novel ensemble learning algorithm, Support Vector Machine Selective Ensemble Learning on Feature Clustering and Genetic Algorithm, is proposed to improve ensemble classifier performance in this chapter. First, features are clustered on weighted average linkage method and feature subsets consisting of the attributes extracted from each category at random serve as the inputs of individual SVM to ensure that there are significant diversities among individual learning machine. Then GA is adopted to optimize SVM parameters to increase individuals' classification performances. Finally, selective ensemble learning enhances the accuracy of ensemble classification system. Through six data sets simulation, the results show that classification effect of the novel algorithm was obviously improved.

Keywords Ensemble learning · Selective ensemble · Feature clustering · GA · SVM

H. Tao (✉) · M. Qiao
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: 65724138@qq.com

M. Qiao
e-mail: qiaomy@foxamil.com

H. Tao · X. Ma · M. Qiao
School of Information and Electrical Engineering,
China University of Mining and Technology, Xuzhou 221116, China
e-mail: Xpma@cumt.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_193,

193.1 Introduction

Support vector machine (SVM) ensemble algorithm, as an ensemble learning method of training multiple individual classifiers and combining their recognition results, can raise classifier's recognition accuracy and further enhance classification system robustness, and become one of the research focuses in pattern recognition and knowledge mining fields, as it takes advantage of information complementary of individual SVM. In order to get a preferable ensemble result, individual learning machine must have significant diversity and high accuracy [1].

There is an ensemble feature selection (EFS) method through disturbing sample feature space to increase individual learning machine's difference. The studies show that the EFS method can not only significantly improve classification system generalization performance, but also effectively avoid "dimension disaster", meanwhile, it reduces the time complexity in the training and classifying process [2]. Therefore, EFS method has become a favorite with many researchers, for instance, Brylla suggested an "attribute bagging" by using random feature subsets [3], Hu raised an ensemble rough subspace [4]. Akay presented an SVM combined with feature selection for breast cancer diagnosis [5]. In this chapter, we propose a new feature subspace ensemble algorithm on feature clustering, which the individual SVM is trained according the reduced feature subsets that are different from one another through feature clustering and selecting in our algorithm. In addition, the ensemble learning performance will be enhanced when parts of individual models are integrated instead of all [6], so selective ensemble learning is adopted where parts of all are combined on the basis of their evaluation results.

The SVM, on the other hand, can gain expected classification accuracy only after its parameters (kernel parameter c and penalty coefficient g) are properly adjusted in classifying, and though grid search can succeed in the optional (c and g) selection through cross validation, sometimes it is very time-consuming in a wide search range. However, genetic algorithm (GA), as a heuristic algorithm that can find the global optimal solution even not traversing all mesh points, is a practical and efficient optimization technology with strong robustness [7]. So GA is used to optimize individual SVM parameters, validation samples accuracy as fitness function, and classification accuracy of individual and ensemble models can all be enhanced.

193.2 GA-SVM Selective Ensemble Learning Algorithm on Feature Clustering

193.2.1 Feature Subset Selection Based on Feature Clustering

For feature clustering, first, the matrix of correlation coefficient between features are calculated according to sample data.

Correlation coefficient between features. Correlation coefficient between feature x_i and x_j is defined as $C_{ij}(2)$.

$$C_{ij}(2) = \frac{\sum_{k=1}^{n}(x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\left[\sum_{k=1}^{n}(x_{ki} - \bar{x}_i)^2\right]\left[\sum_{k=1}^{n}(x_{kj} - \bar{x}_j)^2\right]}}, \; i = 1, 2, \cdots p, j = 1, 2 \cdots p \tag{193.1}$$

where p is the amount of features, n is the amount of samples, $\bar{x}_i$ is the average value of all samples' feature x_i, $\bar{x}_\mathrm{j}$ is the average value of feature x_j.

$$\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki} \tag{193.2}$$

$$\bar{x}_j = \frac{1}{n}\sum_{k=1}^{n} x_{kj} \tag{193.3}$$

Feature clustering on weighted average linkage method. There are three common feature clustering methods including hierarchical clustering, K-means clustering, and fuzzy c-means clustering and the category amount must be determined in advance of the last two. Normally, the category number of features can't be predetermined [8], so we use hierarchical clustering methods to cluster feature.

The distance between x_i and x_j is defined as d_{ij}.

$$d_{ij} = 1 - C_{ij} \tag{193.4}$$

The distance between categories is defined using weighted average linkage(WAL) method, for the method, as a preferable method, takes full advantage of the information about all variables. The square of the distance G_K and G_L is:

$$D_{KL}^2 = \frac{1}{n_K + n_L}\sum_{x_i \in G_K, x_j \in G_L} d_{ij}^2 \tag{193.5}$$

where n_K is the feature amount of G_K, and n_L is the feature amount of G_L.

The recursion formula of the distance square between categories is:

$$D_{MJ}^2 = (1 - \beta)\left[\frac{n_K}{n_M}D_{MK}^2 + \frac{n_L}{n_M}D_{ML}^2\right] + \beta D_{KL}^2 \tag{193.6}$$

where n_M is the feature amount of G_M, $\beta < 1$ and in this chapter β is the default value 0.25.

Feature Subset Extraction. Clustering tree diagram is gotten through hierarchical clustering feature, and in the diagram it is clear that the clustering results and the distance between various features. To ensure the differences between individual SVM, reduced attribute subsets consisting of the features successively extracted from each category that distance is less than the threshold are acted as the input variables of each individual SVM. Thus, the input variables of each

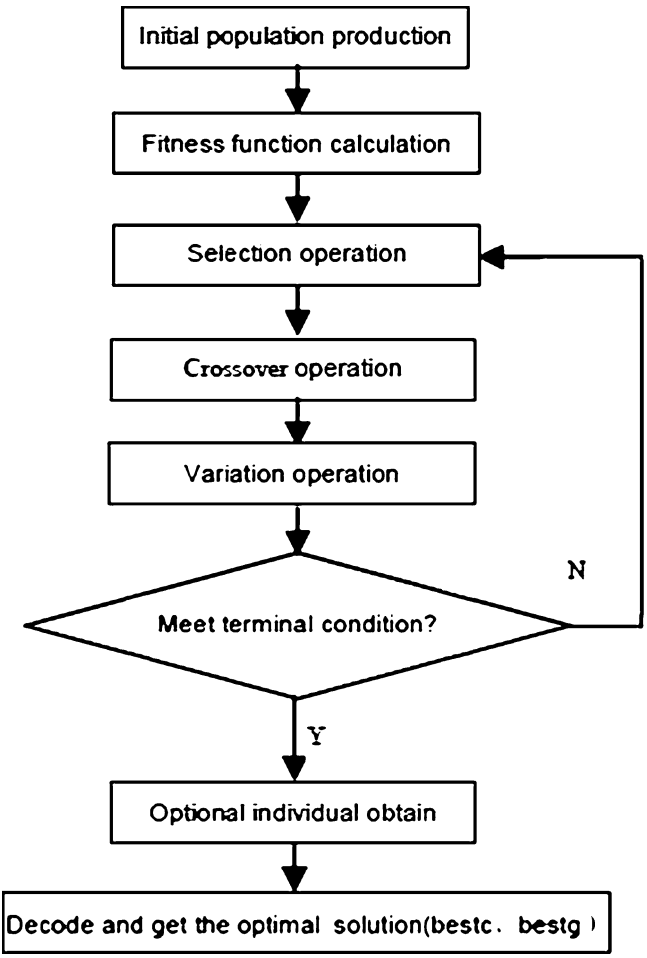


Fig. 193.1 Algorithm flow chart of GA to optimize individual SVM

individual are different, and there must be a certain difference between each individual.

193.2.2 GA-SVM

GA is adopted to optimize individual SVM parameters to increase classification accuracy of individual model [9]. In the genetic algorithm, SVM parameters are coded by binary system that $c \in (0, 100]$ and $g \in [0, 1000]$, fitness function is designed as average classification accuracy on K-fold cross validation that after the

training data are divided equally into K groups we take each subset data as validation set, other K-1 subset as train set, then K models are established, fitness function is the average of classification accuracy of K models validation set.

$$f(X) = \frac{1}{K}\sum_{k=1}^{K} f_1(k) \tag{193.7}$$

where $f_1(k)$ is the classification accuracy of the kth model.

The algorithm flow of GA to optimize individual SVM is shown in Fig. 193.1.

193.2.3 Selective Ensemble Learning

Though there is some diversity among individual SVM established according to different feather subset and individual SVM identifying precision is greatly increased by GA, there is no guarantee that the identifying precisions of all individuals are high because feathers in feather subset are selected from categories after clustering at random. So we employ selective ensemble and its procedures is shown as follows [6].

First step: The GA-SVM average identifying accuracy of cross validation set in the optional parameters is given according to the train produce.

Second step: the individual models are sorted from great accuracy to the little.

Third step: the first l individual models are selected to generate ensemble model.

193.2.4 Algorithm Flow of GA-SVM Selective Ensemble on Feature Clustering

Based on the ideas mentioned above, the algorithm flow of GA-SVM selective ensemble on feature clustering is shown in Fig. 193.2.

193.3 Experimental Simulations

193.3.1 Data Description

For proved the feasibility and effectiveness of the proposed algorithm, the most commonly binary-classification data in UCI Machine Learning Repository is adopted to simulate. The detail information of experimental data set is shown in Table 193.1 [10].

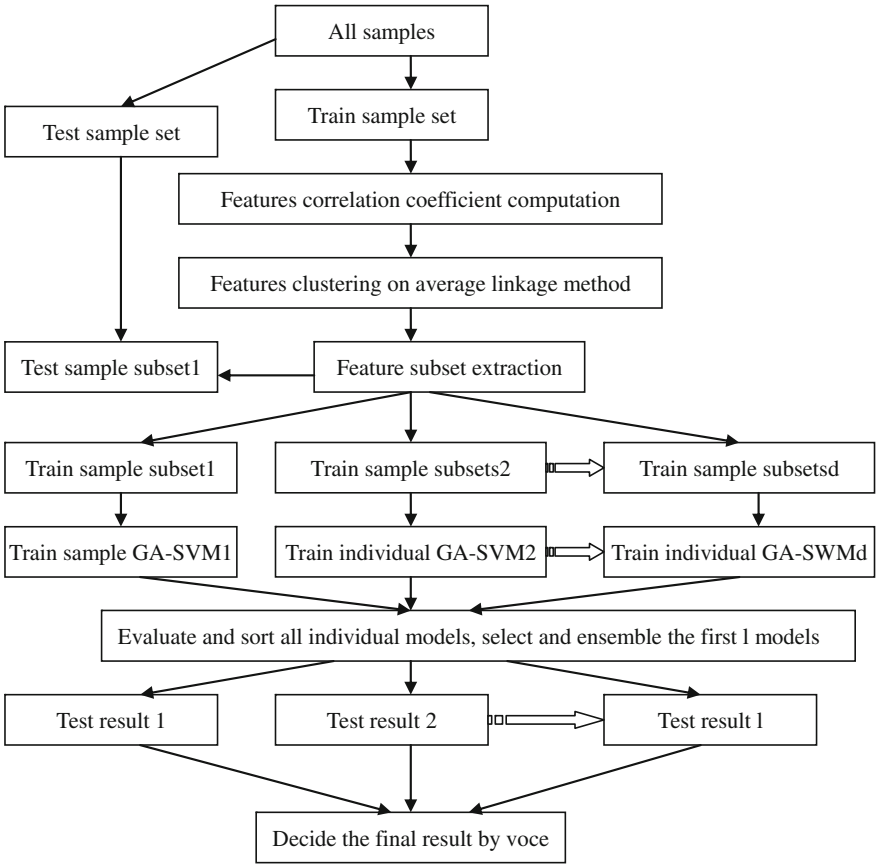


Fig. 193.2 Algorithm flow of GA-SVM selective ensemble on feature clustering

193.3.2 Simulation Results

Now we first take Wdbc data set as an example, and give the step-by-step results in experiment process. The feather clustering tree diagram is shown in Fig. 193.3. The feather category is 15 when the threshold of distance between categories is defined as 0.3 and the 15-dimension feather subset consists of the feathers randomly extracted from each category, then 20 feather subsets is used to train 20 individual GA-SVM and 11 individuals with higher identifying accuracy (shown in Table 193.2) are selected to emsemble.

Table 193.1 Detail information of experimental data set

Data set	Number of samples	Number of attributes	Number of train samples	Number of test samples
Wdbc	569	30	400	169
Heart	270	13	180	90
Ionosphere	351	34	250	101
Sonar	208	60	120	88
German Credit	1,000	24	700	300
Australian	690	14	500	190

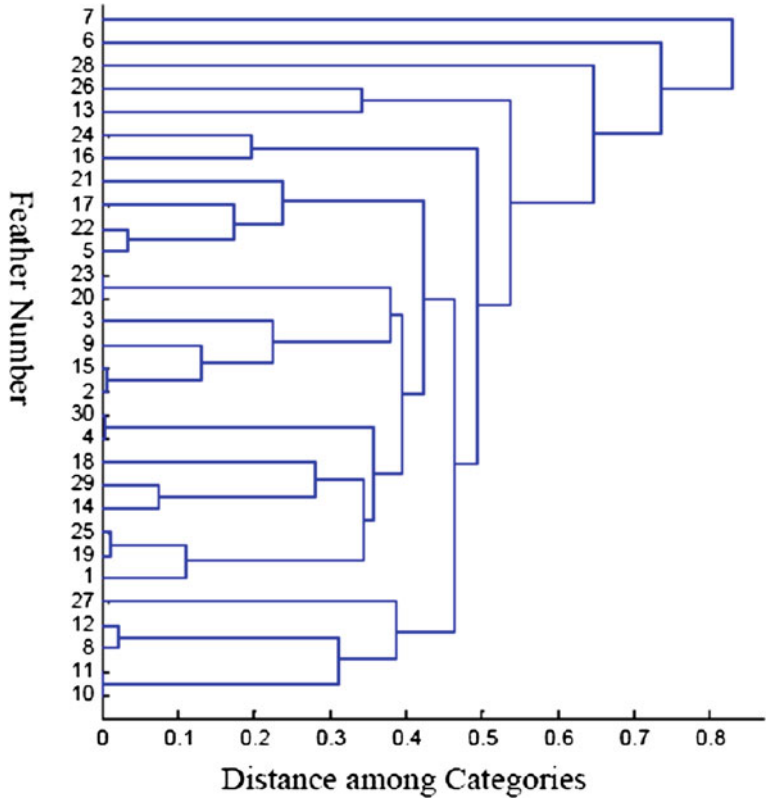


Fig. 193.3 Feather clustering tree diagram of Wdbc

The simulation processes of other 5 data sets are the same as Wdbc. Here we do not give the detailed intermediate results. The classification accuracy comparison of six data sets is shown in Table 193.3 with different algorithm including single GA-SVM, Bagging, AdaBoost, traditional SVM selective ensemble (SVM-SE), all GA-SVM ensemble (GA-SVM-E), and GA-SVM selective ensemble (GA-SVM-SE).

Table 193.2 the information of 11 individual SVM selected

Individual GA-SVM number	Feather number of subset component	Optional parameters (c,g)	Validation accuracy (%)	Test accuracy (%)
1	1,6:10,13,21,23,24,26:30	8.91, 15.24	93.5	92.90
2	4,6:8,10,13,15,19,22,24,26:29	14.79, 5.12	95	95.27
3	1,4,6:9,11,13,16:18,23,26:28	83.44, 15.72	94.25	92.31
4	3,6,7,11:13,18,20,21,24:28,30	21.26, 0.17	95	93.49
5	3,4,6,7,10,12,13,18,21,23:28	2.64, 12.12	94.75	94.08
6	2,6:8,11,13,16,19,21,23,26:30	1.56, 20.49	95.25	94.08
7	1,4,6:8,11,13:15,17,20,24,26:28	7.36, 1.56	94.75	94.08
8	6,7,9,11:13,16,22,23,25:30	5.90, 4.11	97.5	95.27
9	4,6:10,13,14,17,23:28	53.54, 2.73	95.75	92.90
10	1,6,7,11:14,16,22,23,26:28,30	59.55, 20.42	95.25	95.27
11	1,2,5:8,10,13,18,23,24,26:28,30	25.77, 0.11	97.5	95.27

Table 193.3 Classification accuracy (%) comparison of different algorithm

Data set	Single GA-SVM	Bagging	AdaBoost	SVM-SE	GA-SVM-E	GA-SVM-SE
Wdbc	94.67	95.07	95.86	89.25	95.27	97.04
Heart	76.67	82.22	81.11	80.00	82.96	85.50
Ionosphere	90.09	92.07	93.06	91.08	93.06	95.04
Sonar	71.59	76.14	82.95	79.54	84.09	85.22
German Credit	69.67	70.67	71.33	70.00	70.67	71.66
Australian	85.78	86.31	86.31	84.21	85.26	86.84

193.3.3 Discussions

From Table 193.3, it is found that the performances of 5 kinds of ensemble algorithms are significantly improved compared to that of single model, in particular, the classification accuracy of GA-SVM-SE algorithm proposed in this chapter is higher than traditional ensemble algorithm such as AdaBoost and Bagging to some extent. Moreover, based on feather clustering, GA and selective ensemble increase precision of classification system to a certain degree.

193.4 Conclusions

A novel ensemble learning algorithm, GA-SVM selective ensemble learning algorithm on feature clustering, is proposed to improve ensemble classifier performance in this chapter. First, features are clustered on weighted average linkage method and feature subsets consisting of the attributes extracted from each category at random serve as the inputs of individual SVM to ensure that there are significant diversities among individual learning machine. Then GA is adopted to

optimize SVM parameters to increase individuals' classification performances. Finally, selective ensemble learning enhances the accuracy of ensemble classification system. Through six data sets simulation, the results show that the classification effect of the novel algorithm was obviously improved.

Acknowledgments The authors are grateful to the National Natural Science Council of China and the Natural Council of Jiangsu Province for their financial supports (No. 60974126 and BK2009094).

References

1. Kim HC, Pang S et al (2003) Constructing support vector machine ensemble. Pattern Recogn 12:2757–2767
2. Gunter S, Bunke H (2004) Feature selection algorithms for the generation of multiple classifier systems and their application to handwritten word recognition. Pattern Recogn Lett 11:1323–1336
3. Brylia R, Gutirrrez OR, Queka F (2003) Attribute bagging: improving accuracy of classier ensembles by using random feature subsets. Pattern Recogn 6:1291–1302
4. Hu QH, Yu D, Xie ZX et al (2007) EROS-ensemble rough subspaces. Pattern Recogn 12:3728–3739
5. Akay ME (2009) Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst App 36:3240–3247
6. Zhou ZH, Wu JX, Tang W (2002) Ensembling neural networks: many could be better than all. Artif Intell 137:239–263
7. Wang KQ, Yang SC, Dai TH, Bai XB (2009) Method of optimizing parameter of least squares support verctor machine by genetic algorithm. Comput App Softw 7:109–111
8. Zha WX, Li M, Fu YP (2010) Empirical analysis of highway traffic accident in average linkage clustering method. J East China Jiaotong Univ 6:113–118
9. Faruto, Li Y (2009) LIBSVM-farutoUltimateVersion-a toolbox with implements for support vector machines based on libsvm. Software available at http://www.ilovematlab.cn
10. Frank A, Asuncion A (2010) CI machine learning repository. University of California, School of Information and Computer Science, Irvine

Chapter 194
Research on Active Vibration Control of Piezoelectric Intelligent Beam Based on Energy Finite Element Method

Zhen Zhou, Rui Huo and Xiufang Zhang

Abstract The transmission of vibration energy and equilibrium relationship is analyzed. Discrete the energy-balance equation by Galerkin weighted residual method and the Energy Finite Element model of Piezoelectric intelligent cantilever beam system is developed. Energy density wave solution was deduced by classics wave theory, Energy Flow Element model was proved correct by comparing the EFEM solution with exact wave solution. Finally, makes an exploration and analysis on the influence of steady state response and energy density distribution for feedback gain, location and different pairs of piezoelectric elements.

Keywords EFEM · Energy density · Piezoelectric elements · Vibration

194.1 Introduction

Piezoelectric intelligent structures means stick piezoelectric materials on light damping system which anti-interference function is too bad. In recent years, active vibration control systems using piezoelectric materials as distributed sensors and actuators (S/As). Because of the good prospect of application and extension in aerospace, robot, information technology, IT, materials science, and other high technologies, many scholars have made a deep research and gains rich achievements.

Z. Zhou (✉) · R. Huo · X. Zhang
School of Mechanical Engineering, Shandong University,
No.17923 of Jingshi Road, Jinan 250061, China
e-mail: zz2326@163.com

R. Huo
e-mail: huorui@sdu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_194,

As the object of study, piezoelectric intelligent beam's dynamic modeling and optimal design of structure parameter and control low are the problems we concerned about most. External perturbation is always treated as pulse excitation, obtaining the system's state-space equation through modal theory. Taking the smallest vibration and control energy as the objective function, the locations of the piezoelectric sensors and actuators were optimized [1]. In Ref. [2], governing equation was established based on steady stimulations and then made an experimental research on frequency response characteristic. In Ref. [3], the boring bar vibration was controlled by using piezoelectric materials, simulation results suggest that the control effect was well.

Analysis and design of a complicated engineering system structure, usually involving the application of finite element method. In Ref. [4], the method for modeling piezoelectric intelligent structures using finite element method(FEM) was investigated and finite element dynamic equation of piezoelectric intelligent structures was observed. Research on the piezoelectric intelligent structures by FEM has been recently a hotspot in vibration control field. Unfortunately, traditional FEM exists some formidable difficulties, the major cause of numerical result sensitive to the boundary conditions, damping and physical dimension changes is overlapping structural mode [5]. Vibration energy of the subsystem is taken as the variable in EFEM, mode density is not assume at all when establish energy-balance equation and more rigorous than SEA when describe this equation. EFEM has prodigious development foreground in solving high frequency of structure vibration.

In this study, the transmission of vibration energy and equilibrium relationship is analyzed. Discrete the energy-balance equation by Galerkin weighted residual method and the Energy Finite Element model of Piezoelectric intelligent cantilever beam system is developed. Energy density wave solution was deduced by classics wave theory, Energy Finite Element model was proved correct by comparing the EFEM solution with exact wave solution. Finally, makes an exploration and analysis on the influence of steady state response and energy density distribution for feedback gain, location and different pairs of piezoelectric elements.

194.2 Energy Finite Element Model of Piezoelectric Cantilever System

194.2.1 System Model

To begin, one considers a model of piezoelectric cantilever system in Fig. 194.1a, excited in flexure by a harmonically varying force of magnitude F0 applied at its end. The cantilever could be divided into four elements l1, l2, l3 and l1, hp and hb represent the thickness of Piezoelectric Patches and cantilever beam respectively. Figure 194.1b shows the vibration energy inflow and outflow of cubic element, q and πdiss represent power flow through cubic element and the energy dissipated by

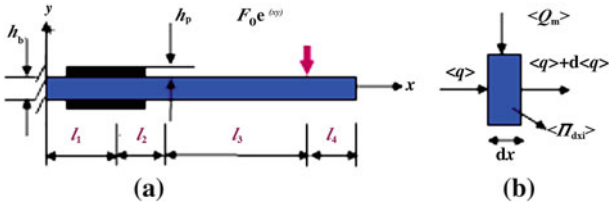


Fig. 194.1 Piezoelectric cantilever system model and energy flow balance

damping respectively, Qin is the energy input from external, At each point of a structure the well known instantaneous energy balance is given by [6]

$$\partial e/\partial t = -\partial q/\partial x - \pi_{\text{diss}} + Q_{\text{in}} \tag{194.1}$$

where ∂e/∂t is the time rate of change of energy density. The relationship between <*q*>, <πdiss> and <*e*> follow as

$$< q > \; = -(c_g^2/\eta\omega) \cdot (d < e > /\mathrm{d}x) < \pi_{\text{diss}} > \; = \eta\omega < e > \tag{194.2}$$

By substituting expression (194.2) in Eq. (194.1) could find the energy balance equation is related by

$$-\left(c_g^2/\eta\omega\right) \times \left(\mathrm{d}^2 < e > /\mathrm{d}x^2\right) + \eta\omega < e > \; = < Q_{\text{in}} > \tag{194.3}$$

where $c_g = 2(\omega^2 EI/\rho A)$ is the group velocity of bending wave, η is the damping loss factor, <> is time-averages of the variables.

194.2.2 System Energy Density Control Equations With No Feedback Control

Harmonic excitation is the only input energy injected into system when there is no feedback control, we can discrete expression (194.3) by Galerkin weighted residual method

$$\boldsymbol{K}^{(n)}\boldsymbol{e}^{(n)} = \boldsymbol{P}^{(n)} + \boldsymbol{Q}^{(n)} \tag{194.4}$$

$$\begin{cases} K^{(n)} = \int\limits_{x_L^{(n)}}^{x_R^{(n)}} \left(-\frac{c_g^2}{\omega\eta}\frac{\mathrm{d}\boldsymbol{N}^T}{\mathrm{d}x}\frac{\mathrm{d}\boldsymbol{N}}{\mathrm{d}x} + \omega\eta\boldsymbol{N}^T\boldsymbol{N}\right)\mathrm{d}x \\ \boldsymbol{P}^{(n)} = \boldsymbol{N}\left(x_R^{(n)}\right)\langle q\rangle\Big|_{x_R^{(n)}} - \boldsymbol{N}\left(x_L^{(n)}\right)\langle q\rangle\Big|_{x_L^{(n)}}, \boldsymbol{Q}^{(n)} = \int\limits_{x_i}^{x_j} \boldsymbol{N}\langle Q_{\text{in}}\rangle\mathrm{d}x \end{cases}$$

where $\boldsymbol{K}^{(n)}$ is a Coefficient Matrix related to stiffness and mass, $\boldsymbol{P}(n)$ is the power flow between nodes, $\boldsymbol{Q}(n)$ is the external energy input, $\boldsymbol{N}(x)$ is the lagrange linear interpolation basic function.

For solving expression (194.4) connection relation between elements is needed. There are two adjacent elments which number are i and $i+1$, $\tau^{(i,i+1)}$ and $\gamma^{(i,i+1)}$ represent transmission efficiency and reflection coefficient between coupling nodes

$$\begin{aligned} langleq^{-}\rangle_R^{(i)} &= \gamma^{(i,i+1)}\langle q^{+}\rangle_R^{(i)} + \tau^{(i,i+1)}\langle q^{-}\rangle_L^{(i+1)}, \\ \langle q^{+}\rangle_L^{(i+1)} &= \tau^{(i,i+1)}\langle q^{+}\rangle_R^{(i)} + \gamma^{(i,i+1)}\langle q^{-}\rangle_L^{(i+1)} \end{aligned} \tag{194.5}$$

The net input power flow for each element

$$\begin{aligned} \langle q\rangle|_{x_R^{(i)}} &= \tau^{(i,i+1)}\langle q^{-}\rangle_L^{(i+1)} - \left[1-\gamma^{(i,i+1)}\right]\langle q^{+}\rangle_R^{(i)} \\ \langle q\rangle|_{x_L^{(i+1)}} &= \tau^{(i,i+1)}\langle q^{+}\rangle_R^{(i)} - \left[1-\gamma^{(i,i+1)}\right]\langle q^{-}\rangle_L^{(i+1)} \end{aligned} \tag{194.6}$$

Simultaneous expression (194.5) and expression (194.6) can obtain

$$\begin{bmatrix} \langle q\rangle|_{x_R^{(i)}} \\ \langle q\rangle|_{x_L^{(i+1)}} \end{bmatrix} = \alpha^{(i,i+1)} \begin{bmatrix} c_g^{(i)} & -c_g^{(i+1)} \\ -c_g^{(i)} & c_g^{(i+1)} \end{bmatrix} \begin{bmatrix} \langle e\rangle_R^{(i)} \\ \langle e\rangle_L^{(i+1)} \end{bmatrix} \tag{194.7}$$

where $\alpha^{(i,i+1)} = 0.5\tau^{(i,i+1)}/\gamma^{(i,i+1)}$.

Cantilever was divided into four elements at least (Fig. 194.2) and $\gamma^{(1,\ 2)} \neq 0$, $\gamma^{(2,\ 3)} \neq 0$, simultaneous expression (194.4)–(194.7)

$$\begin{bmatrix} k_{11}^{(1)} & k_{12}^{(1)} & 0 & 0 & 0 & 0 & 0 \\ k_{21}^{(1)} & k_{22}^{(1)} - \alpha^{(1,2)}c_g^{(1)} & \alpha^{(1,2)}c_g^{(2)} & 0 & 0 & 0 & 0 \\ 0 & \alpha^{(1,2)}c_g^{(1)} & k_{11}^{(2)} - \alpha^{(1,2)}c_g^{(2)} & k_{12}^{(2)} & 0 & 0 & 0 \\ 0 & 0 & k_{21}^{(2)} & k_{22}^{(2)} - \alpha^{(2,3)}c_g^{(2)} & \alpha^{(2,3)}c_g^{(3)} & 0 & 0 \\ 0 & 0 & 0 & \alpha^{(2,3)}c_g^{(2)} & k_{11}^{(3)} - \alpha^{(2,3)}c_g^{(3)} & k_{12}^{(3)} & 0 \\ 0 & 0 & 0 & 0 & k_{21}^{(3)} & k_{22}^{(3)} + k_{11}^{(4)} & k_{12}^{(4)} \\ 0 & 0 & 0 & 0 & 0 & k_{21}^{(4)} & k_{22}^{(4)} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ Q_{\text{in}} \end{bmatrix} \tag{194.8}$$

where $k_{ij}^{(n)}(i,j=1,2)$ are elements of $\boldsymbol{K}^{(n)}$ matrix, $e_1 \sim e_7$ are nodes energy density.

194.2.3 Transform of Feedback Element

Left-hand side of expression (194.8) represent external input energy on elements nodes, when the feedback control working nodes 3 and 4 are both existing external input energy, but they are difficult to obtain. In this study, feedback element will be transformed.

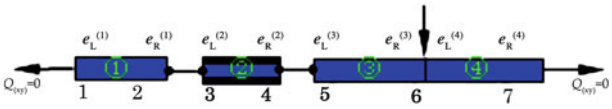


Fig. 194.2 Element finite element model of cantilever beam system

Sensor will generate electric charge $Q_a(x,t)$ when under compression follows as

$$Q_{\mathrm{a}}(x,t) = \int_{l_1}^{l_1+l_2} bd_{31}\left(E_{\mathrm{p}}\varepsilon_1\right)\mathrm{d}x = \frac{1}{2}d_{31}E_{\mathrm{p}}h_{\mathrm{b}}b\left[w_{,x}(l_1+l_2,t) - w_{,x}(l_1,t)\right] \quad (194.9)$$

where b, d_{31},ε_1 and E_p represent width,piezoelectric strain constant,strain and elasticity modulus of piezoelectric element, $w(x, t)$ is the lateral displacement of cantilever beam, bending moment $M_a(t)$ created by feedback voltage $U_a(x,t)$ given by

$$M_a(t) = K_m U_a(x,t) = \left(K_m G/C_p\right)Q_a(x,t) = K\left[w_{,x}(l_1,t) - w_{,x}(l_1+l_2,t)\right] = K\theta \quad (194.10)$$

where K_{m} is transfer coefficient, $K = 0.5d_{31}E_p h_b b K_m G/C_p$ is total feedback gain, θ is outer corner difference between each end of piezoelectric element, C_p capacitance of sensor,.

Taking into accout that feedback control bending moment $M_a(t)$ is in direct proportional with θ, θ also can be expressed as $\theta = l_2/\rho$, where ρ is curvature radius, so the internal section bending moment M could be obtained through bending moment-balance equation

$$M = M_a + EI/\rho = (Kl_2 + EI)/\rho \quad (194.11)$$

Obviously, total bending moment raised the original moment by Kl_2.

194.3 Energy Density Wave Solution of Cantilever Beam System

For comparing with EFEM solution, energy density wave solution of cantilever beam system is deduced as follow. Here cantilever beam is still divided into four elements (Fig. 194.2), $w_i(x,t)$ $(i = 1, 2, 3, 4)$ represents the lateral displacement of each beam element, $E_b I_b$,$\rho_b A_b$,$E_p I_p$,and $\rho_p A_p$ represent bending rigidity and the mass density per unit length of beam and piezoelectric element.

Differential equation of motion for beam elements 1,3 and 4 can be expressed as

$$\rho_b A_b w_{i,tt} + E_b I_b w_{i,xxxx} = 0, (i = 1, 3, 4) \quad (194.12)$$

Differential equation of motion for beam elements 2 can be expressed as

$$\left(\rho_b A_b + 2\rho_p A_p\right) w_{2,tt} + \left(E_b I_b + 2E_p I_p\right) w_{i,xxxx} - M_{a,xx} = 0 \tag{194.13}$$

Let $\lambda_1^4 = \lambda_3^4 = \lambda_4^4 = \omega^2 \rho_b A_b / E_b I_b$ and $\lambda_2^4 = \omega^2(\rho_b A_b + 2\rho_p A_p)/(E_b I_b + 2E_p I_p)$. After Fourier transformation Eqs. (194.12) and (194.13) can be rewriten as follows

$$W_{i,xxxx}(x,\omega) + \lambda_i^4 W_i(x,\omega) = 0\,(i = 1,2,3,4) \tag{194.14}$$

The solution of Eq. (194.14) can be expressed as

$$W_i(x,\omega) = C_{i1}\mathrm{e}^{-j\lambda i x} + C_{i2}\mathrm{e}^{-\lambda i x} + C_{i3}\mathrm{e}^{j\lambda i x} + C_{i4}\mathrm{e}^{\lambda i x}\,(i = 1,2,3,4) \tag{194.15}$$

Coefficient vector of Eq. (194.15) could be defined as

$$C_i^+ = [C_{i1}, C_{i2}]^T,\ C_i^- = [C_{i3}, C_{i4}]^T, \mathbf{e}_i^+(x) = \begin{bmatrix} \mathrm{e}^{-j\lambda_i x} & 0 \\ 0 & \mathrm{e}^{-\lambda_i x} \end{bmatrix}, \mathbf{e}_i^-(x) = \begin{bmatrix} \mathrm{e}^{j\lambda_i x} & 0 \\ 0 & \mathrm{e}^{\lambda_i x} \end{bmatrix},$$

$$\boldsymbol{\lambda}_{Wi}^+ = \begin{bmatrix} 1 & 1 \\ -j\lambda_i & -\lambda_i \end{bmatrix},\ \boldsymbol{\lambda}_{Wi}^- = \begin{bmatrix} 1 & 1 \\ j\lambda_i & \lambda_i \end{bmatrix},\ \boldsymbol{\lambda}_{Fi}^+ = E_i I_i \lambda_i^2 \begin{bmatrix} -1 & 1 \\ j\lambda_i & -\lambda_i \end{bmatrix},\ \boldsymbol{\lambda}_{Fi}^- = E_i I_i \lambda_i^2 \begin{bmatrix} -1 & 1 \\ -j\lambda_i & \lambda_i \end{bmatrix}$$

Through (194.14), (194.15) and boundary conditions, final equation can be obtained as

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} & \boldsymbol{A}_{24} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{A}_{31} & \boldsymbol{A}_{32} & \boldsymbol{A}_{33} & \boldsymbol{A}_{34} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{43} & \boldsymbol{A}_{44} & \boldsymbol{A}_{45} & \boldsymbol{A}_{46} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{53} & \boldsymbol{A}_{54} & \boldsymbol{A}_{55} & \boldsymbol{A}_{56} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{65} & \boldsymbol{A}_{66} & \boldsymbol{A}_{67} & \boldsymbol{A}_{68} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{75} & \boldsymbol{A}_{76} & \boldsymbol{A}_{77} & \boldsymbol{A}_{78} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_{87} & \boldsymbol{A}_{88} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1^+ \\ \boldsymbol{C}_1^- \\ \boldsymbol{C}_2^+ \\ \boldsymbol{C}_2^- \\ \boldsymbol{C}_3^+ \\ \boldsymbol{C}_3^- \\ \boldsymbol{C}_4^+ \\ \boldsymbol{C}_4^- \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{F} \\ \boldsymbol{0} \end{bmatrix} \tag{194.16}$$

where $\boldsymbol{O}$ is the second order null matrix, 0 is the two-dimension null vector,

$$\begin{aligned}
&\boldsymbol{F} = [0\,F]^{\mathrm{T}}, \boldsymbol{A}_{11} = \lambda_{W1}^+, \boldsymbol{A}_{12} = \lambda_{W1}^-, \boldsymbol{A}_{21} = \lambda_{W1}^+ \mathbf{e}_1^+(l_1), \boldsymbol{A}_{22} = \lambda_{W1}^- \mathbf{e}_1^-(l_1), \boldsymbol{A}_{23} = -\lambda_{W2}^+ \mathbf{e}_2^+(l_1),\\
&\boldsymbol{A}_{24} = -\lambda_{W2}^- \mathbf{e}_2^-(l_1), \boldsymbol{A}_{31} = \lambda_{F1}^+ \mathbf{e}_1^+(l_1), \boldsymbol{A}_{33} = K\left[\lambda_{W2}^+ \mathbf{e}_2^+(l_1) - \lambda_{W2}^+ \mathbf{e}_2^+(l_1 + l_2)\right] - \lambda_{F2}^+ \mathbf{e}_2^+(l_1),\\
&\boldsymbol{A}_{34} = K\left[\lambda_{W2}^- \mathbf{e}_2^-(l_1) - \lambda_{W2}^- \mathbf{e}_2^-(l_1 + l_2)\right] - \lambda_{F2}^- \mathbf{e}_2^-(l_1), \boldsymbol{A}_{32} = \lambda_{F1}^- \mathbf{e}_1^-(l_1), \boldsymbol{A}_{43} = \lambda_{W2}^+ \mathbf{e}_2^+(l_1 + l_2),\\
&A_{44} = \lambda_{W2}^- \mathbf{e}_2^-(l_1 + l_2), A_{45} = \lambda_{W3}^+ \mathbf{e}_3^+(l_1 + l_2), A_{53} = \boldsymbol{K}\left[\lambda_{W2}^+ \mathbf{e}_2^+(l_1) - \lambda_{W2}^+ \mathbf{e}_2^+(l_1 + l_2)\right]\\
&\quad - \lambda_{F2}^+ \mathbf{e}_2^+(l_1 + l_2),\\
&\boldsymbol{A}_{54} = \boldsymbol{K}\left[\lambda_{W2}^- \mathbf{e}_2^-(l_1) - \lambda_{W2}^- \mathbf{e}_2^-(l_1 + l_2)\right] - \lambda_{F2}^- \mathbf{e}_2^-(l_1 + l_2), \boldsymbol{A}_{55} = \lambda_{F3}^+ \mathbf{e}_3^+(l_1 + l_2), \boldsymbol{A}_{56}\\
&\quad = \lambda_{F3}^- \mathbf{e}_3^-(l_1 + l_2),\\
&\boldsymbol{A}_{46} = -\lambda_{W3}^- \mathbf{e}_3^-(l_1 + l_2), \boldsymbol{A}_{65} = \lambda_{W3}^+ \mathbf{e}_3^+(l_1 + l_2 + l_3), \boldsymbol{A}_{66} = \lambda_{W3}^- \mathbf{e}_3^-(l_1 + l_2 + l_3), \boldsymbol{A}_{67}\\
&\quad = -\lambda_{W4}^+ \mathbf{e}_4^+(l_1 + l_2 + l_3),\\
&\boldsymbol{A}_{68} = -\lambda_{W4}^- \mathbf{e}_4^-(l_1 + l_2 + l_3), \boldsymbol{A}_{75} = \lambda_{F3}^+ \mathbf{e}_3^+(l_1 + l_2 + l_3), \boldsymbol{A}_{76} = \lambda_{F3}^- \mathbf{e}_3^-(l_1 + l_2 + l_3) \boldsymbol{A}_{77}\\
&\quad = -\lambda_{F4}^+ \mathbf{e}_4^+(l_1 + l_2 + l_3),\\
&\boldsymbol{A}_{78} = -\lambda_{F4}^- \mathbf{e}_4^-(l_1 + l_2 + l_3), \boldsymbol{A}_{87} = \lambda_{F4}^+ \mathbf{e}_4^+(l_1 + l_2 + l_3 + l_4), \boldsymbol{A}_{88} = \lambda_{F4}^- \mathbf{e}_4^-(l_1 + l_2 + l_3 + l_4);
\end{aligned}$$

Table 194.1 Parameter of the collinear jointed beam

	Length	Section area$A$	Momentofinertia I	Elasticity modulus E	Density ρ
Beam	$L_b = 0.5$	$A_b = 4e{-}5$	$I_b = 1.33e{-}11$	$E_b = 2.06e11$	$\rho_b = 7.8e3$
PZ(Case1)	$L_a = 0.1$	$A_a = 2e{-}5$	$I_a = 1.67e{-}12$	$E_a = 0.7e11$	$\rho_a = 7.75e3$
PZ(Case2)	$L_a = 0.05$	$A_a = 1e{-}5$	$I_a = 2.08e{-}13$	$E_a = 0.7e11$	$\rho_a = 7.75e3$

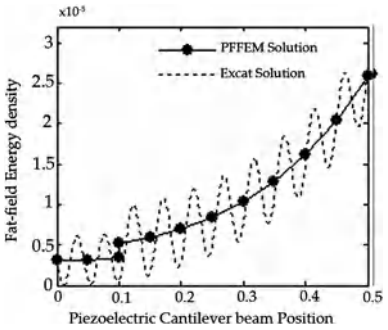


Fig. 194.3 Comparison of exact wave solution with EFEM solution

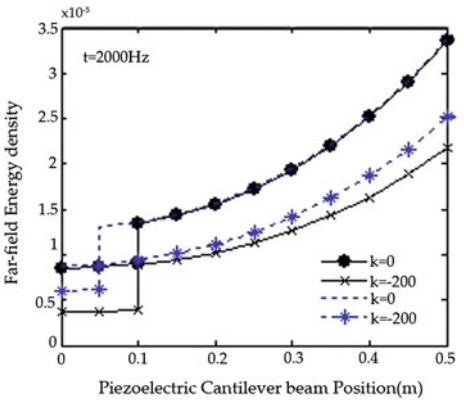


Fig. 194.4 Comparison of different length of piezoceramics by EFEM

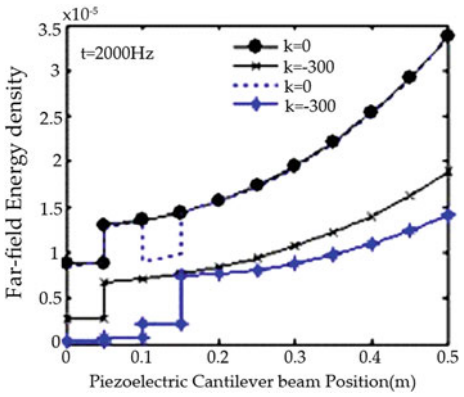


Fig. 194.5 Comparison of different pairs by EFEM

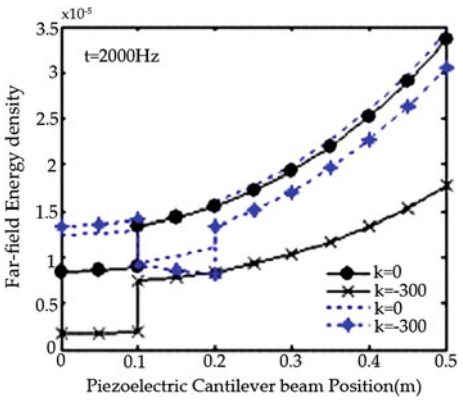


Fig. 194.6 Comparison of different position by EFEM

Wave solution could be obtained by substituting the solution of expression (194.16) in Eq. 194.15. Wave energy density is given by

$$\begin{aligned}\langle e_i \rangle = \frac{\rho_i A_i \omega^2}{2} \Big\{ & |C_{i1}|^2 e^{2k_2 x} + |C_{i2}|^2 e^{-2k_1 x} + |C_{i3}|^2 e^{-2k_2 x} + |C_{i4}|^2 e^{2k_1 x} \\ & + 2\left[\mathrm{Re}(C_{i1} C_{i3}^*) \cos 2k_1 x + \mathrm{Im}(C_{i1} C_{i3}^*) \sin 2k_1 x\right] \\ & + 2\left[\mathrm{Re}(C_{i2} C_{i4}^*) \cos 2k_2 x + \mathrm{Im}(C_{i2} C_{i4}^*) \sin 2k_2 x\right]\end{aligned} \tag{194.17}$$

where k_1 and k_2 represent the real and imaginary part of complex wavenumber, C_{i1}, C_{i3} is the far-field energy densty, C_{i2}, C_{i4} is the near-field energy densty.

194.4 Numerical Analysis

Physical parameters of the piezoelectric cantilever system, results for two cases (Table 194.1)

Figure 194.3 compares the exact wave solution with EFEM solution for the cantilever beam, the result represents well the energy finite model of the system established above is correct. The distribution of energy density along the beam under different length of piezoelectric elements in Fig. 194.4, it can be seen that longer one have better control effect than the shorter. Figure 194.5 shows the comparison of different pairs of piezoceramics by EFEM, obviously, two pairs have better control effect. Figure 194.6 shows the comparison of different stick position, it is found that when the stick position towards end has better control effect.

194.5 Conclusions

In this paper, research is focused on the modeling and analyzing of piezoelectric cantilever beam active vibration control system based on Energy Finite Element Method, proved by wave solution

1. Modern analysis of complex structures relies primarily on FEM, unfortunately, the traditional FEM may become impractical under mid- and high- frequency. EFEM do not exist this problem, but also much more convenient for complex structures than wave theory. The final results obtained by EFEM is time-averages energy, which is more useful than traditional variables like displacement, speed and acceleration in optimal design of active control system.
2. The length and position of the piezoelectric elements have important influences in piezoelectric cantilever beam vibration control system. Appropriately increase the length of piezoelectric elements may obtain better control effect. The more pairs of piezoelectric elements we choose, the better control effect.

References

1. Denglin Z, Rui L, Jie Y (2009) Integrated optimal design of the PZT position size and control of smart cantilever beam[J]. J Mech Eng 45(2):262–267
2. Zhang N, Kirpitchenko I (2002) Modelling dynamics of continuous structure with a piezoelectric sensor/actuator for passive structural control[J]. J Vibr 2439(2):251–261
3. Peng L (2006) Theory an, d experiment research on active vibration control of piezoelectric intelligent boring bar[D]. Jilin university
4. Sze KY, Yao LQ (2000) Modelling smart structures with segmented piezoelectric sensors and actuators[J]. J Vibr 235(3):495–520
5. Xuewen Y, Hongfei C, Xiaojun G (2007) Relevancy among power flow theory statistical energy analysis and energy finite element method[J]. J Ship Mech 11(4):637–646
6. Nefske DJ, Sung SH (1989) Power flow finite element analysis of dynamic system[J]. J Vibr Stress Reliab Des 111(1):94–100

Chapter 195
Pipe Network Optimization for Maximal Utilization Rate of Gravity Head Based on LCA

Yujuan Fu, Yulong Zhang, Huanjie Cai, Dong Chen and Junshi He

Abstract The general optimization design models were aimed at the minimum cost of the gravity distribution network omitted other fees impacts on the system building operation cost and could not ensure the design scheme satisfies the running effect. The objective function in this chapter is based on the gravity water head utilization rate produced by the fall of the head position degree, with two different expressions namely utilization rate of pipe gravity head and utilization rate of path gravity head as network optimization objective functions. Model solution method is with line-up competition algorithm (LCA) between families and inside a family. Compared with the solution method of genetic algorithm aimed at the minimum investment cost, the result showed that the method can obtain the optimal solution and have higher search efficiency.

Keywords Pipe network optimization · Utilization ratio of gravity head · Line-up competition algorithm · Objective function

Y. Fu (✉) · Y. Zhang
Key Laboratory of Agricultural Resource and Environment of Liaoning, Shenyang Agricultural University, Shenyang 110866, China
e-mail: fyj0249@sina.com

H. Cai
The Key Laboratory of Agricultural Soil and Water Engineering, Northwest A&F University, Yangling 712100, Shaanxi, China

Y. Fu · D. Chen
Shenyang Design and Research Institute of Chinese Coal International Engineering Group, Shenyang 110015, China

J. He
College of Water Resources, Shenyang Agriculture University, Shenyang 110866, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_195,

195.1 Introduction

In the optimization design of pipe network system, the aim is to minimize the cost of network investment or the discounted annual fee of operation and management of project investment [1–3]. While, in the process of system construction and operation management, the pipe network investment also includes earthwork excavation fee, machine-shift cost, labor costs, building materials fee and so on, and some cost items can not be accurately calculated in the design stage. That is to say, it is not feasible to accurately calculate the operating system investment costs in the planning and design stage. The general pipe network optimization design is aimed at minimizing the fees of pipe for the system. Although this method simplifies the calculation, it ignores other cost items and their impacts on the optimization results. There are some shortcomings of the optimization model aiming at minimizing the system investment costs, so is not the best choice in all cases. Gosselin L. and Bejan A. [4] presented that it is feasible to build the optimization model aiming at the minimum required pumping power. Currently, multi-objective optimization problem can be solved through converting it into a single objective problem by a variety of strategies. To take full advantage of natural terrain, this paper is objected to make the maximum utilization ratio, and use the evolutionary algorithm of Line-up competition algorithm (LCA), which has less control parameters and fast convergence [5, 6]. Compared the optimization results with the integer programming optimization model of minimizing the cost of genetic algorithm, the results showed that the method is feasible, and offers a new way for the planning programming of gravity water distribution network.

195.2 Model Construction

195.2.1 Objective Function

In this paper, the optimization model is aimed at maximizing the head position's utilization of the whole pipe network. The target function has two expression forms, namely pipe head utilization rate *Fun*1and path head utilization rate *Fun*2:

$$\begin{aligned} \max\ Fun1 &= \frac{\sum_{i=1}^{n} h_{fi}}{\sum_{i=1}^{n} \Delta E_i} \times (1+a) \times 100\% \\ \max\ Fun2 &= \frac{\sum_{i=1}^{n} \tilde{h}_{fi}}{\sum_{i=1}^{n} \Delta \tilde{E}_i} \times (1+a) \times 100\% \end{aligned} \tag{195.1}$$

where, h_{fi}—the head loss along the riser pipe, m; ΔE_i—the elevation difference of the pipeline section, m; $\tilde{h}_{fi}$—the head loss of water along the pipe from the water source to the node i, m;$\Delta \tilde{E}_i$—the pressure head of available water from the water

source to the node i, m; α—the network local head loss coefficient; n-the number of nodes in the network (except the water source).

The graph theory of the minimum spanning tree theory shows that the number of nodes of the tree pipe network is the same as the quantity of the pipe (excluding the water source point). For the purpose of calculation, the nodal number of the ending pipe is taken as the pipe's number.

The head loss h_f along the pipe can be presented as the Brahms Hughes formula:

$$h_f = f\frac{L}{D^b}Q^m \tag{195.2}$$

where: f—the coefficient related with frictional resistance; m—index of discharge; Q—pipeline design flow $m^3{\cdot}s^{-1}$ or $m^3{\cdot}h^{-1}$; L-length of pipe, m; b-index of pipe diameter. The gravity head $\Delta\tilde{E}_i$ from the source to the node i can be shown as:

$$\Delta\tilde{E}_i = E_0 - E_i - E_{gi} \tag{195.3}$$

where: E_0-water elevation, m;E_i-elevation of node i, m; E_{gi}-minimum working head node i, m.

Equations (195.2) and (195.3) were substituted into the Eq. (195.1), two objective functions can be expressed respectively as:

$$\begin{aligned} \max\ Fun1 &= \frac{\sum_{i=1}^{n} f\frac{L_i}{D_{ik}^b}Q_i^m}{\sum_{i=1}^{n} E_i - E_{i-1}} \times (1+\alpha) \times 100\% \\ \max\ Fun2 &= \frac{\sum_{i=1}^{n}\sum_{j=1}^{a} f\frac{L_{ij}}{D_{ijk}}Q_{ij}^m}{\sum_{i=1}^{n} E_0 - E_i - E_{gi}} \times (1+\alpha) \times 100\% \end{aligned} \tag{195.4}$$

where: E_{i-1}—the elevation of the pipe first node i, m; a—the number of pipes which the water flows from the source to the node I; L_i-the length of pipe i, m; k—the diameter of pipe k, for the convenience of calculation, the selection of pipe's diameter is numbered from small to large, 1,2…; D_{ik}—the diameter of pipe i which the material is selected from number k, mm; Q_i—the designed flow of pipe i, $m^3{\cdot}s^{-1}$ or $m^3{\cdot}h^{-1}$; L_{ij}—the j—the pipe's length where the water flows from source to node i, m; D_{ijk}—the diameter of pipe k where the water flows from source to node i through jth, mm; Q_{ij}—the designed flow where the water flows from source to node i through jth, $m^3{\cdot}s^{-1}$ or $m^3{\cdot}h^{-1}$.

195.2.2 Constraint Condition

(1) To ensure the supply pressure of each node, the pressure elevation needs to be higher than the pressure level of service nodes, namely:

$$E_0 - E_i - \tilde{h}_{fi}(1+\alpha) \geq E_{gi}\ i = 1, 2\ldots n \tag{195.5}$$

Equation (195.5) can be transformed to

$$\tilde{h}_{fi} \leq \frac{E_0 - E_i - E_{gi}}{1 + a} \tag{195.6}$$

For each node, the right side of Eq. (195.6) is a constant value, and it can be expressed as $\Delta H = \{\Delta H_1, \Delta H_2 \ldots \Delta H_n\}$, where: $\Delta H_i = \frac{E_0 - E_i - E_{gi}}{1+a}$, for the arranged pipe network, it can be calculated directly.

(2) Diameter constraints:

$$D_i > 0 \tag{195.7}$$

(3) Flow constraints: To prevent water hammer damage, pipe flow should be limited below 3 $m \cdot s^{-1}$; to prevent siltation, pipe flow should not be

$$0.6 \leq V_{ik} \leq 3 \tag{195.8}$$

$$V_i(k) = \frac{4Q_i}{\pi D_{ik}^2} \tag{195.9}$$

So:

$$0.6 \leq \frac{4Q_i}{\pi D_{ik}^2} \leq 3 \tag{195.10}$$

In order to improve the efficiency of computer optimization, the diameter of the alternative pipe can be calculated according to the flow constraints before solving the problem. This method not only reduces one constraint but also lessen the search range of the solution. The upper and lower limit of each pipe's available diameter $D_{i\max}$ and $D_{i\min}$, can be calculated according to Eq. (195.12), selecting the available diameter level in this range, and recording the corresponding pipe number.

$$D \in [D_{i\min}, D_{i\max}] \tag{195.11}$$

$$D_{i\max} = \sqrt{\frac{4Q_i}{0.6\pi}} D_{i\min} = \sqrt{\frac{4Q_i}{3\pi}} \tag{195.12}$$

In summary, the optimization design model, which chooses the selected pipe diameter as the variable and the goal is the highest utilization of the pipe gravity head, can be presented as:

$$\begin{aligned} \max \quad & Fun1 = \frac{\sum_{i=1}^{n} f \frac{L_i}{D_{ik}^b} Q_i^m}{\sum_{i=1}^{n} E_i - E_{i-1}} \times (1 + \alpha) \times 100\% \\ \text{s.t.} \quad & \sum_{i=1}^{k} \underset{fi}{\sim} h (1 + \alpha) \leq E_0 - E_j - E_{gj} \end{aligned} \tag{195.13}$$

Since the model's constraint is an in equation, the penalty function can be used to solve it. The fitness function which is solved by LCA is:

$$max\, val(Fun1) = Fun - ch$$

$$ch = \begin{cases} 0, & \text{If the solution meet the constraints} \\ +\infty, & \text{otherwise} \end{cases} \tag{195.14}$$

The optimization model is aimed at the gravity head's utilization, and its objective function already contains the pressure constraints. Therefore, it can be transformed into unconstrained problem through some algorithm designs. The optimization model is:

$$\max\ val(Fun2) = Fun2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{a} f\frac{L_{ij}}{D_{ijk}^{b}} Q_{ij}^{m}}{\sum_{i=1}^{n}\Delta H_i} \times 100\% \tag{195.15}$$

195.3 Problem Solution

195.3.1 Establishment of Network Vector and Matrix

First, establish the pipe network vectors and matrices based on the graph theory: Node vector: $N = \{n_1, n_2 \ldots n\}$; Node elevation: $E = \{E_1, E_2 \ldots E_n\}$; the elevation difference of both ends of the pipe: $\Delta E = \{\Delta e_1, \Delta e_2 \ldots \Delta e_n\}$; the required working head of each node $H_g = \{H_{g1}, H_{g2} \ldots H_{gn}\}$; vector flux: $Q = \{q_1, q_2 \ldots q_{\mathrm{n}}\}$; pipe network connection matrix: $B = b_{ij(\mathrm{n}+1)\times(\mathrm{n}+1)}$.

where

$$b(i,j) = \begin{cases} = 0, \text{ no pipe connection from node } i \text{ to } j \\ = 1, \text{ connection from } i \text{ to } j, \text{the direction is from } i \text{ to } j \end{cases} \tag{195.16}$$

195.3.2 Model Solution with Line-Up Competition Algorithm

Line-up competition algorithm is selected to solve the problem in this chapter. It is one evolutionary algorithm proposed by Chinese scholar Liexiang Yan. LCA, as one group search process, has similar basic mechanism to GA, which has the operators such as breed, variation, competition and selection. The main difference is that LCA in the evolutionary process always stay independent and has parallel evolution family, each family has one individual and produce posterity through the asexual reproduction. In addition, the competition mechanism of LCA is different from that of GA. There are two competition levels, one is vertical competition which means the progeny of the competition for survival is

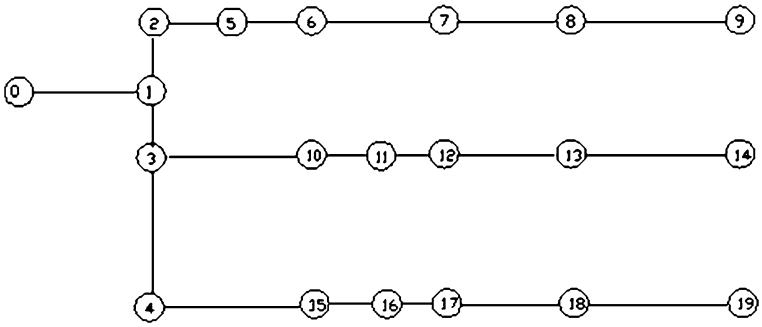


Fig. 195.1 Layout of pipe network of gravity pressure conduit

within the same family and only one of the most outstanding individual can survive and represent the family; another one is horizontal competition, which refers to the position competition between different families and lined up according to the size of the family, the best family will be in the first row and the worst in the last one [5]. The fundamental idea of LCA is to renew or replace the first family in the line through the above competition levels so as to approach the most advantage quickly [6].

The utilization of pipe head as the objective function, the steps of using LCA to calculate the optimization design are as follows:

(1) Calculating water flow path from the water source to each node according to the connection matrix of pipe network, and producing the initial father generation individuals by adjacency matrix of a group of graphs;
(2) Computing the head loss and fitness function of each individual of father generation, ascending the value of fitness function and competing between families;
(3) Producing the offspring individuals by mutating father generation, the mutation times are decided by the position in the ascendant of the father generation;
(4) Computing the fitness value of offspring individuals, and comparing with that of its father generation. The larger value one will be chosen for father generation and participating next computation;
(5) Repeating step 2 ~ 4 until the prescribed generation times are completed.

When solving the optimal design model aiming at the path head utilization rate with LCA, it is necessary to validate the feasibility of each solution, namely, validate whether Eq. (195.17) is true or false. If it is false, then reproduce a new solution, until the number of solution can satisfy the population size. The other steps can repeat the same as former.

$$\frac{\sum_{j=1}^{a} f \frac{L_{ij}}{D_{ijk}^{b}} Q_{ij}^{m}}{\Delta H_i} \leq 1 \tag{195.17}$$

195.4 Case Validation and Comparison

The layout of water distribution gravity irrigation is shown in Fig. 195.1 [7]:

When optimizing with LCA, the models with the population size is 50, evolution times are 300, the objective functions of section gravity head rate and path gravity head rate obtained the optimal solution. As the latter one has no constraint function, less calculation steps, so the speed of optimal solution is faster.

The first objective function expression is $Fun1 = 98.4\%$, the second is $Fun2 = 76.2\%$, the corresponding pipe network investment is 235947.8 RMB.

Meanwhile, for the objective function aimed at the minimum investment cost, the optimal solution is based on the integer encoding of genetic algorithm. The optimal solution can be obtained when the population size is 50 and the evolution times are 500. Therefore, the method is feasible for the gravity water distribution network optimization, and has faster convergence rate and less controlling parameters.

195.5 Conclusions

The previous optimization models of pipe network were aimed at the minimum cost, omitted other impacting factors and the operation reliability. For the gravity distribution system and good use of the gravity water head produced by the natural terrain elevation, this chapter provided gravity water head utilization rate as the system optimization evaluation index, and the objective function aimed at the section and path gravity head. It proposed a new idea for the optimal design of pipeline network. This chapter presented the model solution with LCA method, and projected the corresponding encoding and mutation operators. Compared with genetic algorithm aimed by the case study, the results showed that LCA can obtain the optimum diameter combination of pipeline network and the fast convergence speed. Meanwhile, since the objective function model of the path gravity head utilization rate does not have the constraints limitation and fast calculation speed.

Acknowledgments This research was supported by postdoctoral foundation and youth foundation by Shenyang Agriculture University. The author thanks prof. Yulong Zhang and Prof. Huanjie Cai.

References

1. Lavric V, Iancu P, Plesu V (2006) Cost-based design of wastewater network optimal topology. Res, Res Conserv Recycl 44:53–63
2. Montesinos P, Garcia-Guzman A, Ayuso JL (1999) Water distribution network optimization using a modified genetic algorithm. Water Resour Res 35:3467–3473
3. Keedwell E, Khu ST (2005) A hybrid genetic algorithm for the design of water distribution networks. Eng Appl Artif Intel 18:461–472

4. Gosselin L, Bejan A (2005) Tree networks for minimal pumping power. Int J Therm Sci 44:53–56
5. Yan LX, Ma DX (2001) Global optimization of no convex nonlinear programs using line-up competition algorithm. Comput Chem Eng 25:1601–1605
6. Yan LX (2003) Solving combinatorial optimization problems with line-up competition algorithm. Comput chem Eng 27:251–258
7. Bai D (1995) The optimum design for gravity sprinkle pipe network. Trans Chin Soc Agric Mach 26:43–46

Chapter 196
A New Thinking: Personalized Recommendations on Spatial Information

Mou Naixia, Zhang Lingxian, Li Qing and Liu Wenbao

Abstract The methods for information overload have not been well developed so far, so recommended system was proposed and has been widely used in e-commerce, digital libraries and other areas. With the coming of the new era of geographic information, the situation of spatial information overload is more serious. In this paper, we proposed a new thinking named "personalized recommendation on spatial information", and gave its theoretical framework Three aspects of recommendations on spatial information: user profile, recommended algorithm and quality evaluation of spatial information service were employed to explain the theory. They highlighted location, spatial relations, and relations of time, which were different with other recommended system. The issue of spatial information personalized recommendations was proposed in this paper at the first time, and would be paid more attention by more and more researchers.

M. Naixia (✉) · Z. Lingxian · L. Wenbao
Key Laboratory of Geomatics and Digital Technology of Shandong Province,
Qingdao 266510, China
e-mail: mounaixia@163.com

Z. Lingxian
e-mail: zhang_lingxian@163.com

L. Wenbao
e-mail: liu_wenbao@yahoo.com

M. Naixia · Z. Lingxian · L. Wenbao
Geomatics College, Shandong University of Science and Technology,
Qingdao 266510, China

L. Qing
Geoinformation Science, Faculty of Social Science,
The Chinese University of Hong Kong, Hong Kong,
People's Republic of China
e-mail: kkxxlq@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_196,

Keywords Personalized recommendations · Spatial information · User profile · Recommendation algorithm · Evaluation of service quality

196.1 Introduction

In information age, it is not the lack of information, but how to obtain target information from the disordered heterogeneous information systems. The obvious result of the explosion of information is out-of-information, because users can not access information of their interests in effectively, then information overload problem emerges. So "recommender systems" was put forward in 1997, and widely used in digital libraries, e-commerce and other fields. The main purpose of recommended system is to solve the problem of information overload [1].

With the third wave of internet, 3G emergence, Web 2.0 and the corresponding technical systems (Grid, Ajax, CSS+XHTML) bring new technologies. Users need to get information by interactive, creative, relative, various modes, which is difficult to realize by regular electronic map-based geographic information services [2]. The supply of spatial information was transferred from 4A to 4R, that is, from Anytime, Anybody, Anywhere, and Anything to Right time, Right body, Right place, and Right thing. It meant that not only at any time, in any place, we can get any events but at right time, in right place, we get right information. The transformation from 4A to 4R made the provision of spatial information from "most cost, most benefit" into "least cost, most benefit". New geographical information age is coming. It is different from the traditional geographical information age which marked a new era of person-centered geographic information.

Now more researches concentrated on spatial information services, emphasized how to efficiently merge, integrate and assimilate multi-source information. Compared to other areas of personalized service, spatial information services more emphasis on the participation of people [3]. Not only concerned with user, but with the current location, destination location, time, spatial relations and others factors. Personalized service is the core of recommended systems, namely, humanization and intelligentize of spatial information services are the features of new geographic information age.

196.2 Spatial Information Personalized Recommendations

"Data" is always one core research area of geographic information system. Many of hot topics of GIS in different periods all center on data from centralized data processing through distributed data management to decentralized data processing, such as data acquisition, data management, data analysis, and so on. Then the research focus of GIS shifted to the application of data, such as researching on the

standards and codes of the data sharing in the period from data sharing to data services. During this time, we take more research on standard specification of data sharing. As one part of information technology, the development of GIS is closely and inseparably related to the overall trend of information technology. With the development of information service and gridding technique, the corresponding technical philosophy is also reflected in the development of GIS. For example, GIS services, spatial data mining, grid GIS has become the research hotspot. Generally speaking, however, the GIS development of this period is data-centered as a whole.

The decentralized and heterogeneous phenomenon of spatial data can't be eliminated and further exists. Meanwhile, the amounts of spatial data are growing and requirement diversifies. On the premise of the above condition, providing information services is the better choice instead of data replication. The spatial information service is the process of determining the different user demand of spatial information into different particle-sized services or service chains. A direct consequence of the massive growth of spatial information is "information overload". How to extract the information from huge, decentralized, polygene spatial information source is the problem that needs to be solved urgently.

196.3 Theoretical Framework of Personalized Recommendation on Spatial Information

Like other personalized recommendation systems of spatial information service also involves several aspects: user information acquisition, analyzing and forecasting user behavior, recommendation algorithm of service and evaluation of service quality [4] (Fig. 196.1).

The basic process of personalized recommendations is as follows:

196.3.1 User Profile

User profile plays a central role in the personalized recommendation system, which is the basic component of personalize the recommendation. Croft, Pazzani and others who summarized their experimental results, mentioned the primary factor which affects the efficiency of information processing is information and its related knowledge significance and secondly the algorithm complexity [5]. As the foundation and core of personalized service, the quality of the user profile is directly related to the quality of personalized service. As a result, constructing user model from user information, named user modeling, becomes core and key technology of personalized service.

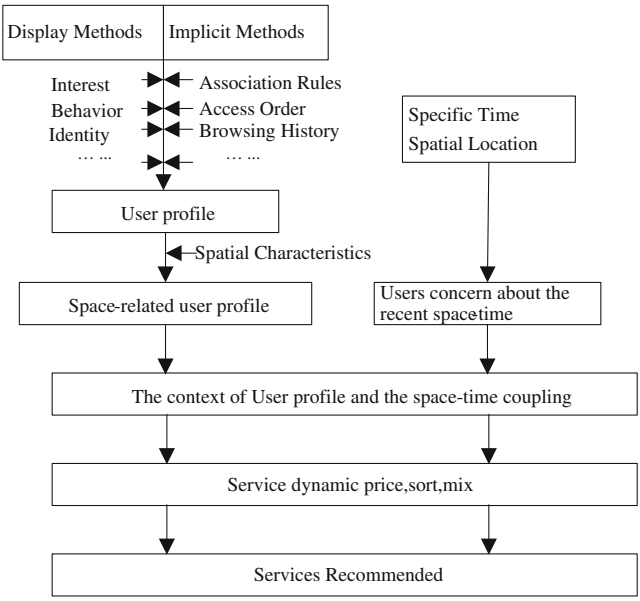


Fig. 196.1 Technical line of personalized recommendations model on spatial information

User's interest model is reflected in: location of the user to select different services and personal interest in ordering has a regulatory role. Traditionally, less interest in the case, users tend to choose easier to reach the place of service. Visible spatial information consumption showed a typical dynamic characteristic of "constructing while running".

196.3.2 Recommendation Algorithm

Recommendation is the object selected objects that users need from multiple candidates, which is information filtering method of high degree of automation. A complete recommendation system includes three parts: the behavior recording module to collect user information, the model analysis module to analyze user preferences and recommendation algorithm module [6], Table (196.1).

196.3.2.1 Collaborative Filtering System

Collaborative filtering system is a widely used recommendation system which was proposed in the first generation, the main idea of which can be divided into two parts: firstly, calculating the similarity among users according to users' historic information; then, predicting the preference degree to a particular product of target

Table 196.1 Comparison with main recommendation technologies and algorithms

Recommended technology	Background conditions	Input	Steps
Collaborative filtering recommendation	Evaluation of U to I	Evaluation Grade of U to I	Identify the user u's neighbors Generated evaluation mark of I according to it
Content-based recommendation	Characteristics attribute of I	Evaluation Grade of U to I	Generated sub-project classification according to the evaluation of U
Recommended based on user statistics	User statistics of U and the evaluation to I	Statistics on user information of U	Identify similar users U Generated evaluation mark of U according to it
Recommend based on utility	Characteristics of I	U of I utility function describing the preferences U to I	Using utility function for each item Generate the purpose sort
Knowledge-based recommendation	Characteristics of I, how it can meet the user's knowledge	Description of needs and interests of U	Compute the degree of matching of project i and user needs
Rule-based recommendation	Browsing or purchasing history of U to I	Browsing buy records	Generate association rules Generate recommendation based on it

U the set of users, I the set of all items, u the user of current to predict, i the project of current to predict

users using evaluation on other products from the neighbor. Then the system recommends special products to target users according to the above preference degree. The biggest advantage of collaborative filtering recommendation system is that it can process objects like music or movie which are hard to text structuring express without special requirement.

Collaborative filtering system is the most used recommendation system, in which Grundy is believed to be the first collaborative filtering system to be put into application. The Grundy system can establish a user interest model and use it recommend related books to every user. While the Tapes-try mail processing system's artificially determining similarities among users will greatly increase the workload and greatly discounted accuracy as the user amount increase. And GroupLens establish user information group making users in each group publish their own information, calculate the similarity among users based on the social information filtering system, in order to collaborative recommend to users in each group. While Ringo recommend music to users by using the same social information filtering method. And there are some other system recommending with

the collaborative filtering method, such as book recommendation system of Amazon.com, jokes recommendation system of Jester, WWW information recommendation system of Phoaks.

196.3.2.2 Content-Based Recommendation

The initial content-based recommendation is the continuation and development of collaborative filtering technology, which does not need the users' suggestion to item, but make the related recommendation according to the selected products' content information from users. As technology like machine learning is being perfected, the current content-based recommendation system can establish configuration file for users and products through analyzing the purchased (or browsed) content, then create or update the user's profile. The system can compare the user and product configuration documents' similarity and recommended the products to users which are the most similar to the configuration profile. For example, in the movie recommendation, content-based system firstly analyzes these films' common that the user has watched and highly graded. Then it recommended other films, which are most similar to users' interested film. Content-based recommendation algorithm is information acquisition and information filtering on earth. Research on text information acquisition and filtering is permitting, so according which many content-based recommendation systems are recommended now.

196.3.2.3 Personalization Recommendation of Spatial Information

User characteristics control service recommendation through making influence on spatial relationship. This can be expressed as:

$$S_\Sigma = \bigcup (\vec{d} \times U\{u_1, u_2, \mathbf{u_3}, \ldots u_n\}, \vec{l} \times U\{u_1, \mathbf{u_2}, u_3, = \ldots u_n\},$$

$$O_\Sigma(U\{u_1, u_2, u_3, \ldots u_n\}) \times g(lay_1 =, lay_2, \ldots))$$

S_Σ	recommended service collection,
$\vec{d}$	direction factor of spatial relations,
$U\{u_1, \mathbf{u_2}, u_3, \ldots u_n\}$	user profile, among in it
u_1	a characteristic of a user profile,
$\mathbf{u_2}$ and $\mathbf{u_3}$	(bold) relative to $\vec{l}$ and $\vec{d}$ the most influential factor (It sometimes has a veto power, such as ethnic restaurants of choice for property),
$\vec{l}$	measure component factor,
O_Σ	geographic objects selected by topology,
$g(\text{lay}_1, \text{lay}_2, \ldots)$	function in the topology of geographic objects.

196.3.3 Service Quality Evaluation of Spatial Information Recommendation

The majority of recommendation systems evaluates whether a recommendation algorithm is good or not by its accuracy. Due to the lack of evaluation standardization, it is difficult to compare different systems of recommendation algorithm. In view of the different systems, the existing accuracy standard includes: predictive accuracy, the accuracy of classification, sorting accuracy, prediction scoring association, distance standardization index and the half-life of utility index.

Service quality evaluation is to test the extent of meets the needs of users and spatial information service is to provide users with more accurate, intelligent and personalized service. The evaluation result often depends on the user's feeling. In fact, the users' satisfaction also reflects the hard parameters of service to some extent. Some existing research indicated that the user feel anxious if a response time exceeded 30 s. But this is not crucial to the problem and research in this area is relatively more after all.

196.4 Conclusion

This paper provides theoretical supports and experimental evidence for personalized services of spatial information in order to advocate the nature of information services, realizing the transformation from "Looking for service" into "Service finding me", which can improve the relevance and intelligence of spatial information services.

Acknowledgments Supported by Special Project Fund of Taishan Scholars of Shandong Province; Shandong Province Natural Science Foundation (ZR2010DM015); Open Research Fund Program of Key Laboratory of Marine Spill Oil Identification and Damage Assessment Technology (201015, 201117); Open Research Fund Program of the Key Laboratory of Surveying and Mapping Technology on Island and Reef, State Bureau of Surveying and Mapping, China (2010B19); China Postdoctoral Science Foundation (20090461255); the National Key Technologies R&D Program (2006BAJ15B02).

References

1. Lihua W, Lu L (2002) User profiling for personalized recommending systems. A Rev 1:55–62
2. Deren L, Zhenfeng S (2009) The new era of geographic information. Sci China Inf Sci 6:579–587
3. Hyunsuk H, Seonghyun S, Kim C (2006) A user-oriented gis search service using ontology in location-based services. In: ICHIT 2006. LNAI 4413 pp 209–218
4. Chun Z, Chunxiao X, Lizhu Z (2002) A survey of personalization technology 10:1952–1960

5. Middleton ES, Shadbolt RN, Roure CD (2004) Ontological user profiling in recommender systems. ACM Trans Inf Syst 1:54–88
6. Jianguo L, Tao Z, Binghong W (2009) The research progress of personalized recommendation system. Prog Nat Sci 1:1–14

Part XIII
Power Engineering and Transmission

Chapter 197
Power System State Estimation Based on GM (1, 1) Model

Yan Donsong, Xiao Tengjiao, Long Yarui and Huang Yuanliang

Abstract According to the contradiction between the static power system state estimation and the actual operation of power system, based on the study of the classical algorithms of power system state estimation, apply the gray system theory of power system based on GM (1, 1) mode to the state estimation of power system, and create a new dynamic power system state estimation algorithm, for example to a four-node power system network test for State estimation simulation. Measurement of a known algorithm used to predict the gray system theory. On the other hand, measurement of unknown quantity, using the result that come from the known quantity, then get all parameters with character of power system state through statistical analysis of simulation results that the normal measurement system, the method is feasible and has a good filtering effect. The power system state estimation algorithm based on the gray system theory of Power System based on GM (1, 1) model has high estimating precision, fast calculation, use small memory and the adaptability of the measurement system is very strong, so it has good prospects.

Keywords Power system · State estimation · GM (1, 1) model · Dynamic estimation

Y. Donsong (✉) · X. Tengjiao · L. Yarui · H. Yuanliang
Electric Automatization Institute of Jinan University,
Zhuhai, Guangdong, China
e-mail: zhjszx@jnu.edu.cn

X. Tengjiao
e-mail: 171180475@qq.com

L. Yarui
e-mail: 491787531@qq.com

H. Yuanliang
e-mail: yoll@jnu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_197,

197.1 Introduction

Because the actual operation of power system is changing, in fact there is no static problem. Only when the time interval is short enough as static can estimate the approximate. It can approximately be regarded as static estimation. If it is possible to trace the changing process of electric power system within a period of time and use a period of time measurement to predict the future of the state amount, the effect is better than the static estimation. This article applies the grey system theory to power system State Model, established the state estimation mathematical model, which is a kind of dynamic estimate, and excels in fast calculation speed, less memory, high precision, using MATLAB program to realize algorithm. However, there are not any records in documents that apply the gray system theory to power system state estimation.

The grey system theory researches on "partially known information and partially unknown information", which is known as "poor information" and is a system of uncertainty. It is based on the generation, development of partly known information to attain the exact description and acquaintance of reality [1]. Power system state estimation is known conditions of real-time power system measurement; node is usually injected into the complex power, branch complex power and voltage amplitude into node. The unknown quantity of state is each node voltage amplitude and phase angle. Power system state estimation belongs to the grey system theory research objects, so we can use the grey system theory to make power system state estimation.

197.2 GM (1,1) Model

The modeling method [2] of Grey model GM (1, 1) is: first, accumulate generation and process the original data sequences, weaken the randomness, enhance its regularity. Then apply single variable first-order differential equations to fit the sequence to get grey predicted model, then get the discrete solutions, eventually predicting the future trend of development.

Set the original sequence

$$x^{(0)} = \left(x^{(0)}(1)\, x^{(0)}(2)\ldots x^{(0)}(n)\right) \tag{197.1}$$

We get AGO sequence accumulate the original data sequences to

$$x^{(1)} = \left(x^{(1)}(1)\, x^{(1)}(2)\ldots x^{(1)}(n)\right) \tag{197.2}$$

Among them:

$$x^{(1)}(1) = x^{(0)}(1)$$
$$x^{(1)}(k) = x^{(0)}(k) + x^{(1)}(k-1),\ k = 2, 3, \ldots, n.$$

Based on the AGO sequence, we get a GM (1, 1) model:

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + ax^{(1)} = u. \tag{197.3}$$

where a, u are unidentified parameter. Set parameter vector $\mathrm{A} = [\mathrm{a}\ \mathrm{u}]^{\mathrm{T}}$, the solution of the gray differential equation (197.3) solution is:

$$x^{(1)}(t) = \left(x^{(0)}(1) - \frac{u}{a}\right)e^{-at} + \frac{u}{a}. \tag{197.4}$$

Equation (197.4) is continuous-time response function. The numerical solution of differential equations is available through the discretization of t,

$$x^{(1)}(k) = \left(x^{(0)}(1) - \frac{u}{a}\right)e^{-a(k-1)} + \frac{u}{a}. \tag{197.5}$$

Using the least squares regression, we get A's approximate solutions:

$$\widehat{A} = \begin{pmatrix} \widehat{a} \\ \widehat{u} \end{pmatrix} = \left(B^T B\right)^{-1} B^T y_n \tag{197.6}$$

where

$$y_n = \left[x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right]^T$$

$$B = \begin{bmatrix} -[x^{(1)}(2) + x^{(1)}(1)]/2 & 1 \\ -[x^{(1)}(3) + x^{(1)}(2)]/2 & 1 \\ \vdots & \vdots \\ -[x^{(1)}(n) + x^{(1)}(n-1)]/2 & 1 \end{bmatrix}.$$

Put estimated value $\widehat{a}$ and $\widehat{u}$ into the type (197.5), then get GM (1,1) predicted model function:

$$\widehat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a}\right)e^{-ak} + \frac{u}{a}. \tag{197.7}$$

By restitute and reduce, we get a fitting function of original sequences:

$$\widehat{x}^{(0)}(k+1) = x^{(1)}(k+1) - \widehat{x}^{(1)}(k). \tag{197.8}$$

197.3 Power System Based on the GM (1, 1) Model

In the power system state estimation, the part of measurement is known, other is unknown, and there exist different measurement errors in the known measurement,

so a power network system can be considered as a complex gray system, and we can apply the concept and method of the gray system theory to static power system state estimation, taking advantage of the known measurement to judge the running state of the power system and make forecast and decision about the future development trend.

Gray modeling needs little information, usually more than four data. And it is not necessary to know the transcendental characteristics of the original data's distribution. Through limited times irregular or discrete smooth original sequence can be switched into regular sequence, which can make model high in precision, keep the model's original characteristics and better reflect the actual situation of the system.

Set the original data of a measurement as:

$$xi(0) = \left[xi^{(0)}(1)\, xi^{(0)}(2)\ldots xi^{(0)}(n)\right]. \tag{197.9}$$

It shows the answer of this measurement for a period of time in equal time interval values, n as sampling frequency, $n \geq 4$.

One accumulation get AGO sequence:

$$xi^{(1)} = (xi^{(1)}(1)\, xi^{(1)}(2)\ldots xi^{(1)}(n)). \tag{197.10}$$

The GM (1,1) Parameters put the parameters into the albino differential equation, according to the original data get the model of power system state estimation prediction model:

$$\tilde{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{ak} + \frac{b}{a}. \tag{197.11}$$

After using the state estimation algorithm based on the grey system theory to estimate the state of the power system, it is necessary to use the predicted results to make the power system's load flow calculation and thus to find all its representation parameters. Its mission is to provide the "working" state of the power system, and physical quantity should accurately meet the Kirchhoff integral theorem. Due to the limits of the hardware measuring equipments, we could not obtain all measurements, so it is indispensable to make the power flow calculation.

197.4 Software Realization and Testing

Take the example of the four nodes in network for power system simulation experiment, the configuration of power system's gauge system is shown in Fig. 197.1.

System consists of four nodes and four branches, node 1, 4 is power plant, node 2, 3 is load, branch 1, 2, 3, 4, is power line, branch 4 is transformer. The system bends with the parameter picture in Fig. 197.2.

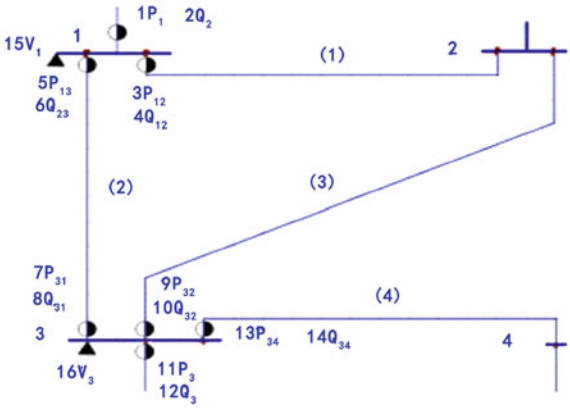


Fig. 197.1 The configuration of power system's gauge system

Line parameters as shown in Table 197.1.

Operation's algorithm routine, interface is shown in Fig. 197.3

In the process of operation, select the measurement V1, V3, P1, Q1, P3, Q3, P12, Q12, P13, Q13, P31, Q31, P32, Q32, P34, Q34, total 16, each measurement sampling ten times.

The measurement applies the grey system theory's prediction algorithm to known quantity, for the unknown quantities measurement, in known quantity measured values of the forecast of the calculation. After getting all the characteristic parameters of the state power system, calculate the parameters of the line loss and system's total loss, to facilitate dispatcher of electric power systems all information of mastery and judgment.

197.5 Analysis of Test Results

Apply these four nodes power system network to the state simulation test for 50 times, including measurement for 16, quantity of state for 7 and the results will be listed in Table 197.2, and statistically analyzed, thus it can evaluate the performance of state estimation procedure.

In the Table, $S_{M,\max,t}S_{E,\max,t}$ and $S_{R,maxt}$ are the largest errors of measurement, the biggest error of estimation and the biggest residual error (they are all absolute value) in t times sampling; $S_{M,t}$, and $S_{E,t}$ are the relative errors of measurement's root mean square value, the relative error of estimate sampling relative measurement error's RMS, $J\left(\widehat{x}\right)$, t is the sampled state estimates of the objective function value in t times.

From the last statistics, we can see

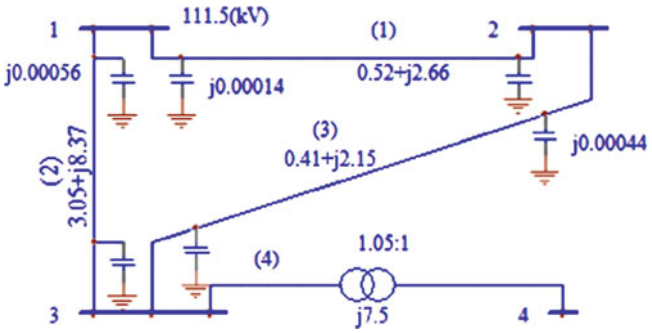


Fig. 197.2 System and line and parameters

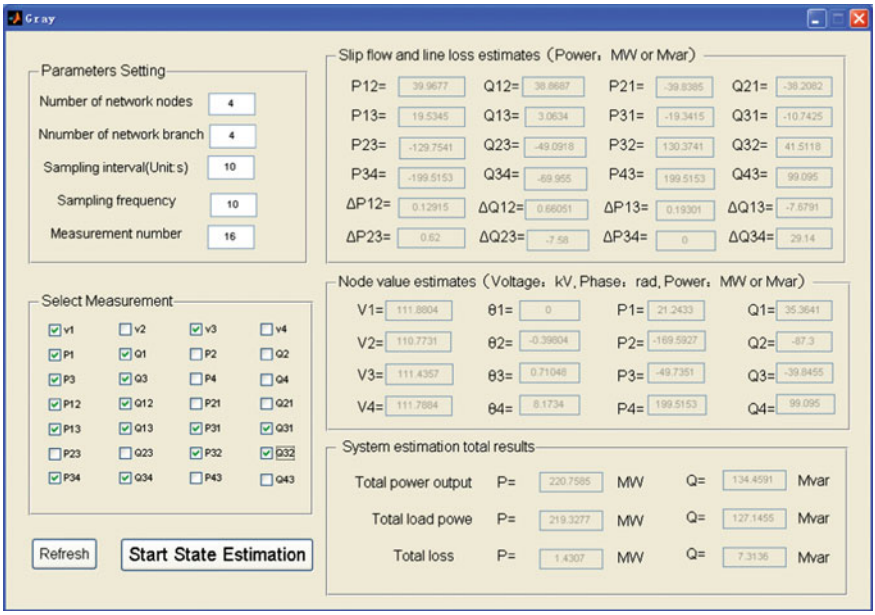


Fig. 197.3 Algorithmic program operation interface

Table 197.1 Line parameters

Branch number	Start number i	End number j	Resistance R(Ω)	Reactance X(Ω)	Susceptance (S)	Ratio K
1	1	2	0.52	2.66	0.00014	–
2	1	3	3.05	8.37	0.00056	–
3	2	3	0.41	2.15	0.00044	–
4	3	4	0.00	7.50	–	1.05

Table 197.2 State estimation results of tests

t	SM,max,t	SE,max,t	SR,max,t	SM,t	SE,t	$J(\hat{x})$, t
1	6.82	3.01	3.85	1.21	0.67	13.86
2	−2.26	2.38	1.92	1.20	0.78	10.39
3	2.99	1.57	2.44	1.16	0.56	16.01
4	2.82	−2.64	1.68	0.99	0.74	6.21
5	−3.59	−1.89	−3.32	0.89	0.73	8.29
⋮	⋮	⋮	⋮	⋮	⋮	⋮
46	−2.10	1.98	2.21	0.93	0.65	6.24
47	3.37	2.19	−2.49	0.76	0.57	15.03
48	3.66	1.59	3.26	0.95	0.49	9.92
49	3.75	0.98	3.47	1.30	0.39	16.20
50	3.12	1.85	−2.61	0.97	0.58	10.38
				0.98	0.65	9.64

50 mean values

$$\bar{S}_M = 0.98 \approx 1;$$

$$\bar{S}_E = 0.65 < 1;$$

$$\bar{J}(\hat{x}) = 9.64 \approx \mathrm{m} - \mathrm{n} = 16 - 7 = 9.$$

Filtering effect can be seen from the ratio of the estimation error $\bar{S}_E$ to the error $\bar{S}_M$.

In results, $\bar{S}_M \approx 1$, measurement system normal, according $\bar{S}_E$ see the filtering quality, $\bar{S}_E < 1$, enhance the state estimation accuracy. The smaller $\bar{S}_E$, the more effective state estimation function is, if $\bar{S}_E > 1$, state estimation doesn't work, filter fail. The results of state estimation whether close to the optimal objective function or not, also can be seen from the mean of the objective function $\bar{J}(x)$, this value should be closer to the redundancy of measurement system [3]. The results of estimation indicate that measurement system simulation is correct. State estimation procedure of good performance. State estimation has good filtering effect.

The advantages of the electric power system state estimation algorithm based on the grey system theory are summarized as follows:

(1) In respect of calculation speed, the electric power system state estimation based on the grey system theory is a small sample of state estimation, and compared with other calculations there is no iterative process and the algorithm is simple, so the superiority of the calculation speed is obvious.
(2) In respect of memory, the overall memory of the state estimation program includes two parts: the program itself and calculation arrays. The complexity of procedure is different on different occasions, for small power system, it is not necessary to use the sparse matrix program, and the procedure is short, so there is not much difference in different algorithms. For large and medium-sized power system, sparse matrix must be adopted, so the length of algorithm

process technology program has bigger difference. Thus, the advantages of the electric power system state estimation algorithm based on the grey system theory is reflected, which only needs to choose appropriate measurement information to carry on the gray prediction. The algorithm is still simple and clear, using much less memory than other algorithms, meanwhile because it is an estimated method, occupying a small amount of space, it has obvious advantages in using memory.

(3) In respect of adaptability to the measure system, it is all useful, regardless of the mixed, branch or nodes injected type. In the practical application, first make power system analysis of every power system and according to the power systems network change the trend of the equation. Once the trend of the equation has been made, this algorithm will quickly and accurately make state estimation of the power systems, and it has strong superiority to all types of measurement.

197.6 Conclusion

The greatest characteristic of this article based on the GM (1, 1) of power system state estimation is the dynamic estimate. It uses the measurement information over a period of time to predict the future of the state quantity, which is different from the existing state estimation methods. Its advantages lie in the fast calculation speed, small memory, high quality estimation, strong adaptability to the measurement system. The algorithm is simple and as long as it has more than four models it can get the modeling, and it is not necessary to know the original data's transcendental characteristics. It has high precision modeling, keep the original system characteristics and better reflect the real situation of the system; making up for the defects that the weighted least-square estimation method for the random noise must be normal distribution, the optimal estimation is consistent and unbiased, etc. Besides, the algorithm in respect of calculation speed and memory is far superior to the least-square state estimation. This algorithm also solves the problem of the limits of the practicality of dynamic estimation because of the large dimension of mathematical model, making dynamic estimation develop from theoretical study to practicality. The higher the redundancy of the measurement of the power system the more obvious its advantage compared with other algorithms. So next it is to study how to achieve the optimal balance between the redundancy of the measurement and the economic loss of the measurement equipment of hardware, thus, making power system state estimation based on the gray system theory more practical.

Although this article has proposed a dynamic model of the state estimation of the power system, it is still a new research field in power system, so in the practical process other problems must arise. Since the real-time state estimation is a comprehensive huge functional system, this article can only make in-depth study of

some problems, not covering every aspect. It is unavoidable for our work to have flaws and defects but we will try to improve them in our later work. Believe that the dynamic estimation of the power system will make progress along with the development of electric power industry in china.

Acknowledgments This paper was supported by National Natural Sciences Foundation of China(51007030).

References

1. Sifeng L, Dangbang G, Dang Yaoguo M (2000) Grey system theory and its application. Science Press, Beijing
2. Wenjuan C, Guoxun J, Liu Jun J (2009) The gray system theory in predicting the application in power. Ind Saf Environ Prot 35(9):1
3. Yu Erkeng M (1985) Power system state estimation. Water Power Press, Beijing

Chapter 198
Communication Channel Reconstruction Technology of Wide-Area Protection Under Extreme Catastrophe

Zhenxing Li, Xianggen Yin, Zhe Zhang and Zhiqin He

Abstract A new scheme of communication channel reconstruction technology (CCRT) for wide-area protection (WAP) under extreme catastrophe scenarios is proposed for the purpose of safeguarding the safety and reliability of relay protection. The communication hierarchy of WAP based on MPLS VPN is obtained according to the characteristic of communication system. Under the damage of optical communication channel, CCRT based on circuitous way is first realized by sturdy optical fiber, carrier, microwave and telecommunication channels and so on, and then CCRT based on wireless sensor network (WSN) is built when the circuitous path cannot reconstruct communication channel for more channels damaged. The key technology of WSN, such as the supply energy pattern, router algorithm and multi-hops way, are also discussed. The test results of multiple schemes show that CCRT based on WSN can be suitable for WAP.

Keywords Wide-area protection · Extreme catastrophe · Channel reconstruction · WSN

198.1 Introduction

Recent research of WAP has clearly demonstrated the advantage to overcome the vulnerability of the interconnected power system when operating beyond its intended design limits [1]. But the research is designed under the normal state

Z. Li (✉) · X. Yin · Z. Zhang · Z. He
State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Electric Power Security and High Efficiency Laboratory, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China
e-mail: lzx2007001@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_198,

of undamaged wide-area measurement and SDH ring network, and WAP system structure is less researched under the channels damaged. Not only were the large power equipments badly damaged and the electric power communication channels also interrupted under the influence of extreme natural disasters, the direct result is that more relaying devices is out of service and malfunction, and even causes cascading trips on the grid which lead to a widespread blackout [2, 3].

The accidents occurred in large areas of the south China early 2008, sustained widespread low temperature, snow and freezing weather, led directly to the power grid to be a damager situation, such as massive lines disconnection, towers collapse, OPGW optical fibers interrupt and relay protection out of service. The damages seriously endangered the safe and stable operation of the power grid [4], at the same time, WAP based on wide-area communication net was disabled. So a strong running communication system must be considered for the research of WAP when single link and more links fail. The CCRT of WAP under the extreme catastrophe has great significance.

WSN is a novel technology in electric communication fields, which has superior characteristics, such as high monitoring accuracy, multi-hops way, self-organizing, tolerant properties, and so on. It is important that WSN can coverage wide-area power grid [5, 6]. In this paper some key technology is probed when WSN applies to WAP system, which aims to offer a new idea on CCRT.

198.2 The Communication Hierarchy of WAP

Current power system communication network is built on SDH ring networks composed of the electric special optical fibers, OPGW and ADSS, laying between substations. So the research around SDH fiber network to construct the wide-area protection communication system is the key.

MPLS technology is integrated on IP and ATM technology. It adopts hardware switching technology, avoids routing query, and accelerates forwarding speed. The technology can be easily realized in the basis of traffic engineering (TE), different service Diff-Service model and virtual private network (VPN).

Therefore, the information transmission scheme of WAP can be realized based on networking technology, IP OVER SDH, and multiple protocol labels switching (MPLS). While the channel congestion is avoided by the traffic engineering (TE), the highest priority of WAP information can be guaranteed using different service model, and the segregation between different requirements can be satisfied with the flexibility and security of VPN. Figure 198.1 shows the communication hierarchy of WAPS. The system is composed of a regional design unit (RDU), six sub stations, four link core routers (LCR), and seven link edge routers (LER). The link bandwidth between LCR and LER is set 2 Mb/s, and the link bandwidth between

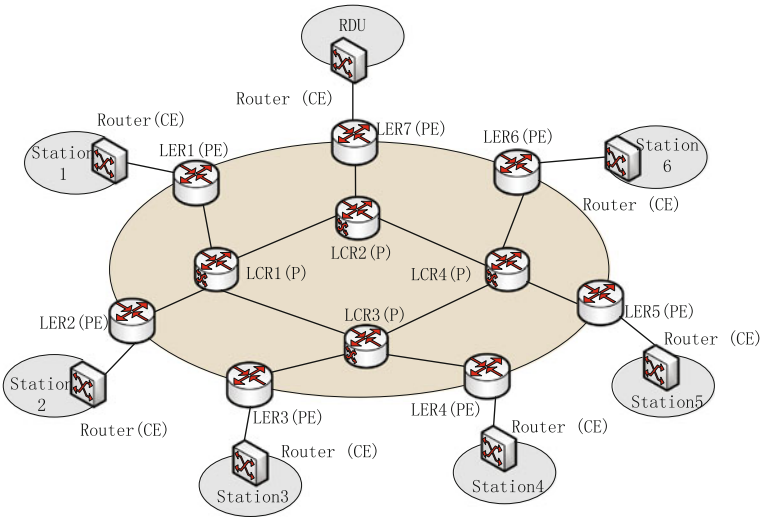


Fig. 198.1 System diagram

LCRs, LCR and RDU are set at 155 Mb/s for more data communication in the main links.

198.3 CCRT of WAP Based on Circuitous Path

CCRT is proposed when single link or more links fail in the communication system. A circuitous channel is built to transmission data information by other sturdy links. As shown in Fig. 198.1, there are a variety of circuitous communication modes when LER2 disconnected, such as the following

LER2–>LER1–>LSR1;
LER2–>LER3–>LSR3–>LSR1;
LER2–>LER1–>LER7–>LSR2–>LSR1.

The above analysis shows that the circuitous channel is not unique when some a link failure. Therefore, a method of router path selection, router algorithm, must be determined. The shortest path algorithm-Dijkstra algorithm [7, 8] is referred to design the router path based on the state of all neighboring routers referring to IP router protocols OSPF. The known conditions of the algorithm are the entire network topology and the length of each link, which the link length just reflects the distance between the two substations according to the known topology structure of the OPGW network. Therefore, the shortest path algorithm based on the circuitous way can be realized in electrical communication network. Note that minimum delay is the main aim to design router path,

and the improvement of relaying delay can be acquired when the length parameter is replaced by the link delay parameter.

198.4 CCRT of WAP Based on WSN

Under extreme natural disasters, more communication channels may be broken down, and the damages give rise to many information isolated islands which do not have any circuitous channels. Even if there is a sturdy circuitous channel, it cannot stand the heavy communication traffic as most of the other channels are ruined and a large amount of data is arranged to be transmitted through it. It is necessary to construct temporary channels for relay protection system introducing new techniques. The wireless communication is the best selection as no wiring is need and no channel interrupting exists.

Figure 198.2 shows a typical communication channel reconstruction system based on WSN. Expanding the wireless communication modules (e.g. CC1000) in protective devices or introducing wireless sensor gateways makes the protective devices a node of WSN. Then according to the requirements of monitoring system, the corresponding sensors nodes can be installed on the towers. Meanwhile, via expanding the wireless communication modules in ECT (MU), the current signal can be transmitted through WSN. Finally all the sensor nods make up WSN in the form of ad hoc networks. When the OPGW channels between Substation-A and Substation-B are detected ruined, the protective devices in the substations immediately switch to WSN and the protective information (current) is transmitted by multi-hop communication. This constructs a reliable emergency channel for protection system. The sampled data can also be transmitted to power communication net by wireless gateway.

198.4.1 Energy Supply of WSN Node

(1) CT-battery supply energy jointly

Figure 198.3 demonstrates the entire structure of energy supply system of WSN nodes. When the transmission lines operates normally, the output voltage of CT energy supply module is larger than the battery voltage, so the diode is off and the energy is supplied by CT. When the transmission lines fault or operate in open circuits, the diode switches on and the battery begin to supply energy. The process of diode's on and off is instant, so the energy supply is ensured in any case. Since transmission line faults can't last for long, the battery need not supply energy for long. Consequently, CT and battery supplying energy jointly can satisfy the energy consume of sensor nodes.

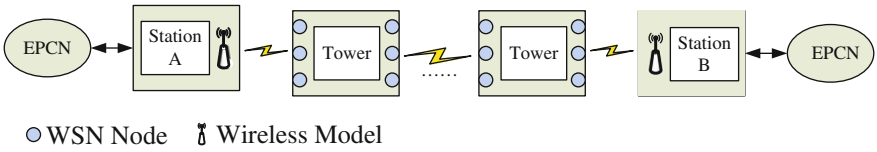
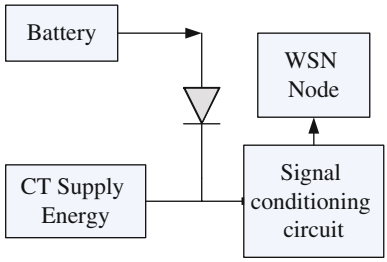


Fig. 198.2 Communication channel reconstruction on WSN

Fig. 198.3 Supply energy system of WSN node



(2) Pulse discharge working pattern

When the battery supplies energy, we should optimize the usage of battery in each node to maximize the life of the whole network. Pulse discharge for each node is an effective method. The energy in each sensor node is mostly consumed by wireless communication modules. The wireless communication modules have four types of working pattern, Sending, Receiving, Idle and sleeping. Therefore, the nodes are switched into sleeping pattern when they are not communicating until they need communication. This pulse discharge working pattern can not only decrease energy consumption but also replenish the battery using the restoration effect of battery during sleeping pattern. These measures can prolong the battery's life greatly.

198.4.2 Routing Algorithm

Routing algorithm is to transmit data packet from source nodes to destination ones. WSN is greatly dependent with its application. No general routing algorithm exists and there may be huge differences between two different routing algorithms. The WSN used as emergency channel in WAP needs to query sampling data in each node, the collecting nodes send querying demands and sensor nodes submits sampling data to collecting nodes, therefore the routing algorithm based on querying is most suitable here. The nodes are arranged along a long and narrow transmission corridor, so the hierarchy querying protocol, for example TopDisc, GAF, and LEACH, is not fit for this case. According to the practical characteristic, based on Directed Diffusion (DD) and Dumor Routing (DR), a routing protocol which is suitable to communication system of WAP has been designed shown in Fig. 198.4.

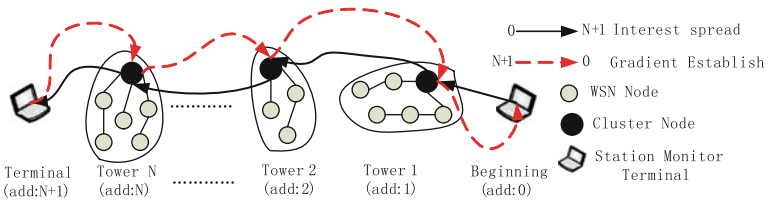
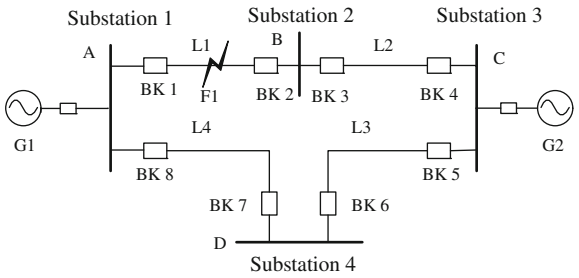


Fig. 198.4 Mechanism of router algorithm

Fig. 198.5 Framework of regional power network



198.5 Principle of Wide-Area Protection Based on Directional Comparison

The study on the algorithm of WAP using fault directional information is the one of main direction [9, 10]. Make F1 single phase earth fault as an example to explain the principle of wide-area protection, as shown in the Fig. 198.5.

The wide-area protection decision-making unit will modify the topology of protected areas in time according to the receiving breaker stature information. According to the Fig. 198.5, if the breakers are closed (1 as close, 0 as is connection). At the same time, relating direction information can be obtained based on the breaker stature information and fault direction information, as shown in Table 198.1.

The system structure relating matrix R can be obtained utilizing the relating information.

$$\begin{array}{c} \text{Line} \quad 1 \quad 2 \quad 3 \quad 4 \\ R = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{substation 1} \\ \text{substation 2} \\ \text{substation 3} \\ \text{substation 4} \end{array}$$

The degree of vertex, which is defined as Additional result of the same row vector, equals to 2 shows the corresponding component is faulty one. DV of every row can be acquired from the matrix R.

Table 198.1 Information of associated direction

	Line 1		Line 2		Line 3		Line 4	
Substation 1	1	1	0	\	0	\	1	-1
Substation 2	1	1	1	-1	0	\	0	\
Substation 3	0	\	1	1	1	-1	0	\
Substation 4	0	\	0	\	1	1	1	1

$$\begin{matrix} \text{Line} & 1 & 2 & 3 & 4 \end{matrix}$$
$$DV = \begin{bmatrix} 2 & 0 & 0 & 0 \end{bmatrix}$$

Therefore, we get a judgment that line 1 has faults.

There is no complicated data calculation for the discriminating principle based on state relation, and it is exact and fast. The other superiority is low data communication traffic. Once the communication channel is damaged, CCRT of WAP is apt to realize.

198.6 Tests

(1) System platform

The simulation of CCRT is test on the platform based on Crossbow series WSN products.

WSN node: IRIS XM2100CA.

WSN gateway: MIB600 (Ethernet interface), MIB520CA (USB interface).

WSN data acquisition card: MTS420CA, MDA320CA.

The tests platform based WAPS is shown in Fig. 198.6. A master station is selected to realize system decision of WAPS, where the decision unit has some abilities of fault component identification, information management, and learning to adapt the running changes of power system. And some other substations are seen as sub stations to achieve the local measurement and control function, where the local measure and control unit can make some serial of judgments and inferences through perception changes of external environment, and then actively cooperate with the decision unit through WSN.

(2) Communication capacity simulation of WAP

According to the bidirectional communications control and data transmission, the test on the basis of the algorithm of direction comparison protection aims to check the influences of different communication traffic.

IRIS node data transmission rate sets 250 kb/s, and data information includes the results of directional protection operations information and breakers state information, which are Boolean values. The maximum transmission bytes go to 60 under single fault. The results show the delay is no more than $0.24 \times N$ ms when N is less than 100.

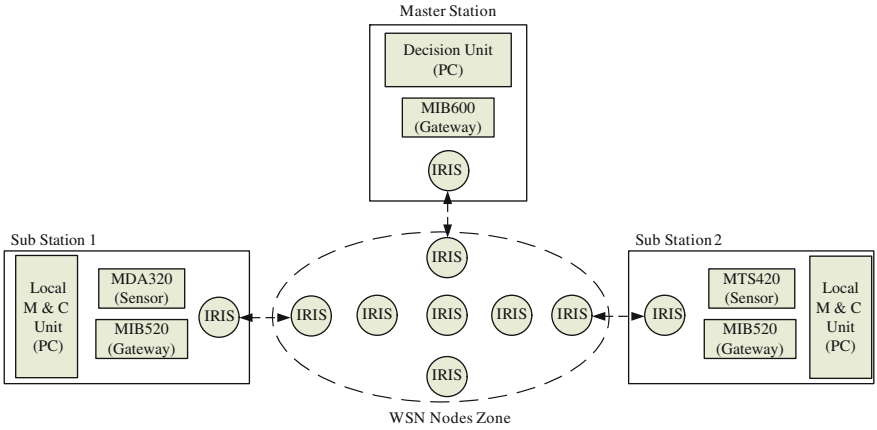


Fig. 198.6 Tests system platform

(3) Multi-hops simulation of WSN

Generally, WSN nodes are installed on the distributed towers or transmission line. Only multi-hops way of nodes can the communication between two terminals of long line be realized. The optimal path selection under multi-hops way is great significant based on the router algorithm.

The results show that the communication hierarchy of WSN under multi-hops way based on DD protocol applies to less than 50 km lines. The longer is line, the longer is also the delay, and the fewer disadvantages are the relaying.

(4) Delay simulation of WSN

Relay protection to delay has strict demands. The long delay directly influence on the rapidity of protection operation and then the stability of power system. The statistic delay tags are obtained by increasing WSN node one by one in the links.

The results show that the communication delay is concerned with communication data size and WSN capacity. The mean delay is less than 0.5 ms when the ratio of communication data size and WSN capacity is less than 0.5.

(5) Battery performance simulation of WSN

The CT supply energy way fails under extreme catastrophe, and the supply energy time is limited by battery. How can the time go to the optimal and the maximum is the key using multi-hops, monitoring and control technology.

The results show that the sleep of battery has obviously affected on battery restore effect, but the time and the restore has not linear relationship. The restore is over when the sleep is more than 6 h. The time gain can go to 12% by sleep.

198.7 Conclusion

The communication hierarchy of WAP is proposed, and the communication networks are built on MPLS VPN technology. The communication channel reconstruction technology based on circuitous way or WSN is proposed, and the key technology is discussed. The results of different tests show that the CCRT can be suited to the WAP communication system.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (50837002, 50877031), National High Technology Research and Development of China (863 Programme) (2009AA05Z208) and the Specialized Research Fund for the Doctoral Program of Higher Education (20090142110055).

References

1. Novosel D, Bartok G, Henneberg G et al (2010) IEEE PSRC report on performance of relaying during wide-area stressed conditions. IEEE Trans Power Delivery 25(1):3–16
2. Tan JC, Crossley PA, Mclaren PG et al (2002) Sequential tripping strategy for a transmission network back-up protection expert system. IEEE Trans Power Delivery 17(1):68–74
3. Eissa MM, Masoud ME, Elanwar MMM (2010) A novel back up wide area protection technique for power transmission grids using phasor measurement unit. IEEE Trans Power Delivery 25(1):270–278
4. PENG X-y, ZHOU H-m (2009) Damage condition of overhead transmission lines in guangdong power grid caused by icing disaster in 2008 and analysis on key impacting factors. Power Syst Technol 33(9):108–112 (in Chinese)
5. Nordma MM, Korhonen T (2006) Design of a concept and a wireless ASIC sensor for location earth faults in unearthed electrical distribution networks. IEEE Trans Power Delivery 21(3):1074–1082
6. Akkaya K, Younis M, Youssef W (2007) Positioning of base stations in wireless sensor networks. IEEE Commun Mag 45(4):96–102
7. Chao Y, Hongxia W (2010) Developed dijkstra shortest path search algorithm and simulation. ICCDA 2010 international conference on computer design and applications 1:116–119
8. Ng F, Butler G, Kay J (1995) An intelligent tutoring system for the dijkstra-gries methodology. IEEE Trans Softw Eng 21(5):415–428
9. Lin X, Ke S, Li Z et al (2010) A fault diagnosis method of power systems based on improved objective function and genetic algorithm-tabu search. IEEE Trans Power Delivery 25(3):1268–1274
10. Düstegör D, Poroseva SV et al (2010) Automated graph-based methodology for fault detection and location in power systems. IEEE Trans Power Delivery 25(2):638–647

Chapter 199
Study on Unbalance Factor of Sequence Impedance for Parallel Multi-Circuit Cable Lines

Xing Deng, Xianggen Yin, Zhe Zhang, Zhiqin He and Zhenxing Li

Abstract Parallel multi-circuit cable lines are wildly used to improve the transmission capacity of urban power grid, but the study on sequence impedance parameters' asymmetry of multi-circuit cable lines is not available. To analyze the asymmetry of multi-circuit cable lines' parameters, a method to measure the unbalance factor is introduced in this paper. Furthermore, take engineering practice as an example, the electromagnetic unbalance factor of two-circuit cable lines under different arrangements are calculated by this method. Finally, a comparison among these unbalance factor is made. The comparative results show that the triangle arrangement has better symmetry than straight-vertical and straight-horizontal arrangement, in which the triangle-vertical arrangement's symmetry is the best.

Keywords Parallel multi-circuit cable lines · Sequence impedance parameters · Arrangement · Unbalance factor · Symmetry

199.1 Introduction

The non transposed or incomplete transposed transmission line may cause the unbalance of electrical parameters. As the result, power equipments such as generator and transformer, cannot normally operation, and even threaten the

X. Deng (✉) · X. Yin · Z. Zhang · Z. He · Z. Li
State Key Laboratory of Advanced Electromagnetic Engineering and Technology,
Electrical Power Security and High Efficiency Lab,
Huazhong University of Science and Technology,
Wuhan 430074, People's Republic of China
e-mail: dengxing333@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_199,

security and stability of power system [1, 2]. In recent years, with the development of power cable technology and urbanization, power cable is widely applied into the construction of power grid because of its unique characteristics. However, the problem of cable parameters' asymmetry has been increasingly outstanding. Additionally, supply load in many cities of large or middle size is on the rapid increase. In order to satisfy larger demands from the supply load and fully utilize transmission capacity of existing cable lines, a feasible and reasonable method is provided which combine the existing power cables and new ones into running in parallel (Or makes the existing cable lines with small capacity running in parallel) [3]. This method can effectively reduce the cost for capacity expanding reform of cable lines and improve the economic benefits. But it also arouses some new problems which need to be solved immediately, such as calculation of parameters and analysis on asymmetry for parallel multi-circuit cable lines. Therefore, study on these aforementioned problems is highly significant in such aspects as guiding the type selection of cables and guaranteeing the security operation of power grid.

Most of power cables which are currently applied to high voltage transmission system have the single core structure. During the capacity expansion and reform of cable lines, the possible laying methods of parallel two-circuit cable lines include cross straight-horizontal, independent straight-horizontal, straight-vertical, triangle-horizontal and triangle-vertical and so on. Different arrangements will affect the calculating results of sequence impedance parameters of cable lines. Meanwhile, because the core wire of cable lines seldom transpose, the parameters of cable lines will become more asymmetric. Moreover, different arrangements are also affect the asymmetry of cable line's parameters. Almost all the studies on parameters' asymmetry of multi-circuit transmission lines are only concerned about overhead lines [4–6]. However, because the metal sheath and differences in laying methods for power cables, the analysis results are different from each other. References [7, 8] presents some algorithms for cable lines' parameters, and analyzes the relations between cable lines' sequence impedance parameters and different grounding schemes. Reference [9] analyzes the symmetry of parameters of high voltage AC cable. But these researches are only focus on the single-circuit cable line.

This paper analyzes the unbalance factor of sequence impedance parameters for parallel two-circuit cable lines under several typical phase sequence arrangements, and presents some suggestions for phase sequence arrangement of parallel two-circuit cable lines.

199.2 Electromagnetic Unbalance Factor

Asymmetry of cable lines' parameters can be represented in unbalance factor, which can be divided into electromagnetic unbalance factor and electrostatic unbalance factor. These two factors describe the unbalance degree of line impedance and capacitance parameters [10].

The cable line is a standard cylindrical capacitor. Its core wire and grounded metal sheath constitute two electrodes. Due to the shielding effect executed by metal sheath on electric filed, different core wires will not affect each other's electric field, and positive-sequence, negative-sequence and zero-sequence capacitance of cable lines are equal to each other. For this reason, the electrostatic unbalance factor will not be discussed in this research.

The analysis on asymmetry of overhead line for reference, the electromagnetic unbalance factor of two-circuit cable lines running in parallel can be divided into two categories: through and circular.

When a positive-sequence voltage is applied to the head of two-circuit cable lines and three-phase is grounding at the end. Under this condition, the zero-sequence and negative-sequence through electromagnetic unbalance factor can be denoted by M_{0T} and M_{2T}. Similarly, the zero-sequence and negative-sequence circular electromagnetic unbalance factor can be denoted by M_{0C} and M_{2C}. These variables can be defined as

$$M_{0T} = \frac{I_{0I} + I_{0II}}{I_{1I} + I_{1II}},\ M_{2T} = \frac{I_{2I} + I_{2II}}{I_{1I} + I_{1II}} \tag{199.1}$$

$$M_{0C} = \frac{I_{0I} - I_{0II}}{I_{1I} + I_{1II}},\ M_{2C} = \frac{I_{2I} - I_{2II}}{I_{1I} + I_{1II}} \tag{199.2}$$

where I_{0I}, I_{1I}, I_{2I}, I_{0II}, I_{1II} and I_{2II} are zero-sequence, positive-sequence and negative-sequence current in circuit I and II during a fault occurs at the end of the cable lines.

199.3 Calculation and Analysis

Phase sequence arrangement of cable lines and grounding scheme of its sheath are the important influential factors for unbalance factor of parameters. The typical laying methods for parallel two-circuit cable lines in the engineering include cross straight-horizontal, independent straight-horizontal, straight-vertical, triangle-horizontal and triangle-vertical, as shown in Fig. 199.1. Different laying methods will directly influence the calculation result of elements in impedance matrix, and thus have larger effect on unbalance factor of cable line's parameters [11]. One-end grounding and two-end grounding scheme of metal sheath are usually applied to cable line which is shorter than 500 m in length. Cable lines with this category are usually short, and therefore the unbalance of electrical quantities caused by asymmetrical parameters barely has the influence on power system.

According to the analysis aforementioned, a case of cable lines in engineer practice is used. The type of this cable is YJLW03-1 $\times$ 1200 mm^2 $-$ 127/220 kV. Based on the methods shown as Eqs. (199.1) and (199.2), Fig. 199.1 displays the electromagnetic unbalance factor of two-circuit cable lines running in parallel under different arrangements. In Fig. 199.1, the distance of d_1 is 300 mm, d_2 is 500 mm, and h is 1,000 mm. The total length of this two-circuit cable lines is

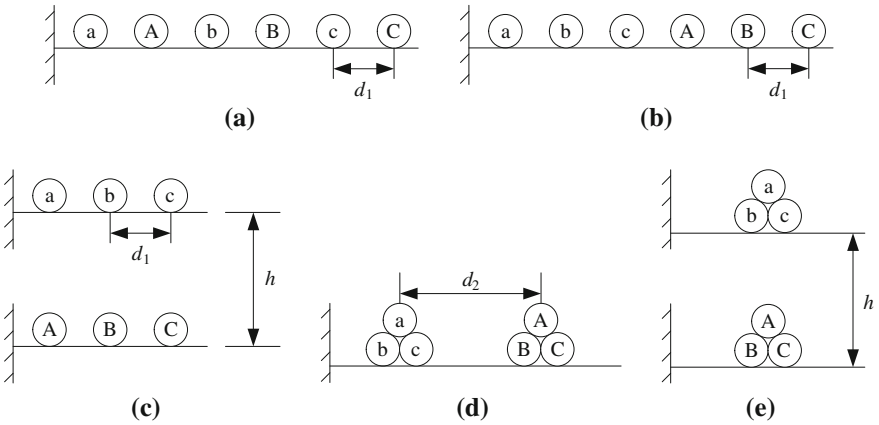


Fig. 199.1 Laying methods of two-circuit cable lines. **a** Cross straight-horizontal. **b** Independent straight-horizontal. **c** Straight-vertical. **d** Triangle-horizontal. **e** Triangle-vertical

Table 199.1 Amplitude of sequence current under cross straight-horizontal arrangement

	I_{1I}	I_{2I}	I_{0I}	I_{1II}	I_{2II}	I_{0II}
I_{am}/kA	24.3040	4.6478	1.9834	23.5925	4.7508	2.0164
P_s/%		19.12	8.16		20.14	8.55

20 km, and the grounding scheme of metal sheath is cross-connection grounding. Applying a positive-sequence voltage of 127 kV at the head of the cable lines (which equals to line voltage of 220 kV), based on the algorithm proposed in Ref. [11], the positive-sequence, negative-sequence and zero-sequence current in circuit I and II under different arrangements are calculated. The detailed calculation results can be found in Tables 199.1, 199.2, 199.3, 199.4, 199.5, where I_{0I}, I_{1I}, I_{2I}, I_{0II}, I_{1II} and I_{2II} are zero-sequence, positive-sequence and negative-sequence current in circuit I and II during a fault occurs at the end of the cable lines, I_{am} is amplitude of current and P_S is ratio between negative(zero)-sequence and positive-sequence current.

It is observed that the P_s under the triangle arrangement is smaller than others, and the P_s under the triangle-vertical arrangement is the smallest. The straight-vertical arrangement has the larger P_s of negative-sequence current compared to the independent straight-horizontal arrangement which has the smaller figure of positive-sequence current. The cross straight-horizontal arrangement has the largest P_s of negative-sequence and zero-sequence current.

According to the definition of electromagnetic unbalance factor, put these sequence currents in Tables 199.1, 199.2, 199.3, 199.4, 199.5, into Eqs. 199.1 and 199.2, the electromagnetic unbalance factor under three different arrangements can be obtained and shown in Table 199.6.

Table 199.2 Amplitude of sequence current under independent straight-horizontal arrangement

	I_{1I}	I_{2I}	I_{0I}	I_{1II}	I_{2II}	I_{0II}
I_{am}/kA	34.7558	2.3920	1.8427	35.1297	2.6027	1.8414
P_s/%		6.88	5.30		7.41	5.24

Table 199.3 Amplitude of sequence current under straight-vertical arrangement

	I_{1I}	I_{2I}	I_{0I}	I_{1II}	I_{2II}	I_{0II}
I_{am}/kA	32.1849	5.3553	0.0786	32.1849	5.3553	0.0786
P_s/%		16.64	0.24		16.64	0.24

Table 199.4 Amplitude of sequence current under triangle-horizontal arrangement

	I_{1I}	I_{2I}	I_{0I}	I_{1II}	I_{2II}	I_{0II}
I_{am}/kA	48.3237	1.0719	0.8212	48.2527	1.0774	0.8212
P_s/%		16.64	0.24		16.64	0.24

Table 199.5 Amplitude of sequence current under triangle-vertical arrangement

	I_{1I}	I_{2I}	I_{0I}	I_{1II}	I_{2II}	I_{0II}
I_{am}/kA	48.0665	0.2379	0.2845	48.0701	0.2416	0.2814
P_s/%		0.50	0.59		0.50	0.59

Table 199.6 Unbalance factor under different arrangements

Arrangement	Through electromagnetic unbalance factor		Circular electromagnetic unbalance factor	
	M_{0T}/%	M_{2T}/%	M_{0T}/%	M_{2T}/%
Cross straight-horizontal	8.35	19.62	−0.0688	−0.2151
Independent straight-horizontal	5.27	7.15	0.0018	−0.3015
Straight -vertical	0.24	16.64	0	0
Triangle-horizontal	1.7	2.23	0	0
Triangle-vertical	0.59	0.5	0.0032	0.0038

There are two main factors may lead to the asymmetry of parallel cable lines' parameters. One is asymmetric arrangement of three phases of single-circuit cable line, and the other is asymmetry of parameters caused by mutual-impedance between circuits. Unbalance factor in Table 199.6 is the comprehensive reflection of these two factors. In order to better analyze the latter factor, unbalance factor of single-circuit cable line's parameters is calculated under different situation and the calculation result is compared to unbalance factor of two-circuit cable lines' parameters. The results are shown in Table 199.7.

Table 199.7 Comparison of unbalance factor between single-circuit and two-circuit cable lines

Arrangement	Two-circuit cable lines		Single-circuit cable line	
	M_{0T}/%	M_{2T}/%	M_{0T}/%	M_{2T}/%
Cross straight-horizontal	8.35	19.62	0.39	12.09
Independent straight-horizontal	5.27	7.15	0.37	14.74
Straight-vertical	0.24	16.64	0.37	14.74
Triangle-horizontal	1.7	2.23	0	0
Triangle-vertical	0.59	0.5	0	0

From the data in Table 199.6, circular electromagnetic unbalance factor under three arrangements are comparatively small, especially under the straight-vertical arrangement, which can be ignored. Therefore, the through electromagnetic unbalance factor under these three arrangements are mainly compared here.

First of all, the triangle arrangement has the better balance than other arrangements and the triangle-vertical arrangement is the best. Among the non-triangle arrangements, the straight-vertical arrangement has small zero-sequence unbalance factor and comparatively larger negative-sequence unbalance factor. Based on Table 199.7, this unbalance factor is caused by asymmetric arrangement of three phases for single-circuit cable line, but two-circuit cable lines running in parallel will execute less influence; negative-sequence unbalance factor under the independent straight-horizontal arrangement is greatly reduced compared to single-circuit. This means negative-sequence asymmetry has been improved for cable lines running in parallel under the independent straight-horizontal arrangement, but the zero-sequence unbalance factor has been increased. From the perspective of influence on grid's operation, negative-sequence component has larger influence on the normal operation. Therefore, the independent straight-horizontal arrangement is better than the straight-vertical arrangement, and the cross straight-horizontal arrangement has the larger negative-sequence current and zero-sequence current as well as the worst balance.

199.4 Conclusion

For the multi-circuit cable lines running in parallel, the laying method will directly affect the symmetry of parameters. Among the different arrangements of two-circuit cable lines, the triangle arrangement has the better symmetry than other arrangements, and its triangle-vertical pattern is best for the balance. For the non-triangle arrangements, the independent straight-horizontal arrangement is better than the straight-vertical arrangement and the cross straight-horizontal arrangement has the worst asymmetry.

Acknowledgments This work was financially supported by the National Natural Science Foundation of China (50837002, 50877031), National High Technology Research and Development of China (863 Programme) (2009AA05Z208) and the Specialized Research Fund for the Doctoral Program of Higher Education (20090142110055).

References

1. Wang X, Lin J, Ban L (2009) Analysis on unbalance factor and transposing modes for 1,000 kV double-circuit transmission lines on the same tower. Power Syst Technol 33(5):1–5 (in Chinese)
2. Ding H, Duan X (2004) Unbalance issue caused by un-transposed transmission lines and its solution. Power Syst Technol 28(19):24–28 (in Chinese)
3. Kong X, Wang Y, Zhang Z (2010) Calculation of induced voltage in metal shield of single-core cable operated in parallel. In: 2010 international conference on power system technology (POWERCON), Hangzhou, China
4. Wei G, Zhang Z, Fang Z (2004) Analysis of imbalance of the parallel multi-circuit transmission line. High Volt Eng 30(10):9–11 (in Chinese)
5. Liu Y, Zhang Z, Yin X (2009) Parameter symmetry of transmission lines of double circuits on same tower. Power Syst Technol 33(S2):192–195 (in Chinese)
6. Wei G, Huang J (1998) Sequence parameters and unbalance coefficients calculation for parallel multi-circuit transmission lines on the same tower. Power Syst Technol 22(10):8–11 (in Chinese)
7. Mao W (2002) Computation and analysis on phase sequence impendence of power cable. Distrib Util 19(4):24–25 (in Chinese)
8. Tang Q, Liu B, Yang H (2009) Effect of the earthing method of the metal sheath in HV cables on line parameters. Electr Wire Cable 2:43–46 (in Chinese)
9. Zhu Z (2007) Influence of electrical parameters of high voltage AC submarine cables on characteristic of symmetry of the grid-link line. High Volt Eng 33(6):177–179 (in Chinese)
10. Anderson JG (1981) Transmission line reference book 345 kV and above. Power Industry Press, Beijing
11. Wang Y, Kong X, Deng X (2010) Calculation and analysis of sequence impedence for parallel multi-circuit cable lines. 2010 international conference on power system technology (POWERCON), Hangzhou, China

Chapter 200
The Design and Application of CAN General Line Based on SJA1000 in the Coal Mine Long Distance Electric Communication System

Guofang Qin, Yingli Lv and Zhongqi Wang

Abstract To describe the application of CAN general line based on SJA1000 in the coal mine long distance electric communication system, which improved the traditional CAN (Controller Area Network, CAN) contact general line driver hardware circuit and gave a hint to design the software. It had been proved in the practice that the system worked well and had a better result. This system acquired a fine result when tried out in units such as no.1 tai hang coal mine and so on and had an extremely high application value.

Keywords SJA1000 · CAN general line · CAN contact · Receiver isolate · Circuit

200.1 Introduction

The field bus was widely adopted because of its low cost, and meeting the communication requirements under the environment of industrial field with the development of industrial control at the site and automation technology on

G. Qin (✉) · Y. Lv
Department of Electrical Engineering, Jiyuan Vocational and Technical College, Jiyuan 459000, Henan, China
e-mail: qinguofang@sohu.com

Y. Lv
e-mail: xalvyingli@163.com

Z. Wang
Jiyuan Power Supply Company of Henan Electric Power, Jiyuan 459000, Henan, China
e-mail: wangzhongqi@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_200,

production. The CAN general line is a kind of serial communication network which is one type of advanced field bus technology, prevalent currently as well as high confidentiality, supporting distributed or real-time control. The CAN general line is currently popular a kind of performance comparison advanced field bus technology, it is a high confidentiality, effective support distributed control or real-time control of serial communication network [1].

In view of the excellent performance of CAN general line, it could find and eliminate hidden security trouble as early as possible, handle online fault in real time and further improve the communication reliability of the system if the information of operation status of voltage, current leakage and vacuum circuit breaker was real-time monitoring shows and real-time alarm, meanwhile, each monitoring information was config. Figure 200.1 flexibly according to customer request and using conditions of various power were inquired on the screen. According to the practical work, CAN general line based on SJA1000 was provided in this paper to solve the remote communication problems in the coal mine power supply system.

200.2 System Design

200.2.1 System Topology Design

A bilaminar bus networks, a real-time network system, was established in order to realize information integration of the power supply in coal mine communication system, as shown in Fig. 200.1. Main function of Management network of Ethernet communication bus based on TCP/IP in the upper covered all of the power equipment of the ground and underground was to finish a lot of work such as integrated monitor, control and management etc. While the field network adopting field bus protocol was to complete data acquisition and control of each equipment at the production site. The communication two layer of network was realized through embedded gateway in Fig. 200.1.

200.2.2 Hardware Design

(1) Hardware Realization of CAN Contact

It was an important link for the communication interface design of CAN-bus which closely related to proper operation of the equipment. Figure 200.2 showed the actual circuit diagram of CAN-bus communications unit whose circuit structure showed as follows:

+5 v power input, the power-on Reset Chip MAX813L not only could keep start system rightly when powered on, but also avoid program coming up “Disorderly

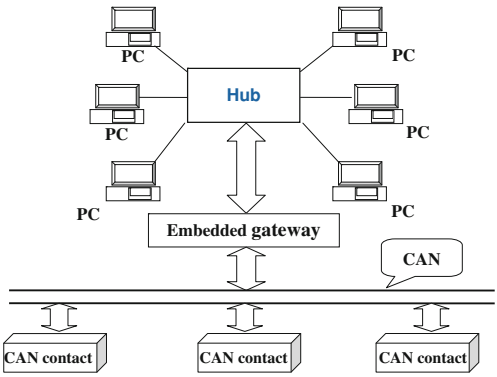


Fig. 200.1 Structure of system topology

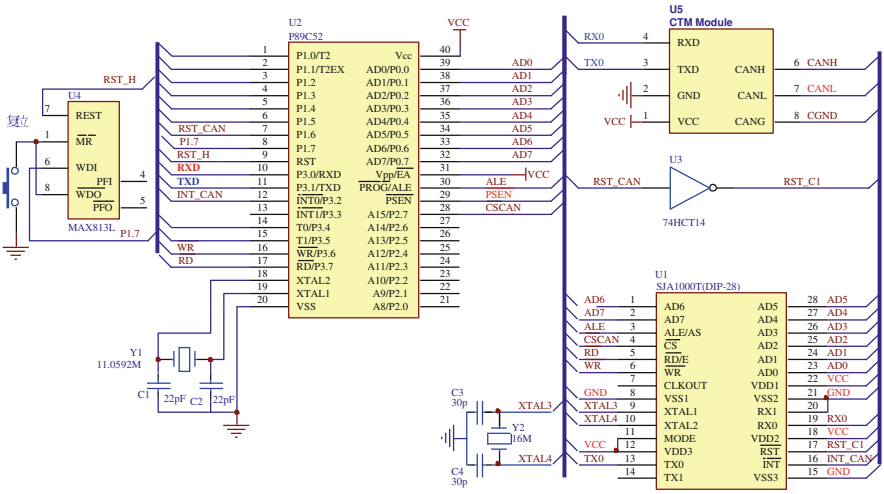


Fig. 200.2 Circuit diagram of CAN-bus communications unit

flying" or turning into" dead circulation", was adopted in order to ensure the healthy and steady operation of system. P89C52SCM of PHILIPS, a derived device of 80C51 microprocessor, manufactured with CMOS technology and the same command system with 80 C51 absolutely, was used by Microprocessor. Reset signal from MAX813L could be transferred by means of phase reversal because low-level RF was needed of SJA1000 [2].

Inverters 74HCT14 with Schmidt Flip-flop used by this circuit design was not only to ensure reset battery correct, but also to reform preferably for wave shape. Crystal Oscillator of CPU and controller SJA1000 for which relevant control lines and 8 bits data lines on read-write could be connected directly was mutual independence which required of setting clock frequency division register files appropriately in writing software. Meanwhile, the new CTM series isolated CAN

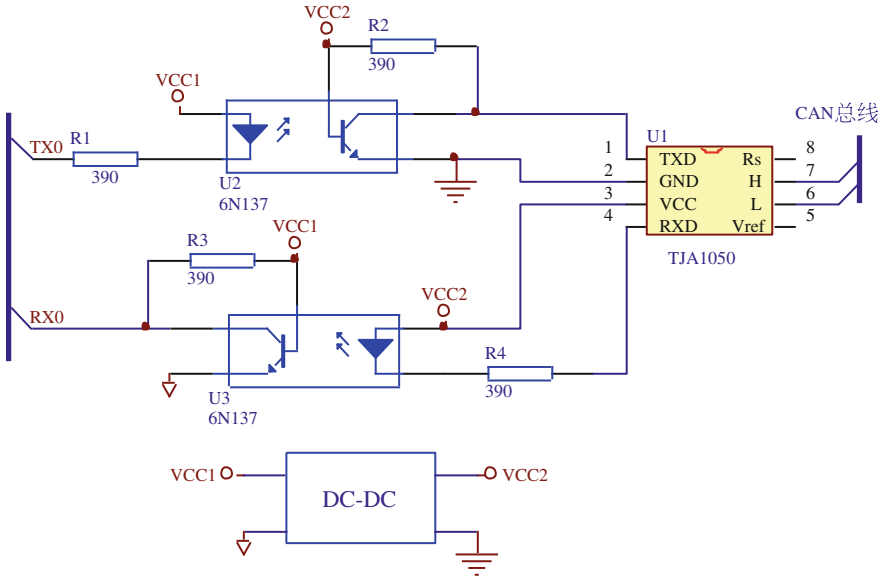


Fig. 200.3 Principle of isolation CAN transceiver circuit

transceiver module was put to use to ensure that controller could run normally when the CAN bus suffered from a severe interference and specific design could be found from transceiver isolating circuit design.

(2) Transceiver Isolating Circuit Design

New CTM series modules which TXD and RXD pins were compatible with +3.3 and +5 V CAN controller with no use of other external components but directly connected sending and receiving pins of +3.3 or +5 V CAN controller and CTM module that were adopted in the transceiver isolating circuit designed, integrated power supply module isolation, electrical isolation, CAN transceiver, CAN bus protection.

The circuit principle diagram of SJA1000 connection of CTM1050 and PHILIPS was shown in Fig. 200.3. From Fig. 200.3, it was found that isolation CAN transceiver module was employed which realized each node electrical and power supply of CAN-bus isolation and independent completely that improved the stability and security of the node.

The "*T*" suffix of CTM series isolated CAN transceiver which have eight model of all means double TVS bus protection components integrated in its inside that can avoid bus error or fault caused by surge and interference as much as possible. Its main advantages are lower cost than discrete components, simplified design, stable and reliable.

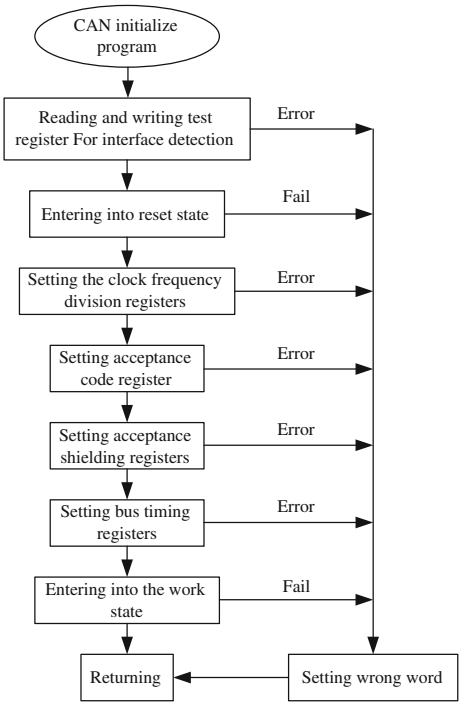


Fig. 200.4 Initialization flow chart of system

200.2.3 Software Design

(1) Introductions of Communication Protocol in System

As a master–slave communication system, it was made up with the ground monitoring center PC and underground CAN node device. The operation state of extension was examined by master control machine by turns unless seriously malfunctions appeared in surveillance area to break out "inspection" arrangement, and to send an emergency message to master control machine. Therefore, specific agreement should be established to obtain parameters the information correctly. Pelican mode and standard frame format, was adopted in this design, and fifth byte of the information frame format represented parameter type such as the network voltage type, three phase, leakage resistance, insulation resistance, power supply switch turntable and so on which would be send from the sixth to eleventh bytes showed the practical parameters information in order to ensure the reliability of data transmission not hesitate to that two bytes content could be said. The chance of data interference could be further reduced according to the principle of "three two out" while master control machine received

(2) Design of Communication Software

The reasonable design of software including such function module as system initialization, field data acquisition and processing, data shows refresh and CAN

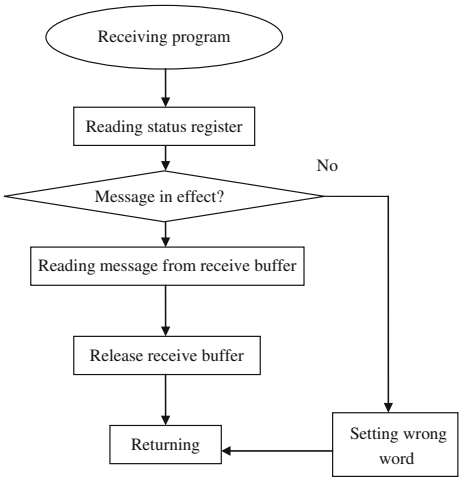


Fig. 200.5 Receiving program flow chart of message

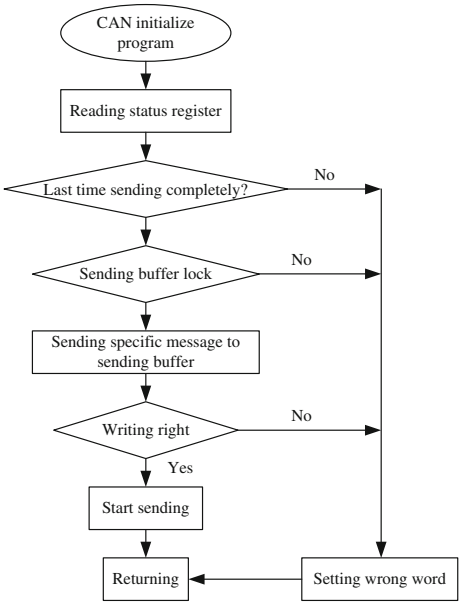


Fig. 200.6 Sending program flow chart of data

information processing was the key of completing the work on data collection and communication [3]. The test to device before CPU got access to the CAN controller is particular important in CAN processing module by means of writing fixed data such as 0AAH or #55H and then reading and checking the correct or not to test register which could discover timely interface fault of SJA1000 [4]. Programming ideas of redundant detection, step judgment, error warning were used to guarantee for the operation of the CPU to SJA1000 controller reliably and

correctly during control mouth initialization of SJA1000 and the process of message sending or receiving which effectively improved stability and healthy of software [5].

Figures 200.4, 200.5, 200.6 were the initialization flow chart of SJA controller, message sending and receiving flow chart respectively.

200.3 Conclusions

CAN bus with SJA1000 controller underground remote electric power communication system was designed in this paper, absorbed the successful application of CAN bus in power, communication and other industry, combined with underground bad working environment in coal mine. The traditional CAN contact drive was reformed on hardware circuit system. Meanwhile, new programming ideas of reliability testing and healthy were raised and carried out successfully in software design. After nearly a year of development, the system acquired a fine result when tried out in units such as no.1 tai hang coal mine and so on and had a extremely high application value.

References

1. Ding E, Zhang L, Meng C, Meng X, Huang H (2011) CAN repeater based on PIC18F258. Coal mine machinery, pp 136–138
2. Cai H, Zhong X (2011) Design of ventilation system based on CAN-BUS in coal mine. Coal technology, pp 37–39
3. Naji Ma (2011) Application of CAN bus in communication system about safety production of coal mine. Coal technology, pp 57–58
4. Sun Y, Li Y, Qian J-s, Ma X (2010) Research on coalmine real-time communication protocol based on CAN bus. Computer engineering, pp 236–238
5. Qian Z, Wu X, Jiang X, Cheng Y (2009) CAN bus communication network of high dependability and real-time for coal mine. Coal mine machinery, pp 203–205

Chapter 201
Transmission Parameters Extraction of Multi-Conductor Cable Based on Transmission Circuit Method

Hu Guoyou, Li Shunming, Xia Pinqi, Men Xiuhua and Guo Haidong

Abstract A new transmission circuit concept is put forward to replace the energy approach by varied equivalent current density using finite element method. This new concept is studied in detail for two-dimension model of multi-conductor and further research are also carried out for temperature-dependent characters. Next, the application of this concept is exploited to determine corresponding transmission parameters, such as impedance, resistance and inductance matrices considering coupling effect. Numerical results demonstrate the correctness of the new approach, and furthermore can provide with the temperature-dependent behaviors of these parameters.

Keywords Transmission parameters extraction · Multi-conductor cable · Transmission circuit method

H. Guoyou · L. Shunming · X. Pinqi · M. Xiuhua · G. Haidong (✉)
Nanjing University of Aeronautics and Astronautics, Nanjing, China
e-mail: haidong1113@126.com

H. Guoyou
e-mail: haidong1113@126.com

L. Shunming
e-mail: smli@nuaa.edu.cn

X. Pinqi
e-mail: haidong1113@126.com

M. Xiuhua
e-mail: haidong1113@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_201,

201.1 Introduction

The important factor of signal distortion is the conductor loss. The losses are associated with the electrical parameters, especially the frequency-dependent electrical signal. Electrical parameters associated with electromagnetic and the external environment [1]. These factors affect the conductor loss, and are also an important factor of signal distortion.

There are two coupling effects—the skin and proximity effect exists. This will cause uneven distribution of current; this representation on the conductor cross-section is non-uniform current density distribution. The non-uniform distribution of current density may leads to change in resistance and inductance parameters with the signal frequency [2, 3], skin effect is the parameters associated with conductor's radius and frequency-related. When in low frequencies, the skin depth δ is much larger than the radius r of the conductor, so the conductor current density redistribution is not obvious, can be simply solved by the Laplace's equation [4]. The skin effect can be ignored at low frequencies, but the coupling effect between conductors in the low frequency effect is significantly greater than the skin effect., this will lead to multi-conductor matrix which cannot be composed only by eigenvalues. Multi-conductor instead, by the same area radius [5], the electrical parameters must be considered [6].

For low frequencies, according to the characteristics that circuit model is easy to adapt to the current distribution, the transmission lines and electrical parameters are solved. At this point the following important factors must be considered: conductor parameters affected by temperature changes; the proximity effect of multi-conductor and the influence of temperature; complete transmission circuit model. Based on the above analysis, a transmission circuit model to obtain accurate electrical parameters and temperature change characteristics are established.

201.2 Transmission Circuit Method

Basic Conception of Signal Transmission. Suppose that there are N conductors in cable, and then all the transmission parameters will become the form of $N \times N$ matrices. Telegraph equation of multi-conductor is given by [7]

$$\begin{cases} \dfrac{\partial [v]}{\partial z} + \left([r] + [L] \dfrac{\partial}{\partial t} \right) [i] = 0 \\ \dfrac{\partial [i]}{\partial z} + \left([g] + [C] \dfrac{\partial}{\partial t} \right) [v] = 0. \end{cases} \tag{201.1}$$

When v and i are sine wave signal, Eq. 201.1 is transformed to the following harmonic equation

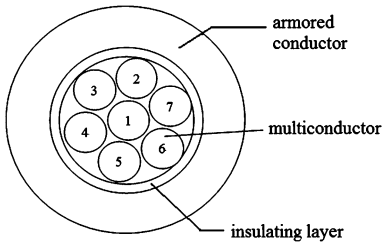


Fig. 201.1 Signal transmission using coaxial cable

$$\begin{cases} \dfrac{\mathrm{d}[V]}{\mathrm{d}z} = -([R] + j\omega[L])[I] \\ \dfrac{\mathrm{d}[I]}{\mathrm{d}z} = -([G] + j\omega[C])[V]. \end{cases} \tag{201.2}$$

In (201.2), V is an $N \times 1$ column vector formed by potentials V_i of each conductor cross section with $i = 1, 2, \cdots, N$. The I is also an $N \times 1$ column vector formed by the total current I_i through each conductor i. $[R]$ and $[L]$ are temperature-dependent matrices. Here, the differential impedance concept is exploited to determine the resistance and inductance matrices of multi-conductor extending along the z-axis. As a matter of fact, V and I take corresponding response with signal frequency and conductor temperature.

Another character of coaxial cable is returned current flows by external armored conductor, which may influence resistance and inductance especially loop values as shown in Fig. 201.1.

Determination of Resistance and Inductance Parameters. Use Laplace's equation $\nabla^2 V(z) = 0$ to get

$$J_S = \sigma E_S = \sigma(-\mathrm{grad}V) = \sigma\left(-\frac{\mathrm{d}V}{\mathrm{d}z}\right). \tag{201.3}$$

Combining (201.2) and (201.3), the current density J_S and different impedance can be expressed as (201.4).

$$[J_S] = -[\sigma]\frac{\mathrm{d}[V]}{\mathrm{d}z} = [\sigma][Z][I]. \tag{201.4}$$

We can have that

$$[R(T)] + \mathrm{j}\omega[L(T)] = \frac{[\rho(\mathrm{T})] \cdot [J_S(T)]}{[I]} \tag{201.5}$$

Note that resistance is a constant at certain temperature, and the most important parameter is equivalent current J_S. Its calculation is extremely effective using finite

element method when enough exact mesh is taken. In this case, resistance and inductance matrices can be calculated when three variants in right end of equation are given.

Assume $I_i = I, \quad I_j = 0$ $(j = 1, 2, \cdots N, j \neq i)$, We have the block matrices are formed by calculations as follows:

$$Z_{ij} = \frac{\rho \cdot J_{Sj}}{I_i} \rightarrow R_{ij} = \mathrm{Re}\left\{\frac{\rho \cdot J_{Sj}}{I_i}\right\};\ L_{ij} = \mathrm{Im}\left\{\frac{\rho \cdot J_{Sj}}{I_i}\right\}/\omega;\ i = 1, 2, \cdots N;\ j = 1, 2, \cdots N. \tag{201.6}$$

Where, the values of resistance and inductance matrices can now be obtained by solving (201.6) N times.

Effect from Return Path. Assume $\mathrm{I}i = \mathrm{I}, \mathrm{I}j = 0,\ IN + 1 = -\mathrm{I}$ $(i, j = 1, 2, \cdots N, j \neq \mathrm{i})$. The equivalent current density $J_{S,N+1}$ is also flowing along the z-direction of conductor $N + 1$ with corresponding resistivity $\rho N + 1$. Now that we can turn to the calculation of resistance and inductance matrices to new configuration as follows:

$$Z_{ij} = \frac{1}{I_i}\left(\rho \cdot J_{Sj} - \rho_{N+1} \cdot J_{S(N+1)}\right) \Rightarrow \begin{cases} R_{ij} = \mathrm{Re}\left\{\rho \cdot J_{Sj} - \rho_{N+1} \cdot J_{S(N+1)}\right\} \\ L_{ij} = \dfrac{\mathrm{Im}\left\{\rho \cdot \mathrm{J_{Sj}} - \rho_{\mathrm{N+1}} \cdot \mathrm{J_{S(N+1)}}\right\}}{\omega} \end{cases} (i = 1, 2, \cdots N; j = 1, 2, \cdots N). \tag{201.7}$$

We can obtain the much effective values by the calculation of Eq. 201.7. Considering all around temperature, it is advantageous to take some discreted solutions for temperature in order to obtain the changes of parameter matrices with respect to temperature ρ or frequency function J_S.

201.3 Numerical Calculation and Analysis

Character of Current Density of Temperature Change. Consider the frequencies at 45, 500 and 5000 Hz, the sine signal used to conductor is 1 A and all the values are obtained using the new developed approach above obtained. Current density J_S of conduct 1 in z-direction at different temperature is shown in Table 201.1.

In Table 201.1, two reverse effects appear due to the increase in temperature and frequency respectively. In the same column at certain frequency, real part and imaginary part of current density decrease with temperature rising, and can be shown approximately as linear functions. Moreover, the results are also shown the increase in the same row with increasing frequency at the same temperature. The numerical

Table 201.1 Current density at different temperature and frequency

$J_s(A/m^2)$ (°C)	$f = 45$ (Hz)	$f = 500$ (Hz)	$f = 5000$ (Hz)
0	2720 + *j*7015	8654 + *j*58845	30504 + *j*527684
50	2473 + *j*5980	7843 + *j*49144	27187 + *j*436976
100	2279 + *j*5231	7226 + *j*42307	24719 + *j*373297
200	1998 + *j*4199	6814 + *j*33283	21253 + *j*289708
300	1812 + *j*3511	5712 + *j*27575	18900 + *j*237216
400	1686 + *j*3015	5247 + *j*23625	17177 + *j*201148

calculating results clearly demonstrate that the real parts and imaginary parts have different change rate with the changes of frequency and temperature. Along with the frequency increasing, the values of imaginary parts increase and exceed one of real parts. However, the imaginary parts also take more decrease than real part for increasing temperature. The electromagnetic induction may be reinforced or weakened along with the increasing or reducing of frequency and temperature.

Conductor Loss. Before we obtaining multi-conductor loss, we first bring (13b) as fundamental calculating $P_d = \frac{1}{2}\iint \rho |J_S(T)|^2 \mathrm{d}S$.

Next consider the seven conductors configuration depicted in Fig. 201.1. All cross sections of conductors are distributed with self and mutual current density. This leads to overall summation of conductor losses, which is more important for parameters calculation and signal transmission. Now we can get certain part of conductor loss as (201.8) when there is sine signal in conductor 1.

$$P_{d,1} = \sum_{i=1}^{7} \frac{\rho(T)}{2} \iint \left|J_{S,1i}(T)\right|^2 \mathrm{d}S. \tag{201.8}$$

Consequently, total loss containing seven conductors can be calculated by (201.8)

$$P_d = \sum_{j=1}^{7}\sum_{i=1}^{7} \frac{\rho(T)}{2} \iint \left|J_{S,ji}(T)\right|^2 \mathrm{d}S. \tag{201.9}$$

Now we have carefully verified that $\rho(T)$ and $|J_S(T)|$ are two important influencing factors for conductor loss. Consider the calculated results in Table 201.1, the same variety condition of conductor loss can be calculated by (201.8) and (201.9).

Parameters Character of Temperature Change. In this term, we consider the detailed behaviors of resistance and inductance for different all around temperature. All the values are obtained by new method developed in the paper and (201.7) that is applied to compute parameters. As is evident form on the above analysis, frequency domain is more suitable to describe the detailed character. In this case, resistance $R11$ and inductance $L11$ results have been plotted in Figs. 201.2 and 201.3 respectively. The frequency range is limited in 0–10 kHz.

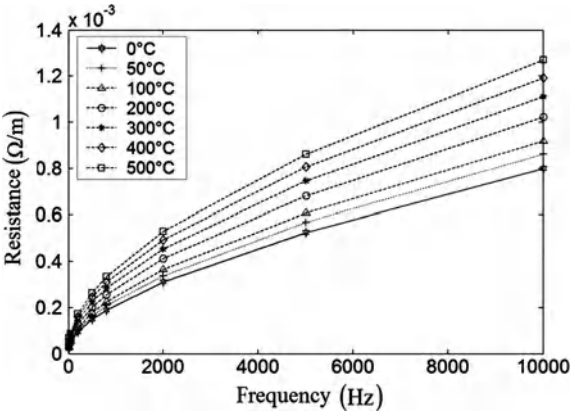


Fig. 201.2 Temperature behavior of self-resistance R_{11}

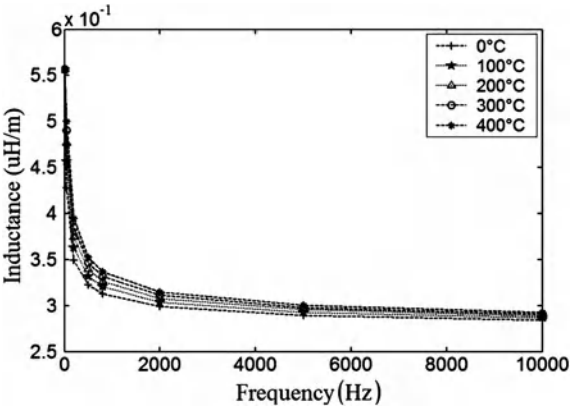


Fig. 201.3 Temperature behavior of self-inductance L_{11}

It can be seen in Fig. 201.2 that higher values of resistance are clearly to be found along with the increases of temperature and frequency. This behavior cannot only be verified using current density character in Table 201.1, and as a matter of fact, the character of resistance is modified thoroughly due to adding another important factor $\rho(T)$. When electric resistivity $\rho(T)$ is limited in 0–500 °C range, we can easily obtain its change magnitude between 1 and 2.15 compared with $\rho(0)$. The total value of $\rho \cdot J$ will be the increment function for temperature. Furthermore, when we continue to increase the frequency, the D-value between two temperatures becomes larger like the plotted results in Fig. 201.2.

From Fig. 201.3, one can also find complete opposite tendency for inductance with increase of frequency, and there is much decrease of values with the increasing frequency. However, for increasing temperature, inductance is also intended to be larger than lower temperature. The upper inductance curve is represented at 400 °C, and sequent to base curve is values at 0 °C. In different frequency range, inductance values between two temperature at $f < 2$ kHz are much larger than D-value at $f > 2$ kHz.

As remarked first, we can confirm that both the values of inductance and resistance are not constant but increase with all around temperature rising. When

applying these varied values to transmission circuit equations, it is necessary to investigate the corresponding change of signal magnitude and phase at the receiving end. Furthermore, this analysis is more important in the coupling of different frequency and temperature fields.

201.4 Conclusions

In the paper, we have presented the important effect caused by temperature changing for electromagnetic analysis and parameters of multi-conductor, and also analyzed the steady and effective parameters characters for signal transmission of coaxial cable model.

For different all around temperature, the change in conductor resistivity has resulted in a corresponding change of equivalent current density, electric field and even conductor loss, which is important for calculating resistance and inductance values compared with classical transmission equations. Even in low-frequency range, current density redistribute on multi-conductor cross section due to vicinity effect with smaller distance between each conductor.

On this basis, a new approach has developed to calculate the resistance and inductance using varied current density where we can effectively avoid complex integral equation. As remarked, all resistance and inductance matrices are represented as $N \times N$ matrices which can be applied to transmission equation to get signal response.

Furthermore, it has been demonstrated that all the values of resistance and inductance are not constant but increase with changing temperature and these various behaviors can be applied in further research to improve the steady analysis of signal transmission.

References

1. Coperich KM, Morsey J, Okhmatovski VI, Cangellaris AC, Ruehli AE (2001) In: IEEE Transactions on microwave theory and techniques, vol 49, pp 1677–1685
2. Tsuk MJ, Kong JA (1991) In: IEEE Transactions on microwave theory and techniques, vol 39, pp 1338–1347
3. Konrad A (1982) In: IEEE Transactions on magnetics vol 18, pp 284–292
4. Zutter DD (2005) In: IEEE Transactions on microwave theory and techniques, vol 53, pp 2526–2538
5. Shishan W, Shengchang J, Yanming L (2003) Numerical analysis of eddy current losses for power transformer windings with round stranded conductors. High Voltage Eng 29:7–10
6. Xuefeng D (2006) Transmission investigation on long cable. Doctoral Dissertation, University of Science and Technology of China. (In Chinese)
7. Zhixun H, Xiaojin W (1996) Theory and operative technology of microwave transmission line. Onomic Science Press, Beijing pp 40–45. (In Chinese)

Chapter 202
Research on the Line Fault Branch Location Method of 10 kV Village Distributed Network

Shan Su and Fuzhong Wang

Abstract According to the characteristics of village distributed network, such as the main line power supply, the multiple branches, the tree distribution network is equivalent to multiple-input multiple-output network, and based on the equivalent, a multi-terminal mathematical model is established. Fault branch is determined through the measurment of branch resistance changes from measurable endpoint voltage changes, and the algorithm is implemented. The algorithm is verified by simulation of the Matlab software. The results show that this algorithm can quickly and accurately locate the fault branches, improve operational reliability and reduce power losses.

Keywords Distributed network · Fault-port · Fault branch identification

202.1 Introduction

The village distributed network uses supply network in a tree, low power corridor, many branches, wide coverage, harsh operating environment, vulnerability to natural disasters and man-made damage, and high chance of failure [1, 2]. When a branch fails, the fault branch is found by switching the method one by one on branch line. A long time power failure likely causes power outages throughout the trunk, and severely affects the industrial and agricultural production and people's

S. Su (✉) · F. Wang
College of Electrical Engineer and Automation, Henan polytechnic University, Jiaozuo 454003, China
e-mail: sushan@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_202,

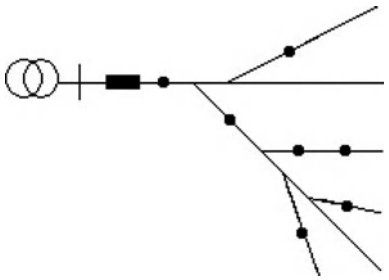


Fig. 202.1 The wiring mode of village distributed network

lives [3]. Therefore, the exact fault location can ensure that in rural branch network security, stability and economic operation, there is enormous social and economic benefits.

The process of the village distributed power network fault location first finds the fault branch, and then finds the point of failure. Therefore, the determination of the fault branch is the key to addressing 10 kV rural network failures. Combining with tree distribution network multi-port equivalent model, this chapter proposes a fault diagnosis method to identify fault branch.

202.2 Multi-port Equivalent Model of 10 kV Village Power Network

Figure 202.1 shows the wiring mode of village distributed network. Tree distribution network with multiple branches can be regarded as a multiple input and multiple output ports network [4]. Figure 202.2 is multiple-port equivalent model with p branches tree distribution network. 1 to l port information can be measured, the said port is measurable (i.e., the end of the branch line). $l + 1$ to p port information is immeasurable (i.e., intersection). Take the port with the port number under the same standard, the reference direction is shown in Fig. 202.2.

Port equation as:

$$\begin{bmatrix} \dot{I}_1 \\ \dot{I}_2 \\ \vdots \\ \dot{I}_p \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1p} \\ Y_{21} & Y_{22} & \cdots & Y_{2p} \\ \vdots & \cdots & \cdots & \vdots \\ Y_{p1} & \cdots & \cdots & Y_{pp} \end{bmatrix} \begin{bmatrix} \dot{U}_1 \\ \dot{U}_2 \\ \vdots \\ \dot{U}_p \end{bmatrix}$$

$$I_p = Y_p U_p \tag{202.1}$$

where YP is port admittance matrix, IP and UP respectively is the port current and voltage column vector.

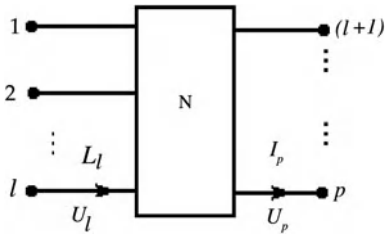


Fig. 202.2 The multiple-port equivalent model with tree distribution network

202.3 Diagnosis of Fault Branches

202.3.1 Principle of Fault Port Diagnosis

On the equivalent model shown in Fig. 202.2, when the branches in the network have fault, in the same excitation conditions, the voltage and current of the fault branches and port will deviate from the normal network of the corresponding value [5]. In the following derivation, the fault network parameters are expressed with corresponding normal network parameters in the form of an incremental increase.

The normal and fault measurable ports are respectively excited at the same current source vector $\dot{I}_f$. The respective port equation can be obtained:

$$Y_p U_p = \begin{bmatrix} \dot{I}_f \\ 0 \end{bmatrix} \tag{202.2}$$

$$(Y_p + \Delta Y_p)(U_p + \Delta U_p) = \begin{bmatrix} \dot{I}_f \\ 0 \end{bmatrix} \tag{202.3}$$

Equation (202.2) was subtracted by Eq. (202.3):

$$Y_p \Delta \dot{U}_p = -\Delta Y_p (U_p + \Delta U_p) \underline{\underline{\Delta}} J_p \tag{202.4}$$

Defining Eq. (202.4) for the port fault current source vector J_p. As can be seen from Eq. (202.4), if the short-circuit transfer admittance increment of one port is not zero, i.e. $\Delta Y_{ij} \neq 0$, the fault current source J_p of the port is not zero. $\Delta Y_{ij} \neq 0$ shows the network failures, so ΔY_{ij} can reflect whether the network is failed. In this meaning, Eq. (202.4) can be called a fault diagnosis equation. The port with the fault branches is the failed port, otherwise trouble-free port. Because J_i is defined as the fault current source port, it can only reflect whether the port is failure.

Trouble-free port corresponds with $\Delta Y_{ij} = 0$, so $J_i = 0$. Therefore, according to the information of measurable port, first locate the non-zero elements J_p, identifying the fault port, and then finding fault branches. From Eq. (202.4) can be obtained:

$$\Delta \dot{U}_p = \begin{bmatrix} \Delta \dot{U}_{lp} \\ \Delta \dot{U}_{hp} \end{bmatrix} = Y_p^{-1} J_p, Y_p^{-1} = \begin{bmatrix} Z_{lp} \\ Z_{hp} \end{bmatrix} \tag{202.5}$$

where $Z_{lp} \in c^{l \times p}$, is the measurable port impedance matrix, $Z_{hp} \in c^{h \times p}$ is the immeasurable port impedance matrix, Z_p is the port impedance matrix of$p$ port network. Then:

$$\Delta \dot{U}_p = Z_p J_p \tag{202.6}$$

and:

$$\Delta \dot{U}_l = Z_{lp} J_p \tag{202.7}$$

Equation (202.7) indicates that J_p can be obtained by the measurable port information $\Delta \dot{U}_l$ and fault network parameters z_{lp}. Thus, Eq. (202.7) also is known as the fault port diagnosis equation.

202.3.2 Algorithm of Fault Port Diagnosis

The basic idea of port diagnostics method is to first find out the fault port, and then find fault branches, so we must first solve the problem of the fault port diagnosis. Once the fault port is identified, fault branches also are easily identified.

Equation solvability conditions are met, due to the characteristics of $p \succ l$, so the answer of J_p will not be unique. The diagnosis of the problem, is the only answer to indicate which port is faulty. In fact, the fault branch number of network is small and the number associated with the fault port is also small. Therefore, only a few components are not zero in answers of $J_p'; cu$. To find this out only J_p, the following conditions should be achieved.

Condition 1 $\quad \operatorname{rank}\left[Z_{lpi}\right] = k+1 \quad \forall 1 \le l_1 \le l_2 \le \cdots \le l_{k+1} \le p$

Condition 2 $\quad \operatorname{rank}\left[Z_{lpi} \vdots \Delta V_l\right] = k+1 \quad \forall 1 \le l_1 \le l_2 \le \cdots \le l_k \le p$

Condition 3 $\quad \exists 1 \le l_1 \le l_2 \le \cdots \le l_k \le p, \quad \ni \operatorname{rank}\left[Z_{lpi} \vdots \Delta V_l\right] = k$

Thus, Z_{lpi} corresponds with k fault port in condition 202.3. Figure 202.3 shows the diagnostic diagram, and is consistent with the diagnostic conditions.

After the fault port was diagnosed, the Eq. (202.7) can be further written as:

$$Z_{lpi} J_p = \Delta V_l \tag{202.8}$$

where Z_{lpi} is the port impedance matrix corresponding with fault port, J_p is the fault port current source vector of non-zero elements. Using the generalized inverse, the following can be uniquely solved:

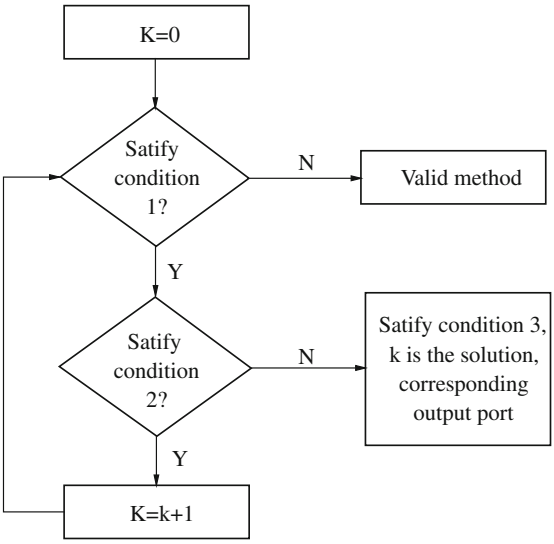


Fig. 202.3 The diagnostic diagram of fault port

$$J_p = (Z_{lpi}^T Z_{lpi})^{-1} Z_{lpi}^T \Delta V_l \tag{202.9}$$

And then, the fault branches can be identified.

202.4 Algorithm Simulation

202.4.1 Fault Network Model

10 kV village distributed network can be equivalent to three branch structure. Figure 202.4 shows the basic parameters of the main electrical components; fault occurs at a distance of 3 km at the D end. O is the supply-side endpoint, the endpoint C, E, F are measurable for the end user, B, D are the branch immeasurable endpoints. Simulation model is established by MATLAB.

202.4.2 Fault Addressing Algorithm Simulation of Village Distributed Network

Figure 202.5 shows the nodes and branch number of Fig. 202.4 village network.

According to distribution network the parameters are:

$$z_0 = 0.58 + j1.59,\ z_1 = z_2 = 0.43 + j0.36$$

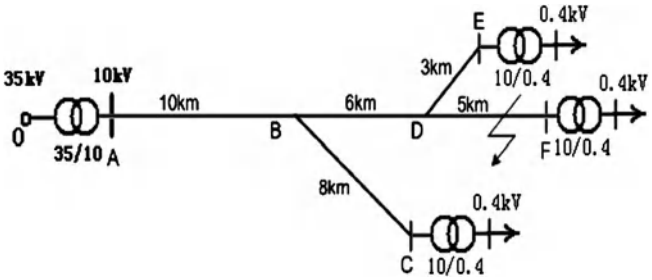
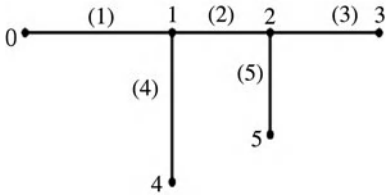


Fig. 202.4 Three branch 10 kV village distributed network

Fig. 202.5 Nodes and branch number of village network



According to the physical meaning of the node impedance parameters, when a unit current injects into node i and the remaining nodes are open circuit, self-impedance value of node i is equal to the voltage of node i. When a unit current injects into node j and the remaining nodes are open circuit, mutual impedance between node i and node j is numerically equal to the voltage of node i.

The node impedance matrix before the fault is shown in Table 202.1.

Nodes 3, 4, 5 is measurable, nodes 1, 2 is intersection and immeasurable. Table 202.2 can be obtained by fault port diagnosis principle.

Fault is metallic single-phase ground fault ($R_f = 0$). Table 202.3 shows the captured voltage of measurable point in a time before fault and fault transience. Due to symmetrical three-phase, only the value of phase A is given here.

The supply voltage is $10.5 \angle 0^\circ$ kV before and after fault. Normal network and fault network is excited at the same current source $[1 \quad 0 \quad 0]^T$. According to the operation results of diagnostic procedures port, when $k = 2$, only:

$$\text{rank}[Z_{l1} Z_{l2} \vdots \Delta V_l] = 2$$

Other combinations of Z_{lpi} are :

$$\text{rank}[Z_{lpi} \vdots \Delta V_l] = 3$$

Therefore, port 2, 3 are the fault ports, corresponding to the l_{23} and l_{DF} which are the fault branches.

Simulation results show that the algorithm is feasible.

Table 202.1 The node impedance matrix of three branches village network

Z_{ij}	1	2	3	4	5
1	5.61 ∠ 39.94°	5.61 ∠ 39.94°	5.61 ∠ 39.94°	5.61∠ 39.94°	5.61∠ 39.94°
2	5.61∠ 39.94°	8.97 ∠ 81.73°	8.97 ∠ 81.73°	5.61 ∠ 39.94°	8.97 ∠ 81.73°
3	5.61 ∠ 39.94°	8.97 ∠ 81.73°	12.05 ∠ 38.87°	5.61 ∠ 39.94°	8.97 ∠ 81.73°
4	5.61 ∠ 39.94°	5.61 ∠ 39.94°	8.97 ∠ 81.73°	10.09 ∠ 39.94°	5.61 ∠ 39.94°
5	5.61 ∠ 39.94°	8.97 ∠ 81.73°	8.97∠ 81.73°	5.61 ∠ 39.94°	10.66∠ 39.94°

Table 202.2 The measurable point impedance matrix of three branches village network

Z_{lp}	1	2	3	4	5
3	5.61 ∠ 39.94°	8.97 ∠ 81.73°	12.05 ∠ 38.87°	5.61 ∠ 39.94°	8.97 ∠ 81.73°
4	5.61 ∠ 39.94°	5.61 ∠ 39.94°	8.97 ∠ 81.73°	10.09 ∠ 39.94°	5.61 ∠ 39.94°
5	5.61 ∠ 39.94°	8.97 ∠ 81.73°	8.97 ∠ 81.73°	5.61 ∠ 39.94°	10.66 ∠ 39.94°

Table 202.3 The voltage of measurable point in a time before fault and fault transient

Type	Voltage of node 3 (kV)	Voltage of node 4 (kV)	Voltage of node 5 (kV)
Pre-fault	5.782 ∠ 2.43°	5.739 ∠ 3.28°	5.764 ∠ 3.03°
Failure	0	2.778 ∠ 5.39°	1.374 ∠ 14.08°
Failure before and after the voltage difference	5.782 ∠ 2.43°	2.952 ∠ 1.30°	4.391 ∠ −0.51°

202.5 Conclusion

The algorithm has been well applied in fault addressing of Wuzhi power network. To occur during the running of several single-phase ground faults, fault branch can be quickly and accurately located, and the power supply reliability is better protected.

References

1. Johns A, Whittard T (1999) New technique for the accurate location of earth faults on transmission system. IEEE Proc Gener Transm Distrib 142(2):73–79
2. Zhang G, Yang M (2005) A virtual value method of locating single-phase grounding fault section of 10 kV feeders in power distribution system. Relay 33(8):34–37
3. Liu J, Zhu Z, Chen G (1999) Branched location method for diagnosing short circuit malfunction of electric net in oilfield. J Daqing Petroleum Inst 23(4):84–86
4. Li M, Wang Y, Wang H (2001) A new approach on detecting the single-to ground fault location on power system with neutral unearthed. Proc CSEE 21(10):6–10
5. Gao R, Wang F, Yu Q (2003) Application of wavelet transform singularity analysis to fault detecting on power distribution lines. Electric Power Sci Eng 4:5–8

Chapter 203
Optimal Allocation of Distributed Generators in a Distribution Network Using Adaptive Multi-Objective Particle Swarm Optimization

Shan Cheng, Min-You Chen, Peter J. Fleming and Xia Li

Abstract This study presents the optimal allocation of distributed generators (DGs) in distribution network based on an adaptive multi-objective particle swarm optimization (AMPSO). In order to enable the DGs to make the best contribution to the distribution network, a multi-objective optimization model addressing power loss and a voltage stability index (VSI) is constructed, in which the capacity factor of DGs is taken into consideration. In the AMPSO, some adaptive techniques including the use of dynamically changing inertia weight, time-varying acceleration coefficients, mutation operation and circular crowding sorting have been integrated to improve the convergence and the diversity performance of the algorithm. The proposed algorithm has been used to determine the optimal allocation of DGs in a IEEE 33-bus distribution network. Simulation results show that the proposed method can identify the optimal solutions with a good tradeoff between the power loss and the VSI.

Keywords Capacity factor · Distributed generation · Multi-objective optimization · Optimal allocation · Power loss · Voltage stability

S. Cheng (✉) · M.-Y. Chen · X. Li
State Key Laboratory of Power Transmission Equipment & System Security and New Technology, School of Electrical Engineering, Chongqing University, Chongqing, China
e-mail: sscheng51@hotmail.com

M.-Y. Chen
e-mail: minyouchen@cqu.edu.cn

X. Li
e-mail: zixingyu1986@126.com

P. J. Fleming
Department of Automatic Control & Systems Engineering, University of Sheffield, Sheffield, UK
e-mail: p.fleming@sheffield.ac.uk

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_203,

203.1 Introduction

Distributed generator is regarded as a natural candidate to be interconnected with the power grid. To enable the DGs for making the best contribution, the key is optimal allocation of the DGs in the distribution network.

During the development of the optimal allocation researchers have paid great attention to the problem described by a single objective and mostly determined either the optimal location or optimal capacity of a given DG. Besides, there are considerable researches in which a multi-objective optimization problem is converted into a single-objective problem. However, single-objective optimization cannot accurately reflect the relationship between the various objectives, and the objectives to be optimized are normally in conflict with respect to each other, which means, the solution to a multi-objective optimization problem is a set of different solutions (so-called Pareto optimal set) instead of a single solution. In fact, multi-objective optimization algorithms can provide flexibility with a variety of diverse choices and such algorithms have been applied to the optimal allocation of DG in recent years. Another aspect that must be addressed is that most researches treated the power output of DG as constant, ignoring the fact that electric power generation systems rarely operate at its rated output, especially those who depend on the climate. Therefore, the concept of capacity factor, which represents the contribution of the DG to the generation adequacy of a power system, has been introduced to estimate the average energy.

Based on these considerations, this study proposes a multi-objective optimization algorithm for determination of optimal DG allocation to minimize the total active power loss and maximize the system's stability. Novel features of the study include: (1) a multi-objective optimization model is constructed, which not only considers economic benefit but also technical aspects and the concept of a capacity factor is introduced into the model; (2) the proposed multi-objective problem is solved by an adaptive multi-objective particle swarm optimization (AMPSO) algorithm, which adopts adaptive techniques including dynamically changing inertia weight, time-varying acceleration coefficients, mutation operation and circular crowding sorting technique; (3) the study proposes optimization of both location and capacity of dispersed DGs rather than either location or capacity.

203.2 Multi-Objective Optimal DG Allocation

The following assumptions are adopted to construct the multi-objective optimization model: (1) the effect of DG is considered as negative real power load at the buses (Bus 1 is assumed to be a slack bus and no DG is installed at this bus.); (2) the power factor (set to be unity) and capacity factor are the same for each DG; (3) the capacity of DGs must be selected from the given capacity candidates (discrete values).

203.2.1 Capacity Factor

Despite the various merits of DG, some types of DG, such as wind power and solar energy would deliver less energy than the rated capacity. To assess the power generation efficiency and ensure the provision of higher quality and more reliable power, a capacity factor [1], which is the ratio of the average power of DG to its rated capacity has been introduced to estimate the contribution of the DG to the generation adequacy of a power system. The capacity factor for each DG can be obtained via a dynamic mathematical model under different weather conditions.

203.2.2 Objective Function

Power loss is a key consideration in the problem of the allocation of DGs. Although it cannot be completely removed, it can be reduced to an acceptable value. The active power loss of the network can be expressed as

$$\min f_{P\text{loss}} = \sum_{k=1}^{N_{\text{bra}}} G_{k(i,j)}[U_i^2 + U_j^2 - 2U_iU_j\cos(\theta_i - \theta_j)] \tag{203.1}$$

where N_{bra} is the total number of branches. $G_{k(i,j)}$ is the conductance of branch k which connects bus i and bus j. U_i and U_j are voltage magnitudes of bus i and bus j, respectively. θ_i and θ_j are voltage angles of bus i and bus j, respectively.

The modern distribution network experiences distinct load changes daily, and under certain critical conditions it suffers from voltage collapse. Thus, considering system security and reliability, minimizing the voltage stability index (VSI) [2] of the system is chosen to be a second objective. A modified VSI is adopted and the VSI of branch k can be expressed as

$$U_{\text{sta},k} = 4[(X_{ij}P_j - R_{ij}Q_j)^2 + (X_{ij}Q_j + R_{ij}P_j)U_i^2] \Big/ U_i^4 \tag{203.2}$$

where R_{ij} and X_{ij} are resistance and reactance of branch k, respectively. P_j and Q_j are total active and reactive power injected to the receiving bus j of branch k, respectively.

Note that $U_{\text{sta},k}$ should be no more than 1.0 and the VSI value of the branch approaches 1.0 is more sensitive to collapse. Thus, the second objective function can be described as

$$\min f_{U_{\text{sta}}} = \max\{U_{\text{sta},1}, U_{\text{sta},2}\ldots U_{\text{sta},N_{\text{bra}}}\} \tag{203.3}$$

203.2.3 Constraints

The multi-objective functions are minimized subject to various operational constraints to satisfy electrical requirements. These constraints are discussed below.

Power balance constraints: The power balance constraints with consideration of DG and its capacity factor C_f can be formulated as follows.

$$\begin{cases} P_i + P_{DGi}C_f - P_{Li} - U_i \sum_{j=1}^{N_{\text{bus}}} U_j(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) = 0 \\ Q_i - Q_{Li} - U_i \sum_{j=1}^{N_{\text{bus}}} U_j(G_{ij}\sin\theta_{ij} + B_{ij}\cos\theta_{ij}) = 0 \end{cases} \tag{203.4}$$

where N_{bus} is the total number of buses in the system. P_i, P_{DGi} and P_{Li} are active power, rated active power of installed DGs and active power load at bus i, respectively. Q_i, and Q_{Li} represent reactive power and reactive power load at bus i, respectively. G_{ij} and B_{ij} represent the conductance and susceptance of the branch from bus i to j. $\theta_{ij} = \theta_i - \theta_j$.

DG penetration capacity constraints: The limit on the total power generated by DG subject to a penetration level:

$$\sum_{i=2}^{N_{\text{bus}}} P_{DGi} \le \eta P_{\text{load}} \tag{203.5}$$

where P_{load} is the total active power load in the system and η represents the penetration rate.

DG-unit size constraints: Based on the assumption (203.2) mentioned above, the DG capacity can be expressed by active power, and each capacity should be no more than the permitted maximum active power.

$$P_{DGi} \le P_{DG\max} \tag{203.6}$$

where $P_{DG\max}$ is the maximum active power of DG-unit size that can be installed at bus i.

The other inequality constraints include voltage operational tolerance limits and feeder transmission capacity constraints.

203.2.4 Variables

The states above clearly show that the state variables include the voltage as well as active and reactive power at each bus. The decision variables, as stated in Section I, include the size and location of the installed DGs which can be represented as $[P_{DG2}, P_{DG3}, \ldots P_{DGN\text{bus}}]^T$. If $P_{DGi} = 0$, $i = 2,3\ldots N_{\text{bus}}$, it means no DG is installed at bus i.

203.3 Adaptive Multi-Objective Swarm Optimization Algorithm

Particle swarm optimization (PSO) is an evolutionary computation technique, in real number space each particle i is associated with its velocity $v_i = [v_{i1}, v_{i2}\ldots v_{iD}]^T$ and position $x_i = [x_{i1}, x_{i2}, \ldots, x_{iD}]^T$, where D stands for the dimensions of the solution space. The best position ever found so far by particle i is recorded as $p_i = [p_{i1}, p_{i2}\ldots p_{iD}]^T$, whose corresponding fitness value is called the particle's best p_{best}. Moreover, the best position found by any particle is recorded as $p_g = [p_{g1}, p_{g2}\ldots p_{gD}]^T$ and its fitness value is called the global best g_{best}. During this evolutionary process, the velocity and position update formula of particle i on the dimension $d|_{d=1,2,\ldots D}$ are described as

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1 * [p_{id} - x_{id}(t)] + c_2 r_2 * [p_{gd} - x_{id}(t)] \tag{203.7}$$

$$x_{id}(t+1) = x_{id}(t) + x_{id}(t+1) \tag{203.8}$$

where t is the current iteration, w is the inertia weight, c_1 and c_2 are positive acceleration coefficients, r_1 and r_2 are random numbers with uniform distribution in the range of [0, 1].

203.3.1 Adaptive Techniques

To improve the performance of PSO for multi-objective optimization, the study developed an adaptive multi-objective particle swarm optimization (AMPSO) algorithm in which the inertia weight can adjust adaptively according to (203.9).

$$w(t) = w_0 + \text{rand} * (1 - w_0) \tag{203.9}$$

where rand is the random operator, and w_0 is a constant value in the range [0,0.5]. Eq. (203.9) ensures that the inertia weight changes in the range of $[w_0, 1]$.

Furthermore, to enhance the global search ability during the run, time-varying acceleration coefficients [3] expressed in (203.10) and (203.11) have been integrated in the AMPSO algorithm.

$$c_1 = 2.5 - t/M_t \tag{203.10}$$

$$c_2 = 0.5 + t/M_t \tag{203.11}$$

In addition, to strengthen the convergence property and improve the diversity performance, a circular crowding sorting (CCS) technique [4] is integrated to select the particles from one iteration to another. CCS seeks to improve upon the non-dominated sorting technique of NSGA-II [5] in terms of diversity preservation. First, the crowding distance of all solutions is computed, and the solution with least the crowding distance is eliminated. Then the crowding distance of the remaining solutions is computed and the solution with the least crowding distance

is also eliminated. This operator is cycled until the required number of non-dominated solutions remain.

Finally, to prevent the premature convergence, a mutation operation has also been used in the proposed AMPSO algorithm.

203.3.2 Algorithm Procedure

The procedure of the AMPSO algorithm can be summarized as follows:

Step 1. Initialization. Set the population number N and iteration number M_t. Initialize the position x_i and velocity v_i of the particles. Set $p_i = x_i$.
Step 2. Evaluation. Compute the values of the objective functions and the associated fitness of the particle according to a weighted aggregation function [6] and update p_i and p_g.
Step 3. Generation of new particles. Generate the new particles with the velocity v_i_new and position x_i_new based on the current position x_i (i = 1, 2,…, N), and compute the values of the objective functions of the newly generated particles. Combine x_i_new and x_i together and store them in templist.
Step 4. Identification of non-dominated solutions. Identify the non-dominated solutions from templist and store them in a matrix ND_list, store the dominated solutions in D_list.
Step 5. Selection of particles for next iteration. a) If the number of particles in ND_list is larger than N, select N particles from ND_list according to CCS and store them as x_new. Otherwise, store templist as x_next and let N_{x_next} denote the number of particles in x_next. b) Identify the non-dominated solutions form D_list and store them as ND_temp and let $N_{ND\text{-}temp}$ represent the number of particles in ND_temp. c) If $N_{ND\text{-}temp} > N_{x\_next}$, select $(N\text{-}N_{x\text{-}next})$ particles from ND_temp according to the CCS technique. Else go to d). d) Add ND_temp to x_next, set $N_{x\text{-}next} = N_{x\text{-}next} + N_{D\_temp}$. Delete the non-dominated solutions from D_list. e) If $N_{x\text{-}next} = N$, stop and set x_next as the current position for next iteration, else, go to b).
Step 6. Mutation operation. If all $|v_i(t)| < 0.2V_{max}$, execute the mutation operation as follows, otherwise go to Step 7. (a) Randomly select 20% of the particles in the current population and store them in x_mut. (b) Evaluate x_mut and find the particles that dominate any particles in ND_list and replace the corresponding particles in the current position. (c) Repeat (a) and (b) k times (k = 1–10).
Step 8. Return to Step 2 until M_t is met.
Step 9. Store the non-dominated solutions from the final population.

203.4 Simulation Analysis of DG Optimal Allocation

The AMPSO algorithm is applied to determine the optimal sizes and sites of the installed DGs in a IEEE 33-bus system [7] as illustrated in Fig. 203.1 for minimizing the total power loss and the VSI of the system. The initial real power loss is

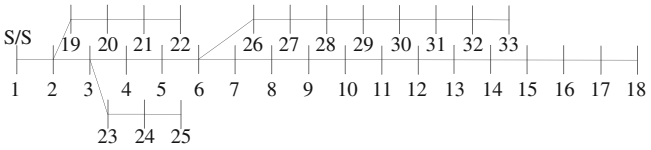


Fig. 203.1 Single line diagram of IEEE 33-bus system

201.53 kW, the VSI is 0.0996. The minimum and maximum voltage magnitudes for each bus are 0.95 and 1p.u., respectively. The maximum capacity of each installed DG and the maximum total capacity of the installed DGs are 1 MW and P_{load}, respectively. Candidates for DG installation include bus 6, 11, 15, 18, 22, 25 and 31. Wind-based DG with $Cf = 0.2209$ is considered.

Figure 203.2a illustrates the VSI values against the power loss values corresponding to the obtained Pareto solutions when Cf equals to 0.2209 and 1, respectively. From Fig. 203.2a, it can be seen that, among the non-dominated solutions, the VSI decreases with an increase in the power loss. It also shows that the objective values are quite different for the two cases. Note that, among the Pareto-optimal (tradeoff) solutions, there are solutions whose total capacity of installed DGs is the same but have different sizes and sites for the installed DGs. Accordingly, the objective values are markedly different. Five solutions are chosen from the Pareto-optimal solutions. Table 203.1 and Fig. 203.2b provide the installation information and corresponding voltage magnitudes, respectively. As shown in Fig. 203.2b, the bus where DG is installed, the nearby voltage magnitudes are prompted significantly and the larger size the better promotion, i.e., bus 31. Figure 203.2b also shows the voltage magnitude of each bus when $Cf = 1$ with the same total capacity, which is an ideal case.

203.5 Conclusions

In the study, the determination of optimal allocation of distributed generations (DGs) with introduction of capacity factor has been solved by an adaptive multi-objective particle swarm optimization (AMPSO) algorithm. The main aim of the research is to determine the optimal sizes and sites of the installed DGs for minimizing the total power loss and the voltage stability index (VSI) of the distribution networks. Multi-objective optimization is well suited to dealing such a problem containing competing objectives. The AMPSO algorithm uses dynamically changing inertia weight and time-varying acceleration coefficients to improve the search capability for obtaining trade-off solutions to this problem. The introductions of the circular crowding sorting technique, the dynamic aggregating function and the mutation operation promotes the diversity of the solutions arising from the algorithm. The proposed method has been also successfully tested on the IEEE 33-bus system. The corresponding results have demonstrated that the

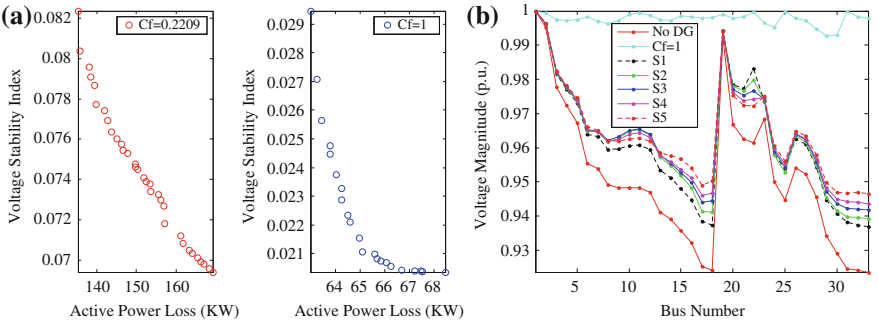


Fig. 203.2 Simulation results. **a** Trade-off solutions when $Cf = 0.2209$ and 1. **b** Voltage magnitudes compared with the cases of DG not being installed and the installed DG operating at rated capacity

Table 203.1 Five candidate solutions having the same capacity and for different installation information

Solution	*floss* (KW)	Improvement (%)		$f_{U\text{sta}}$	Installation information (MW)						
		f_{loss}	$f_{U\text{sta}}$		bus6	bus11	bus15	bus18	bus22	bus25	bus31
S1	139.50	30.78	36.86	0.0629	0	0	0.05	0	2.6	0.4	0.65
S2	131.56	34.72	35.82	0.0639	0	1.15	0.25	0.35	1.55	0.05	0.35
S3	124.05	38.44	32.47	0.0673	0	1.3	0.15	0.6	0.95	0.2	0.5
S4	120.06	40.43	28.95	0.0708	0	1	0.3	0.75	0.6	0.3	0.65
S5	114.62	43.13	26.19	0.0735	0.25	0.3	0.9	0.95	0.3	0.3	0.65

AMPSO algorithm is a feasible and effective way for obtaining trade-off solutions for DG allocation in a distribution network.

Besides active power loss and VSI, other technical and economic objective functions can be considered in the model, and the AMPSO algorithm can, of course, be applied to problems having more than two objectives. For problems having more than three objectives, then other considerations concerning many-objective optimization need to be considered [8, 9].

Acknowledgments This work was supported by Natural Science Foundation Project Natural Science Foundation of China (51177177), State Key Laboratory of Power Transmission Equipment & System Security and New Technology (2007DA10512710205), and the National "111" Project (B08036).

References

1. Wai RJ, Chen YC, Huang LN, Chang YC, Lin CG (2010) Optimization design for installation capacity of hybrid energy generation system. J Taiyuan Univ Technol 5:521–526
2. Liu J, Bi PX, Dong HP (2002) Analysis and optimization of complex distribution networks. China Electric Power Press, Beijing

3. Ratnaweera A, Halgamuge SK, Watson HC (2004) Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. IEEE Trans Evol Comput 3:240–255
4. Luo CY, Chen MY, Zhang CY (2010) Improved NSGA-II algorithm with circular crowded sorting. Control Decis 2:227–231
5. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: Nsga-II. IEEE Trans Evol Comput 2:182–197
6. Chen MY, Zhang CY, Luo CY (2009) Adaptive evolutionary multi-objective particle swarm optimization algorithm. Control Decis 12:1851–1855
7. Baran ME, Wu FF (1989) Network reconfiguration in distribution systems for loss reduction and load balancing. IEEE Power Eng Rev 4:101–102
8. Purshouse RC, Fleming PJ (2007) On the evolutionary optimisation of many conflicting objectives. IEEE Trans Evol Comput 11:770–784
9. Adra SF, Fleming PJ (2011) Diversity Management in evolutionary many-objective optimization. IEEE Trans Evol Comput 15:183–195

Chapter 204
Application of DE-Based SVMs for Fouling Prediction on Thermal Power Plant Condensers

Lianghong Wu, Zhaofu Zen, Xiaoping Zhang and Xuejun Li

Abstract Unexpected fouling in condensers has always been one of the main operational concerns in thermal power plants. This paper describes an approach to predict fouling deposits in thermal power plant condensers by means of support vector machines (SVMs). The periodic fouling formation process and residual fouling phenomenon are analyzed. To improve the generalization performance of SVMs, an improved differential evolution algorithm is introduced to optimize the SVMs parameters. The prediction model based on optimized SVMs is used in a case study of a 300 MW thermal power station. The experiment results show that the proposed approach has more accurate prediction results and better dynamic self-adaptive ability to the condenser operating conditions change than asymptotic model and T–S fuzzy model.

Keywords Fouling prediction · Condensers · Support vector machines · Differential evolution algorithm

204.1 Introduction

Condenser is one of the key equipments in a thermal power plant thermodynamic cycle, and its thermal performance directly impacts the economic and safe operation of the overall plant. Fouling of steam condenser tubes is one of the most important factors affecting their thermal performance, which reduces effectiveness

L. Wu (✉) · Z. Zen · X. Zhang · X. Li
Engineering Research Center of Advanced Mining Equipment,
Ministry of Education, Hunan University of Science
and Technology, Xiangtan, China
e-mail: lhwu@hnust.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_204,

and heat transfer capability with time [1, 2]. It is found that the maximum decrease in effectiveness due to fouling is about 55 and 78% for the evaporative coolers and condensers, respectively [1]. Therefore, the designing and operating of condensers must contemplate and estimate the fouling resistance to the heat transfer. The most commonly used models for fouling estimation are the thermal resistance method and heat transfer coefficient method [3–6]. However, the residual fouling of periodic fouling deposition process and the dynamic changes of heat exchanger operating condition are not considered in these models. Consequently, the estimation error of these methods is very large.

Artificial neural networks (ANNs) are capable of efficiently dealing with many industrial problems that cannot be handled with the same accuracy by other techniques. To eliminate most of the difficulties of traditional methods, ANNs are also used to estimate and control the fouling of heat exchanger in recent years [7–9]. Although the technique of ANNs is able to estimate the fouling evolution of heat exchanger with satisfaction, there are some problems. The selection of structures and types of ANNs depend on experience greatly, and the training of ANNs are based on empirical risk minimization (ERM) principle [10], which aims at minimizing the training errors. ANNs therefore face some disadvantages such as over-fitting, local optimal and bad generalization ability.

Support vector machines (SVMs) are a new machine learning method deriving from statistical learning theory [10]. Since the later 1990s, SVMs are becoming more and more popular and have been successfully applied to many areas such as handwritten digit recognition, speaker identification, chaotic time series forecasting, nonlinear control and so on. Established on the theory of structural risk minimization (SRM) [10] principle, compared with ANNs, SVMs have some distinct advantages such as globally optimal, small sample-size, good generalization ability and resistance to the over-fitting problem [10]. In this paper, the use of SVMs model is developed for the predicting of a thermal power plant condenser. The prediction model was used in a case study of 300 MW thermal power station. The experiment result shows that the prediction model based on SVMs is more precise than the thermal resistance model and other methods, such as T-S fuzzy model [9]. Moreover, to improve the generalization performance of SVMs, differential evolution algorithm is introduced to optimize the parameters of SVMs.

204.2 Periodic Fouling Process in Condenser

In thermal power station condensers, the presence of the fouling represents a resistance to the transfer of heat and therefore reduces the efficiency of the condenser. In order to maintain or restore efficiency it is often necessary to clean condensers. The Taprogge system has been widely applied in the power industry for the maintenance of condenser efficiency [3]. However, the sponge rubber ball system is only effective for preventing the accumulation of waterborne mud, biofilm formation, scale and corrosion product deposition [3]. As for some of

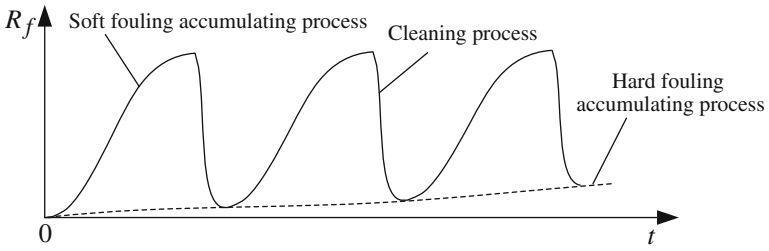


Fig. 204.1 Periodic fouling evolution in power plant condensers

inorganic materials strongly attached on the inside surface of tubes, e.g. calcium and magnesium salts, cannot be effectively reduced by this technique [9]. As a result, at the end of every sponge rubber ball cleaning period, there still exist a lot of residual fouling in the condensers, and the residual fouling will be accumulated continuously with the time. Where the fouling can be cleaned by the Taprogge system is called soft fouling, and those not cleaned by residual fouling is called hard fouling. When the residual fouling accumulates to some degree, the cleaning techniques that can eliminate them, such as chemistry cleaning method, should be used.

Generally, the foul degree of heat exchanger is expressed as fouling thermal resistance, defined as the difference between rates of deposition and removal [3]. In this paper, the corresponding fouling thermal resistance of soft fouling and hard fouling are expressed as R_{fs} and R_{fh}, respectively. Then, the condenser fouling thermal resistance R_f in any time is the sum of soft fouling thermal resistance and hard fouling thermal resistance, expressed as Eq. (204.1) [9].

$$R_f(t) = R_{fs}(t) + R_{fh}(t) = R_f(t_0) + R_{fs}(t - t_0) + R_{fh}(t - t_0). \tag{204.1}$$

where $R_f\,(t_0)$ is the initial fouling.

Figure 204.1 demonstrates the periodic evolution process of fouling in power plant condensers. In fact, the evolution process of fouling in a condenser is very complex, which is related to a great number of variables, such as condenser pressure, cooling water hardness, the velocity of the circulating water and the corresponding inlet and outlet temperatures, the non-condensing gases present in the condenser, and so on. The $R_{fs}(t)$ and $R_{fh}(t)$ expressed a very complex physical and chemical process, their accurate mathematic models are very hard to be obtained. Hence, measurement and prediction of fouling development is a very difficult task. Since the fouling evolution is a very complex nonlinear dynamic process, the traditional techniques based on mathematic analysis, i.e. asymptotic fouling models, are not efficient to describe it [8, 9]. SVMs, as a small sample method to deal with the highly nonlinear classification and regression problems based on statistic learning theory, are expected to be able to reproduce the nonlinear behavior of the process.

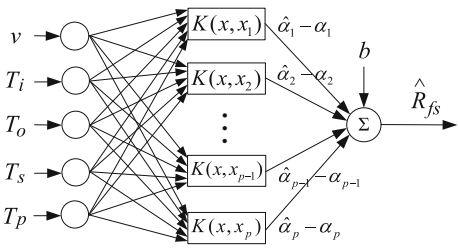


Fig. 204.2 Soft fouling prediction SVM model

204.3 Fouling Prediction Scheme

The formation and development of fouling in condensers is influenced not only by cooling water hardness and turbidity but also by the working conditions of condensers, such as velocity of the cooling water and the corresponding inlet and outlet temperatures, the saturation temperature of steam under entrance pressure of condenser, the non-condensing gases present in the condenser, and so on. According to the previous analysis of periodic fouling process of power plant condensers, the fouling can be classified as soft and hard fouling. Therefore, two SVM models are developed to forecast thermal resistance of soft and hard fouling, respectively. Then, the whole prediction fouling thermal resistance ($\widehat{R}_f$) in condenser is the sum of output of soft fouling prediction model ($\widehat{R}_{fs}$) and output of hard fouling prediction model ($\widehat{R}_{fh}$).

Generally, the evolution of soft fouling is determined by the velocity (v), turbidity (δ), inlet (T_i) and outlet temperatures (T_o) of cooling water, saturation temperature of steam under entrance pressure of condenser (T_s), and prediction time range (T_p) (the running time in a sponge rubber ball cleaning period). Therefore, these variables are chosen as inputs of the soft fouling thermal resistance predictive model. As for hard fouling of the class of calcium and magnesium salts, it is related to the residual fouling at the beginning and the end of previous sponge rubber ball cleaning period (corresponding thermal resistance is $R_{fb,n-1}$, $R_{fe,n-1}$, respectively), hardness of cooling water (σ), saturation temperature of steam under entrance pressure of condenser (T_s), and the accumulating running time of condenser (T_a). Hence, those variables are chosen as the inputs of hard fouling thermal resistance prediction model. The soft and hard fouling prediction model based on SVMs are illustrated in Figs. 204.2 and 204.3, respectively.

The quality of SVMs models strongly depends on a proper setting of parameters and SVMs approximation performance is sensitive to parameters [11, 12]. Parameters to be regulated include hyper-parameters C, ε and kernel parameter σ, if the Gaussian kernel is used [11]. The values of C, σ and ε relate to the actual object model and are not fixed for different data sets. So the problem of parameter selection is complicated.

The values of parameter C, σ and ε affect model complexity in a different way. The parameter C determines the trade-off between model complexity and the

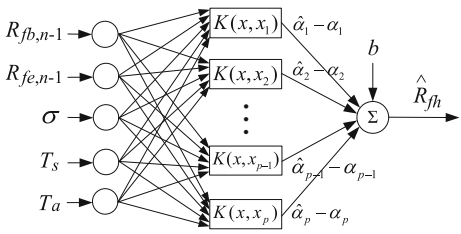


Fig. 204.3 Hard fouling prediction SVM model

tolerance degree of deviations larger than ε. The parameter ε controls the width of the ε-insensitive zone and can affect the numbers of support vectors (SV) in optimization problem. The kernel parameter σ determines the kernel width and relates to the input range of the training data set. Here, parameters selection is regarded as compound optimization problem and differential evolution (DE) algorithm is proposed to select suitable parameters value.

204.4 DE-Based SVMs

DE is a simple but powerful population-based stochastic search technique for solving global optimization problems [13]. DE has three operations: mutation, crossover and selection. The crucial idea behind DE is a scheme for generating trial vectors. Mutation and crossover are used to generate trial vectors, and selection then determines which of the vectors will survive into the next generation. The original DE algorithm is described in the following briefly.

DE operates on D dimensional parameter vectors. The number of vectors is equal to user defined population size. The initial population vector is chosen randomly. A cost function, C is used to rate the individual vector according to their capability to minimize the objective function. The DE process starts from selecting a target vector. Then, it randomly selects two other vectors and generates a difference vector, which is multiplied with a user defined weighting factor F to obtain 'weighted difference vector'. The weighted difference vector and randomly chosen mutation vector are added to create a noisy vector, which is subjected to crossover process with the target vector in order to generate the trial vector. The cost function of trial vector is then compared with the cost function of original target vector. The vector having low cost function is allowed into the new population. The same procedure is adopted for the entire population vector to get a next generation. This process is continued until a termination criterion is satisfied i.e. desired number of generations/constraints. For the space limitation, the detailed algorithm for DE technique may reference literature [13].

The objective of SVMs parameters optimization is to minimize deviations between the outputs of training data and the outputs of SVM, where the mean square error (MSE) is used as the performance criterion. Then the objective of the DE is to search the optimal parameter C, σ and ε to minimize objective function:

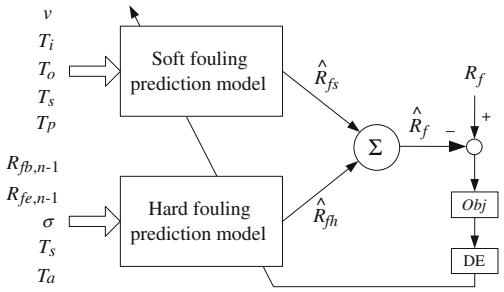


Fig. 204.4 Fouling prediction model based on SVMs optimized by DE

Table 204.1 Fouling thermal resistance prediction results in the first cleaning period

Running time T_{pa} (h)	Operating conditions			Measuring values R_f (K m^2/kW)	Prediction values (K m^2/kW)		
	v (m/s)	T_i (°C)	T_s (°C)		SVMs model	T-S model	Asymptotic model
0	2.0	19.1	33.2	0.0258	0.0260	0.0258	–
5	2.0	18.5	33.3	0.0995	0.0992	0.1018	0.0947
10	2.0	15.6	31.9	0.2028	0.2037	0.2007	0.1872
15	2.0	14.3	31.6	0.2501	0.2494	0.2411	0.2528
20	2.0	15.5	33.5	0.2865	0.2864	0.2830	0.2993
25	2.0	15.5	34.0	0.3174	0.3172	0.3123	0.3323
30	2.0	16.1	34.8	0.3420	0.3393	0.3321	0.3558
35	2.0	14.4	34.6	0.3567	0.3562	0.3497	0.3724
40	2.0	14.2	34.9	0.3722	0.3736	0.3600	0.3842

$$\min F(C,\delta,\varepsilon) = \left(\frac{1}{K}\sum_{k=1}^{K}(y_k - f(x_k,w))^2\right)^{\frac{1}{2}}, \tag{204.2}$$

where K is the number of training data, y_k is the output of the kth training data, and $f(x_k, w)$ is the output of SVMs correspond to input x_k.

Generally, the search range of these parameters is $C \in [1, 1000]$, $\sigma \in (0, 1]$, $\varepsilon \in (0, 0.5]$. For special problems, the search range is changeable. Figure 204.4 illustrates the fouling prediction model using optimized SVMs by DE.

204.5 Experiment Results

In this section, experiments on N-3500-2 condenser (300 MW) in Xiangtan thermal power plant are carried out to prove the effectiveness of the proposed approach. The cooling water of this plant is river water that pumped from the Xiangjiang river. The Taprogge systems are installed in the plant to clean online

Table 204.2 Fouling thermal resistance prediction results of the eighty-fifth cleaning period

Running time T_a (h)	Operating conditions			Measuring values R_f (K m^2/kW)	Prediction values (K m^2/kW)		
	v (m/s)	T_i (°C)	T_s (°C)		SVMs model	T-S model	Asymptotic model
2,830	2.0	2.5	24.9	0.1594	0.1698	0.1632	0.1058
2,835	2.0	2.6	25.4	0.2467	0.2541	0.2345	0.1950
2,840	2.0	2.8	26.3	0.3055	0.3065	0.2827	0.2583
2,845	2.0	2.6	27.1	0.3410	0.3414	0.3177	0.3032
2,850	2.0	2.5	27.7	0.3696	0.3702	0.3548	0.3351
2,855	2.5	2.4	28.1	0.3891	0.3932	0.3776	0.3578
2,860	2.5	2.4	29.2	0.4005	0.4003	0.3883	0.3648
2,865	2.5	2.6	29.6	0.4087	0.4075	0.3979	0.3738
2,870	2.5	2.7	29.4	0.4138	0.4132	0.4062	0.3852

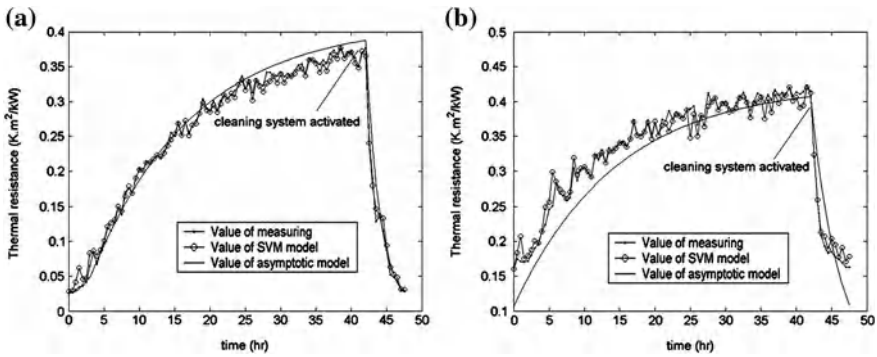


Fig. 204.5 Predicted fouling thermal resistance evolution in the first and eighty-fifth cleaning period: **a** the first cleaning period, **b** the eighty-fifth cleaning period

the condensers. At present, the condenser is cleaned every 2 days using the Taprogge system, and every cleaning time is about 6 h. Obviously, the fitted cleaning period is not optimal, because the fouling accumulating process is dynamic changing with the operating conditions changing.

The experiment system consists of sensors for operating condition parameters measuring, data acquisition system, PC-type computer, etc. A set of 1362 real-time running data in different operating conditions in 84 cleaning periods is collected to train and optimize the SVMs model for fouling prediction, another set of 300 data is chosen for model verification. The control parameters of DE are the following. The number of population is 30, the crossover probability constant CR is 0.5, the mutation factor F is 0.5, and the maximum number of generations is 100. After application of DE, the optimal SVMs parameters of soft fouling prediction model are $C = 848$, $\sigma = 0.513$, $\varepsilon = 0.0117$, the optimal SVMs parameters of hard fouling prediction model are $C = 509$, $\sigma = 0.732$, $\varepsilon = 0.0075$.

The velocity, turbidity, and inlet temperature of cooling water is different in summer and winter, the evolution of fouling in condensers is also different in the two seasons. In the experiments, two sponge rubber ball cleaning periods of the first and eighty fifth period in different seasons are investigated. The hardness and turbidity of cooling water is 56 and 17 mg/L in summer, and is 56 and 29 mg/L in winter.

To demonstrate the effectiveness of the proposed approach, the comparison between the SVMs model, T–S fuzzy logic model [9] and asymptotic model is considered. The asymptotic model is obtained by probability analysis method, and the corresponding expression is as follows [9].

$$R_f(t) = 41.3 \times \left[1 - e^{-(t-1.204)/14.57}\right]. \tag{204.3}$$

Tables 204.1 and 204.2 show the fouling thermal resistance prediction results of the above two models in the first and the eighty-fifth cleaning periods, respectively. From Tables 204.1 and 204.2, we can see that compared with tradition asymptotic model and T-S fuzzy logic model, the SVMs-based prediction model has higher prediction precision. Figure 204.5 shows the predicted fouling thermal resistance evolution based on the optimized SVMs model and asymptotic model. Figure 204.5b clearly shows that the asymptotic model is not able to forecast the fouling evolution process at the beginning stage of the eighty-fifth cleaning period, the reason being that the residual fouling in the periodic fouling formation process is not considered in the asymptotic models.

204.6 Conclusions

The fouling accumulating process in condensers is very complex, which is influenced by a lot of operating conditions. Hence, fouling prediction is a very difficult task, and the traditional techniques, such as asymptotic fouling model and heat transfer coefficient method, are not efficient to describe it. In this paper, the condenser fouling forecasting model based on SVMs was proposed. The parameters of the SVMs model were optimized by differential evolution. Practice application experiment results show that, SVMs-based forecasting model is able to effectively eliminate the influence of residual fouling, and has good dynamic self-adaptive ability. Compared with available models, such as traditional asymptotic model and T–S fuzzy logic model, experiment results present that the proposed optimized SVMs model has higher forecasting accuracy.

Acknowledgments This work was partly supported by the National Natural Science Foundation of China (Grant no. 60835004, 60974048), Hunan Provincial Natural Science Foundation of China (Grant no. 10JJ9007), Hunan Provincial Innovative Platform for Integration of Industries, Universities and Research (Grant no. 2010XK6066), and Aid Program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province.

References

1. Qureshi BA, Zubair SM (2005) The impact of fouling on performance evaluation of evaporative coolers and condensers. Int J Energy Res 29:1313–1330
2. Cristiani P (2005) Solutions to fouling in power station condensers. Appl Therm Eng 25:2630–2640
3. Bott TR (1995) Fouling of heat exchangers. Elsevier Science and Technology, The Netherlands
4. Nebot E, Casanueva JF, Casanueva T et al (2007) Model for fouling deposition on power plant seam condensers cooled with seawater: effect of water velocity and tube material. Int J Heat Mass Transf 50:3351–3358
5. Heo G (2008) Development of a fouling monitor for safety-related heat exchangers. Nucl Eng Des 238:1295–1303
6. Grosfils V, Kinnaert M, Bogaerts P et al (2004) Fouling resistance modeling, identification and monitoring of a thermosiphon reboile. Chem Eng Sci 59:489–499
7. Radhakrishnan VR, Ramasamy M, Zabiri H et al (2007) Heat exchanger fouling model and preventive maintenance scheduling tool. Appl Therm Eng 27:2791–2802
8. Fan S, Wang Y (2008) Fouling soft-sensing in condenser based on feature selection and multiple RBF neural network. Chin J Sci Instrum 29(4):723–728
9. Fan S, Wang Y (2005) The prediction of fouling in condenser based on T-S model. Chin J Sens Actuators 18(1):123–728
10. Vapnik V (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 5:988–999
11. Wang WJ, Xu ZB, Lu WZ (2003) Determination of the spread parameter in the Gaussian kernel for classification and regression. Neurocomputing 55:643–663
12. Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw 17:113–126
13. Rainer S, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11(4):341–359

Chapter 205
Analysis of the Current Return Paths of Power/Ground Plane Segmentation

Yong-Qin Liu, Guo-Qiang Li, Zhi-Ping Zhu and Yan Peng

Abstract The power/ground plane Segmentation suppress the simultaneous switching noise by cutting off the path of SSN. A power/ground Segmentation plane model is established and simulated by software HFSS10.0 to analyze the power/ground plane segmentation current return paths. The simulation results that the slotted plane forced return current to flow back from a distance, and circuit size increases, leading to a potential threat of electromagnetic interference.

Keywords Power/Ground plane segmentation · Simultaneous switching noise · Return path · HFSS

205.1 Introduction

With the trend of digital systems to higher speed and lower power, simultaneous switching noise (SSN) on the power/ground plane has become one of the major bottlenecks for high speed circuit design. Currently, a variety of suppression

Y.-Q. Liu (✉) · G.-Q. Li · Z.-P. Zhu · Y. Peng
School of Physics and Electrical Engineering,
Weinan Normal University, Weinan 714000, China
e-mail: liuyongqin@163.com

G.-Q. Li
e-mail: lgq@wntc.edu.cn

Z.-P. Zhu
e-mail: wnsyzzp@163.com

Y. Peng
e-mail: peyya1984@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_205,

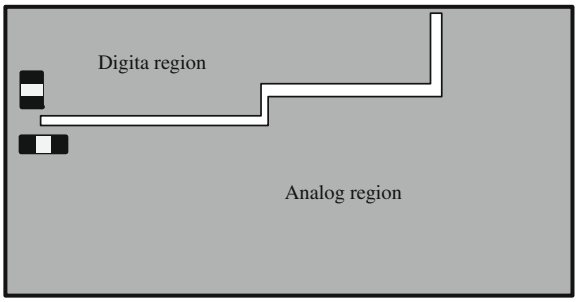


Fig. 205.1 Digital/analog region segmentation

methods has been designed for SSN, which add discrete decoupling capacitors [1–3] and the use of embedded decoupling capacitors [4–6] as the main methods. Due to the noise being mainly through the power/ground plane, a common method for rejection or isolation is the power/ground plane segmentation process. The method had been used in mixed analog-digital design, shown in Fig. 205.1. In the area of digital and analog power plane between the slotted area of supply plane to prevent noise from the digital region into the simulation area. As the digital and analog share the same reference voltage, the two regions link through a narrow band region. Note that, in establishing connectivity to add decoupling capacitors to be effective at preventing noise spread through the connectivity.

The method of power/ground plane segmentation can achieve the purpose by cutting off the noise spread. The noise isolation effect has to rely on full-wave simulation tools, in order to determine the bandwidth in the range of interest to meet the requirements. If it cannot meet the requirements, it is necessary to re-partition or add decoupling capacitors and other measures for improvement.

205.2 Digital/Analog Region Segmentation

Another common power/ground plane segmentation method is called the power island shown in Fig. 205.1. In this practice, a metal plane is commonly used in two different supply voltages, and both the power supply are completely separated to provide a different reference voltage. In the low-frequency high-density PCB, this method is commonly used, its main purpose being to save costs. It is not good when the plane segmentation in the high-frequency noise isolation, especially when the two regions at the same resonance can produce strong coupling. In addition, the continuity of the reference plane is destroyed, the reference signal for its alignment has great influence on the signal integrity in high speed systems. If the plane segmentation is not handled properly it will lead to a series of very difficult signal integrity and electromagnetic interferences.

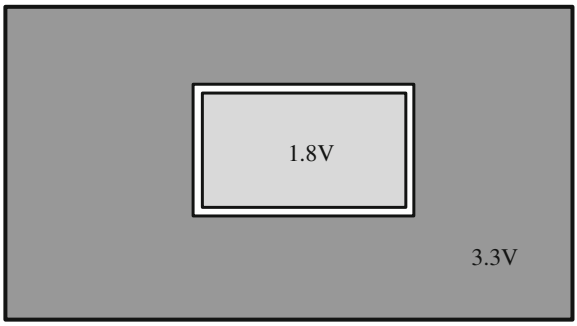


Fig. 205.2 Power islands plane segmentation

205.3 Power Islands Plane Segmentation

The best noise isolation is complete with different power/ground plane supply, shown in Fig. 205.2. At this point, it can not only maintain the best noise isolation for two power supply systems but can also bring great benefits to the system's signal integrity, power integrity and electromagnetic interference.

205.4 Different Power/Ground Power Supply

Shown in Fig. 205.3. The disadvantage is the need for more metal layers, but for high-speed applications, the advantages it brings are much larger than the disadvantages. The current high-performance computer power supply board is using this method, the processor core and I/O circuits were used to complete different power/ground supply, so the system noise management and signal transmission can be properly resolved. With the development of IC, the power supply noise isolation method will be widely used.

205.5 Simulation and Analysis of Plane Segmentation

High frequency structure simulator (HFSS) is recognized as a more accurate three-dimensional full-wave simulation software generic. The software uses FEM algorithm for accurate high-frequency simulation of various structures. The following is established by using HFSS10.0 power/ground plane segmentation model and used to analyze the power/ground plane segmentation return path.

Simulation of the interconnect structure shown in Fig. 205.4, the structure formed by two layers of metal, the medium thickness is 1.3 mm. It can cut off the noise transmission by a slotted flat bottom. Signal traces on the top floor across the slot, a discontinuous interconnection return path is caused by the slotted plane.

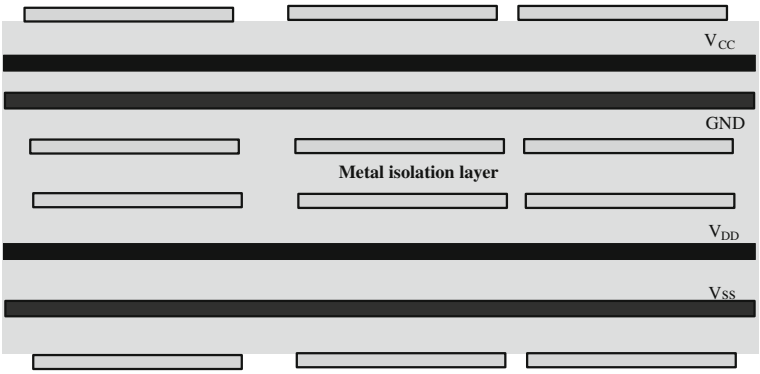


Fig. 205.3 Different power/ground power supply

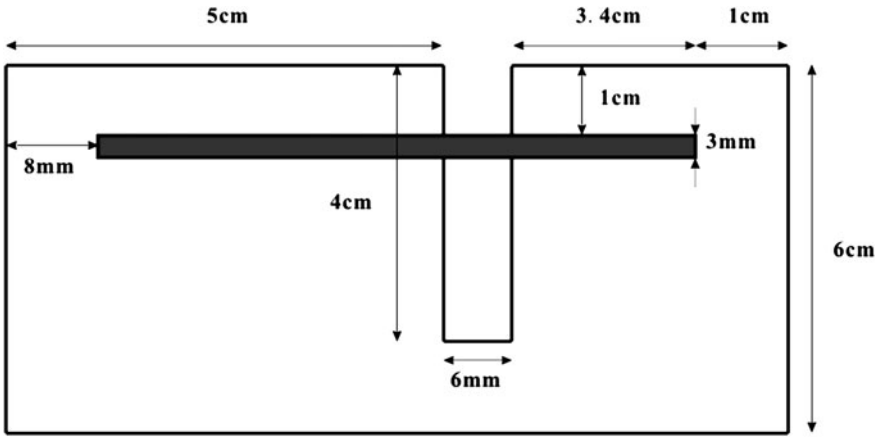


Fig. 205.4 Interconnect structure

Set the type of solution for signal integrity and power integrity simulation, select the Driver terminal type. Then create 3D models, set the media material My_FR4, dielectric constant of 4.5. Create ground level. Ground top and bottom to create two flat Board, Ground under the Board. Cut one Piece from the Ground, named Cutout, forming a continuous ground plane. Shown in Fig. 205.5. Establish the alignment model. Alignment named Trace, trace size dX = 3 mm, dY = 8.2, dZ = 0. Trace distribution of the ideal conductor of electricity. Shown in Fig. 205.6.

Because of PML, a face list needed to generate for the radiation calculation. Analysis setting, set frequency to 2.5 GHz, the impedance is set to 0.1, model checking and simulation. Check whether the convergence of the simulation.

Observe the simulation results, the resulting S-parameters as shown in Fig. 205.7. View the current distribution on the Ground, and the current distribution shown in Fig. 205.8.

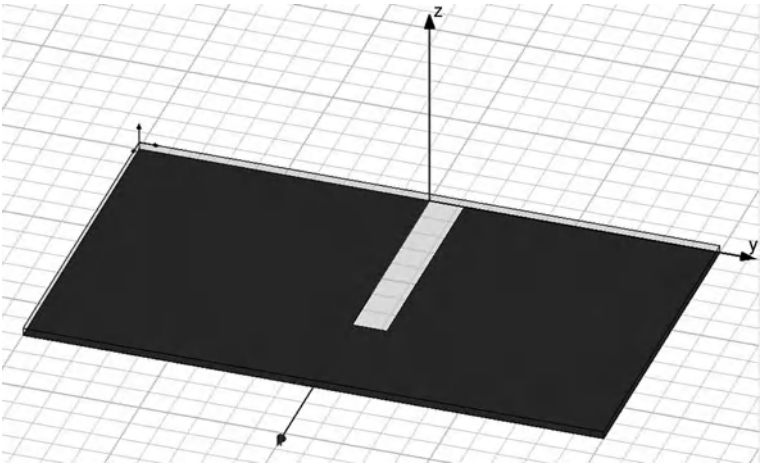


Fig. 205.5 Slotted design graphics

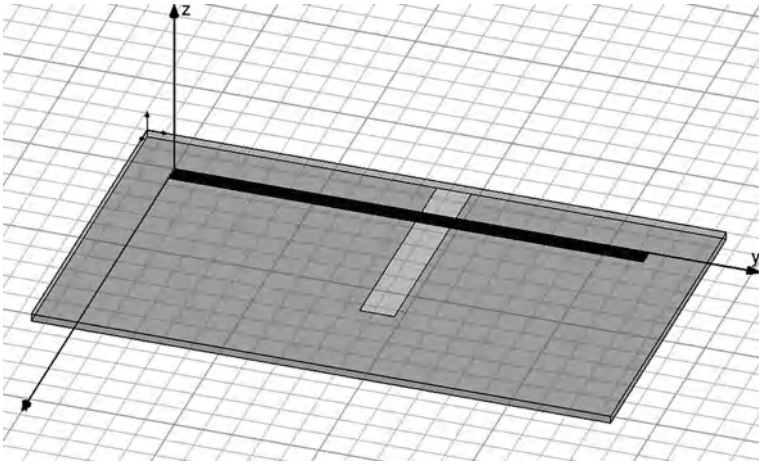


Fig. 205.6 View of structural alignment

Figure 205.8 shows the current distribution, flat slotted force the return current from a distance, and the area increases, leading to a potential threat of electromagnetic interference. The biggest problem of ground plane segmentation is that when the division plane alignment across the slot, how to provide a low impedance return path, which is obviously difficult to do. If not implemented correctly, a ground plane segmentation may bring the issues to be much greater than its benefits. Multilayer PCB power plane split in the more common, however, split power plane and ground plane will also result in similar split the signal integrity and electromagnetic interference. It is relatively easy to maintain the split power

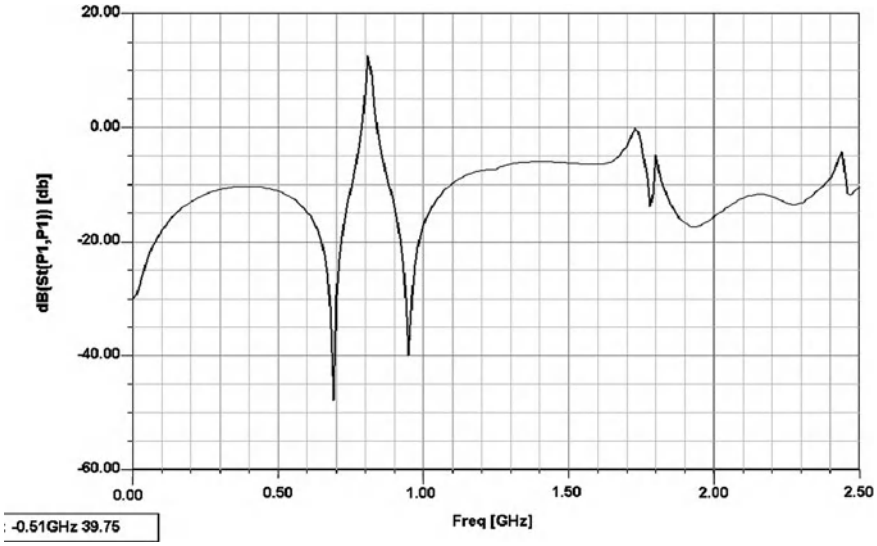


Fig. 205.7 Curve of parameters

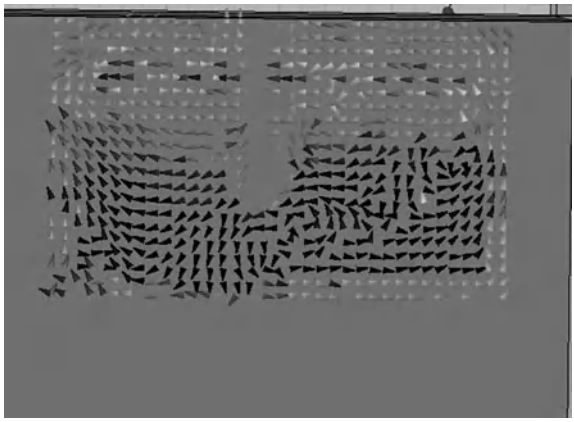


Fig. 205.8 Graphics of current distribution

plane slot not cross the signal line in multi-layer PCB, the signal can be managed across the continuous ground plane on the other. The best slot is located inside the power supply plane layer.

205.6 Conclusion

The biggest problem of plane segmentation is the signal on the slot alignment. Whenever possible, never let the signal traces across the slot, both low-speed signal lines, power lines and high-speed signal lines. It is difficult to control the

return path across the slot alignment, likely to cause signal integrity and electromagnetic interference. If these traces are not avoided, then it must be near the signal traces to provide a return path. Although the two split planes are shorted, it must provide such a return path.

Acknowledgments This work was supported by the Civil-Military Integration Foundation of Shaanxi Province (11JMR07) and the Natural Science Foundation of Education Department of Shaanxi Province (11JK0936).

References

1. Archambeault B, Ruehli A (2001) Analysis of power/ground-plane EMI decoupling performance using the partial-element equivalent circuit technique. IEEE Trans Electromagn Compat 43(4):437–445
2. Beker B, Wallace D (2002) A methodology for optimal value selection of decoupling components for power delivery systems of high-performance ASICs. In: IEEE 11th Topical meeting electrical performance of electronic packaging, pp 195–198
3. Yang X, Chen QL, Chen CT (2002) The optimal value selection of decoupling capacitors based on FDTD combined with optimation. In: IEEE 11th Topical meeting electrical performance of electronic packaging, pp 191–194
4. Hobbs JM, Windlass H, Sundaram V, Chun S, White GE, Swaminathan M, Tummala RR (2001) Simultaneous switching noise suppression for high speed systems using embedded decoupling. In: Proceedings of Electronic Components Technology Conference, pp 339–343
5. Kosaka Y, Takahashi N, Suminaga S (2004) Controlling simultaneous switching noise with built-in decoupling capacitors. In: proceedings of electrical Packaging Technology Conference, pp 652–656
6. Kim H, Jeong Y, Park J, Lee S, Hong J, Hong Y (2003) Significant reduction of power/ground inductive impedance and simultaneous switching noise by using embedded film capacitor. Electr. Perform. Electron. Packag. 129–132

Part XIV
Control System

Chapter 206
Study of Master Manipulator Control System for Teleoperation

Zunyi Deng, Yimin Mo, Xueqin Gao and Yi Ning

Abstract Master manipulator's performance directly affects the controllability and reliability of teleoperation system. Based on the study of the master manipulator's role in the teleoperation, this paper describes a distributed master manipulator control system based on CAN bus, and the design rationale of the joint controller is elaborated. In this control system, a lot of servo control and sensor information processing are handled in each joint controller located physically close to the motor that it is controlling. So the master controller's burden is lightened, and the wiring harness is enormously simplified.

Keywords Teleoperation · Master manipulator · Distributed control · CAN bus

Z. Deng (✉) · Y. Ning
Research Institute of Robotics, Henan University of Technology,
Zhengzhou 450007, China
e-mail: dzyhnxp@163.com

Y. Ning
e-mail: dzyhnxp@163.com

Z. Deng · Y. Mo
School of Mechanical and Electronic Engineering,
Wuhan University of Technology, Wuhan 430070, China
e-mail: dzyhnxp@163.com

X. Gao
Henan Institute of Science and Technology,
Xinxiang 453003, China
e-mail: dzyhnxp@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_206,

206.1 Introduction

Nowadays, due to the limitation of a number of key technologies, including mechanism, control, sensing and intelligence, etc., it is an impossible target to develop a fully autonomous intelligent robot. Telerobot is a realistic and the best choice in the unstructured environments, which combines the advantages of human remote control with the autonomy of industrial robot. Teleoperation technology is using in some special working conditions such as aerospace, military, nuclear industry, deep-sea exploration, and so on. It will play an important role in more and more areas with the development of society and technology [1].

Master manipulator is an important man–machine interface in teleoperation. The bidirectional information exchange and sense reestablishing between man and machine is finished by master manipulator, therefore the load of calculation and communication is heavy. If the centralized control is used, the operational speed of CPU is difficulty to satisfy the performance of real-time and high fidelity. It is propitious to further improve technical performance of master manipulator that dealing with a great lot of servo control and sensor information in each joint controller based on Fieldbus. Moreover, master manipulator has many sensors, and it's wiring harness is complicated, which influences the performance of force telepresence of teleoperation system. A lot of wirings affect use, maintenance and appearance. The paper describes a control system of master manipulator based on CAN bus for the above-mentioned problem. The control system has many features, including decreasing additional interfere, higher anti-jamming ability, simple wiring, concision layout, good real-time and convenient repairs, and so on.

206.2 Control System Main Task

In general, teleoperation system consists of the operator, master manipulator (or master robot), communication system, slave robot and the environment for slave robot. The operator pulls the handle of master manipulator, and control instructions are issued through the communication system, so that slave robot begin to move under the control of master manipulator. At the same time, the interactive force between slave robot and circumstance is converted to the feedback instructions which returns to the master manipulator, so the operator generates the force sense [2].

It can be seen, the master manipulator plays a very important role in the teleoperation system. The two-way information exchange between man and machine and the feel reconstruction requires to complete by master manipulator. The main task of the control system includes the followings:

1. Kinematics solver. In each working cycle, the operator palm's position and attitude parameters are calculated through online testing angular displacement of each master manipulator joint, combined with the geometric parameters of the master manipulate.

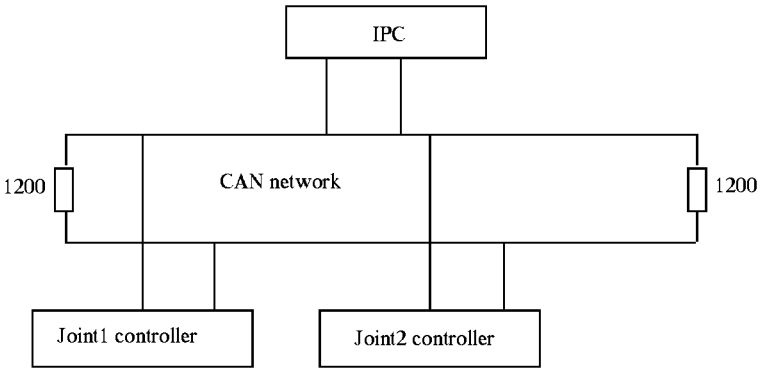


Fig. 206.1 Control system structure

2. Force representation. The interactive force between the slave robot and the circumstance where the slave robot lies in is transferred to the master manipulator, and the force is reasonably assigned to the master manipulator's joints, according to the configuration of the master manipulator. The joint motors are driven based on the force assigned, so the operator generates the force sense through the handle of the master manipulator.
3. Directly send control commands to slave robot. some known environmental information and the prior knowledge can be directly sent control commands to slave robot which complete these operations independently, so the operator's fatigue can be reduced.
4. Manipulator's control state fast switch. In order to make master manipulator be more humane, LIMP function should be established. When the LIMP is on, the data link between the master manipulator and slave robot is automatically cut off, the master manipulator can return with no load. When LIMP is off, the slave robot is control by master manipulator.

206.3 Control System General Structure

The master manipulator in the teleoperation is essentially a robot, which not only has all of the basic general characteristic of industrial robots, but also needs to realize sense reestablishing, and provides the high-fidelity force telepresence to the operator. It's control system is more complex than the control system of industrial robots. Reference [3] the system described in this paper is a distributed control system which is composed of the host computer, the controller for each joint and the communication network. As shown in Fig. 206.1, the host computer is in charge of the kinematics solver, fault diagnosis, and human–computer interaction, and so on. The Joint-to-Joint and Joint-to-Host communication is the CAN bus,

which is a highly reliable standard developed by Robert Bosch. It is a multi-master system, with sophisticated error checking and arbitration, so that any high priority message will always get through first without corruption by other messages. The CAN network is connected via the twisted pair and terminated at either end with 120 Ω resistors, to eliminate signal reflections in the wiring. Each joint controller of the master manipulator is responsible for the joint signal acquisition and transmission, and receives the torque information assigned to each joint, as well drives the joint motor to move.

206.4 Joint Controller Design

Each joint controller is in charge of the acquisition and transmission of joint movement signal, and achieves the force telepresence which is generally finished by the motor. Although, some manipulator's sense reconstruction is not achieved by the motor, which is obtained by a small electromagnetic brake that is installed on the corresponding joint of the master manipulator. The operator obtains the corresponding force feeling through changing the voltage to control the braking torque. This way is relatively simple, and difficult to achieve the precise control of force. So the brushless DC motor is chosen to obtain the force reconstruction, which is simple, large output torque, good performance and a small electrical time constant. The joint controller includes the MCU, movement signal acquisition module, motor drive and control module, communication module, etc.

206.4.1 MCU

The freescale's MC9S12DG128 is selected as MCU, which is a 16-bit device composed of standard on-chip peripherals including a 16-bit central processing unit, 128 K bytes of Flash EEPROM, 12 K bytes of RAM, 4 K bytes of EEPROM, an 8-channel pulse-width modulator, 29 discrete sigital I/O channels, five CAN2.0 A, B software compatible modules. The MCU's main task include movement signal acquisition, completing the direct torque control algorithm, the exchange of information with the host computer, controlling the precision motion controller LM629 to drive motor, etc.

206.4.2 Communication Module

The chosen communication way between the host computer and the joint controllers is the CAN bus, which is a highly reliable communication with sophisticated error checking and arbitration. It is a multi-master system, any one

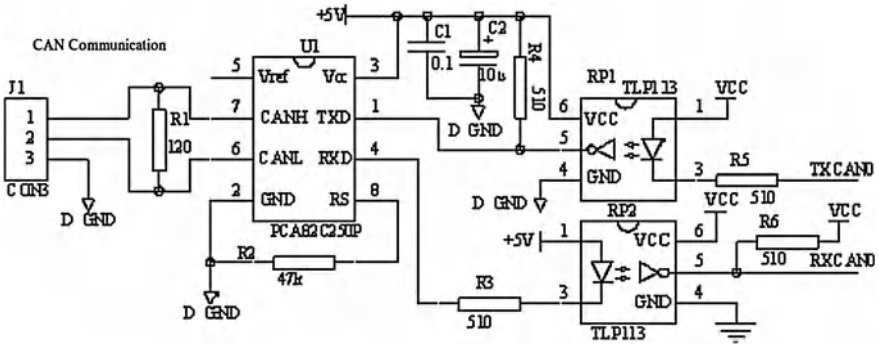


Fig. 206.2 CAN communication interface circuit

node on the network can send information to the other nodes on the network at any one time. It uses short frame structure and transmission time is short, low probability of interference, the data error rate is very low. In the case of serious error, the node can automatically turn off the output function, so that other nodes on the bus are not affected. It uses twisted pair as communication medium, the connection is simple and flexible. The communication interface circuits are shown in Fig. 206.2. The CAN communication software design is implemented through setting various resistors of the MSCAN module, which includes data transmission module, data reception module and error handling module.

206.4.3 Motor Control and Drive Module

The brushless DC motor is chosen to obtain the force reconstruction. The LM629 which is a special all-digital precision motion control processor was used to control the brushless DC motor. A microcontroller, a LM629, a power driver, a DC motor can form a servo system. LM629 has 22 instructions which can be used for MCU to control, transfer data and understand the status information, so as to achieve servo control for motor. Figure 206.3 shows the LM629 control flow.

In the brushless DC motor drive system, when the main circuit voltage is greater than 15 V, the MOSFET drive circuit is more complex and larger. So the brushless DC motor's use in the smaller areas is limited. The joint controllers of master manipulator are located physically close the motors which they are controlling, therefore the size of the joint controller is needed to be small as possible. The SI9979 is a small size and low cost driver chip of motor. It is especially suit for applying in the spatial small situation. So SI9979 is chosen as a motor driver chip which can complete motor speed control, overcurrent protection, undervoltage, etc. with the simple external circuit.

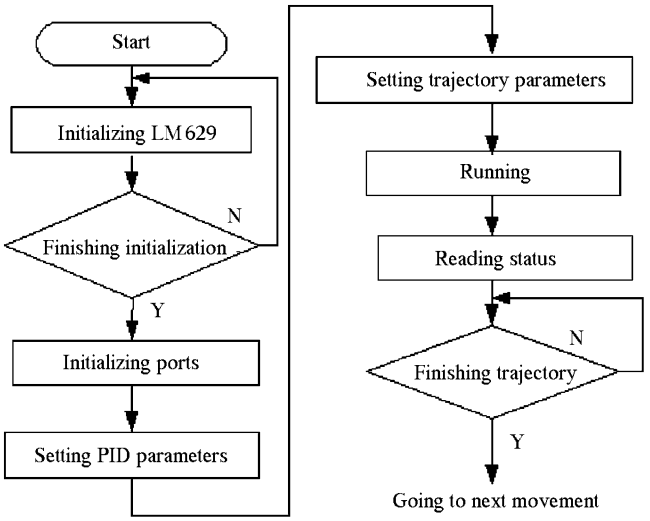


Fig. 206.3 LM629 control flow

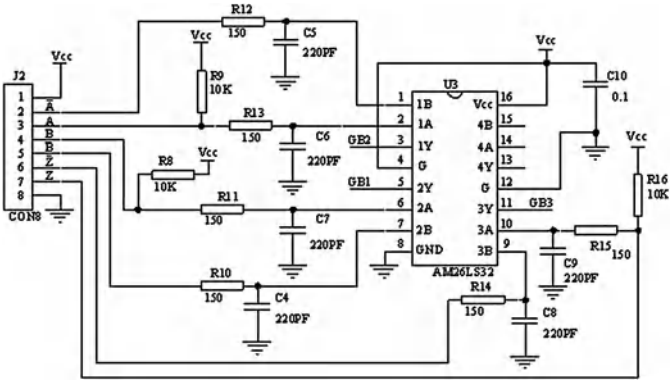


Fig. 206.4 Signal detection and processing circuit

206.4.4 Movement Signal Acquisition Module

Movement signal acquisition module mainly completes the motor encoder signal acquisition, the rotary encoder signal detection and processing circuit is shown in Fig. 206.4. The AM26L32 is the differential signal side bus receiver which transforms the received differential signal into the single end signal, and inputs to LM629.

206.5 Conclusion

Master manipulator is a key equipment for teleoperation system, through which the two-way information exchange and the feel reconstruction between man and robot needs to be done. The control system performance will directly affect the human–computer interaction performance of the whole teleoperation system. In this paper, the distributed control system was built based on a study of the control performance of master manipulator, which deals with a large number of servo control and sensor information about each intelligent node. So it helps to further improve the technical performance of the manipulator. As the master manipulator has more sensors, and the line is complex, so the layout of the wiring harness often directly affects the force-telepresence of the teleoperation. After the distributed control system based on CAN bus is used, the Joint-to-Joint and Joint-to-Host wiring is only the twisted pair wire, which simplifies wirings and reduces the additional interference on the system, the high fidelity force telepresence performance is achieved. From the current technology development trends, the master manipulator based on the fieldbus is needed as soon as possible; the master manipulator based fieldbus in-depth study is the inevitable trend of current technology.

Acknowledgments The authors gratefully acknowledge Prof. Fuan Chen (electrical institute, Henan University of Technology) for enlightening advices and discussion. The authors also thank Shiquan Zhao (Research Institute of Robotics, Henan University of Technology) for providing great technological supports. This work was supported in part by school program (10XZR015) from Henan University of technology.

References

1. Jianfeng W, Aiguo S, Jianqing L (2010) A three-dimensional force reflecting hand controller. Chin J Sens Actuators 23:1417–1420
2. Honglei A, Qing W, Dapeng H (2010) Design of a hybrid force-reflecting hand controller and kinematic simulation. Computer Simul 27:174–177
3. Jianwei C, Aiguo S, Weiyi H (2010) A new 6-DOF universal hand controller. Chin Mech Eng 16:320–324

Chapter 207
Bidirectional Variable Probability RRT Algorithm for Robotic Path Planning

Chen Xiong, Kong Yingying and Han Jianda

Abstract The traditional goal-bias RRT is mentioned to improve the efficiency, but it has an inherent problem, when there are lesser vertexes, the search toward the goal is often invalid; however, when there are more vertexes, the search toward other regions is often unnecessary. To solve this problem, we introduce a kind bidirectional variable probability RRT algorithm. In this paper, we build two trees, and one tree expands toward to the other tree at a variable probability. This probability is in proportion to the coverage of the trees, that is, when there are lesser vertexes, the searches are mainly toward unexplored regions, and when there are more vertexes, we attach more importance to the connection of two trees. The results show the good performance and convergence speed of the proposed algorithm.

Keywords Motion planning · Nonholonomic constraint · RRT · Goal bias · Dynamics models

207.1 Introduction

Rapidly exploring random tree (RRT) has attracted a lot of recent attention. This algorithm can work effectively and does not need to preprocess the map, especially when we take dynamics into account [1].

C. Xiong (✉) · K. Yingying
Intelligent Control Research Lab, Fudan University, Shnaghai, China
e-mail: 09210720032@fudan.edu.cn

K. Yingying
e-mail: chengxiong@fudan.edu.cn

H. Jianda
Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_207,

The initial RRT algorithm is mentioned by LaValle, it is incremental forward-searching [2]. At first, we build a tree T rooted in the initial X_{init}. In each iteration, a random state X_{rand} is selected from the state space, and then the closest vertex X_{near} is found. The next step is to select an input vector $u \in U$ that acts on the X_{near}. By integrating the transition equation over a fixed time interval, X_{new} can be determined.

In order to improve the efficiency, all kinds modified algorithm were mentioned. A kind of di-RRT was mentioned by Lavalle [3]. In this paper, a tree rooted in the initial state and a tree rooted in the goal state were built, and search was done until they connected together. And a kind of goal-bias method is also mentioned which the tree explored toward the goal at a probability P [1], but it has an inherent disadvantage. At the beginning of the search or when there is lesser vertexes, the explore toward the goal is always invalid, especially in the narrow pass environment [4–7]. In 2007, a goal-bias RRT based on the grid method was mentioned [8]. In this paper, the workspace was decomposed into non-overlapping regions, and a tree with the initial grid as its root was built. This method has a good performance, but it was meant for the static environment.

In order to solve the inherent problems of the goal-bias RRT, in this paper, we propose a novel goal-bias RRT. In this paper, two search tree was built, and we connected these two at a probability P, and we use the coverage [9] to control the change of probability p.

This paper is organized as follows. In Sect. 2, we briefly introduce the car model and its non-holonimic constraint. In Sect. 3, we present this new goal-bias RRT. In Sect. 4, the simulations results are analyzed and compared. Lastly our conclusion is presented.

207.2 The Car Model and Its Non-holonimic Constraint

The robotic state transition equation is defined as $\dot{x} = f(x, u)$; in this equation, the vector $u \in U$ is the input vector. $\dot{x}$ denotes the derivative of state. It can be obtained by using Euler Integration, $xnew \approx x + f(x, u)\Delta t$; however, it usually uses a higher order integration technique, such as Runge–Kutta. The formula (207.1) is a three-order Runge–Kutta.

$$\begin{cases} x_{n+1} = x_n + \frac{\Delta t}{6}(k_1 + 4k_2 + k_3),\ k_1 = f(t_n, x_n) \\ \\ k2 = f(t_n + \frac{\Delta t}{2},\ x_n + \frac{\Delta t}{2}k_1),\ k3 = f(t_n + \Delta t, x_n - \Delta t k_1 + 2\Delta t k_2) \end{cases} \tag{207.1}$$

The simplified car model is shown in Fig. 207.1. The car state $X = (x, y, \theta)$, and its transition equation is as follows

$$\dot{x} = V\cos\theta,\ \dot{y} = V\sin\theta,\ \dot{\theta} = \frac{V}{L}\tan\phi \tag{207.2}$$

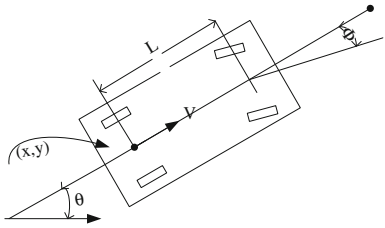


Fig. 207.1 The simplified car model (x, y) is the position of the rear wheels center, and θ denotes the orientation for the car

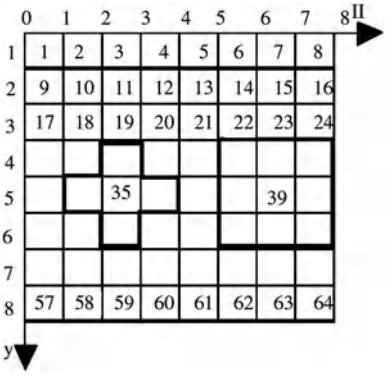


Fig. 207.2 The decomposition of workspace

In this equation, $(V, \phi) \in U$ is the input, V denotes the velocity of the car, and ϕ is the steering angle, and $V \leq V_{\max}, |\phi| \leq \phi_{\max}$.

207.3 Bidirectional Variable Probability RRT

207.3.1 The Calculation of Probability

The probability P is related to the coverage of the tree, when there is a higher coverage, the probability is also larger. To get the coverage of the trees, we first decomposed the workspace into non-overlapping grips, and the coverage refers to the number of the grips that contains vertexes of trees.

(1) *Workspace Decomposition*

Let W denote the 2D workspace on which the robot operates, and builds a coordinate system. The maximum value of the x axis is $X_{\max}$, and the maximum value of y axis is $Y_{\max}$. Take l as the decomposition step. Figure 207.2 gives us an example of decomposition. The grid number of each row is $Nx = X_{\max}/l$ and the grid number of each column is $Ny = Y_{\max}/l$. For an arbitrary grip g, its coordinate value is $g(x, y)$, $x =$ row, $y =$ col. $R = \{1, 2, 3, \ldots, M\}$ is defined as a set of grids order number. The order number of g(1, 1)is 1, g(1, 2) is 2, and the

```
1: D ← workspacedecomposition(w)
2: initializetree T.init and T.goal
3: p ← p_min
4: for i = 1: k do
5:     prand= rand()
6:  if prand > p
7:     xrand = randomstate()
8:     xnew = extend(T.init, xrand)
9:     if connect(T.init,Ygoal);break;end
10:    xnew = extend(T.goal, xrand)
11:    if connect(T.init,Ygoal)
12:    break
13: if prand ≤ p
14:    xrand = randstate(T.goal)
15:    xnew=extend(T.init,xrand)
16:    if xnew = goal;break;end
17: update p
```

```
function:extend(T,Xrand)
1 xnear ← nearest_neighbor(T,xrand)
2 u ← select-input(xrand,xnew)
3 xnew ← new-state(xnear,u,△t)
4 T.add-vetex(xnew)
```

Fig. 207.3 The pseudocode of this paper, the left part is the main function and the right part is the subfunction

relationship between the $g_i = (x_i, y_i)$ and the order number can be represented by the following equation:

$$\begin{aligned} x_i &= (i-1) \bmod Nx + 1, \\ y_i &= (\text{int})((i-1)/Nx) + 1, \end{aligned} \tag{207.3}$$

Proj Function

For a state (x, y, θ), the point workspace decomposition regions g_i: $\forall p \in w$: $proj(p) = g_i$ iff $p \in g_i$ can be located, the coordinate value equation is as follows:

$$xi = (\text{int})(x/l) + 1,\ yi = (\text{int})(y/l) + 1 \tag{207.4}$$

Using this function, we can get the coverage, when there is a new vertex added to the trees, and the coverage should be updated. The probability P can be defined as:

$$p = p_{\min} + \frac{p_{\max} - p_{\min}}{N} \times \text{cov} \tag{207.5}$$

At (207.5), $p_{\min}$ and $p_{\max}$ are the minimum probability and maximum probability. When there are lesser vertexes, the trees are incline to search toward the unexplored regions, and when there are more vertexes, the trees are inclined to search toward the goal state.

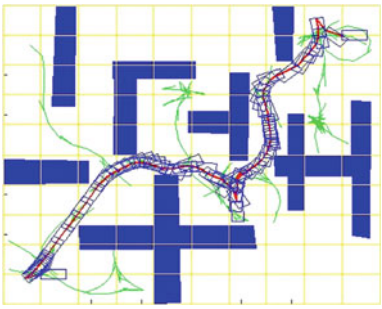


Fig. 207.4 The common obstacle environment

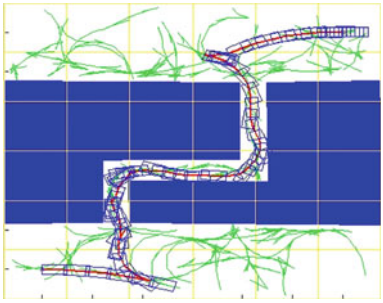


Fig. 207.5 Simple narrow pass environment

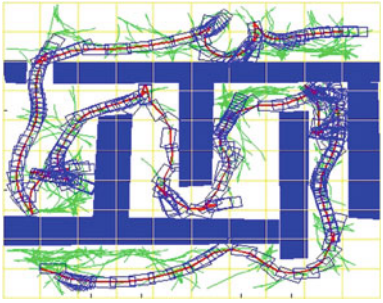


Fig. 207.6 Complex narrow pass environment

207.3.2 The New RRT Algorithm

In this paper, we build two trees and two states x and x' which are considered to be common if $\rho(x, x') < \varepsilon$, and ε is a given constant. The following is the pseudocode (Fig. 207.3).

Step 1: Decompose the workspace, build two trees T.*init* and T.*goal* initialize the probability P.
Step 2: A probability value p_{rand} was selected from [0, 1], if $p_{rand} > p$, select a state $xrand \in X_{free}$, else select a state $xrand \in T.goal$.
Step 3: Expand the T.*init* and T.*goal* using function extend (T, Xrand). When $p_{rand} > p$, just expand T.*init*.

Table 207.1 The comparison between algorithms

	Common obstacle environment		Simple narrow pass environment		Complex narrow pass environment	
	Average time(s)	Average vertexes	Average time(s)	Average vertexes	Average time(s)	Average vertexes
Di-RRTs	6.008	1548	8.0661	3468	19.3918	6811
Goal-bias RRT	21.4216	4326	10.629	3588	26.965	6428
Our algorithm	2.4433	586	3.7576	1578	10.1875	4254

Step 4: After a new state *xnew* was added to a tree, it was tested to see if it lies within a ε-neighborhood of another tree. If so, a path is found and the iteration ends. If not, updates the probability P, and returns to step 2.

207.4 Experimentation

The parameters are set as these: the distance between the front and rear axles: $L = 0.5$, the body length and width are set to L and $0.5L$, for $xrand = (x, y, \theta)$, $x, y \in (-3.75, 3.75)$ and $\theta \in [-\pi, \pi)$, for the input vector $(v, \phi), |V| \leq 3, |\phi| \leq 50°$, the start point is $(-3,-3)$, the goal point is (3, 3) and grip resolution is 10 * 10.

The algorithm was run in three kinds of environment: common obstacle environment, simple narrow pass environment and complex narrow pass environment, and was tested to be effective. The results are as follows (Figs. 207.4, 207.5 and 207.6).

This method is also compared to the traditional goal-bias RRT and di-RRT. Its result is shown in Table 207.1.

From the table we found that the efficiency is improved obviously, the average running times of our improved algorithm are only 40.6, 46 and 53% of the Di-RRTs, and 11, 35 and 38% of the traditional goal-bias RRT in three kinds environments. The invalid vertexes are also reduced obviously.

207.5 Conclusion

The randomized planning techniques have been proved useful in a wide array of applications that includes Robotics. Goal-bias RRT has proved to improve the efficiency of the traditional RRT. But it often makes some invalid search which impacts the path planning. Our paper presents a bidirectional variable probability RRT algorithm. In our paper, we can control the change of probability by the coverage of trees, and then control the main searching regions. Our method is proved to be effective.

References

1. LaValle SM (2006) Planning algorithms. Cambridge University Press, Cambridge
2. LaValle SM (1998) Rapidly-exploring random trees: a new tool for path planning. Computer Science Department, Lowa State University, Ames. Technical Report 98–11
3. LaValle SM (2001) Randomized kinodynamic planning. Int J Robot Res 20(5):378–400
4. Liu H, Yang J, Lu J, Tang Z, Zhao C, Chen W (2006) Review of robotic motion planning with dynamics. Eng Sci 8(1):86–92
5. Plaku E, Kavraki LE, Vardi MY (2007) Discrete search leading continuous exploration for kinodynamic motion planning. In: Robotics: Science and Systems, MIT Press, Atlanta, Cambridge, pp 326–333
6. Ladd AM, Kavraki LE (2004) Fast tree-based exploration of state space for robots with dynamics. In: Proceedings of the international workshop on the algorithmic foundations of robotics, Utrecht/Zeist, The Netherlands, pp 297–312
7. Kuffner JJ, LaValle SM (2000) RRT-connect: an efficient approach to single-query path planning. In: Proceedings of the IEEE international conference on robotics and automation
8. Guo H, Zhu B, Xu S (2007) Rapid-exploring random tree algorithm for path planning of robot based on grid method. J Nanjing Univ 7(2):58–67
9. Plaku E, Kavraki LE, Vardi MY (2010) Motion planning with dynamics by a synergistic combination of layers of planning. IEEE Trans Robot 26(3):469–474

Chapter 208
Design of a Wearable Quantitative Muscle Rehabilitation Evaluation Robotic Hand Device

Xikai Tu, Liguo Yu and Jin He

Abstract There is increasing interest in using robotic devices to assist in improving motor function for stroke patients. Rehabilitation evaluation during the motor rehabilitation process plays a crucial role in instructing robotic or manual therapy. Rehabilitation therapists can take some measure in assessing stiffness or motion range of the patient's joints, but it is difficult for them to evaluate quantitatively and objectively. Furthermore few wearable robotic devices are reported for evaluating the recovery conditions of the patient's hand internal muscles. The development of a pneumatic muscle (PM)-driven quantitative muscle rehabilitation evaluation robotic hand device serves this role. On account of the variance of the patient's muscle recovery active forces and the complex nonlinear dynamics of PMs, the device becomes difficult to control. This paper presents a kind of feasible control strategy applied to this device.

Keywords Muscle rehabilitation evaluation · BP-PID · Pneumatic muscle (PM) · Rehabilitation robotic hand · Stroke

X. Tu (✉) · J. He
Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: tuxikai@gmail.com

X. Tu · L. Yu · J. He
Neural Interface and Rehabilitation Technique Research Center, Huazhong University of Science and Technology, Wuhan 430074, China

L. Yu
Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430074, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_208,

208.1 Introduction

Recent research has indicated that the brain has neural plasticity to recover function after neurological injury. This finding forms the foundation for new therapeutic treatment of patients with neurological diseases: repetitive motor function activities. Stroke is a leading neurological disorder causing severe disabilities or death. Fortunately, stroke patients could partially regain their functional impairment through effective rehabilitation training [1]. Traditionally, physical therapy plays an important role in the mainstream rehabilitation for stroke patients. It involves human therapists for assisting stroke patients to recover their motor ability. However, therapists are currently expensive, and lack consistency and objective assessment [2]. A recent innovation in rehabilitation is robot-assisted rehabilitation. Several research studies have reported significant improvements on robotic rehabilitation outcome based on clinical measures [3].

Rehabilitation evaluation during the motor rehabilitation process plays a key role in instructing robotic or manual therapy. Some research is focused on the multimodal analysis of patient performance, carried out by means of robotic technology and wearable sensors, and aims at providing quantitative measure of biomechanical and motion planning features of arm motor control following rehabilitation [4]. Rehabilitation therapists may take some measure in assessing stiffness or motion range of the patient's joints. However, it is difficult for them to evaluate quantitatively and objectively. For the purpose of objectively quantifying muscle tone, EMG activity was simultaneously recorded during wrist movements imposed by a torque motor [5]. Reference [6] introduces two robot devices for use in the rehabilitation of upper limb movements and reports the quantitative parameters obtained to the rate of improvement, thus allowing a precise monitoring of patient's recovery. Some experts make a study of clarifying the difference in muscle thickness of several muscles by the motor functions used in daily activity in adults with cerebral palsy of different levels of severity involvement [7]. But few wearable robotic devices are used for the recovery condition evaluation of the patient's hand internal muscles [8, 9]. A pneumatic muscle (PM) driven wearable quantitative muscle rehabilitation evaluation robotic hand device is developed and it has the potential of providing objective assessment of muscle function improvement from the therapy.

During the recovery process, the relationship of active and passive components of the patient's muscles may continuously change. Due to the varying relationship and the complex nonlinear dynamics of the device, it is difficult for us to build the overall mathematical model of the device for the model-based control. The BP Artificial Neural Network-based PID (BP-PID) control can solve this problem.

In summary, the contributions of this study include

1. Design a wearable robotic hand for quantitative muscle rehabilitation evaluation;
2. Implement the control of the wearable quantitative muscle rehabilitation evaluation robotic hand by using BP-PID.

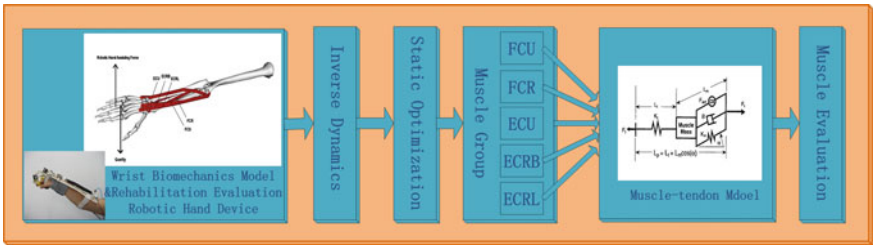


Fig. 208.1 Overall design flow chart of a wearable quantitative muscle rehabilitation evaluation robotic hand device

This paper is organized as follows: Section. 208.2 introduces the overall design of how to evaluate the recovery condition of stroke patients' wrist internal muscle groups, and describes the mechanics of the wearable rehabilitation evaluation robotic hand. Section 208.3 analyzes the control structure of the entire rehabilitation evaluation system. In Section. 208.4, the experimental results of the trajectory tracking control are presented. The results demonstrate the validity of the PM actuator and the BP-PID controller is superior to the conventional PID controller. In the final section, the conclusion of the current work and the possible future improvements are presented.

208.2 Overall Design and Mechanics of Safety Wearable Rehabilitation Evaluation Robotic Hand Device

208.2.1 Overall Design

The design goal is to develop a wearable robotic hand device that provides wrist muscle groups rehabilitation evaluation (Fig. 208.1). The wrist flexion–extension motion is driven by three kinds of forces including wrist five major internal muscles (flexors: FCU, FCR; extensors: ECU, ECRB, ECRL), the gravity of the patient's hand and the robotic hand assisting force. One degree of freedom wrist flexion–extension biomechanics model is built in Opensim [10]. The inverse dynamics determines the wrist joint net force and torque, and by use of static optimization it can resolve the net joint moments into individual muscle force at each instant in time according to the wrist model with muscles. During the rehabilitation process of the stroke patient's wrist, the relationship of active and passive muscle forces will change over time. When the patient gets better, voluntary movement will be stronger and it indicates that the range of active muscle forces becomes larger. Muscle-tendon model is used to investigate the parameter variance of the wrist five major muscles under this device. The quantitative relationship between active and passive forces can be measured [11]. Quantitative muscle rehabilitation evaluation is realized objectively.

Fig. 208.2 Materials of PM actuator and PM actuator made in lab

208.2.2 Mechanic Structure

The "Mckibben" pneumatic actuators used in the joint drive provide compliant actuation and thus augment the safety compared to the traditional rigid drivers (Fig. 208.2).

Figure 208.3 shows the prototype of the wearable rehabilitation evaluation robotic hand worn on the user's right hand.

208.3 Control System

Task-oriented repetitive movements can improve muscular strength and movement coordination in patients with impairments due to neurological diseases. The final goal of control scheme is to provide controllable and quantifiable assistance to measure patients' wrist muscles recovery condition. The sensors can feedback angle, pressure and force information for adaptation of the assistive force and quantitative evaluation of task performance. The angle sensor is used to measure the joint flexion and extension. In addition, there is also a pressure sensor integrated in the electric proportional valve to provide the local feedback. The real-time data acquisition card is responsible for sampling all sensory data. The data is exchanged via a USB connection between the data acquisition card and the host computer. The analog output of the card is connected to the electric proportional valve to control the PM actuator. The air transmission unit is an air-conditioning circuit which contains a filter valve and an electric proportional valve.

The complex nonlinear dynamics of the PM and many uncertainties in the device make this device driven by the PM actuator difficult to control, and it is very difficult for us to build the overall mathematical model of the device for model-based control. To realize the device controlled, a BP-PID controller is taken (Fig. 208.4), where $\Delta u(k)$ is the control increment at the kth sampling time, and e(k) is the deviation of the system at the kth sampling time.

The incremental PID control algorithm is described by (208.1) and (208.2). kp is the proportional gain. $k_i = k_p{*}T/T_i$ is the integral coefficient, and $k_d = k_p{*}T_d/T$ is the derivative coefficient. To achieve a good control effect, the BP-PID controller can find the best nonlinear relationship among k_p, k_i and k_d.

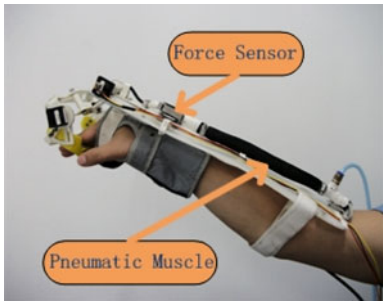


Fig. 208.3 The prototype of the wearable rehabilitation evaluation robotic hand

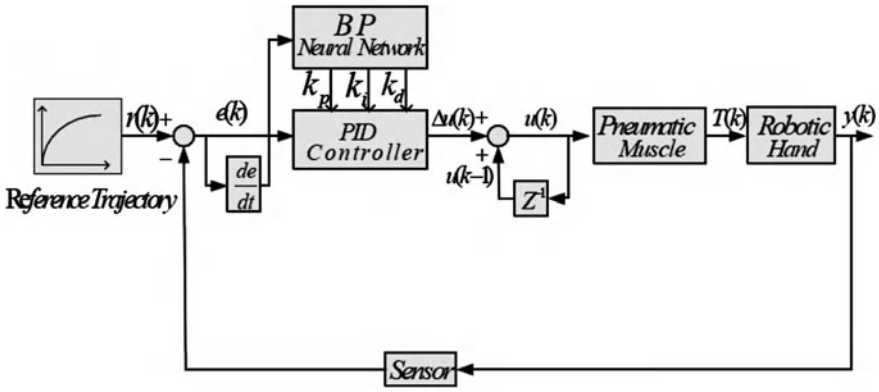


Fig. 208.4 The control structure of rehabilitation evaluation robotic hand with reference trajectory

$$u(k) = u(k-1) + \Delta u(k) \tag{208.1}$$

$$\Delta u(k) = k_p(e(k) - e(k-1)) + k_i e(k) + k_d[e(k) - 2e(k-1) + e(k-2)] \tag{208.2}$$

The BP artificial neural network includes three layers: input layer, hidden layer and output layer (Fig. 208.5).

Neural network weights are modified according to the gradient descent method (208.3), and the performance index function is shown in (208.4)

$$\Delta w(k) = -\eta \frac{\partial E(k)}{\partial w(k)} + \alpha \Delta w(k-1) \tag{208.3}$$

$$E(k) = \frac{1}{2}(r(k) - y(k))^2 \tag{208.4}$$

$g(x) = \dfrac{e^x}{e^x + e^{-x}}$ and $f(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$ are the activation function of output layer and hidden layer, respectively. For the output layer, it can be derived as follows:

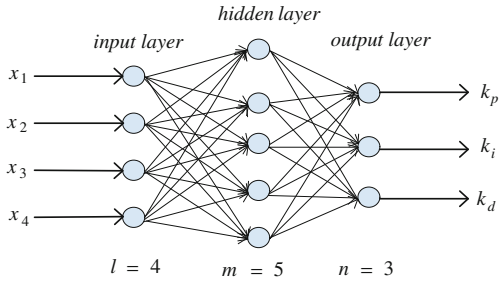


Fig. 208.5 The BP artificial neural network structure

$$\frac{\partial E(k)}{\partial w_{\mathrm{nm}}^{(3)}} = \frac{\partial E(k)}{\partial y(k)} \cdot \frac{\partial y(k)}{\partial \Delta u(k)} \cdot \frac{\partial \Delta u(k)}{\partial O_n^{(3)}(k)} \cdot \frac{\partial O_n^{(3)}(k)}{\partial \mathrm{net}_n^{(3)}(k)} \cdot \frac{\partial \mathrm{net}_n^{(3)}(k)}{\partial w_{\mathrm{nm}}^{(3)}(k)} \tag{208.5}$$

$$\frac{\partial \mathrm{net}_n^{(3)}(k)}{\partial w_{\mathrm{nm}}^{(3)}(k)} = O_m^{(2)}(k) \tag{208.6}$$

$$\frac{\partial \Delta u(k)}{\partial O_1^{(3)}(k)} = e(k) - e(k-1) \tag{208.7}$$

$$\frac{\partial \Delta u(k)}{\partial O_2^{(3)}(k)} = e(k) \tag{208.8}$$

$$\frac{\partial \Delta u(k)}{\partial O_3^{(3)}(k)} = e(k) - 2e(k-1) + e(k-2) \tag{208.9}$$

$$\Delta w_{\mathrm{nm}}^{(3)}(k) = \alpha \Delta w_{\mathrm{nm}}^{(3)}(k-1) + \eta \delta_n^{(3)} O_m^{(2)}(k) \tag{208.10}$$

$$\delta_n^{(3)} = e(k)\mathrm{sgn}\left(\frac{\partial y(k)}{\partial \Delta u(k)}\right) \frac{\partial \Delta u(k)}{\partial O_n^{(3)}(k)} g'\left(\mathrm{net}_n^{(3)}(k)\right) \tag{208.11}$$

Equations (208.10) and (208.11) present the output layer learning algorithm, and the hidden layer learning rule that are similarly available.

208.4 Experiment

208.4.1 Experimental Implementation

Figure 208.3 shows the experimental apparatus of the wearable rehabilitation evaluation robotic hand. The experimental apparatus includes the PM, the electromagnetic proportional valve, pressure regulating valve, force sensor, and the

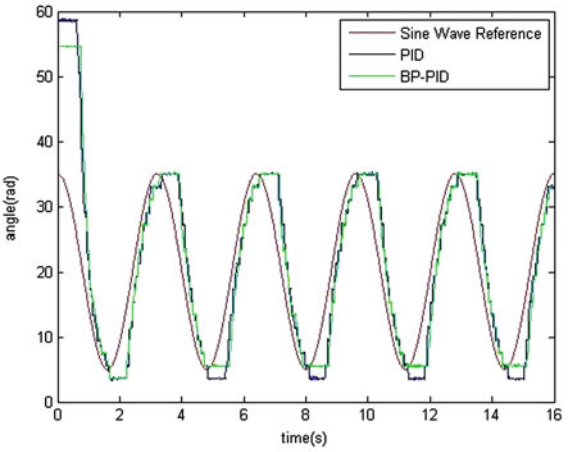


Fig. 208.6 The sine trajectory tracking result of wrist motion using the PID and BP-PID

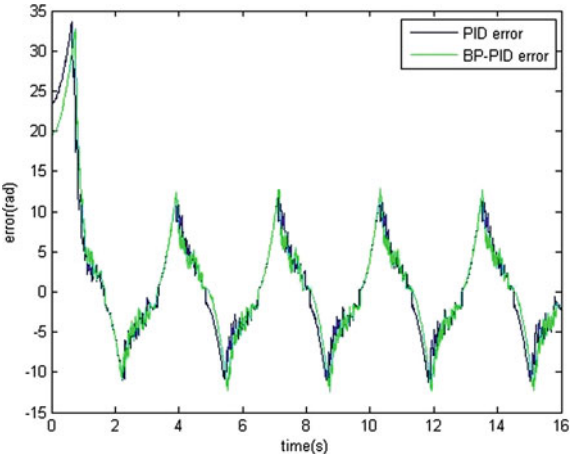


Fig. 208.7 The error tracking result of wrist motion using the PID and BP-PID

acquisition cards. The self-tuning parameters of BP-PID controller are: $k_p = 0.05$, $k_i = 0.005$ and $k_d = 0.02$. The parameters of conventional incremental PID controller are: $k_p = 0.02$, $k_i = 0.004$ and $k_d = 0.01$. The two control algorithms were implemented in the control system, respectively. Two control loops run independently for the PM actuator which drives the wrist joint to rotate. The controlled variable for each control loop is the measured angle at each joint. The software running on the PC was developed using Visual C++ 6.0.

208.4.2 Experimental Results

Some experiments were performed on a healthy subject, the performance of BP-PID and PID controllers were evaluated using the sine wave as the desired joint angle trajectory that should be tracked for the extension/flexion of the wrist joint.

The sine trajectory tracking results are shown in Figs. 208.6 and 208.7. The BP-PID parameters can be realized by using the error, it prevails over the PID in many performances, such as response time, flexibility, adaptability and control precision.

208.5 Discussion and Conclusion

In this paper we introduce a new wearable quantitative muscle rehabilitation evaluation robotic hand device driven by the PM actuator. This device helps rehabilitation therapists to evaluate muscle recovery condition of stroke patients quantitatively and objectively. Two kinds of control algorithms including BP-PID and PID are presented. The complex nonlinear dynamics of the PM and many uncertainties existing in the device make the device driven by PM actuator difficult to control. Therefore it is very difficult for us to build the overall nonlinear dynamics model of the device and realize the model-based control. In the future, many aspects need to be improved. On the one hand, we will search for much better control strategies for operating this device and try to model the nonlinear dynamics of the device. On the other hand, patient experiments will need to be carried out regularly, which will help stroke patients get accurate rehabilitation evaluation feedback and quick recovery.

Acknowledgment This work is partially supported by Stroke Clinic at Miaoling Town, Ezhou City, Hubei Province, China.

References

1. Lo AC, Guarino PD et al (2010) Robot-assisted therapy for long-term upper-limb impairment after stroke. N Engl J Med 362(19):1772–1783
2. Krebs H, Palazzolo J et al (2003) Rehabilitation robotics: performance-based progressive robot-assisted therapy. Auton Robots 15(1):7–20
3. Volpe BT, Lynch D et al (2008) Intensive sensorimotor arm training mediated by therapist or robot improves hemiparesis in patients with chronic stroke. Neurorehabil Neural Repair 22(3):305
4. Zollo L, Rossini L et al (2011) Quantitative evaluation of upper-limb motor control in robot-aided rehabilitation. Med Biol Eng Comput 49:1131–1144
5. Pisano F, Miscio G et al (1996) Quantitative evaluation of normal muscle tone. J Neurol Sci 135(2):168–172
6. Colombo R, Pisano F et al (2005) Robotic techniques for upper limb evaluation and rehabilitation of stroke patients. IEEE Trans Neural Syst Rehabil Eng 13(3):311–324
7. Ohata K, Tsuboyama T et al (2006) Measurement of muscle thickness as quantitative muscle evaluation for adults with severe cerebral palsy. Phys Ther 86(9):1231
8. Rijntjes M, Haevernick K et al (2009) Repeat therapy for chronic motor stroke: a pilot study for feasibility and efficacy. Neurorehabil Neural Repair 23(3):275

9. Timmermans AAA, Spooren AIF et al (2010) Influence of task-oriented training content on skilled arm-hand performance in stroke: a systematic review. Neurorehabil Neural Repair 24(9):858
10. Delp SL, Anderson FC et al (2007) OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans Biomed Eng 54(11):1940–1950
11. Zajac FE (1989) Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. Critical Reviews Biomed Eng 17(4):359

Chapter 209
Online Actor-Critic Learning for Motion Control of Non-holonomic Mobile Robot

Kai Wang, Yingmin Jia, Junping Du and Fashan Yu

Abstract This paper presents a control structure designed for non-holonomic mobile robots by an online algorithm based on policy iteration for learning the continuous-time (CT) optimal control solution with infinite horizon cost. The algorithm learns online in real-time to the solution of Hamilton–Jacobi–Bellman (HJB) equation which has been used for optimal control design. This method finds in real-time suitable approximations of both the optimal cost and control policy, while also guaranteeing closed-loop stability, which implemented as an actor/critic structure involves simultaneous continuous-time adaptation of both actor and critic neural networks (NNs). Simulation examples show the effectiveness of the new algorithm.

K. Wang (✉) · Y. Jia
The Seventh Research Division and the Department of Systems and Control, Beihang University (BUAA), Beijing 100191, China
e-mail: wangkai@ss.buaa.edu.cn

Y. Jia
Key Laboratory of Mathematics, Informatics and Behavioral Semantics (LMIB), Ministry of Education, SMSS, Beihang University (BUAA), Beijing 100191, China
e-mail: ymjia@buaa.edu.cn

J. Du
Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, School of Computer Science and Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
e-mail: junpingdu@126.com

F. Yu
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: yufs@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_209,

Keywords Lyapunov stability · Non-holonomic mobile robots · Neural networks · Actor-critic algorithms · Hamilton–Jacobi–Bellman equation

209.1 Introduction

Mobile robot is one of the well-known systems with non-holonomic constraints. Motion control of non-holonomic mobile robots has been receiving considerable attention for the last twenty years [1]. A detailed analytical study of the kinematics of mobile robots is found in [2]. On the basis of robot kinematics equations, the robot is controlled to follow the arbitrary path reference with a predefined velocity profile. But there is little work about optimal control of the non-holonomic system. From a mathematical view the solution of the optimal control problem is based on the solution of the HJB equation. Until recently, due to the intractability of this nonlinear differential equation for continuous-time (CT) systems, which form the object of interest in this Chapter, only particular solutions were available (e.g. for the linear time-invariant case, the HJB becomes the Riccati equation). For this reason considerable effort has been devoted to developing algorithms which approximately solve this equation [3, 4]. Far more results are available for the solution of the discrete-time HJB equation. Good overviews are given in [5].

In the linear CT system case, when quadratic indices are considered for the optimal stabilization problem, the HJB equation becomes the well-known Riccati equation and the policy iteration method is in fact Newton's method proposed by Kleinman (1968), which requires iterative solutions of Lyapunov equations. In the case of nonlinear systems, successful application of the PI method was limited until Galerkin spectral approximation methods were used to solve the nonlinear Lyapunov equations describing the policy evaluation step in the PI algorithm. Such methods are known to be computationally intensive. These are all offline methods for PI. The key to solving practically the CT nonlinear Lyapunov equations was in the use of NNs [3] which can be trained to become approximate solutions of these equations. In fact the PI algorithm for CT systems can be built on Werbos' actor/critic structure which involves two neural networks: the critic NN, is trained to approximate the solution of the nonlinear Lyapunov equation at the policy evaluation step, while the actor neural network is trained to approximate an improving policy at the policy improving step.

The Chapter is organized as follows. Section 209.2 provides the formulation of the optimal control problem, the online synchronous PI method and shows how to simultaneously tune the critic and actor NNs to guarantee convergence and closed-loop stability. Results for convergence and stability are developed using a Lyapunov technique. Section 209.3 presents simulation examples that show the effectiveness of the online synchronous CT PI algorithm in learning the optimal value and control for mobile robot.

209.2 Kinematic Model and Control Design

The mobile robot shown in Fig. 209.1 consists of a vehicle with two driving wheels mounted on the same axis and a passive self-adjusted supporting wheel, which carries the mechanical structure. The two driving wheels are independently driven by two actuators (e.g. DC motors). It is assumed that the mobile robot under study is made up of a rigid frame equipped with no deformable wheels and that they are moving in a horizontal plane.

209.2.1 Kinematic Model

For the mobile robot system considered here, the pure rolling and non-slipping non-holonomic condition states that the robot can only move in the direction normal to the axis of the driving wheels. The kinematic constraint can be written as

$$A(q)\dot{q} = 0, \quad A(q) = [-\sin\theta \quad \cos\theta \quad 0]. \tag{209.1}$$

The null space of $A(q)$ is given by the matrix $S(q)$. The vector $\dot{q}$ has to lie in this null-space, therefore

$$\dot{q} = S(q)\upsilon \quad S(q) = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix}, \ \upsilon = \begin{bmatrix} v \\ w \end{bmatrix} \tag{209.2}$$

where v and w are the linear velocity of the point C along the robot axis and angular velocity, respectively. System (209.2) is called the kinematic model of the robot.

209.2.2 Tracking Control Problem

Design a control law to control the motion of the mobile robot drive wheels. Suppose the mobile robot is required to follow a reference trajectory.

$$\begin{gathered} \dot{x}_r = \upsilon_r \cos\theta_r \quad \dot{y}_r = \upsilon_r \sin\theta_r \quad \dot{\theta}_r = w_r \\ q_r = [x_r \quad y_r \quad \theta_r]^T \quad \upsilon_r = [v_r \ w_r]^T \end{gathered} \tag{209.3}$$

The tracking error is expressed relative to the local coordinate frame fixed on the mobile robot as

$$e_p = [e_1 \quad e_2 \quad e_3]^T = T_e(q_r - q) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} (q_r - q) \tag{209.4}$$

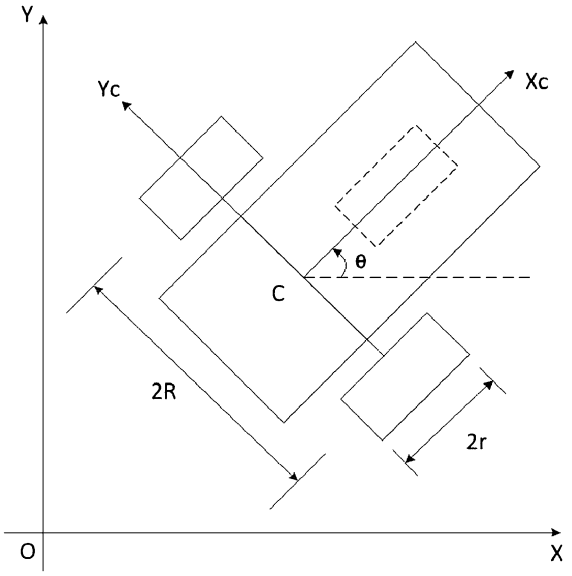


Fig. 209.1 A non-holonomic mobile platform. Both wheels have the same radius denoted by r. Two driving wheels are separated by 2R. Center of mass of the mobile robot is located at point C. The pose of the robot in the global coordinate frame OXY can be completely specified by three generalized coordinates $q = [x \quad y \quad \theta]^T$ where x and y are the coordinates of the point C in the global coordinate frame and θ is the orientation of the local frame CXcYc attached on the robot platform measured from X-axis

Therefore, the tracking error model is obtained by derivative of the error (209.4)

$$\dot{e}_p = \begin{bmatrix} we_2 - v + v_r \cos e_3 \\ -we_1 + v_r \sin e_3 \\ w_r - w \end{bmatrix} \tag{209.5}$$

For convenience, we chose new coordinates and inputs [7] as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} e_3 \\ e_2 \\ e_1 \end{bmatrix} \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} w_r - w \\ v - v_r \cos x_1 \end{bmatrix}$$

Equation (209.5) can be rewritten as

$$\dot{x} = f(x) + g(x)u \tag{209.6}$$

$$f(x) = \begin{bmatrix} 0 \\ w_r x_3 + v_r \sin x_1 \\ -w_r x_2 \end{bmatrix} \quad g(x) = \begin{bmatrix} 1 & 0 \\ -x_3 & 0 \\ 0 & 1 \end{bmatrix}$$

With the new coordinates $(x_1 \quad x_2 \quad x_3)$, the tracking problem is transformed into a stability problem. System (209.6) is referred as the error model of the tracking problem. In the remainder of this Chapter, the coordinates will be used in solving the tracking problem. Note that every coordinate transformation used above is invertible, and $(x_1 \quad x_2 \quad x_3) = (0 \quad 0 \quad 0)$ is equivalent to $x = x_r$, $y = y_r$, $\theta = \theta_r$. Thus, by invertability of coordinate transformation, if it converges to zero then the tracking problem is solved.

209.2.3 Optimal Control and the Continuous-Time HJB Equation

Consider the nonlinear time-invariant affine in system given by (209.6). Define the infinite horizon integral cost

$$V(x_0) = \int_0^\infty r(x(\tau), u(\tau))\mathrm{d}\tau \tag{209.7}$$

where $r(x,u) = Q(x) + u^T R u$ with $Q(x)$ positive definite and R is symmetric positive definite matrix.

Definition 1 *Admissible Policy* A control policy μ(x) is defined as admissible with respect to (209.8) on Ω, denoted by $\mu \in \Psi(\Omega)$, if μ(x) is continuous on Ω, $\mu(0) = 0$. $u(x) = \mu(x)$ stabilizes (209.6) on Ω and $V(x_0)$ is finite.

For any admissible control policy $\mu \in \Psi(\Omega)$, if the associated cost function

$$V^\mu(x_0) = \int_0^\infty r(x(\tau), \mu(x(\tau)))\mathrm{d}\tau \tag{209.8}$$

is $\mathbb{C}^1$, then an infinitesimal version of (209.8) is the so-called nonlinear Lyapunov equation.

$$r(x, \mu(x)) + (\nabla V_x^\mu)^T (f(x) + g(x)\mu(x)) = 0, \quad V^\mu(0) = 0 \tag{209.9}$$

where ∇V_x^μ denotes the partial derivative of the value function V^μ with respect to x.

Defining the Hamiltonian of the problem

$$H(x, \mu, V_x) = r(x(t), \mu(t)) + (\nabla V_x)^T (f(x(t)) + g(x(t))\mu(t) \tag{209.10}$$

the optimal cost function $V^*(x)$ defined by

$$V^*(x_0) = \lim_{\mu \in \Psi(\Omega)} \left(\int_0^\infty r(x(\tau), \mu(x(\tau)))d\tau \right) \tag{209.11}$$

and satisfies the HJB equation

$$\min_{\mu \in \Psi(\Omega)} [H(x, \mu, \nabla V_x^*)] = 0 \tag{209.12}$$

Assuming that the minimum on the right hand side of (209.12) exists and is unique then the optimal control function for the given problem is

$$\mu^*(x) = -0.5\, R^{-1} g^T(x) \nabla V_x^* \tag{209.13}$$

Inserting this optimal control policy in the nonlinear Lyapunov equation we obtain the formulation of the HJB equation in terms of ∇V_x^*

$$Q(x) + (\nabla V_x^*)^T f(x) - 0.25(\nabla V_x^*)^T g(x) R^{-1} g^T(x) \nabla V_x^* = 0, \quad V^*(0) = 0 \tag{209.14}$$

In order to find the optimal control solution for the problem one only needs to solve the HJB equation (209.14) for the value function and then substitute the solution in (209.13) to obtain the optimal control. However, due to the nonlinear nature of the HJB equation finding its solution is generally difficult or impossible

209.2.4 Online Actor-Critic Learning

Assume there exist ideal weights W_1 such that the value function $V(x)$ is approximated as

$$V(x) = W_1^T \phi_1(x) + \varepsilon(x) \tag{209.15}$$

And its derivative

$$\nabla V_x = \left(\frac{\partial \phi_1(x)}{\partial x} \right)^T W_1 + \frac{\partial \varepsilon}{\partial x} = \nabla \left(\phi_1^T W_1 + \varepsilon \right) \tag{209.16}$$

In these equation $\phi_1(x)$ is called the NN activation function vector, $\varepsilon(x)$ is the NN reconstruction error which is bounded. Using the NN value function approximation, considering a fixed control policy $u(t)$, the nonlinear Lyapunov equation (209.9) becomes

$$H(x, u, W_1) = W_1^T \nabla \phi_1 (f + gu) + Q(x) + u^T R u = \varepsilon_H. \tag{209.17}$$

where the residual error is bounded on a compact set due to the function approximation error is

$$\varepsilon_H = -(\nabla \varepsilon)^T (f + gu) \tag{209.18}$$

The weights of the critic NN which provide the best approximate solution for (209.17) are unknown. Therefore, the output of the critic neural network is

$$\hat{V}(x) = \hat{W}_1^T \phi_1(x) \tag{209.19}$$

where $\hat{W}_1$ are the current estimated values of the ideal critic NN weights W_1. The approximate nonlinear Lyapunov equation is

$$H(x, u, \hat{W}_1) = \hat{W}_1^T \nabla \phi_1 (f + gu) + Q(x) + u^T R u = e_H \tag{209.20}$$

Define the critic weight estimation error as $\tilde{W}_1 = W_1 - \hat{W}_1$.

Then $e_H = -\widetilde{W}_1^T \nabla\phi_1(f + gu) + \varepsilon_H$.

Given any admissible control policy u, it is desired to select $\hat{W}_1$. to minimize the squared residual error $E_1 = 0.5\, e_H^T e_H$.

From (209.13) we know that the ideal control policy update is taken as

$$u_1(x) = -0.5R^{-1}g^T(x)\nabla V_1(x) = -0.5R^{-1}g^T(x)\nabla\phi_1^T(x)W_1 \tag{209.21}$$

with W_1 unknown. Therefore, define the control policy in the form of an action neural network which computes the control input in the structured form

$$u_2(x) = -0.5\, R^{-1}g^T(x)\nabla\phi_1^T(x)\hat{W}_2 \tag{209.22}$$

where $\hat{W}_2$ denotes the current estimated values of the ideal NN weights W_1. Define the actor NN estimation error as $\tilde{W}_2 = W_1 - \hat{W}_2$.

Theorem 1 *Let the dynamics be given by* (209.6)*, the critic NN be given by* (209.19) *and the control input be given by actor NN* (209.23)*. Tuning for the critic NN be provided by*

$$\dot{\hat{W}}_1 = -\alpha_1 \frac{\sigma}{(\sigma^T\sigma + 1)^2}\left(\sigma^T\hat{W}_1 + Q(x) + u_2^T R u_2\right) \tag{209.23}$$

where $\sigma = \nabla\phi_1(f + gu_2)$*, and assume that* $\bar{\sigma} = \sigma/(\sigma^T\sigma + 1)$ *is persistently exciting. Let the actor NN be tuned as*

$$\dot{\hat{W}}_2 = -\alpha_2\left[\left(F_2\hat{W}_2 - F_1\bar{\sigma}\hat{W}_1\right) - \frac{1}{4}\bar{D}(x)\hat{W}_2\left(\frac{\sigma}{(\sigma^T\sigma + 1)^2}\right)^T\hat{W}_1\right] \tag{209.24}$$

where $\bar{D}(x) = \nabla\phi_1 g R^{-1} g^T \nabla\phi_1^T$ *and F1* > 0 *and F2* > 0 *are tuning parameters. The closed-loop system state, the critic NN and error actor NN error are UUB.*

209.3 Simulation Results and Conclusion

The reference trajectory is given by $x_r = 1$, $y_r = v_r t$, $v_r = 0.5$ m/s, $\theta_r = \pi/2$. The design parameters Q and R are chosen as identity matrix with appropriate dimension. The initial states of the tracking error system are chosen as $[x_1(0) \quad x_2(0) \quad x_3(0)] = [1 \quad -1 \quad 1]$. Simulations results are provided in Fig. 209.2, to show the effectiveness of the proposed approach.

One can see that after 80 s the convergence of the NN weights in both critic and actor has occurred. This shows that the probing noise effectively guaranteed the PE condition. On convergence, the PE condition of the control signal is no longer needed, and the probing signal was turned off. After that, the states remain very close to zero, as required.

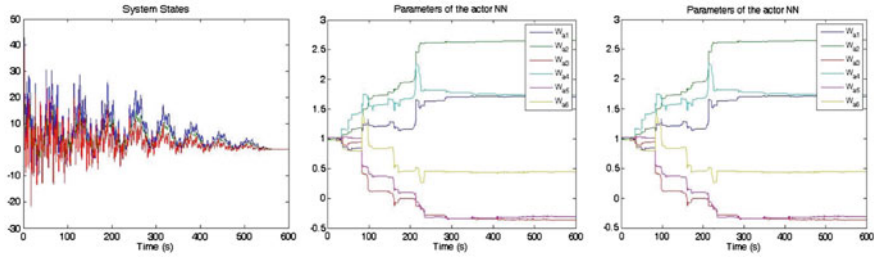


Fig. 209.2 The evolution of the system states, the critic and actor parameters convergence process, respectively

This Chapter presents a control structure designed for non-holonomic mobile robot based on its kinematic model by an online algorithm of for learning the continuous-time (CT) optimal control solution with infinite horizon cost. Our research efforts will be directed toward dynamic systems with modeling uncertainties and disturbances.

Acknowledgments This work was supported by the NSFC (60727002, 60774003, 60921001, 90916024), the MOE (20030006003), the COSTIND A2120061303) and the National 973 Program (2005CB321902).

References

1. Kolmanovsky I, McClamroch NH (1995) Developments in non-holonomic control problems. IEEE Cont Syst Mag 15(6):20–36
2. Alexander JC, Maddocks JH (1989) On the kinematics of wheeled mobile robots. Int J Robot Res 8(5):15–27
3. Abu-Khalaf M, Lewis FL (2005) Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. Automatica 41(5):779–791
4. Murray J, Cox C, Lendaris G, Saeks R (2002) Adaptive dynamic programming. IEEE Trans Syst Man Cybern 32(2):140–153
5. Si J, Barto A, Powel W, Wunsch D (2004) Handbook of learning and approximate dynamic programming. Wiley, Hoboken
6. Ioannou P, Fidan B (2006) Advances in design and control, adaptive control tutorial. SIAM, PA
7. Murray RM, Sastry SS (1993) Non-holonomic motion planning: steering using sinusoids. IEEE Trans Automat Cont 38:700–716
8. Vamvoudakis KG, Lewis FL (2010) Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. Automatica 46:878–888

Chapter 210
Optimization Model and Algorithm for Modern Tram Operation Plan Adjustment

Kuo Guo, Hong-ze Xu and Peng Zhou

Abstract The new concept of virtual station was proposed to describe level crossing. On this basis, modern tram operation plan adjustment model was established with constraints as follows: dwell time; running time between stations; track interval; trams order. Integrated optimization objective of the model contained the total late time, the total number of adjustment and the influence on the traffic order. Improved genetic algorithm was used to solve the model, and the optimal solution obtained was the real-time operation adjustment plan of modern trams. Simulation results show that the model and algorithm can meet operational needs.

Keywords Modern tram · Operation plan adjustment · Optimization model · Genetic algorithm

210.1 Introduction

Modern tram is developed on the basis of traditional tram. With high speed, high flexibility, low construction cost, low noise, no pollution, it is an important direction of urban transport development in the future [1]. Modern tram line is not

K. Guo (✉) · H. Xu · P. Zhou
Beijing Jiaotong University, Beijing 100044, China
e-mail: 09120384@bjtu.edu.cn

H. Xu
e-mail: hzxu@bjtu.edu.cn

P. Zhou
e-mail: peng.zhou1210@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_210,

completely closed, so operation plans have to be adjusted in real time due to a lot of outside interferences. Modern trams are now manually adjusted through the vehicle phones, so safety and operational efficiency are difficult to guarantee. Automatic operation plan adjustment will help improve operational efficiency and reduce the dispatcher's intensity.

Similar to the traditional rail transit, the modern trams run in the fixed tramway in accordance with operation plan. Although the traditional study on automatic train regulation has made some achievements [2–6], these methods are not fully applicable to the automatic plan adjustment of the modern trams. Because modern tram lines will be inevitable to cross the highway, there are level crossings. When operation plans of modern trams are compiled, green phase of level crossing must be fully considered.

Based on the characteristics of the modern tram line, this paper presents the concept of virtual station and describes the level crossing with it. On this basis, operation plan adjustment model is established and solved by improved genetic algorithm. Simulation tests are completed to prove the availability of the model and the algorithm.

210.2 Modern Trams Operation Plan Adjustment Model

Definition 1 If modern tram line goes through a level crossing, and there are traffic lights which command vehicles running on the direction of modern tram line, define the level crossing as a "virtual station".

Definition 2 Define a station in modern tram line as a "real station".

Definition 3 Define the time that modern trams wait for traffic signal changing at a virtual station as "virtual dwell time" of the virtual station.

Definition 4 Define the time that modern trams wait at a real station as "real dwell time" of the real station.

Definition 5 Set n as the total number of the virtual and real stations. Define the n–n matrix P as the station attribute matrix. If the kth station is a real station, then the kth column elements of matrix P are all real number 1; if the kth station is a virtual station, then the kth column elements of matrix P are all pure imaginary number i.

Define A as original plan arrival time matrix, L as original plan departure time matrix, $\widetilde{A}$ as adjusted plan arrival time matrix and $\widetilde{L}$ as adjusted plan departure time matrix. Define

$$X = A \times P \quad Y = L \times P \quad \widetilde{X} = \widetilde{A} \times P \quad \widetilde{Y} = \widetilde{L} \times P$$

where X is original plan arrival time matrix, which contains attributes of stations; Y is original plan departure time matrix, which contains attributes of stations; $\widetilde{X}$ is adjusted plan arrival time matrix, which contains attributes of stations; $\widetilde{Y}$ is

adjusted plan departure time matrix, which contains attributes of stations. They can be expressed as follows:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mn} \end{pmatrix} \quad \widetilde{X} = \begin{pmatrix} \widetilde{x}_{11} & \cdots & \widetilde{x}_{1n} \\ \vdots & \ddots & \vdots \\ \widetilde{x}_{m1} & \cdots & \widetilde{x}_{mn} \end{pmatrix}$$

$$\widetilde{Y} = \begin{pmatrix} \widetilde{y}_{11} & \cdots & \widetilde{y}_{1n} \\ \vdots & \ddots & \vdots \\ \widetilde{y}_{m1} & \cdots & \widetilde{y}_{mn} \end{pmatrix}$$

where m is the number of planned modern trams, x_{ij} describes arrival time at the jth station in the original plan, y_{ij} describes departure time at the jth station in the original plan, $\widetilde{x}_{ij}$ describes arrival time at the jth station in the adjusted plan, $\widetilde{y}_{ij}$ describes departure time at the jth station in the adjusted plan. x_{ij}, $\widetilde{x}_{ij}$, y_{ij} and $\widetilde{y}_{ij}$ contain attributes of stations as follows: If the jth station is real station, they are real numbers; if the jth station is virtual station, they are imaginary numbers.

210.2.1 Objective Function

Essence of the problem modern trams plan adjustment is based on the operation situation of the trams to redefine the operation plan. When the transport sector prepares schedules, the average passenger waiting time, interval capacity and other factors have been fully considered. Therefore, how to make the adjusted plans to run as close as possible the original plans need be considered, while the influence on urban road network traffic order is the smallest. In the adjustment process, the following three factors should be considered.

(1) Total late time

$$Z_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(f_1(\widetilde{x}_{ij} - x_{ij}) + f_1(\widetilde{y}_{ij} - y_{ij}) \right) \tag{210.1}$$

where $f_1(x)$ is a function defined on the complex number field,

$$f_1(x) = \begin{cases} x, & \text{the imaginary part of } x \text{ is } 0 \\ 0, & \text{the real part of } x \text{ is } 0 \end{cases} \tag{210.2}$$

$\widetilde{x}_{ij} - x_{ij}$ represents the late arrival time containing attributes of stations, $\widetilde{y}_{ij} - y_{ij}$ represents the late departure time containing attributes of stations.

(2) Total number of adjustment

$$Z_2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(f_2(f_1(\widetilde{x}_{ij} - x_{ij})) + f_2(f_1(\widetilde{y}_{ij} - y_{ij}))\right) \tag{210.3}$$

where $f_2(x)$ is a function defined on the real number field,

$$f_2(x) = \begin{cases} 1 & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{210.4}$$

(3) Influence on urban road network traffic order

$$Z_3 = \sum_{i=1}^{m}\sum_{j=1}^{n} f_3(\widetilde{x}_{ij} - x_{ij}) \tag{210.5}$$

where $f_3(x)$ is a function defined on the complex number field,

$$f_3(x) = \begin{cases} \big((x \cdot i)\%T_j - G\big)/2, & \text{the real part of } x \text{ is 0 and } (x \cdot i)\%T_j \geq G \\ 0, & \text{others} \end{cases} \tag{210.6}$$

where T_j and G represents the signal period and the remaining green light time of the jth station, which is a virtual station, i is a pure imaginary number, % is remainder sign. Modern trams can request changing traffic lights to gain priority by vehicle-mounted wireless communication devices. Modern trams running on schedule should meet the urban road network signal timing strategies, so that the trams arrive just in the green light phase of level crossings. If some trams arrive early, they cannot gain priority and have to wait for a green light at the crossing; if trams arrive late and the level crossing signal is red at that time, according to the circumstances of traffic flow of each direction, the priority would be provided for the trams with different strategy, including ahead of green light, green light extension and phase insertion. As early or extended green time gets longer, the influence on the traffic order becomes greater.

With the total late time, the total number of adjustment and the influence on the traffic order as the integrated optimization objective, objective function can be expressed as follows:

$$Z = \omega_1 Z_1 + \omega_2 Z_2 + \omega_3 Z_3 \tag{210.7}$$

where w_1, w_2 and w_3 are the corresponding weighting factors of Z_1, Z_2 and Z_3. Operation plan adjustment is equivalent to obtain the minimum value of objective function.

210.2.2 Constraints

Modern Tram operation plan adjustment need to meet the constraints as follows:

(1) Dwell time

a. if the *j*th station is a real station,

$$t_{\min} \leq \widetilde{y}_{ij} - \widetilde{x}_{ij} \leq t_{\max} \tag{210.8}$$

where $t_{\min}$ is the minimum dwell time at the *j*th station, $t_{\max}$ is the maximum dwell time at the *j*th station.

b. if the *j*th station is a virtual station,

$$0 \leq \widetilde{y}_{ij} - \widetilde{x}_{ij} \leq t_{\text{red}} \tag{210.9}$$

where t_{red} is the red light time at the level crossing.

(2) Running time between stations

$$\widetilde{x}_{i(j+1)} - \widetilde{y}_{ij} \geq \frac{S_j}{v_{\max}} \tag{210.10}$$

where S_j is the station spacing between the *j*th station and $(j + 1)$th station, $v_{\max}$ is the maximum speed between the *j*th station and $(j + 1)$th station.

(3) Track interval

$$\widetilde{x}_{(i+1)j} - \widetilde{x}_{ij} \geq \Delta t \tag{210.11}$$

$$\widetilde{y}_{(i+1)j} - \widetilde{y}_{ij} \geq \Delta t \tag{210.12}$$

where Δt is the minimum headway.

(4) Trams order

$$\widetilde{x}_{i(j+1)} - \widetilde{x}_{ij} \geq 0 \tag{210.13}$$

$$\widetilde{y}_{i(j+1)} - \widetilde{y}_{ij} \geq 0 \tag{210.14}$$

210.3 Model Solution

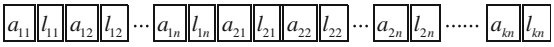
The above analysis shows that the modern tram operation plan adjustment problem is a multi-objective, multi-constraint optimization problem. When the traditional branch and bound method or optimization step-by-step method are applied to solve the problem, either the solution results or the convergence rate are not satisfactory. In this paper, an improved genetic algorithm is applied to solve the model.

210.3.1 Genetic Strategies

(1) Chromosome coding and population initialization

Modern trams operation plans are adjusted frequently, so operation plans of k trams in an hour are prepared for one time. The accuracy of running time between stations is allowed, when it does not affect to distinguish the phase of traffic lights. Taking into account that the minimum green time of traffic safety is 6 s [7], select 6 s as a quantitative unit. The arrival and departure time in operation plans of k trams in an hour are coded with integer encoding.

Define the arrival time of the first tram in the first station as zero, and encode the time offset. For example, if the first tram arrives in the first station at 8:00:00, then 8:05:00 is encoded as 50. Chromosome can be expressed as follows:



Chromosome length is $2 * n * k$.

Modern trams operation plans are adjusted because of trams delays, so the border of the breeding can be controlled according to the original plan and the total late time.

(2) Fitness function

Operation plan adjustment pursues the minimum of the objective function. Therefore, in accordance with single-valued, continuous, non-negative and maximizing design principles, the fitness function is design as equation (210.15) by using the penalty function method.

$$fit(Z) = \frac{1}{1 + Z + n_1F_1 + n_2F_2 + n_3F_3 + n_4F_4} \tag{210.15}$$

where Z is objective function value in the equation (210.7), F_1, F_2, F_3 and F_4 represent, respectively, punishment when they do not satisfy constraints (210.1)–(210.4). n_1, n_2, n_3 and n_4 represent, respectively, the number of times when they do not satisfy constraints (210.1)–(210.4). The values of F_1, F_2, F_3 and F_4 are based on operational condition by weighting the resolution and the calculated quantity of the algorithm.

(3) Population evolution

With the increase of fitness, the probability of reproduction becomes greater. If the total number of individuals in population is C, the probability of an individual to be copied is P_f as equation (210.16).

$$P_f = fit_i \Big/ \sum_{i=1}^{C} fit_i \tag{210.16}$$

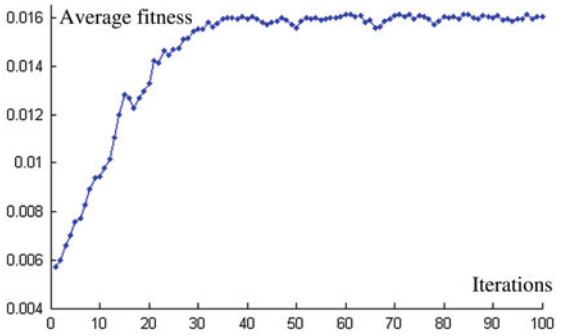


Fig. 210.1 After iteration, the average population fitness began at the 37th generation of convergence

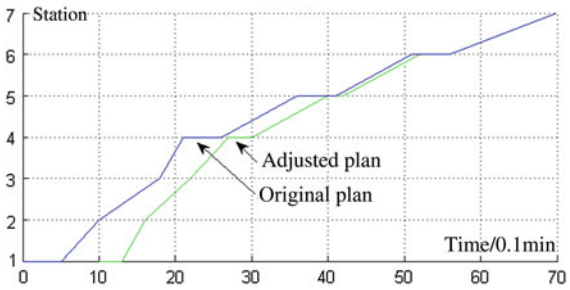


Fig. 210.2 The green line represented the adjusted plan, and the blue line above it represented the original plan. At the sixth station, the modern tram resumed to run on schedule

Make different chromosomes cross according to a certain probability P_c. The highest fitness in this generation of individuals is preserved, and inherits to the next generation without chromosome cross, to prevent damage to fine genes.

Make chromosomes mutate according to a certain probability P_m. It is a random change in some chromosomes, to prevent the initial population missing some important information.

Calculate iteratively until the maximum number of iterations. The individual with highest fitness is selected as the optimal solution, which is the adjusted operation plan.

210.3.2 Algorithm Flow

Step 1 Initialization, set population size C, chromosome length $2*n*k$, the maximum generation D, the crossover probability P_c, mutation probability P_m, the initial total late time D, the minimum headway Δt, the original plan "*Plan*";

Step 2 Set breeding range as [*Plan, Plan* + *T*] and generate initial population size of C randomly;

Step 3 Fitness calculation, calculate each individual fitness by equation (210.15). If $g > D$, the algorithm terminates;

Step 4 Copy operation, select individuals randomly to copy according to the reproduction probability P_f as equation (210.16);

Step 5 Crossover operation, choose from the individual parent to cross-match to get offspring according to crossover probability P_c. Parent of the largest individual fitness becomes offspring directly without crossover.

Step 6 Mutation, make chromosomes mutate according to a certain probability P_m.

Step 7 Set $g = g + 1$ and go to **Step 3**.

210.4 Solution Example

Simulate with the background the operating conditions of Tianjin Binhai line, which has a total of 14 stations and passes through 6 traffic lights. Station attribute can be expressed by a row of matrix P as follows:

$$P_i = [1\ i\ i\ 1\ 1\ 1\ i\ 1\ i\ 1\ 1\ 1\ i\ 1\ 1\ 1\ 1\ 1\ i\ 1]$$

The maximum operating speed of vehicles on the line is 70 km/h, and the running interval in the peak time is 7 min. The eight tram cars in the peak period of 7:00–8:00 will be adjusted. Since the Binhai line is located in the Dongting Road, where a "green wave" technology is used for traffic management, the virtual station dwell time are all 0. Calculate the influence of trams delays on urban road network traffic order, and set signal period $T = 60$ s, and green light time $G = 30$ s.

After some tests and analysis, set weights $w_1 = 1$, $w_2 = 2$, $w_3 = 1$, punishment values $F_1 = 20$, $F_2 = 20$, $F_3 = 10$, $F_4 = 10$, crossover probability $P_c = 0.6$, mutation probability $P_m = 0.01$, population size $C = 100$, minimum headway $\Delta t = 2$ min. The initial state was set to 1 min late in Taida station, so the initial total late time $D = 1$ min. The simulation results were shown in Fig. 210.1 and Fig. 210.2.

The example used the MATLAB programming language, in the configuration of the AMD dual-core processor, clocked at 2.2 GHz, 2G RAM computer running time was 3 s. It could meet the needs of real-time plan adjustment. Optimization model and algorithm were proved to be feasible.

210.5 Conclusion

This paper proposed a new concept of virtual station, established a modern tram plan adjustment optimization model and solved the model with improved genetic algorithm. Tianjin Binhai line was taken as the background to simulate the

situation of tram delays, and operation plans were adjusted automatically according to this model and method. The results show that the optimization model and the solving method meet the needs of automatic operation plan adjustment.

References

1. Xue M, Yang L, Cheng J (2010) Modem trams: characteristics & development both at home and abroad. Urban Transp China 6(6):88–91
2. Zhou L, Qin Z (1994) General algorithm and its computer implementation for train operation plan and adjust. J China Railw Soc 16(3):56–65
3. Araya S, Abe K, Fukumori K (1983) An optimal rescheduling for online train traffic control in disturbed situation. In: Proceedings of 22nd IEEE conference decision and control, IEEE. New York, pp 489–494
4. Zhou X, Zhu Y, Hu J (1999) Research on the algorithm for train regulation based on train running state derivation graph. J China Railw Soc 21(6):1–5
5. Li X, Cha W, Li Y (1999) Optimization for adjusting train operation on busy sections of trunk railway. J China Railw Soc 21(6):10–14
6. Zadeh LA (1994) Fuzzy logic, neural networks, and soft computing. Commun ACM 37(3):77–84
7. Texas Transportation Institute. Passer V. http://www.trafficware.tamu.edu

Chapter 211
Development of Novel Vehicle Operation Control System Elementary Prototype for Modern Tram

Zhibin Yuan, Hongze Xu and Peng Zhou

Abstract By analyzing the business process of modern tram, the functional requirements of vehicle operation control system (VOCS) were given. On this basis, an integrated design of VOCS was proposed, and then the interfaces between the subsystems in VOCS were designed. Based on the PC/104 hardware platform and Vxworks embedded operation system, the functional software of VOCS was developed. Above all, the elementary prototype of VOCS was complete. The debugging result shows that the design of VOCS is reasonable and the elementary prototype meets functional requirements of VOCS.

Keywords Modern tram · Vehicle operation control system · Elementary prototype

211.1 Introduction

Modern tram, which develops from the traditional tram, is a new kind of transport. Because of its low noise, low pollution, low energy consumption, and low construction costs relatively, modern tram has become an important part of urban public transport [1].

Z. Yuan (✉) · H. Xu · P. Zhou
Beijing Jiaotong University, Beijing 100044, China
e-mail: 09120356@bjtu.edu.cn

H. Xu
e-mail: hzxu@bjtu.edu.cn

P. Zhou
e-mail: peng.zhou1210@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_211,

Modern tram system is composed of line system, vehicle system, propulsion power-supply system, and operation control system (OCS) [2]. OCS, which links other three systems together, realizes the functions of operation control, safety protection and dispatching management. The VOCS, which realizes the functions together with other parts of OCS, is an indispensable part of OCS.

At present, research on the VOCS of modern tram is still in its infancy. The current VOCS is divided into two subsystems, automatic tram operation (ATO) subsystem and automatic tram protection (ATP) subsystem. They complete their own function independently and exchange data through communication interface. However, this kind of design increases the complexity of system architecture and reduces the security of system and the sharing of information between subsystems.

By studying the business process of modern tram and analyzing the functional requirements of VOCS, this paper presents a new design of VOCS that integrates ATO subsystem with ATP subsystem. Based on the PC/104 hardware platform and Vxworks embedded operation system, an elementary prototype was developed. Then a testing platform was built to prove the rationality and validity of the design of VOCS, and an ideal result was achieved.

211.2 VOCS Requirement Analysis

Modern tram OCS is composed of central OCS (COCS), wayside OCS (WOCS) and VOCS. The three systems work collaboratively to ensure that modern tram operates safely and efficiently. From the point of view of system design, business process and functional requirements are analyzed.

211.2.1 Business Process Analysis

According to the business scene, modern tram business process is divided into two stages: the tram depot and the main line.

Tram Depot. According to the operation plan, modern tram operates based on the indication given by signals, while VOCS supervises the running of tram all the time to ensure the safety of operation.

Main Line. In view of the fact that modern tram often operates in the block near the bus lane where other means of transportation share with modern tram and wayside signal equipments are relatively simple, the business process of modern tram is different from other rail transportation in main line [3, 4]. The business process is divided into six scenes: departure, close to the switch, far from the switch, close to the crossing, far from the crossing and arrival. The main business process of modern tram in main line is as shown in Fig. 211.1.

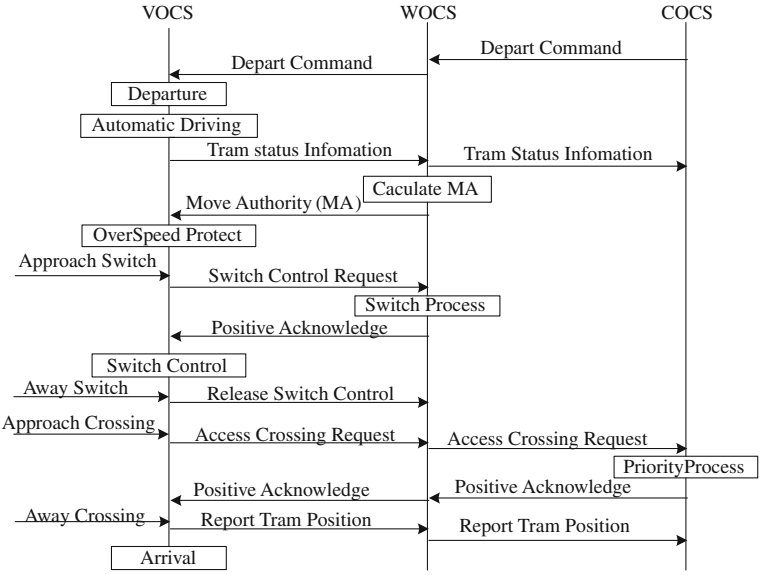


Fig. 211.1 Business process of modern tram in main line

211.2.2 Functional Requirements Analysis

From the analysis of modern tram business process, it is obvious that VOCS plays an important part in safety protection and automatic operation.

Based on the analysis, VOCS has the following functions:

VOCS receives the operation instructions from WOCS, and transfers the status of tram to WOCS in real time to assist WOCS in realizing the supervision of tram.

In view of the running distance, operation time, speed limit of line, characteristics of traction and comfort of passengers, VOCS calculates the optimal driving profile, which is called ATO profile. VOCS regulates the tram speed based on the ATO profile to ensure that the tram can arrive on time.

According to the movement authority from the WOCS and the line speed limit, VOCS calculates the ATP profile. VOCS supervises the speed based on the ATP profile to guarantee the safety of modern tram. If the actual speed exceeds the ATP profile speed at that location, an emergency brake will be taken to make the train stop in front of the danger point.

VOCS supervises the position of switch ahead of the tram in real time, and instructs the switch control system to move the switch directly.

At the crossing, modern tram has the priority to access. VOCS gains the priority by transferring a request to WOCS, which greatly enhance the efficiency of operation.

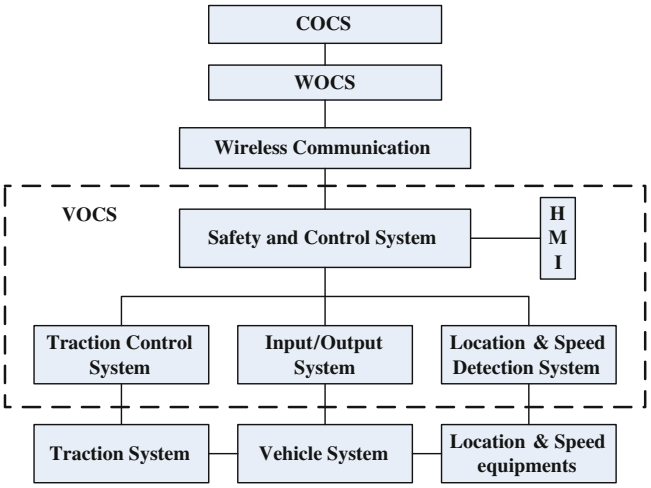


Fig. 211.2 Structure of novel VOCS

VOCS establishes the tram speed and location. Meantime, it also can compensate for the effects of measurement in accuracies on tram location and speed determination.

211.3 Novel VOCS Design

211.3.1 Function Decomposition

The traditional VOCS is divided into two subsystems: ATO subsystem and ATP subsystem. ATO subsystem is in charge of the functions related to the tram control [5], and ATP subsystem is in charge of the functions related to the safety protection [6]. ATO subsystem is restricted to ATP subsystem. The coupling relationship on the function provides a possibility for the integration design of VOCS.

According to the analysis of business process and functional requirements, the novel VOCS can be decomposed into five systems: safety and control system, input/output system, location and speed determination system, traction control system and human–machine interaction (HMI). The structure of novel VOCS is shown in Fig. 211.2.

Safety and control system is the core of the novel VOCS. It synthesizes the information from each system and generates control command to ensure the normal operation of VOCS.

Input/output system, location and speed detection system and traction control system are the executing agencies of safety and control system. Among them, input/output system collects the status of vehicle equipments, which mainly

includes braking system and door control unit, and drives them to execute relational operation according to the instructions of safety and control system; location and speed detection system establishes the speed and location of tram; traction control system regulates the speed of tram under the restriction of safety and control system.

HMI realizes the information interaction between driver and VOCS.

211.3.2 Interface Design

As Fig. 211.2 shows, VOCS has abundant internal and external interfaces, which undertake the information transmission among systems.

Internal interfaces include the interfaces between safety and control system input/output system, location and speed determination system, traction control system and HMI.

Safety and control system connects with vehicle system, location and speed determination system and traction control system via internal safety bus. Using this interface, safety and control system can transfer commands to other systems and collect tram status.

Safety and control system connects with HMI via Ethernet. Using this interface, safety and control system transfers auxiliary driving information, tram status and alarm information to HMI and receives the inputs of driver.

External interfaces include interfaces between VOCS, traction system, vehicle system and WOCS.

Safety and control system connects with wireless communication devices via standard RS232 serial. Using this interface, safety and control system transfers tram status information in real time to WOCS and receives command information from WOCS.

Input/output system connects with vehicle system via digital I/O. Using this interface, vehicle control system collects vehicle status, such as door status and braking status, and drives relational devices.

Traction control system connects with traction system via CAN bus. Using this interface, traction control system transfers traction control command to traction system and receives the status of traction system.

211.3.3 Software Design

Based on the function decomposition and interfaces design, the software of VOCS also consists of safety and control software, vehicle control software, HMI software, location and speed detection software and traction control software.

Safety and control software is the core of VOCS software. According to the functions, it can be divided into six modules: calculation of ATP profile,

calculation of ATO profile, speed supervision, switch control, crossing priority control and door control. Except that speed supervision and calculation of ATP profile are triggered cyclically, other modules are triggered through command.

Input/output software is used to deal with the information about vehicle equipments. One hand, it collects the status of vehicle equipments and transfers it to safety and control system. On the other hand, it drives vehicle equipments according to the instruction issued by safety and control system.

HMI software is designed to display auxiliary driving information (target distance, target speed, etc.), alarm information and tram status, and accept the driver's command.

Location and speed detection software calculates the speed and distance-to-go on the basis of the pulse information from the speed sensor.

Traction control software is to regulate the speed of tram in close-loop PID control by adjusting the size of tractive force and braking force, which is calculated in real time based on the difference between standard speed from ATO profile and actual speed.

211.4 Novel VOCS Elementary Prototype Development

Based on the design of VOCS software, the functional software is developed. Then the simulation platform is built to test and verify the functions of elementary prototype.

211.4.1 Software Development

The VOCS software is developed by using modularization principle. Among them, the functional software is developed based on PC/104 hardware platform and VxWorks embedded operation system; the HMI software is developed based on PC and Windows XP. Calculation of ATP profile and ATO profile are the most important part during the software development. The detailed process of them is given as follows.

ATP profile is calculated based on one-step braking model. The calculation of ATP profile can be classified into three steps: (a) Determine the target point, including target speed and target position; (b) Determine the permanent speed limit profile between target point and current position of tram; (c) Calculate the ATP profile in iterative method from the target point to current point according to one-step braking mode.

ATO profile is obtained by real-time calculation. The calculation of ATO profile can be classified into five steps: (a) Determine the operation level according to the time given by WOCS, and different operation level corresponds to different speed; (b) Establish the speed limit information between target point and current

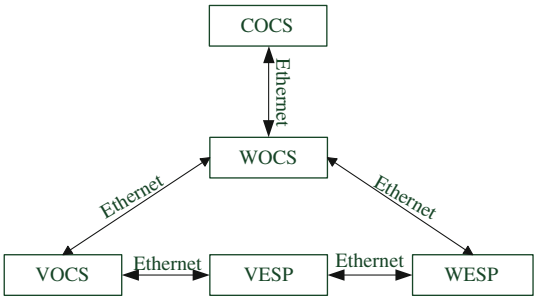


Fig. 211.3 Structure of simulation platform

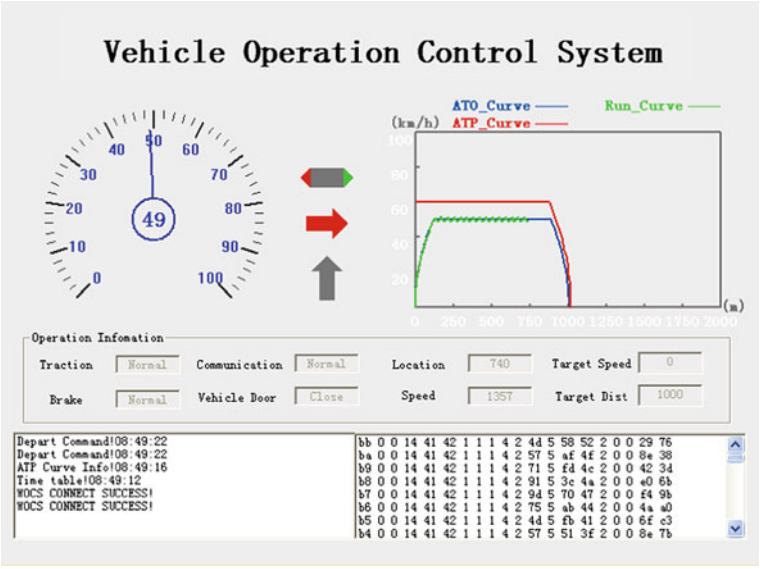


Fig. 211.4 VOCS testing interface

point, and it will generate a speed limit profile; (c) Determine the inlet speed of each segment according to the speed limit information, and it will generate a inlet speed profile; (d) Calculate the stopping brake profile according to the stopping strategies; (e) The ATO profile is generated by getting the minimum value among the above three profiles.

211.4.2 System Debugging

Based on the structure of modern tram OCS, the simulation platform is built for testing the functions of elementary prototype. The simulation platform consists of

COCS, WOCS, vehicle environment support platform (VESP) and wayside environment support platform (WESP), which are connected via Ethernet. The simulation platform is shown in Fig. 211.3.

System debugging is divided into two stages. The communication interfaces and logical functions are tested. On this basis, the VOCS functions are tested together with the simulation platform according to the business process and fault injection is made full use of to ensure the integrity of functions. The debugging results show that the design of novel VOCS is reasonable and complete in functions. VOCS tesing interface is shown in Fig. 211.4.

211.5 Conclusions

This paper presents a novel VOCS design that integrates ATO subsystem with ATP subsystem. On this basis, functional decomposition, interfaces design and software development are completed, and then the simulation platform is built to check the correctness and completeness of the elementary prototype in functions. The test results show that elementary prototype is reasonable and feasible in view of function.

References

1. Xue M (2008) Modern tram characteristics & development both at home and abroad. Urban Transp China 6(6):88–91
2. Wu W (2007) Characteristic and application prospects of tram system. Railw Stand Des 8(8):123–125
3. Wang L (2010) Modern tram signal systems. Railw Signal Commun 45(8):33–34
4. Schindler C, Schwickert M, Simonis A (2010) Structural safety of trams in case of misguidance in a switch. Veh Syst Dyn 48(8):967–981
5. Huang L, Tang T (2002) System analysis and design of automatic train operation on metro. J North Jiaotong Univ 26(3):36–39
6. Song Q, Chen K, Cui S, Xie X (2002) Simulation of overspeed protection of onboard ATP device in UMT. China Railw Sci 23(2):48–51

Chapter 212
Development of Novel Central Operation Control System Elementary Prototype for Modern Tram

Chengyuan Wu, Hongze Xu and Peng Zhou

Abstract From the view of Transportation Organization (TO) and Zone Control (ZC), business process of modern tram system was firstly analyzed. Furthermore, functional requirements of modern tram Operation Control System (OCS) were analyzed, and then an integrated design scheme of TO and ZC was proposed. On this basis, structure, interfaces, and software of novel Central Operation Control System (COCS) were designed. Based on industrial computer and Windows operation system, human–computer interaction software and dispatching management software were developed, and based on PC104 hardware platform and VxWorks embedded operation system, security protection software was developed. Finally, the test of the elementary prototype was completed. The test results show that the design of COCS is reasonable and feasible. The COCS elementary prototype realizes functions correctly and completely.

Keywords Modern tram · Central operation control system · Elementary prototype

C. Wu (✉) · H. Xu · P. Zhou
Beijing Jiaotong University, Beijing 100044, China
e-mail: 09120355@bjtu.edu.cn

H. Xu
e-mail: hzxu@bjtu.edu.cn

P. Zhou
e-mail: peng.zhou1210@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_212,

212.1 Introduction

Modern tram develops from traditional tram, which optimizes the urban traffic structure, and relieves urban traffic pressure. Modern tram is a new mode of transportation, which has advantages of large passenger capacity, low cost, high speed, and pollution-free [1].

Many foreign cities and a few cities in domestic have successively built modern tram system. Modern tram system has profound research significance and wide application prospects. Some experts continue to call for accelerating the development of modern tram system in China, especially in medium-sized cities.

Modern tram Operation Control System (OCS) coordinations with vehicles, traction power supply, line, etc., dispatches trams, ensures running safety and improves operational efficiency. Modern tram OCS is a central system of modern tram system. Transportation Organization (TO) and Zone Control (ZC) of modern tram OCS are present separately, but requirements analysis of modern tram OCS show that the integrated design scheme of TO and ZC is reasonable and feasible, which not only simplifies modern tram OCS, but also shares information and relieves network burden.

Through analyzing business process of modern tram system and functional requirements of modern tram OCS, the article puts forward a integrated design scheme of TO and ZC. On this basis, structure, interfaces, and software of Central Operation Control System (COCS) were designed. Then the COCS elementary prototype was developed based on existing hardware platform. Finally test of the elementary prototype was completed with the help of environment simulation, which was satisfactory.

212.2 The Modern Tram OCS Requirements Analysis

Business process of modern tram system is analyzed; furthermore, functional requirements of modern tram OCS are analyzed. On this basis, a new type of COCS is proposed, which combines TO with ZC.

212.2.1 Business Process Analysis

Figure 212.1 shows the business process of modern tram system. Modern trams drive from depot, and drive into main line through conversion rail, and finally quit operation [2]. The business process of modern tram system is analyzed from the view of TO and ZC, some of business process is omitted.

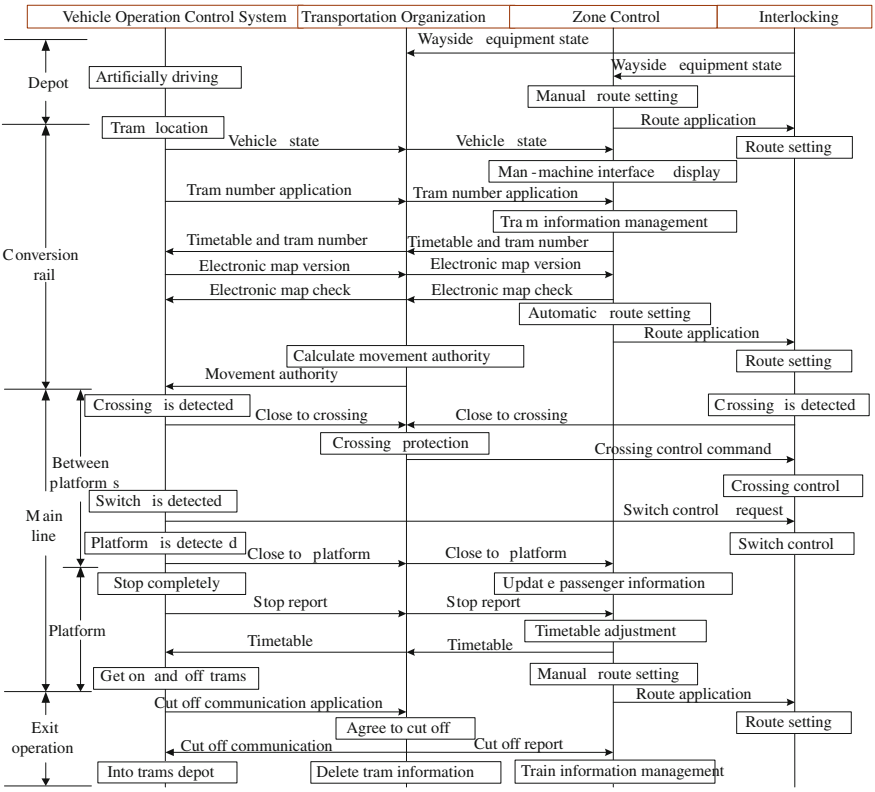


Fig. 212.1 Business process of modern tram system

212.2.2 Functional Requirements Analysis

Traditional tram operation control system is composed of TO, ZC, Vehicle Operation Control System (VOCS), and Interlocking (IL).

TO is the brain of modern tram OCS, which achieves the function of dispatching management, concrete contents are as follows:

TO displays wayside equipments state, tram state, tram diagram, record information, etc.

TO issues commands of manual adjustment, route request, and wayside equipments control. You can also edit the timetable to increase or decrease density of trams.

TO manages all trams about distributing timetable and tram number, checking the electronic map, and updating the passenger information system.

After timetable is deviated or edited artificially, TO regenerates timetable to meet requirements of operation.

According to the current location of trams, tram timetable, and predefined strategies of route, TO sets route automatically, and sends route application to IL.

ZC is the core of modern tram OCS, which achieves the function of safety protection, concrete contents are as follows:

In order to maintain a safe interval between trams, ZC authorizes trams to drive into and through a specific area according to a given operation direction.

ZC protects trams to drive through the crossing safely and effectively according to trams position and priority strategies of signal.

Requirements analysis of modern tram OCS show that the integrated design of TO and ZC is reasonable and effective [3]. Furthermore, modern tram does not have some wayside equipments such as platform screen doors and axle counters, and operation interval is bigger, speed is lower, so the integrated design of TO and ZC is simple and practical.

212.3 The Novel COCS Design

212.3.1 Function Decomposition Scheme

According to the requirements analysis of modern tram OCS and the concept of integration, the COCS is separated into human–computer interaction system, dispatching management system, and safety protection system [4].

Human–computer interaction system monitors the state of all trams and equipments. According to the requirements of operation, operators can edit the timetable, set route, and send vehicle adjustment commands.

Dispatching management system is logical processing mechanism of human–computer interaction system, which manages information of trams, and dispatches trams effectively according to automatic route setting and timetable adjustment.

Safety protection system is security mechanism of COCS, which protects trams to drive through crossing, switch, signal, and platform safely according to crossing protection and movement authority.

In the function decomposition scheme, security functions and unsafe functions are separated, which meets design requirements of safety system. Devices communicate with each other via internal network, so the system realizes information sharing. Dispatching management system and safety protection system ensure operation safely and effectively together.

212.3.2 Structure and Interfaces Design

Structure and interfaces of COCS are shown in Fig. 212.2. This system consists of human–computer interaction computer, dispatching management computer, and security protection computer, which realize the functions of human–computer

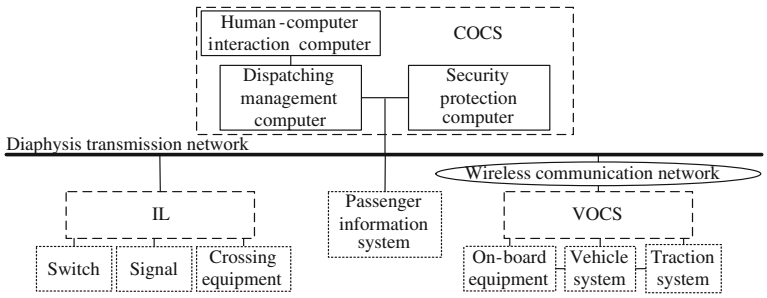


Fig. 212.2 Structure and interfaces of COCS

interaction system, dispatching management system, and safety protection system, respectively. Physical interfaces of COCS all use Ethernet.

Internal interfaces of COCS:

The interface of human–computer Interaction computer and dispatching management computer transmits information of control commands and equipments state.

The interface of dispatching management computer and security protection computer shares information from VOCS and IL.

External interfaces of COCS:

The interface of COCS and VOCS alternates information includes vehicle state, control commands, movement authority, etc.

The interface of COCS and IL is used to receive wayside equipments state and send commands of route application and equipments control.

The interface of COCS and passenger information system is used to update the passenger information.

212.3.3 Software Design

On the Basis of function decomposition scheme, structure and interface design, COCS software includes human–computer interaction software, dispatching management software, and security protection software. Software modules are divided into core functional modules and external interface modules.

The core functional modules of human–computer Interaction software include human–machine interface module and manual operation module; the core functional modules of dispatching management software include tram information management module, timetable adjustment module, automatic route setting module, and operation command processing module; the core functional modules of security protection software include movement authority module and crossing protection module.

Software modular design for dispatching management is as follows. After the commands of manual operation module are processed by operation command processing module, which are sent to the right VOCS or IL according to tram

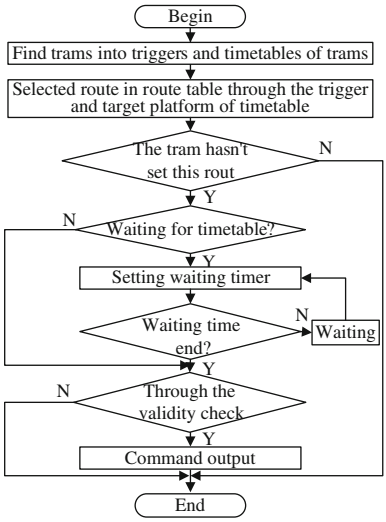


Fig. 212.3 Automatic route setting

information management module. Timetable adjustment module is triggered by time deviation of tram information management module, etc.

Software modular design for safety protection is as follows. After COCS executes automatic route setting module, IL turns and locks switches, and opens signals of switch protection. After COCS executes crossing protection module, IL opens signals of crossing protection. And movement authority module calculates movement authority limit according to equipments state from IL and tram information from tram information management module.

212.4 The Novel COCS Elementary Prototype Development

212.4.1 Software Development

In accordance with principles of modularization and fail-safe, security protection software is developed based on PC104 and VxWorks operating system, human–computer interaction software and dispatching management software are developed based on industrial computer and Windows operating system. Three core functional modules are as follows.

Timetable includes seven data: tram number, moving direction, departure platform, target platform, running time, arrival time, dwell time. COCS loads timetable before the operation of trams. Actual timetable is comparing with project timetable all the time. If timetable is deviation, COCS generates a new timetable according to strategies and algorithm of adjustment. If timetable is edited artificially, COCS also generates a new timetable [5].

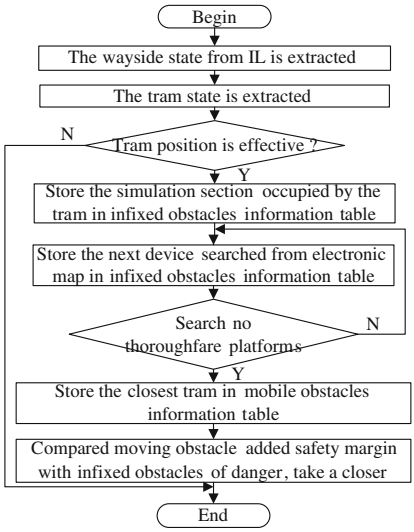


Fig. 212.4 Calculating movement authority limit

The flowchart of automatic route setting module is shown as the Fig. 212.3. Because modern tram has particularity, automatic route setting module is triggered by the position information from VOCS, and route is set between platforms and platforms or signals of switch protection.

COCS distributes movement authority to VOCS cyclically, the flowchart of calculating movement authority limit is shown as the Fig. 212.4. The tracks among platforms, signals, switches, and terminals of lines are defined as simulation sections. Fixed obstacles include simulation sections, platforms, signals, and switches, etc. Moving obstacles is the tram ahead. Fixed obstacles of danger include the no thoroughfare platforms which are not the end of route, parking signals, fault switches, etc. Electronic map is a network topology which is composed of simulation sections, platforms, signals, and switches.

212.4.2 Simulation Test

In order to test functions of COCS, simulation test platform was built, as shown in Fig. 212.5. Simulations of peripheral environment includes vehicle simulation system, and wayside simulation system developed with Visual C++. After the test of interface and logic functions of software, functions and business process of COCS were tested with the help of simulations of peripheral environment. The test results prove that the design of COCS is reasonable and feasible, and the elementary prototype meets functional requirements of COCS.

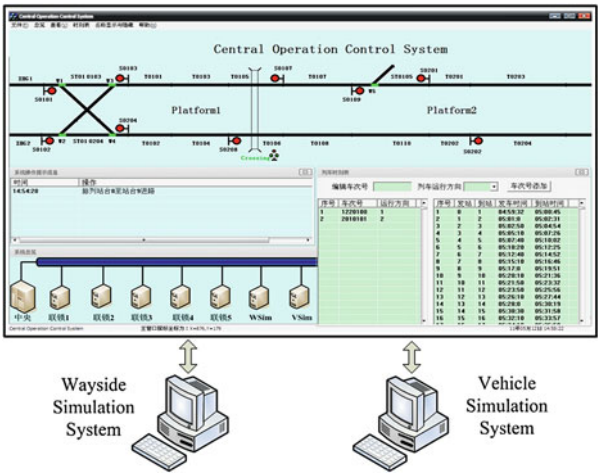


Fig. 212.5 Simulation test platform of COCS

212.5 Conclusions

This article puts forward the design scheme of the new type of COCS, and developed the COCS elementary prototype according to this scheme. The test results show that the design scheme is reasonable and feasible, the elementary prototype realizes functions of COCS correctly and completely. But this system has shortcomings, it is necessary to perfect the system function, optimize the system structure, and further research and practice.

References

1. Xue M (2008) Modem trams: characteristics & development both at home and abroad. Urban Transp China 6:88–91
2. Wu H (2009) Design and implementation of modern tram central operation control system elementary prototype. Beijing Jiaotong University, Beijing
3. Rosen D (2008) Concept of siemens modern tram. Metro and Light Rail, pp 1–4
4. Ambegoda ALATD, De Silva, WTS (2008) Centralized traffic controlling system for Sri Lanka railways. Information and automation for sustainability. Institute of Electrical and Electronics Engineering Computer Society, Piscataway, pp 145–149
5. Kanacilo EM, Verbraeck A (2007) Assessing tram schedules using a library of simulation components. Simulation conference, institute of electrical and electronics engineers incorporated, New York, pp 1878–1886

Chapter 213
Development of Modern Tram Ground Operation Control System Elementary Prototype

Zuoze Ma, Hongze Xu and Peng Zhou

Abstract For developing modern tram ground operation control system (GOCS), the business process of modern tram was analyzed. On this basis, functional requirements of the GOCS were analyzed. The design scheme of the GOCS, which included function decomposition, structural design, interfaces design and functional software design, was put forward. According to the design scheme, the GOCS elementary prototype that realized trams operation safety protection was developed. Through debugging with other subsystems of modern tram, expected results were gained which showed that the design scheme of the GOCS was rational and feasible.

Keywords Modern tram · Ground operation control system · Elementary prototype

213.1 Introduction

Developed on the basis of traditional tram, modern tram uses more low floor multi-modules hinged vehicles, resilient wheels, electric traction, has strong start-stop ability and is suitable for small curve radius and large gradient driving. For these reasons, modern tram is one of the most important developing directions of urban public transportation [1].

Z. Ma (✉) · H. Xu · P. Zhou
Beijing Jiaotong University, Beijing 100044, China
e-mail: 09120354@bjtu.edu.cn

H. Xu
e-mail: hzxu@bjtu.edu.cn

P. Zhou
e-mail: peng.zhou1210@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_213,

Except for the vehicle system, rail line system, traction power supply system and modern tram contains operation control systems, which is responsible for all trams operation control and safety protection. Modern tram operation control system includes central operation control system (COCS), vehicle operation control system (VOCS) and the ground operation control system (GOCS). The GOCS, which connects the COCS, the VOCS and trackside equipments, is responsible for safe tram interval and information forwarding.

At present, the domestic research for the GOCS is still in the initial stage, most tram lines prefer fixed block way to guarantee safe tram intervals. The operational efficiency is low.

Based on the business process of the modern trams, this paper analyzed functional requirements of the GOCS first, and then proposed the design scheme of the GOCS. According to the design scheme, the elementary prototype was developed. Through debugging with other subsystems of modern tram, anticipative results were gained.

213.2 Function Requirements Analysis of the GOCS

The COCS, the GOCS and the VOCS coordinates to realize the modern tram operation management and safety protection [2]. This section analyzed the business process of modern tram around the GOCS first, and then emphatically analyzed the functional requirements of the GOCS.

213.2.1 Business Process Analysis

As shown in Fig. 213.1, the business process of modern tram can be divided into three stages: from depot to main line, in the main line and from main line to depot.

From depot to main line. The GOCS establishes communications with the VOCS, and applies for train number to the COCS for trams.

In the main line. The GOCS transponds trams' status information received from the VOCS to the COCS, calculates moving authority limit (MAL) for the trams, handles with switch moving command and crossing the right-of-way request.

From main line to depot. The VOCS informs the GOCS that the tram will exit main line. The GOCS cuts off communication with the VOCS and deletes the information of the tram.

213.2.2 Function Requirements Analysis

The analysis of the business process of modern tram shows that the GOCS is the core of the modern tram operation control system. The GOCS is mainly

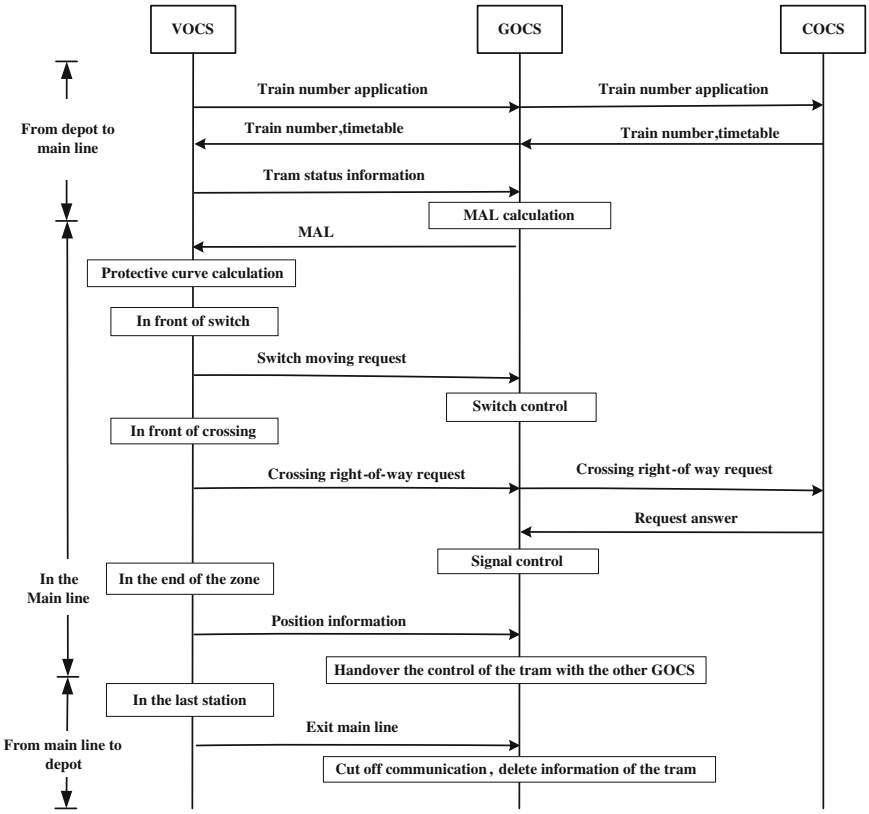


Fig. 213.1 Business process of modern tram

responsible for safe tram interval division, and information transponding. Specific content is as follows:

Train number application. The GOCS applies train number for every tram from depot, and manages all trams within the same zone according to train numbers.

Vehicle status information receive and forwarding. The GOCS transponds all trams status information that received from the VOCS to the COCS.

Trackside equipments status information acquisition and sending. The real-time GOCS supervises trackside equipments, and transponds those status information to the COCS for show.

MAL calculation. The GOCS calculates MAL for trams based on trams' status information and trackside equipments' status information, and in real-time detects the status of the trackside equipments within the scope of the MAL, and withdraws the end point of the MAL while obstacles were found. Obstacles include fault switches, red signal and so on.

Switch control. After receiving the switch moving command, the GOCS moves the switch to proper position and then locks it, while conditions have been met.

Crossing right-of-way request processing. After receiving crossing right-of-way request, the GOCS transponds the request to the COCS, and realizes the control of the signal according to the response from the COCS [3].

Tram control handover. When the line is long, the whole line is divided into multizone to be managed. Through the secure handover of the tram control, the GOCS realizes the management of all trams.

213.3 Design of the GOCS

213.3.1 Function Decomposition of the GOCS

Based on functional requirements of the GOCS and the characteristics of modern tram, the GOCS can be divided into three parts: control and safety protection system (CSPS), trackside equipment control system (TECS) and human–machine interaction system (HMIS).

The CSPS guarantees safe tram interval, processes commands from the COCS and the VOCS and transponds the status information of the trams and the trackside equipments. The GOCSs from adjacent zones hand over the control of the tram that trans-regional travels, and realize the management of all trams [4].

The TECS is responsible for handling the requests about switch movement and crossing right-of-way, and real-time supervision switches status and signals display.

The HMIS provides users one platform for artificial commands execution and the supervision of trams and trackside equipments, which is helpful for manual operation and more intuitively accesses rail lines and trams' status information.

213.3.2 Structural Design of the GOCS

According to function decomposition of the GOCS, the GOCS designed in this paper is composed of control and safety protection computer (CSPC), distributed I/O computer (DIOC) and human–machine interactive computer (HMIC), as shown in Fig. 213.2. The CSPC realizes the tram safe protection function, the DIOC takes charge of trackside equipments control and status information acquisition and the HMIC provides one human–machine interface.

While realizing the core control functions, the GOCS takes charge of a lot of commands and status information forwarding. These require it to be equipped with abundant internal and external interfaces.

Internal interfaces

Internal interface1. The CSPC connects with the HMIC and the DIOC through internal security information bus, realizes the interaction of command information, rail line information and trams status information.

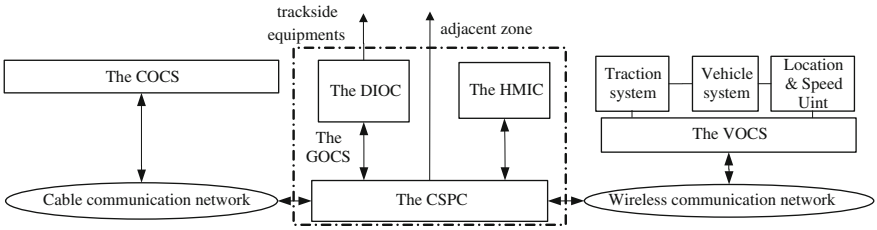


Fig. 213.2 Structure and interface diagram of the modern tram operation system

Internal interface2. Connected by independent internal security information bus, the adjacent zoning GOCSs realize the handover of the tram control.

External interfaces

Interface with the COCS. Connecting with the COCS by Ethernet, the GOCS receives commands from the COCS and transponds status information of trackside equipments and trams.

Interface with the VOCS. Connecting with the wireless communication system through the CAN bus, the GOCS realizes the information exchange with the trams, including commands, trams status information, MAL and so on.

Interface with Trackside equipments. Connecting with switches and signals by relay interface, the DIOC realizes the status collecting and control of the switch and the signal.

213.3.3 Software Design of the GOCS

According to the function decomposition and structure design of the GOCS, the GOCS functional software can be divided into core control functional software, human–machine interactive functional software and input/output software.

Core control functional software. Core control functional software realizes the major functions. Specific content as follows: train number application, trams status information receiving and forwarding, trackside equipments status information acquisition and sending, MAL calculation, switch control, crossing right-of-way request processing and tram control handover.

Human–machine interactive functional software. Human–machine interactive functional software is responsible for tram and rail line information display and artificial commands acquisition. Display information including: all trams status information, MAL information, executive order, trackside equipments status information and so on.

Input/output software. The input/output software drives corresponding systems to control trackside equipments to perform specified actions according instructions issued by the CSPC, and transfers the trackside equipments current status information to the CSPC.

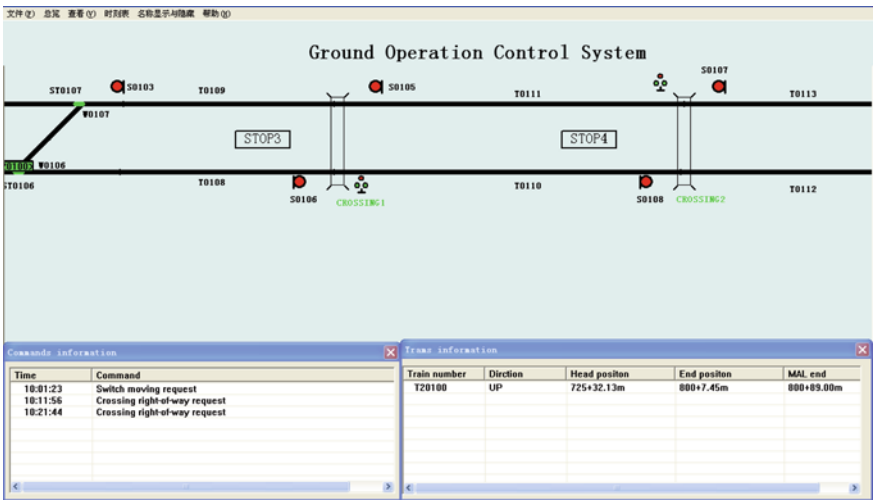


Fig. 213.3 Human–machine interface of the GOCS

213.4 Development of the GOCS Elementary Prototype

In this section, the GOCS functional software was developed according to the GOCS software design scheme. Through debugging with other parts of modern tram simulation platform established in lab, the simulation test of the GOCS elementary prototype was completed.

213.4.1 Software Development

Following software modular development principles, core control functional software, which was based on the hardware platform, minimize embedded system PC104, and suitable for Vxworks operating system, was developed to realize the core control function of the GOCS; human–machine interactive functional software, which was based on Personal Computer and suitable for Windows operating system, was developed to realize human–machine interactive function of the GOCS. Figure 213.3 shows the human–machine interface [5].

Now, introduce the development of the GOCS functional software with "MAL calculation" as example, which belongs to core control function. MAL calculating process is as follows:

As Fig. 213.4 shows, after communication with the GOCS is established, the VOCS begins to send tram current status information to the GOCS periodically. At the same time, the GOCS collects and supervises trackside equipments status real time. Based on the status information of the trams and trackside equipments, the

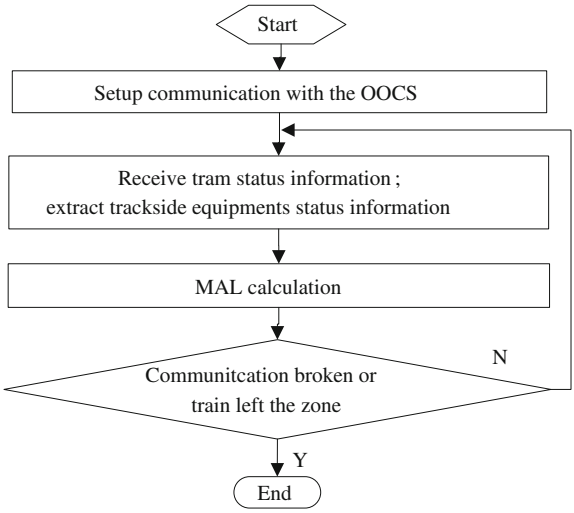


Fig. 213.4 MAL calculation flowchart

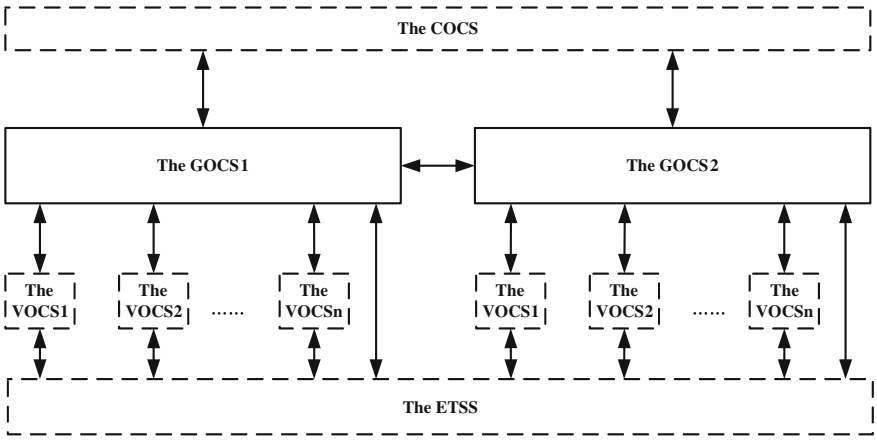


Fig. 213.5 Modern tram simulation platform

GOCS calculates MAL for trams until the communication between the GOCS and the VOCS were broken or the train left the zone.

213.4.2 Simulation Test

Modern tram simulation platform is established on the basis of business process of modern tram, as shown in Fig. 213.5. The actual operation parameters of trackside equipments and trams are provided by environment and tram simulation system

(ETSS). The GOCS elementary prototype developed by this paper connects with the ETSS, the VOCS and the COCS through cable. During debugging process of modern tram simulation system, interfaces and the logic function of the GOCS were tested first; second, the business completeness of the GOCS was tested. Business includes: train number application, MAL calculation, trackside equipments control and so on.

Simulation test results showed that the proposed design scheme of the GOCS was rational and feasible.

213.5 Conclusion

First of all, this paper analyzed the business process of modern tram around the GOCS, and then determined the functional requirements of the GOCS. On that basis, proposed the design scheme of the GOCS, and completed the development of the GOCS elementary prototype. Through debugging with other subsystem of modern tram gained the expected results which showed that the design scheme was rational and feasible.

References

1. Beijing Infrastructure Investment Co., Ltd (2010) Beijing modern trams technical standards [S] 2:24–29
2. Wang L (2005) New tram signal system [J]. Railw Signal Commun 45(1):33–34
3. Rosen D (2008) Concept of siemens modern tram [J]. Subways Light Rail (2):1–4
4. Li S (2005) Control of modern tramway and road system [J]. Urban Mass Transit 8(4):43–46
5. Xue M et al (2008) Modern trams: characteristics and development both at home and abroad [J]. Urban Transp China 6(6):88–96

Chapter 214
The Intelligent Multimedia Interpretation System Based on Embedded Technology

Jie Zeng, Meng Zhang and Hu Sheng

Abstract A novel multimedia interpretation system based on radio frequency identification (RFID) technology, embedded control technology and multimedia technology is provided in this paper. The novel system consists of three parts: tourists' visiting cards, RFID module and media player module. This paper describes the design work of hardware circuit, the control software of the RFID module and the media player module. This system was tested using different radio frequency (RF) cards. For different RF cards, the system outputs the corresponding responses: the buzzer outputs beep and the liquid crystal display (LCD) displays the results. The test results show that the intelligent multimedia interpretation system can realize the automatic recognition of tourists' identities, automatic selection of language, automatic playing of multimedia interpretation files. The system is characterized by self-assistance, convenient operation, intelligence, low power consumption and environmental protection, so it can help to improve the interpretation pattern and raise the intelligence and digital levels. This system can be widely used in tourism and exhibition.

Keywords Multimedia interpretation system · Radio frequency identification · ARM · WinCE operating system

J. Zeng (✉) · H. Sheng
School of Electrical and Information, Dalian Jiaotong University,
Dalian 116028, China
e-mail: zengyize@hotmail.com

M. Zhang
School of Graduate, Dalian Jiaotong University,
Dalian 116028, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_214,

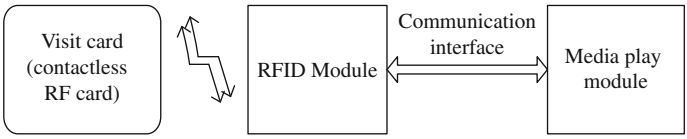


Fig. 214.1 Architecture of intelligent multimedia interpretation system

214.1 Introduction

Interpretation, an important way to impart information about exhibits and view spots, is a critical part of tourism and exhibition. Tourists' explanation is the most widespread interpretation mode, but this mode has the following drawbacks: single language, a stack of repetitive work, restrictions of tourists' autonomy. Motivated by enriching interpretation content, increasing languages and facilitating visit, an intelligent multimedia interpretation system is provided in this paper. Based on radio frequency identification (RFID) technology, embedded control technology and multimedia technology, the novel system can be widely used in exhibitions in buildings and museums.

214.2 System Architecture

Figure 214.1 shows the three parts of the intelligent multimedia interpretation system. The system can be divided into three parts: tourists' visiting card, RFID module and media player module. Visiting card is a radio frequency (RF) card which stores visitor's authorization and information about classification of language. It is an identity or a key to the multimedia interpretation files. RFID module can read the information from the visiting card and realize the media play function. The system consists of RF circuit and card reading control circuit. RF circuit, including 13.56 MHz RC522 chip and antenna, can realize the communication between RFID module and RF card. The card reader uses ATMEGA8 microcontroller as the control chip, which controls MF RC522 chip and communicates with the media player module.

The media player module uses ARM9 series S3C2440 chip as the core chip and consists of some peripheral equipments. Media player module, including multilingual interpretation files, can communicate with RFID module, and play the corresponding interpretation file, which is stored in the peripheral equipment.

The working process is as follows: after turning on the power button and starting the system, the media player module starts up the media player and initializes the RF card reader. When visitors put the RF card near the RF card reader, the RF card will be charged, and the wireless data link between RF card

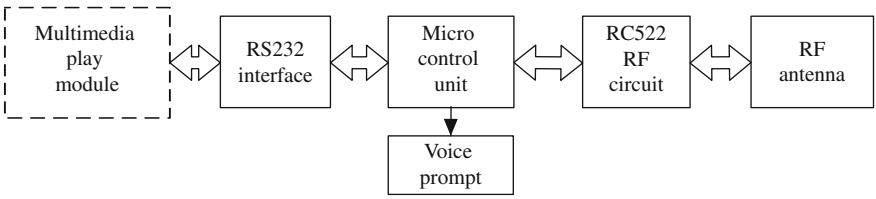


Fig. 214.2 Architecture of RFID module

and the card reader will be established. Then, the information inside the card will be transmitted to RFID module, and the RFID module will send the data to the media player module through a serial interface. After validating the visitor's identity, the system finishes reading and opens the corresponding media. After finishing the above operations, the system waits for the next visitor.

214.3 Hardware Design

214.3.1 Hardware Design of RFID Module

The hardware of RFID module includes control circuit, MF RC522 RF circuit, RF antenna, serial communication interface, and voice prompt circuit. The architecture of RFID module is provided in Fig. 214.2. The core part of the module is MF RC522 RF circuit, which is responsible for the reading and writing operation of RF card, including RF signal generation, modulation and demodulation, authentication, and overlap prevention [1].

The RF antenna circuit, including electromagnetic capability (EMC) high frequency filtering circuit and antenna matching circuit, is responsible for transmitting and receiving waves between the card reader and the RF card. RC522 RF circuit is the key part of the RF card reader. MF RC522, a product of Philips company, communicates with RF card and operates the RF cards under the control of the microcontroller.

Micro-control circuit is the control center of the RF card reader. ATMEGA8 microcontroller is employed to initialize the read and write module, initialize the serial communication module, communicate with the host computer by controlling serial communication interface, and receive commands from host computer. By doing so, the RF card can read and write automatically. Micro-control circuit connects with RC522 RF circuit using serial peripheral interface (SPI). The SP3232EEA chip is used in the serial communication circuit in order to establish the connection between visitor identification module and multimedia module. Furthermore, the voice prompt circuit is designed to prompt the visitor operating the RF card successfully [2].

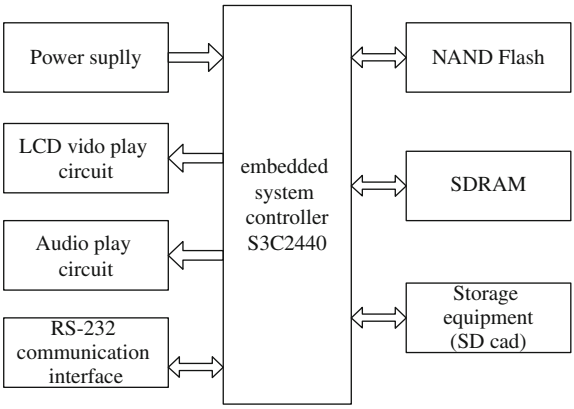


Fig. 214.3 Hardware architecture of multimedia play module

The working principle of the RFID module is as follows: after turning on the system, the microcontroller initializes each pin and beeps the buzzer. When entering the card searching stage, the microcontroller connects with the RC522 RF circuit by SPI interface. RC522 RF circuit transmits 13.56 MHz carrier wave by controlling RF antenna. A wireless data link will be established by transmitting the carrier wave when responded by RF card. RF antenna will return the information to RC522 after finishing the operation. Then, RC522 decodes and sends the data to microcontroller through SPI interface. The microcontroller sends the data to media player through serial communication interface.

214.3.2 Hardware Design of Media Play Module

S3C2440 chip with ARM920T core was selected as the controller of the media player module, which includes embedded controller S3C2440, power circuit, SDRAM memory, off-chip data memory NAND Flash, media player circuit, liquid crystal display (LCD) video player circuit, outside memory interface circuit, and so on. The hardware architecture of media player module is shown in Fig. 214.3.

S3C2440 chip is the kernel chip of the embedded controller. The chip possesses ARM920T core, SDRAM controller, NAND Flash controller, IIS bus controller, IIC bus controller, LCD controller, and serial communication interface. S3C2440 chip has good performance, so it can meet the requirements of intelligent language selection and multimedia interpretation.

Single chip 256M NAND Flash ship was used and two HY57V561620 chips with 64M SDRAM were selected in the system to provide sufficient space [3]. Furthermore, external storage devices, such as SD card, were used to provide massive memory space.

Power circuit provides stable power supply for S3C2240 system kernel, peripheral devices and RFID module card reader. In power circuit, LM117 and

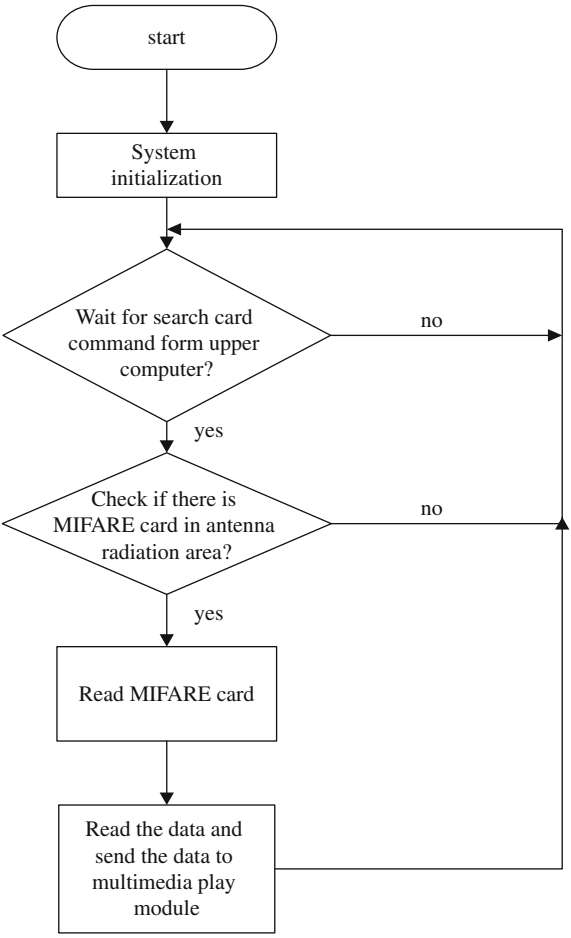


Fig. 214.4 Software flowchart of RFID module

MAX8860 were used to convert the voltage to 3.3 and 1.25 V, respectively. LCD video player circuit is a tool to display multimedia interpretation files. The audio player circuit uses the IIS bus to connect with UDA1341 voice chip to play the voice interpretation files. Besides, external multimedia module communicates and controls RFID module using serial communication interface.

214.4 Software Design

214.4.1 Software Design of RFID Module

RFID module, which automatically identifies visitors and selects the language, is the front part of the intelligent multimedia interpretation system. We programmed the RFID module to control the Atmega8 microcontroller, initialized the MF

RC522 chip and card reader, and operated the RF card. Figure 214.4 provides the software flowchart of the RFID module. After turning on the system, Atmega8 microcontroller and the MF RC522 chip are initialized, and the system finishes the self-test. When receiving the command from the host, the system switches to searching mode and waits for the visitors. When visitors put the RF card near the RF card reader, the card reader connects with the RF card and finishes the anti-collision, card selection, and the password authentication. Then, the card reader reads the data inside the E2PROM of visitor's RF card, and finishes the reading operation. The RFID module sends the data to media player module through a serial interface. After finishing the above operations, the system waits for the commands from the host.

214.4.2 Software Design of Multimedia Player Module

The software design of the multimedia player module includes BootLoader developing, WinCE platform building and transforming, and driver and application software developing. In this system, the WinCE5.0 was selected to operate the multimedia player module, and the Nbot was used as the BootLoader of the system. The functions, such as control RF module, receiving data from serial interface, and playing the interpretation files, are realized by transforming the WinCE operating system and developing the driver and application software [4].

The working process of the multimedia player module is as follows: the system starts the WinCE core from BootLoader, loads the file system, starts the media player, and initializes the serial interface. Then, the system controls the RF card reader using serial interface to search the RF card. The multimedia player module receives data from the card reader using serial interface, and then, controls card reader to stop searching, and switches to energy saving mode. The corresponding interpretation file is played by the media player. After playing the interpretation file, the multimedia player module controls the card reader to search the new RF card, and waits for the new valid data. The software flowchart of media play module is shown in Fig. 214.5.

214.4.3 Software Flowcharts of the System

The software flowchart of the system is provided in Fig. 214.6. After turning on the power of the system, RF module and the multimedia player module are initialized, and the multimedia player module sends the card searching command to RF module. Then, the card reader of the RF module starts searching the RF card. When visitors put the RF card near the RF card reader, the card reader responds immediately, then connects with RF card and reads the information from the card. The card reader sends the data to the multimedia player module of the host.

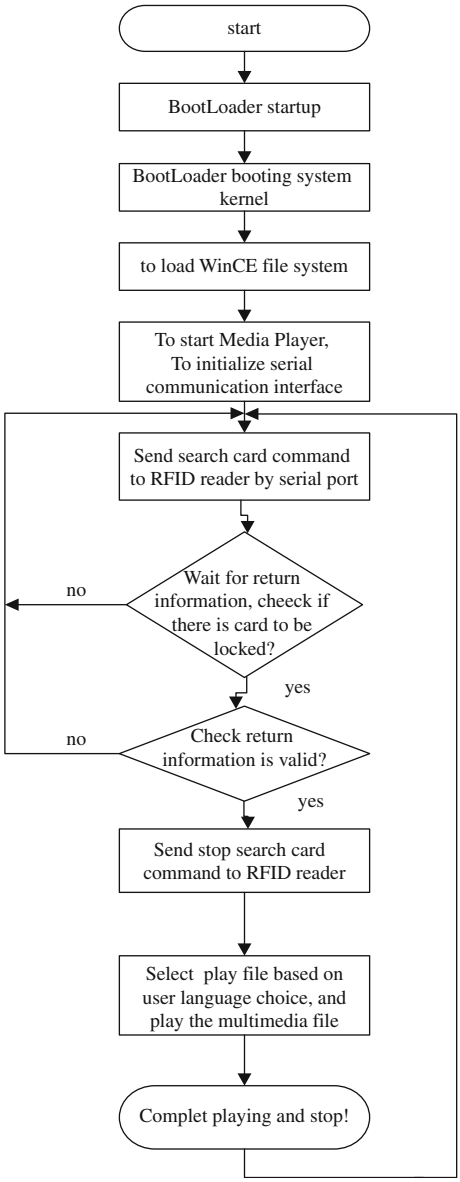


Fig. 214.5 Software flowchart of MP module

The multimedia player module receives data, and estimates the information. If the information is valid, then the multimedia player module sends the command to stop card searching. The module selects the corresponding language and plays the interpretation file according to the visitor's information. After playing the interpretation file, the system restarts, and switches to RF card searching state. The software design of each module is provided as follows.

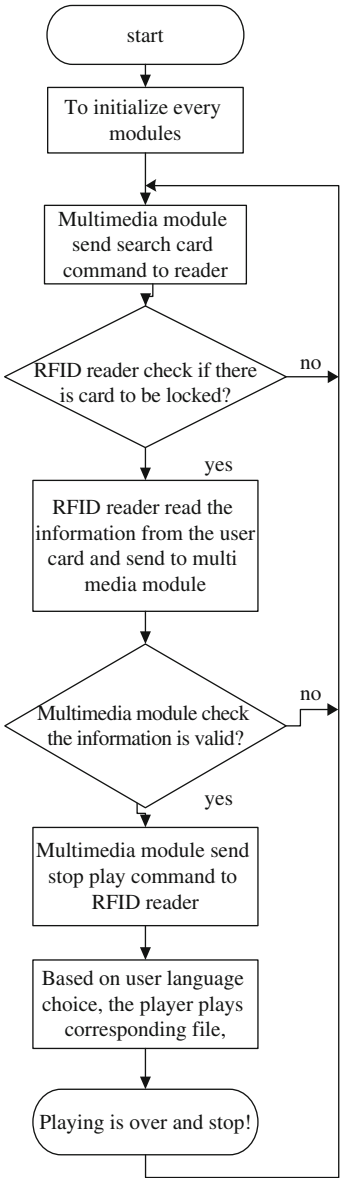


Fig. 214.6 Software flowchart of system

214.5 Experimental Verification

The intelligent system was tested using RF cards. The RFID module and multimedia player module were connected using serial interface, and the 0 card and 1 card were used to test the entire system. The physical model of the intelligent multimedia interpretation system is shown in Fig. 214.7 [5]. In the figure, part 1 is

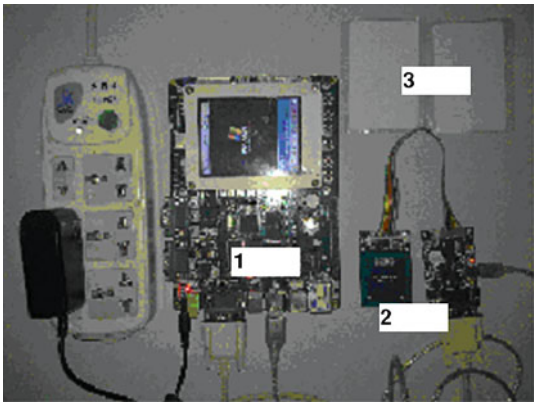


Fig. 214.7 Physical model of intelligent interpretation system

the multimedia player module, which is the embedded system with S3C2440 core; part 2 is the RFID module, which includes micro-controller circuit, MF RC522 RF circuit and antenna circuit; part 3 are two MIFARE RF cards, which store language information. Two parts are connected using serial interface, and the 3.3 V power is provided by the USB interface of the multimedia player module.

System testing method is provided as follows:

(1) After turning on the system, the initialization processes of multimedia player module and RFID module are tested. The initialization processes of multimedia player module is starting BootLoader, loading BootLogo, starting WinCE operating system, loading the media player, and then finishing the initialization. The beep of the buzzer suggests the completion of initialization.
(2) After initialization, the functions of RFID module are tested. The RF card moves slowly toward the antenna of the RF card reader. The distance between RF card and the antenna of the card reader is tested, and the response time of the RFID module is also recorded. After testing, the coverage area of the antenna is about 70 mm. When RF card is moved in this range, the buzzer beeps, the multimedia module starts working, and the interpretation file is played. The test results show that the RF module works normally.
(3) The multimedia player module is tested as follows: the RF card is moved toward the card reader, and the LCD displays the visitors' information. When RF card reader finds the RF card, the multimedia player module responds immediately, and the LCD displays the information about the interpretation file. The test results show that the multimedia player module works normally.

The above test results show that the intelligent multimedia interpretation system can basically realize the functions of tourists' identity recognition, language automatic selection, and multimedia interpretation files automatic playing. The system is stable and reliable, so it has goof use-value.

214.6 Summary

The intelligent multimedia interpretation system in this paper integrates advanced RFID technology, embedded control technology, and multimedia technology. The system realizes the functions of tourists' identity recognition, language automatic selection, and multimedia interpretation files automatic playing. Because of these new technologies, the novel system has advantages of intelligence, low power consumption, and environmental protection. The developing and the application of the system will improve the interpretation pattern and raise the intelligence and digitalization levels.

References

1. Zhanqing Y, Sujian L (2004) Theory and application of wireless RFID. Publishing House of Electronics Industry[M], Beijing, pp 8–12
2. Zeng J, Zhang M, Li G, Chen S (2010) Development of multilanguage selected automatic voice playing system based on RFID technology. In: International conference on intelligent control and information processing[C], Dalian
3. Chunlei D (2003) ARM architecture and programming. Tsinghua University Press[M], Beijing, pp 5–7
4. Dongquan Z et al (2006) Windows CE practical technology. Publishing House of Electronics Industry[M], Beijing, pp 2–30, 48–75
5. Zhang M (2011) Development of intelligent interpretation system based on ARM9[D]. Dalian Jiaotong University, Dalian

Chapter 215
Embedded Video Server Based on H.264 ASIC

Shimei Su, Yan Wang and Yi Liu

Abstract In order to achieve an embedded video server's flexibility and cut down the cost, a new solution based on Application Specific Integrated Circuit was introduced in this paper. Its hardware design and software hierarchical structure were introduced. An adaptive transmission strategy was proposed to fit the transmission networks. The results of experiments show that its main performance meets the practical needs and it can be used in some fields such as network video surveillance and hard disk video recorders.

Keywords H.264 · Video server · Embedded system · Adaptive transmission strategy

215.1 Introduction

Embedded video server becomes a trend of video monitoring development because of its high reliability, high speed, low cost, small volume, low power consumption, strong adaptability to environment, etc. [1–6]. Designing an embedded video

S. Su (✉) · Y. Wang
School of Electrical Engineering, Zhengzhou University,
Zhengzhou, China
e-mail: smsu@zzu.edu.cn

Y. Wang
e-mail: wystudyhard@126.com

Y. Liu
School of Information Engineering, Zhengzhou University,
Zhengzhou, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_215,

server system which can control flexibly, at the meantime can complete high-performance image compression with low cost becomes a significant issue.

This paper introduces a video server solution based on H.264 ASIC (Hi3510, embedded ARM and dual-core DSP). An adaptive transmission strategy for H.264 code stream is detailed.

215.2 The Video Server Architecture

Figure 215.1 illustrates the hardware architecture of the system. With a core of video processing chip Hi3510, it consists of a video input module, an audio input module, a storage module, a local display module, an external communication interface module and an outward expansible real-time clock module. The most important for the system design is exerting the processing ability of Hi3510, and the reliability of the system is taken into consideration at the meantime.

In video input module, four channels of video analog signals are imported to the video controller TW2834 by the form of composite video broadcasting signals (CVBS). TW2834 completes four-channel video signal acquisition and video analog signal decoding. Since only a 16 KB code memory and an 8 KB data memory are embedded in Hi3510, the storage module is designed in an expandable dynamic storage space hooked up two slice of SDRAM with an external bus interface (EBI). One slice is used for a flash memory of the operating system and application software, and the other is used for SDRAM storing encoded video data from Hi2510.

The external communication interface module, including RS232, RS485, Ethernet interface, USB interface, etc., is connected to a PTZ-camera controller. The video stream is sent by Ethernet interface and transferred through the Internet. Users can use a PC Client software or a Web browser to decode and play it.

215.3 Design of Embedded Server Software

The embedded server software can be divided into the operating system and the application software. An embedded Linux is used as the operation system.

With the increasing complexity of embedded systems, the software development may account for more than 70% of all the workload [7, 8]. To improve the embedded software module reusability and robustness, a hierarchical model is introduced in this system as Fig. 215.2.

The data layer supports database access, including reading, writing, deleting of video data index and modifying system configuration files, etc. It also provides audio and video codec API, including video decoding and video output, audio and video composite coding, etc.

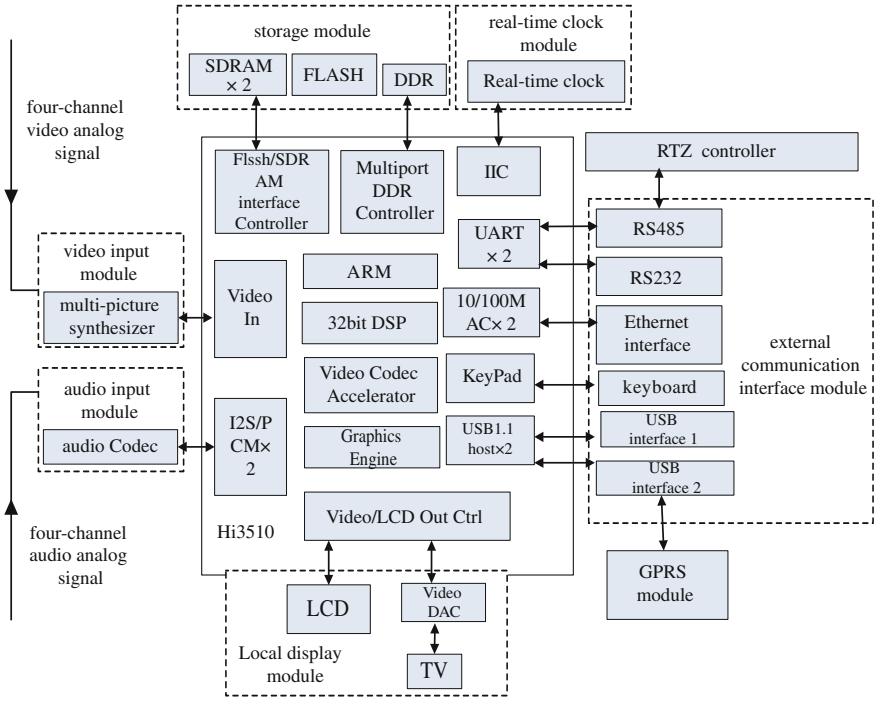


Fig. 215.1 Hardware architecture of video server

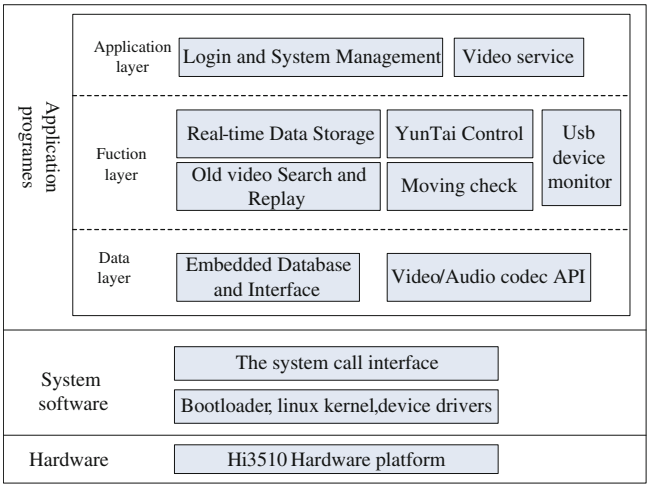


Fig. 215.2 Overall structure of system software

Function layer implements the core functions, including the real-time video data storing, the indexing and playback of the local storage videos, PTZ-camera controls and anomaly detections, etc.

Application layer is responsible for communication with the clients, and provides services such as browsing the current or the stored videos, sending abnormal alarming pictures, etc.

215.4 Adaptive Stream Transmission Strategy

Because of the diversity of clients (PC, mobile phone, PDA, etc.) and networks (LAN, ADSL, wireless network, etc.), the code rate control of the compressed video streams is necessary to adapt the different network conditions [9–12]. To realize smooth transmission, a frame-skipping control strategy is used to adjust the frame number of video coding.

215.4.1 Real-Time Protocol (RTP) Transmission Model

The RTP/RTCP protocols are employed to transfer video streams. RTP is responsible for media data transmission and RTCP for the test and control commands transmission. The RTP/RTCP transmission control model is shown as Fig. 215.3.

The client sends the receiving reports (RR) to the server periodically. The server analyses the RTCP feedback information and evaluates the current state of the network and client reception, then adjusts sending frame rate according to the following frame rate adjustment strategy.

215.4.2 Frame Rate Adjustment Strategy

The packet loss is used to evaluate the current network condition [13]. According to the feedback packet loss parameter from a client, the system chooses an appropriate transmission rates for it. The formula of packet loss is computed as follows:

$$F(n) = (C(n) - C(n-1))/((H(n) - H(n-1))) \tag{215.1}$$

In formula (215.1), $C(n)$ is the total number of lost packets after the n transmissions, $H(n)$ is the highest sequence number of the received packets after the n transmissions, $F(n)$ is the packet loss ratio.

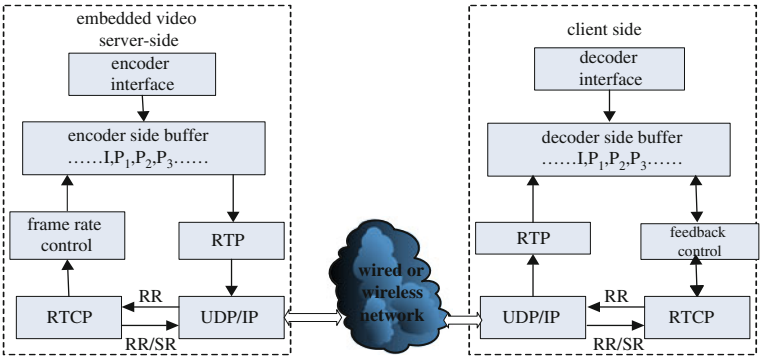


Fig. 215.3 RTP transmission control model

Table 215.1 Frame-skipping strategy

Adjusting levels	Frame rate (fps)	Frame sequence (1 s)	Note
1	1	I	Send one I frame
2	2	IP1	Send one I frame,P1
3	4	IP1P2P3	Send one I frame,P1…3
4	8	IP1P2……P7	Send one I frame,P1…7
5	16	IP1P2……P15	Send one I frame,P1…15
6	25	IP1P2……P24	Send one I frame,P1…24

$$f(n) = F(n)/(NTP(n) - NTP(n-1)) \tag{215.2}$$

In formula (215.2), packet loss rate *F(n)* is equal to the packet loss ratio divided by the Network Time Protocol (NTP) timestamp difference between the two RR.

Due to small changes in the background of video monitoring site, GOP format used in the encoder is IP frame format. It produces 25 frames per second (fps) and basic order is $IP_1P_2P_3P_4P_5…P_{24}$. Divide frame-skipping into six levels (as Table 215.1), and send RTP packets at different levels to terminals according to network performance. Selectively discarding some frames can improve availability of terminals' received data, then it can improve the subjective quality of images decoded.

According to packet loss rate, divide network states into three classes: light load, full load and congestion. The system maps them into three markings—increase, hold or decrease of frame rate. According to the markings, it can appropriately adjust sending frame rate. Set frame rate adjustable interal $[f_{min}, f_{max}]$, actual sending rate $f_{send} \in [f_{min}, f_{max}]$. Under heavy package loss, if the system reduces the sending rate quickly, it can improve sending frame rate.

Table 215.2 The experimental data from the system tests

Networks	Clinet device	Image resolution	Frame rate (fps)	Code rate (kbps)	Real time
LAN	PC	1 way D1(720*576)	25	2097	Y
		2 way D1(720*576)	25	2097*2	N
		1 way Half D1(720*288)	25	1048	Y
		2 way Half D1(720*288)	25	1048*2	Y
		4 way Half D1(720*288)	25	1048*4	N
		1 way CIF(352*288)	25	512	Y
		2 way CIF(352*288)	25	512*2	Y
		4 way CIF(352*288)	25	512*4	Y
Wireless LAN	Nokia N80	1 way CIF(352*288)	25	512	Y
		1 way CIF(352*288)	16	256	N

If (flag == DECREASE)

$$f_{\text{adjust}} = \max\{\alpha \times f_{\text{send}}, f_{\min}\}$$

Else if (flag == INCREASE)

$$f_{\text{adjust}} = \max\{\beta \times f_{\text{send}}, f_{\max}\}$$

Else $f_{ajust} = f_{send}$

where f_{adjust} is the frame rate to be adjusted, f_{send} the actual sending frame rate. Set the maximum frame rate $f_{\max} = 25$ fps, the minimum frame rate $f_{\min} = 1$ fps, diminishing factor $\alpha = 0.5$, increasing factor $\beta = 2$.

215.5 Experimental Results

The system has been completed and runs stably. Table 215.2 shows test results. When it works on the modes of 1 way D1, 2 way Half D1 and 4 way CIF, respectively, the fluent pictures can be received from a client. When it works on the modes of 2 way D1and 4 way Half D1, a small delay can be perceived (less than 2 s). When it works with 1 way CIF, under wireless LAN, the receiving delay on a cell phone (Nokia N80) becomes obvious with 25 fps because of the limits of the wireless channels and the decoding ability of the mobile phone.

215.6 Conclusion

A solution to the embedded network video server is proposed in this paper, which has flexible configuration, powerful expandability, high operation stability, etc. A hierarchical infrastructure is designed to improve development efficiency

and enhance each module's reusability. An adaptive code rate transmission control strategy is employed to adjust the video data transmission of different terminals with varied network states. With the above characteristics, this video server has a broad application prospect in the fields such as network video surveillance and hard disk recorders.

Acknowledgments Sponsored by nature science project of the education department of Henan province.

References

1. Li B, Ge BS, Li W (2004) A general DSP based multi-modal video encoder. Chin J Comput 27(12):1648–1656
2. Lu BS, Chen QM, Ding SJ (2006) Implementation of embedded network video server based on TMS320DM642. Comput Eng Des 27(13):2362–2364
3. Xiao ZH, Han QF, Zhu LQ (2006) Design and implementation of a remote video surveillance system based on ARM and DSP. Comput Eng Sci 29(9):53–55
4. Zhang J, Zhou DX, Hao FZ, Zeng HL (2006) Design and implementation of network video server based on AT2042. Comput Eng Des 27(24):4797–4809
5. Liang XR, Yan GP (2005) Introduction and analysis of typical schemes for embedded video surveillance server system. TV Eng 10:93–96
6. Wang F, Chen QB, Li L, Long KP (2006) The Introduction and analysis of typical schemes for digital video surveillance system. Ch Digit Cable TV 23:2289–2293
7. Xu XW, Wang ZY, Cao XY (2007) A kind of software architecture for embedded system development. Comput Sci 34(8):258–274
8. Liu G, Shao ZQ, Xiao LZ (2007) Some key issues of component based on development of embedded software. Appl Res Comput 24(3):62–65
9. Guo Y, Li GL, Xu Y (2004) Architecture and analysis about video transcoder. J Image Gr 9(10):1153–1159
10. Yu SF, Chen ZP, Zhuang ZW (2007) Review of video transcoding for network video applications. J Commun 28(1):111–118
11. Xv L, Gao W, Ji XY, Zhao DB (2008) A rate control algorithm for scalable video coding. Chin J Comput 31(7):1176–1180
12. Zhou B, Bi QQ (2004) A network adaptive transmission strategy of MPEG video stream. Comput Eng 30(16):160–162
13. Gao X, Shen SB, Gu GQ (2000) Analysis of network multimedia real-time transport protocol. Appl Res Comput 1(17):6–8

Chapter 216
A SoC Design with Delta MPU Cores

Limin Liu

Abstract The SoC design with delta MPU cores is a typical intelligent solution for embedded systems. In this paper, a design of SoC with multi-core is discussed. There are three MPU cores, operation MPU, backup MPU and monitor MPU, with delta structure in the SoC. The self-repairing function of the SoC is schemed through the back-up MPU to replace operation MPU. In order to transfer two MPU cores for operation of SoC, the input and output of the operation core and back-up core should be similar. The SoC is also called DRSoC, Dynamically Reconfigurable SoC. For some high reliability applications, the SoC design is useful.

Keywords Embedded systems · SoC · Reconfiguration · Multi-core

216.1 Introduction

The concept of SoC, System on a Chip, is to integrate some application system on one semiconductor chip [1]. When a system is integrated on a chip, but not one or more PCB, the system should be with smaller size, higher reliability, faster operation and lower cost for some applications. SoC, therefore, is the evolution and advanced organization of embedded systems. In fact, SoCs may be considered as hardware [2]. Hardware to replace some software is a useful feature of SoC. For some specific applications, the replacement would obtain better operating efficiency.

There are two main fields, design and organization, of SoC research. The design is normally concerned to IC design and microelectronics. The organization research may

L. Liu (✉)
IT School, Institute of Embedded Systems, Huzhou University,
Huzhou 313000, Zhejiang, China
e-mail: liulimin@ieee.org

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_216,

be to build SoC in various architectures. For the organization, some functions or parameters of SoC can be modified. The SoC is called reconfigurable SoC, RSoC [3, 4].

Some SoC applications, such as aerospace, robots and control systems, ask higher security operations. The self-repairing, self-tuning or self-learning solution is required. In order to guarantee reliable operation of systems, dynamically updating concept for SoC organization or functions can be employed. A new kind of SoC, therefore, Dynamically Reconfigurable System on a Chip, DRSoC, is issued.

In order to reconfigure internal functions dynamically, hardware of RDSoC has to be built on PLD devices. The more popular PLD devices are FPGA, Field Programmable Gate Arrays and CPLD, Complex Programmable Logic Devices. With ISP, in-system programming, and IAP, in-application programming, some functions on PLD devices can be updated more easily.

When the dynamic reconfiguration is to repair some components or functions in SoC, the reconfiguration is a typical self-repairing procedure. There are different ways to approach DRSoC. Here, a self-repairing SoC with multi-core architecture is discussed.

216.2 The System on a Chip

A SoC is built on one IC chip. MPU cores are central nervous system of SoC [5]. The IPs, Intellectual Property modules, are essential function modules of SoC. The MPU cores and IPs can be designed by HDL on one PLD chip. HDL is hardware description language, a kind of software. PLD is some programmable hardware. The HDL programming on PLD may produce a new hardware with new functions [6]. When there is a good structure of the system and cooperation of HW and SW, a SoC solution can get a better benefit. Here, the architecture of a SoC with multi-MPU cores is discussed.

A common organization of SoC is illustrated as Fig. 216.1. A SoC is composed of MPU, memory and IPs. There are various IPs in a SoC, such as functional IP, interface IP and other IPs.

IPs are important components of SoC. For most application, SoC is a special system. Actually, MPU and memory are common. The special features of a SoC are shown from the IPs [7]. Many functions of SoC are designed with different functional IPs. The communication between the SoC and its peripherals is set up through some interface IP. When IP modules replace common circuits or software, the efficiency of SoC operation is better. The reason is the IP to be designed for an application specially. For example, MPU to write a number into external RAM, a peripheral, may need four or five instructions of software at least. The instructions include setting writing state, giving address, outputting the number and removing writing state. Sometimes, NOP instructions are required to match time between different components. If an IP of ex-RAM control is issued, similar operation just uses one instruction. The operation efficiency can be increased several times. Therefore, IPs are essential parts of SoC [8].

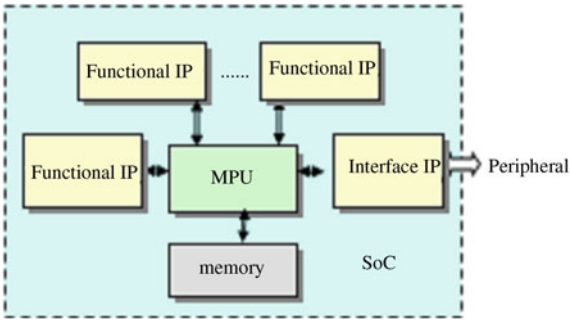


Fig. 216.1 Common organization of a SoC

Although IPs are important, the MPU or MPUs are cores of SoC. It is because IPs are driven by instructions. The MPU decides execution of instructions. Hence, some reconfiguration of SoC based on multi-MPU is important.

Sometimes, the reconfiguration can be considered as updating or repairing. There are several different updating, MPU core, interface, instruction and synthesized updating, for a SoC.

MPU core updating is to change or modify the embedded MPU core in a SoC. There are various MPU or MCU cores to be chosen, such as ARM, NIOS, R8051, CZ80 and so on. For a SoC based on FPGA, to replace the cores is essentially depended on the resource of FPGA [9].

Interface updating means to modify the interface of the SoC with user's requirements. When to add an I^2C interface for an employed SoC, for example, interface updating is used. In fact, the modification is concerned with to add or remove some interface IP for the SoC.

Instruction updating is a design for some instruction of SoC to be changed. Sometimes, users want to update instruction set for new functions. When to increase or decrease the size of instruction set for a SoC, instruction updating is required [10].

Synthesized updating can involve two or more above updating items. This updating shows more modifications for a SoC. With more updating contents, synthesized updating may be a common reconfigurable method for most SoC applications.

If to divide some levels for a SoC updating, synthesized updating takes the highest level; MPU core updating is middle; interface updating and instruction updating may be at the lowest position. Here, the MPU core updating is discussed.

216.3 SoC with Multi-MPU Cores

As common knowledge, the triangle is a shape with a wide base and cannot be easily knocked over. The triangle is a stability structure.

In order to set up a high reliability SoC with multi-core, a delta structure with three MPU is concerned. An organization of SoC with delta MPU cores is shown in Fig. 216.2.

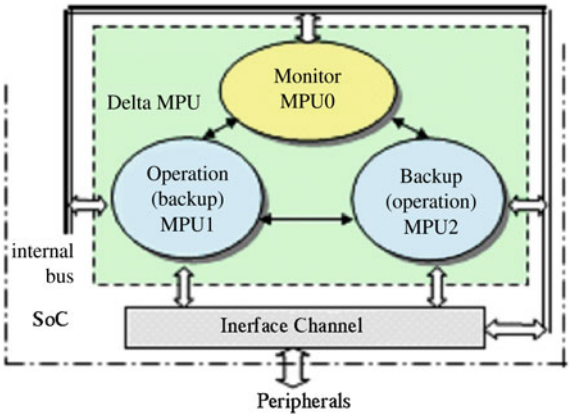


Fig. 216.2 An organization of SoC with delta multi-MPU-core

There are three support points in a delta structure. For this case, the support points are MPU0, MPU1 and MPU2. For peripherals, the SoC operation is managed by one MPU core. Other two cores, in fact, are arranged to guarantee system updating in operation. It also can be called dynamically reconfiguring, self-repairing or self-updating.

MPU0 is a monitor. It is an internal controller. MPU0 connects with internal bus and other two MPU cores. In operation, the MPU0 get system information from interface channel through internal bus. The information includes output parameters, status information of the system and demands from upper controllers. When some self-repairing algorithm is embedded into MPU0, MPU0 may control the SoC to take self-repairing automatically.

MPU1 and MPU2 are operation/backup cores. When MPU1 is in operation, the MPU2 is a backup core. In contrast, if MUP2 is an operation core, MPU1 can be in backup. The input and output signals for the two MPU cores are similar. At least, the numbers of signals are similar. In operation, if the functions of MPU core are required modifying or updating, the monitor controls backup core to reconfigure or update in structure, functions or instructions. The backup core becomes a new MPU after updating. Then, according to bumpless switching algorithm, the monitor core may choose a proper timing point, and transfers the new core into operation. At the same time, the original operation core alternates as a backup role.

All Communication with outside of SoC is taken through interface channel. In fact, the interface is composed of some IPs. In order to carry out the switching between operation and backup cores, some other IP modules could be required.

The delta core structure is fond of high reliability. In this case, the triangle MPUs are able to manipulate the core updating for a SoC in applications. On the other hand, the cost of dynamic updating with high stability and reliability is two more MPU cores. Therefore, the delta core structure is a design for some special solutions.

When a SoC is built by FPGA, the updating or reconfigurable procedures are easier. It is because of FPGA is a kind of PLD device to be easily modified. New products of FPGA can be made with about 180,000 LE (Logic Element) on a chip. This means more than one hundred MPU core, each to take 400-1700 LE, may embed into one FPGA chip [11]. And, to build multi-MPU SoC is possible. On the other hand, new FPGA products supply enough resource for DRSoC yet [12].

The frame of a DRSoC based on VHDL is as follows.

```
Library
Entity drSoC_swt is
Port();
End drSoC_swt;
Architecture struct of drSoC_swt t is
    Signal
    Constant
Begin
Res: process;
Clk: clk_div;
SWITCH: process;
    Begin
    If init then
    Initialization;
Elsif sg then
    CheckSignal;
Elsif mtst then
    Match State;
    Elsif switch then
    Switch MPU;
End if;
End process;
Us: ustimer;
Ms: mstimer;
End struct;
```

There are some analysis and decision algorithms in memory of a SoC. They are software to be run in monitor MPU. According to the algorithms, monitor MPU determines reconfiguration of backup MPU and switches it into operation.

Therefore, a solution of the DRSoC with delta multi-core must be a cooperation of software and hardware designs.

216.4 Conclusions

The SoC with delta MPU-core structure can be called DRSoC, Dynamically reconfigurable SoC. It can update its functions or organization dynamically. The "dynamically" means the updating with no stopping, no resetting, bumpless switching or no ex-behavior changing for an operation. The idea can guarantee

better security operations for some systems, such as self-repairing or self-tuning systems.

There are three MPU, monitor, operating and backup MPU, in the structure. It is a steady structure. When the SoC is required updated or repaired, the monitor controls backup MPU to be modified with some proper functions, then through a bumpless switching, the backup MPU becomes a new operation core. So, the self-repairing of the SoC is set up.

The design is based on FPGA with VHDL. It is a typical co-design of hardware and software. For most cases of DRSoC applications, cooperative design of hardware and software is the more efficient way.

Some applications required high reliability; the self-repairing scheme of SoC with delta multi-core is feasible and useful.

Acknowledgments This research was supported in part by the National Natural Science Foundation of China under grant 60872057, by Zhejiang Provincial Natural Science Foundation of China under grants R1090244, Y1101237, Y1110944 and Y1100095. We are grateful to NSFC, ZJNSF and Huzhou University.

References

1. Saleh R, Wilton S et al (2006) System-on-chip: reuse and integration. Proc IEEE 94(6): 1050–1069
2. Sifakis J (2009) Embedded systems design—scientific challenges and work directions. In: Proceedings of DATE '09, vol 4. IEEE Press. Nice, France, p 2
3. Liu L (2010) Advanced manufacturing of china and embedded systems. In: Proceedings of CCIE 2010, Wuhan, China, vol 6, pp 81–84
4. Ostua E, Viejo J et al (2008) Digital data processing peripheral design for an embedded application based on the Microblaze soft core. In: Proceedings of 4th south conference programmable logic, San Carlos de Bariloche, Argentina, vol 3, pp 197–200
5. Wallner S (2005) A configurable system-on-chip architecture for embedded and real-time applications: concepts design and realization. Elsevier J Syst Archit 51:350–367
6. Claasen TM (2006) An industry perspective on current and future state of the art in system-on-chip (SoC) technology. Proc IEEE 94(6):1121–1137
7. Smit G, Schuler E, Becker J, Quevremont J, Brugger W (2005) Overview of the 4S project. In: Proceedings of 2005 international symposium on system-on-chip, Tampere, Finland, vol 11, pp 70–73
8. Kay D, Chung S, Mourad S (2005) Embedded test control schemes using iBIST for SOCs. IEEE Trans Instrum Meas 54(3):956–964
9. Dimond RG, Mencer O, Luk W (2006) Combining instruction coding and scheduling to optimize energy in system-on-FPGA. In: Proceedings of 14th annual IEEE symposium on field-programmable custom computing machines, Napa, CA, USA, vol 4, pp 175–184
10. Liu L (2010) A hardware and software cooperative design of SoC IP. In: Proceedings of CCIE 2010, Wuhan, China, vol 6, pp 77–80
11. Kilts S (2007) Advanced FPGA design: architecture implementation and optimization. Wiley, New Jersey
12. Pan Z, Wells B (2008) Hardware supported task scheduling on dynamically reconfigurable SoC architectures. IEEE Trans VLSI Sys 16(11):1465–1474

Chapter 217
Design and Implementation of White LED Lighting Automatic Control System

Fashan Yu, Mingjie Zong and Xifang Tian

Abstract A white LED automatic control system is designed to maintain the illumination intensity of object region. AT89S52 is used as the controller and light senor IC BH1750 with I^2C bus is used to measure the illumination of object region. PWM signals are generated to control the RMS of the current to regulate the brightness of white LED. Using numeral PID as the control algorithm the constant lighting control was implemented. The hardware design and software programing of white LED lighting dimming is finished in this paper. The experiment shows the steady-state error is less than 2%, the dynamic characteristics are excellent and the system can reject the interference to some extent.

Keywords White LED · Luminosity control · I^2C bus · Numeral PID · PWM

217.1 Introduction

In the background of the global energy shortage concerns rising again, white LED lighting in the global market is to be given more attention. White LED used for lighting has many distinct advantages. Its luminous efficiency is far higher than incandescent lamp, no pollution, long service life, resistant to shake, and it is the

F. Yu · M. Zong (✉) · X. Tian
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: zongmingjie2011@163.com

F. Yu
e-mail: yufs@hpu.com

X. Tian
e-mail: txf16888@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_217,

real green light source [1]. At present, most public lighting systems are artificially controlled or timing controlled, and do not have a natural light intensity and the ability to automatically adjust the lighting intensity. It can cause lighting conditions in discord with human body comfort, waste of electric energy, light pollution and so on. This article introduces the design of automatic control system of white LED lighting, that can provides a convenient and low cost solution. Because of white LED with a good characteristics: the white LED light brightness and white LED through the positive current size basically proportional relationship [2], it can adjust white LED through the positive current and easily realize the effective value of the light. This article through the light sensor BH1750 acquisition the light information of target zone, and comparison, the use of deviation and setting signal digital position type digital (PID) operation produce of appropriate control quantity, through the change of pulse width modulation (PWM) signal of empty adjust LED lamp brightness of the value, so that the light intensity maintain target zone in near set value. So it not only can build a comfortable environment, and light can make full use of natural light to save electric energy effect.

217.2 Design of System

As mentioned, the white LED light brightness and white LED through the positive current size basically are proportional to the relationship, so this can adjust white LED through the positive current size and easily realize dimmer. In some existing light in technology, from the simple to the complex variable resistor of PWM, each kind of methods all have its advantages and disadvantages. The highest efficiency of light PWM, current control is also the most accurate [3], so in white LED light applications, the PWM light way to get universal application.

This design goal is to realize the target zone constant value and be able to control light manually setting illumination references. In order to achieve this goal, this paper is used the control closed-loop control strategy. Light sensor BH1750 real-time data acquisition target zone of light intensity, feedback to the controller, controller performs PID algorithm, adjust the PWM signal of empty, control than white LED through the positive current, and then adjust the effective value of white LED light, in order to maintain the brightness of the light value in the target zone near the set value, and achieve the target area light intensity automatic control. White LED lighting of the automatic control system principle block diagram shown as shown in Fig. 217.1.

White LED lighting automatic control system design diagram as shown in Fig. 217.2, the controller adopts AT89S52 single chip microcomputer (SCM), light sensor is BH1750, its constitute the module of collecting light, the keyboard and the LCD monitor convenient human–computer interaction, relay control circuit to stop a white LED lamps and lanterns, constant current driving circuit for LED lamps and lanterns is stable and reliable power supply, to move light PWM circuit to adjust LED the brightness of lamps and lanterns.

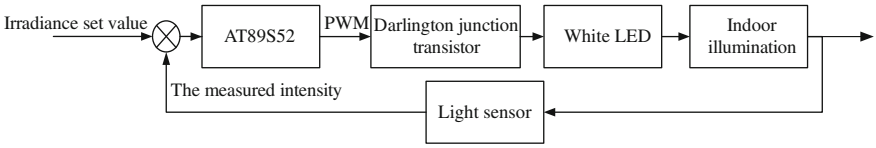


Fig. 217.1 Schematic diagram of white LED lighting automatic control system

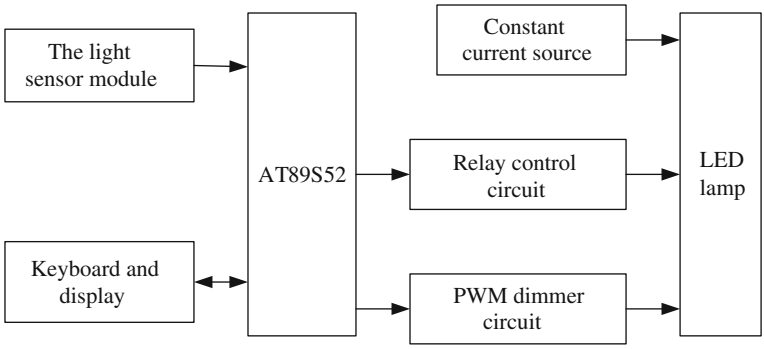


Fig. 217.2 System diagram of white LED lighting automatic control system

217.3 Design of System Hardware

217.3.1 Module Circuit Design of BH1750 Light Sensor

Chip of BH1750 light sensor can directly change the light intensity into digital, inhibit the 50 Hz/60 Hz light interference, with wide measurement range, high accuracy, small infrared effects etc., and broad prospects in the field of meter-ingbroad. As communicating with the SCM by I^2C protocol, the BH1750 only need two transmission lines (clock line SCL and signal line SDA) to connect SCM I/O port, greatly simplifying the hardware design. Considering the chip working voltage of SCM AT89S52 and light sensor BH1750 are not consistent, so it needs to add level conversion circuit. This part of the circuit diagram is shown in Fig. 217.3. In the figure, the three integrated voltage stabilizer U1, can convert +5 to +3.3 V for light sensor module circuit. As SCL is completly controlled by SCM, the diode circuit as shown in Fig. 217.3 can realize the one-way level conversion. In the Fig. 217.3 while normally working, the forward voltage drop about diode D1, schottky diode, is about 0.23 V. In signal line SDA, the data transmission is bidirectional, so it is needed to realize two-way level transformation. In this paper we realize it through driver N channel enhanced field effect transistors. After the test, this circuit design can complete work.

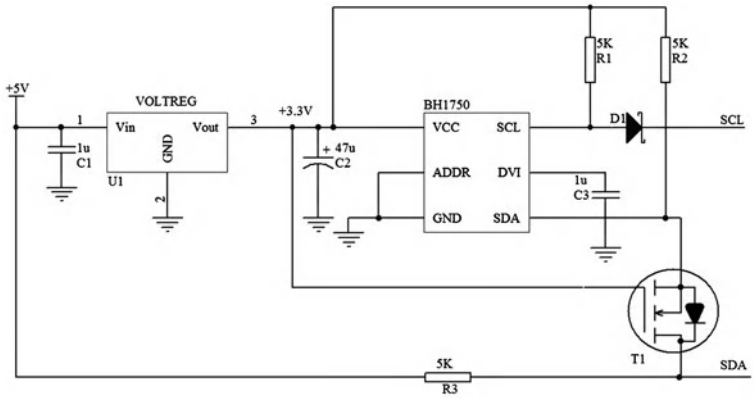


Fig. 217.3 Circuit of light senor

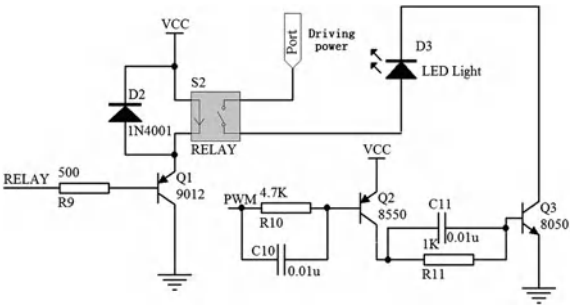


Fig. 217.4 Electromagnetic relay control circuit and PWM dimming circuit

217.3.2 Relay Control Circuit and PWM Light Circuit

In this paper, we adopt electromagnetic relay to control LED lamps. As output current of SCM ports cannot direct drive electromagnetic relay, we use the PNP transistor as amplifier to drive coil closure. Due to the relay of the coil is the perceptual load, for protection in the relay, ends the parallel a IN4001 as free-wheeling diode to protect. This paper has 1 W white LED lamp in the rated working status, rated current of about 300 mA, so light circuit design became of lyndon transistor structure to increase drive, to meet the needs of the controlled objects. Specific circuit as shown in Fig. 217.4.

217.3.3 Constant Current Driver Circuit

White LED is essentially a diode structure, the voltage and current flows through a typical relationship of the index. The device at conduction, small changes in the terminal will cause more electricity flow guide the change, and current changes

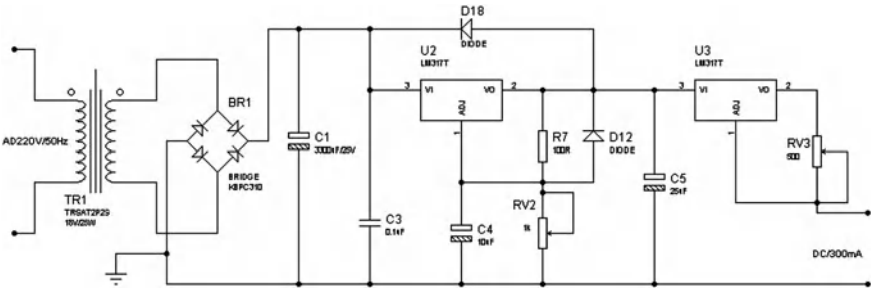


Fig. 217.5 Circuit of steady current driver

affect not only brightness, can also affect the quality of white light, cause color shift, this is we should avoid in application. Therefore, white LED to the constant current drive mode will be better than the constant pressure drive mode [4]. In this paper, the LM317 construct a constant current driver circuit. LM317 belongs to the three adjustable output is the voltage regulator, after configuration proper peripheral circuit that can be as a constant voltage source and constant current source to use. First, after voltage adjust, rectification, filter, the LM317 connect to a constant voltage source, adjust the resistance of the RV2, making the C5 on both ends of the voltage stability in +12 V, and then another LM317 connect to a constant current output, according to load current adjust the feedback resistance RV3 requirements of resistance. The feedback resistance of the resistance can such as the formula (217.1) calculation.

$$R_{\text{sense}} = \frac{1.25}{I_{\text{out}}} \tag{217.1}$$

The formula in the value of the feedback resistance for asked white LED the rating of the current work. The power circuit as shown in Fig. 217.5.

217.4 System Software Design

217.4.1 Control Algorithm

This paper adopts the PID controller to realize constant illumination control. Light adjusting system belongs to the fast system, actuator inertia is small, anti-interference ability is weak, therefore the sampling period should be as short as possible, integral time constant can be greatly [5]. The system of the sampling period is 40 ms. Control system based on PID performance depends largely on the PID controller parameter selection, selection of parameters can be the first reference of empirical formulas about range, according to the practical test results do adjustment. In this paper the proportion coefficient selected for $K_{\text{p}} = 0.25$, another test found that when introducing an integral or differential system stability and response speed is improved, and the stability is decreased, so $K_{\text{p}} = 0.25$, $K_{\text{i}} = \infty$, $K_{\text{d}} = 0$.

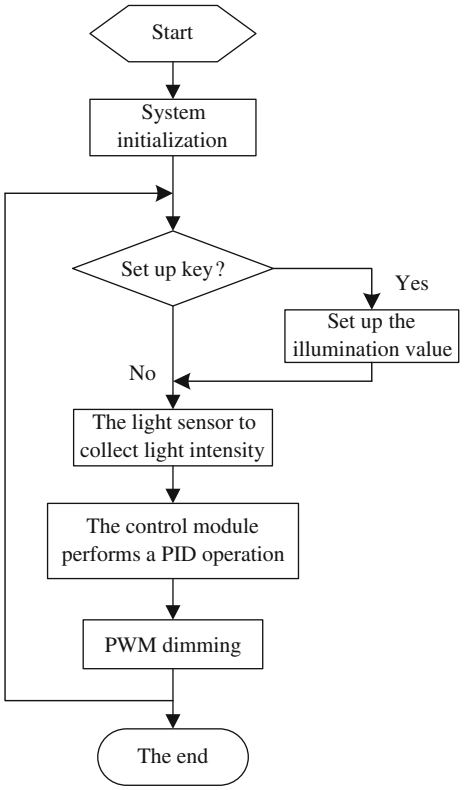


Fig. 217.6 Program flowchart

217.4.2 Mode of Light Adjustment

The light adjusting mode using PWM, because AT89S52 does not have hardware PWM module, so the use of timers to analog PWM signal. This design adopts square wave period, but only by changing the duty ratio of the scheme to implement the software PWM.PWM dimming is to use the principle of people's visual residual effect, when the LED light dark frequency greater than 100 Hz, people would feel flicker, and that the lights are always bright. This paper selects the bright and dark frequency 200 Hz.

217.4.3 Program Flow

Power system, can be manually set the desired illumination values, then the light sensor BH1750 collecting light intensity signal, in the AT89S52 MCU in operative PID operation, output suitable control volume, to reduce the deviation, after several cycle, system in steady state, the target area light intensity was stable near the enactment value in. Program flow diagram as shown in Fig. 217.6.

Table 218.1 System performance when set value changes and interference is fixed

The given value/Lux	Transition time/ms	Steady-state error/Lux
600	335	6
800	485	4
900	635	3

Note The system initial target area light value is zero

Table 218.2 System performance when set value is fixed and interference changes

Interference/Lux	Transition time/ms	Steady-state error/Lux
100	108	4
200	151	5

217.5 System Performance Analysis

(a) When $K_p = 0.25$, given the value change, interference invariant, the measured data is shown in Table 218.1.
Through the analysis found that system regulation time for millisecond level, dynamic performance, and high steady state accuracy.

(b) When $K_p = 0.25$, given a constant value (set to 600Lux), were added step interference signal, the measured data is shown in Table 218.2.
Through the analysis found that adding step interference signal, the system can quickly to suppress interference, the anti-interference ability is strong.

217.6 Conclusion

This paper achieves the PID tuning of the white LED lighting automatic control system. The proposed scheme has excellent static control performance, low cost and wide prospect. The white LED lighting automatic control system can not only be able to satisfy people's harsh lighting conditions, but also to maximize the use of natural light and save a lot of large amount of electricity. In the era of growing energy crisis, white LED lighting automatic control system has opened up a new path for people, which use efficient lighting approach.

References

1. Yang H (2008) LED lighting drive circuit design and selected example. China Electric Power. Press, Beijing
2. Song S, Liu T, Ai Y, Xu H (2009) Based on AVR SCM LED adaptive light system. Comput Appl 31(1):33–34
3. Kopp U (2008) High brightness LED dimming technology. Foreign Electron Compon 2:64–67

4. Fu J (2007) The boost dual mode PWM + LED drived chip design. Zhejiang University, Hangzhou
5. Xing S, Zhuang Y, Liang G (2010) The intensity of illumination with the automatic control system based on PID control. Comput Digit Eng 38(5):71–72

Chapter 218
LPV Modeling of the Missile Attitude Control System Based on Small Deviation Equation

Shaobo Ni, Jianwu Wu, Jiayuan Shan, Lei Liu, Yongji Wang and Kun Hu

Abstract With a concern about the missile attitude control system, this paper attempts to solve the online calculation method of the coefficient of the small deviation equation, on which a LFT-based LPV model of the missile attitude control system is set up.

Keywords Missile control · Attitude control · Variable gain control · Linear variable parameter

218.1 Introduction

It is a common method to devise the missile attitude control system which considers it as a time-invariant system with the 'freezing' coefficient method, designs a linear controller at each second point using the traditional controller design method and interpolates among these linear controllers with specified scheduling rules to get a whole nonlinear control law. However, this method has two

S. Ni · J. Shan
The College of Aerospace, Beijing Institute of Technology, Beijing 100081, China
e-mail: nsbxx@hotmail.com

J. Wu
Beijing Institute of Aerospace Automatic Control, Beijing 100854, China

L. Liu (✉) · Y. Wang · K. Hu
Key Laboratory of Ministry of Education for Image Processing and Intelligent Control,
Department of Control Science and Engineering,
Huazhong University of Science and Technology, Wuhan 430074, Hubei, China
e-mail: lei.liu.chn@gmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_218,

disadvantages. Firstly, the parameters of the controller vary in an open-loop way without the feedback of the performance of the close-loop system. Secondly, there is no enough theoretical improvement of the stability of this controller.

The variable gain control method based on the Linear Parameter-Varying (LPV) system which was introduced by Shamma and Athans from MIT in the 1990s can overcome the two disadvantages of the traditional variable gain control method [1, 2]. The research on the LPV system-based variable gain control is mainly based on Linear Fractional Translation (LFT) [3, 4] and the Lyapunov function [5–9] which are both solved by the Linear Matrix Inequalities (LMIs). The LMI could be solved with the standard convex optimization algorithms.

In order to apply the LPV system-based variable gain control method to the design of the ballistic missile attitude control system in the active flying segment, this paper considers the small deviation kinematic equation, studies the online calculation problem of the time-variant parameter and sets up the LPV model of the missile attitude control system based on LFT.

218.2 The Small Deviation Kinematic Model of the Missile

To analyze the problem in a more convenient way, this paper just considers the rigid motion of the missile in the pitch channel. If we choose the state variables as $X = [\Delta\theta \quad \Delta\varphi \quad \Delta\dot{\varphi}]^T$ ($\Delta\theta$ is the deviation of the flight path angle, $\Delta\varphi$ is the deviation of the pitch angle), the outer disturbance as $w = [\alpha_w \quad \bar{F}_y \quad \bar{M}_{z1}]^T$ (α_w is the attack angle caused by the wind, $\bar{M}_{z1}$ is the disturbance moment coefficient in the normal direction and $\bar{F}_y$ is the disturbance force coefficient in the normal direction) and the control input as $u = \delta_\varphi$(δ_φ is the nozzle swing angle in the normal direction) [10, 11], the state equation of the pitch channel is:

$$\dot{X} = AX + B_1 w + B_2 u, \tag{218.1}$$

where

$$A = \begin{bmatrix} c_2 - c_1 & c_1 & 0 \\ 0 & 0 & 1 \\ b_2 & -b_2 & -b_1 \end{bmatrix},\ B_1 = \begin{bmatrix} c_1' & 1 & 0 \\ 0 & 0 & 0 \\ -b_2 & 0 & 1 \end{bmatrix},\ \text{and } B_2 = \begin{bmatrix} c_3 \\ 0 \\ -b_3 \end{bmatrix}$$

in which,

$$b_1 = 57.3 m_{dz1}^{\omega} q S_m l_k^2 / J_{z1} V,\ b_2 = 57.3 m_{z1}^{\alpha} q S_m l_k / J_{z1},$$
$$b_3 = \left[F_p(x_R - x_z) + m_R \dot{W}_{x1} l_R\right] / \sqrt{2} J_{z1}, c_1 = \left(4F_p \cos\alpha + 57.3 C_y^{\alpha} q S_m\right) / m\mathrm{V},$$
$$c_1' = 57.3 C_y^{\alpha} q S_m / m\mathrm{V},\ c_2 = g \sin\theta / \mathrm{V},\ c_3 = F_p / \sqrt{2} m\mathrm{V}$$

218.3 The Online Calculation Method of the Parameter

In order to get the LPV model of the missile attitude control system, it is necessary to choose a certain group of time-variant parameter which could be measured online, and write the coefficient matrix A, B_1 and B_2 in the small deviation equation (218.1) as explicit functions with the group of parameters as the variables. A probable scheme is to designate the group of parameters as b_1, b_2, b_3, c_1, c_1', c_2 and c_3. We can judge from above that if these variables are chosen as parameters, the coefficient matrix A, B_1 and B_2 will be the affine function of the group of parameters. The advantage of the method is that we will only need to solve finite number of groups of LMIs when we apply the Lyapunov function to design the LPV controller. On the other way, it is convenient to write the system state equation as the LFT form and use the LFT-based LPV variable gain control method.

The theoretical formulas to calculate the group of parameters $b_1, b_2, b_3, c_1, c_1', c_2$ and c_3 are provided. In order to apply them in the online calculation of the parameters, the paper will analyze the models of the mass (m), thrust (FP), centroid position (xz), moment of inertia (Jzl) and other parameters.

218.3.1 Model of Mass Center

During the flying process the missile will satisfy

$$-v\frac{\mathrm{d}}{\mathrm{d}t}m + S_a(P_a - P_h) + F_q = m\dot{W}_{x1} \tag{218.2}$$

in which, v is the injection rate of the gas relative to the missile body, S_a is the sectional area of the engine nozzle, P_a is the static pressure of the engine nozzle gas, P_h is the pressure of atmosphere at the height of h, F_q is aerodynamic drag. The first item on the left side of equation (218.2) is the dynamic thrust of the engine, the second item is the static thrust of the engine. If we hypothesize $m_1(t)$ as the solution of equation (218.2), then

$$\begin{aligned} &-(\ln m_1(t) - \ln m_1(t_0)) + \int_{t_0}^{t} \frac{S_a(P_a(\tau) - P_h(\tau)) + F_q(\tau)}{m_1(\tau)v}\mathrm{d}\tau \\ &= \frac{1}{v}(W_{x1}(t) - W_{x1}(t_0)) \end{aligned} \tag{218.3}$$

Suppose the solution of the equation

$$-v\frac{\mathrm{d}}{\mathrm{d}t}m = m\dot{W}_{x1} \tag{218.4}$$

is $m_2(t)$, then

$$-(\ln m_2(t) - \ln m_2(t_0)) = (W_{x1}(t) - W_{x1}(t_0))/v \tag{218.5}$$

If we choose $m_1(t_0) = m_2(t_0)$, then we can get from Eq. (218.3) and (218.5) that

$$m_2(t) = m_1(t)\ \exp\left[-\int_{t_0}^{t}\frac{S_a(P_a(\tau) - P_h(\tau)) + F_q(\tau)}{m_1(\tau)v}\mathrm{d}\tau\right]. \tag{218.6}$$

Since the dynamic thrust of the engine is much larger than the sum of its static thrust and aerodynamic drag, there must exist a certain positive constant ζ which satisfies that with any $t \in [t_0, t_f]$

$$\left|S_a(P_a(t) - P_h(t)) + F_q(t)\right| \le \zeta\left(-v\frac{\mathrm{d}}{\mathrm{d}t}m_1(t)\right) \tag{218.7}$$

which means that

$$\begin{aligned}\exp\left[\zeta\int_{t_0}^{t}\frac{\dot{m}_1(\tau)}{m_1(\tau)}\mathrm{d}\tau\right] &\le \exp\left[-\int_{t_0}^{t}\frac{S_a(P_a(\tau) - P_h(\tau)) + F_q(\tau)}{m_1(\tau)v}\mathrm{d}\tau\right] \\ &\le \exp\left[-\zeta\int_{t_0}^{t}\frac{\dot{m}_1(\tau)}{m_1(\tau)}\mathrm{d}\tau\right]\end{aligned} \tag{218.8}$$

namely

$$\left(\frac{m_1(t)}{m_1(t_0)}\right)^{\zeta} \le \exp\left[-\int_{t_0}^{t}\frac{S_a(P_a(\tau) - P_h(\tau)) + F_q(\tau)}{m_1(\tau)v}\mathrm{d}\tau\right] \le \left(\frac{m_1(t_0)}{m_1(t)}\right)^{\zeta} \tag{218.9}$$

connected with equation (218.6), we can again get

$$\frac{|m_2 - m_1|}{m_1} \le \max\left\{1 - \left(\frac{m_1(t_f)}{m_1(t_0)}\right)^{\zeta}, \left(\frac{m_1(t_0)}{m_1(t_f)}\right)^{\zeta} - 1\right\}. \tag{218.10}$$

Considering that $0 < m_1(t)/m_1(t_0) \le 1$ and $\zeta \ll 1$ will cause $|m_2 - m_1|/m_1 \approx 0$, namely $m_2 \approx m_1$, so we get the relation between the missile mass m and the sight acceleration $\dot{W}_{x1}$ from Eq. (218.4) approximately as,

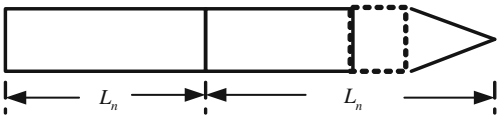


Fig. 218.1 The missile body diagram

$$m(t) = m(t_0)\exp\left[-\frac{1}{v}\int_{t_0}^{t}\dot{W}_{x1}(\tau)\mathrm{d}\tau\right] \tag{218.11}$$

218.3.2 Model of Thrust

On the basis of the model of the missile mass, we can get the model of thrust as,

$$F_P = m(t)\dot{W}_{x1} - F_q \tag{218.12}$$

in which the aerodynamic drag F_q can be calculated with the following formula,

$$F_q = -C_{x1}qS_m \tag{218.13}$$

In the formula, C_{x1} is the drag coefficient around axis $x1$.

218.3.3 Model of Centroid Position

Suppose m_{n0} and $m_{(n+1)0}$ are the overall mass of the missile at the initial moment of the n th and $(n+1)$th flying segment ,respectively, mn is the mass of the missile shell of the nth segment, Ln and $\tilde{L}_n$ are the length of the missile body after the nth and $(n+1)$th segment being separated, x_n is the length of the residual fuel of the nth segment, x_{zn} is the range between the centroid of the part without the nth fuel segment and the theoretical tip (Fig. 218.1).

During the the nth flying segment of the missile, the centroid position x_z will satisfy

$$\left(m - m_n - m_{(n+1)0}\right)\left(x_n/2 + L_n + \tilde{L}_n - x_z\right) + \left(m_n + m_{(n+1)0}\right)\left(x_{zn} - x_z\right) = 0 \tag{218.14}$$

Suppose the fuel segment is uniformly thick, namely

$$x_n = k_n\left(m - m_n - m_{(n+1)0}\right) \tag{218.15}$$

Then, considering the Eq. (218.14), we can get

$$\begin{aligned} x_z = {} & \frac{k_n}{2}\cdot m + \left(m_n + m_{(n+1)0}\right)\left[\left(x_{zn} - \tilde{L}_n\right)\right. \\ & \left. + \frac{1}{2}k_n\left(m_n + m_{(n+1)0}\right)\right]\cdot\frac{1}{m} + \left[\tilde{L}_n - k_n\left(m_n + m_{(n+1)0}\right)\right] \end{aligned} \tag{218.16}$$

which means that the centroid position x_z has the form

$$x_z = \left(c_{x_z}\right)_1 \cdot m + \left(c_{x_z}\right)_2 \cdot \frac{1}{m} + \left(c_{x_z}\right)_3 \tag{218.17}$$

218.3.4 Moment of Inertia Model

In the *n*th flying segment of the missile, the moment of inertia J_{z1} can be divided into the *n*th fuel, the *n*th missile shell and the missile body without the *n*th segment and be calculated, respectively, namely

$$\begin{aligned} J_{z1} = & m\left[\left(x_n + \tilde{L}_n - x_z\right)^2 - \text{sign}\left(\tilde{L}_n - x_z\right)\left(\tilde{L}_n - x_z\right)^2\right]/2 + m_n\left[\left(L_n + \tilde{L}_n - x_z\right)^2\right. \\ & \left. - \left(x_n + \tilde{L}_n - x_z\right)^2\right]/2 + m_{(n+1)0}\left[x_z^2 + 2\text{sign}\left(\tilde{L}_n - x_z\right)\left(\tilde{L}_n - x_z\right)^2\right. \\ & \left. - \left(x_n + \tilde{L}_n - x_z\right)^2\right]/2 \end{aligned} \tag{218.18}$$

Considering the relation between the centroid position x_z and mass m provided by the Eq. (218.17), there exists relationship between the moment of inertia J_{z1} and mass m,

$$J_{z1} = \left(c_{J_{z1}}\right)_1 \cdot m^3 + \left(c_{J_{z1}}\right)_2 \cdot m^2 + \left(c_{J_{z1}}\right)_3 \cdot m = \left(c_{J_{z1}}\right)_4 = \left(c_{J_{z1}}\right)_5 \cdot \frac{1}{m} + \left(c_{J_{z1}}\right)_6 \cdot \frac{1}{m^2} \tag{218.19}$$

218.4 The LFT Model of the Missile

In order to apply the LPV system-based variable gain control method which is based on Lyapunov function in the article [9], it is just necessary to consider the overall data and estimate the change rate of the group of parameter b2, b3, c1, c1′, c2 and c3. However, to use the LFT-based LPV system variable gain control method in article [4], it is also necessary to write the missile attitude control system equation as the LFT form. Choosing both the measured output y and the evaluation index z as $\Delta\varphi$ in the system (218.1), and using the graphical method in article [12], we can get the LFT model of the missile pitch channel as:

$$\begin{aligned} \begin{bmatrix} z \\ y \end{bmatrix} &= F_u(G(s), \Theta)\begin{bmatrix} w \\ u \end{bmatrix} \\ &= \left(\begin{bmatrix} G_{zw} & G_{zu} \\ G_{yw} & G_{yu} \end{bmatrix} + \begin{bmatrix} G_{z\theta} \\ G_{y\theta} \end{bmatrix}\Theta(I - G_{\theta\theta}\Theta)^{-1}\begin{bmatrix} G_{\theta w} \\ G_{\theta u} \end{bmatrix}^T\right)\begin{bmatrix} w \\ u \end{bmatrix} \end{aligned} \tag{218.20}$$

the system matrix of which can be depicted as

$$\begin{cases} \begin{bmatrix} q_\theta \\ z \\ y \end{bmatrix} = \overbrace{\begin{bmatrix} G_{\theta\theta}(s) & G_{\theta w}(s) & G_{\theta u}(s) \\ G_{z\theta}(s) & G_{zw}(s) & G_{zu}(s) \\ G_{y\theta}(s) & G_{yw}(s) & G_{yu}(s) \end{bmatrix}}^{G(s)} \begin{bmatrix} n_\theta \\ w \\ u \end{bmatrix} \\ n_\theta = \Theta q_\theta \end{cases} \tag{218.21}$$

in which

$$\Theta = \mathrm{diag}\left\{ \widetilde{c_2 - c_1}, \widetilde{c_1}, \widetilde{c_1'}, \widetilde{c_3}, \widetilde{b_2}, \widetilde{b_2}, \widetilde{b_2}, \widetilde{b_3} \right\} \tag{218.22}$$

The state space realization of $G(s)$ in the equation (218.23) is

$$G(s) = \begin{bmatrix} D_{\theta\theta} & D_{\theta 1} & D_{\theta 2} \\ D_{1\theta} & D_{11} & D_{12} \\ D_{2\theta} & D_{21} & D_{22} \end{bmatrix} + \begin{bmatrix} C_\theta \\ C_1 \\ C_2 \end{bmatrix} (sI - A)^{-1} \begin{bmatrix} B_\theta \\ B_1 \\ B_2 \end{bmatrix}^T \tag{218.23}$$

in which (in the Eq. (218.23), the matrix which are not given below are all zero matrix),

$$A = \begin{bmatrix} \overline{c_2 - c_1} & \overline{c_1} & 0 \\ 0 & 0 & 1 \\ \overline{b_2} & -\overline{b_2} & 0 \end{bmatrix},$$

$$B_\theta = \begin{bmatrix} \widehat{c_2 - c_1} & \widehat{c_1} & \widehat{c_1'} & \widehat{c_3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \widehat{b_2} & -\widehat{b_2} & -\widehat{b_2} & -\widehat{b_3} \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0 & 1 & 0 \\ \overline{c_1'} & 0 & 0 \\ -\overline{b_2} & 0 & 1 \end{bmatrix},\ B_2 = \begin{bmatrix} \overline{c_3} \\ 0 \\ -\overline{b_3} \end{bmatrix},\ C_1 = C_2 = [0 \quad 1 \quad 0]$$

$$C_\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ D_{\theta 1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ D_{\theta 2} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

In the these matrix above, the functions $\overline{*}$; $\widehat{*}$ and $\widetilde{*}$ are defined respectively as: $\overline{x} = (x_{\max} + x_{\min})/2$; $\widehat{x} = (x_{\max} - x_{\min})/2$ and $\widetilde{x} = (x - \overline{x})/\widehat{x}$ (b_2, b_3, c_1, c_1', c_2 and c_3 (b_2, b_3, c_1, C_1', c_2 and c_3 are all normalized in the function $\widetilde{*}$).

218.5 Conclusion

This paper provides the online calculation method of the coefficient of the small deviation equation of the ballistic missile attitude control system in the active flying segment on the basis of which a LPV model is set up, which lays a basis for the application of the LPV system variable gain control method.

Acknowledgments This work was supported in part by the National Nature Science Foundation of China, No 60975058, Nature Science Foundation of Hubei NO.2010CDB01904, 2010 Aerospace Technology Support Foundation, 2010 Defense Innovation Research Foundation of Huazhong University of Science and Technology.

References

1. Shamma JS, Athans M (1990) Analysis of nonlinear gain-scheduled control systems. IEEE Trans Autom Control 35:898–907
2. Shamma JS, Athans M (1991) Guaranteed properties of gain scheduled control for linear parameter-varying plants. Automatica 27:559–564
3. Packard A (1994) Gain scheduling via linear fractional transformations. Syst Cont Lett 22:79–92
4. Apkarian P, Gahinet P (1995) A convex characterization of gain-scheduled H_∞ controllers. IEEE Tran Autom Cont 40:853–864
5. Becker G, Packard A (1994) Robust performance of linear parametrically varying systems using parametrically-dependent linear feedback. Syst Cont Lett 23:205–215
6. Becker G (1996) Additional results on parameter-dependent controllers for LPV systems. In: Proceedings of the 13th IFAC world congress, pp 351–356
7. Wu F, Yang XH, Packard A, Becker G, Induced L_2 norm control for LPV systems with bounded parameter variation rates. Int J Rob Non Cont 6:983–998
8. Yu J, Sideris A (1997) H_∞ control with parametric Lyapunov functions. Syst Cont Lett 30:57–69
9. Apkarian P, Adams R (1998) Advanced gain scheduling techniques for uncertain system. IEEE Trans Cont Syst Technol 6:21–32
10. Xu Y (1999) Design and analysis of the control system of the ballistic missile and launch vehicle. Aerospace Press, Auburn
11. Xu Y (1989) Control system. Aerospace Press, Auburn
12. Guo Di, Rugh WJ (1995) A stability result for linear parameter-varying systems. Syst Cont Lett 24:1–5

Part XV
Electric Machines and Electric Machines Control

Chapter 219
Adaptive Terminal Sliding Mode Control for Chaos Oscillation of Power System

Yuhui Zhang, Zhihua Zhang and Tianyun Li

Abstract An adaptive Terminal sliding mode control method is put forward to restrain chaotic oscillation when power system is suffered from periodic load disturbance. First, the use of Lyapunov index proved that the system will occur with chaotic oscillation under certain circumstances. Then through the design of Terminal sliding mode, the convergence of error is sped up to achieve good control effect. Using adaptive law to estimate the changes in disturbance, it makes the controller have strong robustness. According to Lyapunov theory, the stability of closed-loop system is proved. Simulation results show the effectiveness of the control method.

Keywords Chaos oscillation · Adaptive · Terminal sliding mode · Lyapunov exponent

219.1 Introduction

Power system is a typical nonlinear system [1–3], it contains many of the dynamic behavior of complex nonlinear phenomena, such as low-frequency oscillations, sub-synchronous oscillation, bifurcation and chaos. When the amplitude of

Y. Zhang (✉) · Z. Zhang · T. Li
Northeast Dianli University, Jilin, Jilin, China
e-mail: zyh586@163.com

Z. Zhang
e-mail: 148050374@qq.com

T. Li
e-mail: tyleee@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_219,

periodic load disturbance meets certain conditions, it will have chaotic oscillation. Chaos is a form of random motion of deterministic nonlinear systems and it is extremely sensitive to initial values, long-term unpredictability.

As the chaotic oscillations may cause the system to instability and seriously endanger the safe operation of the system, many scholars have conducted an effective exploration to the chaos control in recent years. Literature [4] eliminates chaotic oscillations of power system by using time-delayed feedback. Literature [5] achieves chaos control of power system by using neural networks. Literature [6, 7] design chaos controller by using fuzzy sliding mode variable structure control method. Literature [8] inhibition of the chaotic oscillation of power systems by using a variable feedback. Literature [9] uses adaptive method to control chaos of system which is under certain disturbances and unknown disturbances and achieves the purpose of eliminating chaos. Literature [10] uses LS-SVM to study the dynamics behavior of power system to obtain the trained LS-SVM model of power system and realize the control of power system chaos oscillation.

Although the above methods can effectively control the chaotic oscillation of system, the system output track the desired signal is too long. The actual power system will subject to random noise Inevitably. Literature [11] has pointed out the role of boundary noise can increase the system's chaotic region and makes power systems easier to generate Chaotic Motion. None of the above-mentioned studies consider the impact of noise disturbance on the system. On the basis of literature [12], this chapter combines adaptive control with Terminal sliding mode control to propose an adaptive Terminal sliding mode chaos control. Using Lyapunov index the existence of chaotic oscillations of power system is proved. Designing Terminal sliding mode to allow the system to converge rapidly and giving the adaptive law of disturbance. Simulations show that the controller can effectively calm chaotic oscillation power system.

219.2 The Dynamic Behavior of Simple Power Systems Under Periodic Load Disturbances

219.2.1 Mathematical Model

This paper uses nonlinear mathematical model of synchronous motor and ignores the impact of excitation dynamic. Assuming that the generator mechanical power in the transient process has remained unchanged in the transient process and excluding the transient saliency effect of generator. This chapter uses the second-order simplified model [13]:

$$\begin{cases} \dot{\delta} = \omega \\ \dot{\omega} = -\dfrac{1}{H} P_s \sin\delta - \dfrac{D}{H}\omega + \dfrac{1}{H} P_m + \dfrac{1}{H} P_e \cos\beta t \end{cases} \tag{219.1}$$

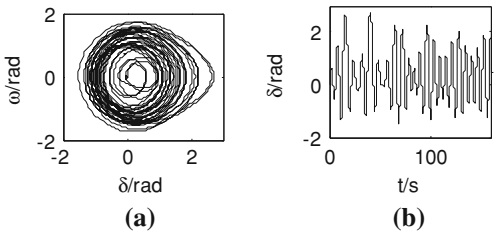


Fig. 219.1 Phase map and $\delta(t)$ curves of the simulated system when $P_e = 29.6$

Among it, δ and ω are generator power angle and relative speed. P_s and P_m are electromagnetic power and mechanical power. H and D are equivalent moment of inertia and damping coefficient. P_e is Load disturbance amplitude. β is disturbance load frequency.

Make $\alpha = P_s/H,\ \gamma = D/H,\ \rho = P_m/H,\ F = P_e/H$, then Eq. (219.1) becomes

$$\begin{cases} \dot{\delta} = \omega \\ \dot{\omega} = -\alpha \sin\delta - \gamma\omega + \rho + F\cos\beta t \end{cases} \tag{219.2}$$

Make $[x_1, x_2] = [\delta, \omega]$, then the following equation is controlled closed-loop system equations:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\alpha \sin x_1 - \gamma x_2 + \rho + F\cos\beta t + u \end{cases} \tag{219.3}$$

219.2.2 The Generation of Chaotic Oscillations Result from the Changing of Load Disturbance Amplitude

Make $H = 100,\ P_s = 100,\ D = 2,\ P_m = 20,\ \beta = 1$ that is $\alpha = 1,\ \gamma = 0.02,\ \rho = 0.2$, study the dynamic behavior of power system when P_e changes [13].

When $\beta = 1,\ P_e = 0$, that is the system only have periodic load disturbance except damping. $\delta(t)$ eventually restored to the steady state after a period of oscillation. When the amplitude of periodic load disturbance P_e is increased to 29.6, the system will appear with chaotic oscillations. The phase trajectory of system appears with chaotic attractor and orbital motion is ergodic. Phase diagram in Fig. 219.1 shows irregular oscillations generator power angle. When P_e has a further small increase (such as $P_e = 29.613$), system evolved from the chaotic oscillation to divergent oscillation and its stability has completely lost. Therefore chaotic oscillations have a great harm to the system's safe operation (Fig. 219.2).

System has three Lyapunov exponents looking from Fig. 219.3 and one of the three Lyapunov exponents is positive. It proves that when the system parameters $P_e = 29.613$, it will produce chaos phenomenon.

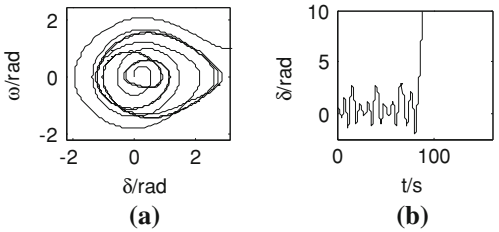


Fig. 219.2 Phase map and $\delta(t)$ curves of the simulated system when $P_e = 29.613$

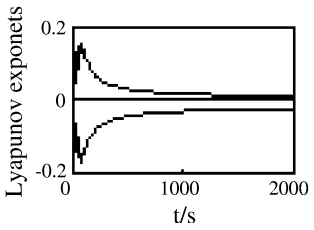


Fig. 219.3 Exponent map of the simulated system when $P_e = 29.613$

219.3 The Design of Terminal Sliding Mode Controller

Terminal sliding mode control theory can be as follows: first, nonlinear function is introduced in the design of sliding hyperplane; second, construct Terminal sliding mode surface so that the tracking error in the sliding surface can converge to zero within a limited time.

Make

$$s = cx_1 + x_2, \quad c > 0 \tag{219.4}$$

Derivative of formula (219.4) is

$$\dot{s} = cx_2 - \alpha \sin x_1 - \gamma x_2 + \rho + F \cos \beta t + u \tag{219.5}$$

construct Terminal sliding mode:

$$\dot{s} + \alpha s + \beta s^{q/p} = 0, \ \alpha, \quad \beta > 0, \tag{219.6}$$

not only p and q is positive odd number, but also $p > q$.

Nonlinear control law as follows:

$$\begin{aligned} u &= -\alpha s - \beta s^{q/p} - cx_2 + \alpha \sin x_1 + \gamma x_2 - \rho - \hat{F} \cos \beta t \\ \dot{\hat{F}} &= rs \cos \beta t \end{aligned} \tag{219.7}$$

r is adaptive gain coefficient.

The proof of stability of the system as follows:

The result of type (219.7) into type (219.5) is

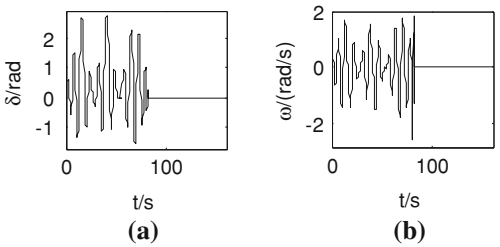


Fig. 219.4 The dynamic response curves of the controlled system when $P_e = 29.6$

$$\dot{s} = -\alpha s - \beta s^{q/p} - \hat{F}\cos\beta t + F\cos\beta t \tag{219.8}$$

construct Lyapunov function:

$$V = \frac{1}{2}s^2 + \frac{1}{2r}(\hat{F} - F)^2 \tag{219.9}$$

Its derivation is

$$\begin{aligned}\dot{V} = {} & s\dot{s} + \frac{1}{r}\left(\hat{F} - F\right)\dot{\hat{F}} = -\alpha s^2 - \beta s^{(q+p)/p} - s\hat{F}\cos\beta t + sF\cos\beta t \\ & + s\left(\hat{F} - F\right)\cos\beta t \\ = {} & -\alpha s^2 - \beta s^{(q+p)/p}\end{aligned} \tag{219.10}$$

As p and q are positive odd numbers, then $\dot{V} < 0$, therefore the system is stable.

219.4 Numerical Simulation Analysis

To verify adaptive Terminal sliding mode control method which is proposed by this chapter is effective, simulate the entire system In Matlab/Simulink environment.

Select the initial state $x_0 = [0,\ 0]$.

Consider the following two situations:

(1) system without noise:

When $t = 80\,s$ and $P_e = 29.6$ introduce control item. Simulation curve of system is shown in Fig. 219.4. The figure shows that the system is in chaotic state of motion before the introduction of control item. Chaotic oscillation of the system is quickly suppressed after the introduction of the control item and system status comes back to the initial balance.

When $P_e = 29.613$, input into the controller before the system loses stability, dynamic response curves of the controlled system is shown in Fig. 219.5. Simulation results show that, under the action of the controller, the system can quickly shift from the chaotic running state to stable state and stabilize equilibrium point.

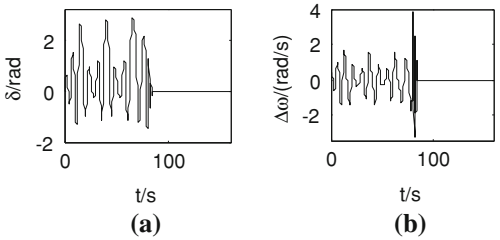


Fig. 219.5 The dynamic response curves of the controlled system when $P_e = 29.613$

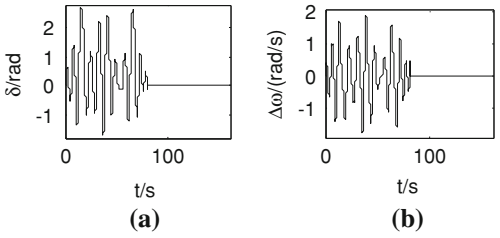


Fig. 219.6 The dynamic response curves of the controlled noisy system when $P_e = 29.6$

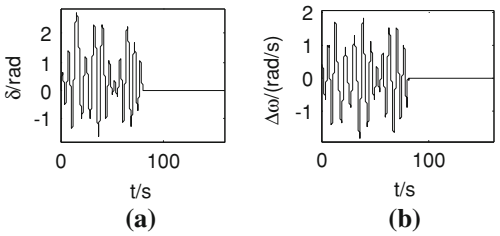


Fig. 219.7 The dynamic response curves of the controlled noisy system when $P_e = 29.613$

(2) system with noise:

Adding white noise into the above system, the white noise's mean is zero and variance is 0.1.

Figures 219.6 and 219.7 show that, when added white noise into the system, the controller is still able to calm the chaotic attractor, inhibit chaotic oscillation. We can see that system added controller has a certain anti-disturbance ability and good robustness.

219.5 Conclusion

This chapter combines adaptive control with Terminal sliding mode control to control the chaotic oscillations of power system. First, using Lyapunov exponents proved that system will produce chaos in a certain amplitude of the disturbance. Then apply adaptive Terminal sliding mode control method to inhibit the chaotic attractors. The controller has the following advantages:

(1) Terminal sliding mode control can guarantee that the system converge to the equilibrium state rapidly;
(2) Consider the impact of noise on the system and verify the robustness of the controller;

Simulation shows that the controller can eliminate chaotic oscillation of the system, verify the effectiveness of method proposed by this paper.

References

1. Zhu F, Zhao H, Liu Z et al (2007) The influence of large power grid interconnected on power system dynamic stability. Proc CSEE 27(1):1–7 (in Chinese)
2. Lan Z, Ni Y,Gan D (2005) A survey on transient stability control of modern power systems. Power Syst Technol 29(15):40–50 (in Chinese)
3. Yuan B, Sun Q (1994) Using bifurcation in power system. Power Syst Technol 18(4):13–17 (in Chinese)
4. Zhang Q, Wang B (2004) Controlling power system chaotic oscillation by time-delayed feedback. Power Syst Technol 28(7):23–26
5. Dou C, Zhang S, Yuan S (2003) Chaos forecast and control of power systems based on chaos neural network model. Electr Power 36(7):23–26
6. Liu M, Piao Z, Wu X (2009) Fuzzy sliding mode variable structure control for chaos oscillation of synchronous generator. Electr Power Autom Equip 29(7):85–88
7. Yuan L, Wu H, Tu J (2009) Integral sliding mode control for chaos oscillation of synchronous generator. Electr Power Autom Equip 29(7):85–88
8. Zhang Q, Wang B (2002) Chaos control with variable feedback in power system. Electr Power Autom Equip 22(10):22–25
9. Wang B, Zhang Q, Yang C (2003) Chaotic oscillation control of electric power system based on adaptive Backstepping. Electr Power Autom Equip 23(11):9–12
10. Tan W, Li Z, Zhang M (2010) Chaotic oscillation control of electric power system based on least square support vector machines. J Hunan Univ Sci Technol (Nat Sci Ed) 25(3):59–62
11. Tan Y (2010) Study on nonlinear dynamics in power system under noise perturbation. Guangxi Normal University, Guangxi
12. Zhang N, Sun Y, Yu L (2005) Terminal sliding mode control for class of chaotic systems. J Jilin Univ (Inf Sci Ed) 23(6):635–638
13. Song Y, Zhao G, Qi D (2006) Adaptive compensation control for chaos oscillation of power system. Proc CSU-EPSA 18(4):5–8

Chapter 220
Gain Self-Tuning of PI Controller for PMSM Servo System

Lili Cui, Ruibin Wu and Dongliang Liu

Abstract The PMSM servo system is a nonlinear, time-varying, and complex system. The tradition PI control is very difficult to realize, which is a line control and it needs the accurate mathematical model. Fuzzy control is proposed in this chapter, which synthesizes the fuzzy control and PI control. Using fuzzy inferences method, the parameters can be adjusted adaptively. For PMSM servo system, the vector control is used in current loop and fuzzy PI control is used in speed loop instead of traditional PI control. Results from numerical simulations and practical implementation show that the fuzzy PI controler has better performance and robustness than convertional PI controller in the PMSM servo system.

Keywords Fuzzy control · PI control · PMSM · Simulation

220.1 Introduction

Permanent magnet synchronous motor has been widely used in high performance drive applications for its advantages, such as compactness, high efficiency, reliability [1]. In the PMSM vector control system, the controller design plays important role in the system performance. Therefore, many scholars and experts are working on the motor control methods.

The PI control is used in the traditional PMSM servo system [2]. PI control which is simple in structure, and has good robustness and high reliability,

L. Cui · R. Wu (✉) · D. Liu
Hangzhou Dianzi university, Hang Zhou 310018, China
e-mail: wuruibin@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_220,

is suitable for a deterministic system which has accurate mathematical model. The system performance depends on PI parameters.

Two new methods are proposed to adjust the parameter in [2]. The results from experiment show that these methods improve the profermance of traditional PI control. But it relies on the accurate mathematical model of PMSM. PMSM is a nonlinear, time-varying,and complex system, and it is difficulte to built the accurate mathematical model of controlled object. So the tarditional PI controller is not suitable to higher performance.

These problems can be overcome by the fuzzy logic controllers [3], which do not require any mathematical model and are based on the linguistic rules obtained from the experience of the system operator. The results from experiment show that the performance of the fuzzy controller as compared to the PI controller is superior. It has good dynamic performance and good robustness. But, there are certain steady state error and the precision is poor the fuzzy control about the fuzzy control.

In [4], a fuzzy neural network controller is designed through synthesizes the fuzzy control and neural network control. The results from research show that the fuzzy neural network control system since learning ability, and has good fuzzy identification and generalization ability, can be used for nonlinear control accuracy requirements of high speed of special occasions. But the neural network, the network model express complex training time long.

In [5], genetic algorithm (GA) is used as optimization algorithms in permanent magnet synchronous motor system, and obtained a better optimization result. The GA is invented to mimic some of the processes observed in natural evolution. The GA does not have much mathematical requirements about the optimization problems. But it does not solve complexes constraint problems easily. And huge evaluations for generations and populations sometime are time-consuming.

In many intelligent control methods, the fuzzy control is used widely. In this chapter, the fuzzy control and PI control are synthesized. Using fuzzy inferences method, the parameters can be adjusted adaptively.comprehensive both combined the advantage of fuzzy control make PI, according to system performance requirements for the parameters of the setting, new all-digital fuzzy control method of vector control system applied to. Results from numerical simulations and practical implementation show that the fuzzy PI controler has better performancer in the PMSM servo system.

220.2 The Model of PMSM

The voltage equation of PMSM d-q axes is given by:

$$\frac{di_d}{dt} = -\frac{R}{L}i_d + P\omega i_q + \frac{1}{L}u_d \tag{220.1}$$

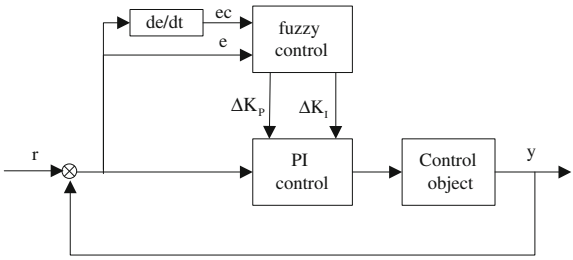


Fig. 220.1 Block diagram of fuzzy PI controller

$$\frac{di_q}{dt} = -\frac{R}{L}i_q - P\omega i_d - \frac{P\phi_f}{L}\omega + \frac{1}{L}u_q \tag{220.2}$$

$$\frac{d\omega}{dt} = \frac{3P\phi_f}{2J}i_q - \frac{B}{J}\omega - \frac{T_L}{J} \tag{220.3}$$

The torque equation of PMSM can be given in (220.4)

$$T_e = \frac{3}{2}p_n[\Psi_f i_q + (L_d - L_q)i_d i_q) \tag{220.4}$$

where u_d and u_q are the d-q axes stator voltages, respectively; i_d and i_q are the d-q axes stator currents; ω_r is the rotor electric angular velocity; L_d and L_q are the d-q axes stator inductances, respectively; ψ_f is the rotor magnetic flux linking the stator; p is the differential operator(d/dt).

220.3 Fuzzy-PI Controller Design

220.3.1 The Configuration of System

The PMSM servo control system is two closed loops control system. The PI controller is used in current loop. The fuzzy PI controller is adopted in speed loop instead of the traditional PI controller [2].

The parameters of PI can be adjusted on-line, which can be changed through the inquiry to fuzzy control rules table. The block diagram of control system is shown in Fig. 220.1.

220.3.2 Fuzzy Rules Regulation

In the fuzzy controller, speed error e and error change rate e_c are used as fuzzy input and the proportional factor Δk_p, the integral factor Δk_i are used as fuzzy output.

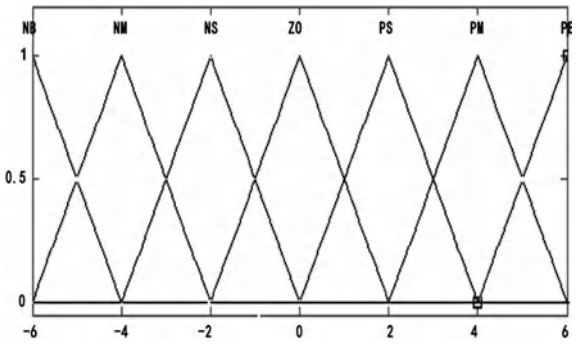


Fig. 220.2 Membership function of e, e_c, ΔK_P and ΔK_I

The fuzzy variable of input variables e and e_c is E and EC, the fuzzy variable of output variable Δk_p, Δk_i is ΔK_P, ΔK_I. The fuzzy sets of e, e_c, ΔK_P, ΔK_I, are all defined as $\{NB, NM, NS, ZO, PS, PM, PB\}$, where NB, NM, NS, ZO, PS and PM represent negative big, negative medium, negative small, zero, positive small, positive medium, positive big, respectively. The quantum rank of e, e_c, ΔK_P, ΔK_I is chosen as thirteen ranks e, e_c, ΔK_P, $\Delta K_I = \{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}$. The membership function of e, e_c, ΔK_P, ΔK_I is the triangle distribution. The membershio function of each variable is shown in Fig. 220.2. According to manual experience, the fuzzy rules are generalized as Table 220.1, Table 220.2 according to the expert experiment in PMSM servo system and simulation analysis of the system.

(2) Design of fuzzy control rules

220.3.3 The Quantum Factor

In this paper, the reference speed ω^* is 2,000 rad/s, keep the speed error in 3%, we can get the fuzzy domain of e is taken as $[-60, 60]$. The quantum factor of e is $K_e = 0.1$. The speed error derivative is kept in 5% of speed error. We can get the fuzzy domain of e_c is taken as $[-3, 3]$. The quantum factor of e_c is $K_{ec} = 2$. The quantum factor of ΔK_P and ΔK_I is $K_{up} = 0.1$, $K_{ui} = 0.5$.

220.3.4 Fuzzy Inference and Fuzzy Decision-Making

The Mamdani inference method is used as the fuzzy inference mode. The inference can be written as "If e is A and ec is B, then ΔK_P is C and ΔK_I is D".

According to fuzzy control rules in Table 220.1, a fuzzy set of output can be obtained by used Mamdani inference method when inputs are given:

If e is NB and ec is NB, then ΔK_P is PB and ΔK_I is NB.

Table 220.1 The control rulers of ΔK_P

ΔK_P		e_c						
		NB	*NM*	*NS*	*ZO*	*PS*	*PM*	*PB*
e	*NB*	*PB*	*PB*	*PM*	*PM*	*PS*	*ZO*	*ZO*
	NM	*PB*	*PB*	*PM*	*PS*	*PS*	*ZO*	*NS*
	NS	*PM*	*PM*	*PM*	*PS*	*ZO*	*NS*	*NS*
	ZO	*PM*	*PM*	*PS*	*ZO*	*NS*	*NM*	*NM*
	PS	*PS*	*PS*	*ZO*	*NS*	*NS*	*NM*	*NM*
	PM	*PS*	*ZO*	*NS*	*NM*	*NM*	*NM*	*NB*
	PB	*ZO*	*ZO*	*NM*	*NM*	*NM*	*NB*	*NB*

Table 220.2 The control rulers of ΔK_I

ΔK_I		e_c						
		NB	*NM*	*NS*	*ZO*	*PS*	*PM*	*PB*
e	*NB*	*NB*	*NB*	*NM*	*NM*	*NS*	*ZO*	*ZO*
	NM	*NB*	*NB*	*NM*	*NS*	*NS*	*ZO*	*ZO*
	NS	*PM*	*NM*	*NS*	*NS*	*ZO*	*PS*	*PS*
	ZO	*NM*	*NM*	*NS*	*ZO*	*PS*	*PM*	*PM*
	PS	*NM*	*NS*	*ZO*	*PS*	*PS*	*PM*	*PB*
	PM	*PS*	*ZO*	*PS*	*PS*	*PM*	*PB*	*PB*
	PB	*ZO*	*ZO*	*PS*	*PM*	*PM*	*PB*	*PB*

If *e* is *NB* and *ec* is *NM*, then ΔK_P is *PB* and ΔK_I is *NB*.

...

If *e* is *PB* and *ec* is *PB*, then ΔK_P is *NB* and ΔK_I is *PB*.

220.3.5 Fuzzy-PI Controller

The parameters of PI controller can be given by:

$$K_P = K_P^{'} + \Delta K_P \tag{220.5}$$

$$K_I = K_I^{'} + \Delta K_I \tag{220.6}$$

The $K_P^{'}$ is set to 0.4 and the $K_I^{'}$ is set to 10, and it is got from the traditional PI controller.

220.4 Simulation Results

The MATLAB software is used to the whole system simulation. The parameters of PMSM areas follows: rated power $P = 2{,}000$ W; rated phase voltage $U = 220$ V; rated speed $n = 2{,}000$ rpm; *d.q* axes inductances $L_d = L_q = 0.0153$ H; stator

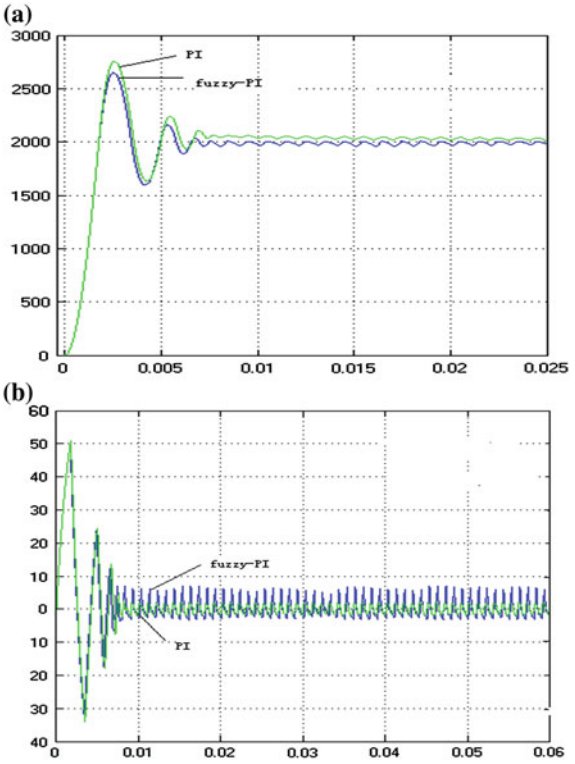


Fig. 220.3 Permanent magnet synchronous motor system dynamic response curve **a** Speed response curve of fuzzy PI control **b** Torque response curve of fuzzy PI control

resistance $Rs = 0.56\,\Omega$; number of pole-pairs $p_n = 4$; moment of inertia $J = 0.00045$ kg·m. The motor starts at rated speed at 2,000 r/min.

The simulation results compare of the fuzzy PI controller and the PI controller is shown in Fig. 220.3. It is obvious that the dynamic and steady performances of fuzzy PI control system are better than the PI control system.

220.5 Conclusion

This chapter designs a fuzzy PI controller, which has the advantages of PI controller and fuzzy controller. The controller is used as the speed control in PMSM control system, which can adjust the controller parameters online according to the speed error and the speed error change rate. The simulation results show that the fuzzy PI controller has faster response, and is stronger and robust compared to the traditional PI controller.

References

1. Hong W, Yong Y, Dianguo X (2004) The position servo system of PMSM [J]. In: Proceedings of the Chinese society for electical engineering
2. Haijun B, Qingding G (2003) Gain self-tuning methods of PI controllers for PMSM drivers [J]. J Shenyang Univ Technol
3. Sant AV, Rajagopal KR (2009) PM synchronous motor speed control using hybrid fuzzy-PI with novel switching functions [J]. IEEE Trans Magn 45(10):4672–4675
4. Cong S, Song R (2000) Animproved Bspline fuzzy neuralnetwork controller [C]. Intell Control Autom 3:1713–1717
5. Jiancheng L, Shoudao H, Wenjiao S, Jinping Z (2010) Optimal design of permanent magnet synchronous motors based on gene handing genetic algorithm [J]. Small Special Electr Mach
6. Sheng Y, Shoudao H, Gaolin L (2010) The application of fuzzy control in PMSM servo system [J]. Power Electron
7. Xiaochong C, Wanping W (2004) The emulation of PID controller and fuzzy self-adjusted PID controller [J]. Mach Tool Hydraul
8. Xudong G, Jidong Z (2007) Fuzzy-PID control for PMSM servo system [J]. J Harbin Univ Commer
9. Jingzhou Z, Weijing Y Research on fuzzy self-adaptive PID control and its emulation [J]. Comput Simul
10. Bangying Z, Wei S, Penghui H, Huiyuan Y (2008) Fuzzy PI control on permanent magnet synchronous motor [J]. J Jishou Univ

Chapter 221
Compositive Diagnosis Method of Turbine-Generator Air-Gap Eccentric Fault

Li Yonggang, Zhou Guowei, Wan Shuting and Li Heming

Abstract The characteristics of generator stator and rotor radial vibration and the characteristics of stator winding parallel branches circulating current on the air-gap eccentric fault are respectively analyzed. Three types of air-gap eccentric faults are studied, such as static eccentric, dynamic eccentric, and mixed eccentric composed by the previous two. Considering the differences in the stator and rotor vibration characteristics and circulating current characteristics caused by the faults, a compositive diagnosis method of air-gap eccentric fault based on the characteristics of vibration and circulating current is proposed. Finally, the data of a SDF-9 type generator on the fault condition are measured in the laboratory to verify the theoretical analysis presented above.

Keywords Air-gap eccentric fault · Compositive characteristics; generator · Stator winding parallel branches circulating current · Vibration

221.1 Introduction

As the core power generation equipment, the reliability and the security of the generator operation are pretty important. So it is pretty necessary to study the mechanisms and the methods of the generator fault diagnosis. In the groovy

L. Yonggang (✉) · Z. Guowei · L. Heming
Department of Elecrical Engineering, North China Electric Power University,
Box 20 Baoding 071003, China
e-mail: lygzxm0@163.com

W. Shuting
Department of Mechanical Engineering, North China Electric Power University,
Box 238 Baoding 071003, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_221,

monitoring and diagnosis system, it is emphasized to measure and analyze the rotor vibration signal. Since generators are different from the normal rotary machines, meanwhile, rotor and stator would link through the air-gap field, an air-gap eccentric fault would cause the air-gap magnetic with a deformation and make different electromagnetic force wave acting on rotor and stator, which not only makes the rotor vibrate, but also makes the stator vibrate. What's more, circulating current with somewhat characteristics would be created in the stator winding parallel branches at the same time. So studying the vibration characteristics and the circulating current characteristics on an air-gap eccentric fault condition may get a more comprehensive cognition to the effect that the generator faults act on the generator vibration and the stator winding parallel branch circulating current, and provide a more comprehensive omen for the kind of fault diagnosis to improve the sensitivity and the precision based on the current achievements.

Many current researches focus on the characteristics of rotor vibration and electricity on air-gap eccentric fault. The diagnosis, fault factors, preventive maintenance, unbalanced magnetic pulls and radial vibration of the rotor caused by the air-gap eccentric fault were first investigated [1–3]. Then the rotor radial vibration and axial vibration of this fault were presented [4]. The finite element method and the modified winding function theories have been used as powerful tools to calculate the electrical parameters [5, 6] and magnetic field parameters [7, 8] caused by air-gap eccentric fault. Based on these, the on-line monitoring for air-gap eccentric fault diagnosis developed well [8].

Based on the idea of these references above mentioned, this paper detailedly analyzes the rotor and stator vibration characteristics and the stator winding parallel branch circulating current characteristics of three types of air-gap eccentric faults, such as static eccentric, dynamic eccentric, and mixed eccentric composed by the previous two. Then put forward an air-gap eccentric fault compositive diagnosis method based on the rotor and stator vibration characteristics and the stator winding parallel branch circulating current characteristic (i.e. electromechanical compositive characteristic).

221.2 Analysis of the Air-Gap Field

When on a normal operation state (while $p = 1$), the magneto-motive force of the air-gap can be expressed as:

$$\begin{aligned} f(\alpha, t) &= F_r \cos(\omega_r t - \alpha) + F_s \cos(\omega_r t - \alpha - \psi - \frac{\pi}{2}) \\ &= F_1 \cos(\omega_r t - \alpha - \beta) \end{aligned} \tag{221.1}$$

where: F_r—the main magneto-motive force
F_s—the armature reaction magneto-motive force

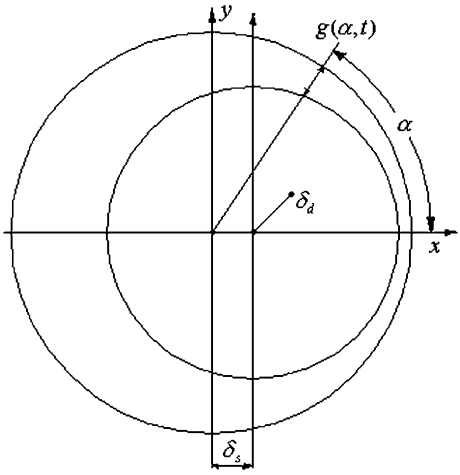


Fig. 221.1 Eccentric Air-gap of generator

$\omega_r = 2\pi f$—the mechanical angular frequency of the rotor
f_r—the mechanical frequency of the rotor (for turbine-generators, $\omega_r = \omega$, $f_r = f$)
α—the mechanical angle of the stator
ψ—the inner power-angle of the generator

$$F_1 = \sqrt{F_s^2 \cos^2 \psi + (F_r - F_s \sin \psi)^2}$$
$$\beta = arctg \frac{F_s \cos \psi}{F_r - F_s \sin \psi}$$

Considering the air-gap eccentric fault, choosing the coordinate as Fig. 221.1 and adopting the point that has the minimal air-gap as the grid origin, the air-gap is:

$$g(\alpha, t) = g(1 - \delta_s \cos \alpha - \delta_d \cos(\alpha - \omega_r t)) \tag{221.2}$$

where: g—the specific air-gap
δ_s—the relative static air-gap eccentricity
δ_d—the relative dynamic air-gap eccentricity

Since the air-gap eccentricity is little, to deploy the magnetic conductance with Fourier and ignore the high-order component, it can be expressed as:

$$\Lambda(\alpha, t) = \frac{1}{g(\alpha, t)} = \Lambda_0 - \Lambda_s \cos \alpha - \Lambda_d \cos(\alpha - \omega_r t) \tag{221.3}$$

where: Λ_0—the constant component of the air-gap magnetic conductance
Λ_s—the magnetic conductance component caused by the static eccentricity
Λ_d—the magnetic conductance component caused by the dynamic eccentricity

221.3 Analysis of the Rotor Vibration Characteristics

According to Ref. [9], the magnetic pull per unit area q and the unbalanced magnetic pulls (F_x, F_y) of x-axis and y-axis acting on the rotor can be expressed as:

$$\begin{aligned} q(\alpha,t) = \frac{F_1^2}{4\mu_0}\Bigg\{ & \left[\Lambda_0^2 + \frac{\Lambda_s^2}{2} + \frac{3\Lambda_d^2}{4} + \frac{\Lambda_s\Lambda_d}{2}\cos\alpha - 2\Lambda_0\Lambda_s\cos + \frac{\Lambda_s^2}{2}\cos 2\alpha\right] \\ & + \left[\frac{3\Lambda_s\Lambda_d}{4}\cos\omega_r t - 3\Lambda_0\Lambda_d\cos(\omega_r t - \alpha) + \Lambda_s\Lambda_d\cos(\omega_r t - 2\alpha)\right] \\ & + \left[\frac{\Lambda_s\Lambda_d}{2}\cos(2\omega_r t - \alpha) + (\Lambda_0^2 + \Lambda_d^2 + \Lambda_s^2\cos^2\alpha - 2\Lambda_0\Lambda_s\cos\alpha)\cos(2\omega_r t - 2\alpha)\right] \\ & - \Lambda_0\Lambda_d\cos(3\omega_r t - 3\alpha) + \frac{\Lambda_s\Lambda_d}{2}\cos(3\omega_r t - 4\alpha) + \frac{\Lambda_d^2}{4}\cos(4\omega_r t - 4\alpha)\Bigg\} \end{aligned} \tag{221.4}$$

$$\begin{cases} F_x = -\dfrac{\mathrm{LRF}_1^2\pi}{4\mu_0}(2\Lambda_0\Lambda_s + 4\Lambda_0\Lambda_d\cos\omega_r t + \Lambda_0\Lambda_s\cos 2\omega_r t) \\ F_y = -\dfrac{\mathrm{LRF}_1^2\pi}{4\mu_0}(2\Lambda_0\Lambda_s + 4\Lambda_0\Lambda_d\sin\omega_r t + \Lambda_0\Lambda_s\sin 2\omega_r t) \end{cases} \tag{221.5}$$

The characteristics of the rotor radial vibration on static eccentric fault ($\Lambda_s \neq 0$, $\Lambda_d = 0$), dynamic eccentric fault ($\Lambda_s = 0$, $\Lambda_d \neq 0$), and mixed eccentric fault ($\Lambda_s \neq 0$, $\Lambda_d \neq 0$) can be gotten:

(1) The unbalanced magnetic pulls caused by the static eccentricity can be expressed as $F_x = -\dfrac{\mathrm{LRF}_1^2\pi\Lambda_0\Lambda_s}{4\mu_0}\cos 2\omega_r t$ and $F_y = -\dfrac{\mathrm{LRF}_1^2\pi\Lambda_0\Lambda_s}{4\mu_0}\sin 2\omega_r t$ would cause the rotor vibration at $2f_r$.
(2) The unbalanced magnetic pulls caused by the dynamic eccentricity can be expressed as $F_x = -\dfrac{\mathrm{LRF}_1^2\pi\Lambda_0\Lambda_d}{\mu_0}\cos\omega_r t$ and $F_y = -\dfrac{\mathrm{LRF}_1^2\pi\Lambda_0\Lambda_d}{\mu_0}\sin\omega_r t$ would cause the rotor vibration at f_r.
(3) The mixed eccentric fault would cause the rotor vibration at f_r and $2f_r$.

221.4 Analysis of the Stator Vibration Characteristics

The stator core is made up of lots of overlapped silicon sheets, and its mechanical model is a hollow elastic shell cylinder. The essentially influential factor for the stator vibration characteristics is the pulsating peculiarity of the magnetic pull

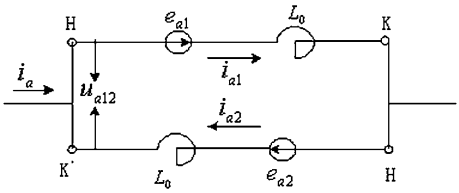


Fig. 221.2 Stator winding parallel-connected branches

acting on the stator. Therefore, the stator vibration characteristics can be studied via the magnetic pull per unit area.

According to Eq. (221.4) and Ref. [9], the characteristics of the stator vibration on static eccentric fault, dynamic eccentric fault, and mixed eccentric fault can be gotten:

(1) When the static eccentric fault occurs, the magnetic pull per unit area on stator there are constant components and pulsating components of $2f_r$, and cause the stator vibration at $2f_r$.
(2) When the dynamic eccentric fault occurs, the magnetic pull per unit area on stator there are constant components and pulsating components of f_r, $2f_r$, $3f_r$ and $4f_r$, and cause the stator vibration at f_r, $2f_r$, $3f_r$ and $4f_r$.
(3) When the mixed eccentric fault occurs, it will cause the stator vibration at f_r, $2f_r$, $3f_r$ and $4f_r$.

221.5 Analysis of Stator Winding Parallel Branches Circulating Current Characteristics

The stator windings are double-wye connected in turbogenerators, as shown in Fig. 221.2, each phase has two windings, which are the generator's stator parallel branches. The electromotive force difference of the two branches can be expressed:

$$u_{a12} = e_{a1} + i_{a1}L_0 + i_{a2}L_0 + e_{a2} \tag{221.6}$$

Under normal condition, $w_{a1} = w_{a2}$ and $\Lambda_{a1} = \Lambda_{a2}$, the odd-time harmonic components magnetic force would not induce electric potential difference to the stator winding parallel branches. so that there would not be any circulating current in the branches.

When the static eccentricity fault occurs, $\Lambda(\alpha,t) = \Lambda_0 - \Lambda_s\cos\alpha$. Because α of the first branch is not equal to that of the second branch, and the electromotive force difference and circulating current will occur. Because there is only odd mmf of air-gap on normal rotor winding, especially fundamental harmonic, the fundamental circulating current will occur when static eccentricity fault occurs.

Table 221.1 Characteristics of the rotor vibration and circulating current

Eccentric type	Rotor vibration characteristics	Stator vibration characteristics	Circulating current characteristics
Static eccentric	$2f_r$	$2f_r$	Fundamental harmonic circulating current increases
Dynamic eccentric	f_r	f_r , $2\,f_r$, $3\,f_r$, $4\,f_r$	Fundamental harmonic circulating current increases
Mixed Eccentric	f_r, $2f_r$	f_r , $2\,f_r$, $3\,f_r$, $4\,f_r$	Fundamental harmonic circulating current increases

When the dynamic eccentricity fault occurs, $\Lambda(\alpha,t) = \Lambda_0 - \Lambda_d \cos(\alpha - \omega_r t)$, if for the first branch $\alpha = 0$, then $\alpha = \pi$ would be for the second branch. Since the change of the induced electric potential difference caused by a dynamic eccentricity in the first branch is:

$$\begin{aligned}\delta_{a1} &= 3.99 f w k_{wr} l \tau F_{\gamma m} \Lambda_d \cos(\omega t) \cos\gamma(\omega t) \\ &= 2.0 f w k_{wr} l \tau F_{\gamma m} \Lambda_d \cos(\omega t) \cos\gamma(\omega t)\end{aligned} \tag{221.7}$$

The change of the induced electric potential difference caused by a dynamic eccentricity in the second branch would be:

$$\begin{aligned}\delta_{a2} &= 3.99 f w k_{wr} l \tau F_{\gamma m} \Lambda_d \cos(\omega t - \pi) \cos\gamma(\omega t - \pi) \\ &= 2.0 f w k_{wr} l \tau F_{\gamma m} \Lambda_d \cos(\omega t - \pi) \cos\gamma(\omega t - \pi)\end{aligned} \tag{221.8}$$

When γ is even, $\delta_{a1} + \delta_{a2} = 0$, therefore the electromotive force difference and the circulating current of even harmonic is equal to zero. When γ is odd, $\delta_{a1} + \delta_{a2} = 2\delta_{a1} > 0$, the electromotive force difference and the circulating current of odd harmonics aren't equal to zero. There is only an odd mmf of air-gap on normal rotor winding, especially fundamental harmonic. So the fundamental electromotive force difference and circulating current will occur when the dynamic eccentricity fault occurs.

221.6 Compositive Fault Diagnosis Method for Air-Gap Eccentric Fault

Based on the above analysis, the characteristics of rotor and stator vibration and the characteristics of stator winding parallel branches circulating current under static eccentric fault, dynamic eccentric fault and mixed eccentric fault can be gotten (as Table 221.1). Obviously the air-gap eccentric cannot be distinguished by traditional method only according with circulating current or rotor and stator vibration characteristics.

Considering the differences of the stator and rotor vibration characteristics and stator winding parallel branches circulating current characteristics caused by three

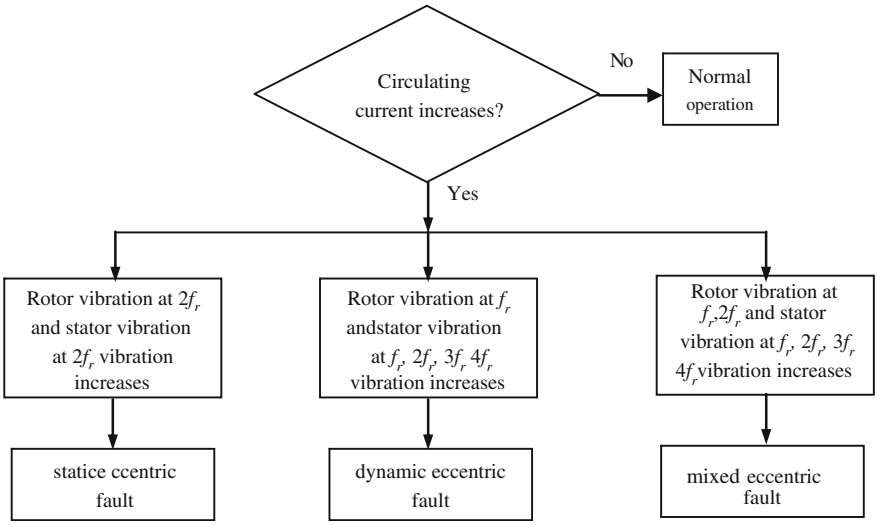


Fig. 221.3 Diagnosis flow diagram

types of air-gap eccentric faults, such as static eccentric, dynamic eccentric, and mixed eccentric composed by the previous two, a new compositive diagnosis method of turbine-generator air-gap eccentric fault based on the characteristics of the stator and rotor vibration and circulating current is put forward (viz. electro-mechanical compositive characteristic). The diagnosis flow diagram is shown in Fig. 221.3.

221.7 Experiment Analysis

The experimental generator is a SDF-9 fault simulation generator (as Fig. 221.4a), and a static eccentric fault can be simulated (as Fig. 221.4b and c).

The parameters are as follows: rated capacity 7.5 kVA; rated voltage 400 V; rated rotational speed n_r = 3000 r/min, with number of pole-pairs p = 1; stator winding is double layer short-range winding; air-gap is 0.8 mm.

Figure 221.5 is the rotor vibration spectrum, and it is obvious that the rotor vibration at $2f_r$ would increase when the generator has a static eccentricity.

Figure 221.6 is the stator vibration spectrum, it is obvious that the stator vibration at $2f_r$ would increase when the generator has a static eccentricity.

Figure 221.7 is stator winding parallel branches circulating current spectrum. Under normal condition, the circulating current should not exist in the parallel branches in theory. However, it dose exit in the experiment system due to the asymmetry inside the generator. And it is obvious that fundamental harmonic circulating current increases more than other harmonic components when the generator has a static eccentricity.

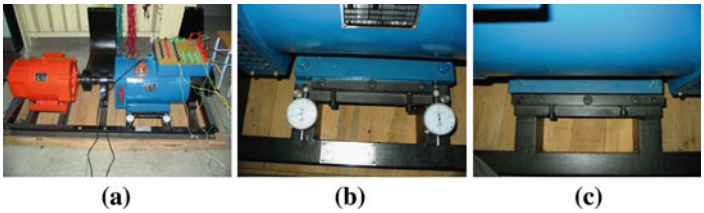


Fig. 221.4 SDF-9 fault simulation generator

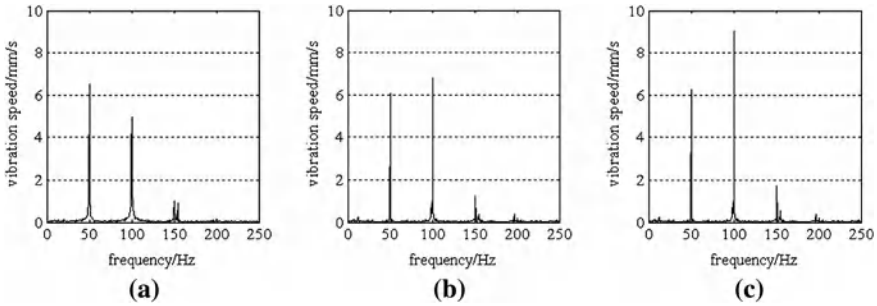


Fig. 221.5 Vibration spectrum of rotor. **a** Normal operation. **b** 0.2 mm static eccentricity. **c** 0.4 mm static eccentricity

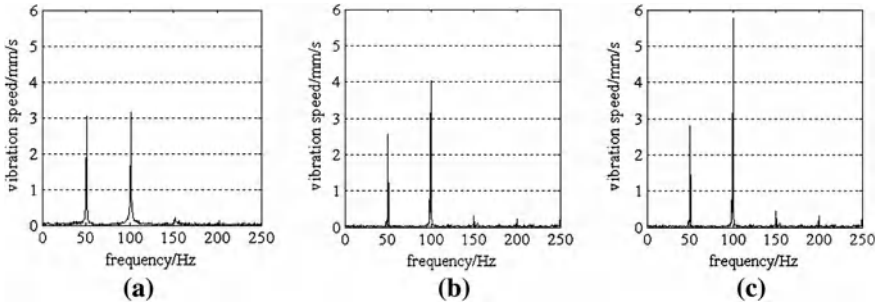


Fig. 221.6 Vibration spectrum of stator. **a** Normal operation. **b** 0.2 mm static eccentricity. **c** 0.4 mm static eccentricity

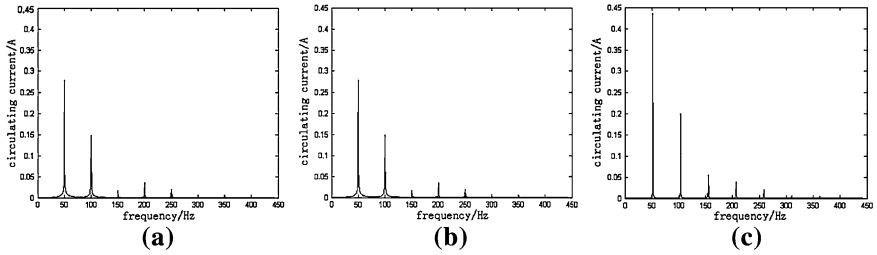


Fig. 221.7 Stator winding parallel branches circulating current spectrum. **a** Normal operation. **b** 0.2 mm static eccentricity. **c** 0.4 mm static eccentricity

Limited by the experimental condition, the experiment of the dynamic eccentric fault has not been successfully completed. From the experimental data above, it can see that the experiment results comply with the theoretical analysis.

221.8 Conclusion

This paper investigates the characteristics of generator stator and rotor radial vibration and the characteristics of stator winding parallel branches circulating current under normal condition and the air-gap eccentricity fault condition. The experiment verifications are also provided. The conclusions drawn from the theoretical and experimental investigation can be given as follows:

(1) The air-gap static eccentric fault would cause the rotor vibration at $2f_r$, the dynamic eccentric fault would cause the rotor vibration at f_r, and the mixed eccentric fault would cause the rotor vibration at f_r and $2f_r$.
(2) The air-gap static eccentric fault would cause the stator vibration at $2f_r$, the dynamic eccentric fault and the mixed eccentric fault would cause the stator vibration at f_r, $2f_r$, $3f_r$ and $4f_r$.
(3) The fundamental harmonic components of the stator winding parallel branches circulating current will increase when the air-gap eccentric fault occurs, including static eccentric fault, dynamic eccentric fault and mixed eccentric fault.
(4) Considering the differences in the stator and rotor vibration characteristics and circulating current characteristics caused by three types of air-gap eccentric faults, a compositive diagnosis method of air-gap eccentric fault based on the characteristics of vibration and circulating current is gained.

References

1. Wang L, Cheung RW, Ma Z, Ruan J, Peng, Y (2008) Finite-element analysis of unbalanced magnetic pull in a large hydro-generator under practical operations. IEEE Trans Magn 44(6):1558–1561
2. Perers R, Lundin U, Leijon M (2007) Saturation effects on unbalanced magnetic pull in a hydroelectric generator with an eccentric rotor. IEEE Trans Magn 43(10):3884–3890
3. Keller S, Tu Xuan M, Simond J-J, Schwery A (2007) Large low-speed hydro-generators-unbalanced magnetic pulls and additional damper losses in eccentricity conditions. IET Electr Power Appl 1(5):657–664
4. De Canha D, Cronje WA, Meyer AS, Hoffe SJ (2007) Methods for diagnosing static eccentricity in a synchronous 2 pole generator. 2007 IEEE Power Technology Conference Proceedings, Lausanne, Switzerland
5. Jiahui Z, Arui Q, Guo T (2008) Branch voltage of a salient pole synchronous generator with eccentric rotor and skewed slots. J Tsinghua Univ (Sci Technol) 48(4):453–456

6. Toliyat HA, Al-Nuaim NA (1999) simulation and detection of dynamic air-gap eccentricity in salient-pole synchronous machines. IEEE Trans Ind Appl 35(1):86–93
7. Tabatabaei I, Faiz J, Lesani H, Nabavi-Razavi MT (2004) Modeling and simulation of a salient-pole synchronous generator with dynamic eccentricity using modified winding function theory. IEEE Trans Magn 40(3):1550–1555
8. Bruzzese C, Giordani A, Santini E (2008) Static and dynamic rotor eccentricity on-line detection and discrimination in synchronous generators By No-Load E.M.F. space vector loci analysis. 19th international symposium on power electronics, electrical drives, automation and motion, Ischia, Italy
9. Wan S, Li H, Li Y (2005) Analysis on vibration characteristics of generator with the fault of eccentric air-gap. J Vib Shock 24(6):21–23

Chapter 222
The Vector Control Research of Permanent Magnet Linear Synchronous Motor

Jun Zhu, Xu Dong Wang, Bao Yu Xu and Hai Chao Feng

Abstract In order to reduce the speed and torque ripples, a vector control scheme for permanent magnet linear synchronous machine (PMLSM) based on PWM control is proposed in this paper. The proposed control method uses the space vector coordination transformation method to change the stator abc coordination current into *dq* synchronous rotating coordination, and the *q* axis voltage component was obtained to control the torque directly. By establishing the PWM vector control model based on $i_d = 0$, the simulation experiment was conducted. It is showed that the speed and torque ripples are greatly reduced and the quick dynamic response characteristic is still retained, and that the PWM vector control based on $i_d = 0$ scheme can well control the PMLSM obviously.

Keywords Vector control · PWM · PMLSM · dq rotating coordinate

J. Zhu (✉) · X. D. Wang · H. C. Feng
School of Electrical Engineering and Automation, Henan polytechnic University, No.2001, Century Avenue, Jiaozuo, China
e-mail: Zhujun@hpu.edu.cn

X. D. Wang
e-mail: wangxd@hpu.edu.cn

H. C. Feng
e-mail: fenghc@hpu.edu.cn

B. Y. Xu
School of Mechanical and Power Engineering, Henan polytechnic University, No.2001, Century Avenue, Jiaozuo, China
e-mail: xubaoyu@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_222,

222.1 Introduction

The basic working principle of permanent magnet linear synchronous motor (PMLSM) is same as the permanent magnet rotating synchronous motor (PMRSM), so the PMLSM can be seen as a PMRSM cut along the radial and stretched the linear style. As a kind of linear drive unit and direct driver system, PMLSM makes the load connect to the rotor directly, the control object and motor construct an integrated structure. The direct driver model of PMLSM, which converts electrical energy directly into mechanical energy without any other intermediate drive gearing, can eliminate the advance effects of PMRSM connected through link bar, such as backlash, high friction, high inertia, etc. Because of the simple direct driven form, high stiffness structure, and direct driver model, the PMLSM can obtain more higher speed and accelerate than the traditional PMLSM. It has a lot of advantages, for example, high force density, lower inertia, high positioning precision and accuracy, high dynamic performance, unlimited travel, and so on. Then the PMLSM can be widely used in electromechanical servo system, high-speed and high precision digital control system, precision instrument, precision machining, horizontal and vertical and transportation systems.

However, the direct driver model without intermediate drive transmission makes the PMLSM sensitive to internal system perturbation and external disturbances. It has the problem of end effect, nonlinear model and parameter uncertainty to bring the thrust ripple, slot, and end effects, which can reduce the servo performance of PMLSM. The disadvantages of PMLSM make the system control more difficult, so a much higher requirement was proposed to design the controller to improve the system performance of PMLSM [1, 2].

The vector control is a kind of usual control method for AC servo motor and it is being used in controlling PMLSM. According to the multivariable, nonlinearity, and strong coupling of PMLSM, the vector control converts the stator current of PMLSM into rotor synchronous current by using the vector conversion technology. That separates the flux and torque currents so as to linearly control the output torque of PMLSM, so that the thrust ripple and end effect can be reduced, and the servo performance was improved for the PMLSM servo system.

222.2 The Mathematical Model of PMLSM

The paper proposed that the PMLSM has a sinusoidal back-EMF, it has symmetric three-phase sine waveform of stator voltage and current, its magnet circuit is unsaturated, excluding the eddy current and hysteresis losses, without damper winding and zero conductivity [3–5]. The space vector axis of PMLSM is shown in Fig. 222.1

According to the above assumption, the vector control of PMLSM can be converted into the rotor *dq0* axis, and the mathematical model was obtained [6].

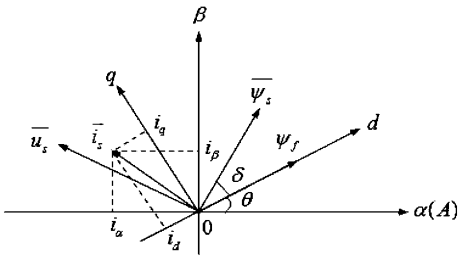


Fig. 222.1 The space vector axis of PMLSM

The equations of voltage, flux linkage, torque,and motion for PMLSM are as follows:

The voltage equations are:

$$u_d = R_s i_d + p\psi_d - \omega_r \psi_q \tag{222.1}$$

$$u_q = R_s i_q + p\psi_q + \omega_r \psi_d. \tag{222.2}$$

In Eqs. (222.1) and (222.2), p is differential operator, R_s is armature winding resistance, and ω_r is rotor angular velocity. They show d and q components of the stator voltage in the rotor synchronous $dq0$ axis.

The flux linkage equations are:

$$\psi_d = L_d i_d + \psi_f \tag{222.3}$$

$$\psi_q = L_q i_q. \tag{222.4}$$

In Eqs. (222.3) and (222.4), ψ_f is flux linkage constant of permanent magnet, L_d and L_q are dq coil self-inductances.

The electromagnetic torque equation is:

$$T_e = \frac{3}{2} p_n(\psi_d i_q - \psi_q i_d) = \frac{3}{2} p_n[\psi_f i_q + (L_d - L_q) i_d i_q] \tag{222.5}$$

In Eq. (222.5), T_e is electromagnetic torque, and p_d is motor pole pairs.

The mechanical motion equation is:

$$J\frac{\omega_r}{\mathrm{d}t} = T_e - T_L - B\omega_r. \tag{222.6}$$

The variable J is moment inertia, and B is friction coefficient.

In the static two-phase $\alpha\beta$ axis, the voltage, flux linkage, and torque are as follows:

$$\begin{bmatrix} u_\alpha \\ u_\beta \end{bmatrix} = \begin{bmatrix} R_s + pL_a & 38; pL_{\alpha\beta} \\ pL_{\alpha\beta} & 38; R_s + pL_\beta \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} + \omega_r \psi_f \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \tag{222.7}$$

$$\psi_\alpha = \int (u_\alpha - R_s i_\alpha)\mathrm{d}t \tag{222.8}$$

$$\psi_\beta = \int (u_\beta - R_s i_\beta)\mathrm{d}t \tag{222.9}$$

$$T_e = \frac{3}{2} p_n \left(\psi_\alpha i_\beta - \psi_\beta i_\alpha\right) \tag{222.10}$$

In Eqs. (222.7)–(222.10), u_α and u_β are $\alpha\beta$ axis components of the stator voltage. i_α and i_β are the stator current $\alpha\beta$ axis components. ψ_α and ψ_β are $\alpha\beta$ axis components of the stator flux linkage.

For the vector control of PMLSM, three kinds of coordinate changes are used, they are stator three phase *abc* coordinate, stator two-phase static $\alpha\beta$ coordinate, and two-phase *dq* coordinate based on the rotating magnetic field of rotor. They have the following conversion relationship.

$$\begin{bmatrix} F_d \\ F_q \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \cos\theta & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \\ \sin\theta & \sin\left(\theta - \frac{2\pi}{3}\right) & \sin\left(\theta + \frac{2\pi}{3}\right) \end{bmatrix} \begin{bmatrix} F_a \\ F_b \\ F_c \end{bmatrix} \tag{222.11}$$

$$\begin{bmatrix} F_\alpha \\ F_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} F_a \\ F_b \\ F_c \end{bmatrix} \tag{222.12}$$

$$\begin{bmatrix} F_d \\ F_q \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} F_\alpha \\ F_\beta \end{bmatrix} \tag{222.13}$$

In Eqs. (222.11)–(222.13), F can be expressed voltage, current, and flux linkage, respectively.

222.3 The Vector Control of PMLSM

222.3.1 The System Structure of Vector Control

The vector control can reduce the torque ripple, end effect and improve the dynamic response performance for the torque control research. The vector control strategy of PMLSM is same as PMRSM. There are some mainly control methods: (a) $i_d = 0$ control method, (b) maximum torque-current ratio control method, (c) $\cos\varphi = 1$, constant flux linkage control methods, etc.

The $i_d = 0$ control method is a simple and high efficient current control scheme. It does not have a demagnetizing effect without *d* axis demagnetizing component for the armature reflect. The performance of PMLSM will not be deterioration for the demagnetizing phenomenon, so that the electromagnetic torque is proportional to the armature current, that is $T_e = p_n \psi_f i_q$. The system principle of vector control for PMLSM is based on $i_d = 0$ as given in Fig. 222.2.

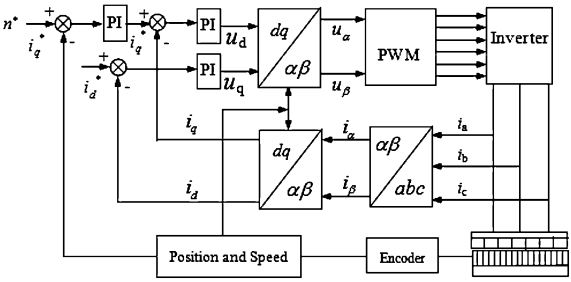


Fig. 222.2 $i_d = 0$ vector control principle of PMLSM

The control scheme mainly contains several control modules. They are stator current detection, rotor position and speed detection, speed loop regulator, Clark conversion, Park transformation and inverse transformation, and voltage PWM control modules. It is realized as the following procedures, the rotor space position is detected by using sensor, the rotor speed and electrical angle are calculated, the q axis reference i_q^* of stator current is output from the PI speed regulator. The stator phase current is detected through current sensor. It is decomposed to obtain the dq axis components i_d and i_q. The space vector voltage dq axis components u_d and u_q are forecasted through the current regulator, the PWM signal drives the inverter to provide the voltage to PMLSM, so that we can obtain the high efficient control method for PMLSM.

222.3.2 The Simulation Model of Vector Control for PMLSM

According to the vector control principle, the PMLSM vector control model was established based on PWM, it contains PMLSM module, the coordinate transformation, and inverse transformation from stationary coordinate to synchronous rotating coordinate, speed regulator loop, and current regulator loop, then a dual closed loop vector control system is built for PMLSM based on PWM control. Its simulation model is given in Fig. 222.3

222.3.3 The Simulation Experiment of Vector Control for PMLSM

The vector control of PMLSM based on PWM can be verified through the experiment simulation. The system simulation parameters are given as follows:

Number of pole pairs is 3, permanent magnet flux ψ_f is 0.2324 Wb, rotor resistance R_s is 1 Ω, d axis inductance L_d is 13.91 mh, q axis inductance L_q is 13.91 mh, viscous friction coefficient B is 0.1 N s/m, rotor mass is 96 kg, pole pitch τ is 39 mm.

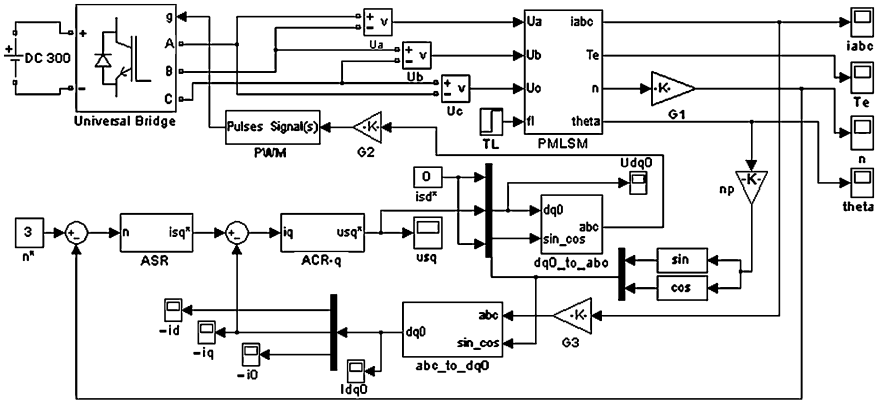


Fig. 222.3 The vector control simulation model of PMLSM

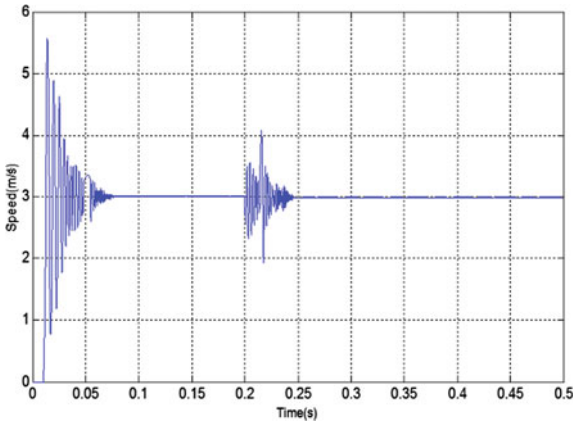


Fig. 222.4 Speed response of PMLSM

In the experiment, the given load torque was 1000 N M and velocity was 3.0 m/s, at 0.2 s the load torque was changed as 1000 N M, after 0.5 s, then the variable state of speed, torque, stator *abc* axis current, and *dq* axis current can be obtained along the whole loading process, their response state, respectively as Figs. 222.4, 222.5, 222.6, and 222.7.

Although the simulation experience results have some overshoot, the responses of speed and torque are very quick, they can recover the setting variety in time and reduce the ripple successfully when the PMLSM is starting and changing the load. It showed that the vector control can provide a greatly dynamic performance for PMLSM in Figs. 222.4 and 222.5.

In Figs. 222.6 and 222.7, we can see the stator current at static *abc* coordination and synchronous rotating $\alpha\beta$ coordination respectively. It showed that the q axis

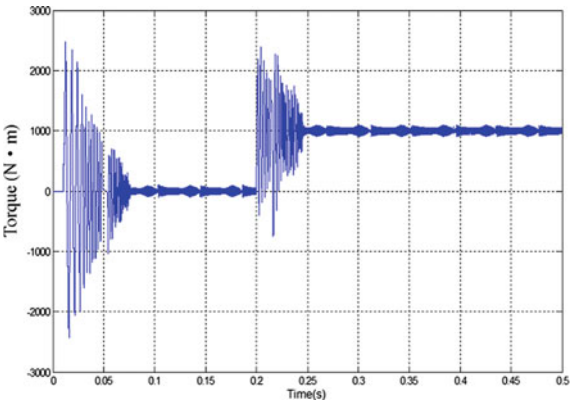


Fig. 222.5 Torque response of PMLSM

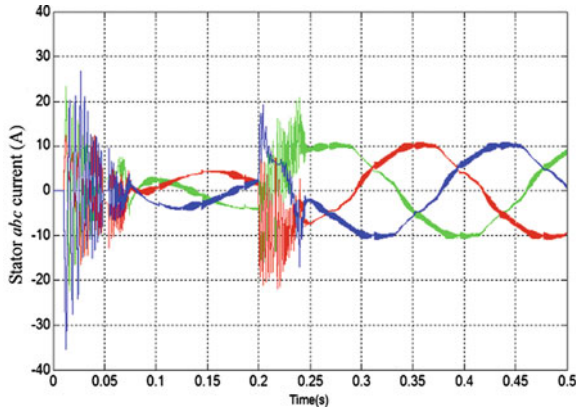


Fig. 222.6 Stator *abc* axis current

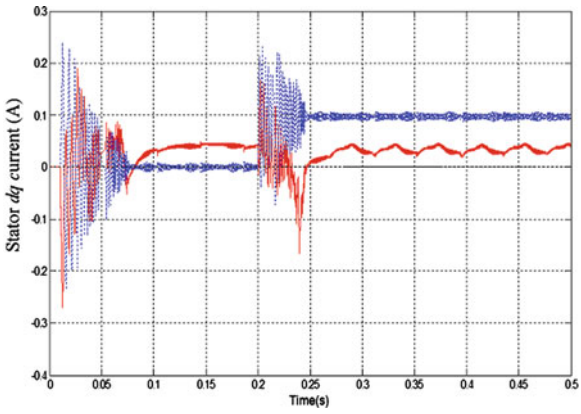


Fig. 222.7 Stator *dq* axis current

current component can separate successfully, so that we can control the speed and torque directly and obtain perfectly dynamic performance for the PMLSM through vector control.

222.4 Conclusion

In this paper, a vector control model based on PWM for PMLSM has been proposed. The detailed vector controller model for PMLSM is given, and the $i_d = 0$ vector control method is simulated. It is known that the proposed method can separate the q axis current component accurately through *abc*/*dq* coordinate transformation, and control the electromagnetic torque directly without the effect of d axis component. From the experiment, although the speed and torque had some ripples when the load was changing, it is seen that the speed and toque can be followed perfectly. With the vector control scheme of PMLSM, it shows that the flux and torque ripples can be reduced obviously, the PMLSM drive system still keeps the quick dynamic response characteristic.

Acknowledgments The paper is supported by the Dr. Foundation of henan polytechnic university under grant B2011-104 and National Science Foundation of China (61074095).

References

1. Wang H, Zhang ZJ, Liu CY (2010) Compensation methods of longitudinal endeffects in permanent-magnet linear synchronous motor. Proc CSEE 30(36):46–52
2. Yang JY, Ma H, Guan LR, Yang S (2010) Two-dimensional segmented synthesis iterative learning control of permanent magnet linear motor. Proc CSEE 30(30):74–80
3. Chen R, Yan YG (2004) Control project research on permanent magnet synchronic servo system. Electric machines and control 8(3):205–208
4. Deng GM (2010) Simulation of control system on permanent magnet synchronous motor. J Longyan Univ 28(5):28–31
5. Sun HY, Huang XD, Hong RJ, Pei L (2010) Simulation research on vector control system of permanent magnet synchronous motor. Mach Des Manuf 3:122–124
6. Lin WJ (2005) The control strategy research of permanent synchronous motor servo system. Ph.D. thesis of Zhejiang University

Chapter 223
Finite Element Analysis and Optimization for Magnetic Circuit System of Moving-Coil Linear Compressor

Peng Zhao, Shulian Liu and Shuiying Zheng

Abstract In order to solve the calculation problem which was caused by the large mesh quantity and convergent issue when finite element model was established according to the real magnetic circuit system structure of moving-coil linear compressor, a new modeling method was proposed. Based on ANSYS parametric design language, a parametric finite element model for magnetic circuit system of moving-coil linear compressor was established and the magnetic analysis was carried out. The validity and reliability of the finite element model was verified through magnetic experiment. According to the magnetic analysis results, the structural optimization design for magnetic circuit system was accomplished and a reasonable schematic was achieved. The optimization results show that not only the magnetic induction of gas gap is increased by 28.16%, but a permanent magnet is also operating at the optimal point and the magnetic saturation of iron core has been removed thoroughly.

Keywords Moving-coil linear compressor · Magnetic circuit system · Parametric modeling · Magnetic analysis · Optimization design

P. Zhao (✉) · S. Zheng
Institute of Chemical Machinery,
Zhejiang University, Hangzhou 310027, Zhejiang, China
e-mail: zpzju1984@126.com

S. Zheng
e-mail: zhengshuiying@zju.edu.cn

S. Liu
Department of Electromechanical Engineering,
Zhejiang University of Science and Technology,
Hangzhou 310012 , China
e-mail: liushul21@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_223,

223.1 Introduction

Performance and efficiency of moving-coil linear compressor are directly affected by characteristic of magnetic circuit system, as a result, the optimization for the system must be carried out to design the optimal structural parameters. Zhao ke [1] and Liu Xiaohui [2] have carried out the magnetic analysis of a magnetic circuit system by the 2D finite element method. The magnetic characteristics of a linear compressor were studied using traditional magnetic circuit analysis method by Min-Fu Hsieh [3] and Chen [4]. However, the influence that is caused by iron coil's and permanent magnet's shape on magnetic circuit system cannot be taken into account by both the 2D finite element method and the traditional magnetic circuit analysis method.

There was a gap in iron core of the moving-coil linear compressor. If the model was established according to the real structure, the mesh quantity would be very large, and the calculation could not be converged. Consequently, a new modeling method was proposed in this paper. Based on ANSYS parametric design language, a parametric finite element model was established and the magnetic analysis was carried out. The validity and reliability of the finite element model was verified through magnetic experiment. In terms of the problem and the corresponding reason of original magnetic circuit system, the optimal structural parameters were obtained by optimization design.

223.2 Finite Element Theory of Electromagnetic Field

Electromagnetic field theory is described by maxwell equations, which are composed of the ampere circuital theorem, the faraday electromagnetic induction law, the gauss law of electric flux and the gauss law of magnetic flux [5, 6]. The maxwell equations are shown as follows.

$$\begin{cases} \oint_{\Gamma} H\mathrm{d}l = \iint_{\Omega}\left(J + \dfrac{\partial D}{\partial t}\right)\mathrm{d}S \\ \oint_{\Gamma} E\mathrm{d}l = -\iint_{\Omega}\left(J + \dfrac{\partial B}{\partial t}\right)\mathrm{d}S \\ \oiint_{S} D\mathrm{d}S = \iiint_{V} \rho \mathrm{d}v \\ \oiint_{S} B\mathrm{d}S = 0 \end{cases} \tag{223.1}$$

From above, Γ is the boundary of surface Ω, $\boldsymbol{H}$ is the magnetic field intensity, $\boldsymbol{J}$ is the vector density of conduction current, $\boldsymbol{D}$ is the electric flux density, $\boldsymbol{E}$ is the electric ield intensity, $\boldsymbol{B}$ is magnetic induction, ρ is the volume density of charge, and V is the volume enclosed by closed surface S.

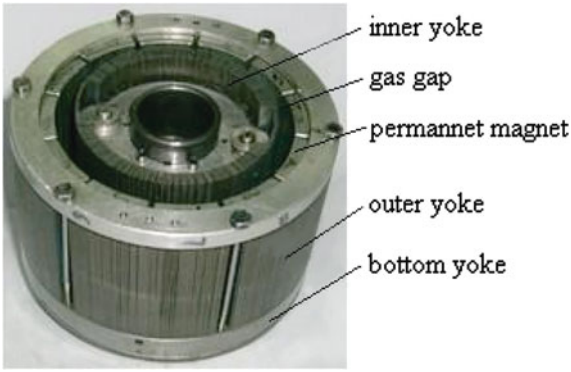


Fig. 223.1 Magnetic circuit system prototype

The electromagnetic analysis is based on maxwell equations and the object is divided into finite elements in ANSYS. According to vector magnetic potential or scalar electrical potential, the magnetic flux satisfying certain boundary and initial conditions can be solved, and other quantities will be derived from it, such as energy loss, magnetic force a magnetic moment, and so on.

223.3 Parametric Model of Magnetic Circuit System

According to the structural characteristics, the magnetic circuit system model could be parametric in order to avoid lot of duplicate work. The finite element analysis process will be made automatic by changing each parameter, as a result, the efficiency of optimization design is greatly increased.

The magnetic circuit system of moving-coil linear compressor is mainly composed of a cylinder, a inner core of inner yoke, the inner yoke, the gas gap, the permanent magnet, a inner core of outer yoke, the outer yoke and the bottom yoke, which is shown as Fig. 223.1.

The 3D finite element model of magnetic circuit system was built. However, based on its axisymmetric structure, 1/4 of model was only required. Air model was also built around the system in order to study the flux magnetic leakage. According to the experiment, there is nearly no magnetic flux at the place of 20 mm outer the system, as a result, the air model could be built within the bounds of 20 mm. Inner and outer yoke were formed by silicon steel sheets superimposed in the circumferential direction , and then there has gap in iron core. If the model was built according to the real structure, the mesh quantity would be very large, and even the calculation could not be converged. The gap among silicon steel sheets was concentrated in the finite element model on the basis of equal magnetic flux area. The parametric model was between as Fig. 223.2.

Edge element method was adopted and the type element of SOLID117 was used in this model. There are several mainly important results, such as the

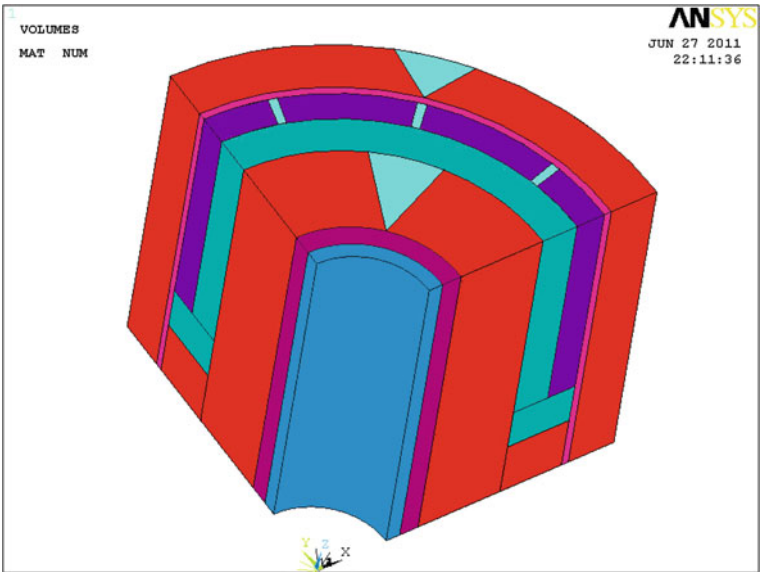


Fig. 223.2 Parametric model of magnetic circuit system

distribution of magnetic induction in gas gap, the magnetic saturation of iron core and the working point of permanent magnet. Consequently, the mesh density at the above regional must be increased. The mesh model is shown in Fig. 223.3. The boundary condition of magnetic flux parallel was loaded to all the external surfaces of the model.

223.4 Finite Element Analysis of Magnetic Circuit System Model

Finite element model of the magnetic circuit system was built based on the original size of moving-coil linear compressor in our laboratory. In order to study the influence of parameters on the magnetic circuit system, the finite element analysis and optimization design have been done with the purpose of making the permanent magnet work at the best operating point, the average gas magnetic induction reach 0.5T and the magnetic saturation of iron core be removed. The distribution of magnetic induction of original magnetic circuit system is shown in Fig. 223.4.

According to the average gas magnetic induction, the finite element result was 0.4148 T and the experimental data was 0.4227 T. They had an error of no more than 1.869%, which showed that the accuracy of the model was proved. There are two disadvantages on the original magnetic circuit system, which were that the average gas magnetic induction was too small and the permanent magnet working

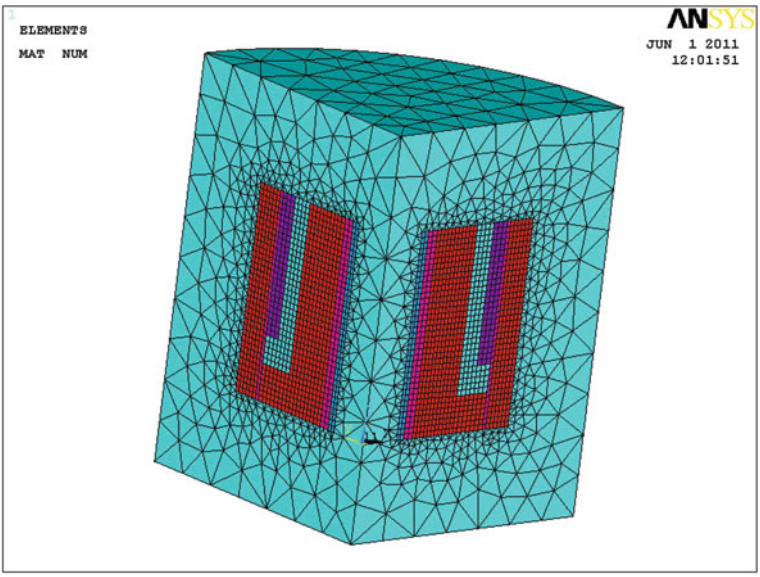


Fig. 223.3 Mesh model (including air model)

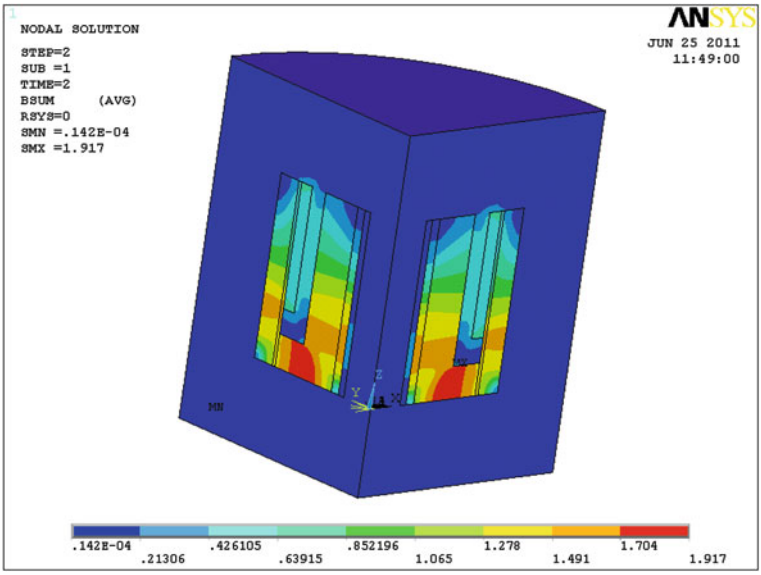


Fig. 223.4 Magnetic induction distribution of original magnetic circuit system

point was 0.5409T differing from the best operating point 0.6540 T. The reasons shown as follows could be obtained from Fig. 223.4. (1) The magnetic flux area of bottom yoke was too small and magnetic induction at its most regional was as more as 1.917 T, which caused a serious magnetic saturation. (2) The serious magnetic saturation also happened at the joint regional between inner yoke and bottom yoke. (3) The magnetic induction of outer yoke was only 1.587 T, that is to say, the utilization ratio of outer yoke was low. (4) The magnetic induction at the upper regional of inner yoke and outer yoke was too small, which not only caused a serious waste of iron core materials, but also increased the weight of the linear compressor.

223.5 Optimization Design of Magnetic Circuit System

223.5.1 Optimization Model of Magnetic Circuit System

Owing to the disadvantage of the original linear compressor, the optimization sought increases gas magnetic induction and permanent magnet working point, and avoids magnetic saturation in iron core. The zero-order optimization method was adopted, by which most engineering optimization problems can be solved.

(1) Objective function

The goal is the minimum value in ANSYS, however, the maximum gas magnetic induction B_g is the optimization seek. Consequently, a new parameter $\mathrm{OPT\_}B_g = \mathrm{C} - B_g$ has been created as the objective function by means of a constant C.

(2) State variables

The numbers of state variables must be enough according to all the constraint conditions. (a) axialdirection height H: Because of the geometrical limitation, its degree constraint was [73, 77 mm]. (b) radialdirection thickness W: Because of the geometrical limitation, its degree constraint was [65, 70 mm]. (c) the maximum magnetic induction of outer yoke $B_{o\max}$: The magnetic saturation value of silicon steel is 1.78 T. In order to increase the utilization of outer yoke and avoid the magnetic saturation, the degree constraint was [1.60 and 1.78 T]. (d) the maximum magnetic induction of inner yoke $B_{i\max}$: As same as the outer yoke, the degree constraint was [1.60 and 1.78 T]. (e) the maximum magnetic induction of bottom yoke $B_{b\max}$: As same as the outer yoke, the degree constraint was [1.60 and 1.78 T]. (f) the average magnetic induction of permanent magnet B_m: The best operating point is 0.6540 T and then the degree constraint was [0.6500 and 0.6580 T].

(3) Design variables

The structural parameters were use as the design variables. (a) axial direction height increment of bottom yoke $H_{b\mathrm{add}}$: Because of the geometrical limitation, its

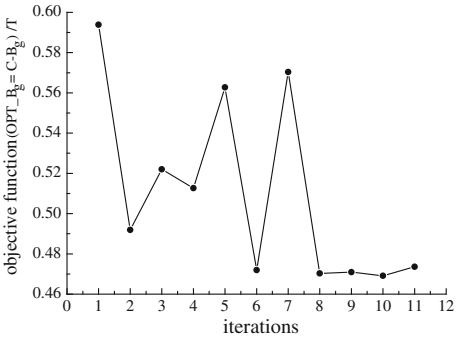


Fig. 223.5 The change curve of OPT_B_g

optimization region was [0, 9 mm]. (b) radial direction thickness increment of inner yoke $W_{i\text{add}}$: If $W_{i\text{add}}$ is more than 3 mm, magnetic induction of inner yoke will be very low. Consequently, the optimization region was [0, 3 mm]. (c) radialdirection thickness increment of outer yoke $W_{o\text{add}}$: As same as the inner yoke, the optimization region was [0, 3 mm]. (d) axialdirection height reduction of gas gap $H_{g\text{sub}}$: When $H_{g\text{sub}}$ decreases, the stitch length will be smaller. The optimization region was [0, 5 mm]. (e) the radialdirection thickness increment of permanent magnet $W_{m\text{add}}$: the magnetic potential will increase as $W_{i\text{add}}$ increases, however, in consideration of high cost, the optimization region was [0, 2.5 mm].

223.5.2 Optimization Process and Result Analysis

The iterations of optimization loop were 30 and the optimal design schematic was obtained at the tenth iteration step. The change curves of objective function $OPT\_B_g = \mathrm{C} - B_g$ and the average magnetic induction of permanent magnet B_m with iterations are shown as Figs. 223.5 and 223.6. The curves of the maximum magnetic induction of outer yoke, inner yoke and bottom yoke with iterations are shown as Fig. 223.7.

From Fig. 223.5 we can see that the optimal result of objective function $OPT\_B_g = \mathrm{C} - B_g$ is 0.46909 T. In the optimization program the constant C equals 1, as a result, the optimal value of gas magnetic induction is 0.53091 T, which increases by 27.99% relative to the one 0.4148 T of the original magnetic circuit system.

From Fig. 223.6 we can see that the optimal result of the average magnetic induction of permanent magnet is 0.65403 T, which increases of 20.92% relative to the one 0.5409 T of the original magnetic circuit system.

From Fig. 223.7 we can see that in the optimal design schematic the maximum magnetic induction of outer yoke is 1.7584 T, the maximum magnetic induction of inner yoke is 1.7784 T and the maximum magnetic induction of bottom yoke is

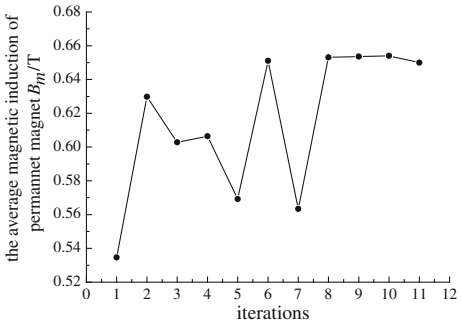


Fig. 223.6 The change curve of B_m

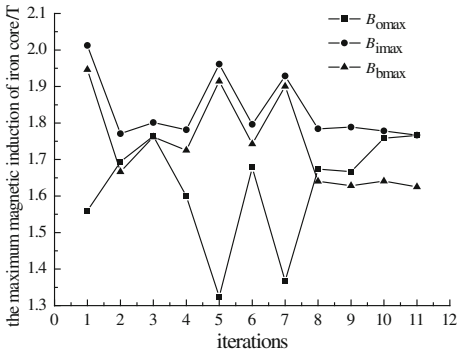


Fig. 223.7 The change curve of the maximum magnetic induction of iron core

Table 223.1 Original and final structural parameters of magnetic circuit system

Structural parameters (mm)	Original	Final
The axial direction height of bottom yoke H_b	13	22
The radial direction thickness of inner yoke W_i	18	19.5
The radial direction thickness of outer yoke W_o	10	11
The axial direction height of gas gap H_g	50	45
The radial direction thickness of the permanent magnet W_m	6	8

1.641 T, which are all lower than the magnetic saturation value of silicon 1.78 T, that is to say, the magnetic saturation of iron core in the original system has been removed thoroughly.

According to the optimal design variables, the final structural parameters of magnetic circuit system are shown in Table 223.1.

Based on the final structural parameters, the magnetic analysis was carried out. The gas magnetic induction is 0.5316 T, which increases of 28.16% relative to the one of original magnetic circuit system. The working point of permanent magnet is 0.6543 T, when it has the greatest magnetic energy. The maximum magnetic induction of iron core is 1.777 T and the magnetic saturation has been removed thoroughly.

223.6 Conclusions

(1) In order to solve the calculation problem which was caused by the large mesh quantity and convergent issue when finite element model was established according to the real magnetic circuit system structure of moving-coil linear compressor, a new modeling method was proposed.
(2) Design efficiency can be increased greatly by parametric modeling which is the foundation of the optimization analysis and improvement.
(3) The validity and reliability of the finite element model was verified through magnetic experiment. The finite element result and the experimental data only had an error of no more than 1.869%.
(4) The optimization for magnetic circuit system was carried out. Not only magnetic induction of gas gap is increased of 28.16%, but also permanent magnet is operating at the optimal point and the magnetic saturation of iron core has been removed thoroughly.

References

1. Ke Z, Tao J, Xiaoping Y (2008) Analysis and test of moving magnet linear compressor's electromagnetic characteristics. J Eng Thermophys 9:1499–1502
2. Xiaohui L, Zhihai L, Shuiying Z (2008) Dynamic characteristics of moving-magnet linear compressor based on finite element method. Electric Mach Control Appl 6:1–4
3. M-F Hsieh, K-H Hu (2008) Analysis of a tubular linear motor with soft magnetic composites for reciprocating compressors. J Appl Phys 7:112–124
4. Chen N, Tang YJ, Wu YN, Chen X, Xu L (2007) Study on static and dynamic characteristics of moving magnet linear compressors. Sci Direct 47:457–467
5. Zhang Q, Hu R, Kang S (2010) ANSYS 12.0 electromagnetics finite element analysis. China Machine Press, Beijing
6. Guoqiang L, Lingzhi Z, Jiya J (2005) Ansoft finite element analysis of engineering electromagnetic field. Publishing House of Electronics Industry, Beijing

Chapter 224
A Method of Neural Networks Controller Design for Electric Steering Actuator

Huang Zhiyi, Zhang Weiguo, Chen Kang and Su Jianqiang

Abstract To overcome the effect of nonlinear disturbance, the electric steering actuator controller design method based on neural networks is introduced. Firstly, the actuator control system is designed according to the requirement of flight control systems. Then, considering the dynamics of actuators, the nonlinear mathematical model of actuators is built, and the controller is designed with proportion-integral–differential and neural network control method. Finally, two control methods are contrasted. The simulation result demonstrates that NN controller can do it better in removing the influence of uncertainty and disturbance to achieve satisfying control effect.

Keywords Electric steering actuator · Controller · Proportion–integral–differential control (PID) · Neural network control

224.1 Introduction

Steering loop is the key element of a flight control system. It receives and calculates the control information from a flight control computer and then outputs the steering control signal according to control law's request. Steering actuators handle the control surface to deflect according to the signal offered by Steering loop to achieve the stability and control of aircraft. With the higher growth rates of power electronics technology, the electric steering actuators are used largely instead of traditional hydraulic servo in a flight control system. Electric steering takes the electric power as energy. It consists of motor, speed detector, position

H. Zhiyi (✉) · Z. Weiguo · C. Kang · S. Jianqiang
College of Automation, Northwestern Polytechnical University, Xi'an, China
e-mail: yizhi_lz@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_224,

sensor, gear transmission, and safety protection devices. As the actuator of steering loop, electric steering actuator's working drag load (hinge moment) changes along with the flying state. It has a very big effect on the steering engine's work. Traditional control method uses PID controller design which can satisfy the dynamics of system. But with the increase of flight mission and system structure complexity, the traditional control method could not satisfy the request of system. In recent years, there are a large number of documents which research the electric actuator controller design based on intelligent control technology. Yu Lu adopts two NN structures to do the actuator nonlinear compensate design [1]. Masahiro Ohka uses multilayer forward NN to design the controller for micro actuator [2], Nardenio designs the NN dynamic controller on sliding modes to do the robot electric actuator control [3]. Zhongcai Pei adopts RBF neural network PID control method to design electro-hydraulic actuator system controller [4]. Mohammed Y designs NN PI and NN PID controller to control pneumatic actuators and compares the effectiveness of the two methods [5].

From the above, neural network control technology can avoid the uncertainty and nonlinear disturbance of complex process to improve the robustness, fault tolerance, and adaptive and learning ability of controller parameters. Considering the electric steering actuator modeling error and nonlinear influence, the neural network PID controller is designed which combines the advantage of traditional PID control with the NN control. The simulation result demonstrates that the proposed method is effective.

224.2 The Electric Steering Actuator Control System

224.2.1 The Control System Configuration

In this chapterr we choose a linear displacement electric steering actuator that consists of dc motor, position sensor, rotating transformer and ball screw pair. The steering engine controller uses the control system structure in which DSP chip is the processor. The voltage changes to DSP's acceptable range through signal regulating circuit. And then A/D converting unit of DSP converts the analog signal to digital signal. At the same time, position sensor converts the current position signal of airplane control surface to analog signal, and then offers into an A/D converting unit of DSP through the signal regulating circuit. DSP gives the reference value of the speed through particular control algorithm after getting feedback position signal. Motor's feedback speed can be obtained from speed feedback signal of rotary transformer. Motor feedback current can be obtained through the current sampling circuit to A/D module of DSP. Then this current goes through the current closed-loop, combined with rotating transformer's position signal by DSP to produce the corresponding switch signals (Fig. 224.1).

According to the electric steering actuator's control characteristics, this system uses digital three closed-loop control strategy, speed, and position closed-loop

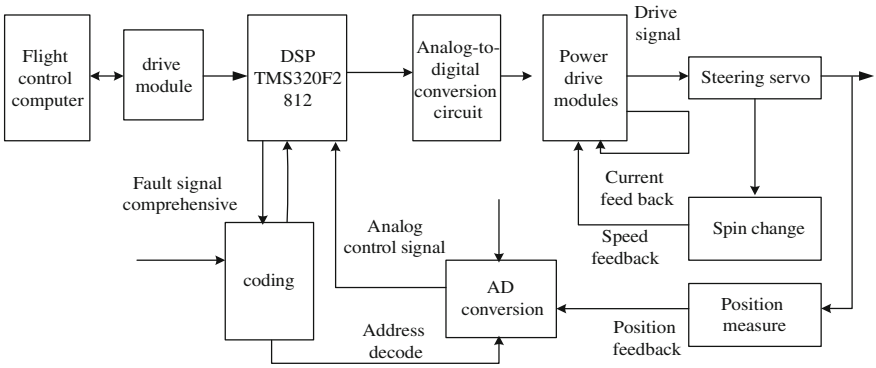


Fig. 224.1 The structure of the electric steering actuators

feedback control. Position feedback, which is the out loop of actuator loop, can ensure that position following with floating and determines the accuracy of control surface position control. Speed feedback, which is the middle loop, determines the speed of control surface's movement from the current position to the target position and plays a key role on system's rapidity and stability. Current feedback is the inner-loop of actuator and its input is speed loop's output.

224.2.2 The Mathematical Model of Electric Steering Actuator

Flight control system's actuator consists of a basic servo motor control system. Its working principle is to satisfy armature voltage balance equations, electromagnetic torque equation, back-EMF (Electromotive Force) equation, and motor torque balance equation with motor voltage. The four equations and steer surface load torque equations construct the mathematical model of an electric steering actuator.

$$\begin{cases} u = E_b + i_a R_a + L_a \dfrac{\mathrm{d}i_a}{\mathrm{d}t} \\ E_b = K_E \dfrac{\mathrm{d}\theta_m}{\mathrm{d}t} \\ T_M = K_T i_a \\ T_B = T_M - T_L = J_G \dfrac{\mathrm{d}^2\theta_m}{\mathrm{d}t^2} \end{cases} \tag{224.1}$$

Here, L_a is inductance (h); R_a is resistance (Ω); T_M is motor torque ($\mathrm{N} \cdot \mathrm{m}$); T_L is load torque converted to motor shaft ($\mathrm{N} \cdot \mathrm{m}$); T_B is acceleration torque ($\mathrm{N} \cdot \mathrm{m}$); J_G is total moment of inertia converted to motor shaft ($\mathrm{kg} \cdot \mathrm{m}^2$); K_E is back electromotive force coefficient ($\mathrm{V} \cdot \mathrm{s}$ / rad); is torque coefficient ($\mathrm{N} \cdot \mathrm{m}$ /A).

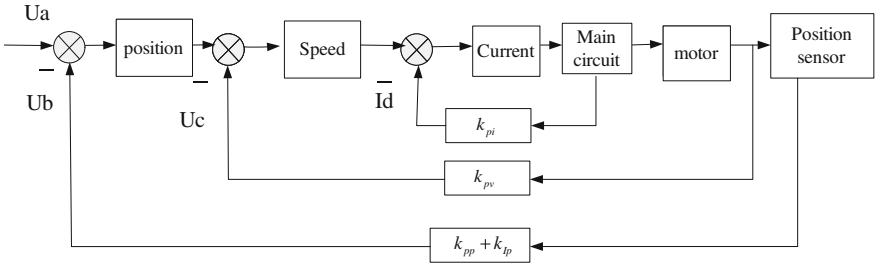


Fig. 224.2 The PID controller

Through mathematical derivation, in the no-load conditions $T_L = 0$, the transfer function of the servo motor is as follows.

$$\frac{\theta_m(s)}{u(s)} = \frac{K_T}{s(J_G L_a s^2 + J_G R_a s + K_T K_E} \tag{224.2}$$

224.3 The Electric Steering Actuator Controller

As the static error exists when motor control speed, the open loop control technology cannot remove the static error and satisfy the design request of control system, in the practical project application steering actuator control system uses closed-loop control technology which includes current, voltage, and position closed-loop feedback to achieve improvement in system dynamic and static performance.

224.3.1 PID Controller Design

Conventional PID controller has the advantages of simple and fast algorithm. When the electric steering actuator control system is designed by using PID control, the control method that is adapted in inner-loop current feedback gradually changes to out-loop speed and position feedback. As an electric steering actuator consists of a reducer, it always shows the system characteristics of big inertia, the position feedback of system adopting PI or PID control that can achieve control without static error to ensure position control's accuracy. Speed feedback, which is the middle loop, determines the speed of control surface's movement from the current position to the target position and plays a key role on system's rapidity and stability. Adopting the proportion(P) control can achieve the rapidity request and is profitable to system. Current feedback is the inner-loop of a steer loop and its input is speed loop's output. The current loop indicates the current provided to a motor when the motor runs at a pace. As its main effect is damping, it

can regulate the control parameter P to satisfy system's rapidity and stability request. The concrete control system structure is shown in Fig. 224.2.

After the structure of PID controller is setting up, the controller parameters are adopted by using the root locus method from inner-loop to out-loop until there is a satisfactory control effect.

224.3.2 NN PID Control

As long as we select controller parameters, the conventional PID control effect is very good. But when the system has strong external disturbance and uncertainty, adopting NN control will get a better control effect.

In order to achieve better system tracking, we design the PID control construction based on NN parameter adjusting. Its basic design idea is that according to system running status, PID controller parameter is adjusted in real-time to achieve optimal model tracking, and the output of NN corresponds to the three parameters k_p, k_i, and k_d of PID controller. Through network learning and adjustment, the neural network output corresponds to PID control parameters under some optimal control law.

In this paper, the single hidden layer structure BP neural network is used to adjust PID parameters. The transfer functions of hidden layer and output layer use Sigmoid function $f(x) = 1/(1 + e^{-x})$, output error of system is defined as $e(k) = \text{ref}(k) - y_{\text{out}}(k)$, $y_{\text{out}}(k)$, denotes system output, and $\text{ref}(k)$ denotes output of reference model. Network hidden layer's output is $h_j = f(\sum_i w_{ji} x_i)$, network output layer is $y_l = f(\sum_j w_{lj} h_j)$, $(l = 1, 2, 3)$ here, $y_1(k) = k_p$, $y_2(k) = k_i$, $y_3(k) = k_d$. So according to incremental PID control algorithm, the output of PID controller is

$$\begin{aligned} u(k) = {} & u(k-1) + k_p(e(k) - e(k-1)) + k_i e(k) + k_d(e(k) \\ & - 2e(k-1) + e(k-2)) \end{aligned} \tag{224.3}$$

In order to improve searching accuracy, appending the momentum which can make the searching converge to global optimum and using the momentum network learning algorithm are

$$\begin{cases} \Delta x_{k+1} = \eta \Delta x_k + \sigma(1 - \eta) g_k \\ x_{k+1} = x_k + \Delta x_{k+1} \end{cases} \tag{224.4}$$

Here, g_k is gradient of $f(x)$ at x_k, η is momentum factor ($\eta \in [0, 1]$ and), σ is learning efficiency.

The learning algorithm of output layer weight is

$$\Delta w_{lj}(k) = \eta \Delta w_{lj}(k-1) + \sigma(1 - \eta)\left(\frac{\partial E}{\partial w_{lj}}\right) \tag{224.5}$$

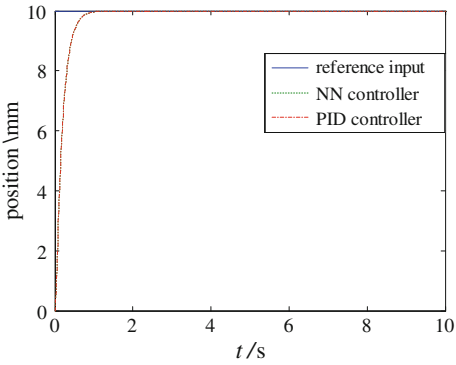


Fig. 224.3 Step response curve

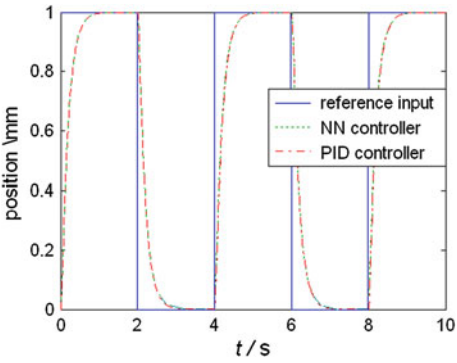


Fig. 224.4 Square response curve

According to error function, the gradient of output layer weight is

$$\frac{\partial E}{\partial w_{lj}} = -ef'\left(\sum_j w_{lj}y_j\right)h_j e_l, \quad (l = 1,2,3), \tag{224.6}$$

Here,

$e_1 = e(k) - e(k-1)$, $e_2 = e(k)$, $e_3 = e(k) - 2e(k-1) + e(k-2)$,

In the same way, the gradient of hidden layer weight is

$$\Delta w_{ji}(k) = \eta \Delta w_{ji}(k-1) + \sigma(1-\eta)f'\left(\sum_i w_{ji}x_i\right)\sum_l \frac{\partial E}{\partial w_{lj}}w_{ji}(k) \tag{224.7}$$

According to network output and controller parameter adjustment algorithm above, BP NN can be used as adaptive adjustment signal to make the output of system tracking reference model so as to achieve satisfactory dynamics.

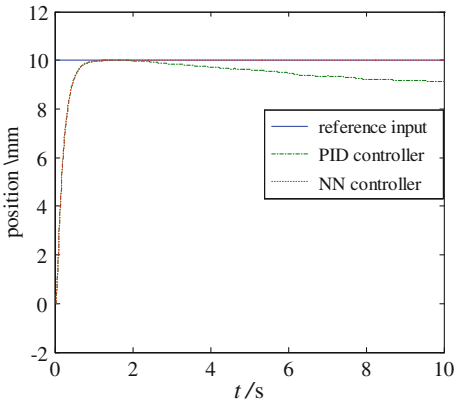


Fig. 224.5 Step response with disturbance

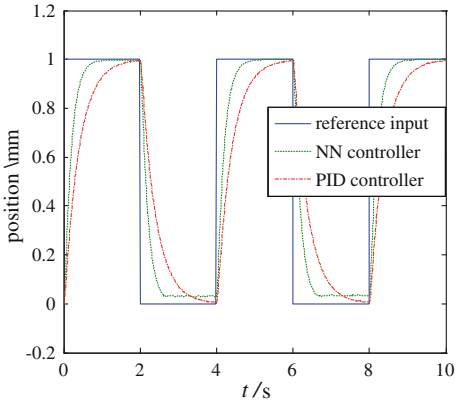


Fig. 224.6 Square response with disturbance

224.4 The Simulation Analyses

According to the proposed control method above, the simulation analyses are introduced by using the mathematical model of electric steering actuator. The simulation time is 10 s, sampling period is 0.01. The step signal and rectangular square signal are input which are 1 V amplitude and 2 Hz frequency. The simulation results are shown in Figs. 224.3 and 224.4. In order to test the NN controller effect on nonlinear and external disturbance, disturbance signal and add nonlinear factors are input such as transmission device position gap, speed saturation and position saturation. The simulation results are shown in Figs. 224.5 and 224.6. In the figure above, full line is reference input, dotted line is NN control output, and chain dotted line is traditional PID control output.

We can see from the simulation results that NN and PID controller have the same control effect, with even PID control effect being better when tracking rectangular square signal. But when the nonlinear and external disturbance are added, NN controller shows yet a good performance. PID controller deviates from

reference signal under order step response. Traditional PID control result signal has obvious time lag under rectangular square wave signal and its regulation time extends obviously.

224.5 Conclusion

In this paper, an electric steering actuator controller of flight control system is designed. According to characteristic, the structure of control system is set up. The mathematical model is built that is based on the dynamic characteristics of electric steering actuator systems. Then the controller is designed by using NN and traditional PID control method, and the two controllers are analyzed. The simulation result demonstrates that NN controller has advantages in handling control problems under nonlinear and uncertain situation.

References

1. Lu Y, Liu JK (2006) Actuator nonlinearities compensation using RBF neural networks in robot control system. In: IMACS multiconference on "computational engineering in systems applications"(CESA), Oct 4–6
2. Ohka M, Sawamoto Y, Matsukawa S, iyaoka T, Mitsuya Y (2007) Parallel type two-axial actuator controlled by a multi-layered neural network. In: International symposium on micro-nanomechatronics and human science
3. Martins NA, Bertol DW (2008) Neural dynamic control of a nonholonomic mobile robot incorporating the actuator dynamics. In: Iinternational conference on computational intelligence for modelling control & automation
4. Zhongcai P, Yanfang Z, Zhiyong T (2007) Model reference adaptive PID control of hydraulic parallel robot based on RBF neural network. In: Proceedings of the IEEE international conference on robotics and biomimetics
5. Hassan MY (2010) Comparison between neural network based PI and PID controllers. In: 7th international multi-conference on systems, signals and devices

Chapter 225
Research on Rotor Resistance Estimation for Induction Machines Based on Reactive Power Reference Model

Wenxiang Wei, Guorong Liu and Guanghui Zhu

Abstract According to the problem of reducing the robust control for indirect-field-oriented control of induction machines with uncertain rotor resistance, an estimation method of rotor resistance based on reactive power reference model is addressed. The estimation procedure is carried out independently of the state resistance, load torque, and rotor speed changes. The paper analyzes the stability of the estimation method last. Finally, the validity and practicability are verified by the simulation results.

Keywords Induction machine · Indirect-field-oriented control · Reactive power reference model · Rotor resistance estimation

W. Wei (✉) · G. Liu
College of Electrical and Information Engineering,
Hunan University, Changsha 410082, Hunan, China
e-mail: weiwenxiang@sohu.com

G. Liu
e-mail: lgr@mail1.hnie.edu.cn

W. Wei
Ministry of Education, Engineering Research Center of Advanced
Mining Equipment, Hunan University of Science and Technology,
Xiangtan 411201, Hunan, China

G. Zhu
Xiangtan Electric Manufacturing Group Co Ltd,
Xiangtan 411104, China
e-mail: zhuguanghui863@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_225,

225.1 Introduction

The vector control based on field-oriented control has traditionally been used in high-performance drive of the induction motor (IM) for variable-frequency speed regulation applications. IFOC is a very simple solution worth considering in many applications. Unfortunately, the calculation of the angular position of the rotor-flux vector is an important factor to get high-performance drive in IFOC, and which is sensitive to errors in rotor resistance. The rotor resistance of IM undergoes unknown time-varying with temperature, saturation, and skin effect under variations running period and states, and which may vary up to 100% of its nominal value to rotor windings hearting and this variation can be hardly recovered using a thermal model combined with temperature sensors [1]. Moreover, in all electric drivers, the load torque is typically unknown and flux observer used in the currently IM control rely on a good knowledge of the rotor resistance. Therefore, the IFOC control algorithms rely on a good knowledge of the rotor resistance of IM.

In recent years, a great effort has been made in the development of rotor-resistance estimation methods in the literature [2, 3] those methods assumed that rotor resistance of IM is constant and is available or properly estimated. The later research improved the former problem In [4, 5, 6] unfortunately, a new problem has arisen in the following researches.

In this article, a RPRM is proposed for IM rotor resistance which is derived from [7]. Measurements of the stator currents, rotor speed, calculations of the stator reference voltages, and rotor flux position to estimate the rotor resistance were used in the estimation method to estimate the rotor resistance; the proposed method is derived using a new formulation of the parameter adaptation law and using in an IFOC scheme.

225.2 Induction Machine Model and Reactive Power Reference Model

Assuming that linear magnetic circuits, parasitic effects such as hysteresis, eddy currents, and others of IM are generally neglected. The state-variable model of IM that referred to an arbitrarily rotating reference frame to the rotor can be obtained from [8] and the angular speed of the reference frame ω_1 can be obtained as.

$$\omega_1 = \omega_r + \frac{L_m R_r}{L_r \psi_{\mathrm{rd}}} i_{\mathrm{sq}}. \tag{225.1}$$

Assuming that the reference frame can be chosen so that ψ_{rq} is always equal to zero in [8], the d and q components can be called T and M components. This is the so-called field-oriented rotating reference frame based on rotor. The new state-variable model in M–T reference frame can be written as

$$\begin{cases} \dfrac{di_{\mathrm{sm}}}{\mathrm{d}t} = -K_1 i_{\mathrm{sm}} + \omega_1 i_{\mathrm{st}} + K_2 \psi_r + \dfrac{1}{\sigma L_s} u_{\mathrm{sm}} \\ \dfrac{\mathrm{d}i_{\mathrm{st}}}{\mathrm{d}t} = -K_1 i_{\mathrm{st}} - \omega_1 i_{\mathrm{sm}} - K_4 \omega_r \psi_r + \dfrac{1}{\sigma L_s} u_{\mathrm{st}}, \\ \dfrac{\mathrm{d}\psi_r}{\mathrm{d}t} = -\dfrac{R_r}{L_r}\psi_r + \dfrac{L_m R_r}{L_r} i_{\mathrm{sm}} \\ \dfrac{\mathrm{d}\omega_r}{\mathrm{d}t} = K_3 \psi_r i_{\mathrm{st}} - \dfrac{n_p}{J} T_L \end{cases} \tag{225.2}$$

To simplify notations we use the reparametrization:

$$K_1 = \frac{R_s L_r^2 + R_r L_m^2}{\sigma L_s L_r^2},\ K_2 = \frac{L_m R_r}{\sigma L_s L_r^2},\ K_3 = \frac{n_p^2 L_m}{J L_r},\ K_4 = \frac{L_m}{\sigma L_s L_r},$$

where R_s and R_r are, respectively, the stator resistance and rotor resistance, L_s and L_r are, respectively, the stator and rotor self-inductance, L_m is the mutual inductance, i_{sm}, i_{st}, u_{sm}, and u_{st} are, respectively, the M and T components of the stator current and voltage. The angular speed of the reference frame is ω_1, while ω_r is the rotor speed in electric rad/s. ψ_r is the rotor flux, σ is the total leakage factor ($\sigma = 1 - L_m^2/L_s L_r$), p is the number of pole pairs, and J is the total motor and load inertia. Then, ω_1 can be rewritten in M–T reference frame as

$$\omega_1 = \omega_r + \frac{L_m R_r}{L_r \psi_{\mathrm{rm}}} i_{\mathrm{st}}. \tag{225.3}$$

225.2.1 Reactive Power Reference Model Based on IFOC

Assuming that the IM speed-control system with IFOC is current controlled and rotor flux vector modulus is kept constant, ω_1 can be estimated from (225.2) as

$$\hat{\omega}_1 = \frac{i_{\mathrm{sm}}^* u_{\mathrm{st}} - i_{\mathrm{st}}^* u_{\mathrm{sm}}}{L_s\left(i_{\mathrm{sm}}^{*2} + \sigma i_{\mathrm{st}}^{*2}\right)} = \frac{Q}{L_s\left(i_{\mathrm{sm}}^{*2} + \sigma i_{\mathrm{st}}^{*2}\right)} \tag{225.4}$$

Note that Q in (225.4) is reactive power of IM [7], the i_{s} and σ in (225.8) do not change with temperature [9], and do not contain R_r which is affected by the temperature. Consequently, Eq. (225.4) can observe ω_1 accuracy. By using (225.1) and taking the estimation of rotor resistance $\hat{R}_r$ to replace R_r, the following results can be obtained from the IM speed-control system with IFOC.

$$\omega_1 = \omega + \frac{\hat{R}_r}{L_r}\frac{i_{\mathrm{st}}}{i_{\mathrm{sm}}}, \tag{225.5}$$

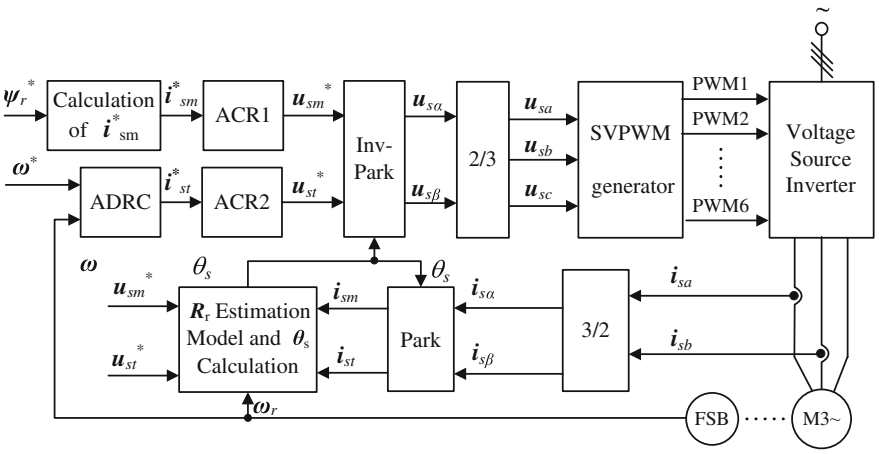


Fig. 225.1 Indirect field-oriented control scheme

$$\theta_1 = \int \omega_1 \mathrm{d}t. \tag{225.6}$$

Perfect rotor flux vector oriented will be achieved as the $\hat{R}_r$ is equal to R_r. Thus we can select (225.4) as reference mode and (225.5) as target mode, and the integral of error can be used for the R_r estimation. This is the so called RPRM based on IFOC.

225.2.2 Rotor Speed Control of IM Based on IFOC

IFOC of an IM with Active Disturbances Rejection Controller (ADRC) is depicted in Fig. 225.1. The rotor speed, ω_r, is controlled by ADRC, which is changed by load torque and IM parameters. ADRC has perfect control with large-scale external disturbance and is not sensitive to IM parameters [10].

225.3 The Rotor Resistance Estimation Method Based on RPRM

The algorithms to calculate θs and the rotor resistance estimation are depicted in Fig. 225.2. The RPRM described in (225.4) is used to estimate the rotor speed ω_r, which is compared to the rotor speed of the reference frame described in (225.5) imposed by the voltage source to form the error e. This error $e = \hat{\omega}_1 - \omega_1$ is used to calculate the estimated rotor resistance $\hat{R}_r$ by means of the estimation

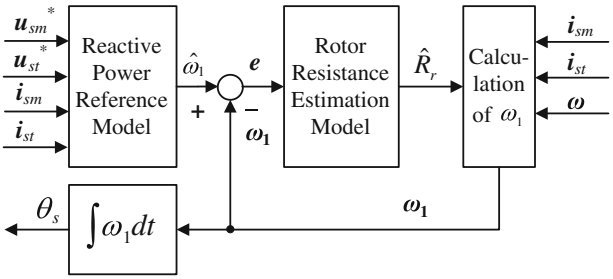


Fig. 225.2 Rotor resistance estimation model based on reactive power reference model

algorithm which is taken into (225.5) for updating, and then rotor resistance estimation model with closed feedback mode is presented. The calculation of θs depicted in (225.6) will achieve perfect field orientation when the $\hat{R}_r$ is equal to R_r [7].

The stator reference voltage u_{sm} and u_{st} can be obtained, in steady state, by calculating them from (225.2) and talking them to (225.4). Finally, the rotor resistance $\hat{R}_r$ estimation algorithm can be obtained, one can write

$$\hat{\omega}_1 = \omega_1 + \frac{L_m^2 i_{sm}^* i_{st}^*}{L_s L_r^2 (i_{sm}^{*2} + \sigma i_{st}^{*2})}(\hat{R}_r - R_r), \tag{225.7}$$

Assuming that

$$\alpha = \frac{L_m^2 i_{sm}^* i_{st}^*}{L_s L_r^2 (i_{sm}^{*2} + \sigma i_{st}^{*2})} = \frac{(1-\sigma) i_{sm}^* i_{st}^*}{L_r (i_{sm}^{*2} + \sigma i_{st}^{*2})}, \tag{225.8}$$

The estimated angular speed can be rewritten as $\hat{\omega}_1 = \omega_1 + \alpha(\hat{R}_r - R_r)$, assuming that

$$e = \hat{\omega}_1 - \omega_1 = \alpha(\hat{R}_r - R_r). \tag{225.9}$$

According to the Popov hyperstability theory, adaptive of the rotor resistance convergence rate can be obtained [1]:

$$\frac{\mathrm{d}\hat{R}_r}{\mathrm{d}t} = \left(K_P + \frac{K_i}{p}\right) e = \left(K_P + \frac{K_i}{p}\right)(\hat{\omega}_1 - \omega_1). \tag{225.10}$$

The rotor resistance convergence rate in (225.10) is regulated by the error described in (225.9), and PI regulation rules are used in (225.10), where K_p is proportion factor, K_i is integration factor, and p is differential operator. The T components of stator reference current i_{st}^* in α described in (225.8) contains the stator resistance R_s which decreases the calculation of i_{st}^*, therefore, the difficulty is encountered in selecting K_p and K_i.

Table 225.1 Simulation parameters of induction machine

Rated output power	22 KW	Stator resistance	0.284 Ω
Rated voltage	380 V	Rotor resistance	0.258 Ω
Rated frequency	50 Hz	Stator self-induction	0.228 H
Rated speed	1450 rpm	Rotor self-induction	0.234 H
Pole pair	2	Mutual induction	0.220 H
Rate flux	1.0 Wb	Load inertia	0.268 kg.m^2

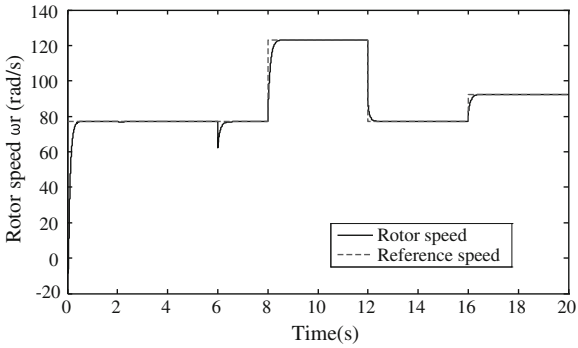


Fig. 225.3 Rotor speed simulation result with variable load torque

In this paper, an adaptive method in selecting proportion factor K_{adj} is presented, the integration factor in (225.10) is eliminated in this method, and accuracy rotor resistance estimation model will be achieved by the method but not depend on the change of R_s. Finally, the new novel rotor resistance update rules with adaption can be written as

$$\frac{\mathrm{d}\hat{R}_r}{\mathrm{d}t} = K_{\text{adj}} e = \gamma \frac{(1-\sigma) i_{\text{sm}}^*}{L_r} (\hat{R}_r - R_r), \tag{225.11}$$

where $K_{\text{adj}} = (i_{\text{sm}}^{*2} + \sigma i_{\text{st}}^{*2})/i_{\text{st}}^*$ is adaptive proportion factor and γ is a positive constant that affects the convergence rate. L_r and ω_1 is not dependent on the $\hat{R}_r$ with IM control based on constant rotor flux vector modulus and i_{sm}^* is constant corresponding the control model [11].

225.4 Simulation Results

A simulation (in [12]) was carried out to verify the rotor resistance estimation by implementing the control scheme depicted in Fig. 225.1. The parameters of IM used in the simulation are listed in Table 225.1.

Figures 225.3, 225.4 and 225.5 show the simulation results with variable load torque. The response of the rotor speed, ω_r, is shown in Fig. 225.3, the rotor speed

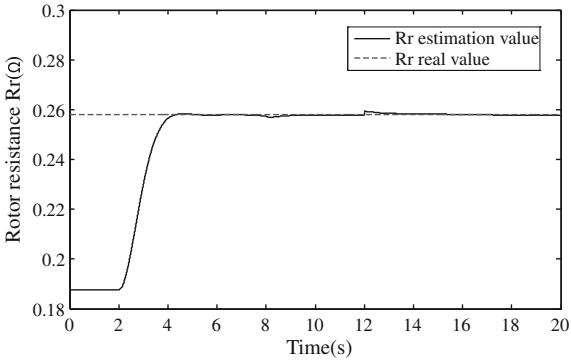


Fig. 225.4 Rotor resistance estimation result with variable rotor speed and load torque

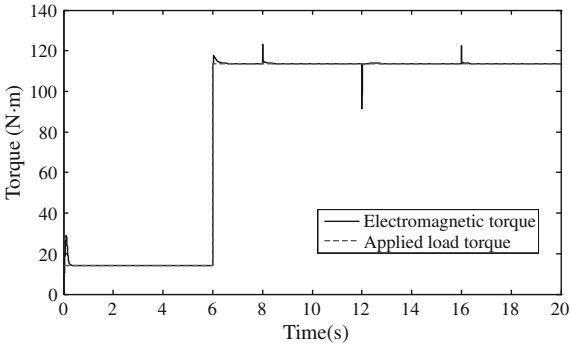


Fig. 225.5 Load torque simulation result with rotor resistance estimation

tracks its reference value in steady state closely and with no overshoot in dynamic state. Fig. 225.4 shows the rotor resistance estimation process, the algorithm estimates the rotor resistance; changes in the rotor speed reference or in the load torque produce small transients in the estimation response. The applied load torque and the calculations of the electromagnetic torque are plotted in Fig. 225.5, which showing that, when the estimation of rotor resistance is accuracy, the electromagnetic torque equal to the load torque in steady state and convergent quickly.

225.5 Conclusion

This paper has presented a method to estimate the rotor resistance for IM based on the RPRM. A standard IFOC with ADRC speed controller and PI current controller have been used for the IM controlling. The estimation model independently of the state resistance, load torque and a rotor speed changes. The simulation results showing that the estimation model is accurate and improves the restriction ability over the parameter disturbance. However, the rotor speed is required in the estimation method, which is not achieved in the speed sensorless control of IM.

Acknowledgments This work has been Project Supported by National Natural Science Foundation of China (51177040); National Research Program of China (2006BAB11B03).

References

1. Du SF, Shen YX, Ji ZC (2008) Research on induction motor rotor resistance identification based on FMRAS. Small Special Electr Mach 10:51–54
2. Castaldi P, Geri W, Montanari M, Tilli A (2005) A new adaptive approach for on-line parameter and state estimation of induction motors. Control Eng Pract 13(1):81–94
3. Marino R, Peresada S, Tomei P (2000) On-line stator and rotor resistance estimation for induction motors. IEEE Trans Control Syst Technol 8(3):570–579
4. Wang AY, Ling ZH (2009) Auto-tuning rotor resistance identification of induction motor considering iron loss. 2009 3rd international symposium on intelligent information technology application, pp 586–589
5. Minh TC, Hoang LH (2008) Rotor resistance estimation using fuzzy logic for high performance induction motor drives. 2008 3rd international symposium on intelligent information technology application, pp 303–308
6. Naft Sard MS, Benbouzid MEH (1999) Induction motors direct field oriented control with robust on-line tuning of rotor resistance. IEEE Trans Energy Conv 14(4):1038–1042
7. Chen BS (2003) Electrical towage automatic control system. Mechanical Industry Press, Beijing, pp 200–214
8. Pedro RS, Aurelio GC, Vicente FB (2007) Rotor-resistance estimation for induction machines with indirect-field orientation. Control Eng Pract 15:1119–1133
9. Rehman HU, Derdiyok A, Guven MK, Xu L (2001) A MRAS scheme for on-line rotor resistance adaptation of an induction machine. IEEE Power Electr Special Conf 2(4):817–822
10. Wei WX, Liu GR (2010) Control and simulation analysis of the electric wheel dump truck based on active disturbances rejection controller. J Hunan Univ (Nat Sci) 37(10):47–49
11. Jin H, Huang J (2006) Adaptive flux estimation and parameters identification of induction motors based on model reference approach. Trans China Electrotech Soc 21(1):65–69
12. Xue DY, Chen YQ (2002) System simulation technology and application based on MATLAB/Simulink. Tsinghua University Press, Beijing, p 350

Chapter 226
Simulation of Double-Loop DC Motor Speed-Regulation System of Fuzzy Control

Sumin Han and Quanyou Wang

Abstract To solve the weak robustness and strong oscillation of current nearby the equilibrium point in the double-loop DC drive system, speed loop (ASR) is introduced to fuzzy control. First, according to the ranges of input and output variables, fuzzy controller was designed and fuzzy rules were generated. Fuzzy control simulation model of double-loop speed regulation system was established and simulation was carried on by matlab. Compared with analysis waveform of the traditional system, it is concluded that apart from achieving rapid and stable starting current, small overshoot, small downhill of steady-state, the fuzzy speed-regulation system can improve the current oscillation.

Keywords Double-loop · Fuzzy control · DC · Speed-regulation · PI

226.1 Introduction

DC motor is widely used in rolling and electric locomotives for the advantages of on starting torque, wide speed range. As a major way of the DC motor speed-regulation, speed and current double closed-loop system has many properties as well stability, high-speed performance, etc. Being relatively mature both in theory and practice, it plays a more important role in the field of drag. But the traditional

S. Han (✉) · Q. Wang
Institute of Electrical Engineering, Henan Polytechnic University,
Century Avenue.2001, Jiaozuo 454000, China
e-mail: hansumin@hpu.edu.cn

Q. Wang
e-mail: wqy119@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_226,

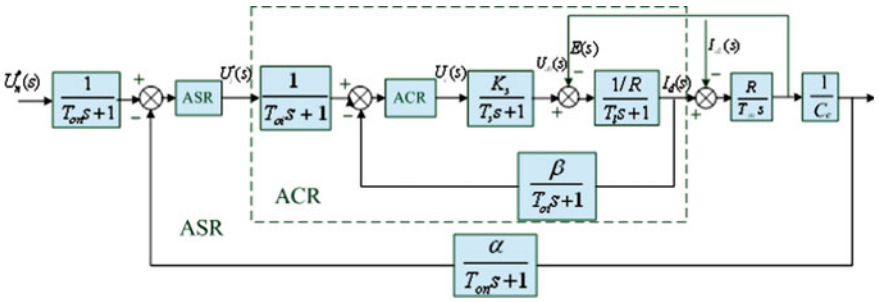


Fig. 226. 1 Double-loop speed control system block diagram of the dynamic

double-loop system is hard to balance rapidity and stability, to achieve a well control performance for the change of control object, especially the current oscillation is more severe nearby the equilibrium point. It is possible to realize precise control of the change of control object, fast response, robustness and other advantages when fuzzy control is applied to the speed-regulation system [1, 2].

226.2 Traditional Double-Loop Speed-Regulation System

Figure 226.1 is dynamic block diagram of double-loop speed-regulation system. The figure in the dashed box is the current regulator (ACR), and the larger ring is speed regulator (ASR). In this graph, T_{oi}, T_{on} are respectively filter time constant of ASR, ACR in the filter inertia presented by first-order inertia. To balance this filter delay effect, the given filter link is increased which was represented by the first-order inertial link [3].

The main circuit of DC double loop speed-regulation system consist of three phase AC voltage source, thyristor bridge rectifier, DC motor, speed loop, current loop and other components in Fig. 226.2. In the simulation system diagram is the output of ACR control trigger device of thyristor rectifier. Setting the limitary output of ASR and adjusting the gain of the speed feedback channel, we can gain the best way to start, to brake, to restrain the disturbance from power and load fluctuations.

Simulation parameters are set as follows. DC : $U_N = 220$ V, $I_N = 136$ A, $n_N = 1460$ r/ min, $C_e = 0.132\ V \cdot \text{min}/\text{r}$; Allowing overload multiple 1.5; flywheel inertia:$GD^2 = 22.5\ \text{N} \cdot \text{m}^2$; amplification factor of the thyristor device: $T_l = 0.03\ s\ K_s = 40$; Total resistance of the armature circuit: $R = 0.5\ \Omega$; Time constant: $T_l = 0.03$ s, $T_m = 0.18$ s; current feedback factor: $\beta = 0.05$ V/A (≈ 10 V/1.5^{I_d}); speed feedback factor: $\alpha = 0.007\ \text{V} \cdot \text{min}/\text{r}$ ($\approx 10\ \text{V}/^{n_{\max}}$).

Both ASR and ACR use no static errors PI controller, output limit: [−10, 10], rated load torque: $T_L = 9.55\ C_e I_N \approx 171.4$ Nm. Simulation is divided into two

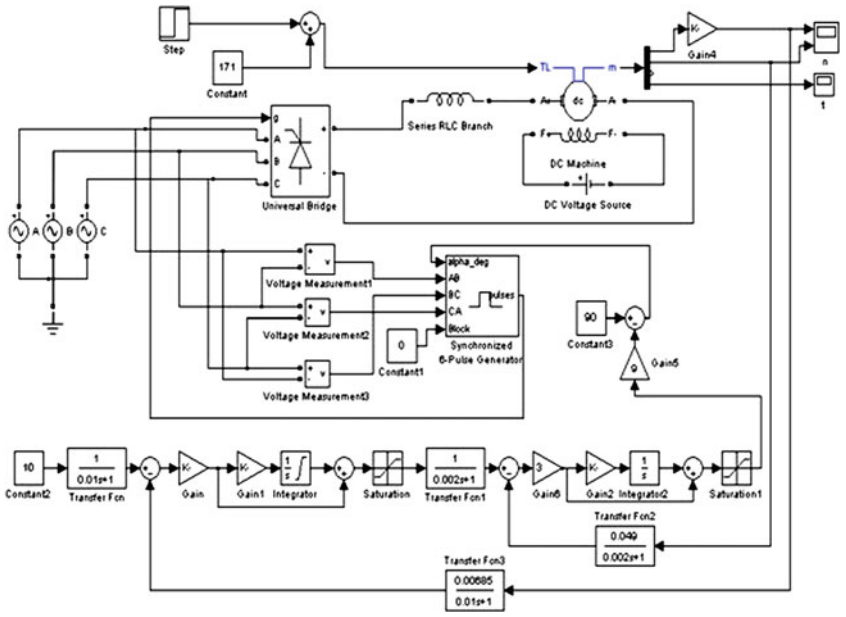


Fig. 226.2 Double-loop speed control system simulation diagram

stages: (1) given the sudden start; (2) add disturbance which has 30% of the rated torque when 1.5 s.

Figure 226.3 shows the simulation results. From the speed and armature current curve, it is apparent that the motor start at high constant current and torque, also has the advantages of no static error and well steady speed performance. When disturbance was added after 1.5 s, the speed decreased slightly and then return to the rated speed rapidly, while the current curve reach steady state value finally after small oscillation. To sum up the above arguments, we can see double-loop DC speed control system has good static and dynamic characteristics, and anti-disturbance performance. But we can see also the poor dynamic nature of current curve. When disturbance signals was added suddenly and the speed reaches the maximum, the current oscillation is obvious.

Especially on the moment of given value added suddenly current exceeds the overload value of 204–230 A or so. Thus introducing fuzzy control to ASR can improve the current oscillation at equilibrium point.

226.3 Double-Loop Fuzzy Control System Design

The biggest advantage of fuzzy control is precise which does not depend on charged object, be able to overcome the non-linear and other factors. By membership function inaccurately described language can be turned into fuzzy language. After fuzzy rule base are established, fuzzy inference can be done

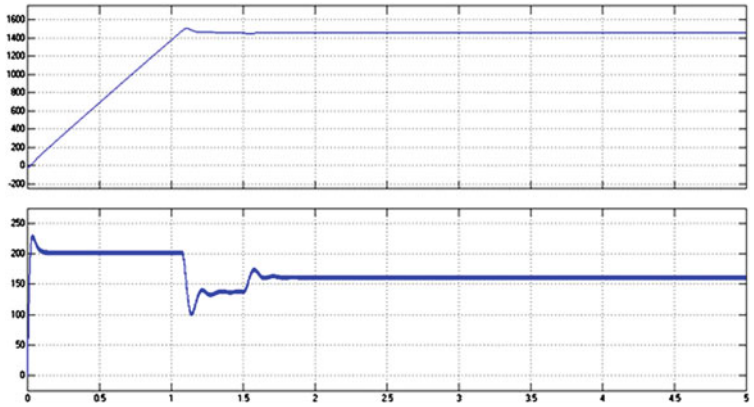


Fig. 226.3 Current speed of graphs double-loop

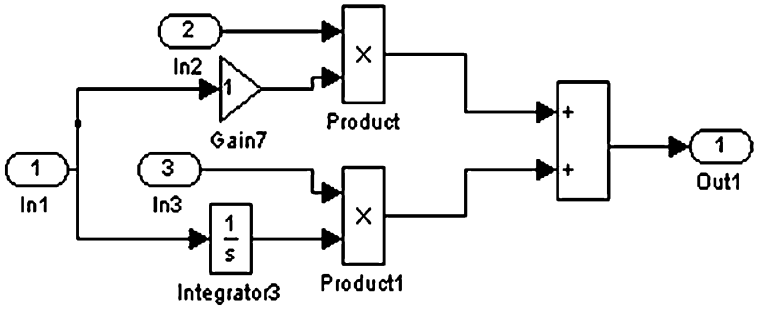


Fig. 226.4 PI subsystem model

according to the state of input variables, fuzzy control may be realized. Base on the above analysis, introduced in speed loop, fuzzy controller take speed deviation E and the derivative of its EC as input variables, coefficients of fuzzy PI control as output variables: KP, KI, and perform operations according to PI subsystem showed in Fig. 226.4 [1, 2, 4, 5].

Fuzzy controller was established by matlab, the domains of E, EC were set: [−1, 1], [−5, 5], of KP, KI: [5, 30], [5, 30]. Membership function is uneven trigonometric curve. Fuzzy subset of E, EC is {NB, NM, NO, PO, PM, PB}, fuzzy subset of KP, KI is {NB, NM, NS, ZO, PS, PM, PB}. 36 fuzzy rules were generated as shown Table 226.1, and the results were saved to be prepared for the model simulation.

Fuzzy controller model (Fig. 226.5) instead of the ASR in Fig. 226.2 as shown in the blue box. The input and output of the proportional and integral coefficient can be further setted in Fig. 226.4. The simulation results show that speed curve is still rising rapidly and smoothly in Fig. 226.6, it rapidly decline to given value when 5% overshoot appearing, and it adjusted quickly to the given value after

Table 226.1 Fuzzy rule

E	EC					
	NB	NM	NO	PO	PS	PM
NB	PB	PB	PM	PS	ZO	ZO
NM	PB	PB	PM	PS	ZO	NS
NO	PM	PM	PM	ZO	NS	NS
PO	PS	PS	ZO	NS	NM	NM
PS	PS	ZO	NS	NM	NM	NB
PB	ZO	ZO	NM	NM	NB	NB

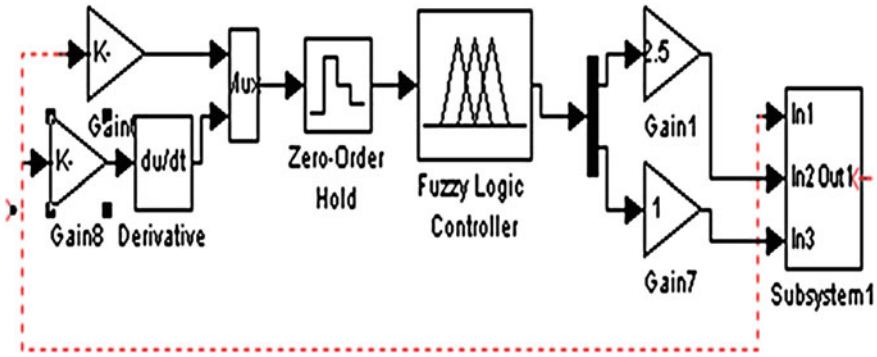


Fig. 226.5 The simulation model of fuzzy controller

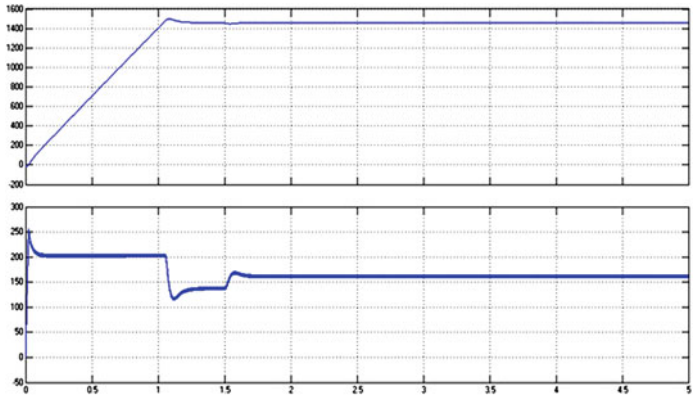


Fig. 226.6 Fuzzy control double-loop speed and current curve

disturbance torque was added at 1.5 s. While dynamic process of the current curve is significantly improved, and start current below the overload value compared with Fig. 226.3.

226.4 Conclusion

Compared to traditional double-loop speed-regulation system, fuzzy control system is able to achieve the characteristics of rapidly stable start, small overshoot, small steady downhill. Importantly, it does not depend on the nature of the load, does not establish a precise mathematical model, is not impacted from uncertainties such as parameter drift, and realize fuzzy human control, and improves the robustness of the system. In addition, it can significantly improve the dynamic process of equilibrium point. The article introduces fuzzy control applied to specific objects, which will play a role in its actual application.

Acknowledgments The paper was sponsored by Science Technology Research Project of He'nan Educational Committee (2011A470004).

References

1. Ou W-B (Mar 2010) Simulation of DC double closed-loop speed-regulation system and its fuzzy control. Ind Mine Autom 3:57–60
2. Dong Z, Zhu Z (2010) Application research of a dual closed loop controller of DC motor based on fuzzy PID control. Electr Autom 32(5):14–16
3. Ruan Y, Chen W (2006) Motion control systems. Tsinghua University Press, Beijing, pp 54–105
4. Du S-F et al (2005) Intelligent control theory and applications. China Agricultural University Press, Beijing, pp 70–80
5. Zhou X, Wang Q et al (2010) Fuzzy contol simulation for DC motor based on MATLAB/SIMULINKI. Motor Technol 1:8–10

Chapter 227
Study on Direct Force Control for PMLSM Based on SVPWM and the Flux Dynamic Compensation

Haixing Wang, Jikai Si and Haichao Feng

Abstract Focused on directing force control (DFC) of permanent magnet linear synchronous motor (PMLSM) existing thrust ripple, the space voltage vector pulse width modulation technology (SVPWM) will be introduced into PMLSM DFC strategy in order to reduce thrust ripple and improve system control performance. Full consideration of the motor running process with the rise of temperature, the change of the primary winding resistance caused the error between the flux observed value and the actual value. A new flux dynamic compensation algorithm was proposed. The dynamic compensator was built to improve the observation accuracy of the primary flux, which solved effectively the wrong voltage vector choice and the system control performance failure due to the error of the primary flux. The experimental results show that the PMLSM DFC based on SVPWM and the flux dynamic compensation has better control performance.

Keywords SVPWM · Dynamic compensation · PMLSM · Direct force control

H. Wang (✉) · J. Si · H. Feng
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454003, China
e-mail: wanghaixing@hpu.edu.cn

J. Si
e-mail: sijikai@hpu.edu.cn

H. Feng
e-mail: fhc@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_227,

227.1 Introduction

Direct force control (DFC) has advantages such as, clear thought control without complex coordinate transformation, small requirements for the motor parameters, strongly robustness, fast response, and so on. But the thrust ripple and the control effect non-ideal are the existing problems for DFC. How to reduce the speed and thrust fluctuations are the key application problems for the PMLSM DFC. Several factors result in the thrust ripple of PMLSM DFC. The discontinuous space voltage vector switching and the observation errors of primary flux are the important [1–4].

In this chapter, we focused on the existing thrust ripple of the DFC, the SVPWM will be introduced into the PMLSM DFC. At the same time, full consideration of the motor running process with the rise of temperature, a new flux dynamic compensation algorithm is proposed. The dynamic compensator is built to improve the observation accuracy of the primary flux, which maybe the cause for wrong choice of the voltage vector and result in the system control performance deterioration, even failure. The proposed flux dynamic compensator solves effectively the wrong voltage vector choice and the system control performance failure due to the error of the primary flux, improving the control performance.

227.2 Space Voltage Vector Pulse Width Modulation

227.2.1 Principle Analysis of SVPWM

SVPWM divides each sector into a number of corresponding switching period between the cell interval, in which the size of the switching period determined by the switching frequency [5, 6]. Each cell interval is composed of six effective space vector and two zero vector, which form different linear combination as the equivalent reference vector u_{ref} to make motor flux track approaching circular. In any short period, the inverter output and the reference voltage u_{ref} output are consistent. As shown in Fig. 227.1, in a calculation period T_p, the voltage vector is assumed in the third sector, and the voltage vector u_{ref} is the linear combination of u_4 and u_6.

After the action time of each voltage vector determined, then the next is to determine the action order of each voltage vector, which has to follow the principle of switching frequency least and minimize the maximum deviation of each phase current vector. SVPWM has lots of advantages, such as effectively inhibit harmonic component, reduce the current waveform distortion, and greatly improve the utilization of DC voltage, and make the rotating magnetic field close to circular, so as to achieve higher control performance.

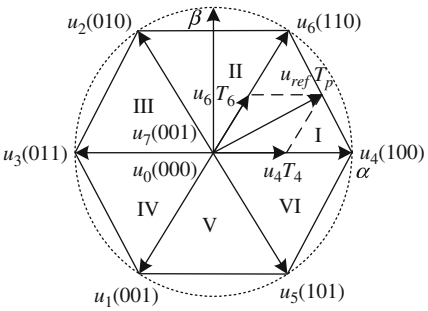


Fig. 227.1 Resolution graph of voltage vector

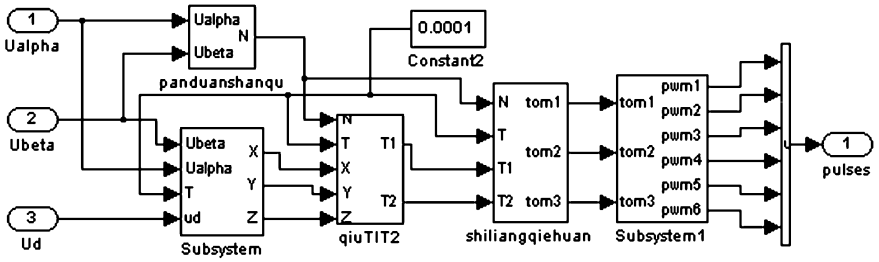


Fig. 227.2 Simulation module of SVPWM

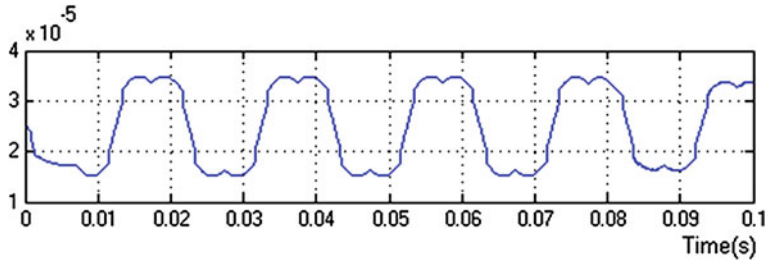


Fig. 227.3 Modulation waveform of T_{cm1}

227.2.2 SVPWM Simulation

The SVPWM simulation needs the judgment of the sector, the action time calculation, the distribution of the voltage vector, and the decision of the switching vector points, and so on. It is known that SVPWM can be realized long as the _two-dimensional static coordinate component u_α and u_β of the reference voltage vector u_{ref}, and the calculation period of the PWM T_s.

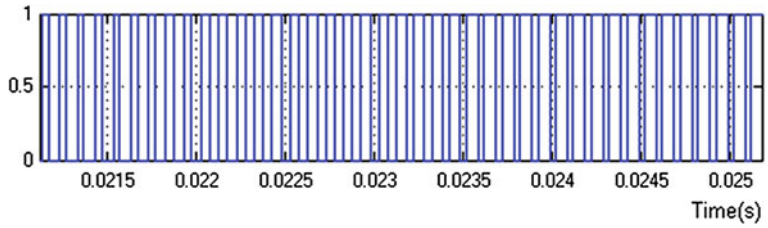


Fig. 227.4 Waveform of SVPWM

The simulation model of SVPWM shown in Fig. 227.2, mainly including the module of the judgment of the sector, the module of the X/Y/Z calculation, the module of the action time of T_1 and T_2 calculation, the module of the vector switching point, and the pulse generator module.

The simulation results are shown in Figs. 227.3 and 227.4.

Figure 227.3 is a modulation waveform of the vector switching point T_{cm1} Fig. 227.4 shows pulse waveform of SVPWM.

It can be seen from Fig. 227.3, the modulation waveform T_{cmx} ($x = 1, 2, 3$) are three-phase saddle, so it has the advantages to optimize the PWM inverter.

227.3 Design of Flux Dynamic Compensator

227.3.1 Compensator Design

If the primary resistance R_s remains constant, the observed value and actual value of the flux are consistent. In fact, the resistance changes with the motor temperature and current frequency. Any resistance variation may bring integration error and the actual value of flux does not match with the observed value, which may lead to the error choice of the voltage vector and system control performance deterioration [7, 8]. The variation of the resistance variation can be seen as the uncertainty of the system from the point of view of the control theory. For the actual uncertainty, whether it is linear, nonlinear, or no matter how complex, one point, which is self-satisfied or changing rate bounded, is basically satisfied. Based on this point, the design of a new flux dynamic compensator is designed, according to the advantages and disadvantages of the PI regulator and the adaptive observer.

The ideal value of flux is ψ_s^*, so

$$\dot{\psi}_s = u_s - (R_s + \mathrm{d}R_s)(i_s + \mathrm{d}i_s) = u_s - R_s i_s + \mu \tag{227.1}$$

or $\psi_s = \int (u_s - R_s i_s + \mu)\mathrm{d}t + \psi_{s0}$, where, $\mu = -R_s\,\mathrm{d}i_s - \mathrm{d}R_s \cdot i_s - \mathrm{d}R_s \cdot \mathrm{d}i_s$.

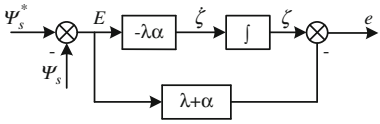


Fig. 227.5 Structure schematic diagram of flux compensator

The nominal value is defined as

$$\dot{\psi}_s^* = u_s - R_s i_s \tag{227.2}$$

In order to achieve the following equation

$$\lim_{t\to\infty}(\psi_s(t) - \psi_s^*(t)) = 0 \tag{227.3}$$

Equation (227.1) should be added the compensation *e(t)*, then the following equation is obtained

$$\dot{\psi}_s = u_s - R_s i_s + \mu + e \tag{227.4}$$

or $\psi_s = \int (u_s - R_s i_s + \mu + e)dt + \psi_{s0}$.

Equations (227.4 and 227.3) subtraction gets error equation:

$$\dot{E} = \dot{\psi}_s - \dot{\psi}_s^* = \mu + e \tag{227.5}$$

In the control system, μ is as the uncertainty of the unknown disturbance, and e can be seen as the control input.

In order to inhibit the effect of the uncertainty μ on flux observer, the following dynamic compensator is structured.

$$\begin{cases} \dot{\xi} = -\lambda\alpha E \\ e = -(\lambda+\alpha)E + \xi \end{cases} \tag{227.6}$$

From the above analysis, dynamic compensator block diagram is shown in Fig. 227.5.

227.3.2 Simulation of Compensator Flux Compensation

The simulation model is shown in Fig. 227.6.

The sine signal is used as the original e, and the [−0.3, 0.3] random signal as the interference signal. The given flux is 0.2324 Wb. The result is shown in Fig. 227.7.

In the Fig. 227.7, the observation error is within 2% after compensation, having better effect, shown that the compensator is feasible to improve the PMLSM DFC.

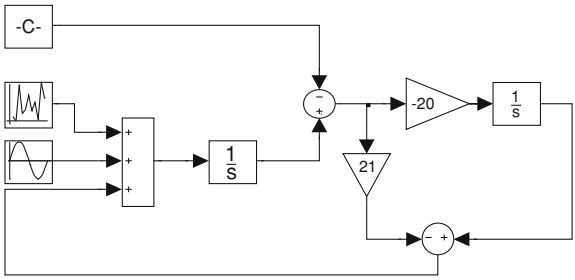


Fig. 227.6 Simulation model of flux-linkage compensator performance

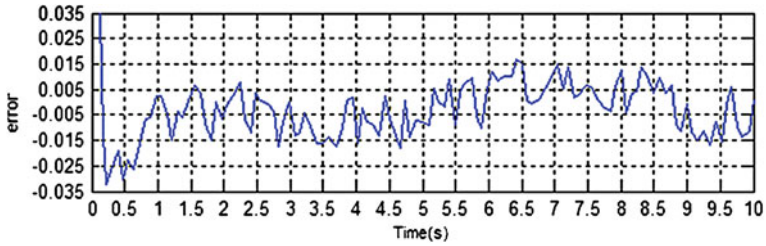


Fig. 227.7 Simulation of dynamic flux-linkage compensator

Fig. 227.8 Experimental platform of the PMLSM DFC

227.4 Experiment

227.4.1 Experimental Platform

The experimental platform is shown in Fig. 227.8.

227.4.2 Experimental Results

Figure 227.9 shows the experimental results, including the flux-linkage locus waveform, the thrust force waveform, the speed waveform, and A phase current

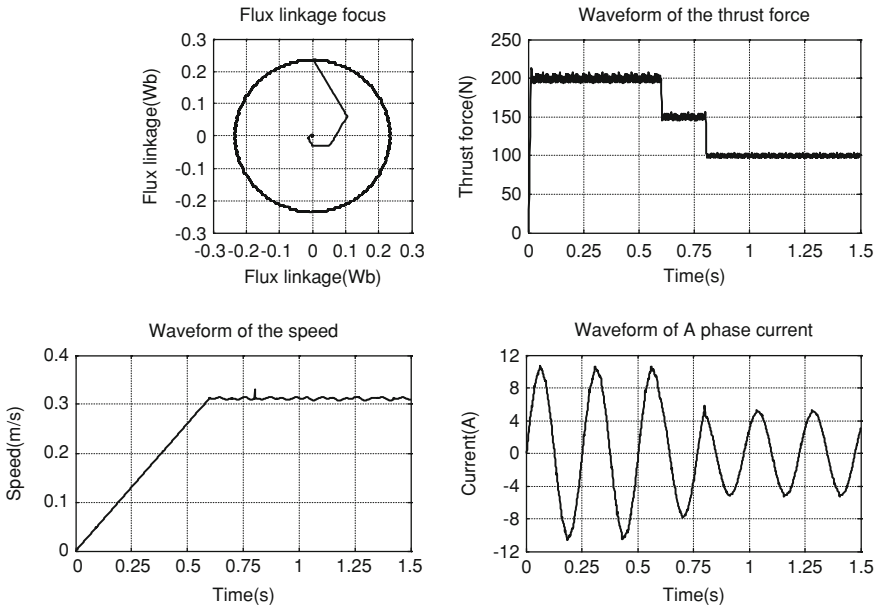


Fig. 227.9 Testing waveforms of the flux-linkage trace, thrust force, speed, and A phase current

waveform. The PMLSM starts at the load 150 N, and then reduce the load to 100 N at 0.8 s.

From the point of view on the primary flux-linkage locus, thrust variation, speed change, and current trends, the primary flux-linkage locus has less fluctuation and closer to the circular. The response of the thrust force is fast and always following the given value. In the speed response curve, the disturbance of speed is obviously smaller when the sudden load reduction. The current curve is similar to the sine. Particularly, the current distortion has been significantly improved when the motor running state changed.

The experimental results show that SVPWM DFC in the PMLSM control based on the flux dynamic compensation is feasible. The flux dynamic compensation algorithm is correctness and useful to improve the dynamic performance of DFC, significantly reducing the fluctuation of the flux and thrust force, further improving the steady-state performance of the PMLSM DFC.

227.5 Conclusion

In this paper, the reasons of the thrust force fluctuation for the conventional PMLSM DFC are analyzed. The SVPWM is introduced into PMLSM control system in order to make the inverter having the continuous voltage vector and

reduce the fluctuation of the primary flux and the thrust force. In addition, based on the deep analysis of flux observer methods, full consideration of the primary winding resistance change during the motor running process, a new flux dynamic compensation algorithm is proposed. Then the primary flux dynamic compensator is designed to improve the accuracy of the flux observation and ensure the accuracy selection of the voltage vector. The experimental results show that the PMLSM DFC based on the SVPWM and flux dynamic compensation, maintains the excellent dynamic performance, and significantly reduces the fluctuation of the flux and the thrust force, shown better control performance.

Acknowledgment This work was supported by the Foundation for University Key Teacher by Henan Province (2010GGJS-230), and Doctor Foundation of Henan Polytechnic University (2011065), and National Natural Science Foundation of China (NO. 61074095), and Ministry of Education Scientific Research Foundation for Chinese Overseas Returnees and Research Fund for the Doctoral Program of Higher Education (NO. 20104116120001).

References

1. Zhong L et al (1997) Analysis of direct torque control in permanent magnet synchronous motor drives. IEEE Trans Power Electr 12(3):528–536
2. Junyou Y, Guofeng H, Jiefan C (2007) Analysis of PMLSM direct thrust control system based on sliding mode variable structure. In: Conference proceedings-IPEMC: CES/IEEE 5th international power electronics and motion control conference, Vol 1, pp 662–666
3. Jiefan C (2006) Study on thrust force of permanent magnet linear synchronous motor and its direct thrust force control system. Shenyang University of Technology, Shenyang
4. Hua W, Gang X, Li W, Gent Y (2005) Variable structure torque control of induction motors using space vector modulation. Electr Eng 87:93–102
5. Hong N (2010) Modeling and simulation of power electronics and motor control system [m]. China Machine Press, Beijing
6. Weiping Z, Zhengguo W, Jinsong T, Daming L (2006) A novel algorithm of SVPWM and the study on the essential relationship between SVPWM and SPWM. Proc CSEE 26(2):133–137
7. Junyou Y, Jiefan C, Guofeng H (2007) Direct thrust control of PMLSM using SVM and sliding model variable structure. Trans China Electrotech Soc 22(6):24–29
8. Liming S, Ke W, Yaohua L (2008) Direct thrust control of linear induction motor based on improved flux observer. Trans China Electrotech Soc 23(9):45–50

Chapter 228
Study on Multi-Motor Power Balance of Controlled Speeder for Belt Conveyor

Xinhuan Wang, Hongwei Zhang and Wei Xiong

Abstract In this paper, the dynamic and static influencing factors of multi-motor power balance for belt convey was discussed and thus the corresponding engineering solution was put forward through analyzing the power balance course of active control for belt conveyor. The new solution which collects the current and speed signal simultaneity as the control bases to solve the power balance for belt conveyor was put forward and the correctness and feasibility of the principle was discussed and then the hardware and software design of electric control system were introduced in detail. The experiment test showed that the electric control system was able to accurately control mechanical drive and hydraulic devices and realize the functions of soft start and soft stop and the belt conveyor had characteristics such as steady power, quick response-time, extensive tape speed range and high transmission efficiency.

Keywords Belt conveyor · Controlled speeder · Electrical control system · PLC · Power balance

228.1 Introduction

With the modernized mine and coal-mining technique developing, the applications of long carrying distance and large carrying ability belt conveyor are becoming increasingly common. How to reasonably define the driving schemes for big-scale belt conveyor is becoming one of the key techniques in the popularizing and

X. Wang (✉) · H. Zhang · W. Xiong
School of Electrical Engineering and Automation,
Henan Polytechnic University, No. 2001 Century Avenue,
Jiaozuo 454000, Henan, China
e-mail: wang_xh@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_228,

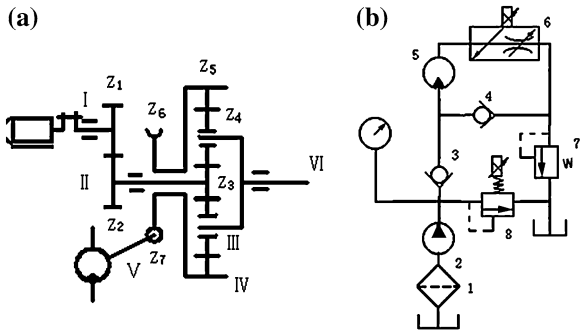


Fig. 228.1 Structure of mechanical system and hydraulic system. **a** Structure of mechanical system (Z_1 driving gear, Z_2 driving gear, Z_3 central gear, Z_3 planetary gear, Z_5 internal gear, Z_6 worm gear, Z_7 worm lever). **b** Structure of hydraulic system (1 oil-absorptive filter, 2 hydraulic pump, 3 and 4 one-way valve, 5 hydraulic motor, 6 electro-hydraulic proportional throttle valve, 7 relief valve, 8 electro-hydraulic proportional relief valve)

applying course. The influencing factors of belt conveyor driven by multiple soft start such as the difference of outer performance of every soft start motor, the error of equipment manufacturing and installation, the difference of elastic elongation of belt and the varied load etc. which often cause the unbalance of motor load power and then bring the power unbalance problems of belt conveyor, the above problems must be solved in actual application although the start and running are controlled, or the service life of belt conveyor will be reduced dramatically and the whole conveyor system will be scraped in a serious circumstance. In this paper, a new power balance solution is introduced through analyzing the power balance course of active control for belt conveyor and the new solution realizes the power balance of multi-motor or multi-point driving through controlling fuel delivery of hydraulic motor so as to control speed of output axles of mechanical system.

228.2 Whole Structure of Controlled Speeder for Belt Conveyor

Active controlled speeder for belt conveyor is a typical electromechanical and hydraulics integration product which is made up of mechanical system, hydraulic system and electric control system and the device is the organic combination of the three above systems and was distributed rationally.

228.2.1 Structure and Principle of Mechanical System

The mechanical system is made up of a gear retarder and a 2 K–H planetary speeder, as shown in Fig. 228.1a. The gear retarder plays a part in deceleration and

preliminary extending torque and the 2 K–H planetary speeder not only further extend torque and reduce speed but also realizes compound motions and overload protection.

If internal gear is first driven by hydraulic motor and the speed attains certain numerical value when belt conveyor starts up, both the output *axis* VI and input *axis* VI are possible run according to the working principles of planetary transmission. The resistance torque of the output *axis* VI is heavy as the axis is jointed together with rollers, but the resistance torque of input *axis* VI is small and thus the rotors of main motor roll with input *axis* VI Adjust the hydraulic system appropriately to ensure rotatable steering of input *axis* VI according with the working steering of main motor and then start main motor to realize the purpose of no-loading start. The hydraulic motor becomes a load of main motor when the speed of main motor reaches the rated speed and controlling hydraulic motor to reduce speed according to a certain rule, then the start speed of the output *axis* VI increases according to corresponding rule to realize the purpose of soft start which main motor starts at no-loading first, then the rollers evenly and steadily raise speed to the rated speed.

228.2.2 Composing of Hydraulic System

As the internal gear has a little load and requires continuously variable speeder etc. controlled speeder for belt conveyor adopts hydraulic system which is made up of electro-hydraulic proportional valve as power source for internal gear. The structure of hydraulic system is shown in Fig. 228.1b and the main hydraulic components include hydraulic pump, hydraulic motor, electro-hydraulic proportional throttle valve and electro-hydraulic proportional relief valve etc. Hydraulic motor is power source for worm lever of mechanical system and realizes continuously variable speeder for belt conveyor by adjusting the flow of hydraulic motor through electro-hydraulic proportional throttle valve 6. Electro-hydraulic proportional relief valve 8 is used to adjust working pressure of hydraulic system to ensure the hydraulic system always working in optimal state and thus to realize the design purpose of the highest efficiency and the lowest oil temperature.

228.3 Study on Control Strategy of Multi-Motor Power Balance

To ensure the output of every driving device evenly when the long carrying distance and large carrying ability belt conveyor driven by multi-motor, the line speed of rollers should be the same in theory. As the error of equipment machining and the difference of mechanical performance of motor, the line speed of rollers are not the same and the actual output of motor cannot keep balance. Even if there is no difference between mechanical performance of motor and the diameter,

the speed of rollers are the same, the situation which the line speed of rollers and the actual load are not the same may also occur because the actual diameter of roller are not same as the dirt are bound in running. Therefore, the power balance of multi-motor or multi-point driving is an important aspect of judging working performance of driving device for belt conveyor.

In nature, realize the power balance of multi-motor or multi-point driving through adjusting speed to achieve. The new method of power balance of multi-motor driving includes the factors of speed and current of motor into the changes of power, that is to say, collecting the current and speed signal simultaneity as the bases to control power balance. If the load is heavy, through controlling the open degrees of electro-hydraulic proportional throttle valve 6 to control fuel delivery of hydraulic motor so as to make hydraulic motor to drive internal gear turning, controlling speed of output *axles* VI of mechanical system to reduce speed of roller and the load of motor then come down along, then the power of multi-driving devices realize power balance under the action of belt. The control strategy of multi-motor power balance considers the two factors of current and speed simultaneity and combines them as the control bases of multi-motor power balance and overcomes the shortcoming of existing methods of multi-motor power balance which only consider current and the control precision is low.

The motors of the same performance parameter are generally selected when belt is initially designed. When every motor runs within rated load, the same voltage power supply and power factor is also approximately equal and we can take the load current of every motor as the control bases to keep power balance. The motor drive and generation condition can be distinguished through comparing synchronous speed of motor with detected motor speed. When the motor speed is larger than synchronous speed, the motor works in generation condition, otherwise, it works in motor state. When the motor works in motor state, the given motor current is positive and otherwise, the given motor current is negative. Electronic control system takes the average current collected simultaneity of every motor as criterion to adjust and the basic principle is that current signals I_1,I_2,……In are inputted in controller and the controller calculates current I_p as Eq. (228.1).

$$I_p = \frac{1}{n}\sum_{j=1}^{n} I_j \tag{228.1}$$

Compare I_p and I_j and get corresponding deviation current as Eq. (228.2).

$$\Delta I_1 = I_1 - I_p,\ \Delta I_2 = I_2 - I_p,\ \cdots,\Delta I_n = I_n - I_p \tag{228.2}$$

When $\left|\frac{\Delta I_j}{I_p} \times 100\%\right| < 5\%$, speed of corresponding output axles are not adjusted and when $\left|\frac{\Delta I_j}{I_p} \times 100\%\right| > 5\%$, controlling speed of output *axis* VI of mechanical system to realize power balance of multi-motor driving through controlling the open degrees of electro-hydraulic proportional throttle valve 6 to control fuel delivery of hydraulic motor. The controlling on open degrees of

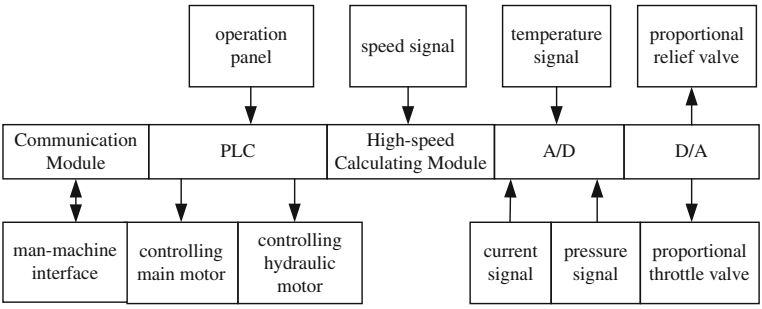


Fig. 228.2 Electric control system structure of controlled speeder for belt conveyor

electro-hydraulic proportional throttle valve need be combined with motor speed to confirm the working condition of motor.

228.4 Realization of Multi-Motor Power Balance

The electric control system structure of controlled speeder for belt conveyor, as shown in Fig. 228.2, is made up of PLC, sensors, man–machine interface and driving module of hydraulic control valve, specially explaining that PLC contains basic module, high-speed calculating module, A/D and D/A module. The PLC model is selected as FX2N32MT which adopts photoelectric separation techniques and has characteristics such as strong anti-jamming capability, high reliability and cost effectiveness. The electric control system controls the start/stop of main motor and hydraulic motor and the open degrees of proportional valve to realize the functions of soft start, soft stop and power balance of motor for belt conveyor through collecting and computing the output signals of pressure transmitter, speed sensor and current sensors by PLC. The electric control system is installed in explosion protection box to satisfy the safety request of mine control device and the window is set at the front of explosion protection box to install touch screen which contributes to operators to examine running parameters, execute fault alarm and alter system parameter of belt.

The driving of hydraulic control valve is the key part of electric control system and the E-ME-AC electronic amplifier which offers a switching current for electromagnet is adopted to drive electro-hydraulic proportional relief valve, that is to say, proportionally control adjustment measure of valve to the input current signal. The E-ME-T electronic amplifier is adopted to drive electro-hydraulic proportional throttle valve and this sort of amplifier which proportionally control the position of valve spool to changing input signal is used to proportional valve with valve spool position feedback and finally realize continuously variable speeder through offering a suitable current for electromagnet to adjust the flow of hydraulic motor. The driving circuit of hydraulic control valve is shown in Fig. 228.3.

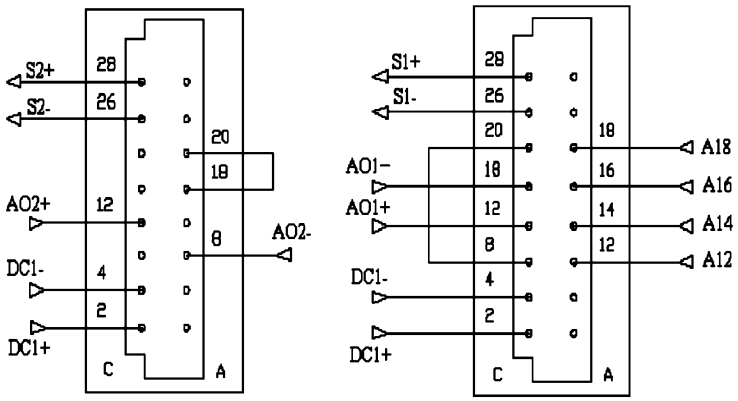


Fig. 228.3 Driving circuit of hydraulic control valve

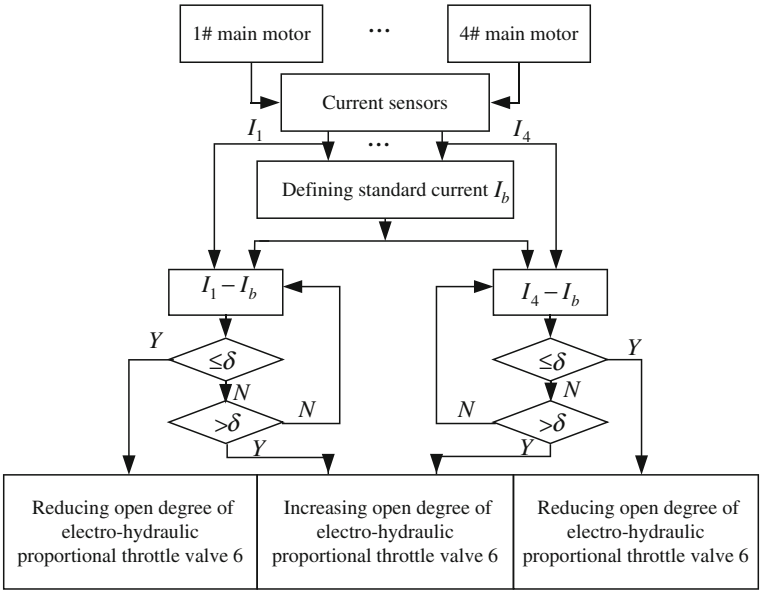


Fig. 228.4 Flow of power balance module

According to the control strategy of power balance, a standard current is defined first and then compares with the detected current signals of four motors from current sensors and adjusts the open degree of electro-hydraulic proportional throttle valve according to the compared values. Current sensor adopts hysteresis comparators and δ is the hysteresis width of current sensors and the flow of power balance module is shown in Fig. 228.4.

228.5 Conclusion

Controlled speeder for belt conveyor is an electromechanical product which is the organic combination of mechanical system, hydraulic system and electric control system and the electric control system adopts advanced PLC and matching modules as the main control components and has characteristics such as sound functions and low cost etc. The experiment tests show that the electric control system designed in this paper is able to accurately control mechanical drive and hydraulic devices and realize the functions of soft start and soft stop and the belt conveyor has characteristics such as steady power, quick response-time, extensive tape speed range and high transmission efficiency. The electric control system is able to meet the requirements of coal production and is an ideal driving device for long carrying distance and large carrying ability belt conveyor in our country.

References

1. Wang F-Z, Wang X-D, Wang S (2007) Hardware design of control system of belt conveyor driven by multi-motor based on PLC. Ind Mine Autom 1:37–40
2. Wang G-B (2003) Study on belt conveyor controlled start speeder equipment. J China Coal Soc 28(3):316–319
3. Fu Z, Zhao J, Shuai H (2009) Power balance control of multi-motor for belt conveyor system. J Xi Univ Sci Technol 29(4):478–481
4. Wang Z-G, M-S Ou Yang, Nie W (2006) Research of fuzzy controller on power balance for multi-motor driving belt system based on DSP. Coal Mine Machin 27(4):654–656
5. Dong-jin H, Ping L, Wei-liang J (2006) Research of power balance while using multi-point differential fluid viscidity timing clutch in conveyor belt system. J China Coal Soc 31(6): 829–832

Chapter 229
Research for Fuzzy PID Control Strategy in Linear Motor Feeding System

Fuzhong Wang and Panpan Li

Abstract In order to solve the problem of speed fluctuation in the linear motor drive feeding system. We introduced fuzzy theory on the basis of the traditional PID control algorithm, constituting fuzzy PID control algorithm self-setting PID parameter. Introduce the composing principle and design process of fuzzy PID controller. The MATLAB simulation experiment results show that this control algorithm response speed faster, good robustness, and overshoot is only 5%, satisfy the actual requirements of system.

Keywords Linear motor · Feeding system · Fuzzy PID

229.1 Introduction

Permanent magnet linear synchronous motor has double merit including permanent magnet motor and linear motor. It has a series of features [1] such as thrusting large, low loss, time constant small, fast response etc. But the permanent magnet linear synchronous motor special movement way, caused the edges at effect, the system internal variables of the association between serious, existing in the process of nonlinear time-varying, big, delay characteristics, and the influence of gravity and friction, so the motor running reliability, stability are poor and control

F. Wang (✉) · P. Li
Henan Polytechnic University,
Jiao zuo 454000, Henan, China
e-mail: wangfzh@hpu.edu.cn

P. Li
e-mail: libao19880828@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_229,

precision are low. In order to solve this problem, we must take effective control strategy to curb system itself parameters and external environment changes in the disturbance, which requires the control method has strong robustness [2]. Based on the traditional PID control algorithm, introduce the fuzzy set theory. Realize PID parameters of the setting, can improve the control effect better.

229.2 The Composing Principle of Fuzzy PID Controller

Linear motor feeding system aims at a nonlinear time-varying control object. Because single fuzzy control puts only deviation and deviation rate as input fuzzy quantity, not introduced integral mechanism, makes it not to eliminate steady-state error, the control accuracy is not high, and the independent PID control is powerless to the nonlinear time-varying object.

Fuzzy PID control comprehensive fuzzy control and PID control, ability to adapt to the advantages of good, strong anti-jamming capability, for nonlinear, large delay, and strong coupling of the controlled object control effect is very ideal. By constantly testing different time speed and the actual speed of a given error e and error change ec, online adjustment parameters, and so that was control object has good dynamic and static performance. The system structure as shown in Fig. 229.1.

This fuzzy PID controller will put speed error e and error rate ec as input of fuzzy controller. First through blur into fuzzy amount of e and ec, and then reasoning in the fuzzy reasoning machine according to the fuzzy rule, gets fuzzy output of three parameters of PID. And then after solving fuzzy into precise output, as input of PID controller.

Algorithm for the position of digital PID controller is as follows:

$$u(k) = K_p e(k) + K_i \sum_{j=0}^{k} e(j) + K_d[e(k) - e(k-1)]$$

K_p is ratio, K_i is integral coefficients, K_d is differential coefficient, $e(k)$ is the deviation, $u(k)$ is the output of the PID controller. So because every time the PID control parameters according to the specific control effect of real-time fixed amount, control for a change, and thus be able to do the optimal control function [3].

229.3 Design of Controller

In the specific design process of controller, because the input amounts have the speed error e and error change ec. Output is K_p, K_i, K_d. So the controller need second input and three output.

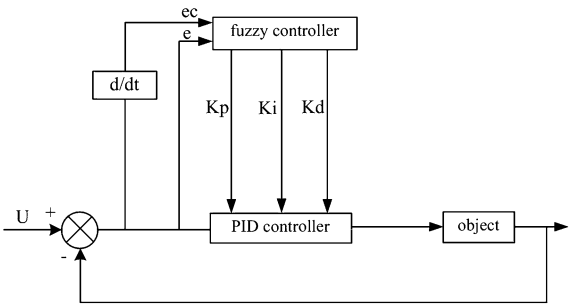


Fig. 229.1 Fuzzy PID controller system

Input and output of the fuzzy set are for {negative big (NB) negative middle (NM) negative small (NS) zero (ZE) small (PS) middle (PM) big (PB)};

Domains are for {6, 5, 4, 3, 2, 1, 0, 1, 2, 3, 4, 5, 6}; total of 13 level. Using fuzzy control toolbox in MATLAB, we can get all the variables of membership functions. Considering the domain to cover degree and sensitivity, stability, and robustness principle, In the domain of the left by "Z" type membership functions, the right end "S" type membership functions, other part adopts triangle membership functions.

PID parameters setting must consider the function of three parameters at different moment and the interaction relationship. We can get fuzzy control rule table shown in Table 229.1.

Fuzzy control lookup table got by the fuzzy reasoning and solving fuzzy, solving fuzzy by gravity method [4]. The fuzzy control lookup table are given in the Table 229.2, element is the amount of by the input domain to determine the amount of the control elements quantization value, and then take the appropriate scale factor, as the output.

229.4 Mathematical Model

d-q shaft voltage balance equation [5] of Motor as shown in type 1 below.

$$\begin{bmatrix} V_d \\ V_q \end{bmatrix} = \begin{bmatrix} R_s + pL_d & -\omega L_q \\ \omega L_d & R_s + pL_q \end{bmatrix} \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} 0 \\ \omega \psi_{PM} \end{bmatrix} \tag{229.1}$$

V_d, V_q — d shaft voltage, q shaft voltage
i_d, i_q —d shaft current, q shaft current
ψ_d, ψ_q — d shaft, q shaft Magnetic flux chain
R_s — The armature winding resistance
ω — The translation of the permanent magnet linear motor speed converted into Rotating motor speed

Table 229.1 Fuzzy control rules of K_P, K_i, K_d

E	EC						
	NB	NM	NS	ZO	PS	PM	PB
NB	PB,NB,PS	PB,NB,NS	PM,NB,NB	PM,NB,NB	PS,NS,NB	ZO,ZO,NM	ZO,ZO,PS
NM	PB,NB,PS	PB,NB,NS	PM,NM,NB	PS,NS,NM	PS,NS,NM	ZO,ZO,NS	NS,ZO,ZO
NS	PM,NB,ZO	PM,NM,NS	PM,NS,NM	PS,NS,NM	ZO,ZO,NS	NS,PS,NS	NS,PS,ZO
ZO	PM,NM,ZO	PM,NM,NS	ZO,NS,NS	NS,ZO,NS	NS,PS,NS	NM,PM,NS	NM,PM,ZO
PS	PS,NM,ZO	PS,NS,ZO	ZO,ZO,ZO	NS,PS,ZO	NS,PS,ZO	NM,PM,ZO	NB,PB,ZO
PM	PS,ZO,PS	ZO,ZO,PS	NS,PS,PS	NM,PS,PS	NM,PM,ZO	NM,PB,PS	NB,PB,PB
PB	ZO,ZO,PB	ZO,ZO,PM	NM,PS,PM	NM,PM,PM	NM,PM,PS	NB,PB,PS	NB,PB,PB

Table 229.2 Fuzzy lookup table of K_P

E	EC												
	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6
−6	6	5	6	5	6	6	3	3	1	0	0	0	0
−5	5	5	5	5	5	5	5	3	3	1	1	0	0
−4	6	5	6	5	6	6	6	3	3	1	0	0	0
−3	5	5	5	5	5	5	5	2	1	0	−1	−1	−1
−2	3	3	3	4	3	3	3	0	0	0	−1	−1	−1
−1	3	3	3	4	3	3	1	0	0	0	−2	−2	−1
0	3	3	3	4	1	0	0	−1	−1	−1	−3	−3	−3
1	2	2	2	2	0	0	−1	−3	−3	−3	−3	−3	−3
2	1	1	1	−1	0	−2	−3	−3	−3	−2	−3	−3	−3
3	0	0	0	−1	−2	−2	−5	−5	−5	−5	−5	−5	−5
4	0	0	0	−1	−3	−3	−6	−6	−6	−5	−6	−5	−5
5	0	0	0	−1	−3	−3	−5	−5	−5	−5	−5	−5	−5
6	0	0	0	−1	−3	−3	−6	−6	−6	−5	−6	−5	−6

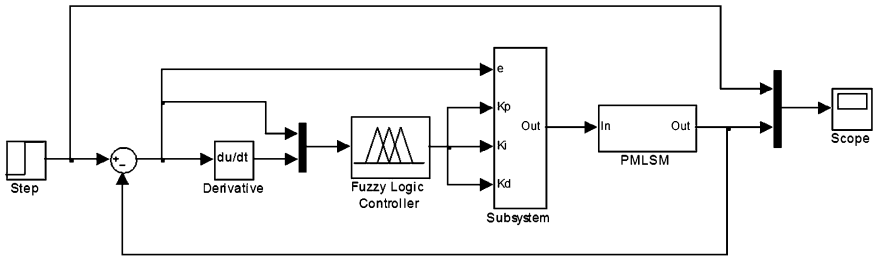


Fig. 229.2 Simulation schematic of fuzzy PID control

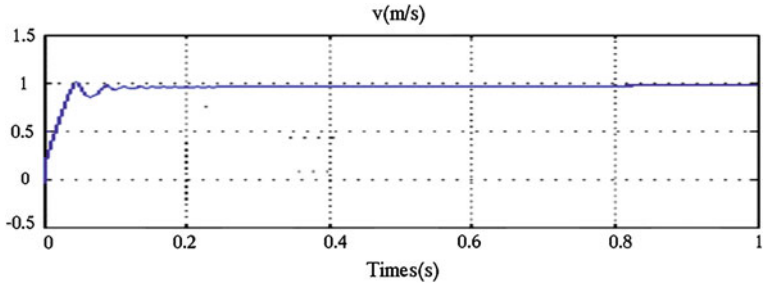


Fig. 229.3 Speed simulation waveform

ψ_{PM} — Magnetic chain of permanent magnet
p — Differential operator $\mathrm{d}/\mathrm{d}t$

Thrust formula of Linear motor as shown in type 2.

$$F_e = \frac{3\pi}{2\tau}\left[\psi_{\mathrm{PM}} i_q + (L_d - L_q) i_d i_q\right] \tag{229.2}$$

In addition, the mechanical motion equation of motor as shown in type 3.

$$F_e = F_L + B_v v + m\frac{\mathrm{d}v}{\mathrm{d}t} \tag{229.3}$$

F_L — Load resistance
B_v — Viscous friction coefficient
m — The quality of the moving parts

229.5 The Simulation Analysis

Permanent magnet linear synchronous motor simulation parameters adopted: $R_s = 2.64\ \Omega$, $L_d = L_q = 25.9$ mH, $\Psi_{PM} = 0.3\ \text{V} \cdot \text{s}$, $\tau = 16$ mm, $m = 28$ kg, $B_v = 4\ \text{N} \cdot \text{s/m}, F_L = 50$ N. MATLAB simulation model as shown in Fig. 229.2.

Through the corresponding parameter Settings, when given speed for 1 m/s, the sudden increase load 200 kg, electric motor speed curve as shown in Fig. 229.3.

Fuzzy PID controller can improve the system performance index, make the system more robust, the system responds faster, Overshoots is only 5%, and meet the demands of the system.

229.6 Conclusion

According to the research we get the conclusion as follows:

The effect of fuzzy PID control strategies in permanent magnet linear synchronous motor feeding system is better, improving the stability and reliability of the system.

Fuzzy PID controller which introduced fuzzy theory based on the traditional PID, can real-time adjust PID parameters, reach the required effect.

Make full use of fuzzy logic toolbox of MATLAB software, establish the simulation model of system, and the simulation results is close to the actual situation of system, have a great reference value.

References

1. Hongwei Z, Liucheng J, Runsheng K et al (2004) Fuzzy PID control for hoisting system drive by PMLSM. Micromotor China 37(6):34–36
2. Hongwei Z, Liucheng J, Xinhuan W, Fuzhong W, Runsheng K (2005) The power-angle control strategy research of the permanent magnet linear synchronous motor. Coal J 8:529–533
3. Shieh KY, Chang M (1994) A fuzzy controller improving a linear model following controller for motor drives fuzzy systems. IEEE Trans Electr Appl 21(3):194–202
4. Yunyue Y, Kaiyuan L (2001) PID control and fuzzy control in linear induction motor. Trans China Electrotech Soc 16(3):12–15
5. Huacai L, Yuetong X, Weimin Y, Zichen C (2007) Fuzzy PID control of linear motor feeding system. Electr Technol J 4:59–63

Chapter 230
Research on the Four-Switch Three-Phase Inverter Fed PMSM-FOC System

Xiaodan Wang, Hongli Qin and Duan Li

Abstract The Four-switch Three-phase Inverter (FSTPI) based AC drive system has attracted much attention because it can decrease the total system cost and simplify the system topology. It can also realize a fault tolerant topology for a post-fault reconfiguration for the traditional Six-switch Three-phase Inverter (SSTPI). In this paper, the Field Oriented Control (FOC) of the Permanent Magnet Synchronous Motor (PMSM) fed by the FSTPI was studied. First, the space vector pulse width modulation method for FSTPI was discussed and three equivalent zero-vector methods were investigated. Second, the control principle of the PMSM-FOC system was discussed as well. Based on the above analysis, the control model for the system was simulated and the results indicate that the FSTPI-PMSM-FOC drive system can operate steadily and has satisfied dynamic torque feature.

Keywords PMSM · Four-switch three-phase inverter · SVPWM · Torque control

X. Wang (✉) · D. Li
School of Electrical Engineering and Automation,
Henan Polytechnic University, Henan 454000, China
e-mail: wxd@hpu.edu.cn

D. Li
e-mail: zaosi007@163.com

H. Qin
Jiaozuo Construction Economy School,
Henan 454000, China
e-mail: liduan@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_230,

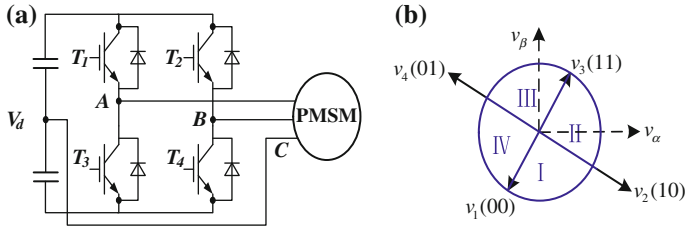


Fig. 230.1 The topology of FSTPI fed PMSM system and the vector distribution in the plane for the PMSM-FOC system

230.1 Introduction

The field oriented control (FOC) algorithm for induction machine, featured as simple implementation, good robustness to machine parameter variation, rapid dynamic response, was studied widely in the world [1, 2]. Due to the high power density, high efficiency and high torque output in the low speed, permanent magnet synchronous motor (PMSM) has wide perspective in electric vehicles and ship propulsion areas and the FOC method was widely used.

The conventional PMSM-FOC system is fed by six-switch three-phase inverter (SSTPI), which is the most fragile part in the system and the fault tolerant method is one of the hot topics in the research area. There are two general solutions: (1) using redundancy power electronics devices; (2) using post-fault configuration of SSTPI, named as four-switch three-phase inverter (FSTPI). The first solution increases the system hardware cost. The second solution can not only reduce costs, but also can simplify the system topology [3–7]. Compared with SSTPI, FSTPI only employs two legs and four switches and allows generating four active vectors in the $\alpha-\beta$ plane. So the space vector pulse width modulation (SVPWM) method is the key point for FSTPI.

In this chapter, the Field Oriented Control of PMSM fed by the FSTPI was studied. First, the space vector pulse width modulation method for FSTPI was analyzed. Second, the control principle of the PMSM-FOC system was discussed. Finally, the above analysis was validated by simulation results.

230.2 Space Vector Pulse Width Modulation Method for FSTPI

Figure 230.1a depicts the topology of FSTPI fed PMSM system. Compared to conventional SSTPI, the topology of FSTPI reduces two fully controlled power electronics switches and increases one electrolytic capacitor in the DC side.

Define S_a, S_b as the switching variable, in which $S_a = 0$ and $S_a = 1$ mean the conduction state of the lower and upper power electronics devices (IGBT or the

Table 230.1 Voltage vectors in the α–β plane for the PMSM-FOC system

S_a	S_b	$V = v_\alpha + jv_\beta$
0	0	$v_1 = (V_d/\sqrt{6})e^{-j2\pi/3}$
1	0	$v_2 = (V_d/\sqrt{6})e^{j\pi/3}$
1	1	$v_3 = (V_d/\sqrt{2})e^{-j\pi/6}$
0	1	$v_4 = (V_d/\sqrt{2})e^{j5\pi/6}$

anti-parallel diode) of phase A. So, there are four-switching states for FSTPI and every state associates with certain three phase instantaneous voltage value, which is different with the load connection pattern. For the PMSM-FOC system with the winding connected in wye, Table 230.1 shows the voltage vectors in the α–β plane and Fig. 230.1b depicts the vector distribution.

The characteristic of the switching vectors is as follows: (1) Their amplitude is different, $|v_2| = \sqrt{3}|v_1|$; (2) There is no zero vector; (3) There are only four vectors, which decrease the control freedom compared to SSTPI.

230.2.1 Space Vector Pulse Width Modulation Algorithm

Define v^* as the reference voltage to be synthesized by the SVPWM algorithm for FSTPI within a switching period T_s.

According to the pulse equivalent principle,

$$v^* T_s = v_1 t_1 + v_2 t_2 + v_3 t_3 + v_4 t_4 \tag{230.1}$$

in which the dwell time t_1, t_2, t_3, t_4 are restricted to,

$$T_s = t_1 + t_2 + t_3 + t_4 \tag{230.2}$$

In the α–β plane, v^* can be expressed as $v^* = v_\alpha^* + jv_\beta^*$, and Fig. 230.1b shows that $v_1 = -v_3 = v_{1\alpha} + jv_{1\beta}$, $v_2 = -v_4 = v_{2\alpha} + jv_{2\beta}$. So Eq. (230.2) can be simplified as and $t_{13} = t_1 - t_3$, $t_{24} = t_2 - t_4$.

$$v^* T_s = v_1 t_{13} + v_2 t_{24} \tag{230.3}$$

Rewriting Eq. (230.3) in the $\alpha\beta$ components, we get,

$$\begin{aligned} v_\alpha^* T_s &= v_{1\alpha} t_{13} + v_{2\alpha} t_{24} \\ v_\beta^* T_s &= v_{1\beta} t_{13} + v_{2\beta} t_{24} \end{aligned} \tag{230.4}$$

From Table (230.1), $v_{1\alpha} = -\sqrt{6}V_d/12$, $v_{1\beta} = -\sqrt{2}V_d/4$, $v_{2\alpha} = \sqrt{6}V_d/4$, $v_{2\beta} = \sqrt{2}V_d/4$ and t_{13}, t_{24} can be computed from Eq. (230.4),

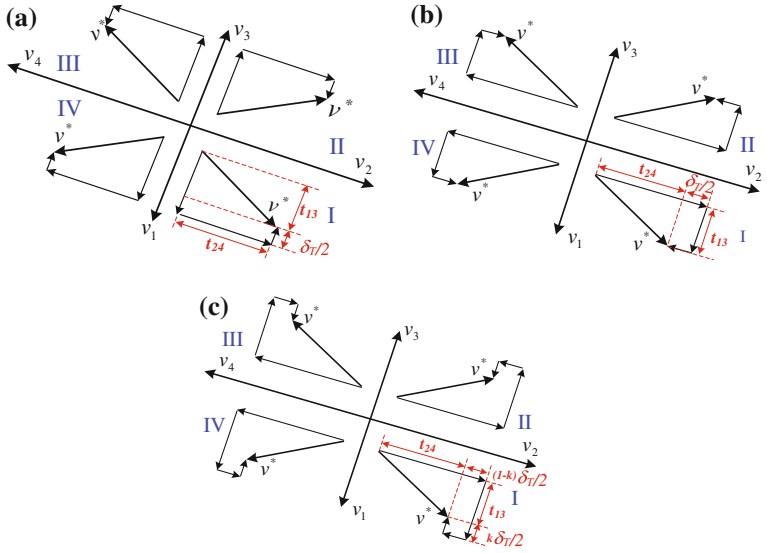


Fig. 230.2 Three different ways to alternate zero vector for FSTPI

$$\begin{cases} t_{13} = -\sqrt{\dfrac{3}{2}}(v_\alpha^* + \sqrt{3}v_\beta^*)\dfrac{T_s}{V_d} \\ t_{24} = \sqrt{\dfrac{3}{2}}(v_\alpha^* - \dfrac{1}{\sqrt{3}}v_\beta^*)\dfrac{T_s}{V_d} \end{cases} \tag{230.5}$$

As it can be seen that from (230.2), (230.5), there are four unknowns but only three equitions. So the dwell time t_1, t_2, t_3, t_4 cannot be calculated. And there is no zero vector, from (230.2), the acting time is,

$$\delta_T = T_s - |t_{13}| - |t_{24}| \tag{230.6}$$

From the relationship of the switching vector in Fig. 230.1b, $v_1 = -v_3$, $v_2 = -v_4$ during the time δ_T, the zero vector can be equivalent to the following three alternative ways:

(1) Vector v_1, v_3 are applied, seeing in Fig. 230.2a;
(2) Vector v_2,v_4 are applied, seeing in Fig. 230.2b;
(3) All the switching vectors are applied, seeing in Fig. 230.2c. The zero acting time δ_T is divided into two parts, $k\delta_T$ and $(1-k)\delta_T(0 \le k \le 1)$, in which a partion factor k is introduced. The calculted dwell time t_1, t_2, t_3, t_4 in different sector are shown in Table 230.2.

From Table 230.2, the value of partion factor k indicates the way to alternate zero vector for FSTPI. If $k = 0$ the first way is selected and if $k = 1$ the second way is selected. If $0 < k < 1$ all four vectors are employs. The calculated dwell time t_1, t_2, t_3, t_4 satisfied the Eqs. (230.2), (230.5).

Table 230.2 The calculated dwell time t_1, t_2, t_3, t_4 in different sector with the partion factor k

Sector	t_{13} , t_{24}	
I	$t_{13} > 0,\ t_{24} > 0$	$t_1 = t_{13} + \frac{k}{2}\delta_T; t_2 = t_{24} + \frac{1-k}{2}\delta_T;\ t_3 = \frac{k}{2}\delta_T;\ t_4 = \frac{1-k}{2}\delta_T;$
II	$t_{13} \leq 0,\ t_{24} > 0$	$t_1 = \frac{k}{2}\delta_T;\ t_2 = t_{24} + \frac{1-k}{2}\delta_T;\ t_3 = -t_{13} + \frac{k}{2}\delta_T;\ t_4 = \frac{1-k}{2}\delta_T;$
III	$t_{13} < 0,\ t_{24} \leq 0$	$t_1 = \frac{k}{2}\delta_T;\ t_2 = \frac{1-k}{2}\delta_T;\ t_3 = -t_{13} + \frac{k}{2}\delta_T;\ t_4 = -t_{24} + \frac{1-k}{2}\delta_T;$
VI	$t_{13} \geq 0,\ t_{24} < 0$	$t_1 = t_{13} + \frac{k}{2}\delta_T;\ t_2 = \frac{1-k}{2}\delta_T;\ t_3 = \frac{k}{2}\delta_T;\ t_4 = -t_{24} + \frac{1-k}{2}\delta_T;$

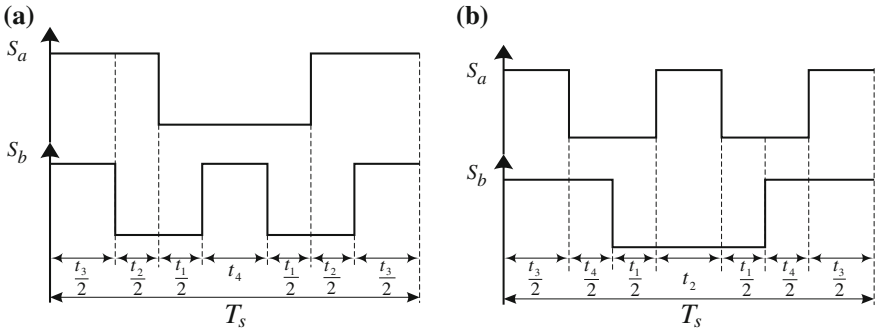


Fig. 230.3 Switching sequence for FSTPI. **a** $t_{24} < 0$. **b** $t_{24} < 0$

230.2.2 Space Vector Switching Sequence

With the space vector selected and their dwell time calculated, the next step is to arrange the switching sequence. For the SSTPI, the switching sequence satisfy the requirement that the transition from one switching vector to the next only involves two switches in the same inverter leg, one being switched on and the other switched off, which is realized by insert redundant zero vector $v_0(000)$ and $v_7(111)$. i.e. in the sector I, the space vector $v_1(001)$ and $v_2(011)$ are selected as the switching vector, so the switching sequence is: $v_0(000) \rightarrow v_1(001) \rightarrow v_2(011) \rightarrow v_7(111) \rightarrow v_2(011) \rightarrow v_1(001) \rightarrow v_0(000)$.

For FSTPI, the switching sequence also should satisfy the above requirement, a possible arrangement is shown in Fig. 230.3.

230.3 Control Principle of the FSTIP Fed PMSM-FOC System

The machematical model of PMSM in the synchronous frame is [7],

(1) Voltage equation:

$$\begin{aligned} u_d^r &= R_s i_d^r + pL_d i_d^r - \omega_e L_q i_q^r \\ u_q^r &= R_s i_q^r + pL_q i_q^r + \omega_e L_d i_d^r + \omega_e \psi_m \end{aligned} \tag{230.7}$$

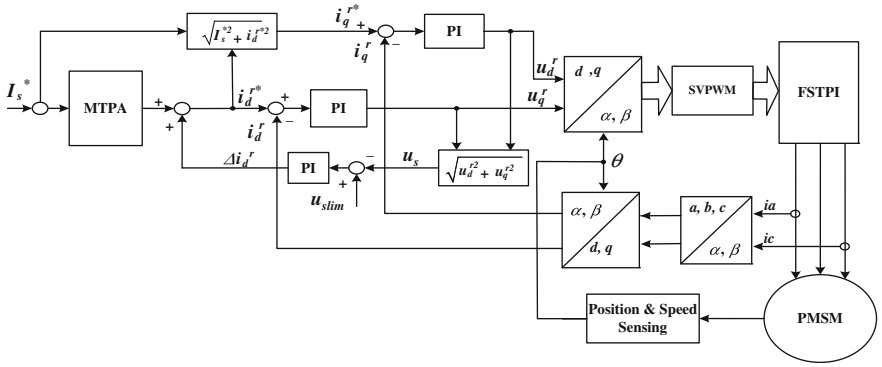


Fig. 230.4 Control principle frame of the FSTPI fed PMSM-FOC system

(2) Electromagnetic torque equation:

$$T_e = \frac{3P}{2}[i_q^r \psi_m + (L_d - L_q) i_d^r i_q^r] \tag{230.8}$$

in which, 'p' is differential operator, R_s is the stator resistance, L_d, L_q are the d- and q-axis inductance, P is the motor pole pairs, ψ_m is the permanent magnet flux linkage.

The control principle frame of the FSTPI fed PMSM-FOC system is depicted in Fig. 230.4, in which I_s^* is reference current, obtained by the electromagnetic-current table. According to the maximum torque per ampere (MTPA) curve, the reference current I_s^* is divided into d- and q-axis command current and the PI current regulator control the feedback current to track it quickly, accurately. The four-switch turning on/off signal of FSTPI is given by the modulation of reference voltage u_d^{r*}, u_q^{r*}, which is the output of current regulator.

For the SSTPI and FSTPI, the maximum output line-to-line voltage U_{slim} is limited by the DC link voltage. Flux-weakening control becomes necessary because rotor speed inceases, the back EMF also increases, which may exceed the maximum output voltage U_{slim}. In Fig. 230.5, the flux-weakening current component is generated by the unidirectional PI regulator, in which the input is the difference between u_{slim} and the voltage reference u_s.

230.4 Simulation

The control method of the FSTPI fed PMSM-FOC system is validated by simulation, parameters setting are as follows:

(1) PMSM: $R_s = 0.107\ \Omega$, Ld = Lq = 2.5 mH, $\psi_m = 0.175$, $P = 4$;
(2) FSTPI: $C_1 = C_2 = 3300\text{e-}6\text{uF}$, $Vdc = 330\ V$, switching frequency fsw = 10 kHz;

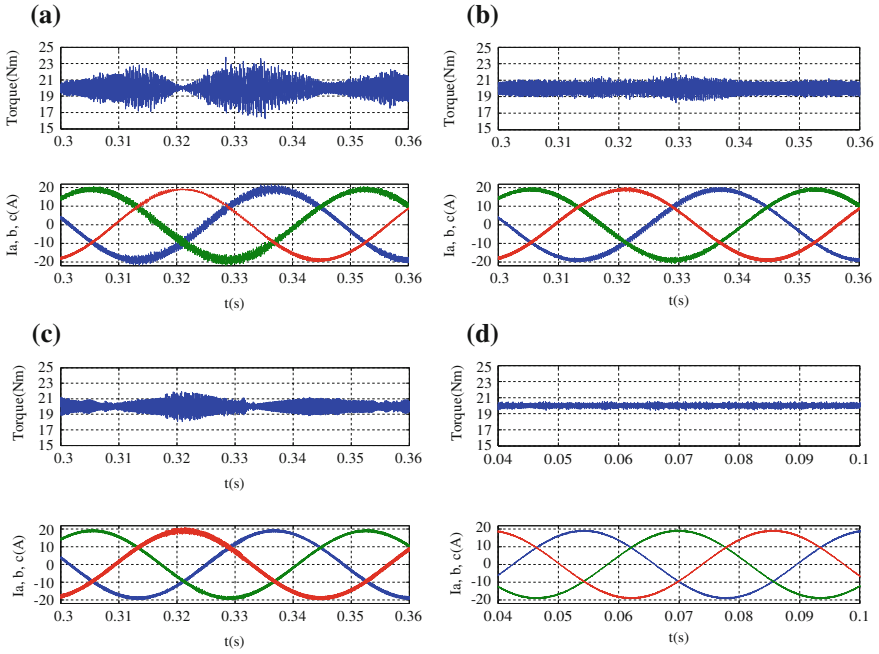


Fig. 230.5 Eletromagnetic torque and stator current of the PMSM-FOC system. **a**. FSTPI with $k = 0$. **b**. FSTPI with $k = 0.5$. **c**. FSTPI with k=1. **d**. SSTPI

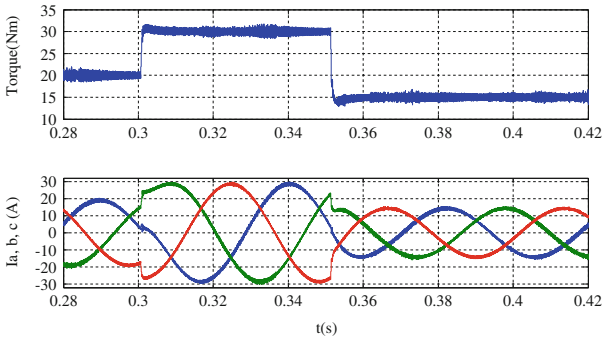


Fig. 230.6 Electromagnetic torque response curve of the FSTPI fed PMSM-FOC system

Figure 230.5a, b, c shows the simulation results of the FSTPI fed PMSM-FOC system with different partion factor value, in which the torque reference is sett as 20 Nm, the motor speed as 320 rpm. It indicates that when the partion factor $k = 0.5$, the electromagnetic ripple and stator current total harmonic distortion (THD) are significantly better than the other two cases. Fig. 230.5d depicts the

simulation results of the SSTPI fed system. Compared with Fig. 230.5b, the electromagnetic ripple and stator current THD decreases. It means there is a compromise between cost reducing and performance improving.

Figure 230.6 shows the electromagnetic torque response of the FSTPI fed PMSM-FOC system. At the simulation time $t = 0.3$ s, the torque reference sharply increases to 30 Nm and at $t = 0.35$ s, it decreases to 15 Nm. The system output electromagnetic torque can track the torque reference variation quickly.

230.5 Conclusion

In this paper, the FSTPI was applied in the PMSM-FOC system and three equivalent zero-vector methods were investigated, the analysis indicate that:(1) With the partion factor $k = 0.5$, the method that all four vector to synthesize zero vector improve the electromagnetic ripple and stator current total harmonic distortion; (2) The FSTPI-PMSM-FOC drive system can operate steadily and has satisfied dynamic torque feature.

References

1. Blaschke F (1972) The principles of field orientation as applied to the new closed-loop control system for rotating field machines. Siemens Rev 34:217–220
2. Kastha D, Bose BK (1994) Investigation of fault modes of voltage-fed inverter system for induction motor drive. IEEE Trans Ind Appl 30(4):1028–1038
3. Salazar L, Vasquez FC, Wiechmann E (1993) On the characteristics of a PWM AC controller using four switches. PESC' 93 11:307–313
4. Jacobina CB et al (1995) Vector and scalar control of a four switch three phase inverter. IAS' 95 3:2422–2429
5. Hong-Hee L, Dzung PQ (2009) The development of space vector PWM for four switch three phase inverter fed induction motor with DC—link voltage imbalance. ICIT 10–13:1–6
6. Beltrao de Rossiter Correa, etc. (Nov 2006) A general pwm strategy for four-switch three-phase inverters. IEEE Trans Power Electr 21(6):1618–1627
7. Jun H, Bin W New integration algorithms for estimating motor flux over a wide speed range. IEEE Trans Power Electr 13(5):969–977

Chapter 231
Research of Active Disturbances Rejection Control in Electric Wheel Dump Truck Speed Regulation System

Wenxiang Wei and Xiaoping Zhang

Abstract According to the poor robust control and large-scale parameter perturbation in electric wheel dump truck speed regulation system, a kind of controller based on active disturbances rejection controller (ARDC) is researched in rotor-flux-oriented vector control speed regulation system, analyzing the compensation of ARDC to perturbation of parameters and external disturbance, and dynamic decoupling of rotor flux and torque is achieved. The simulation results show the essential features of the presented method.

Keywords Electric wheel dump truck · ADRC · Parameter perturbation · Dynamic decoupling · Robust control

231.1 Introduction

The electric wheel autonomous dump truck is widely used in open mine transportation field, and owns great proportion in coal and iron ore mining, respectively, with 40 and 90%. The transportation cost in this field may vary up to 50% of its total cost, improving the transportation efficiency and reducing the cost is an important task. Therefore, the high-performance motor drive will be achieved in improving the development of the electric wheels autonomous dump truck [1].

W. Wei (✉) · X. Zhang
Engineering Research Center of Advanced Mining Equipment, Ministry of Education, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China
e-mail: weiwenxiang@sohu.com

X. Zhang
e-mail: zxp836@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_231,

Induction machines (IM) have several advantages with respect to other motor machines. In recent years, many control strategies are presented to extend the use of IM in those applications. One of these techniques is the well-known field-oriented control (FOC).

FOC is a very simple solution worth considering in many applications. Unfortunately, the calculation of the angular position of the rotor-flux vector is an important factor to get high-performance drive in FOC, and which is sensitive to errors in IM parameters especially the rotor resistance, and the rotor flux is not usually measurable; and the rotor resistance varies considerably due to heating. At present, control strategies based on PID is widely used in IM applications for its simplicity of construction [1]. However, PID can lead to the overshoot of output, and derivative of PID is not realized physically. A new-type active disturbance rejection controller (ADRC) proposed by Han [2], which can estimate and compensate the unknown internal and the external disturbance. The ADRC is a high-performance controller and is widely used in controlling of IM applications [3–6].

In this paper, a new strategy includes two ADRC proposed for IM speed control which is derived from [2]. The strategy reported here restructures the IM model and treats the parameters variation as the unknown internal disturbance and the variation of speed and torque is treated as the external disturbance.

231.2 Mathematic Model of ADRC

ADRC consists of three parts: Tracking differentiator (TD), Extended state observer (ESO) and Nonlinear state error feedback control law (NLSEF). The structure of a second-order is illustrated in Fig. 231.1.

For the sake of simplicity, consider a second-order plant.

$$\ddot{y} = f(t, y, \dot{y}, w(t)) = b(t)u, \tag{231.1}$$

where w(t), unknown extended disturbance, $f(t, y, \dot{y}, w(t))$ and b(t), unknown function, including the unknown internal disturbance and external disturbance. Assuming that $b_2 > b(t) > b_1 > 0$, and $b(t)$ selects a constant value b_0 which belongs to that rang. Thus, (231.1) can be rewritten as

$$\ddot{y} = f(t, y, \dot{y}, w(t)) + (b(t) - b_0)u + b_0 u, \tag{231.2}$$

where $(b(t) - b_0) \cdot u$ can be treated as a new disturbance, the extended state $z_3(t)$ of ESO can estimate the actual effect perfectly as the $|b(t) - b_0|$ remains in proper range, which can be written as

$$a(t) = f(t, y, \dot{y}, w(t)) + (b(t) - b_0)u, \tag{231.3}$$

Then $b_0 u$ can be used to realized the ADRC, which is equal to use $z_3(t)/b_0$ to compensate $a(t)$ derived from [2].

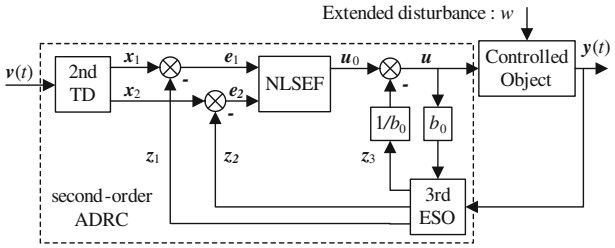


Fig. 231.1 The structure chart of second-order ADRC

231.3 Induction Machines Model Based on ADRC

Due to fact that parameters with large perturbation during low speed period in IM application of the electric wheels autonomous dump truck. A new strategy based on ADRC is presented, which illustrates the dependence on parameters compare to PI regulator.

A variable-frequency speed regulation system with SVPWM voltage convertor is discussed in the paper, the IM state-variable model written as [7]

$$\begin{cases} \dot{i}_{sd} = -K_1 i_{\mathrm{sd}} + \omega_1 i_{\mathrm{sq}} + K_2 \psi_{\mathrm{rd}} + \dfrac{1}{\sigma L_s} u_{\mathrm{sd}} \\ \dot{i}_{sq} = -K_1 i_{\mathrm{sq}} - \omega_1 i_{\mathrm{sd}} - K_4 \omega_r \psi_{\mathrm{rd}} + \dfrac{1}{\sigma L_s} u_{\mathrm{sq}} \\ \dot{\psi}_{\mathrm{rd}} = -\dfrac{1}{T_r}\psi_{\mathrm{rd}} + \dfrac{L_m}{T_r} i_{\mathrm{sd}} \\ \dot{\omega}_{\mathrm{r}} = K_3 \psi_{\mathrm{rd}} i_{\mathrm{sq}} - \dfrac{n_p}{J} T_L \end{cases}, \tag{231.4}$$

To simplify notations we use the reparametrization:

$$K_1 = \frac{R_s L_r^2 + R_r L_m^2}{\sigma L_s L_r^2}, \quad K_2 = \frac{L_m R_r}{\sigma L_s L_r^2}, \quad K_3 = \frac{n_p^2 L_m}{J L_r}, \quad K_4 = \frac{L_m}{\sigma L_s L_r},$$

where R_s and R_r are, respectively, the stator resistance and rotor resistance, L_s and L_r are, respectively, the stator and rotor self-inductance, L_m is the mutual inductance, i_{sd}, i_{sq}, u_{sd} and u_{sq} are, respectively, the d and q components of the stator current and voltage. The angular speed of the reference frame is ω_1, while ω_r is the rotor speed in electric rad/s. ψ_r is the rotor flux, σ is the total leakage factor ($\sigma = 1 - L_m^2/L_s L_r$), p is the number of pole pairs, J is the total motor and load inertia.

Estimation of the rotor-flux vector module ψ_{rd} and its angular position φ is extremely important, which is the precondition of decoupling the flux subsystem and rotor subsystem. The IM parameters time varying and perturbation are the main factor to reduce the decoupling ability of the two subsystems. Therefore, the IM model restructure based on ADRC described in Sect. 231.1 is a suitable solution.

231.3.1 Flux Subsystem of IM Based on ADRC

IM model described in (231.4) can extend state according to (231.1)–(232.3), and the standard form of ADRC can be obtained. ψ_{rd} is a constant value as IM work during steady state in speed subsystem, assuming that $x_1 = \omega_r, x_2 = K_3\psi_{\mathrm{rd}}i_{\mathrm{sq}} - n_pT/J$, the following equation can be obtained as

$$\begin{aligned}\dot{x}_2 &= K_3\psi_{\mathrm{rd}}\dot{i}_{\mathrm{sq}} = K_3\psi_{\mathrm{rd}}\left(-K_1 i_{\mathrm{sq}} - \omega_1 i_{\mathrm{sd}} - K_4\omega_r\psi_{\mathrm{rd}} + \frac{1}{\sigma L_s}u_{\mathrm{sq}}\right)\\ &= -K_1\left(K_3\psi_{\mathrm{rd}}i_{\mathrm{sq}} - \frac{n_p}{J}T_L\right) - K_1\frac{n_p}{J}T_L - K_3\psi_{\mathrm{rd}}\cdot\omega_1 i_{\mathrm{sd}} - K_3\psi_{\mathrm{rd}}\cdot K_4\psi_{\mathrm{rd}}\omega_r + K_3\psi_{\mathrm{rd}}\cdot\frac{1}{\sigma L_s}u_{\mathrm{sq}}\end{aligned} \tag{231.5}$$

$$\begin{aligned}\dot{x}_2 = &-K_1x_2 - K_3K_4\psi_{\mathrm{rd}}^2x_1 - K_3\psi_{\mathrm{rd}}\omega_1 i_{\mathrm{sd}} - K_1\frac{n_p}{J}T_L + \frac{K_3}{\sigma L_s}(\psi_{\mathrm{rd}} - \frac{1}{2}\psi_{\mathrm{rd}}^*)u_{\mathrm{sq}}\\ &+\frac{K_3}{2\sigma L_s}\psi_{\mathrm{rd}}^*u_{\mathrm{sq}}\end{aligned} \tag{231.6}$$

The following expression can be written as

$$\begin{cases} f_1(t,x_1,x_2,w_1) = -K_1x_2 - K_3K_4\psi_{\mathrm{rd}}^2x_1 - K_3\psi_{\mathrm{rd}}\omega_1 i_{\mathrm{sd}} - K_1\dfrac{n_p}{J}T_L\\ b(t) - b_0 = \dfrac{K_3}{\sigma L_s}(\psi_{\mathrm{rd}} - \dfrac{1}{2}\psi_{\mathrm{rd}}^*)\\ b_0 = \dfrac{K_3}{2\sigma L_s}\psi_{\mathrm{rd}}^* \end{cases} \tag{231.7}$$

where ψ_{rd}^* is the reference rotor-flux value of IM, b_0 is half of the $b(t)$, thus, following result can be obtained.

$$\dot{x}_2 = f_1(t,x_1,x_2,w_1(t)) + (b(t) - b_0)u_{\mathrm{sq}} + b_0u_{\mathrm{sq}}, \tag{231.8}$$

231.3.2 Speed Subsystem of IM Based on ADRC

Restruction of the speed subsystem controller is similar to Sect. 231.3.1. Assuming that $x_3 = \psi_{\mathrm{rd}},\ x_4 = (L_mi_{\mathrm{sd}} - \psi_{\mathrm{rd}})/T_r$,the following equation can be obtained as

$$\begin{cases} f_2(t,x_3,x_4,w_2(t)) = -\left(K_1 + \dfrac{1}{T_r}\right)x_4 + \dfrac{1}{T_r}(K_2L_m - K_1)x_3 + \dfrac{L_m}{T_r}\omega_1u_{\mathrm{sq}}\\ b'(t) - b_1 = K_4\left(R_r - \dfrac{1}{2}R_r^*\right)\\ b_1 = \dfrac{1}{2}K_4R_r^* \end{cases} \tag{231.9}$$

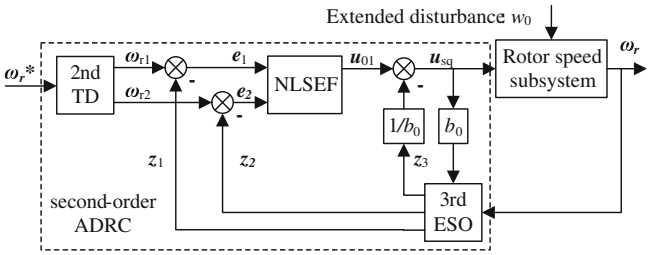


Fig. 231.2 The ADRC structure chart of flux subsystem

where R_r^* is the real rotor resistance value of IM, assuming that $b(t)$ is equal to $K_4 R_r$, thus, following result can be obtained.

$$\dot{x}_4 = f_2(t, x_3, x_4, w_2(t)) + (b^{'}(t) - b_1)u_{sd} + b_1 u_{sd} \tag{231.10}$$

Equations (231.8) and (231.10) are the typical form of second-order ADRC derived from [2], which can be used to regulate the flux subsystem and speed subsystem.

231.4 Design IM Variable-Frequency Speed Regulation System

According to ADRC theory and its construction method introduced in Sect. 231.3, the two subsystem based on ADRC can be obtained, the speed subsystem is depicted in Fig. 231.2, the structure of the flux subsystem is similar to Fig. 231.2. The IM variable-frequency speed regulation system with FOC based on ADRC is depicted in Fig. 231.3.

231.5 Simulation Results

To verify the performance of the system depicted in Fig. 231.4. The proposed control system-based ADRC compared with PI controller has been simulated under large inertia and rotor resistance variations, and external torque disturbance in Matlab/Simulink [9]. The IM parameters are listed in Table 231.1.

The simulation results have been obtained by implementing the control scheme depicted in Fig. 231.3. The reference rotor-flux vector ψ_{rd}^* is 1 Wb, the reference rotor speed is 120 rad/s, the initial torque is rated torque, and it varies 30 N m at $t = 2.5$ s. The rotor resistance varies at $t = 2$ s with 50% increase and $t = 2.5$ s with 100% increase.

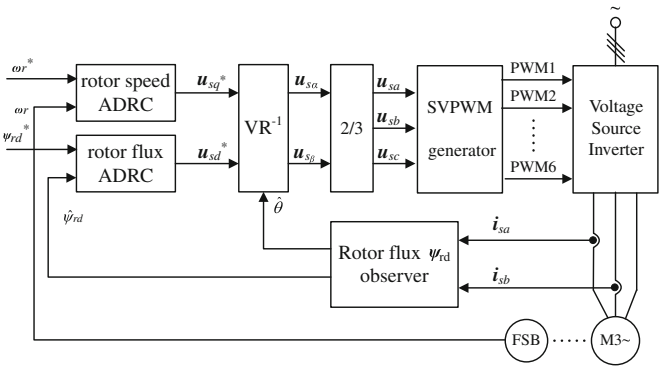


Fig. 231.3 The induction-motor variable-frequency speed regulation system based on ADRC

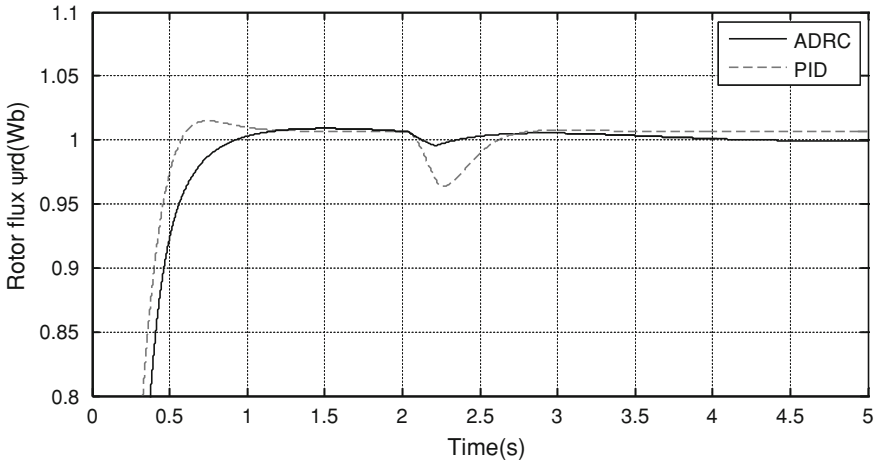


Fig. 231.4 The rotor-flux waveform of control system based on ADRC and PI regulator

Table 231.1 Simulation parameters of induction machine

Rated output power	22 KW	Stator resistance	0.284 Ω
Rated voltage	380 V	Rotor resistance	0.258 Ω
Rated frequency	50 Hz	Stator self-induction	0.228 H
Rated speed	1,450 rpm	Rotor self-induction	0.234 H
Pole pair	2	Mutual induction	0.220 H
Rate flux	1.0 Wb	Load inertia	0.268 kg m^2

The response of the rotor flux, ψ_{rd}, is shown in Fig. 231.4: the overshoot of ψ_{rd} is over 2.34% with PI regulator while ADRC [8] is equal to zero; the rotor-flux perturbation is over 4.55% with PI regulator as the rotor resistance varies at

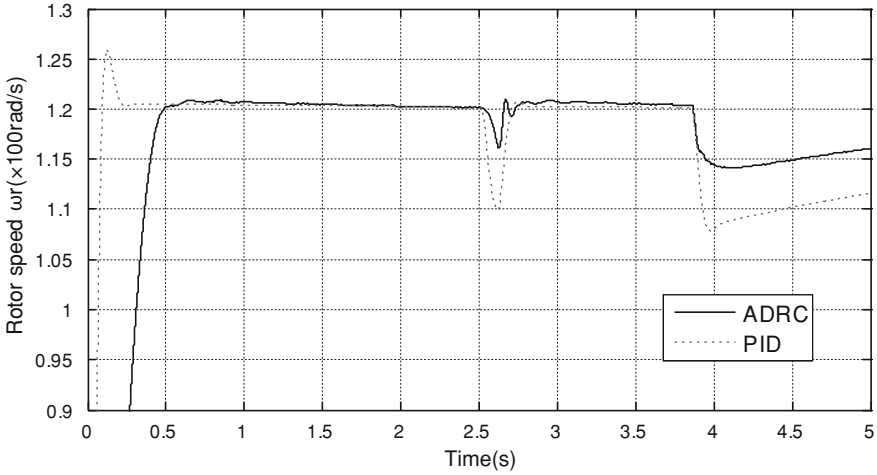


Fig. 231.5 The rotor speed waveform of control system based on ADRC and PI regulator

$t = 2$ s with 50% increase, while ADRC is less than 1.0%, and the recovery time of ψ_{rd} is obviously fast with ADRC.

The response of the rotor speed, ω_r, is shown in Fig. 231.5: the overshoot of ω_r is over 5.0% with PI regulator while ADRC is equal to zero; the rotor speed overshoot is over 8.33% with PI regulator as the rotor resistance varies at $t = 2.5$ s with 100% increase, while ADRC is less than 3.0%, and the control ability of the system with ADRC is better as external torque 30 N m disturbs at $t = 2.5$ s.

231.6 Conclusion

This paper has presented a new-type structure controller for IM with FOC based on the ADRC. The control strategy is suitable for application of controlling the electric wheels autonomous dump truck, with complex work conditions of lager speed various rang, frequent starting and braking, large-scale parameters perturbation at low speed, etc. The simulation results show that it has strong robustness, better performance with uncertain disturbance, and it do not depend on the accuracy IM model. It promotes a new method on researching different motor control.

Acknowledgments This work has been Project Supported by Hunan Province Innovative Platform for Integration of Industries, Universities and Research (Grant no. 2010XK6066); Hunan Provincial Natural Science Foundation of China (10JJ9007).

References

1. Zhu GH (2009) Suggestion of the "twelfth five years" solid mineral resources exploitation efficient and utilization of major technical equipment electric wheel dump truck. China Electr Equip Ind 4:70–73
2. Ren YF, Liu G, Zhao M (2009) Application of active disturbance rejection control in direct torque control of induction motor. Electr Drive 39(6):15–18
3. Shao LW, Liao XZ, Zhang YH (2006) Active disturbance rejection controller scheme for permanent magnetic synchronous motor. Electr Drive 36(10):3–6
4. Bai J, Li HD, Hao ZH (2005) Induction-motor variable-frequency speed regulation system implemented by auto-disturbances-rejection controller. Trans China Electrotechn Soc 20(6):73–76
5. Su WF, Sun XD, Li FH (2004) Vector control of induction motors with active disturbance rejection control. J Tsinghua Univ (Sci Technol) 44(10):1329–1332
6. Han JQ (1998) Auto-disturbances-rejection controller and its applications. Control Decis 13(1):19–23
7. Chen BS (2003) Electrical towage automatic control system. Mechanical Industry Press, Beijing, pp 200–214
8. Liu M, Shao C (2003) Auto-disturbances-rejection controller and its parameter adjusting for asynchronous motor. Control Decis 18(5):540–544
9. Xue DY, Chen YQ (2002) System simulation technology and application based on MATLAB/Simulink. Tsinghua University Press, Beijing, pp 350–362

Part XVI
Specific Motor and Drives

Chapter 232
Back-EMF Estimation Based on Extended Kalman Filtering in Application of BLCD Motor

Dongliang Liu, Yao Zhang and Lei Pan

Abstract This paper proposes a new back-EMF detection method for BLDC motor, which is based on the extended Kalman filter. Through setting the stator current and the back-EMF voltage as the state variables, the model of the extended Kalman filter could be established from the derivation of the BLDC motor's mathematical model. Accurate speed and position control could be carried out then, and the experiment results show the feasibility and accuracy of this method.

Keywords BLDC motor · Extended Kalman filter · Back-EMF estimation · Zero-crossing detection

232.1 Introduction

The BLCD motor has advantages such as simple structure, easy to control, high reliability, great power density and high efficiency which has been applied in various industrial fields. In recent years, sensorless control of BLCD motor has been the hot research at home and abroad. Common methods of rotor position signal detection have back-EMF signal detection method, the stator inductance method, free-wheeling diode method, magnetic flux linkage estimation method and the state observer method, etc. [1–7]. The most practical and effective method

D. Liu (✉) · Y. Zhang
Hangzhou University of Electronic Science and Technology,
Institute of Automation, Hang Zhou 310018, Zhejiang, China
e-mail: liudl@hdu.edu.cn

L. Pan
Department of Electronic and Information Engineering,
Tianjin Institute of Urban Construction, Tianjin 300384, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_232,

is back-EMF estimation among these methods. The method e.g. [8] is through detect back-EMF respectively in pwm shut off and opening phrase and get back-EMF zero crossing point using two different reference voltage in the low speed and high speed without position sensors and current sensors, but increasing the complexity of the hardware circuit. The method in [9] which gets back-EMF zero crossing by comparing the terminal voltage of impending phase windings and the midpoint voltage of Inverter dc circuits does not need reconstructing motor central point and using filter circuit, but needs hardware circuit to get zero crossing point. The method of Extended Kalman Filtering in Application of BLCD Motor [10] has great lag and poor usability.

In this paper, a new method of back-EMF detection based on Extended Kalman Filtering of BLCD Motor is proposed. The proposed method which makes phase current and back-EMF as state variables is based on the mathematical model of BLDC motor, and it obtained the model of Kalman filter. The method can control the speed and position, to a large extent, it overcomes the problems of directly back-EMF's poor anti-interference ability, delay calculation and phase shift of low-pass filtering and so on [1–6].

232.2 The Basic Principle of Back-EMF Detection

Figure 232.1 shows ordinary BLCD motor which adopts three-phase voltage type inverter to supply power, the stator winding is star connected. Because of BLCD motor midpoint n general is not derivation, so it is difficult to accurately construct phase voltage equation [7–9], therefore considering to build line voltage equation, and subtract phase voltage with each other can get line voltage expression [5, 7]:

$$\begin{cases} u_{ab} = R(i_a - i_b) + Lp(i_a - i_b) + (e_a - e_b) \\ u_{bc} = R(i_b - i_c) + Lp(i_b - i_c) + (e_b - e_c) \\ u_{ca} = R(i_c - i_a) + Lp(i_c - i_a) + (e_c - e_a) \end{cases} \tag{232.1}$$

In the formula (232.1), p: differential operator, d/d$_t$; R: stator resistance; Ls: stator phase windings self-induction; M: stator phase windings mutual-inductance; $L = \mathrm{Ls} - M$.

Then consider the difference value of line voltage, for example, u_{ab}–u_{bc}:

$$u_{ab} - u_{bc} = R(i_a + i_c - 2i_b) + Lp(i_a + i_c - 2i_b) + (e_a + e_c - 2e_b) \tag{232.2}$$

Assuming a phase and c phase are breakover, b is shut off, then $e_a + e_c = 0$ (ideal back-EMF and current curve as Fig. 232.2 shows). By three-phase brushless DC motor stator winding coupling star, we can see:

$i_a + i_b + i_c = 0$, then formula (232.2) can be simplified as:

$$u_{ab} - u_{bc} = -3\mathrm{R}i_b - 3\mathrm{L}pi_b - 2e_b \tag{232.3}$$

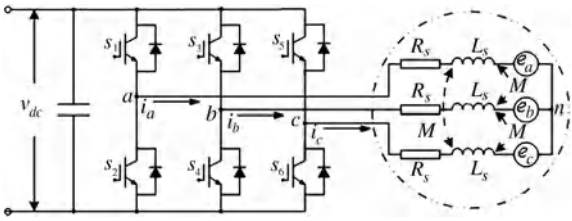


Fig. 232.1 Three-phase brushless DC circuit diagram machine

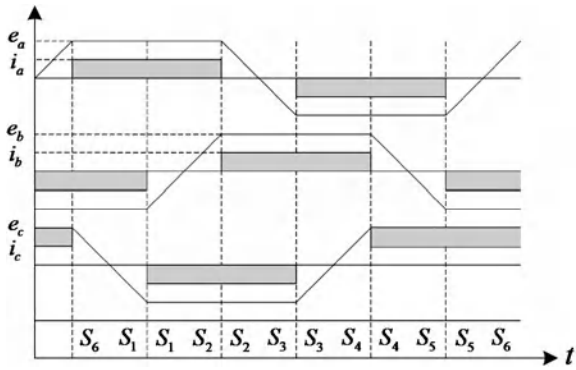


Fig. 232.2 Ideal back-EMF voltage and current waveform

Namely:

$$\dot{i}_b = -\frac{3R}{3L} i_b + \frac{1}{3L}(u_{bc} - u_{ab}) - \frac{2}{3L} e_b \tag{232.4}$$

232.3 The Principle of Extended Kalman Filter

The Kalman filtering method is system noise normal distribution, the filter gives the minimum variance estimation, state when not normal situations, the filter gives state linear minimum variance estimation. Kalman filtering is developed on the basis of linear system, however, actual system tend to be nonlinear. Motor system is a nonlinear system. For the application of the nonlinear system, the key problem is nonlinear equation linearization problem. On this basis, discrete linear system of Kalman filtering estimation formula can be applied in the prototype nonlinear system [11]. Set of nonlinear differential equations as follows:

$$\dot{x} = f(x, U) \tag{232.5}$$

In the formula (232.5), x is the state vector for the system, U is interferential function vector for the system. Then two Jacobin matrix are respectively as follows:

$$A_{[i.j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(x, U, 0) \tag{232.6}$$

$$B_{[i.j]} = \frac{\partial f_{[i]}}{\partial U_{[j]}}(x, U, 0) \tag{232.7}$$

Then, $\dot{x} = f(x, U)$ can be written as:

$$\dot{x} = A(x, t)\, x(t) + B(x, t)\, u(t) \tag{232.8}$$

Namely linear interference equation [11–14]. When it has the initial condition $x(t_0) = x_0$ the solution of nonhomogeneous linear differential equation is:

$$x = \phi(t, t_0) \cdot x_0 + \int_0^t \phi(t, t_0) \cdot G(t) \cdot u(t)dt \tag{232.9}$$

The state transition matrix $\phi(t, t_0)$ is the solution of the equation $\mathrm{d}\phi(t, t_0)/\mathrm{d}t = F(t) \cdot \phi(t, t_0)$. The initial condition is $\phi(t, t_0) = I$, make $\Gamma(t_1, t_0) = \int_0^t \phi(t_1, t) \cdot G(t)\mathrm{d}t$, Then the formula can be written as the form of difference equation as follows:

$$x(t_1) = \phi(t_1, t_0) \cdot x(t_0) + \Gamma(t_1, t_0) \cdot u(t_0) \tag{232.10}$$

In order to measure the equation's linearization, suppose:

$$y = h(x) \tag{232.11}$$

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(x, 0) \tag{232.12}$$

Then formula (232.11) can be written as:

$$y = H(t) \cdot x(t) \tag{232.13}$$

Obviously, add white noise sequence W_{k-1} and V_k to formula (232.8) and (232.13), the nonlinear problem of Kalman filter [11–14] is completed.

The following is the general steps to construct extended Kalman filter [10, 11]:

(1) Calculate the value of state prediction

$$\hat{x}(k+1)/k = F_k(\hat{x}k/k) \cdot \hat{x}k/k + G_k(\hat{x}k/k) \cdot u_k \tag{232.14}$$

where: $\hat{x}k/k$ is the updated value at the moment of $t_k = kT$, $\hat{x}(k+1)/k$ is the predicted value at the moment of t_k to $t_{k+1} = (k+1)T$.

(2) State error covariance matrix

$$\hat{P}_{(k+1)/k} = F'_k(\hat{x}_{k/k}) \cdot \hat{P}_{k/k} \cdot F'^T_k(\hat{x}_{k/k}) + Q_k \tag{232.15}$$

$$F'_k = \frac{\partial[F_k(x_k)x_k + G_k(x_k)u_k]}{\partial x_k}|_{x_k = x_{k/k}} \tag{232.16}$$

where: $\hat{P}(k+1)/k$ is the predicted value of state covariance at the moment of t_k to $t_{k+1} = (k+1)$ [11].

(3) Gain of Kalman filter

$$\begin{aligned} K(k+1) &= \hat{P}[(k+1)/k]H[x(t_{k+1})]^T \\ &\times [H[x(t_{k+1})]\hat{P}(k+1)/k\, H[x(t_{k+1})]^T + R_k]^{-1} \end{aligned} \tag{232.17}$$

where: R_k is noise covariance, $K(k+1)$ is the gain of Kalman filter at the moment of $t_{k+1} = (k+1)T$.

(4) The updates of state error covariance matrix

$$P_{(k+1)/(k+1)} = \{I - K(k+1)H[x(t_{k+1})]\} \cdot \hat{P}_{(k+1)/k} \tag{232.18}$$

where: y_{k+1} is the measured value of observed equation at the moment of t_{k+1}.

(5) The updates of state predicted value

$$\hat{x}_{(k+1)/(k+1)} = \hat{x}_{(k+1)/k} + K_{k+!}\{y_{k+1} - H[\hat{x}(k+1/k)]\hat{x}(k+1/k)\} \tag{232.19}$$

where: y_{k+1} is the measured value of observed equation at the moment of t_{k+1}.

232.4 Back-EMF Estimation Based on Kalman Filter

Because i_b, u_{bc} and u_{ab} in formula (232.4) could be measured, so which are taken as the known state variables; and because e_b could not be measured directly, so which is taken as the unknown state variables. Formula (232.4) can be expressed as:

$$\dot{x}_b = A_b x_b + B_b u_b + F_b w_b \tag{232.20}$$

$$y_b = C_b x_b \tag{232.21}$$

where:

$$A_b = \left[-\frac{3R}{3L}\right],\ B_b = \left[\frac{1}{3L}\right],\ F_b = \left[-\frac{2}{3L}\right],\ x_b = [i_b],\ u_b = [u_{bc} - u_{ab}],$$
$$w_b = [e_b],\ C_b = [1]$$

In the formula above, the back-EMF is taken as unknown disturbing signal, and this disturbing signal can be described by difference equation as follows:

$$\dot{z} = Dz \tag{232.22}$$

$$w = Hz \tag{232.23}$$

where:

$$D = \begin{bmatrix} 0_{(\delta-1)\times 1} & I_{(\delta-1)} \\ 0_{1\times 1} & 0_{1\times(\delta-1)} \end{bmatrix}, H = [I_1 \quad 0_{1\times(\delta-1)}], w = \sum_{i=0}^{\delta} a_i t^i,$$

I is unit matrix; δ is degree of polynomial w, and $\delta \geq 1$; a_i is the coefficient of unknown vector.

When there is no disturbance, make $a_i = 0$. The methods above can be described most disturbance signal and some unknown disturbance signal by changing the degrees of polynomial. Therefore, without loss of generality, the formula (232.22) and (232.23) of disturbance model is completely considerable. So the back-EMF equation can be described by difference equation of disturbance introduced. Such as formula (232.24) below:

$$\dot{x}_b = A_b x_b + B_b u_b \tag{232.24}$$

$$y_b = C_b x_b \tag{232.25}$$

where:

$$A_b = \begin{bmatrix} A & FH \\ 0 & E \end{bmatrix} = \begin{bmatrix} -\frac{3R}{3L} & -\frac{2}{3L} \\ 0 & 1 \end{bmatrix}, x_b = \begin{bmatrix} i_b \\ e_b \end{bmatrix}, B_b = \begin{bmatrix} B \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3L} \\ 0 \end{bmatrix},$$

$$u_b = [u_{bc} - u_{ab}], C_b = [C \quad 0] = [1 \quad 0]$$

Similarly, we can obtain the state model of the other two phase, and finally the system state model can be achieved:

$$\dot{x} = Ax + Bu \tag{232.26}$$

$$y = Cx \tag{232.27}$$

where:

$$x = [i_a \quad e_a \quad i_b \quad e_b \quad i_c \quad e_c]^T;$$

$$A = \begin{bmatrix} -\frac{3R}{3L} & -\frac{2}{3L} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{3R}{3L} & -\frac{2}{3L} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{3R}{3L} & -\frac{2}{3L} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} \frac{1}{3L} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{3L} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3L} \\ 0 & 0 & 0 \end{bmatrix},$$

$$u = \begin{bmatrix} u_{ab} - u_{bc} \\ u_{bc} - u_{ab} \\ u_{ca} - u_{ab} \end{bmatrix}, y = \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Add white noise sequence W_{k-1} and V_k to formula (232.26) and (232.27), we can complete the Kalman filtering nonlinear problem of back-EMF detection of BLCD Motor.

Use the method of Sect. 232.2 to make iterative operation, the back-EMF of running motor can be obtained. The motor can be controlled by reversing direction when the phase current nearly equal to 0 and the back-EMF equal to 0.

232.5 Calculation of the Velocity and the Position

The value of the back-EMF of each phase is able to be estimated by the Kalman Filter according to the content above, which can be used to estimate the velocity and the position of the motor through a simple calculation [15–18], the details are displayed as follows.

$$E = K_e \omega_e \tag{232.28}$$

where E means the amplitude of the back-EMF, K_e means the constant, ω_e means the angular velocity of the motor. And $\omega_e = \frac{E}{K_e}; \omega_m = \frac{2}{P}\omega_e$, in this formula, ω_m means the mechanical velocity of the motor, P means the number of the pole pairs, then the position of the rotor of the motor.

$$\theta = \int \omega_e \mathrm{d}t + \theta_0 \tag{232.29}$$

θ means the angle of the rotor's position, θ_0 means the angle of the initial position.

232.6 The Experiment and the Analysis of the Results

The whole experiment framework of the velocity control of the brushless DC motor is shown as Fig. 232.3. The parameters of the brushless DC motor of the experiment system are displayed as Table 232.1. The maximum velocity variation set in the program is 2,000 rmp/s.

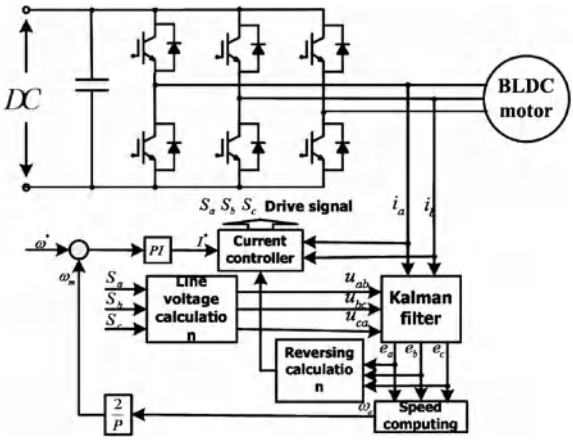


Fig. 232.3 The framework of the control system

Table 232.1 Parameters of the brushless DC motor

Rated voltage	V	310 V
Rated torque	Te	15 Nm
Rated speed	Nr	1,000
Stator resistance	Rs	1.5 Ω
Stator inductance	L	4.22 mH
Rotary inertia	J	0.08 kg m^2
Back- EMF constant	Ke	0.175
Magnetic pole logarithm	Pn	2

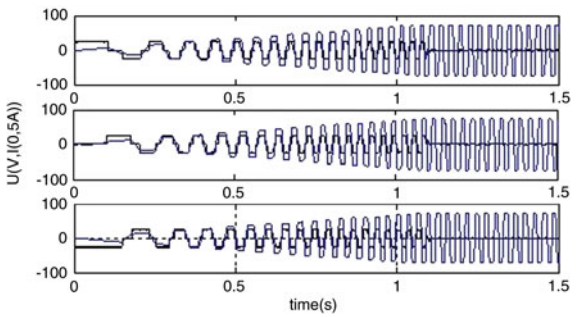


Fig. 232.4 A, B, C the current of three phase (*black*) and the curve of the back-EMF(*blue*)

The experiment of the system is consisted of three parts:

(1) When the given velocity reaches 1,000 rpm, the results of no-load are shown as Figs. 232.4, 232.5, 232.6.
It can be seen from Fig. 232.4 that the curve is similar to the curves of the ideal back-EMF and the current, which makes that the phase current is nearly

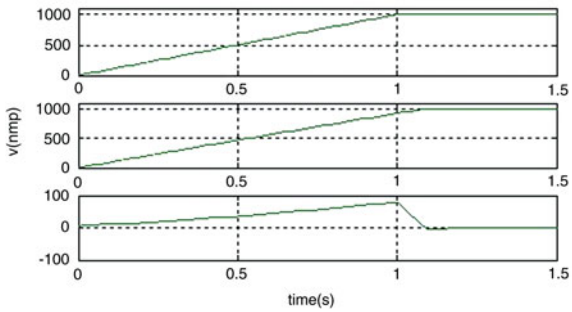


Fig. 232.5 The curves of the given velocity, actual velocity and the velocity error

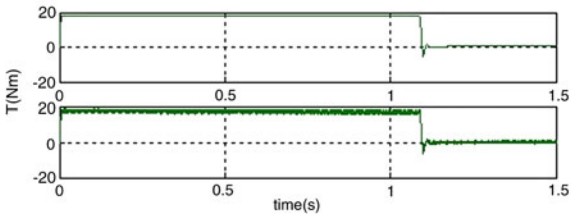


Fig. 232.6 The curves of the given torque and the actual torque

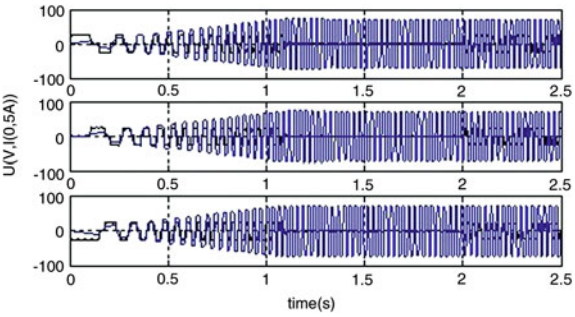


Fig. 232.7 A, B, C the current of three phase (*black*) and the curve of the back-EMF (*blue*)

equal to 0 and the examination of the over-zero of the back-EMF come true and there is no obvious after flow phenomenon. And Fig. 232.5 shows that the velocity tracing error can be considered as a linear increase approximately at the moment of the startup of the motor. It will fall rapidly with a little overstrike. The transition progress will last about 0.1 s, then the system will come into steady state, at that time there is no overstrike. Figure 232.6 describes the situation of the torque tracking. And it is concluded that the situation of the torque tracing is very well without obvious deviation.

(2) When the given velocity is 1,000 rpm, we will add 15 Nm torque to the system and the results are shown at Figs. 232.7, 232.8, 232.9.

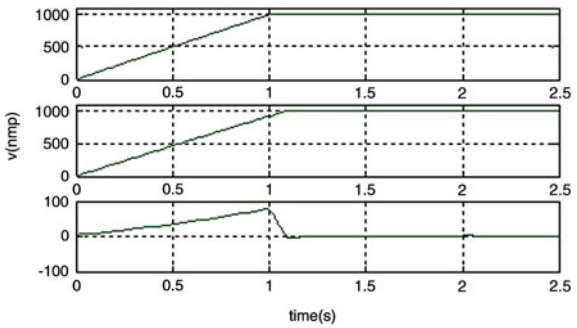


Fig. 232.8 The curves of the given velocity, actual velocity and the velocity error

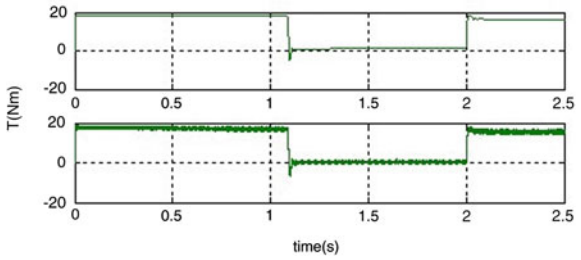


Fig. 232.9 The curves of the given torque and the actual torque

It can be concluded from Fig. 232.7 that the curve is very close to the curves of the ideal back-EMF and the current, which makes that the phase current was nearly equal to 0 and the examination of the over-zero of the back-EMF come true and there was no obvious after flow phenomenon. With the increase of the torque, the current and the back-EMF increased at the beginning. When the torque reached the value of the given torque, the current and the back-EMF would not change. But the torque suddenly changed after 2 s, and the current and the back-EMF increased with the change and came into another steady state after 0.2 s.

Figure 232.8 shows that the velocity error increased linearly approximately with the increase of the velocity after the system started. It reached the maximum value, that is 80 rmp, after 1 s and then it fell rapidly. With the increase of the velocity the transition progress was short with a little overstrike, and the system came into the steady state after 0.2 s. The system was added a sudden torque about 15 Nm, and the velocity decreased a little and it became steady after 0.2 s. It can be concluded from Fig. 232.9 that the actual output torque could do well track the given torque.

Figure 232.10 shows the curve is similar to the curves of the ideal back-EMF and the current, which makes that the phase current was nearly equal to 0 and the examination of the over-zero of the back-EMF come true and there was no obvious after flow phenomenon. With the increase of the torque, the current and the back-EMF increased at the beginning. After 1.2 s when the torque reached the value of

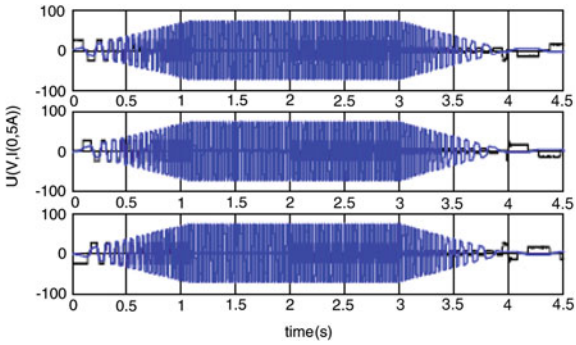


Fig. 232.10 A, B, C the current of three phase (*black*) of the back-EMF (*blue*)

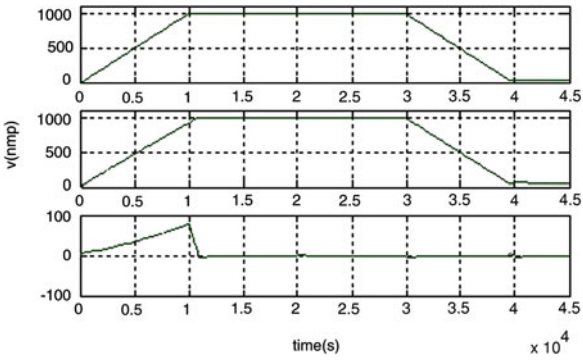


Fig. 232.11 The curves of the given velocity, and the curve actual velocity and the velocity error

the given torque, the current and the back-EMF would not change. But the torque suddenly changed after 2 s, and the current and the back-EMF increased with the change and came into another steady state after 0.2 s. When the time was 3 s, the given velocity decreased, and the current and the back-EMF fell rapidly and came into steady state after a short transition progress. But the torque reversed direction suddenly when the time was 3.8 s, and the current reversed quickly as well and became steady about 0.4 s later.

Figure 232.11 shows the velocity error increased with the increase of the velocity after the system started. The velocity error reached the maximum value 1 s after started, and then it decreased and it became steady after 1.2 s. The velocity fell resulted from that plus torque suddenly changed, but the magnitude of the decrease was small.

When the time was 3 s, the velocity instruction changed from 50 to 1,000 rpm. There was some velocity error appearing at the moment of the change, but the system could track the given torque very well 0.2 s later. The velocity decreased a little immediately as the torque changed suddenly when the time was 3.8 s and it

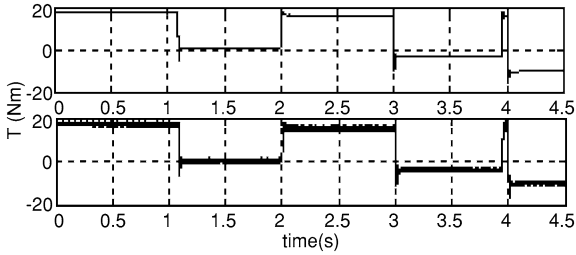


Fig. 232.12 Torque given and actual torque curve

became steady 0.3 s later. Figure 232.12 shows that the actual out torque was able to track the given torque very well.

232.7 Conclusion

This paper put forward a method on the back-EMF zero detection algorithm based on the Kalman filter of the brushless DC motor. This algorithm considers the back-EMF as unknown state variables to make estimations, the results show that it can efficiently estimate the value of the back-EMF of each phase and reverse the direction accurately. The details are displayed as follows:

It can leave out the work to detect the peripheral circuit unnecessarily through this algorithm.
It can conduct the accurate reverse at the steady state and the transient state through this algorithm.
It is able to control the velocity of the motor precisely by this algorithm.
It is allowed to control the position of the brushless DC motor with this algorithm, which is relatively simple.
It is not sensitive to the noise signal through Kalman filter by this algorithm, so it can be used in the industrial controlling occasions.

References

1. Hailong S, Ming Y, Yu F et al (2002) Review of sensorless control of brushless DC motor. Electr Mach Control 6(3):208–212
2. Yanqing Z, Huajie Y, Changqing Y (2007) Position detection for position sensorless brushless DC motor. Electr Mach Control Appl 34(7):1–5
3. Kim TH, Ehasani M (2004) Sensorless control of the BLDC motors from near-zero to high speeds. IEEE Trans Power Electron 19(6):1635–1645
4. Jianwen S, Nolan D (2002) A novel direct back EMF detection for sensorless brushless DC (BLDC) motor drives. In: applied power electronics conference and exposition, 17 annual IEEE, vol 1, pp 10–14, 33–37

5. Lei Z, Qu W, Lu H et al (2006) A novel sensorless control system of brushless DC motors. Trans China Electro Techn Soc 21(10):26–30
6. Hongwei W, Hui L (2005) DSP control of brushless DC motor without position sensor. Power Electron 39(6):102–104
7. Jianwen S (2006) An improved microcontroller-based sensorless brushless DC (BLDC) motor driver for automotive applications. IEEE Trans Ind Appl 42(5):1216–1221
8. Li Z, Xia C, Chen W (2010) A position sensorless control strategy for BLDCM based on line back-EMF. Trans China Electrotechn Soc 25(7):38–44, 48–53
9. Quan J, Chao B (2005) A new phase-delay-free method to detect back EMF zero-crossing points for sensorless control of spindle motors. IEEE Trans Magn 41(7):2287–2294
10. Wang H, Liu H, Zhang Y, Liu P (2009) Sensorless BLDC drive allowing for neutral point voltage. Trans China Electrotechn Soc 24(7):46–51
11. Jing X (1973) Kalman filter and its applications. China National Defence Industry Press, Beijing
12. Kim T-S et al (2008) A new approach to sensorless control method for brushless DC motors. Int J Control Autom Syst 6(4):477–487
13. Gunabalan R et al (2009) Sensorless control of induction motor with extended Kalman filter on TMS320F2812 processor. Int J Recent Trends Eng 2(5):14–19
14. Lin H, Yan W, Lin Y et al (2010) Applying dual extended Kalman filter (DEKF) theory to sensorless control of brushless DC motor (BLDCM). J North west Polytechn Univ 28(2):197–201
15. Jang GH, Kim MG (2006) Optimal commutation of a BLDC motor by utilizing the symmetric terminal voltage. IEEE Trans Magn 42(10):3473–3475
16. Zou J, Jiang S, Zhang H (2009) A novel method of detecting for rotor position of a sensorless brushless DC motor. Trans China Electrotechn Soc 24(4):48–53
17. Wu X, Zhou B, Song F (2009) A new control method to correct position phase for sensorless brushless DC motor. Trans China Electrotechn Soc 24(4):54–59
18. Yen SL, Fu SS (2004) Novel back-EMF detection technique of brushless DC motor drives for whole duty-ratio range control. The 30th annual conference of the IEEE industrial electronics society. Busan, Korea, vol 3, pp 2729–2732

Chapter 233
Study on Modeling and Simulation of Four-Phase Switched Reluctance Machine Based on Ansoft Maxwell2D

Renguo Xi, Shoujun Song and Weiguo Liu

Abstract According to the real structural parameters of the motor and its external circuit, the complete simulation model of a four-phase 8/6 poles switched reluctance motor (SRM) is built based on Maxwell2D of Ansoft Corporation. The basic performances and the starting process of SRM are analyzed based on this model. The simulation results include the distribution of magnetic field at various rotor positions, static electromagnetic characteristics and dynamic performance. These results agree well with the basic running performances, and can benefit for the design and optimization of SRM and its control system.

Keywords Switched reluctance motor (SRM) · Ansoft · 2-D magnetic field · Finite-element method (FEM)

233.1 Introduction

Switched reluctance machine (SRM) drive system consist of switched reluctance machine, power inverter, digital controller and position sensor [1, 2]. It has the advantages of both asynchronous and DC machine drive system [3]. Its market developing potential is great, and it has been widely used in electric vehicles, household appliances, general industry, aviation industry, servo system and so on.

The finite-element analysis (FEA) software named Maxwell2D is developed by Ansoft Corporation [4]. It is very suitable for the analysis of static and transient

R. Xi (✉) · S. Song · W. Liu
Northwestern Polytechnical University, 127 Youyi Xilu,
Xi'an 710072, Shaanxi, China
e-mail: sunnyway@nwpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_233,

electromagnetic fields. Furthermore, it has good ability for post-processing [5]. It is a convenient, efficient and practical software for electromagnetic analysis.

It is well known that SRM is a nonlinear, multivariable and strong coupling electromechanical system [6, 7]. It is very difficult to analyze the SRM system accurately by traditional analytical method. In this paper, Maxwell2D is used to analyze the basic operating characteristics and starting performances of the SRM system, and several simulation results are given. The simulation results are very important for the design and control of SRM system.

233.2 Simulation Model

SRM is a typical electromechanical integration product, and there is strong coupling among machine, power inverter and controller, so the simulation model consists of three parts accordingly.

233.2.1 Machine

The modeling process of SRM in Maxwell2D is as follows:

(1) Select the plane for solving;
(2) Determine the physical dimensions of the machine and draw the model;
(3) Determine the material properties for each part of the machine.

In this paper, a four-phase, 8/6-pole SRM is studied. The material of the stator and rotor is D23 steel, the shaft is defined as air, and the windings are made up by copper.

The physical parameters of the SRM are shown in Table 233.1.

According to the parameters in Table 233.1, the 2-D model of SRM can be drawn in Maxwell2D as shown in Fig. 233.1.

233.2.2 Power Inverter

In this paper, a four-phase asymmetric half-bridge inverter is adopted [8]. Figure 233.2 shows its model. Where S1–S8 are voltage controlled switches and W1–W8 are current controlled switches. W1–W8 are controlled by phase current IA–ID to implement current chopping control (CCC). By CCC method, the large peak value of current and flux linkage can be avoided and constant torque mechanical characteristics can be obtained.

Table 233.1 Physical parameters of the SRM

Parameters	Values
Stator outside diameter	210 mm
Rotor outside diameter	113 mm
Core length	135.5 mm
Stator pole arc	21°
Stator yoke height	13.5 mm
Rotor pole arc	23°
Rotor yoke height	15.75 mm
Stator slot depth	34.6 mm
Shaft diameter	50 mm
Air gap length	0.4 mm
Rating power	1.8 KW
Rating voltage	270 VDC
Rating speed	1250 r/min

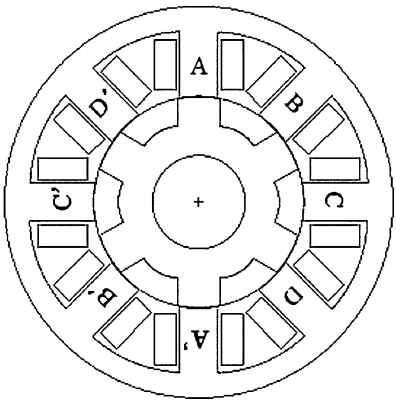


Fig. 233.1 2D model of the SRM in Maxwell2D

233.2.3 Driving Signal Generating Circuit

During the operation of the SRM system [9], the correct switching signal for each phase can be generated based on position information of the rotor. In Maxwell2D, rotor position information can be obtained by position variable P. Figure 233.3 shows the simple circuit in Maxwell2D to generate driving signal.

Voltage controlled switches S1–S8 in power inverter are controlled by voltage sources VA−VD in driving signal generating circuit. VA–VD are functions of position variable P. Figure 233.4 shows the pulse voltage signal obtained by the circuit in Fig. 233.3.

Before the simulation with the model of the SRM system, following steps should be completed:

(1) Determine the boundary conditions of finite-element calculation and complete mesh;

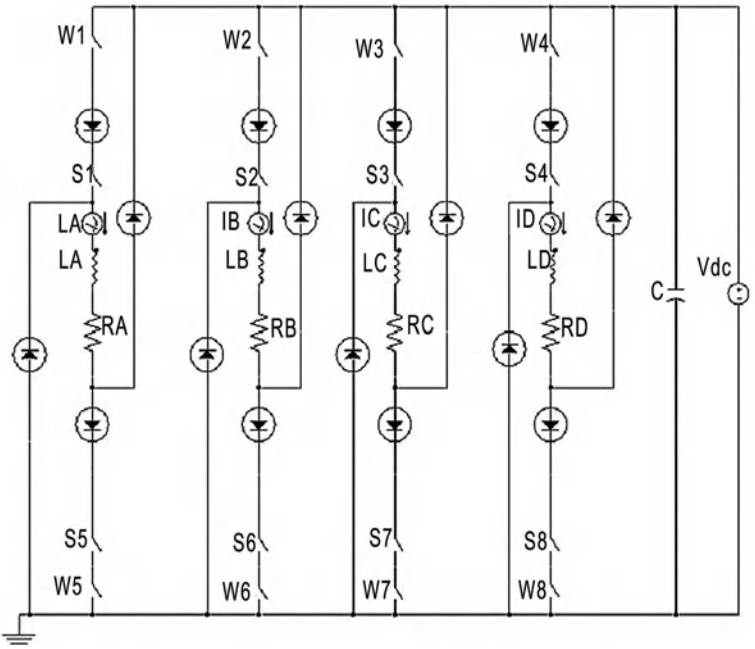
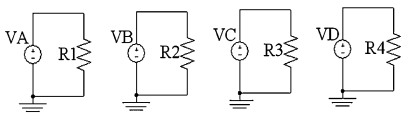


Fig. 233.2 Model of the four-phase asymmetric half-bridge inverter

Fig. 233.3 Model of driving signal generating circuit



(2) Determine whether various kinds of loss or not during the calculation;
(3) Determine dynamic parameters include moving boundary, applied load, step size and so on.

Figure 233.5 shows the result of the finite-element mesh.

233.3 Simulation Results

In this paper, the SRM operates as a motor. The rotor position shown in Fig. 233.1 is defined as the initial position where the mechanical angle is 0°. In the simulation, the turn-on sequence is C–D–A–B–C. The conduction interval of each phase is 15° and the conduction period is 60°.

The simulation is conducted and the magnetic field distribution, static electromagnetic parameters and dynamic performances are obtained.

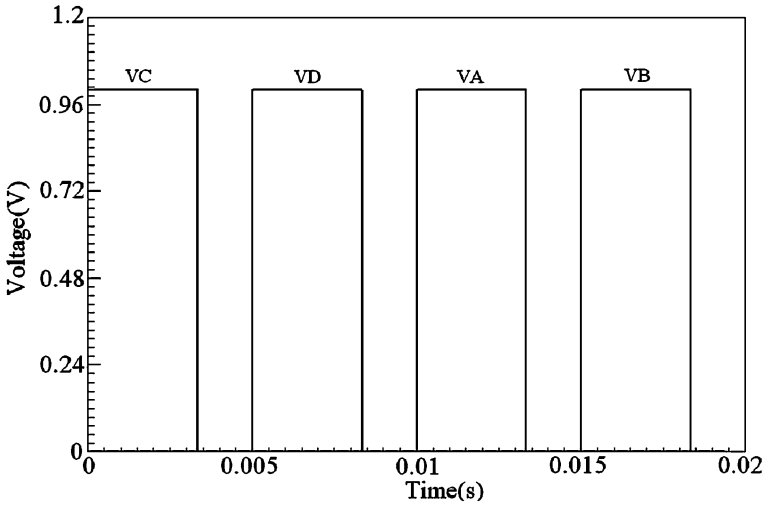
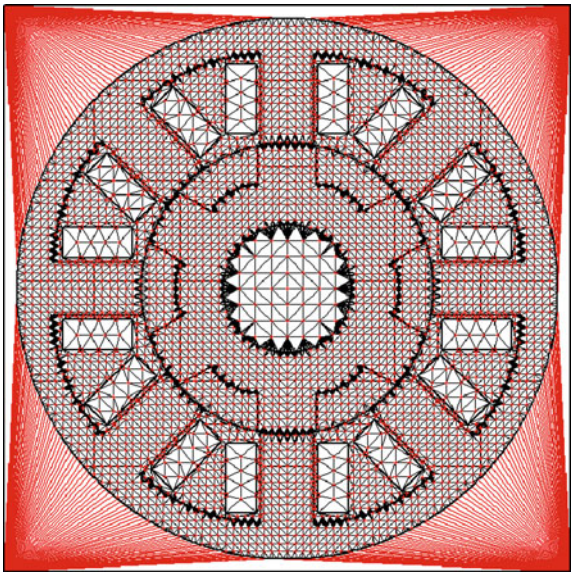


Fig. 233.4 Pulse voltage signal

Fig. 233.5 Finite-element mesh of the SRM



233.3.1 Magnetic Field Distribution

Figure 233.6 shows the distribution of the magnetic field in SRM. During the simulation, different rotor position has different distribution. In this paper, only one situation is shown.

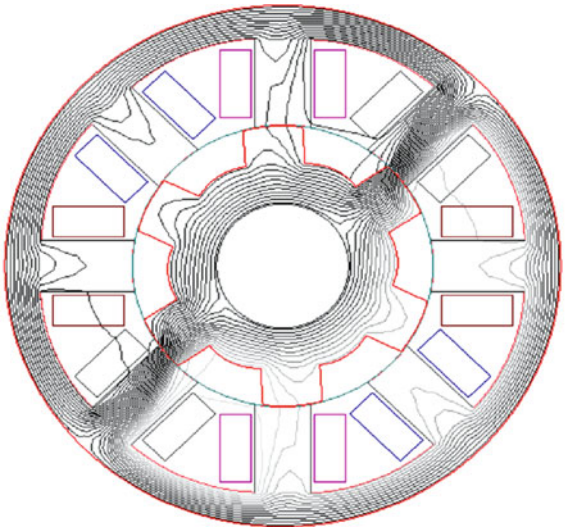


Fig. 233.6 Distribution of the magnetic field in SRM

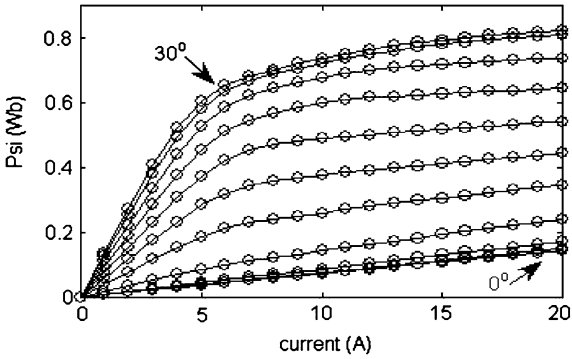


Fig. 233.7 Flux characteristics

233.3.2 Static Electromagnetic Parameters

Based on the static magnetic field module in Maxwell2D, the static electromagnetic characteristics can be obtained by parametric analysis. Figure 233.7 shows the flux characteristics and Fig. 233.8 shows the static torque characteristics.

233.3.3 Dynamic Performances

In this paper, two control methods named angular position control (APC) and current chopping control (CCC) are analyzed by simulation respectively. The waveforms of phase current, total torque, flux linkage and back-emf are obtained.

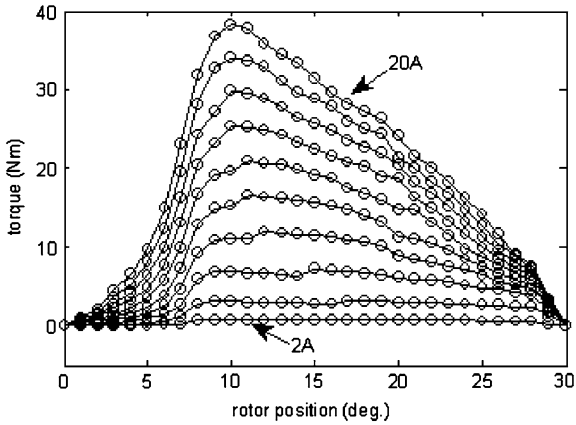


Fig. 233.8 Static torque characteristics

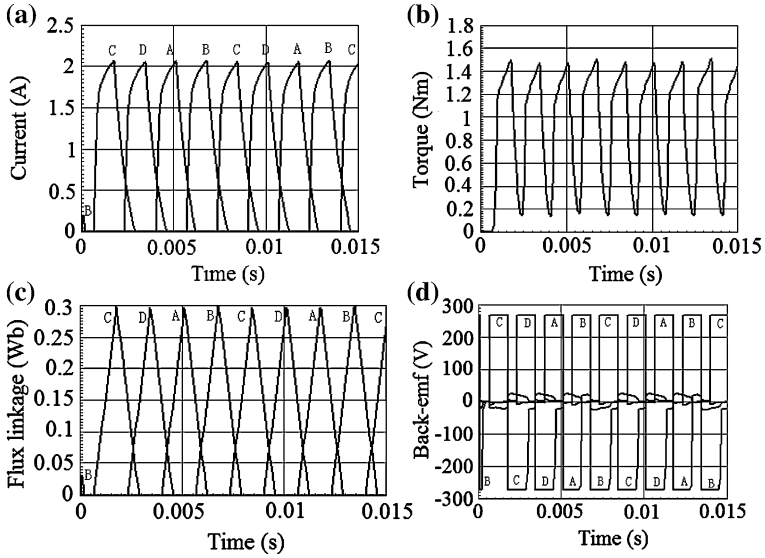


Fig. 233.9 Simulation results with APC method **a** Phase current **b** Total torque **c** Flux linkage **d** Back-emf

APC: Turn-on angle θ_{on} is 6°, turn-off angle θ_{off} is 16° and speed is 1500 r/min. Figure 233.9 shows the waveforms of phase current, total torque, flux linkage and back-emf obtained by APC.

CCC: Turn-on angle θ_{on} is 6°, turn-off angle θ_{off} is 18.5°, speed is 500 r/min, reference current is 4.5A and hysteresis width is 0.2A. Figure 233.10 shows the waveforms of phase current, total torque, flux linkage and back-emf obtained by CCC.

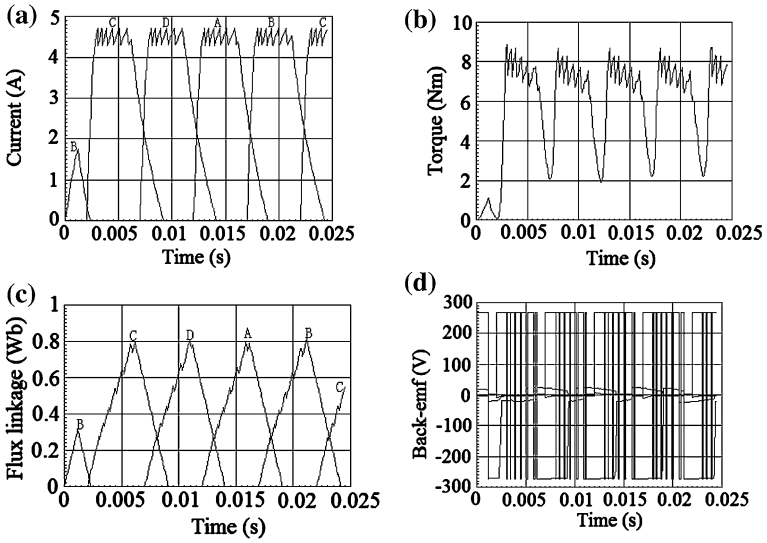


Fig. 233.10 Simulation results with CCC method **a** Phase current **b** Total torque **c** Flux linkage **d** Back-emf

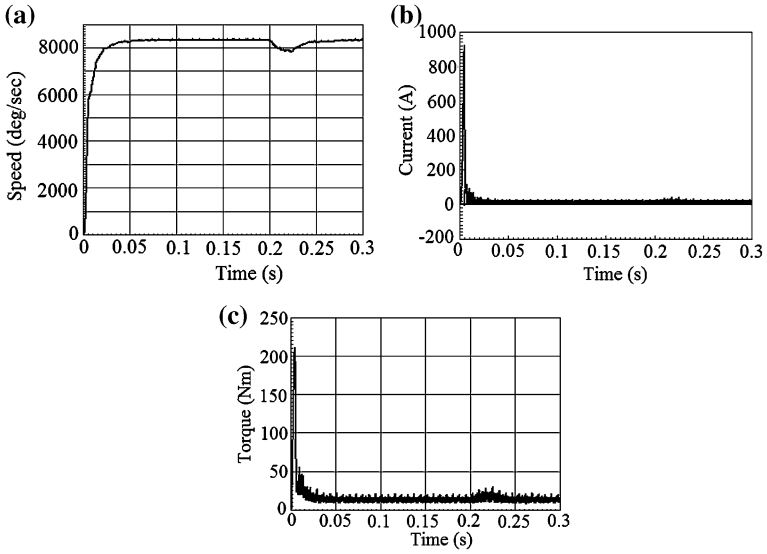


Fig. 233.11 Simulation results during starting **a** Speed **b** Current**c** Torque

233.3.4 Starting Performances

In this paper the starting process of the SRM system is simulated to verify its starting and adjusting performances. Figure 233.11 shows the waveforms of speed, current and torque during starting. Turn-on angle is 6° and turn-off angle is 21°. The system is started with idle load. 5 Nm load is added at 0.2 s and removed at 0.225 s. In Fig. 233.11, it can be found that the system researches its steady speed at about 0.05 s and the speed is adjusted quickly during the change of the load. The system has good starting and adjusting performances.

233.4 Conclusion

In this paper, the finite-element model of a four-phase, 8/6-pole SRM is built in Maxwell2D and the external driving circuit is considered as well. The simulation results reflect the distribution of the magnetic field in SRM accurately and provide a basis for the optimization design of the structural parameters. The static electromagnetic parameters obtained in this paper are very important to build the nonlinear model of SRM in other software, such as MATLAB. The results of dynamic performances simulation are very important for the optimization of turn-on and -off angles. Finally, the starting and adjusting performances of the SRM system are verified by simulation of starting process, and the simulation results are very useful for the study on optimization design of the SRM, decrease of the torque ripple and increase of the starting torque.

Acknowledgments This work is supported by the research plan of Natural Science Foundation of Shaanxi Province (no. 2011GQ7001).

References

1. HongHua W (1997) Switched reluctance machine speed regulating system. Mechanical Industrial Press, Beijing
2. Jianhua W (2001) Switched reluctance machine design and application. Mechanical Industrial Press, Beijing
3. Huijun Z, Wen D, Zhenmin Y (2005) Simulation and analysis of switched reluctance motor. Micromachine 6:10–12
4. Wengang J, Baojin C, Qiang L, Li Q (2005) Simulation and analysis of switched reluctance motor applied in the automobile based on ansoft. Agric Mech Res 5:271–273
5. Junyong L, Deliang L, Xiangyang F (2004) Simulation and analysis of linear PM brushless DC motor based on ANSOFT. Middle Minitype Mot 1:18–20
6. Genhua Z, Deliang L, Junyong L (2003) Performance analysis of multi-phase and multi-pole motor based on Ansoft. Small and special electric machine 6:15–16

7. Stanton S (2005) Analysis of a dual channel switched reluctance motor/generator. Ansoft corporation
8. Huijun Z, Wen D, Zhenmin Y (2007) Simulation and analysis of switched reluctance motor based on Ansoft. Small Special Electr Mac 3:11–12
9. Guiling Y, Wen D (2010) Simulation and experiment for single phase switched reluctance generator based on Ansoft. Electr Mac Cont Appl 7:11–15

Chapter 234
Key Performance Analysis of Six-Phase Induction Machine Driving by Trapezoidal Phase Current Waveform with Magnet 6.27 Software

Beibei Xie, Xing Yuan and Yongle Ai

Abstract This paper researches key performance of the six-phase induction machine using Magnet 6.27 software driving by six-phase trapezoidal phase current. First, the machine model has been drawn with AutoCAD and imported into Magnet which adding the related magnetic material and setting machine parameters. The stator current source has been configured trapezoidal waveform and applied to six-phase induction machine. Second, the performance of air-gap magnetic field and electromagnetic torque has been analyzed with the Magnet 6.27. The analysis results indicate that the trapezoidal phase current realizes the separate control of field and torque current. Finally, the machine simulation model has been added motion unit to realize 2D dynamic simulation under the condition of no-load and load starting. The simulation results indicate that the starting speed of six-phase induction machine driving by trapezoidal phase current waveform is fast and the operation is smooth.

Keywords Magnet · Finite element analysis · 2D transient simulation · 2D dynamic simulation

B. Xie (✉) · X. Yuan · Y. Ai
Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: xiebeibei@hpu.edu.cn

X. Yuan
e-mail: yuanxingbluesky@163.com

Y. Ai
e-mail: aiyongle@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_234,

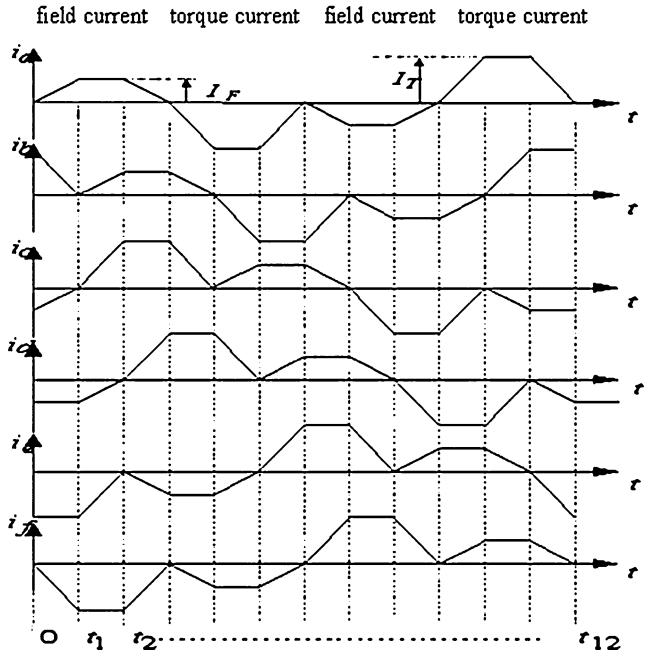


Fig. 234.1 Six-phase trapezoid current waveform

234.1 Introduction

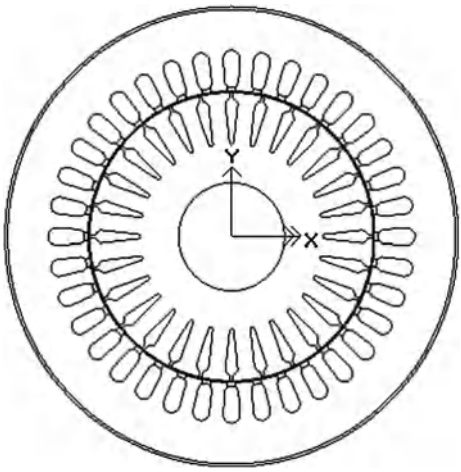
With the rapid development of power electronic technology and control theory, power converters can eliminate the limit of three-phase sine power supply. Now, multiphase machine can be driven by multiphase power inverter [1]. Compared to the three-phase machine, the multiphase machine has some advantages. But its control system is much more complicated [2]. This paper adopts the control strategy of trapezoidal phase current proposed in [3] and uses Magnet software to study the key performance of six-phase machine. The six-phase trapezoidal current supplied to the six-phase machine makes the air-gap magnetic field similar to the DC machine. Then the six-phase trapezoidal current can complete independent control of field and torque current. Consequently, it omits the Park and its inverse transformation in traditional vector control. It can truly imitate the controlling of DC machine. The phase current of six-phase trapezoidal waveform is shown in Fig. 234.1 [3].

This paper uses Magnet 6.27 software to set the trapezoid phase current waveform and builds 2D simulation model of six-phase induction machine based on the proposed current. There is the simulation of 2D transient magnetic field for the electromagnetic field distribution and starting performance of six-phase induction machine. The key performance of electromagnetic torque and speed has been analyzed under the condition of machine's no-load and load.

Table 234.1 The machine parameters

Machine phase number	Machine pole pairs	Stator slot number	Rotor slot number	Rated speed (rad/min)	Stator diameter (mm)	Rotor diameter (mm)	Core length (mm)
6	1	36	28	3,000	161	102	135

Fig. 234.2 Machine model



234.2 The Finite Element Analysis of Six-Phase Induction Machine

The six-phase induction machine parameters used in this paper are shown in Table 234.1 [4]. The machine model has been built with AutoCAD software, and then this paper studies the key performance of six-phase induction machine using magnet 6.27 software driving by six-phase trapezoidal phase current.

234.2.1 The Construction of Six-Phase Machine's 2D Simulation Model

The model of six-phase induction machine has been built using AutoCAD software as shown in Fig. 234.2.

The machine model has been imported into Magnet software adding the appropriate magnetic materials. The six-phase trapezoidal current waveform is shown in Fig. 234.1 which supply stator winding. The field current's amplitude is equal to 3.5 A, and the amplitude of torque current is variable. The torque current can control machine's steady-state electromagnetic torque through adjusting torque current's amplitude. Then current source has been set six-phase trapezoidal

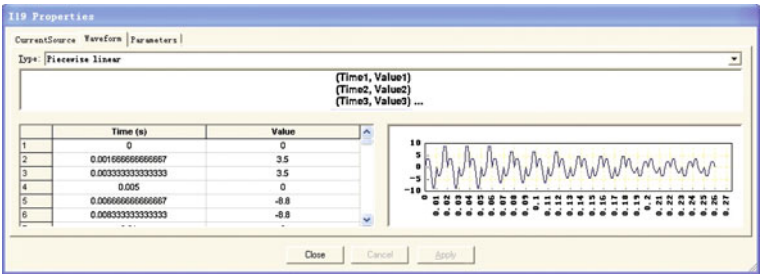


Fig. 234.3 A-phase trapezoidal current waveform

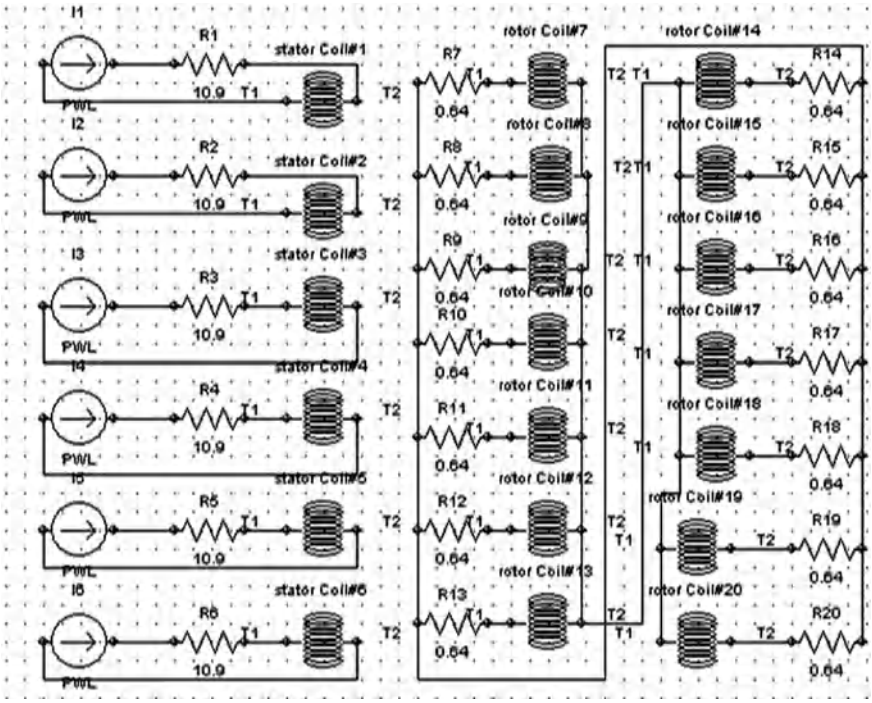


Fig. 234.4 Stator and rotor circuit

waveform applied to six-phase machine in Magnet software. The A-phase trapezoidal current waveform is shown in Fig. 234.3.

The parameters of the stator and rotor winding have been configured in Magnet software. The windings number of stator is 83; the resistance of stator is 10.9 Ω; the windings number of rotor is 28; the resistance of rotor is 0.64 Ω in the study. Finally, the connection of stator circuit has been completed. The stator and rotor circuit in the Magnet is shown in Fig. 234.4.

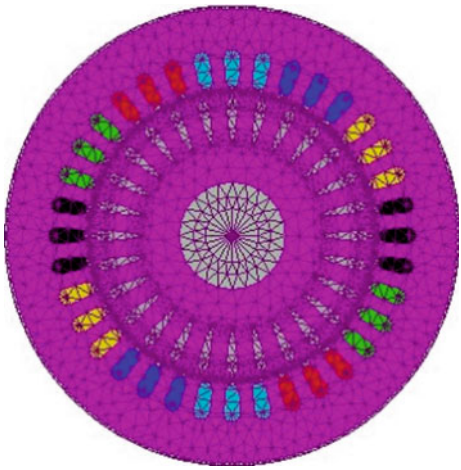


Fig. 234.5 Finite element grid figure of the six-phase induction machine

The finite element simulation model of six-phase induction machine has been built after adding magnetic materials and setting the parameters of the circuit. In the simulation model the units are divided into small region. The calculation of Maxwell's Equations costs much time and the analysis result of simulation is accurate [5]. The finite element grid figure of machine model has been divided into three as shown in Fig. 234.5.

This current waveform is derived considering the DC machine control principle with separate field and armature currents, as well as the advantage of a rectangular flux density waveform in the air gap of the induction machine [6].

234.2.2 The Simulation Analysis of Electromagnetic Field Simulation Analysis

The transient 2D simulation of electromagnetic field has been operated using the Magnet. The distribution of electromagnetic field is shown in Fig. 234.6.

From Fig. 234.6 it can be seen, the distribution of magnetic field is regular. It is useful to operate under the condition of no-load and load.

234.2.3 The 2D Dynamic Analysis of Machine Simulation Model

In order to determine whether the machine torque and speed can reach to the predetermined value, it is necessary to add motion unit to the original machine simulation model, and proceed with 2D dynamic simulation. The change of electromagnetic torque with time is shown in Figs. 234.7 and 234.8 which the run

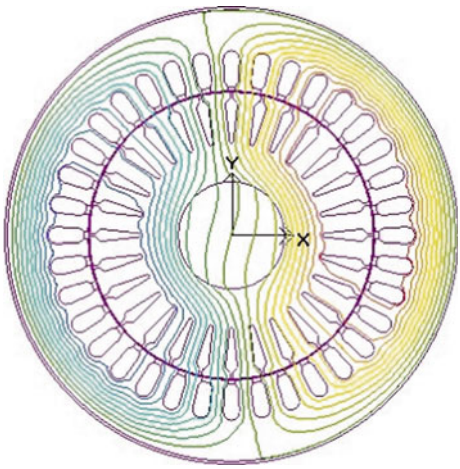


Fig. 234.6 Distribution of electromagnetic field in six-phase induction machine

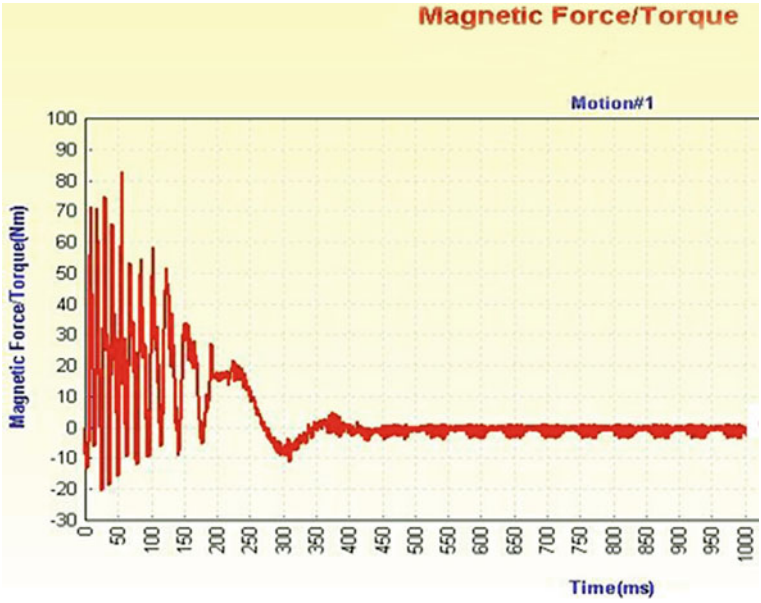


Fig. 234.7 Electromagnetic torque curves when the machine no-load operation time is 1 s

time of no-load starting is 1 s and that of load starting is 0.5 s after no-load starting.

From Fig. 234.7, it can be seen, when the machine starts up under the condition of no-load, 0.3 s later, the amplitude of electromagnetic torque is smooth and tends to 0. The electromagnetic torque curves are no obvious torque ripple from Fig. 234.7. The result of simulation shows that the electromagnetic torque

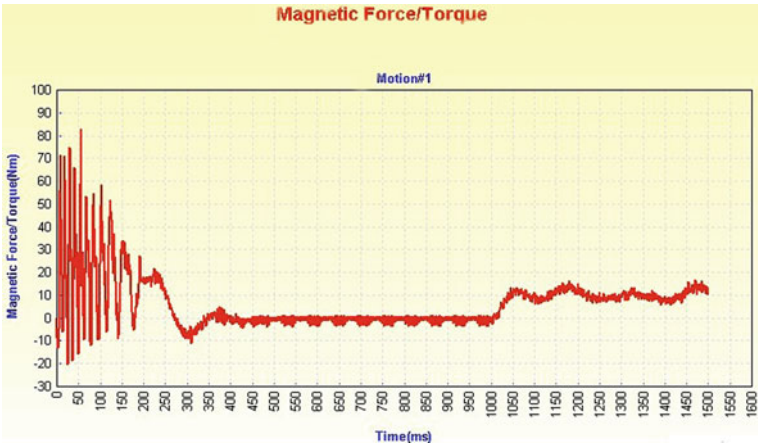


Fig. 234.8 Electromagnetic torque curves of adding 10 N m load after no-load running 1 s

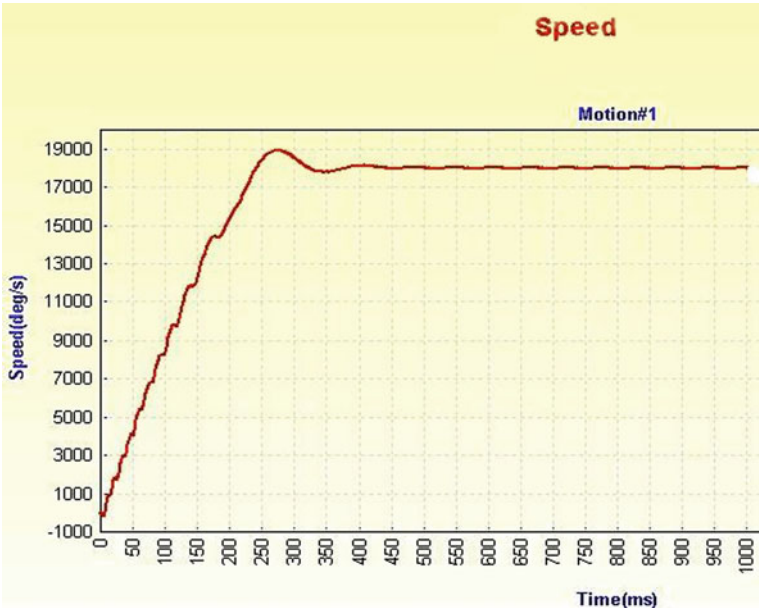


Fig. 234.9 Speed curves when the machine no-load running time is 1 s

response quickly, then the needed time of stable operation is short on no-load. The electromagnetic torque waveform is shown in Fig. 234.8 that the machine adds 10 N m load after 1 s. The amplitude of electromagnetic torque trends to 10 N in 1.1 s. The results show that the no-load starting is fast and stable, then load starting has fast response and electromagnetic torque's ripple is small.

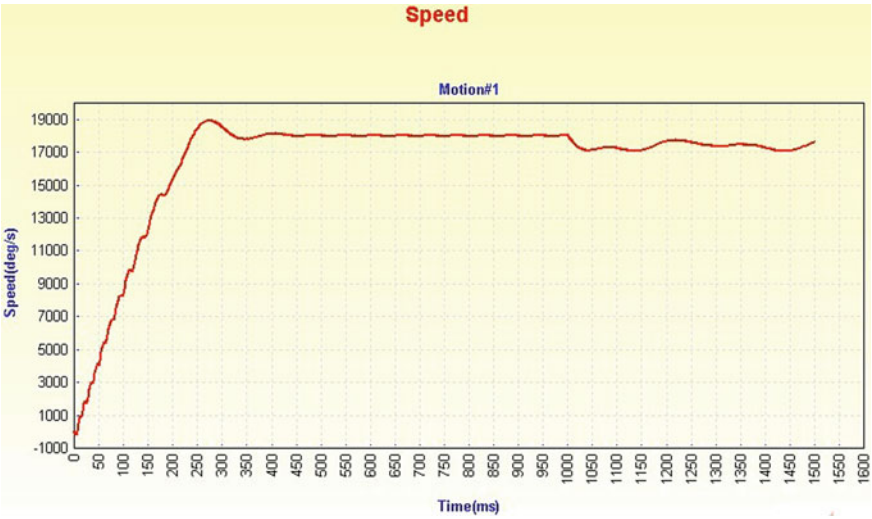


Fig. 234.10 Speed curve of adding 10 N m load after the run time of the no-load is 1 s

The angular velocity ω is shown in Fig. 234.9 under the condition of machine no-load starting. The variation of parameters must be considered when the machine starts up. Because it simulates the starting process, the initial velocity is 0. The formula of machine speed is $n = \omega \cdot 60/(2\pi)$. From Fig. 234.9, it is found that the speed of the machine reach the rated speed at about 3,000 rad/min after 0.3 s. The speed curve of adding 10 N m load is shown in Fig. 234.10 after 1 s of no-load. The operation of machine becomes stable when its speed reduces about 1% after adding load for 0.05 s. The speed has small ripple and responses quickly. It can be seen the speed changes little when the machine's load is regulated from 0 to 10. The simulation results show that no-load and load starting have fast response, and the electromagnetic torque and speed have no obvious fluctuations. It illustrates the machine analyzed is of good performance.

234.3 Conclusion

In this paper the simulation model of six-phase induction machine has been researched with finite element method of Magnet based on trapezoidal phase current waveform. In the study the field and torque current of phase current have been decoupled due to trapezoidal phase current. In other words, it simulates the control of DC machine. The paper uses the Magnet 6.27 software to simulate the 2D transient magnetic field. The original model adds the motion unit to simulate and analyze the performance of the machine starting, which simulate the changing processes of electromagnetic torque and the speed under the condition of no-load and load.

Acknowledgments The work was supported by the National Research Foundation(NFR) in South African and Foundation of Henan Polytechnic University Doctor(Grant No:648207).

References

1. Li S, Wang M, Liu H (2005) Six-phase induction motor based on DSP vector control system design. Big Mot Technol 3:23–26
2. Hou L, Su Y, Chenlin (2001) The development present situation and application prospects of multiphase induction motor in speed control system. Micro-motor 5:42–44
3. Ai Y, Ai B, Niu L (2009) The electromagnetic torque of six-phase induction motor on the driving of six-phase trapezoidal current waveform. Big Mot Technol 1:10–13
4. Ai Y, Kamper MJ, Wang Y (2009) The novel control strategy research of six-phase ac induction motor, vol 8. China Power Press, Beijing, pp 21–25
5. Mark F (1996) Mesh generation. Edinburgh Parallel Computing Centre the University of Edinburgh, Edinburgh
6. Ai Y, Kamper MJ, Wang Y (2004) Torque performance investigation of double three-phase motor using special current waveform. In: Proceedings of 4th conference 2004 international electronic motion control (IPEMC2004). Xi'an, pp 1673–1678

Chapter 235
Design and Finite Element Analysis of Brushless Doubly Fed Machine with Radially Laminated Magnetic Reluctance Rotor

Feng-Ge Zhang, Qun Yang and Xiu-Ping Wang

Abstract For the problem of the weaker coupling capability between the two sets stator winding of brushless doubly fed machine (BDFM), the lower power density and poor performance. This paper researches a novel 8 + 4 pole radially laminated magnetic reluctance rotor BDFM, the finite element method is adopted to calculate magnetic filed. The best project can be obtained through the determination of pole-arc factor, conducting magnetic and magnetic obstacle layer width radio and the width and number of conducting magnetic layer. The results show: radially laminated magnetic reluction rotor can significantly improve the coupling capability.

Keywords BDFM · Radially laminated magnetic reluctance rotor · Coupling capability · Finite element method

235.1 Introduction

Brushless doubly fed machine (BDFM) is formed by two self-cascaded induction machine [1, 2]. It is a new type of excellent performance AC motor. The rotor of BDFM cancels brush and slip ring, so it has good performance and stable operation

F.-G. Zhang (✉) · Q. Yang · X.-P. Wang
Education Ministry and Province Key Laboratory of Special Motor and High Voltage Apparatus, Shenyang University of Technology, Shenyang 100870, China
e-mail: zhangfg@sut.edu.cn

Q. Yang
e-mail: yq1154944609@163.com

X.-P. Wang
e-mail: wangxpmail@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_235,

characteristics. In stator, there are two sets of different pole windings called power windings and control windings, respectively, the pole pairs are P_p and P_q. Rotor can use cage or reluctance structure [3]. Through the rotor modulate the rotating magnetic field of two sets stator windings current to realize mechanical and electrical energy conversion.

As the structure and principle of BDFM is different from common AC motor, there are many questions on BDFM, which are shown as [4–6]:

(a) The weaker coupling capability of traditional rotor, resulting in the lower power density and poor performance.

(b) The control windings is connected to frequency converter and much harmonic magnetic field exists in BDFM which will result in polluted power grid.

(c) The efficiency of BDFM is low that is because the existence of much harmonic magnetic field and the iron loss of rotor.

In view of above problem, in this paper, how to raise coupling capability of reluctance rotor is a starting point, studying radially laminated magnetic reluctance rotor BDFM of 8 + 4 pole.

235.2 Magnetic Field Modulation Principle of Reluctance Rotor

For reluctance rotor, the assumption is that the magnetic circuit is linear. If the number of tooth is P_r. When the motor is rotating, the expression of air gap late permeance relative to stator coordinate is

$$\lambda_g(\varphi, t) = \lambda_0 + \sum_{\upsilon=1}^{\infty} \lambda_\upsilon \cos \upsilon P_r(\varphi - \theta_{r0} - \omega_{\mathrm{rm}} t) \tag{235.1}$$

where, λ_0 is the average component of air gap permeance, λ_υ is the amplitude of order υ harmonic. θ_{r0} is the angle between rotor and axis of power windings magnetomotive force. ω_{rm} is the mechanical angular velocity of rotor.

As the maximum component of permeance harmonic is order 1 harmonic, the higher order permeance harmonic can be neglected, so the expression of air gap late permeance is:

$$\lambda_g(\varphi, t) = \lambda_0 + \lambda_1 \cos[P_r(\varphi - \theta_{r0} - \omega_{\mathrm{rm}} t)] \tag{235.2}$$

So, the flux density of gap can be expressed as:

$$B(\varphi, t) = \left[f_p(\varphi, t) + f_q(\varphi, t)\right] \lambda_g(\varphi, t) \tag{235.3}$$

where, $f_p(\varphi, t)$, $f_q(\varphi, t)$ are respectively the expression of power windings and control windings magnetomotive force.

$$B(\phi, t) = B_{p0} + B_{q0} + B_{p1(+)} + B_{p1(-)} + B_{q1(+)} B_{q1(-)} \tag{235.4}$$

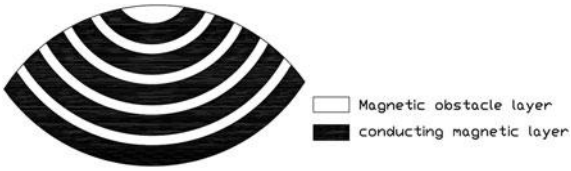


Fig. 235.1 Structure of radially laminated magnetic reluctance rotor

Where, $B_{p0} = F_{\mathrm{pm}}\lambda_0 \cos\left(P_p\varphi - \omega_p t\right)$

$$B_{q0} = F_{\mathrm{qm}}\lambda_0 \cos\left(P_q\varphi - \alpha - \omega_q t\right)$$

$$B_{p1(+)} = F_{\mathrm{pm}}\left(\frac{\lambda_1}{2}\right)\cos\left[\left(P_p + P_r\right)\varphi - \left(\omega_p + P_r\omega_{\mathrm{rm}}\right)t - P_r\theta_{r0}\right]$$

$$B_{p1(-)} = F_{\mathrm{pm}}\left(\frac{\lambda_1}{2}\right)\cos\left[\left(P_p - P_r\right)\varphi - \left(\omega_p - P_r\omega_{\mathrm{rm}}\right)t - P_r\theta_{r0}\right]$$

$$B_{q1(+)} = F_{\mathrm{pm}}\left(\frac{\lambda_1}{2}\right)\cos\left[\left(P_q + P_r\right)\varphi - \left(\omega_q + P_r\omega_{\mathrm{rm}}\right)t - \alpha - P_r\theta_{r0}\right]$$

$$B_{p1(-)} = F_{\mathrm{pm}}\left(\frac{\lambda_1}{2}\right)\cos\left[\left(P_q - P_r\right)\varphi - \left(\omega_q - P_r\omega_{\mathrm{rm}}\right)t - \alpha - P_r\theta_{r0}\right] \quad (235.5)$$

It can be seen there are six kinds of component in air gap magnetic. Only $B_{p1(-)}$ and $B_{q1(-)}$ have the same frequency with control windings and power windings, respectively which is useful for BDFM [7].

235.3 The Design of Radial Laminated Magnetic Reluctance Rotor

Radially laminated magnetic reluctance rotor is characterized by better coupling capability, more standardized magnetic circuit [8]. And its structure is shown in Fig. 235.1.

235.3.1 The Determination of Pole-Arc Factor

Pole-arc factor affects the magnetic circuit and the reluctance between the adjacent teeth. The effect on the coupling capability of rotor is shown in Fig. 235.2. With the increase of pole-arc factor, the coupling capability of rotor decreases. But little pole-arc can result in magnetic saturation, so the pole-arc factor in this paper is 0.7.

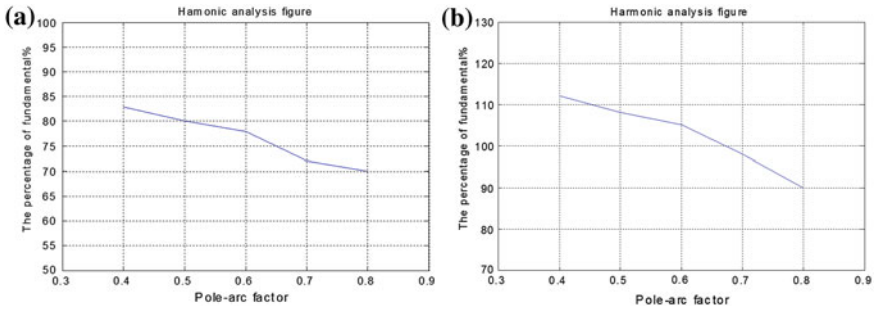


Fig. 235.2 The effect of pole-arc factor on the coupling capability

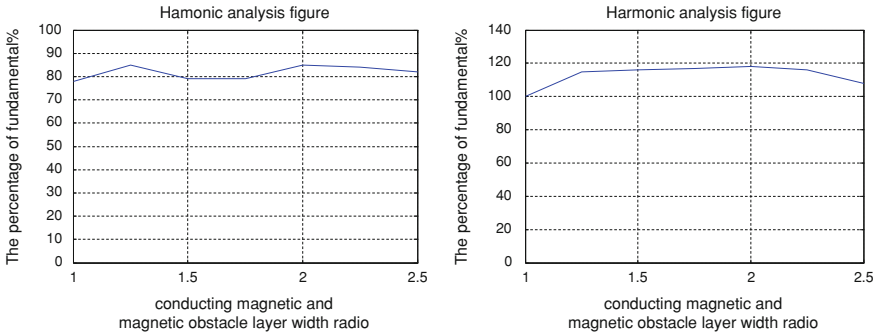


Fig. 235.3 The effect of conducting magnetic and magnetic obstacle layer width radio on the coupling capability

235.3.2 The Determination of Conducting Magnetic and Magnetic Obstacle Layer Width Radio

The magnetic obstacle layer can hinder the harmonic magnetic field, so the reasonable selection of conducting magnetic and magnetic obstacle layer width radio is necessary. The influence of width radio on the coupling capability of rotor is shown in Fig. 235.3. According to it, the width radio in this paper is 2:1.

235.3.3 The Determination of Conducting Magnetic Layer Width

In order to get the best coupling capability, optimization determination must be performed on the conducting magnetic layer width. The influence of width on the

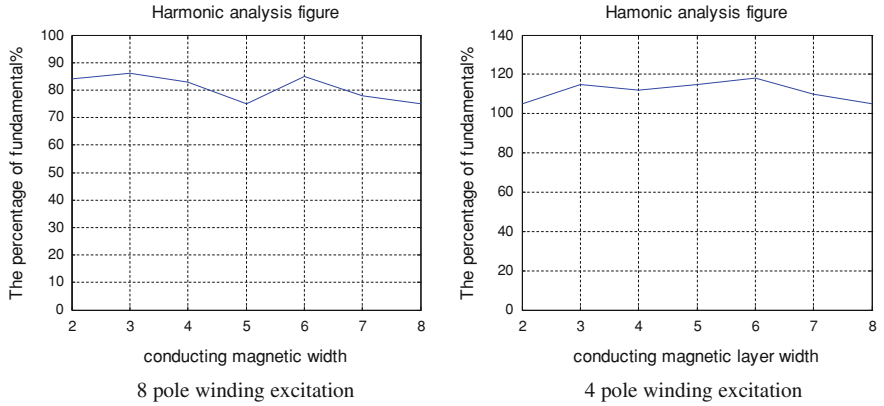


Fig. 235.4 The effect of conducting magnetic width on the coupling capability

Table 235.1 Design of conducting magnetic number

Percentage (%)	8 pole winding excitation (%)	4 pole winding excitation (%)
1 conducting magnetic layer, the percentage of fundamental	75	105
2 conducting magnetic layer, the percentage of fundamental	72	108
3 conducting magnetic layer, the percentage of fundamental	80	115
4 conducting magnetic layer, the percentage of fundamental	85	120
5 conducting magnetic layer, the percentage of fundamental	85	121

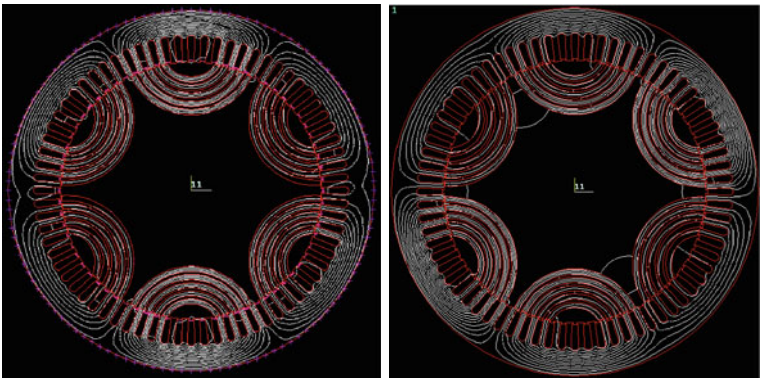
coupling capability of rotor is shown in Fig. 235.4. So in this paper, the most reasonable width of conducting magnetic layer is 6 mm.

235.3.4 The Determination of Conducting Magnetic Layer Number

The analysis results are shown in Table 235.1. In this paper, conducting magnetic layer number is 5.

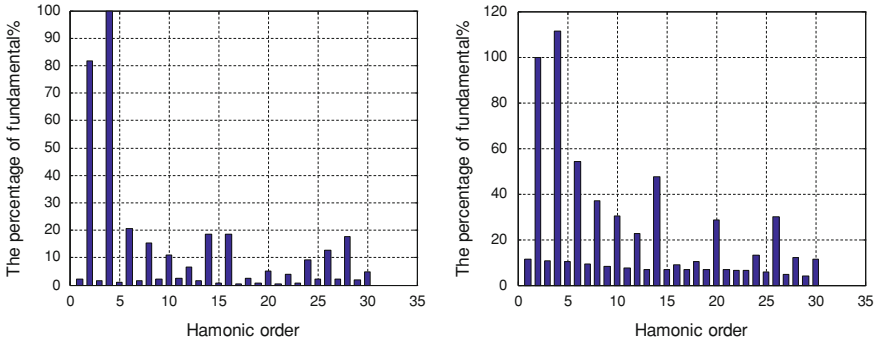
235.4 Finite Element Analysis

The magnetic field analysis of radially laminated magnetic reluctance rotor BDFM is performed under different excitation. Stator plate of Y250 M-6 is selected. The pole-arc is 0.67, conducting magnetic and magnetic obstacle layer width radio is 2, conducting magnetic layer width is 6 mm and conducting magnetic layer number



8 pole winding excitation 4 pole winding excitation

Fig. 235.5 Flux of radially laminated magnetic barrier rotor BDFM



8 pole winding excitation 4 pole winding excitation

Fig. 235.6 Harmonic content of radially laminated magnetic barrier rotor BDFM

is 5. The results are shown in Fig. 235.5. The FFT is used to investigate the harmonic components of air gap flux density quantitatively. The analysis results are shown in Fig. 235.6. It shows that when power windings (pole pairs $P_p = 4$) excite along, the percentage of magnetic field which has the same pole pairs with control windings is 82%; when control windings (pole pairs $P_q = 2$) excite along, the percentage of magnetic field which has the same pole pairs with power windings is 113%.

235.5 Conclusion

This paper studies the coupling capability of radially laminated reluctance rotor BDFM. The finite element method is adopted to calculate magnetic filed of different rotor structures radially laminated magnetic reluctance rotor BDFM under

different excitation. Furthermore, spectral analysis is performed on air gap magnetic density to investigate the coupling capability quantitatively. The results show that: magnetic circuit of radially laminated magnetic reluctance rotor BDFM is more standardized, radially laminated magnetic reluctance rotor has the better coupling capability.

Acknowledgments This work is supported by "Liaoning Bai Qian Wan Talents Program" (2008921038).

References

1. Hunt LJ (1914) The cascade induction motor. J IEEE 52:406–426
2. Creedy F (1921) Some developments in multi-speed cascade induction motors. J IEEE 59:551–552
3. Boger MS, Wallace A, Spée R et al (1995) General pole number model of the brushless doubly-fed machine. IEEE Ind Appl 31(5):1022–1028
4. Zhang FG, Wang FX, Wang Z (2002) Comparative experiment study on the performance of brushless doubly-fed machine with different rotor structures. China Soc Electr Eng 4:52–55
5. Xu L, Liang F, Ye L (1994) Comparison study of rotor structures of doubly excited brushless reluctance machine by finite element analysis. IEEE Trans Energy Convers 9(1):165–172
6. Xu L, Wang F (1997) Comparative study of magnetic coupling for a doubly fed brushless machine with reluctance and cage rotors. IEEE thirty-second IAS annual meeting, IAS 97. Conference record of the 1997 IEEE
7. Wang FX, Zhang FG (2004) Field modulation brushless doubly-fed machine. Jilin University Press, Changchun
8. Xu L, Liang F, Lipo TA (1991) Transient model of a doubly excited reluctance motor. IEEE Trans Energy Convers 6(1):126–133

Chapter 236
Electromagnetic Design and 3D Finite Element Analysis for Axial Sectional Claw Pole Machine with Permanent Magnet Outer Rotor

Zhang Feng-ge, Wang Bo and Liu Guang-wei

Abstract In order to improve work efficiency of the motor in modern automobile ISAD system and solve running problem which the motor copes with under different terrain, the electric and magnetic circuit of claw pole motor of outer rotor permanent magnet of axial split phase is designed and studied in this paper. First, the electric and magnetic circuit of claw pole motor is designed by equivalent magnetic circuit method using MATLAB software. Second, simulation analysis for parameters of no-load back electromotive force, cogging torque and other parameters are conducted by finite element software. Finally, through comparative analysis of experimental and simulation results of the prototype, it is proved that this way of design of the electric and magnetic circuit can improve the motor working performance.

Keywords Claw pole motor · Axial sectional · Electromagnetic design · Finite element analysis

Z. Feng-ge (✉) · W. Bo · L. Guang-wei
Education Ministry and Province key Laboratory
of Special Motor and High Voltage Apparatus,
Shenyang University of Technology, Shen yang 100870, China
e-mail: zhangfg@sut.edu.cn

W. Bo
e-mail: beyond.29@163.com

L. Guang-wei
e-mail: liu-guangwei@hotmail.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_236,

236.1 Introduction

Axial sectional claw pole machine (ASCPM) with permanent magnet outer rotor has a series of outstanding advantages, such as it has a simple structure, low manufacturing cost, brushless reliable, high torque density, high power density, easy to make a multi-polar and suitable for operation at medium or higher frequencies. It makes automotive startup, power generation, electrical and brake integration (the ISAD system of the modern automobile) achievable due to the external rotor structure. Drag by the inverter the traditional electric excitation claw pole motor will generate huge iron loss. Increase the motor losses and reduce efficiency, and the motor temperature will increase, the insulation aging faster, resulting in decreasing motor life. ASCPM will solve these problems well. Since the motor uses soft magnetic composite materials (SMC) as the stator claw pole, it can effectively reduce the iron loss at high frequencies, and no brushes and slip rings, make the motor safe and reliable [1].

236.2 Axial Sectional Claw Pole Machine with Permanent Magnet Outer Rotor

ASCPM as a new motor, it has the following characteristics:

a. The motor adopt three-stage independent structure. There is no coupling between the stators, and can be independently analyzed. In addition, the winding form is simple and convenient, also facilitate to design to be the multi-phase structure, to make it has good control characteristics;
b. The motor has the characteristics of transverse flux. To make the magnetic circuit separated, compared to traditional motor it has high torque density and power density;
c. The motor use SMC, it will have low iron loss while high-speed and high frequency, the mold casting process is relatively simple and can be mass production, to lower the manufacturing costs [2].

The stator and rotor structure of the ASCPM as shown in Figs. 236.1 and 236.2, respectively. In addition to these structural characteristics, the use of soft magnetic composite materials, ASCPM more suitable for working in the high frequency and can be approximated that is only the hysteresis loss and nearly no eddy current loss, compared with traditional motor. Therefore, the motor is more suitable for using inverter to supply power. And achieve high-speed operation, so it has high torque density and high power density. Between the section and paragraph staggered 120 electrical degree, the cogging torque has been weakened to some extent due to the axial section [3].

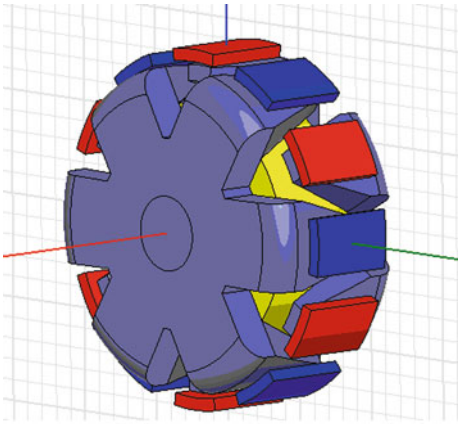


Fig. 236.1 Stator of ASCPM

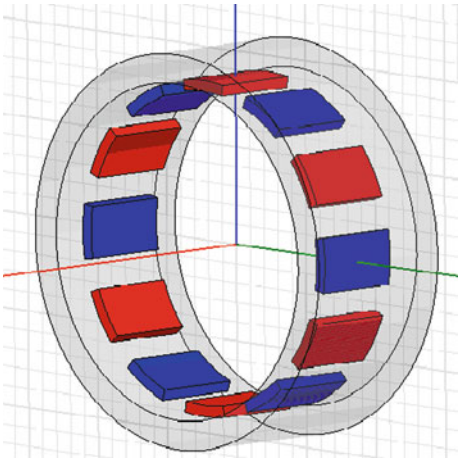


Fig. 236.2 Rotor of ASCPM

236.3 Electromagnetic Design Methodology

236.3.1 Equivalent Magnetic Circuit Model

Equivalent magnetic circuit method used in this paper established a claw pole permanent magnet external rotor motor equivalent circuit model. The basic principle of the equivalent magnetic circuit method is to equivalent magnetic circuit of the motor permeability. Finally in order to solve the problem, the equivalent motor magnetic field is transformed to specific circuit . After transformation, according to the similarity of electric and magnetic circuits, solving each part of the magnetic bits and the flux of each branch, corrected with a correction factor. So that the magnetic circuit of each segment corresponding potential difference is equal to

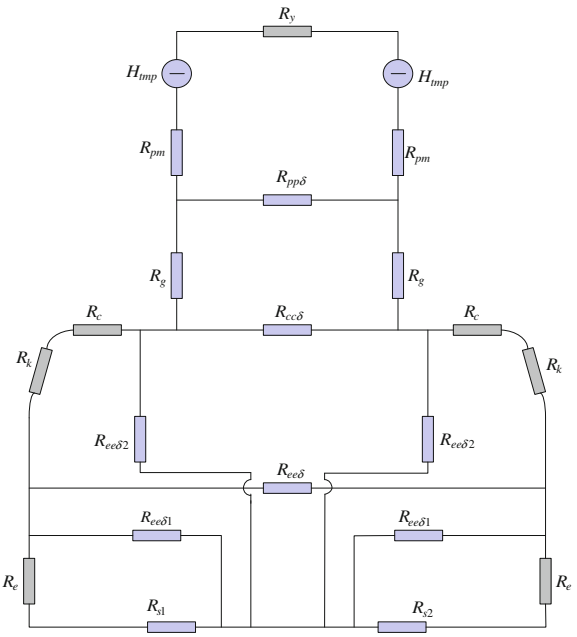


Fig. 236.3 The model of magnetic circuit

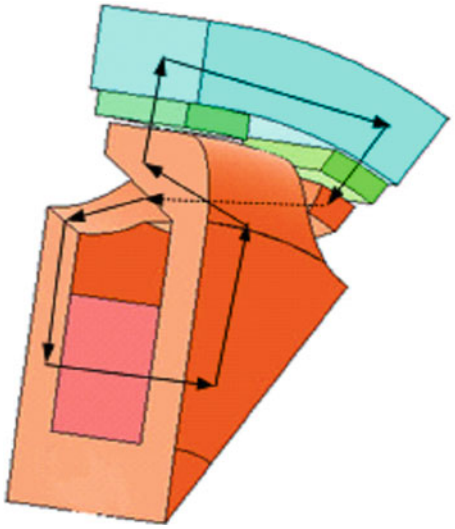


Fig. 236.4 The diagram of flux

the magnetic potential difference between the corresponding points, further obtain the motor parameters and performance. This greatly simplify the calculation and reduces the computational time (Figs. 236.3, 236.4).

According to the modules in the diagram: core axis R_s, Department of claw extreme R_e, Claw knee R_k, Claw part R_c; as well as part of the magnetic flux leakage: Claw leakage flux between the extreme surface $\Lambda_{mcc\sigma 1}$, claw pole magnetic flux leakage between the side $\Lambda_{mcc\sigma 2}$, Department claw extreme R_{eesgm}:

$$R_s = \rho_{smc} \frac{4pL_s}{\pi(D_{s0}^2 - D_{si}^2)} \tag{236.1}$$

$$R_e = \rho_{smc} \frac{p}{2\pi b_e} \ln\left(\frac{D_e}{D_{s0}}\right) \tag{236.2}$$

$$R_k = \rho_{smc} \frac{\pi h_{c2}}{4 b_e b_{c2}} \tag{236.3}$$

$$R_{eesgm} = \frac{4\ p\ L_r}{\mu_0 \pi\left(D_e^2 - D_{s0}^2\right)} \tag{236.4}$$

$$R_c = \rho_{smc} \frac{L_c}{b_{cm} h_{cm}} \tag{236.5}$$

$$R_{cc\sigma} = \frac{1}{\Lambda_{mcc\sigma 1 + \Lambda_{cc\sigma 2}}} \tag{236.6}$$

236.3.2 Determination of Main Dimensions

According to Sect. 236.3.1, design computing platform, the permanent magnet width is the independent variable. Based on the above parts of Claw formula, to derive Claw pole magnetic circuit model of the motor, MATLAB software is programed to cycle through the operation, selecting the optimal size of the permanent magnet parameters. According to the results of operations, the optimal width of the magnet block is $b_{pm} = 13.3$ mm. Based on this calculation and the saturation phase diagram of the ASCPM, derive the input power and the electromagnetic power of Claw pole motor. As the project adopts external permanent magnet rotor, external rotor yoke set with Permanent magnet block. Therefore, according to the salient pole motor $X_d = X_q$ derive input power P_1 and electromagnetic power P_{em} of claw pole motor.

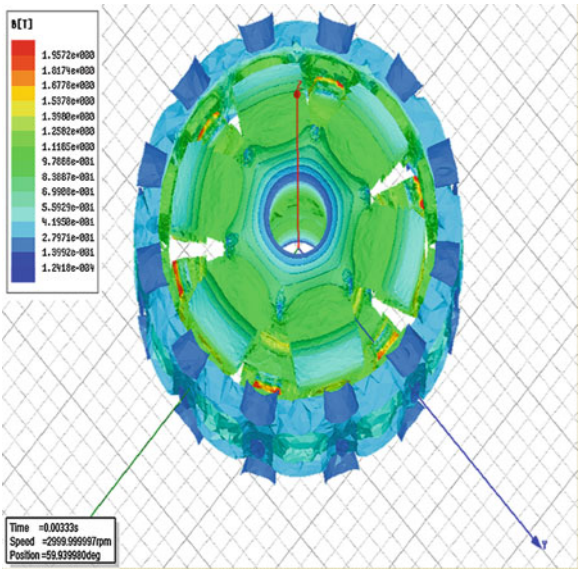


Fig. 236.5 Distribution of magnetic density

236.4 Finite Element Simulation Analysis of Claw Pole Motor

236.4.1 Magnetic Field Distribution in a Motor

In order to reflect the changes of no-load back electromotive force (EMF) and the cogging torque for a motor, the region of the whole motor is soluted with 3D finite element. The complexity of computation is increased, but calculating the boundary conditions is simple. Furthermore, simulation results approximate to the realities of the situation. Due to phase noncoupling of three-section structure the analysis progresses for single stage stator.

Figure 236.5 shows flux density distribution of stator without load. It can be seen from the figure that the end region of claw pole perhaps enter saturation condition, the border of claw pole lightly approach to saturation, and flux density distribution of root is less. Therefore, SMC material can be put into full application and flux density distribution is reasonable. Figure 236.6 shows flux orientations of external rotor. It can be seen from the figure that magnetic circuit is still mutually independent between each item, and magnetic flux leakage is less after surface permanent magnet is adopted. Hence the scheme of external rotor is feasible. Flux orientations: if they start from the N-pole of permanent magnet, they will go through air gap, claw pole, knee, end, stator yoke, the other side of end, knee and claw pole, and then they will pass the S-pole of permanent magnet, rotor yoke. And finally they will reach the N-pole permanent magnet to form a main magnetic flux circuit.

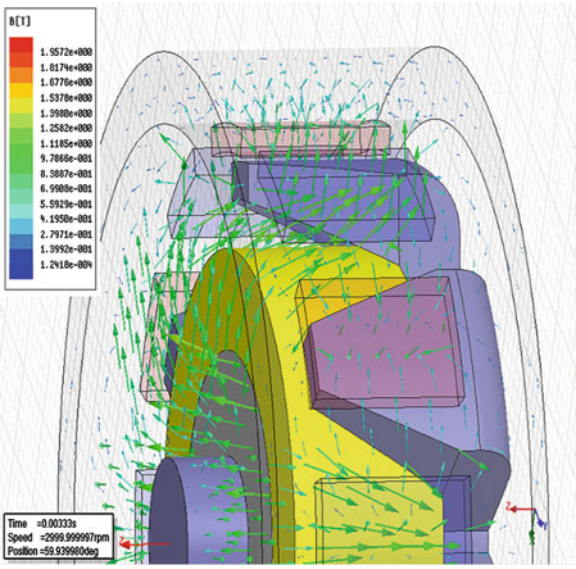


Fig. 236.6 Flux distribution

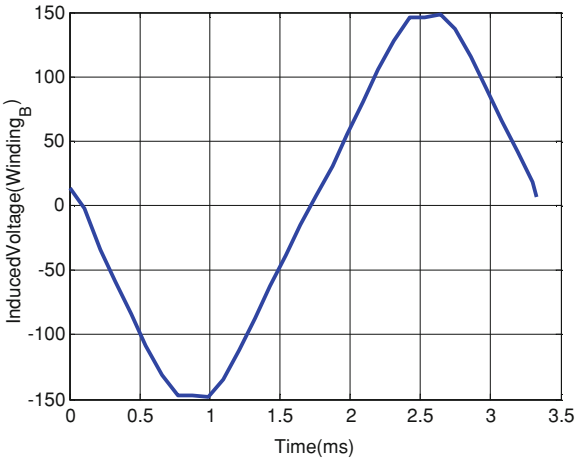


Fig. 236.7 Single-phase of EMF

236.4.2 Characteristic Analysis of the Motor Without Load

When the no-load motor revolve at speed of 3,000 RPM, the characteristics curves are obtained as follows: Fig. 236.7 shows EMF waveform of the single stage claw pole motor. The peak values of no-load EMF is 137.5 V, and close to the sinusoidal distribution. Therefore, harmonic content of back EMF waveform is lower.

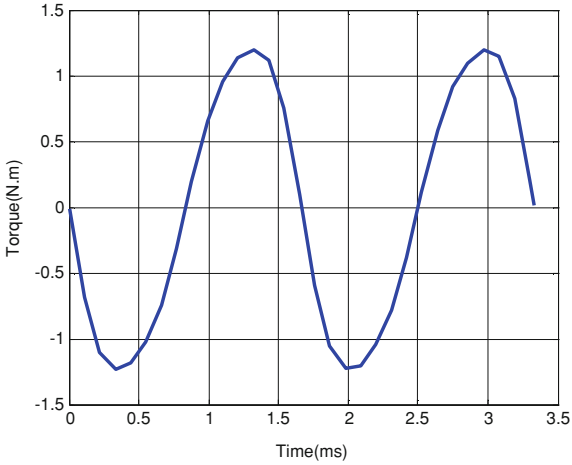


Fig. 236.8 Single-phase of cogging torque

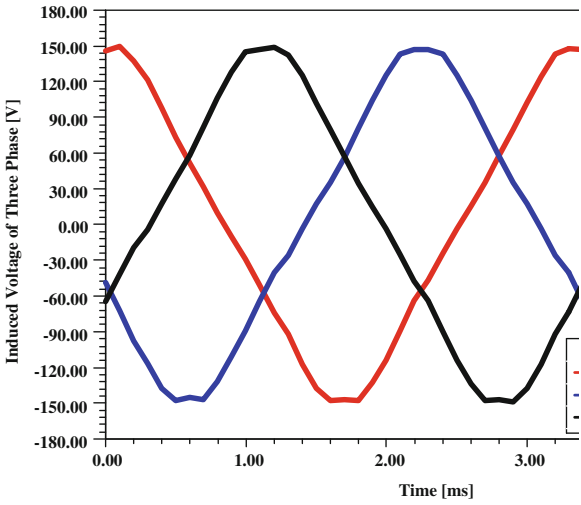


Fig. 236.9 Three-phase of EMF

Figure 236.8 shows single-phase cogging torque of claw pole motor. The peak values is 1.5 Nm. As Figs. 236.9 and 236.10 shows the three-phase, the EMF of three-phase is 149.5, the cogging torque of three-phase is 1.3 Nm.

Therefore, compared with other single stage claw pole motors, axial split phase motor cause cogging torque to be weakened to some extent because there is electric angle of 120 electrical degree between the three sections of claw pole. However, compared with other motors cogging torque is still very large because

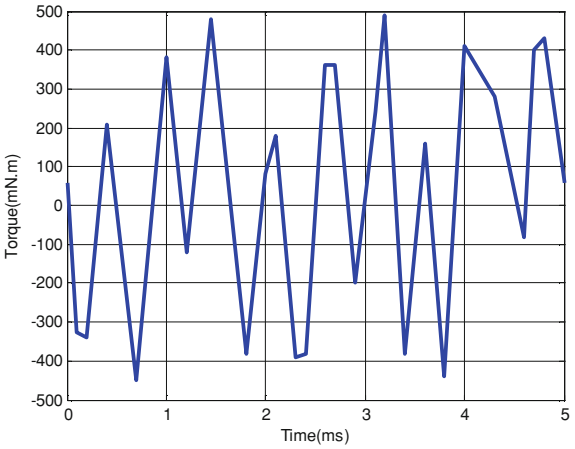


Fig. 236.10 Three-phase of cogging torque

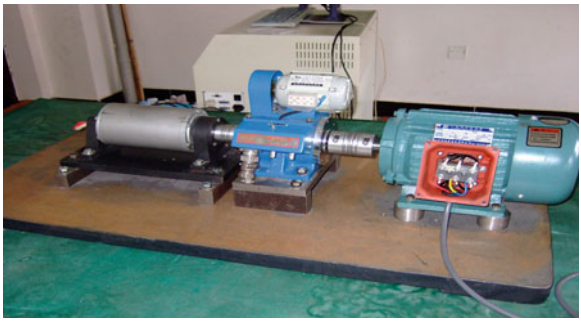


Fig. 236.11 Laboratory bench

the motor belong to pole-to-pole type. Consequently the torque ripple and noise caused by cogging torque still exist.

236.5 Experiment

Figure 236.11 shows laboratory furniture of the project, including experimental prototype, converter, torque and measurement instrument, and three-phase induction motor. The measurement of no-load EMF usually adopts "reverse drawing method" to measure [4]. In other words the prime mover and the tested motor are mechanically connected. And then the prime mover draws the tested motor to run without load and to measure the output voltage of the tested motor. The experiment uses the three-phase induction motor draws the experimental prototype at speed of 3,000 RPM.The waveform of EMF is shown in Fig. 236.12,

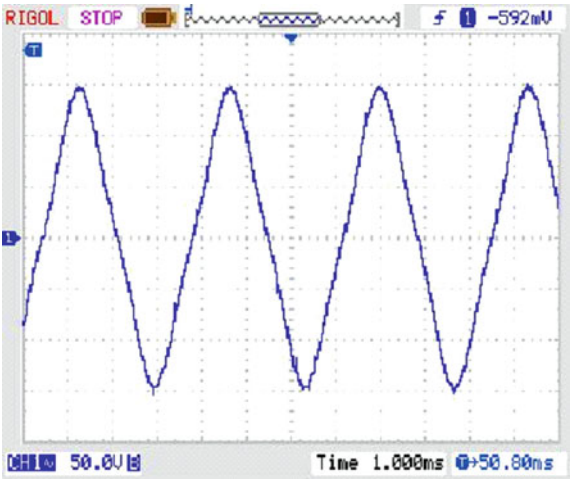


Fig. 236.12 Laboratory Furniture

which is measured from experiment. It can be seen that the waveform are close to sine-wave which is in accord with the result of calculation.

236.6 Conclusion

Based on equivalent magnetic circuit model of claw pole motor, the electromagnetic design of claw pole motor design platform software is develped, the calculation program is derived, the best width of the permanent magnet is obtained by using the MATLAB, and then through the ANSOFT finite element analysis software 3D model simulation, calculated no-load EMF and cogging torque. Finally, the experimental data of ASCPM is given out, which proved that the ASCPM is suitable for the modern automobile ISAD system.

Acknowledgments Project of innovation team of education department of Liao Ning province (2009T075).

References

1. Guo YG, Zhu JG (2002) Magnetic field calculation of claw pole permanent magnet machines using magnetic network method. J Electr Electron Eng 22(1):69–75
2. F Zhang, S Zhang, H Bai (2006) Magnetic field analysis and performance calculation for new type of claw pole motor with permanent magnet outer rotor. IPEMC, Shanghai, pp 827–831
3. Guo YG, Zhu J, Watterson PA (2003) Improved design and performance analysis of a claw pole permanent magnet SMC motor with senseless brushless DC drive. PEDS'2003, vol 1, pp 704–709
4. Zhang F, Bai H (2007) Hans Peter Gruenberger. Comparative study on claw pole electrical machine with different structure. ICIEA, Harbin, pp 636–640

Chapter 237
Simulation Analysis on Temperature Field for Doubly Fed Wind Generator Used in High Altitude Condition

Ding Shuye, Ge Yunzhong and Lü Xiangping

Abstract In this paper, a 1.5 MW doubly fed wind generator was taken as an example. Its three-dimensional heat transfer physical model and mathematical model were established based on giving basic assumptions. Meanwhile, the corresponding boundary conditions of the solved region were given, and the fluid field and temperature field were analyzed by using finite volume method when the generator working in altitude area of 4,000 m. The temperature rise distribution characteristics of solid and fluid inside generator were obtained. The coupling calculating results were compared with the measured results when the generator was in rated operation in plain area. At last, the effect of the altitude to the temperature inside generator was analyzed, which would provide theoretical basis for the operation of wind generator.

Keywords Wind generator · Fluid–solid coupled · High altitude · Temperature field · Fluid field

D. Shuye (✉) · G. Yunzhong
College of Electrical and Electronic Engineering, Harbin University of Science and Technology, Harbin city 150040, China
e-mail: geyunzhong717@126.com

L. Xiangping
AC/DC Generator Company of Harbin Electrical Machinery Plant Limited Liability Company, Harbin city 150040, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_237,

237.1 Introduction

The land elevation above 1,000 m is accounting for 60% of the land area of China, the land elevation above 2,000 m accounting for 33% and the land elevation above 3,000 m accounting for about 16% of the area. Because of the big elevation difference in our country, there is a wide application of wind driven generator. For the generator working at high altitude, the influence of the altitude to the temperature in wind generator must be considered.

The compulsory air-cooled and liquid-cooled is widely used in the traditional wind turbine generator [1]. As the elevation rises, the air pressure and the density decrease which result in the drop of the cooling capacity of the traditional electrical equipment.

Many experts and scholars all over the world have researched internal physical field in generator recent years [2–5]. But few of which have researched the generator working at high elevation [6]. Literature [7] analyzed the internal fluid field and temperature field of a 3 MW large permanent magnet wind generator in detail, which pointed out the distribution characteristics of the internal fluid flow and temperature in the generator.

In this paper, a 1.5 MW doubly-fed wind generator with double medium in 4,000 m altitude region was taken as example. The three-dimensional fluid field and temperature field were solved directly using finite volume method, and the distribution characteristic of temperature field and fluid field was obtained. The coupling calculating result was compared with the measured results when the generator was in rated operation in plain area. At last, the influence of the altitude to the temperature rise in generator was analyzed.

237.2 Climatic Characteristics of High Altitude Areas

237.2.1 The Relationship of Altitude with Air Pressure and Density

With the gradual increasing of altitude, the air pressure and density will reduce gradually. The changing relation of the air pressure and density with the altitude are list in Table 237.1.

237.2.2 The Relationship Between Air Temperature and Altitude

With the increasing of altitude, the environmental temperature reduces. The relationship between temperature and altitude is list in Table 237.2.

Table 237.1 Relationship between air pressure, air density and altitude

Altitude/km	0	1	2	3	4	5
Air density (kg/m^3)	1.292	1.1667	1.0504	0.9432	0.8437	0.7532
Air pressure (kPa)	101.0	90.0	79.5	70.1	61.7	54.0

Table 237.2 Relationship between temperature and altitude

Altitude/km (°C)	1	2	3	4
Highest temperature	40	35	30	25
Average temperature	20	15	10	5

237.3 Mathematical and Physical Model of the Fluid Flow and Heat Transfer in the Generator

237.3.1 The Structure of Cooling System

The 1.5 MW double-fed wind generator was using dual dielectric cooling way which was divided into inner and external cooling ways. The inside ventilation system means that a kind of fan was set up inside the wind turbine. The air was forced ventilated cooling by the fan in the generator which formed the closed-end cycle. The air circulated along the closed-end cycle, and then the heat in the air transferred to the external fluid, finally the temperate of cooling fluid was decreased. The outside ventilation system means that the fluid cycle system was assembly on generator casing, and this cycle system the cooling water transfer the energy with the hot air of the first cooling. The heat generated by the loss during the generator continuous running was transferred outside the generator.

237.3.2 Mathematical Description

The fluid flow and heat transfer in cooling system of generator according to principles of mass, momentum and energy conservation. If the flow was in disturbance state, it also complied with additional turbulent transport equations.

237.3.2.1 The k − ε Mathematical Model of Disturbance Flow

This paper used the most widely standard $k-\varepsilon$ two equations model. The standard $k-\varepsilon$ model is a semi-empirical formula based on turbulent kinetic energy and dissipation rate. The k equation is an exact equation, while the ε equation is deduced by empirical formula [4].

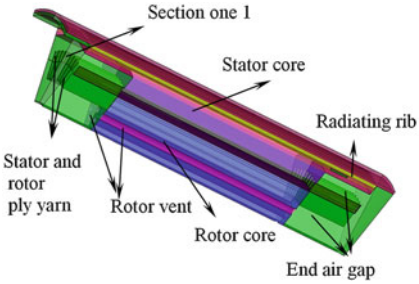


Fig. 237.1 Physical model of solution domain

237.3.2.2 Three-Dimensional Heat Transfer Equation

The steady-state temperature field of the generator was solved, and thermal control equation selected the thermal control equation in which had internal heat source three-dimensional steady-state anisotropic medium. The equation in Descartes coordinates is as follows:

$$\frac{\partial}{\partial x}\left(k_x \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k_y \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k_z \frac{\partial T}{\partial z}\right) = -q \tag{237.1}$$

where, T was the temperature of the object; λ_x, λ_y, λ_z were respectively the coefficient of thermal conductivity at x, y, z direction $W/(m.K)$; q was internal heat source density W/m^3.

237.3.3 The Physical Model in Solving Domain

According to the fluid flow and heat transfer characteristics in generator wind ventilation path system and chassis flow systems, the fluid flow and heat transfer coupling solution model of the generator cooling system could be determined as a 1/8 area in circumference direction. The solving domain physical model was shown in Fig. 237.1.

237.3.4 Basic Assumptions and Boundary Conditions

237.3.4.1 Basic Assumptions

According to its structure and the fluid characteristics in the generator, the hypotheses were as follows:

In solving process, the specific characteristic of cooling medium was set for 4,000 m altitude;

In the internal flow field of the generator, the velocity of fluid is far less than the velocity of sound, so the fluid was taken as incompressible fluid;

Because the Reynolds number of the fluid in the generator was very large ($Re > 2{,}300$) which belonged to the turbulent, the method of turbulence model is adopted in the flow field;

The heating process in actual generator was very complex, so all the generator parts were seen as isothermal solid. The power loss fevered in its all parts regardless of other influences and it was average distribution, thus taken the generator as an ideal heating element;

The entrance of the model adopted speed entrance and the export of the model used pressure export.

237.3.4.2 Boundary Conditions

When the direct coupling calculating method was adopted to calculate fluid field and temperature field in the stator and the rotor, the surfaces between bodies were directly iteratively computed by the way of coupling boundary in solving regional style. In the convective heat transfer process, the thermal boundary condition was not given in advance which was constrained reciprocally by the fluid and the mural face.

The flow of air was 1.45 m^3/s. Combined with the entrance area of the model, the velocity of air flow at the entrance was 3.18 m/s; the flow of water was 55 L/min; the velocity of water flow was 1.43 m/s. The stator, the rotor core, the stator and the rotor thread were taken as the heat source in this model.

237.4 The Numerical Calculation and Analysis of Fluid and Temperature Field

The finite volume method was adapted to couple solving the three-dimensional coupled solution region, and the differential implicit iteration convergence format was used. Through calculation, the distribution characteristics and laws of flow and temperature could be obtained.

237.4.1 Fluid Field in Solving Domain

The finite volume method was adopted to couple solving the three-dimensional fluid field and the coupling field in the generator, the flow characteristics and law of the fluid field was shown in Fig. 237.2.

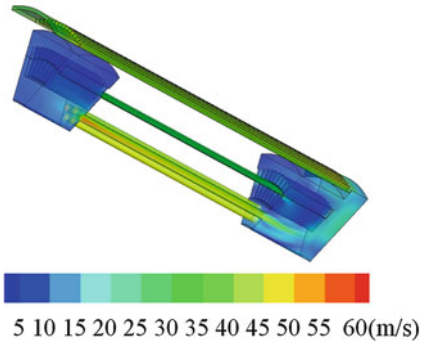


Fig. 237.2 Air velocity distribution profile in three-dimensional flow field

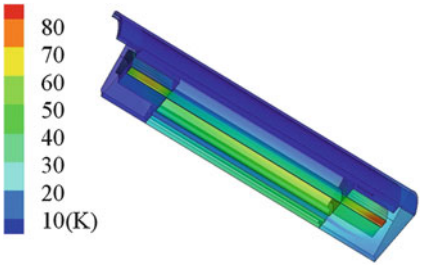


Fig. 237.3 The temperature distribution of the solving field

From Fig. 237.2, it can be seen that the maximum velocity of air was 62.09 m/s in solving domain, which distributed in the rotor shaft. In the end gas cavity, the wake appeared in the rotor shaft. With the increasing gas chamber at the end, this phenomenon disappeared gradually. When the air arrived at the end area, the space greaten, the air velocity reduced and the flow velocity was relatively flat. So the air can exchange heat sufficiently with the end winding to reduce the temperature of the end winding.

237.4.2 The Calculation Results for Temperature Field

When the generator was in rated load operation, the completely domain temperature rise distribution was shown in Fig. 237.3. It can be seen that the maximum temperature rise distribution was in the stator thread in solving the domain and the maximum value was 82.3 K.

For the purpose of detailed analysis of the temperature rise distribution of the solving domain, a cross-section was intercepted in the solution model as shown in the Fig. 237.4. This section is one radial direction section along the stator and the rotor centerline.

Figure 237.4 was the radial temperature rise distribution of the solving domain. From Fig. 237.4, the axial temperature rise distribution in the generator can be

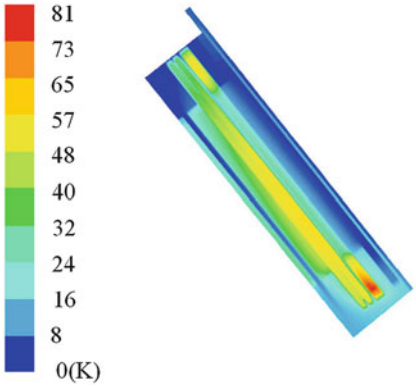


Fig. 237.4 The radial temperature distribution of the solving domain

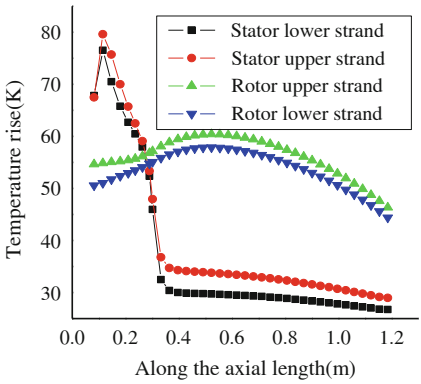


Fig. 237.5 Thread along the axial length temperature distribution

seen. At fan end, the thread at the end of the stator had the highest temperature and the rotor thread within the core was higher. The temperature rise distribution of the upper and the lower thread of the stator and the rotor from the air inlet of the core to the end thread of the fan were shown in Fig. 237.5. From Fig. 237.5, the axial temperature rise variation of the thread can be seen clearly.

237.4.3 Comparison and Analysis of Calculated Temperature for the High Altitude Areas at Measuring Position with Experimental Temperature for the Plain Region

The temperature sensors were placed at the end of the stator thread in running, and six temperature rise measuring element were uniform layout. The experimental value was namely the average value read from these six temperature sensors.

The temperature rise was 75.7 K when the wind turbine generator was running on rated condition at 4,000 m altitude. And the temperature rise was 62 K when the wind generator was running on rated condition at the plain. The temperature rise of the wind generator at 4,000 m altitude was 13.7 K higher than that at the plain in rated condition. This phenomenon was due to the higher altitude which lowered the pressure and the air density. The volume of cooling air which took away the heat of generator was not changed. But the air density decreased, and the heat capacity is also reduced accordingly which made the decreased cooling capacity of generator and higher temperature rise.

The literature [6] had a detailed analysis about a 750 KW wind generator using at high altitude. The result was that when the altitude was 4,000 m, the temperature rise increased about 16 K. In other words, when the same generator ran at the same condition, the temperature rise at 4,000 m altitude was 16 K higher than that at plain. The temperature rise of the generator was about 78 K at high altitude which was 2.9% different to the calculated value.

Higher altitude and reduced ambient temperature rise had a compensation effect for the radiate of wind turbines, and which required the analysis of environmental conditions of the actual operation of the generator. When the raised temperature of the generator should be solved in terms of actual environmental conditions, the relationship between the temperature rise and altitude above sea level in Table 237.2 was not influenced by other factors in addition to the sea level elevation. Therefore, more important considering was reduced pressure and air density, and for ambient temperature rise compensation is based on the physical environment temperature rise of the generator.

When the wind generator ran in the cabin, the temperature rise in the cabin was namely the environment temperature of the generator. When calculated the temperature rise of wind generator, the influence of the reducing atmospheric pressure and air density with the increasing altitude should be considered and the environment in the cabin where the generator placed in should also be considered.

237.5 Conclusions

(1) Considering the influence of the environment at high altitude on the heat dissipation of generator, the fluid and temperature distribution law of this generator was obtained which can provide guidance for the generator working at high altitude.
(2) For air-cooling system, the influence that high altitude acted on its temperature rise is mainly the reducing atmospheric pressure and air density with the increasing elevation. The compensation for environmental temperature was based on the cabin environment.
(3) Quantity of heat in rotor must be transferred out by two convective heat transfers; the cooling conditions of the rotor part were relatively bad.

(4) The highest temperature rise of the rotor thread appeared in the core, and the temperature rise of rotor thread was significantly higher than the temperature rise of stator thread.

References

1. Yuan WW, Jiang YL (2007) Cooling systems in wind turbine. World Sci-tech R&D 29:80–85
2. Yao RP, Rao FQ (2003) Research on stator thermal field of evaporation cooling hydro-generator. Proc CSEE 23:87–90 (in Chinese)
3. Lu YP, Zheng GL, Li JT et al (2010) Uniformity of multi-path ventilation in rotor of air inner-cooled turbo generator. Proc CSEE 30:99–104 (in Chinese)
4. Li WL, Ding SY, Jin HY (2005) Numerical calculation of large synchronous generator stator temperature fields based on coupled fields. Proc CSEE 25:129–134 (in Chinese)
5. Georg TS, Richard ZM, Alexander S (2010) Cooling air flow, losses, and temperatures in large air-cooled synchronous machines. IEEE Trans Ind Electron 57:172–180
6. Zhang F, Zhu CJ, Zhang XL (2010) 75 KW of high altitude wind generators type application. Electrotech Appl 29:44–48
7. Jiang YM, Ding SY, Ge YZ (2010) Numerical analysis of fluid field & thermal field inside large wind-generator. Power Convert Technol 4:24–27 (in Chinese)

Chapter 238
The CAD System of the Wound-Rotor Brushless Doubly Fed Generator

Fei Xiong and Xuefan Wang

Abstract For analyzing and calculating the steady-state performance of the brushless doubly fed generator (BDFG) accurately and quickly, based on the electromagnetic design program of the traditional three-phase induction machines and the equivalent circuit of the BDFG, the computer aided design (CAD) system of the wound-rotor BDFG is developed by using the software of Visual C++ 6.0 in this paper. The machine performance under different rotor speeds and different control voltages can be calculated in the CAD system. The comparison of the calculated results and the experimental data indicates that the results of the CAD system are accurate and it is effective and convenient for the design of the wound-rotor BDFG by using the computer.

Keywords Brushless doubly fed generator (BDFG) · Computer aided design (CAD) system · Equivalent circuit · Steady-state performance

238.1 Introduction

The brushless doubly fed generator (BDFG) is a special induction machine which has two stator windings of different pole pairs and one rotor. One stator winding, the power winding (PW), is directly connected to the grid or the load and it

F. Xiong (✉) · X. Wang
Department of Electrical Machinery, Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: hust_xf@126.com

X. Wang
e-mail: xuefanwang@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_238,

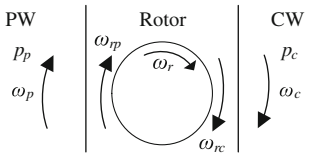


Fig. 238.1 BDFG synchronous mode

provides the power; and the other one, the control winding (CW), is fed by a converter with variable voltage at variable frequency connected to grid or the load and the converter is bidirectional power-flow [1, 2]. The rotor structure can be cage, reluctance or wound rotor [3–5]. To avoid direct coupling between the stator windings, each of the stator windings generates a magnetic field and the rotor is designed to couple to both magnetic fields. The normal operation of the BDFG is in a doubly fed and synchronous mode, and the shaft speed has a fixed relationship to two excitation frequencies [6].

Compared with traditional doubly fed slip-ring wound–rotor induction generator, the absence of the slip rings and brushes in the BDFG increases reliability and reduces maintenance. This is particularly important as more and more installations are being constructed offshore and in difficult-to-reach places [7]. Thanks to significant benefits of the BDFG, it has wide application in variable speed generation (VSG) fields, such as wind power generation and ship shaft generator systems [8].

For calculating the performance of wound-rotor BDFG accurately and quickly, the computer aided design (CAD) system of the machine is developed in this paper. The electrical parameters, such as the resistance and inductance, and the steady-state performance of the machine can be calculated when the basic machine design parameters are input in the CAD system.

238.2 Basic Operation of the BDFG

The BDFG can be operated as a grid-connected or stand-alone generator, and the doubly fed and synchronous mode is used for controlling the PW voltage, frequency and power. Stator and rotor quantities are shown for the synchronous mode in Fig. 238.1. The rotor angular velocity ω_r is given by

$$\omega_r = \frac{\omega_p + \omega_c}{p_p + p_c} \tag{238.1}$$

and the slips for two windings s_p. and s_c can be defined as

$$s_p = \frac{\omega_p - p_p\omega_r}{\omega_p} = \frac{\omega_{rp}}{\omega_p} \tag{238.2}$$

$$s_c = \frac{\omega_c - p_c\omega_r}{\omega_c} = \frac{\omega_{rc}}{\omega_c} \tag{238.3}$$

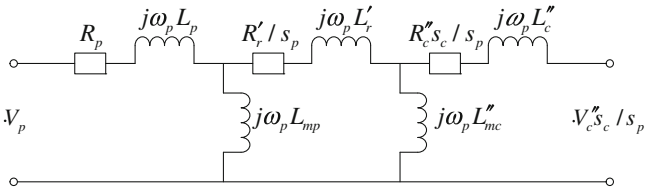


Fig. 238.2 Equivalent circuit for the BDFG

where ω_p and p_p are the angular velocity and pole pairs of the PW, and ω_c and p_c are the angular velocity and pole pairs of the CW.

The angular velocities of the magnetic fields due to the PW and CW in the rotor reference frame ω_{rp} and ω_{rc} are

$$\omega_{rp} = \omega_p - p_p \omega_r \tag{238.4}$$

$$\omega_{rc} = \omega_c - p_c \omega_r \tag{238.5}$$

In the synchronous mode, from (238.1), (238.4) and (238.5), $\omega_{rp} = -\omega_{rc}$. Hence from (238.2) and (238.3), it can be derived that

$$\frac{s_c}{s_p} = -\frac{\omega_p}{\omega_c} \tag{238.6}$$

From (238.1), it is shown that when the rotor speed is a constant, the PW voltage and current frequency can be changed by modifying the CW voltage or current frequency.

238.3 Equivalent Circuit for the BDFG

The equivalent circuit is a simple method of representing the steady-state performance of the machine. Since the meaning of the parameters has a clear physical interpretation, the equivalent circuit can be helpful for the machine understanding, design and optimization. The equivalent circuit is used to calculate the voltages, currents, torque, efficiency, power factor and other steady-state values of the machine.

The BDFG can be conceived as two induction machines with different pole pairs in the same frame, intended for the operation in the self-cascaded mode. Therefore a per-phase equivalent circuit can be developed to represent the BDFM as two connected induction machines. The equivalent circuit for the BDFG with omitting iron losses is shown in Fig. 238.2, where $\dot{V}_p$ and $\dot{V}''_c$ are the voltages of the PW and the CW respectively, and all the parameters are referred to the PW. The parameters are defined in Table 238.1.

Table 238.1 Definition of the parameters

Parameters	PW	CW	Rotor
Resistance	R_p	R''_c	R'_r
Inductance	L_p	L''_c	L'_r
Magnetizing inductance	L_{mp}	L''_{mc}	

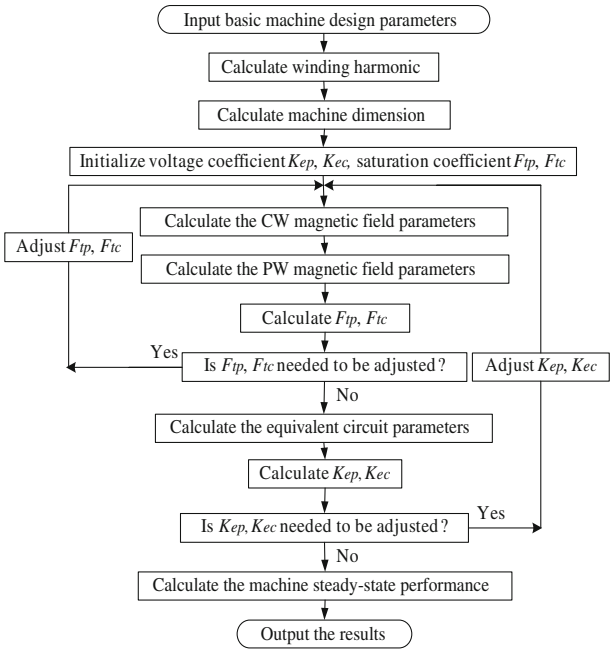


Fig. 238.3 Main program flow chart of wound-rotor BDFG

Figure 238.2 can be recognized as two normal induction machine equivalent circuits with rotors connected together, which is conceptually how the BDFM operates. Similar to induction machines, when the machine dimension and winding parameters are derived, the parameters of the equivalent circuit can be calculated.

238.4 CAD System of the Wound-Rotor BDFG

Based on the equivalent circuit of the BDFG and the improvement of the electromagnetic design program of three-phase induction machines, the main program flow chart is shown in Fig. 238.3.

The main program flow chart have four parts: winding harmonic calculation, machine dimension calculation, magnetic field calculation and equivalent circuit

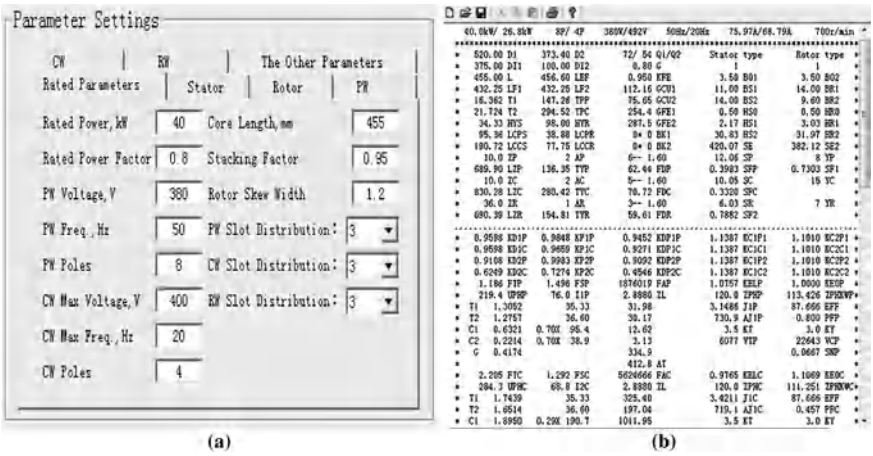


Fig. 238.4 Software interface of the CAD system. **a** Software interface of the CAD system. **b** Output interface

calculation. Because there is no direct coupling between the stator windings, the calculation idea is that: first, the BDFG is considered as two normal induction machines in the magnetic field calculation and the magnetic field parameters of two machines are computed; and then, the equivalent circuit parameters can be obtained after the magnetic field calculation; finally, the machine steady-state performance are calculated by using the equivalent circuit of the BDFG. There are four assumed values: voltage coefficients K_{ep}, K_{ec}, and saturation coefficients F_{tp}, F_{tc}. The assumed values should be regulated by the absolute error between the calculated and assumed values until two values are very close.

According to the main program flow chart, the CAD system is developed by using the software of Visual C++ 6.0. The user just only inputs the basic machine design parameters or loads the input files in the operation interface of the CAD system, and then the CAD system can give the output results and files. The CAD system has the friendly software interface, shown in Fig. 238.4.

238.5 Calculation Results

The experimental research has been made for a prototype BDFG with 4 and 2 pole pairs and the calculation results and experimental data are shown in Table 238.2, where the phase voltage of the PW is kept at 220 V/50 Hz.

From Table 238.2, the comparison of the calculated results and the experimental data indicates that the results of the CAD system are rather accurate, and the CAD system can satisfy the accuracy design requirement of the BDFG. The

Table 238.2 Comparison of the calculation results and experimental data

Speed (rpm)	Calculated results			Experimental data		
	PW power (kW)	CW phase voltage (V)	CW phase current (A)	PW power (kW)	CW phase voltage (V)	CW phase current (A)
400	39.8	159.4	127.1	39.8	173	124
450	39.7	85.9	125.1	39.7	95	119
480	40.8	44.5	125.2	40.8	52	122
510	40.3	22.8	125.5	40.3	24	116
700	41.7	326.7	129.3	41.7	359	126

CAD system provides an effective and convenient method for the design of the wound-rotor BDFG by using the computer.

238.6 Conclusions

The CAD system of wound-rotor BDFG is developed in this paper and it has been used to calculate the steady-state performance of the machine. The calculation results are in good agreement with the experimental data.

References

1. Willianson S, Ferreira AC, Wallace AK (1997) Generalised theory of brushless doubly-fed machine. Part 1: analysis. IEE Proc Electr Power Appl 144:111–122
2. McMahon RA, Roberts PC, Wang X, Tavner PJ (2006) Performance of BDFM as generator and motor. IEE Proc Electr Power Appl 153:289–299
3. Roberts PC, McMahon RA, Tavner PJ, Maciejowski JM, Flack TJ, Wang X (2004) Performance of rotors for the brushless doubly-fed (induction) machine (BDFM). In: 16th international conference of electrical machines. IEEE Press, New York, pp 450–455
4. Xu L, Liang F, Lipo TA (1991) Transient model of a doubly excited reluctance motor. IEEE Trans Energy Convers 6:126–133
5. Wang X (2003) A new brushless doubly-fed machine with a wound-rotor changing-pole winding. In: Proceedings of the CSEE. 23, pp 108–111 (in Chinese)
6. Broadway ARW, Burbridge L (1970) Self-cascade machine: a low speed motor or high frequency brushless alternator. Proc Inst Electr Eng 117:1277–1290
7. Shao S, Abdi E, McMahon RA (2007) Stable operation of the brushless doubly-fed machine (BDFM). In: 7th international conference on power electronics and drive systems. IEEE Press, New York, pp 897–902
8. Wang F, Zhang F, Xu L (2002) Parameter and performance comparison of doubly fed brushless machine with cage and reluctance rotors. IEEE Trans Ind Appl 38:1237–1243

Chapter 239
Contour Controller Design for Direct Drive XY Table System Based on Friction Observer

Li-Mei Wang, Kai Su and Bing Li

Abstract The contour accuracy of XY table directly driven by two permanent magnet linear synchronous motors is influenced by the nonlinear dynamics. To enhance the free-form contour following performance, real-time contour error algorithm was adopted. The friction observer can overcome the influences of the nonlinear dynamic of the single axis, the velocity mismatch between two axis is eliminated by the contour controller. The simulation results show that the designed controller makes the system of XY table have high contour machining precision and strong robustness.

Keywords PMLSM · Direct drive XY table · Real-time contour error method · Friction observer · Contour controller

239.1 Introduction

High efficiency and high precision are the basic features of high speed machining process. It is promoted that the linear motor is used widely in CNC machine tool, industrial robot and industrial automation [1].

For the high precision direct drive XY table mainly focuses on two kinds of methods: reduce the tracking error of the single axis and improve the contouring accuracy of the system. The major factor that influences tracking accuracy is nonlinear dynamic friction force. The existence of the nonlinear friction force deteriorates the precision performance of a single axis and also deteriorates the

L.-M. Wang (✉) · K. Su · B. Li
Shenyang University of Technology, Shenyang 110870, China
e-mail: libinger521@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_239,

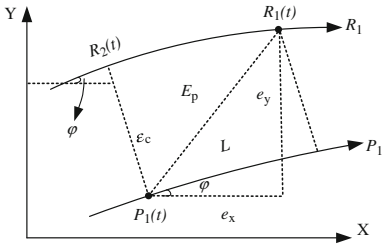


Fig. 239.1 Real-time contour error

overall contouring performance. To achieve higher precision of motion, the friction force can be adequately compensated. So, a friction observer is designed to compensate the influences of the nonlinear disturbance, but this method cannot ensure the contouring accuracy because of the influences of velocity mismatch between two axes. For it, Koren proposed the cross-coupled controller (CCC) dedicated to improve biaxial dynamic performance and reduce the contour error, but the analysis and design only limit to the linear contour error [2]. Chen proposed an equivalent errors method that applicable for multi-axis nonlinear motion contour following control system, but it wasn't perfect for the free-form contour error model [3]. Chen proposed a real-time contour error method, which particularly suited for free-form contour following tasks [4].

So, this paper employs a different strategy to achieve higher precision. The proposed friction observer as the single axis controller compensates the nonlinear frictional characteristics of the system. The designed contour controller entirely improves the contouring accuracy and forces the contour error approach zero in finite time. Moreover, this chapter employs the real-time contour error estimate algorithm to reduce the modeling error that is caused by the time variant trajectory.

239.2 Dynamic Model of Direct Drive XY Table

XY table is driven by two permanent-magnet linear synchronous motors. In free-form contour following tasks, an accurate contour error model can be exploited to improve contouring accuracy. So, real-time contour error algorithm applicable for multi-axis nonlinear motion system was adopted.

239.2.1 Real-Time Contour Error Model

For high precision machining, the model accuracy directly influences contour performance. And for the free-form contour following task, it is very difficult ensure the model accuracy due to the time variant trajectory. This paper adopts the

real-time contour error estimate algorithm to establish the contour error model as shown in Fig. 239.1.

Where R_1 and P_1 denote desired trajectory path and actual trajectory path; $R_1(t)$ and $P_1(t)$ represent the command position and the actual position respectively, $R_2(t)$ are the points on the reference path. Where the length L is the projection of current tracking error E_p on the tangent vector at $P_1(t)$, the distance between $R_2(t)$ and $R_1(t)$ on the tangent vector at $R_1(t)$ is approximate to length L. By definition, the contour error ε_c is the shortest distance from the current actual contact position $P_1(t)$ to the reference trajectory.

The contour error ε_c is calculated by the distance between the position $P_1(t)$ and $R_2(t)$ as:

$$\varepsilon_c = -e_x \sin\varphi + e_y \cos\varphi \tag{239.1}$$

where e_x and e_y denote the axial tracking errors of x and y axes, φ denotes the angle between the tangential line to the horizontal x-axis, and can be denoted as:

$$\varphi = \tan^{-1}\left(\frac{R_{1y}(t) - R_{2y}(t)}{R_{1x}(t) - R_{2x}(t)}\right), \tag{239.2}$$

where $R_{1y}(t)$ and $R_{1x}(t)$ are the components of the command position point $R_1(t)$ of the x and y axes; $R_{2y}(t)$ and $R_{2x}(t)$ are the components of the position point $R_2(t)$ of the x and y axes. Based on the previous formulations, the real-time contour error model of the free-form trajectory is calculated as the Eqs. (239.2) and (239.3). Based on the above analysis, the reduction in contour error cannot guarantee the following performance of the system, therefore constitute the tracking error and contour error as a new state variable ε_p:

$$\varepsilon_p = \begin{bmatrix} \varepsilon_c \\ \varepsilon_t \end{bmatrix} = Te_p = \begin{bmatrix} -\sin\varphi & \cos\varphi \\ \cos\varphi & \sin\varphi \end{bmatrix} \begin{bmatrix} e_x \\ e_y \end{bmatrix} \tag{239.3}$$

239.2.2 XY Table System Dynamic Base on Lugre Model

To represent all the friction force dynamics, this chapter employ the Lugre friction model, the Lugre model is written as:

$$\frac{dz}{dt} = v - \frac{|v|}{g(v)} z \tag{239.4}$$

$$g(v) = \{f_c + (f_s - f_c)\exp[-(v/v_s)^2]\}/\sigma_0 \tag{239.5}$$

$$f_r = \sigma_0 z + \sigma_1 \frac{dz}{dt} + \sigma_2 v \tag{239.6}$$

where the average deflection of the bristles is denoted by z, v denotes the relative velocity between two surfaces, v_s represents the Stribeck velocity, f_r denotes the overall friction force, f_c denotes the Coulomb friction, f_s denotes the maximum static friction force, σ_0 denotes the structural stiffness, σ_1 denotes the damping coefficient, σ_2 denotes the viscous coefficient, $g(v)$ represents the partial nonlinear static friction force function when the velocity of the system keep constant.

Based on the previous the model, XY table dynamic model can be written as:

$$\begin{cases} m_x \ddot{x}_1 = u_x - f_{rx} \\ m_y \ddot{x}_2 = u_y - f_{ry} \end{cases} \tag{239.7}$$

where x_1 and x_2 denote the displacement of the mover, m_x and m_y denote the total mass of the mover and the load of x and y axes, respectively, u_x and u_y denote the control input of x and y axes; f_{rx} and f_{ry} denote the overall friction force of x and y axes, then using both variables the state equation can be written as:

$$F_r = \begin{bmatrix} f_{rx} \\ f_{ry} \end{bmatrix} = \begin{bmatrix} \sigma_{0x} z_x + \sigma_{1x} \dfrac{dz_x}{dt} + \sigma_{2x} v_x \\ \sigma_{0y} z_y + \sigma_{1y} \dfrac{dz_y}{dt} + \sigma_{2y} v_y \end{bmatrix} = \Sigma_0 z + \Sigma_1 \dot{z} + \Sigma_2 v \tag{239.38}$$

where z_x and z_y denote the average deflection of the bristles of x and y axes, v_x and v_y denote the velocity components of x and y axes; σ_{0x}, σ_{0y}, σ_{1x}, σ_{1y}, σ_{2x} and σ_{2y} denote three kinds of friction coefficients, where Σ_0, Σ_1 and Σ_2 are all diagonal with the coefficients σ_{0i}, σ_{1i} and σ_{2i} on their diagonals, correspondingly, and the index $i = x, y$ denote the x-axis and the y-axis, respectively.

239.3 Contour Controller Design Base on Friction Observer

In this study, a contour controller base on the friction observer is proposed and synthesized in the following.

239.3.1 Friction Observer Design

The average deflection of the bristles z that describes the friction dynamic is immeasurable, so utilize the state observer to solve the problem. Under such circumstances, take the inaccessibility of the velocity vector v and the bristle state z as the objective of the observer. Therefore, an observer-type controller is proposed as:

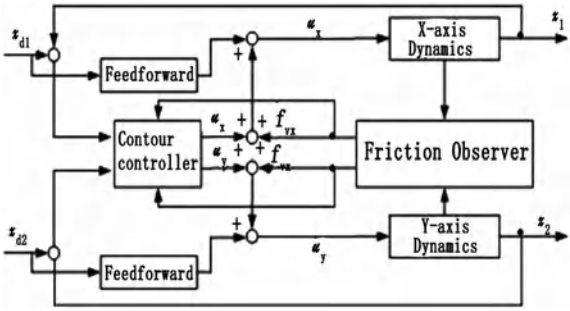


Fig. 239.2 XY table control system

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{v}} \\ \dot{\hat{z}} \end{bmatrix} = \begin{bmatrix} 0 & I & 0 \\ 0 & -M^{-1}(\Sigma_1 + \Sigma_2) & -M^{-1}\Sigma_0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{v} \\ \hat{z} \end{bmatrix} + \begin{bmatrix} 0 \\ M^{-1} \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ M^{-1}\Sigma_1 \\ -I \end{bmatrix} g(\hat{v}, \hat{z}) + L(x - \hat{x}) \tag{239.9}$$

where $\hat{x}$, $\hat{v}$ and $\hat{z}$ denote the result of observer of the state variable x, y and z, the control input matrix is denoted by $u = [u_x \quad u_y]^T$, observer gain matrix is denoted by $L = [L_1^T \quad L_2^T \quad 0]^T$, $e_p^0 = x - \hat{x}$, $e_v^0 = v - \hat{v}$ and $e_z^0 = z - \hat{z}$ denote the displacement observation error, velocity observation error and average deflection of the bristles observation error. When lead the friction observer into the system, we must consider the bounded theory of the observation error e_v^0 and e_z^0, the analysis of the observer design can be stated by the proposition of the document [5]. According to boundary of the state $\bar{e}_z^0$, if the values of σ_{0i} representing the structural stiffness of the bristle material are large, then the bound of $||\bar{e}_z^0||$ can be sufficiently small. σ_{0i} values of steel are larger since for the most multiracial stage systems are made by steel, this implies that the ultimate bounds of $\bar{e}_p^0$ and $\bar{e}_v^0$ states can be very small and satisfy the objective of state observation.

239.3.2 XY Table Contouring Controller

Supposing the reference command xd which is at least twice differentiable and tracking error vectors of the system can be denoted by $e_p^t = x - x_d$ and $e_v^t = \dot{e}_p^t = \dot{x} - \dot{x}_d$, respectively. The control system structure diagram is represented as shown in Fig. 239.2.

Where the position reference command can be denoted as x_{d1} and x_{d2}, the biaxial system can be denoted as $\bar{u}_x$ and $\bar{u}_y$, $\bar{f}_{rx}$ and $\bar{f}_{ry}$ is estimation value of the friction observer. So, the overall contouring control input u is written as:

$$u = \hat{F}_r + M\ddot{x}_d + \bar{u} \tag{239.10}$$

Table 239.1 Parameter setting of XY table

Parameters	X axis	Y axis
m (Kg)	5.8	1.4
f_s (N)	10.82	13.61
f_c (N)	7.5	9.5
σ_0 (N/m)	1×10^5	1×10^5
σ_1 (Ns/m)	316.23	316.23
σ_2 (Ns/m)	350	380

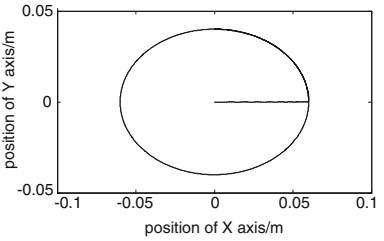


Fig. 239.3 Elliptic desired path and output path

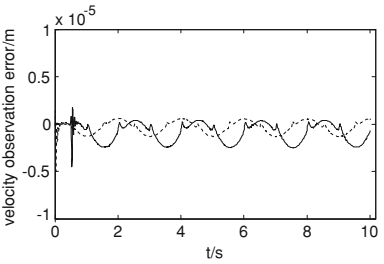


Fig. 239.4 Observation error of x and y axis of the feedrate system

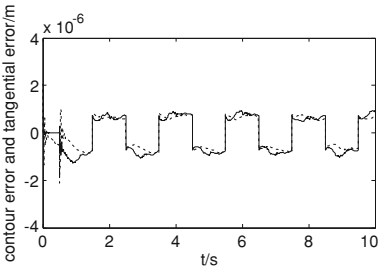


Fig. 239.5 Contour error and tangential error of elliptic path

Where $\bar{u} = [\,\bar{u}_x \quad \bar{u}_y\,]^T$; $\hat{F}_r = \Sigma_0\hat{z} + \Sigma_1\dot{\hat{z}} + \Sigma_2\hat{v}$, the contour error and tangential error dynamic equation can be rewritten as:

$$\bar{u} = MT(2T\dot{T}\hat{\varepsilon}_V + T\ddot{T}\varepsilon_p - K_p\varepsilon_p - K_v\hat{\varepsilon}_V) \tag{239.11}$$

where K_p and K_v denote two constant matrices $\hat{\varepsilon}_v$ denote estimation of ε_v.

239.4 Simulation and Analysis

The X–Y table parameters of X and Y axis were shown in Table 239.1.

Matlab Simulink 7.0 was adopted to do simulation experiment. Fig. 239.3 showed the command track of X and Y axis, when the external disturbance and nonlinear friction both existed.

It can be known that the command track and the actual output track were almost coincident. Figures 239.4 and 239.5 showed respectively the observation error of federate and bristle deflection of the system.

The evolutions of e_v^0 and e_z^0 of Fig. 239.4 are very small because the most multiracial stage system are made by steel and the σ_0 value of steel are large. These results are consistent to the analysis of Proposition1 and reveal the effectiveness of the proposed observer design. Contour error and tangential error of XY table are shown in Fig. 239.5. The contour error converges to a small quantity.

239.5 Conclusions

In this chapter, to reduce the model error, real-time contour error algorithm applicable for multi-axis nonlinear motion system was adopted. The friction observer eliminates the nonlinear dynamic friction forces of the system. For the velocity mismatch between two axes, contour controller is designed by three kinds of state variables that are from the friction observer. Theoretical analyses, numerical simulations results verify the effectiveness of the entire observe-based contouring control scheme.

References

1. Wang XK, Cheng DJ, Wu D (2001) Overview of linear motor used in machine tool feeding system [J]. Manufact Technol Mach Tool 8:18–21
2. Feng L, Koren Y, Borenstein J (1993) Cross-coupling motion controller for mobile robots [J]. IEEE Control Syst Mag 13:35–43
3. Chen SL, Tsai YC (2008) Contouring control of a parallel mechanism based on equivalent errors [C]. In: American control conference, Washington, pp 2384–2388
4. Cheng MY, Lee CG (2007) Motion controller design for contour-following tasks based on real-time contour error estimation [J]. IEEE Trans on Ind Electr 54:1686–1695
5. dewit Canudas C, Olsson H, Astrom KJ (1995) A new model for control of systems with friction [J]. IEEE Trans on Autom Control 40:419–425

Chapter 240
Maximum Torque Control of Permanent Magnet Linear Synchronous Motor Based on the Hamiltonian

Zhiping Cheng and Liucheng Jiao

Abstract Based on the brief introduction of the port-controlled Hamiltonian (PCH) system, the PCH system model of non-salient permanent magnet linear synchronous motor is established. The control laws are presented in load known conditions. Simulation results show that the method of port-controlled Hamiltonian is easy to be realized to control permanent magnet linear synchronous motor (PMLSM) and the physical meaning is clear and the effect is ideal.

Keywords Hamiltonian · Permanent magnet linear synchronous motor · Modeling · Control · Stability

240.1 Introduction

As PMLSM is not periodic in structure, the primary iron core and windings are disconnected. Tests proved that the ideal steady-state model and continued rotating motor control methods are not suitable for PMLSM. Therefore, some advanced control theory have been applied into the control of PMLSM and the abundant achievements have been obtained [1–3].

In recent years, a new nonlinear control system with energy-shaping methods appears [4]. Its main feature is that the system has a port-controlled Hamiltonian structure. According to the peculiar feedback stabilization method, controller

Z. Cheng (✉) · L. Jiao
College of Electric Engineering, Zheng zhou University,
Zheng zhou 450001, China
e-mail: zpcheng@zzu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_240,

design and stability analysis are easier. This method has been used by Professor Yu Hai-Sheng etc. to research the permanent magnet motor [5].

PMLSM can be considered as a two-port energy conversion device. Therefore, the PCH system model of the PMLSM is necessary by energy-shaping methods and passive control, and the control of PMLSM with load constant and known cases is of realistic significance, which also provides a new way for permanent magnet linear driver in the industrial application of AC servo system.

240.2 Basic Forms of Port-Controlled Hamiltonian (PCH) system

PCH system [6] is described as:

$$\dot{x} = f(x) + g(x)u = J(x)\frac{\partial H}{\partial x}(x) + g(x)u, \quad y = h(x) = g^T(x)\frac{\partial H}{\partial x}(x) \tag{240.1}$$

The energy dissipation concept is introduced into the PCH system [7] and the PCH dissipative models:

$$\dot{x} = [J(x) - R(x)]\frac{\partial H}{\partial x}(x) + g(x)u, \qquad y = g^T(x)\frac{\partial H}{\partial x}(x) \tag{240.2}$$

240.3 PCH System Model of Non-Salient Pole of PMLSM

The model of non-salient pole of PMLSM can be expressed as [3]:

$$\begin{aligned} L_d\frac{di_d}{dt} &= u_d - ri_d + \frac{\pi}{\tau}vL_qi_q \\ L_q\frac{di_q}{dt} &= u_q - ri_q - \frac{\pi}{\tau}v\psi_r - \frac{\pi}{\tau}vL_di_d \\ m\frac{dv}{dt} &= \frac{\pi}{\tau}\psi_r i_q - F_L + F_d + f \end{aligned} \tag{240.3}$$

Defined:

$$\begin{aligned} x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} L_d i_d \\ L_q i_q \\ mv \end{bmatrix} = \begin{bmatrix} L_d & 0 & 0 \\ 0 & L_q & 0 \\ 0 & 0 & m \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ v \end{bmatrix} = D \begin{bmatrix} i_d \\ i_q \\ v \end{bmatrix} \\ u = \begin{bmatrix} u_d & u_q & F_d - F_L + f \end{bmatrix}^T \\ y = g^T(x) \frac{\partial H(x)}{\partial x} = \begin{bmatrix} i_d & i_q & v \end{bmatrix}^T \end{aligned} \tag{240.4}$$

Take the sum of electrical energy and mechanical energy as Hamiton function of PMLSM.

$$H(x) = \frac{1}{2} x^T D^{-1} x = \frac{1}{2} \left[\frac{1}{L_d} x_1^2 + \frac{1}{L_q} x_2^2 + \frac{1}{L_q} x_3^2 \right] \tag{240.5}$$

Then the system PCH form of PMLSM can be written as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -r & 0 & \frac{\pi}{\tau} L_q i_q \\ 0 & -r & -\frac{\pi}{\tau}(L_d i_d + \psi_r) \\ 0 & \frac{\pi}{\tau}\psi_r & 0 \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ v \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_d \\ u_q \\ F_d - F_L + f \end{bmatrix} \tag{240.6}$$

$$\begin{aligned} R(x) = \begin{bmatrix} -r & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad g(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ J(x) = \begin{bmatrix} 0 & 0 & \frac{\pi}{\tau} L_q i_q \\ 0 & 0 & -\frac{\pi}{\tau}(L_d i_d + \psi_r) \\ 0 & \frac{\pi}{\tau}\psi_r & 0 \end{bmatrix} \end{aligned} \tag{240.7}$$

the feedback control $u = \alpha(x)$ can be found in order to make the closed-loop system as follows:

$$\dot{x} = [J_d(x) - R_d(x)] \frac{\partial H_d}{\partial x}(x) \tag{240.8}$$

where $J_d(x)$ is interconnected to damping matrix. $R_d(x)$ is damping matrix. In addition:

$$\begin{aligned} J_d(x) = J(x) + J_a(x) = -J_d^T(x) \\ R_d(x) = R(x) + R_a(x) = R_d^T(x) \geq 0 \end{aligned} \tag{240.9}$$

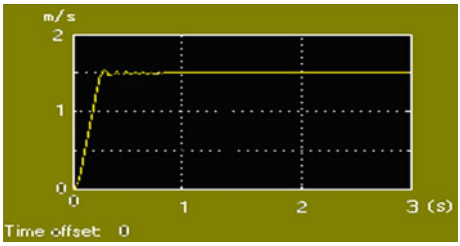


Fig. 240.1 Speed curve (240.1)

240.4 Design and Stability Analysis of System Controller

Take a non-salient PMLSM as example. $L_d = L_q$ in the non-salient PMLSM, the expression of $i_d = 0$ meets "maximum torque control". The expectation point of balance can be got in Eq. (240.6) for a given v_0. According to the "maximum torque" control principle and load constant. There is:

$$x_0 = [x_{10} \quad x_{20} \quad x_{30}]^T = \left[0 \quad \frac{L_q(F_L - F_d - f)}{\frac{\pi}{\tau}\psi_r} \quad mv_0\right]^T \tag{240.10}$$

Take the expected Hamiltonian function:

$$H_d(x) = \frac{1}{2}(x - x_0)^T D^{-1}(x - x_0) \tag{240.11}$$

For

$$\begin{gathered} \frac{\partial H_d(x)}{\partial x} = D^{-1}(x - x_0) \quad \frac{\partial^2 H_d(x)}{\partial x^2} = D^{-1} \\ \frac{\partial H(x)}{\partial x} = D^{-1}(x), \quad K(x) = \frac{\partial_a H(x)}{\partial x} = \frac{\partial_d H(x)}{\partial x} - \frac{(x)}{\partial x} = -D^{-1}(x_0) \end{gathered} \tag{240.12}$$

Equation (240.12) met the positive definite. So the closed-loop system designed is stable at the balance point. Take:

$$J_a(x) = \begin{bmatrix} 0 & -J_{12} & -J_{13} \\ J_{12} & 0 & -J_{23} \\ J_{13} & J_{23} & 0 \end{bmatrix}, \quad R_a(x) = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & r_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{240.13}$$

where J_{12}, J_{13}, J_{23}, R_1, R_2 are undetermined interconnections and damping parameters respectively. Substitute Eq. (240.12) into Eq. (240.13), there is:

$$-[J_d(x) - R_d(x)]D^{-1}x_0 = -[J_a(x) - R_a(x)]D^{-1}x + g(x)\alpha(x) \tag{240.14}$$

The expression of u can be obtained from Eq. (240.12):

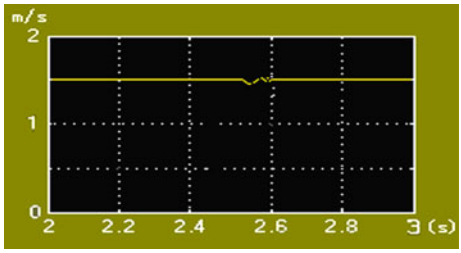


Fig. 240.2 Speed curve (240.2)

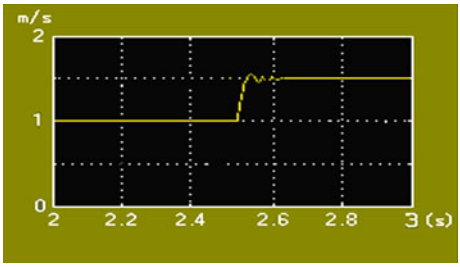


Fig. 240.3 Speed curve (240.4)

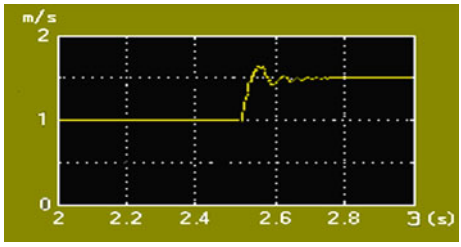


Fig. 240.4 Speed curve (240.5)

$$u = \alpha(x) = [\, u_d \quad u_q \quad -\tau_L \,]^T \tag{240.15}$$

Substitute Eqs. (240.9) and (240.13) into Eq. (240.14), Eq. (240.20) can be deduced with Eq. (240.10).

$$-[J_d(x) - R_d(x)]D^{-1}x_0 = -[J_a(x) - R_a(x)]D^{-1}x + g(x)\alpha(x) \tag{240.16}$$

As

$$u = \alpha(x) = [\, u_d \quad u_q \quad F_d - F_L + f \,]^T \tag{240.17}$$

$$\left\{\begin{bmatrix}0 & 0 & \frac{\pi}{\tau}L_q i_q\\ 0 & 0 & -\frac{\pi}{\tau}(L_d i_d+\psi_r)\\ 0 & \frac{\pi}{\tau}\psi_r & 0\end{bmatrix}+\begin{bmatrix}0 & -J_{12} & -J_{13}\\ J_{12} & 0 & -J_{23}\\ J_{13} & -J_{23} & 0\end{bmatrix}\right\}\begin{bmatrix}x_{10}\\ x_{20}\\ x_{30}\end{bmatrix}=$$
$$\left\{-\begin{bmatrix}-r & 0 & 0\\ 0 & -r & 0\\ 0 & 0 & 0\end{bmatrix}-\begin{bmatrix}-r_1 & 0 & 0\\ 0 & -r_2 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{L_d} & 0 & 0\\ 0 & \frac{1}{L_q} & 0\\ 0 & 0 & \frac{1}{m}\end{bmatrix}\right\}$$
$$\left\{\begin{bmatrix}0 & -J_{12} & -J_{13}\\ J_{12} & 0 & -J_{23}\\ J_{13} & J_{23} & 0\end{bmatrix}-\begin{bmatrix}r_1 & 0 & 0\\ 0 & r_2 & 0\\ 0 & 0 & r_3\end{bmatrix}\right\}\times\begin{bmatrix}\frac{1}{L_d} & 0 & 0\\ 0 & \frac{1}{L_d} & 0\\ 0 & 0 & \frac{1}{m}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}-\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}u_d\\ u_q\\ F_d-F_L+f\end{bmatrix} \tag{240.18}$$

There is Eq. (240.19) from Eq. (240.18):

$$\frac{J_{13}}{L_d}x_{10}+\frac{J_{23}+\frac{\pi}{\tau}\psi_r}{L_q}x_{20}=\frac{J_{13}}{L_d}x_1+\frac{J_{23}}{L_q}x_2-(F_d-F_l+f) \tag{240.19}$$

x_0 is substituted into Eq. (240.19).Then:

$$J_{23}=-\frac{L_q x_1}{L_d(x_2-x_{20})}J_{13} \tag{240.20}$$

Take:

$$J_{13}=\frac{L_d}{L_q}(x_2-x_{20})\quad J_{23}=-x_1\quad J_{12}=-Kx_3 \tag{240.21}$$

K is a arbitrary parameter. It is easy to prove that K does not affect the stability of system. So the controller will be the follow:

$$\begin{aligned}u_d&=-r_1 i_d+kmvi_q-\left(\frac{\pi L_q}{m\tau}\right)+\frac{\tau K L_q(F_l-F_d-f)}{\pi\psi_r}i_q\\ u_q&=\left(\frac{\pi v_0 L_d}{\tau}-L_d v_0\right)i_d+(L_d-km)vi_d-r_2 i_q+\frac{\pi\psi_r v_0}{\tau}\end{aligned} \tag{240.22}$$

240.5 Simulation Results and Analysis

SPWM is used. Simulation voltage is 35 V. Inverter frequency is 4 Hz. $K = 18$, Model hysteresis precision is 0.025. Friction is equivalent to the load resistance.

Figure 240.1 is the starting speed simulation curve during 3 s with $r1 = r2 = 1.3$. The maximum overshoot is about 2.8% and response time is about 0.35 s. Figure 240.2 is the speed response curve, in which the speed of rotor is 1.5 m/s, the load disturbance of 75 N is added at 2.5 s. Figure 240.3 is speed

curve, which the speed of rotor is 1 m/s. The speed disturbance of 0.5 m/s is added at 2.5 s. For verification system performance, a conventional PID controller is given, as shown in Figure 240.4 with the same conditions. We can come to a conclusion that the control of PMLSM based on energy can significantly reduce system parameter variations and external disturbances. The system has stronger robustness and practicability.

Acknowledgments It is a project supported by the National Natural Science Foundation of China (61075071), the Natural Science Foundation of Henan Province (2010A470006/2011A470010). Zhengzhou Science and Technology Bureau, Henan Province (10PTGG380-1).

References

1. Lan Y, (2007) Study on Robust Control for Permanent Magnet Linear Motor Servo System. Shenyang University of Technology, China (in Chinese)
2. Guan L, Sun S (2009) Research on adaptive slide mode variable structure direct thrust speed control for permanent magnet linear motor. Trans Shenyang Ligong Univ (02):45–47 (in Chinese)
3. Zhu X, Zeng Q, Jiao L (2006) Study and simulation to PMLSM direct torque control. Min Process Equip 34(5):88–91 (in Chinese)
4. Wang Y, Cheng D, Li C et al (2003) Dissipative Hamiltonian realization and energy-based l2-disturbance attenuation control of multimachine power systems. IEEE Trans Autom Control 48(8):1428–1433
5. Yu H, Wang H, Zhao K (2006) Hamiltonian modeling and IDA passivity-based control of permanent magnet synchronous motor. Electr Mach Control 10(3):229–233 (in Chinese)
6. Van Der Schaft AJ (2000) L2-Gain and Passiving Techniquess in Nonlinear Control. Springer, London
7. Ortega R, Van der Schaft AJ, Mareels I et al (2002) Interconnection and damping assignment passivity-based contron of port-controlled Hamiltionian systems. Automatica 38(4):585–596

Chapter 241
Calculation of Winding Factor and Analysis of Armature MMF of a PMSM with Fractional-Slot and 5-Phase Winding

Shangguan Xuanfeng, Zhang Jiaolong and Zhang Wenli

Abstract In this paper, a permanent magnet synchronous machine with fractional-slot and 5-phase winding is studied. After the superposition of the back EMF (electro-motive force) phasors of all coil-sides in a phase winding, the winding factor of the machine is calculated. The method has a clear idea and is fast. On the basis of the calculating armature MMF, the contents of its harmonics are analyzed. For three kinds of coil spans, the comparisons of the winding factors and MMF harmonics are made. The drawn results in this chapter can be applied to guide machine designs.

Keywords Winding factor · Magnetic-motive force · Fractional-slot · 5-phase winding

241.1 Introduction

Fractional-slot permanent magnet synchronous machines have been gaining interest over the past decade [1, 2]. The machines have several advantages such as high-power density, high efficiency, low cogging torque, short end winding, high

S. Xuanfeng · Z. Jiaolong (✉)
School of Electrical Engineering and Automatization,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zjl@hpu.edu.cn

S. Xuanfeng
e-mail: sgxf@hpu.edu.cn

Z. Wenli
Li Feng Power Plant, Henan Province Coal and Chemical Group,
Jiaozuo 454000, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_241,

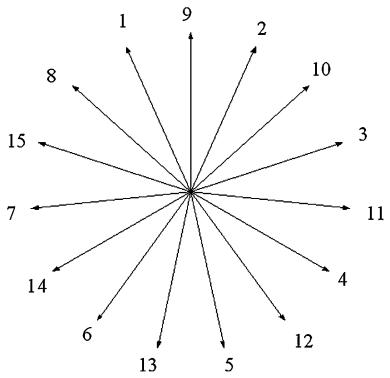


Fig. 241.1 Star of slot

flux-weakening capability. Multiphase motor drives possess many advantages over the traditional 3-phase motor drives [3, 4], these include fault tolerant and reducing the amplitude of torque pulsation, reducing the stator current per phase without increasing the voltage per phase. So, Multiphase motor drives are in applications where high reliability is demanded such as electric/hybrid vehicles, aerospace applications, ship propulsion and high power applications. This chapter will discuss a five-phase, fractional-slot, surface-mounted permanent magnet synchronous machine with 15 slots and 4 poles. Three kinds of coil spans are used, they are 4-solt pitch, 3-solt pitch, and 2-solt pitch. This chapter covers the calculation of winding factors and armature MMFs, harmonic analysis of armature MMFs.

241.2 Star of Slot and Winding Layout

The prototype is a permanent magnet machine with 4 poles, 15 slots, and 5-phase double-layer winding. The surface-mounted permanent magnet poles on the rotor are radially magnetized. The electrical angle between adjacent two slots is obtained,

$$\alpha = \frac{p \cdot 360^\circ}{Q} = \frac{720^\circ}{15} = 48^\circ$$

Where, p is the number of pole pairs, Q is the number of the stator slots.

The number of slot per pole per phase

$$q = \frac{Q}{2pm} = \frac{15}{2 \times 2 \times 5} = 0.75$$

The star of slot is shown as Fig. 241.1.

Three kinds of coil spans adopted are respectively 4, 3 and 2 slot-pitches. By the star of slot, winding layouts are determined in Fig. 241.2.

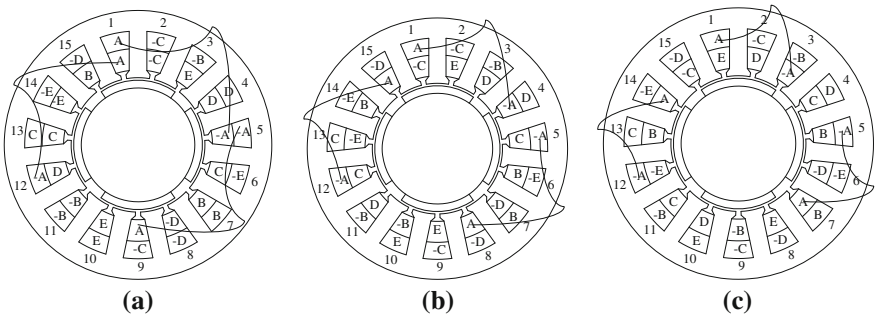


Fig. 241.2 Winding layouts of five-phase SMPM motors with 15 stator slots **a** Coil span = 4. **b** Coil span = 3. **c** Coil span = 2

241.3 Calculation of Winding Factor

Winding factor depends on three sub factors (pitch factor, distribution factor and skewing factor). Each of the sub factors can be easily calculated for conventional integer slot distributed windings with well-known formulas. However, these factors are hard to visualize in fractional-slot windings. In this chapter, the method to calculate winding factor by the back EMF of each coil-side from the star of slot has some advantages such as ideas are clear, calculation easy.

241.3.1 Winding Factor Calculation of 4-Solt Span Coil

According to the each coil layout of phase-*A* winding shown in Fig. 241.2a, the upper coil side in slot-1 and the lower coil side in slot-5 form the first coil of phase-*A* winding, the upper coil side in slot-5 and the lower coil side in slot-9 form the second coil, the upper coil side in slot-12 and the lower coil side in slot-1 form the third coil. It is shown the phase angle difference between the two sides of each coil is 192° electrical degree ($4\alpha = 4 \times 48^\circ = 192^\circ$). The three coils of each phase winding are in series. The back EMF phasors of six coil sides of phase-*A* winding are shown in Fig. 241.3a. The back EMF of the first coil is the resultant of $\dot{E}_{1A+}$ and $\dot{E}_{5A-}$. The back EMF of the second coil is the resultant of $\dot{E}_{5A-}$ and $\dot{E}_{9A+}$. The back EMF of the third coil is the resultant of $\dot{E}_{12A-}$ and $\dot{E}_{1A+}$. The phase angle differences of the back EMF of the coils are 12° electrical degree, not in the same phase. So, it isn't feasible that the three coils are in parallel.

The back EMF phasor of *i*th coil side is expressed with $\dot{E}_i$, and then the winding factor of phase-*A* winding can be calculated (the three coils of each phase winding in series)

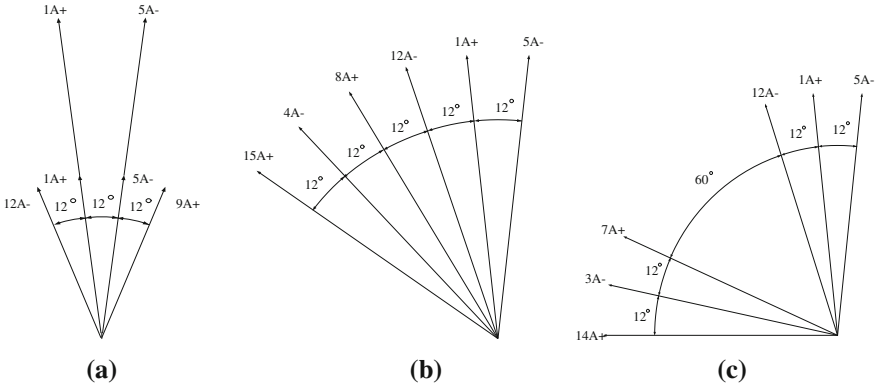


Fig. 241.3 Back EMF phasors of the coil sides in phase-*A* winding **a** Coil span = 4. **b** Coil span = 3. **c** Coil span = 2

$$k_{w1} = \frac{\left|\sum_{i=1}^{6} \dot{E}_i\right|}{6E} = \frac{\left(\sqrt{2 \times 2^2 + 2 \times 2^2 \times \cos 12^\circ} + \sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 36^\circ}\right)E}{6E} = 0.98$$

241.3.2 Winding Factor Calculation of 3-Solt Span Coil

According to Fig. 241.2b, the upper coil side in slot-1 and the lower coil side in slot-4 form the first coil of phase-*A* winding. The upper coil side in slot-5 and the lower coil side in slot-8 form the second coil. The upper coil side in slot-12 and the lower coil side in slot-15 form the third coil. It is shown the phase angle difference between the two sides of each coil is 144° electrical degree ($3\alpha = 3 \times 48^\circ = 144^\circ$). The back EMF phasors of the six coil sides in phase-*A* winding is shown in Fig. 241.3b. The back EMF of the first coil is the resultant of $\dot{E}_{1A+}$ and $\dot{E}_{4A-}$. The back EMF of the second coil is the resultant of $\dot{E}_{5A-}$ and $\dot{E}_{8A+}$. The back EMF of the third coil is the resultant of $\dot{E}_{12A-}$ and $\dot{E}_{15A+}$. In Fig. 241.3b, the phase angle differences of the back EMF of the coils are 12° electrical degree.

Winding factor of phase-*A* winding can be obtained

$$k_{w1} = \frac{\sum_{i=1}^{6} \dot{E}_i}{6E} = \frac{\left(\sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 12^\circ} + \sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 36^\circ} + \sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 60^\circ}\right)E}{6E} = 0.9371$$

241.3.3 Winding Factor Calculation of 2-Solt Span Coil

According to Fig. 241.2c, the upper coil side in slot-1 and the lower coil side in slot-3 form the first coil. The upper coil side in slot-5 and the lower coil side in slot-7 form the second coil. The upper coil side in slot-12 and the lower coil side in slot-14 form the third coil. It is shown the phase angle difference between the two sides of each coil is 96° electrical degree ($2\alpha = 2 \times 48^\circ = 96^\circ$). The back EMF phasors of the coil sides of phase-*A* winding are shown in Fig. 241.3c. The back EMF of the first coil is the resultant of $\dot{E}_{1A+}$ and $\dot{E}_{3A-}$. The back EMF of the second coil is the resultant of $\dot{E}_{5A-}$ and $\dot{E}_{7A+}$. The back EMF of the third coil is the resultant of $\dot{E}_{12A-}$and $\dot{E}_{14A+}$. In Fig. 241.3c, the phase angle differences of the back EMF of the coils are 12° electrical degree.

Winding factor of phase-*A* winding can be obtain

$$\begin{aligned} k_{w1} &= \frac{\sum_{i=1}^{6} \dot{E}_i}{6E} \\ &= \frac{\left(\sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 60^\circ} + \sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 84^\circ} + \sqrt{2 \times 1^2 + 2 \times 1^2 \times \cos 108^\circ}\right)E}{6E} \\ &= 0.7322 \end{aligned}$$

The highest winding factor can be obtained with 4-solt span coil (the coil span is closest to pole pitch 3.75), the lowest 2-solt coil span.

241.4 MMF of Armature Reaction

241.4.1 Armature Currents and MMF of 5-Phase Machines

Armature currents are expressed by

$$\begin{cases} i_A = I_m \sin(\omega t) \\ i_B = I_m \sin(\omega t - 2\pi/5) \\ i_C = I_m \sin(\omega t - 4\pi/5) \\ i_D = I_m \sin(\omega t + 4\pi/5) \\ i_E = I_m \sin(\omega t + 2\pi/5) \end{cases} \tag{241.1}$$

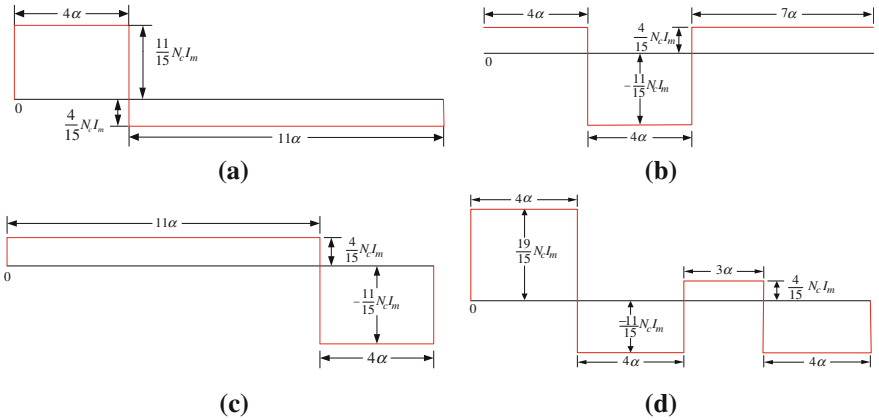


Fig. 241.4 MMF of phase-A winding. **a** Coil-1, **b** Coil-2, **c** Coil-3, **d** Resultant MMF

The armature MMF of 5-phase machines

$$f(\theta, i) = f_A(\theta, i_A) + f_B(\theta, i_B) + f_C(\theta, i_C) + f_D(\theta, i_D) + f_E(\theta, i_E) \tag{241.2}$$

The MMF of phase-M winding (M is A, B, C, D, E)

$$f_M(\theta, i_M) = \sum_{i=1}^{n} f_{Mi}(\theta, i_M) \tag{241.3}$$

Where, $f_{Mi}(\theta, i_M)$ is the MMF of i-coil in phase-M winding, n is the number of the coils in per phase winding.

$$f_{Mi}(\theta, i_M) = \begin{cases} \dfrac{(Q-S)}{Q} i_M N_c & \theta \in (\theta_{i1} \sim \theta_{i2}) \\ \dfrac{S}{Q} i_M N_c & \text{other} \end{cases} \tag{241.4}$$

Where, N_c is the turn number of per coil, S is coil span in number of slots, i_M is the current of phase-M winding. $\theta_{i1}$and θ_{i2} express the angular positions of the two sides of the i-coil.

241.4.2 MMF of Phase-A Winding

At $\omega t = \pi/2$, phase-A current reaches the biggest value.

$$i_A = I_m,\ i_B = 0.309 I_m,\ i_C = -0.809 I_m,\ i_D = -0.809 I_m,\ i_E = 0.309 I_m.$$

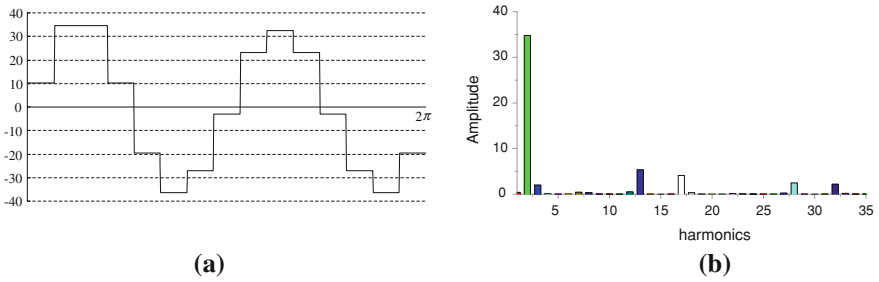


Fig. 241.5 Resultant MMF and its harmonic contents of 5-phase winding (coil span is 4). **a** Armature MMF, **b** Harmonics

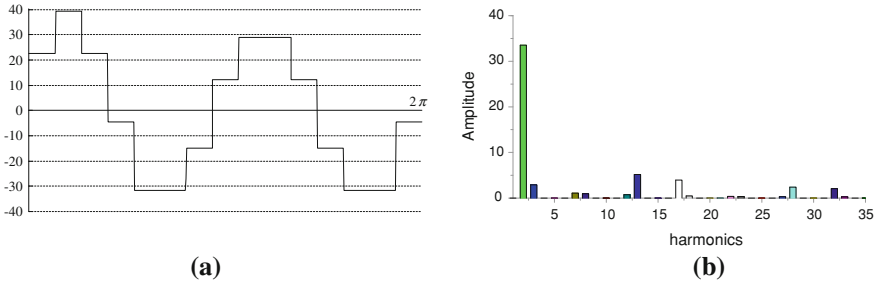


Fig. 241.6 Resultant MMF and its harmonic contents of 5-phase winding (coil span is 3). **a** Armature MMF, **b** Harmonics

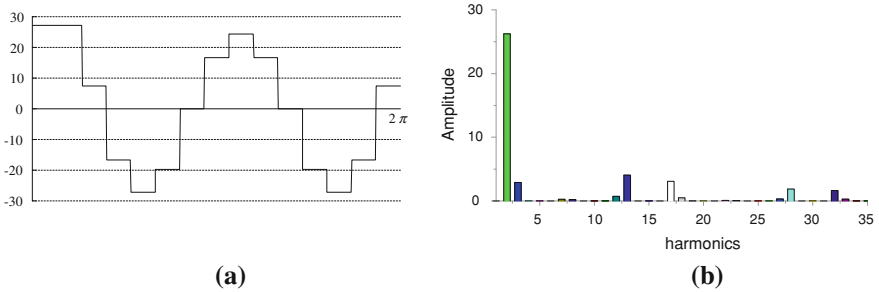


Fig. 241.7 Resultant MMF and its harmonic contents of 5-phase winding (coil span is 2). **a** Armature MMF, **b** Harmonics

The MMFs of three coils in phase-*A* winding are expressed respectively in Fig. 241.4a–c. The resultant MMF of the three coils is expressed in Fig. 241.4d. In Similar way, The MMFs of the other four phase windings can be obtained.

241.4.3 Resultant MMF of the 5-Phase Windings

Figure 241.5a and b show respectively the resultant MMF and its harmonics. Per unit of vertical axis expresses $N_c I_m/15$ ampere-turns.

When coil span are 3 and 2, the armature reaction MMFs and the harmonic contents of the MMFs are shown in Fig. 241.6 and Fig. 241.7.

The first order harmonic has one cycle along air-gap, main harmonic has two cycles. The amplitude of main harmonic is largest for 4-solt span coil, smallest for 2-solt span coil. Main harmonic is the component of the working wave. There are other influential harmonics such as thirteenth harmonic, seventeen harmonic, and twenty-eighth harmonic. Though the amplitude of the main harmonic is slightly smaller for 3-solt span coil than 4-solt span coil, 3-solt span coils has shorter end winding, reducing the use of wire and copper loss. 2-solt span coil is not a good choice.

241.5 Conclusion

This paper calculates the winding factors by the back EMF of each coil-side. The idea is clear and operation convenient. The armature reaction MMFs are calculated, the harmonic contents are analyzed. For three kinds of coil spans, the comparisons of winding factors and MMF harmonics are made. The suitable winding layout can be selected.

References

1. EL-Refaie AM (2010) Fractional-slot concentrated-windings synchronous permanent magnet machines opportunities and challenges. IEEE Trans Ind Electron 57(1):107–121
2. Bianchi N, Fornasiero E (2009) Impact of MMF space harmonic on rotor losses in fractional-slot permanent-magnet machines. IEEE Trans Energy Convers 24(2):323–328
3. Parsa L, Toliyat H, Goodarzi A (2007) Five-phase interior permanent-magnet motors with low torque pulsation. IEEE Trans Ind Appl 43(1):40–46
4. Bianchi N, Bolognani S, Pré MD (2008) Impact of stator winding of a five-phase permanent-magnet motor on postfault operations. IEEE Trans Ind Appl 55(5):1978–1987

Chapter 242
Neural Modeling and Control of Hysteretic Dynamics in Ultrasonic Motors

Xinliang Zhang, Jingjing Du and Yonghong Tan

Abstract Ultrasonic motors operate on the inverse piezoelectric effect and the frictional driven force. Complex nonlinear characteristics such as hysteresis and friction nonlinearities are involved. Due to the inherent undifferentiable property and complex contact mechanism, to design a high performance controller is a real challenge. To combat this problem, a novel neural modeling and control strategy is presented. The motor velocity is adopted to allow for the nonlinear friction effect, and the multi-valued hysteresis of the motor is transformed into a one-to-one mapping with the introduction of a special hysteretic operator based on the expanded input space method. Based on the neural model, an adaptive neural controller is designed, where a so-called generalized gradient of the hysteretic operator is introduced to get the sensitivity information. Finally, the simulation results are presented for the validation of the proposed approach.

Keywords Ultrasonic motor · Hysteresis · One-to-one mapping · Generalized gradient · Neural networks

X. Zhang (✉) · J. Du
School of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo 454000, China
e-mail: zxldq@hpu.edu.cn

J. Du
e-mail: hzdujing@163.com

Y. Tan
College of Mechanical and Electronic Engineering,
Shanghai Normal University, Shanghai 201814, China
e-mail: tany@shnu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_242,

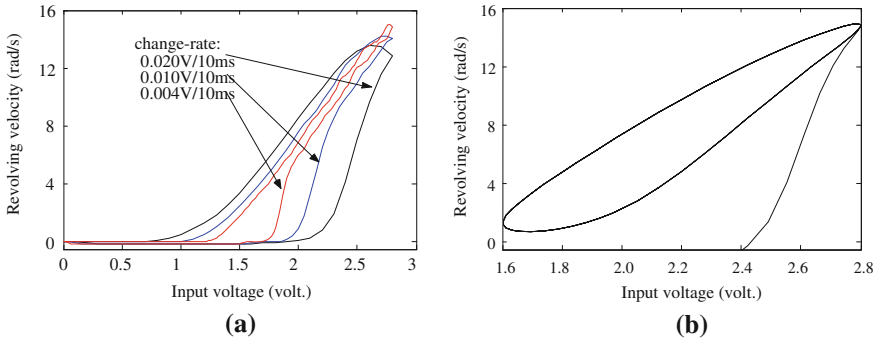


Fig. 242.1 **a** Rate-dependent property of RTWUSM. **b** The structure of a Wiener model

242.1 Introduction

The ultrasonic motor (USM) exhibits excellent advantages compared to the conventional AC/DC servo-motors, such as high retention torque, no electro-magnetic interferences and quick response etc. It has been used widely in ultra-precision servo-systems. Its modeling and control have become one of the hot topics in engineering [1]. The rotary traveling wave ultrasonic motor (RTWUSM) operates based on the ultrasonic vibration force, i.e. a traveling bending wave is generated on the stator surface due to the inverse piezoelectric effect and the rotor is driven by frictional forces generated by the contact layer due to the preload force [2].

Until today, there have been some theoretical or experimental schemes for modeling the RTWUSM such as Hamilton's variation principle [3] or the finite element analysis [4]. However, those methods are rather complex in computation since a lot of parameters in models need to be determined. Reference [5] had studied the online modeling and control for the RTWUSM using the neural networks. However, the hysteresis within the motor is not considered in the paper.

Due to the fact that the driven force is generated from the inverse piezoelectric effect, there exists an obviously hysteretic phenomenon with RTWUSM. Also, under the effect of the friction in the contact layer, the hysteresis shows a special behavior. As shown in Fig. 242.1a, when it is excited by the triangular wave-type voltages with different change-rates, a rate-dependent property is observed as well as the dead-zone nonlinearity. Moreover, when a sinusoidal excitation voltage is fed, an asymmetric characteristic is presented as shown in Fig. 242.1b. These characteristics of the hysteresis in the RTWUSM differ from the Preisach-type hysteresis [6].

Recently, some researches have considered the neural modeling approaches for hysteresis [7–9]. Reference [7] employed two neural networks to approximate the ascent and descent sub-loops of the hysteresis respectively. Reference [8] proposed a hysteretic operator from the classical Preisach model to decompose the multi-valued mapping of the hysteresis into a one-to-one mapping. However, these

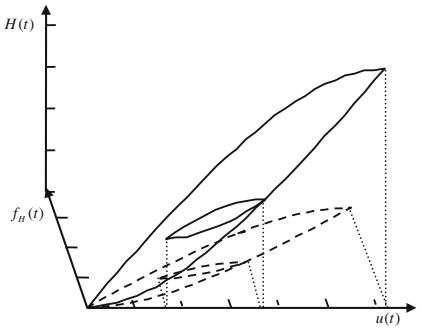


Fig. 242.2 The relation between the hysteretic operator and the hysteresis

models were mainly used to model the rate-independent hysteresis; it did not include the rate-dependent effect of the hysteresis. Reference [9] proposed a neural hysteresis model based on the expanded input space, where a generalized gradient to detract the moving direction of hysteresis and the input derivative to represent the input frequency effect were included.

In this paper, a neural model is proposed to describe the phenomenon of rate-dependent and asymmetric hysteresis existing in the RTWUSM. Because the hysteresis is a nonlinear system with multi-valued mapping and memory, the conventional neural network is impossible to model its behavior directly. Hence, an expanded input space method proposed in [10] is employed. With the introduction of a special hysteretic operator, the multi-valued mapping of the hysteresis is transformed into a one-to-one mapping. Simultaneously, to compensate the asymmetric hysteretic phenomenon in the RTWUSM, the revolving velocity is introduced as an additional input component for the model. Then, based on the expanded input space, a neural model of the RTWUSM is derived. Subsequently, an adaptive neural controller is designed whose weights are updated according to the neural model by use of the so-called generalized gradient method.

242.2 Neural Model of the Hysteresis

It had been proved theoretically that the neural networks are sufficiently complete to approximate any continuous function with reasonable accuracy. However, the neural networks can not be used directly to approximate the hysteresis which is characterized of multi-valued mapping, undifferentiable property and memory. To transform the multi-valued hysteresis into a one-to-one mapping, a special hysteretic operator is introduced as shown in Fig. 242.2. The hysteretic operator describes the 'shape' of hysteresis $H(t)$ on an additional plane (i.e. $u(t)$–$fH(t)$ plane), such as the change-tendency at the ascending and descending sub-loops and the turning points. Then with the introduction of such hysteretic

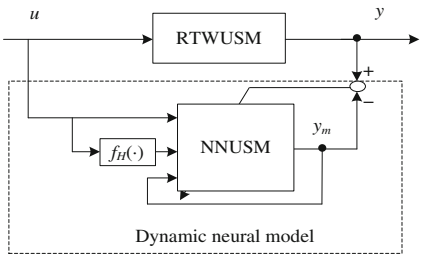


Fig. 242.3 the neural model structure for the RTWUSM

operator, the output of the hysteresis is uniquely determined on an expanded input space which is composed of $fH(t)$ and $u(t)$.

The introduced hysteretic operator can be described as,

$$f_H[u](t) = \big(u - u_p\big)\big(1 - \exp(-|u - u_p|)\big) + f_H[u_p] \tag{242.1}$$

where $u(t)$ is the current input; $f_H[u](t)$ is the current output of the operator; $f_H[u_p]$ is the output of the operator at the current dominant extremum u_p.

It can be proven that if the introduced hysteretic operator keeps a wiping-out property [6], it is a continuous function to input $u(t)$. Also, the multi-valued mapping $H(u(t))$ between the output of hysteresis and its input can be transformed into a one-to-one mapping $\Gamma(u(t), f_H(t))$ on an expanded input space with the introduced hysteretic operator [10]. Moreover, the obtained mapping is a continuous one-to-one mapping on a compact set; the feedforward neural networks can be used to approximate this mapping, i.e.

$$\|\Gamma(u(t), f_H[u](t)) - NN(u(t), f_H[u](t))\| < \varepsilon \tag{242.2}$$

where $NN(\cdot)$ is the neural network realization of the approximation of the one-to-one mapping. ε is the approximation error and for any $\varepsilon_N > 0$, there exists $\varepsilon < \varepsilon_N$.

It should be noted that the hysteretic behavior of the RTWUSM exhibits rate-dependent and asymmetric performance due to the interaction of the components in the RTWUSM as shown in Fig. 242.1a and b. In the case where the input voltage is lower than a certain value, the speed response of the motor presents 'dead-zone' nonlinearity. On the other hand, when the motor is excited by a sinusoidal-wave voltage, e.g. $u(t) = 2.2 + 0.6\sin(2.5t)$, the up and down sub-loops of the input–output curve takes an asymmetry behavior at low velocities due to the nonlinear friction force. Although there have been some research results for modeling the friction, there is a common consensus that the friction is a static nonlinear function of velocity [11]. So in this paper, the revolving velocity of the RTWUSM is adopted as an additional input component to compensate for the effect of the friction resulting in the asymmetry. Then the neural model of RTWUSM is derived as shown in Fig. 242.3.

The output of the model can be described mathematically as,

$$y_m(k) = N\text{NUSM}(u(k), f_H(k), y_m(k-1)) \tag{242.3}$$

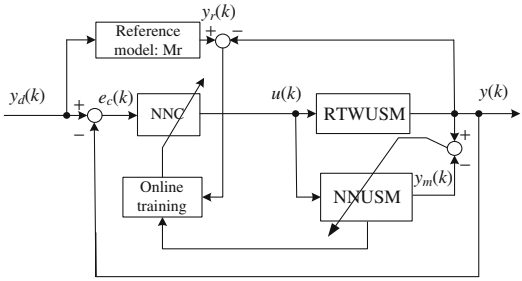


Fig. 242.4 The on-line modeling and control scheme of RTWUSM

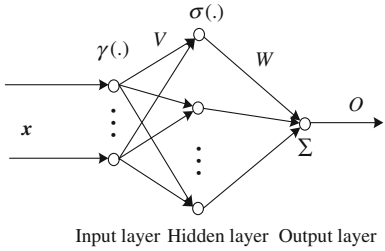


Fig. 242.5 The architecture of the used BP neural networks

where u is the input of the motor; y is the velocity output, while $f_H(k)$ is the introduced hysteretic operator as discussed above.

242.3 Online Neural Modeling and Control of RTWUSM

Due to the driving principle of the RTWUSM, a complex mathematical analysis is usually involved for the construction of an exact model. Thereby, it makes the design of modern controllers such as adaptive, variable structure and robust controllers a very tough problem. To overcome this problem, an adaptive neural controller is designed as shown in Fig. 242.4. In the proposed control system, the sensitivity information of the NNC is provided by the neural model of RTWUSM from Eq. (242.3), which is identified online to accommodate the nonlinear uncertainties.

The structure of the neural networks is shown in Fig. 242.5, which is composed of n_1 input nodes, n_2 hidden nodes and one output node. $\gamma(\cdot)$ and $\sigma(\cdot)$ are the active functions for the input and hidden layers respectively, which are both chosen as Logsig functions. The output of the neural networks is expressed by,

$$O = \boldsymbol{W}^{\mathrm{T}}\bar{\boldsymbol{\sigma}}\left(\boldsymbol{V}^{\mathrm{T}}\bar{\boldsymbol{\gamma}}(\boldsymbol{x})\right) \tag{242.4}$$

where $\boldsymbol{W} = [W_j]^{\mathrm{T}}$ is the weight matrix between hidden and output layers; $\boldsymbol{V} = [V_{ij}],\ i = 0, 1, \ldots, n_1,\ j = 0, 1, \ldots, n_2$, is the weight matrix between input and hidden layers; $\bar{\sigma}(\cdot) = [1, \boldsymbol{\sigma}(\cdot)]^{\mathrm{T}} = [1, \sigma_1(\cdot), \sigma_2(\cdot), \ldots, \sigma_{n_2}(\cdot)]^{\mathrm{T}}$ is the output of the

hidden nodes; $\bar{\gamma}(\cdot) = [1, \boldsymbol{\gamma}(\cdot)]^{\mathrm{T}} = [1, \gamma_1(\cdot), \gamma_2(\cdot), \ldots, \gamma_{n_1}(\cdot)]^{\mathrm{T}}$ is the output of the input nodes; $\boldsymbol{x} = [x_1, x_2, \ldots, x_{n_1}]^{\mathrm{T}}$ is the input vector.

Given the cost function of the approximation process as,

$$E(k) = \frac{1}{2}(O_d(k) - O(k))^2 = \frac{1}{2}e^2(k) \tag{242.5}$$

The weights are updated according the following equations,

$$W(k+1) = W(k) + \Delta W(k) = W(k) - \eta_j \cdot \frac{\partial E(k)}{\partial W(k)} \tag{242.6}$$

$$\boldsymbol{V}_i(k+1) = \boldsymbol{V}_i(k) + \Delta \boldsymbol{V}_i(k) = \boldsymbol{V}_i(k) - \eta_i \cdot \frac{\partial E(k)}{\partial \boldsymbol{V}_i(k)}, i = 0, 1, 2, \ldots, n_1 \tag{242.7}$$

where η_j and η_i are the learning rates.

Define a discrete Lyapunov function as,

$$V(k) = \frac{e^2(k)}{2} \tag{242.8}$$

Considering the property of the Logsig function, it derives that when the learning rates η_j and η_i are chosen as [5],

$$\eta_j = \frac{1}{n_2 + 1} \tag{242.9}$$

$$\eta_i = \frac{16}{n_1 n_2 \|\boldsymbol{W}^*\|_\infty^2 \left(\|\boldsymbol{\gamma}\|_\infty^2 + 1\right)} \tag{242.10}$$

where $\boldsymbol{W}^* = [\boldsymbol{W}_j]^{\mathrm{T}}, j = 1, \ldots, n_2$, there exists,

$$\Delta V(k+1) = V(k+1) - V(k) < 0 \tag{242.11}$$

i.e. the convergence of the weights tuning can be guaranteed.

Define the cost function of the control system as,

$$E_m(k) = \frac{1}{2}(y_r(k) - y(k))^2 = \frac{1}{2}e_r^2(k) \tag{242.12}$$

where $y_r(k)$ and $y(k)$ are the reference signal and actual output of motor respectively.

According to the universal approximation of the neural model for the RTWUSM, it gives,

$$y_m(k) \cong y(k) \tag{242.13}$$

where $y_m(k)$ is the output of the neural model.

Then the sensitivity information for updating the weights between the hidden and output layers in the neural controller is derived as,

$$\frac{\partial E_m(k)}{\partial \boldsymbol{W}^c} = \frac{\partial E_m(k)}{\partial u(k)} \frac{\partial u(k)}{\partial \boldsymbol{W}^c} \simeq -e_r(k) \frac{\partial y_m(k)}{\partial u(k)} \frac{\partial u(k)}{\partial \boldsymbol{W}^c} \tag{242.14}$$

where,

$$\frac{\partial y_m(k)}{\partial u(k)} = \frac{\partial y_m(k)}{\partial O^m(k)} \frac{\partial O^m(k)}{\partial \boldsymbol{\sigma}^m(k)} \frac{\partial \boldsymbol{\sigma}^m(k)}{\partial \boldsymbol{\gamma}^m(k)} \left(\frac{\partial \boldsymbol{\gamma}^m(k)}{\partial x_1^m(k)} + \frac{\partial \boldsymbol{\gamma}^m(k)}{\partial x_2^m(k)} \frac{\partial x_2^m(k)}{\partial u(k)} \right) \tag{242.15}$$

and $x_1^m(k) = u(k)$, $x_2^m(k) = f_H(k)$ is the input components of the neural model.

For the local Lipschitz function [12] $f_H[u](k)$ as described by Eq. (242.1), an generalized gradient is introduced to search its changing tendency [9], i.e.

$$\partial f_H[u](k) = \begin{cases} \dfrac{f_H[u](k) - f_H[\bar{u}](k)}{u(k) - \bar{u}(k)}, & \text{if} (u(k) - u(k-1))(u(k-1) - u(k-2)) < 0 \\ \nabla f_H[u](k), & \text{else} \end{cases} \tag{242.16}$$

$$\nabla f_H[u](t) = \frac{f_H[u](k+1) - f_H[u](k)}{u(k+1) - u(k)} \tag{242.17}$$

$$\bar{u}(k) = \frac{u(k) + u(k-1)}{2} \tag{242.18}$$

Then we get,

$$\frac{\partial E_m(k)}{\partial \boldsymbol{W}^c} \cong -e_r(k) \varsigma \frac{\partial O^c(k)}{\partial \boldsymbol{W}^c} \tag{242.19}$$

$$\varsigma = \frac{\partial y_m(k)}{\partial O^m(k)} \frac{\partial O^m(k)}{\partial \boldsymbol{\sigma}^m(k)} \frac{\partial \boldsymbol{\sigma}^m(k)}{\partial \boldsymbol{\gamma}^m(k)} \left(\frac{\partial \boldsymbol{\gamma}^m(k)}{\partial x_1^m(k)} + \frac{\partial \boldsymbol{O}_1^m(k)}{\partial x_2^m(k)} \partial f_H(k) \right) \tag{242.20}$$

and learning rate is determined by,

$$\eta_j^c = \frac{1}{(n_2^c + 1)\varsigma^2} \tag{242.21}$$

Also, for the weights matrix between the input and hidden layers, it gives,

$$\frac{\partial E_m(k)}{\partial \boldsymbol{V}_i^c} \cong -e_r(k) \varsigma \frac{\partial O^c(k)}{\partial \boldsymbol{V}_i^c}, \; i = 0, 1, 2 \tag{242.22}$$

And the learning rate is,

$$\eta_i = \frac{8}{n_2^c \| \boldsymbol{W}^{c*} \|_\infty^2 \left(\| \boldsymbol{\gamma}^c \|_\infty^2 + 1 \right)} \tag{242.23}$$

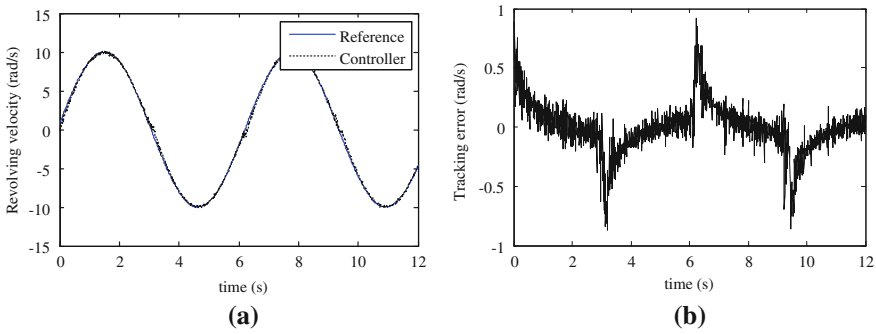


Fig. 242.6 **a** The tracking results of the control system for the periodic sinusoidal command. **b** The tracking error of the control system

242.4 Simulations

The proposed approach is applied to the modeling of the RTWUSM (TRUM-60). The sampling frequency is 100 Hz. The input, used to stimulate the RTWUSM to generate the input/output data for the neural model initialization, is a sinusoidal voltage with time-varying amplitude, frequency and phase, i.e.

$$u(t) = A(t)\sin(\omega(t)t + \phi(t)) + b(t) \tag{242.24}$$

The number of the hidden nodes for NNC and NNUSM are both set to 10. Fig. 242.6a gives the tracking results to the desired trajectory and Fig. 242.6b gives the tracking errors of the control system. It can be seen from the experimental results that the proposed approach has obtained good tracking performance.

242.5 Conclusions

In this paper, a neural network based approach for modeling and control of the hysteresis of the RTWUSM is proposed. To apply the neural networks for the identification of the RTWUSM, an expanded input space based method is adopted with the introduction of a hysteretic operator and rotor velocity into the input space. Then the neural model of the motor is derived based on the expanded input space. Subsequently, an adaptive neural controller is designed with the sensitivity information derived from the neural model online. Simulation results have validated the effectiveness of the proposed approach.

Acknowledgments This research is partially supported the Innovation Program of Shanghai Municipal Education Commission (Grant No.:09ZZ141), the Advanced Research Grant of Shanghai Normal University (Grant No.: DYL200809), the Natural Science Foundation of Shanghai (Grant No.: 09ZR142340), and the Doctors' Funds (B2011-007, B2010-88) of Henan Polytechnic University.

References

1. Hagood NW IV, Andrew JM (1995) Modeling of a piezoelectric rotary ultrasonic motor. IEEE Trans Ultrason Ferroelectr Freq Control 42:210–224
2. Sun D, Liu J, Ai X (2002) Modeling and performance evaluation of traveling-wave piezoelectric ultrasonic motors with analytical method. Sens Actuators A 100:84–93
3. Zhu M (2004) Contact analysis and mathematical modeling of traveling wave ultrasonic motors. IEEE Trans Ultrason Ferroelectr Freq Control 51:668–679
4. Frangi A, Corigliano A, Binci M, Faure P (2005) Finite element modeling of a rotating piezoelectric ultrasonic motor. Ultrasonics 43:747–755
5. Lin F, Wai R, Hong C (2001) Identification and control of rotary traveling wave ultrasonic motor using neural networks. IEEE Trans Control Syst Technol 9:672–680
6. Hu H, Mrad RB (2002) On the classical Preisach model for hysteresis in piezoceramic actuators. Machatronics 13:85–94
7. Hwang CL, Jan C, Chen YH (2001) Piezomechanics using intelligent variable-structure control. IEEE Trans Ind Electron 48:47–59
8. Li C, Tan Y (2004) A neural networks model for hysteresis nonlinearity. Sens Actuator Phys A112:49–54
9. Dong R, Tan Y, Chen H, Xie Y (2008) A neural networks based model for rate-dependent hysteresis for piezoceramic actuators. Sens Actuator Phys A143:370–376
10. Zhang X, Tan Y, Su M (2009) Modeling of hysteresis in piezoelectric actuators using neural networks. Mech Syst Signal Process 23:2699–2711
11. Armstrong B, Neevel D, Kusik T (2001) New results in npid control: tracking, integral control, friction compensation and experimental results. IEEE Trans Control Syst Tech 9:399–406
12. Mayergoyz ID (1988) Dynamic Preisach models of hysteresis. IEEE Trans Magn 24:2925–2927

Chapter 243
Conducting Bar Loss Analysis of Squirrel Induction Motor in Steady Operation

Xu Dong Wang, Peng Fei Li, Xiao Zhuo Xu and Ying Hua Cui

Abstract When squirrel induction motor is steadily running with full load, with layered superposition principle the formulas of conducting bar harmonic loss is deduced. The harmonic loss, the total loss of the conducting bar and the loss ratio caused by harmonic current in the total losses are calculated. In this chapter, the harmonic loss of the induction motor Y160L-4 is analyzed and calculated when it is in steady operation.

Keywords Induction motor (IM) · FEM (finite element Simulation) · Harmonic · Conducting bar loss

243.1 Introduction

Squirrel induction motor is a kind of widely used motor with advantages of simple structure, reliable operation and can be easily controlled. When the motor is running, main loss includes copper loss of stator winding, aluminum loss of

X. D. Wang (✉) · P. F. Li · X. Z. Xu · Y. H. Cui
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan, China
e-mail: wangxd@hpu.edu.cn

P. F. Li
e-mail: evamylike@126.com

X. Z. Xu
e-mail: xxz@hpu.edu.cn

Y. H. Cui
e-mail: cuiyinghuafriend@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_243,

rotor, iron loss of stator and rotor [1]. Generally, when steadily operated, conducting bar current contains ultraharmonics except for fundamental wave. Conducting bar loss is analyzed when non-sinusoidal current flows through the bar in Ref. [4]. Using the methods of finite element analysis and analytic computation, rotor aluminum loss is calculated when steadily operated with full load, and the influence of harmonic current in conducting bar on motor aluminum loss is analyzed.

243.2 Rotor Loss Analysis

When induction motor is steadily operated, loss of bars will be influenced by ultraharmonic current flowing through conducting bar, according to Refs. [2–4].

First, divide the conducting bar into countless laminas, analyze current and loss in every layer. Then composite all the losses to get total losses of the whole conducting bar. Finally, the practical design formula of loss will be obtained.

To calculate conveniently, rectangle conducting bar is divided into N layers rectangle slice with equal height. Consuming that the sum of harmonic current i_j flowing through the j layer:

$$i_j = \sum_n \sqrt{2} I_{nj} \sin(\omega_n t - \alpha_n) \tag{243.1}$$

where, I_{nj}, α_n, ω_n are effective value, initial phase angle and angular velocity of the j layer n sequence harmonic current respectively. The current of that conducting bar: $I = \sum_{j=0}^{\infty} i_j$.

Suppose the j layer rectangle slice resistance:

$$R_j = \rho_j \frac{l}{a_j h_j} \tag{243.2}$$

where ρ_j is electrical resistivity of the j layer slice, which equals ρ, electrical resistivity in every layer, due to the same material is adopted. a_j, h_j are width and height of that layer, and $h_j = \frac{h}{n}$, where h is conducting bar depth, l is axial direction length of the bar.

According to (243.1) and (243.2), average loss of the slice in one period can be expressed as:

$$p_j = \frac{1}{T} \int_0^T i_j^2 R_j \mathrm{d}t \tag{243.3}$$

where T represents one period of fundamental current, total losses of rotor bar can be got with the loss of every layer using principle superposition.

$$p = \sum_{j=1}^{N} p_j = \sum_{n} \sum_{j=1}^{N} I_{nj}^2 R_j = \sum_{n} \sum_{j=1}^{N} I_{nj}^2 R_j \tag{243.4}$$

In formula (243.4), $\sum_{j=1}^{N} I_{nj}^2 R_j$ represents loss of one bar caused by n sequence harmonic current, according to Ref. [4] which introducing resistance increase coefficient, convert AC resistance of every harmonic in one conducting bar to DC resistance of that corresponding harmonic. As rotor conducting bar is _supposed to rectangle, resistance increase coefficient can be calculated according to Ref. [5]:

$$k_{nr} = \xi \frac{\mathrm{sh}2\xi + \sin 2\xi}{\mathrm{ch}2\xi - \cos 2\xi} \tag{243.5}$$

In the formula:

$$\xi = \sqrt{\frac{\omega_n \mu}{2\rho}} h \tag{243.6}$$

where μ, ρ are permeability and resistivity of conducting bar in the slot respectively.

$$p = \sum_{n} k_{nr} I_n^2 R_d \tag{243.7}$$

where R_d is DC resistance of whole conducting bar, I_n is efficient value of n sequence harmonic current which flowing through the bar.

Formula (243.7) is the conducting bar loss solving equation. Loss in whole bar caused by every sequence harmonic can be calculated and then composited according to formulas above.

243.3 FEM Analysis of Three-Phase IM

The formula is verified and analyzed in the model of Y160L-4. Internal magnetic field which changes slightly along with axial when motor steadily operated can be approximately regarded as two dimension problem.

243.3.1 Building Simulation Model

The main parameters of the motor are shown in Table 243.1.

Figure 243.1. shows the axial cross section of the model and the rotor slot type of that model.

Table 243.1 Parameters of motor model

Rated voltage/ V	Rated velocity/ r/min	Stator slot	Rotor slot	Stator outer diameter/ mm	Stator inner diameter/ mm	Line number per slot	Wiring pattern	Air gap width/ mm
380	1460	36	26	260	170	22	△	0.5

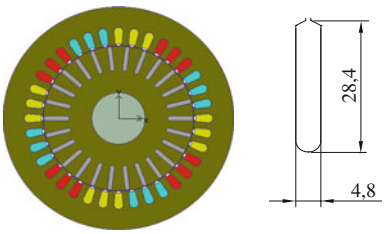


Fig. 243.1 Axial cross section of IM model and rotor slot type

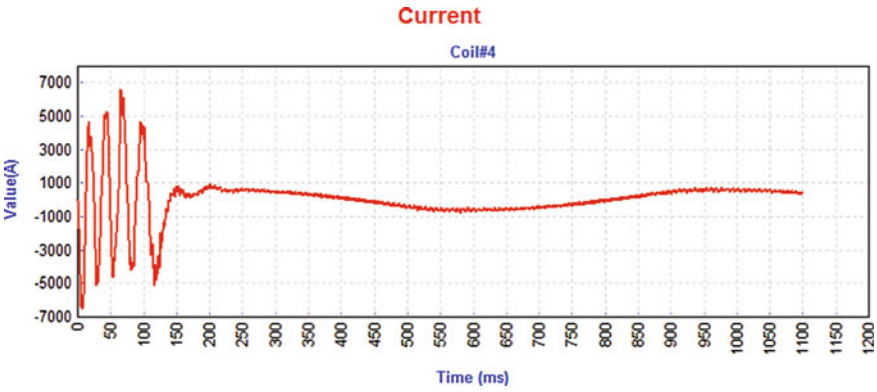


Fig. 243.2 Current curve in conducting bar

243.3.2 Comparison of FEM Simulation and Calculation

In this chapter, FEM simulate software is used to get accurate conducting bar current when motor steadily operated, which are compared with the calculation value. As the simulation is 2 dimension, end resistance and inductance cannot be concerned, so these parameters should be added on the side. The value of end resistance and inductance added can be calculated according to Refs. [6, 7], or with the methods of finite element analog simulation. The calculation value of resistance and inductance are 2.35592463853e-006 Ω and 2.726163345109e-008H, respectively. The model is built with these paremeters.

Current frequency of conducting bar is little when motor steadily operated, so it can be regarded as actual current. Figure 243.2 shows one conducting bar current waveform during motor starting and steady operation.

The stable velocity of motor is 1460 r/min, slip ratio s equals 0.0267, so the fundamental frequency is 4/3 Hz. Take 1 ms as one step, extract 750 dots of current when motor steadily operated, make Fourier decomposition with these data. Figures 243.3 and 243.4 show the harmonic analysis of bar current.

From the Figs. 243.3 and 243.4 above, other harmonics except for fundamental wave can be seen. Take some harmonic which have large amplitude for analysis, such as sequence 2–11, 88–97, 172–176, 217–224. The results are shown as follows:

$$
\begin{aligned}
I = {} & 11.21393 + 647.23356\sin\left(\frac{8}{3}\pi t + \alpha_1\right) + 11.67854\sin\left(\frac{16}{3}\pi t + \alpha_2\right) \\
& + 8.674\sin(8\pi t + \alpha_3) + 2.708\sin\left(\frac{32}{3}\pi t + \alpha_4\right) + 19.724\sin\left(\frac{40}{3}\pi t + \alpha_5\right) \\
& + 5.139\sin(16\pi t + \alpha_6) + 6.113\sin\left(\frac{56}{3}\pi t + \alpha_7\right) + 2.585\sin\left(\frac{64}{3}\pi t + \alpha_8\right) \\
& + 3.597\sin(24\pi t + \alpha_9) + 2.408\sin\left(\frac{80}{3}\pi t + \alpha_{10}\right) + 2.310\sin\left(\frac{88}{3}\pi t + \alpha_{11}\right) \\
& + 3.518\sin\left(\frac{704}{3}\pi t + \alpha_{88}\right) + 5.93805\sin\left(\frac{712}{3}\pi t + \alpha_{89}\right) + 25.285\sin(240\pi t + \alpha_{90}) \\
& + 19.232\sin\left(\frac{728}{3}\pi t + \alpha_{91}\right) + 34.309\sin\left(\frac{736}{3}\pi t + \alpha_{92}\right) + 20.120\sin(248\pi t + \alpha_{93}) \\
& + 26.537\sin\left(\frac{752}{3}\pi t + \alpha_{94}\right) + 12.109\sin\left(\frac{760}{3}\pi t + \alpha_{95}\right) + 6.353\sin(256\pi t + \alpha_{96}) \\
& + 3.092\sin\left(\frac{776}{3}\pi t + \alpha_{97}\right) + 7.445\sin\left(\frac{1456}{3}\pi t + \alpha_{172}\right) + 6.359\sin(488\pi t + \alpha_{173}) \\
& + 11.244\sin\left(\frac{1472}{3}\pi t + \alpha_{174}\right) + 3.908\sin\left(\frac{1480}{3}\pi t + \alpha_{175}\right) + 5.024\sin(496\pi t + \alpha_{176}) \\
& + 4.382\sin\left(\frac{1736}{3}\pi t + \alpha_{217}\right) + 31.269\sin\left(\frac{1744}{3}\pi t + \alpha_{218}\right) + 11.107\sin(584\pi t + \alpha_{219}) \\
& + 87.654\sin\left(\frac{1760}{3}\pi t + \alpha_{220}\right) + 7.876\sin\left(\frac{1768}{3}\pi t + \alpha_{221}\right) + 21.633\sin(592\pi t + \alpha_{222}) \\
& + 4.020\sin\left(\frac{1784}{3}\pi t + \alpha_{223}\right) + 11.828\sin\left(\frac{1792}{3}\pi t + \alpha_{224}\right)
\end{aligned}
$$

where α_i is initial phase angle of every sequence harmonic current.

Figure 243.1 shows that bar height is 28.4 mm and width is 4.8 mm. To simplify the calculation, rotor bar shape is regarded as rectangle, which electrical resistivity is $4.34 \times 10^{-8}\Omega * m$, magnetic conductivity is $4\pi \times 10^{-7}\ H/m$, axial length of motor is 195 mm. Considering influence of frequency, bar ohmic losses calculated with layering calculational method are shown in Table 243.2.

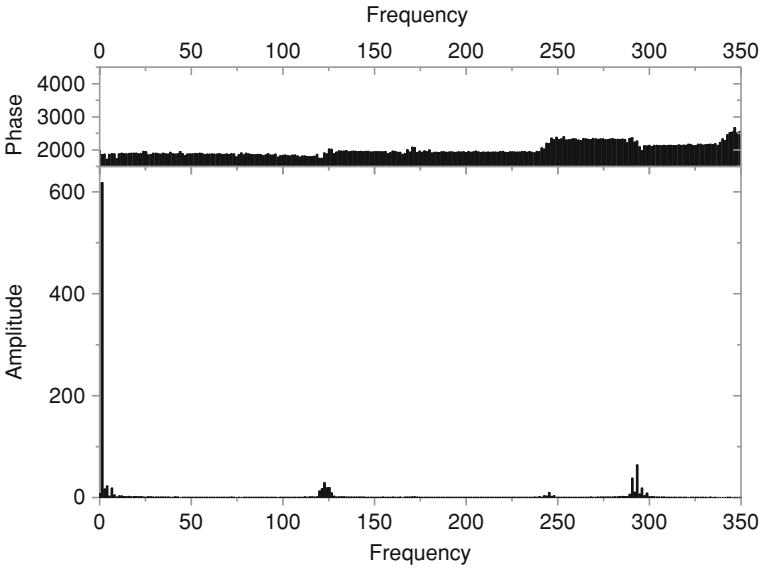


Fig. 243.3 Harmonic analysis of bar current

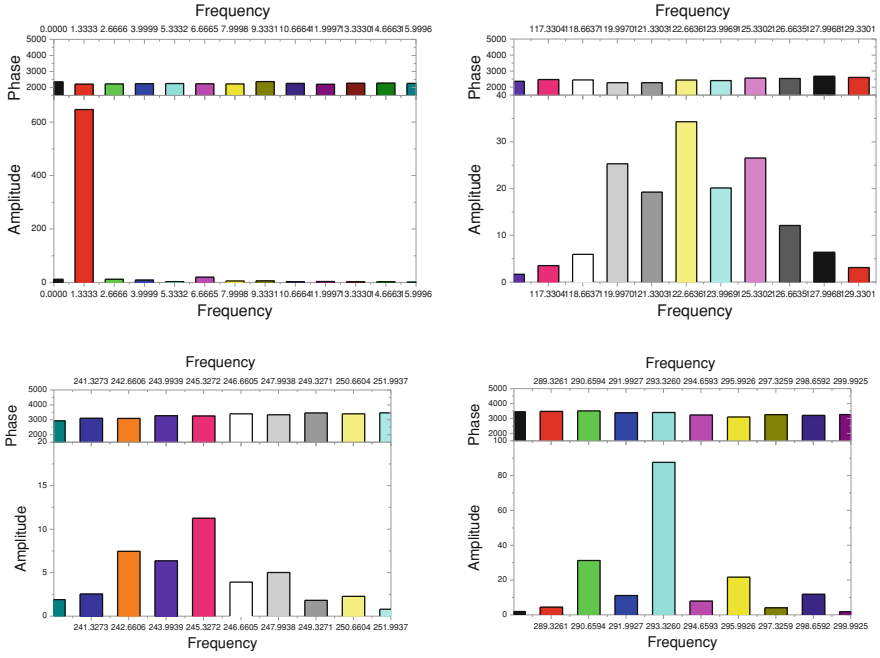


Fig. 243.4 Fractionated gain of bar current harmonic analysis

Table 243.2 Analytic computation of bar losses with different frequency

Frequency (Hz)	Loss (W)	Frequency (Hz)	Loss (W)	Frequency (Hz)	Loss (W)	Frequency (Hz)	Loss (W)
0	0.0078	12	4.3e-4	376/3	0.0665	868/3	0.0027
4/3	13.015	40/3	2.0e-4	380/3	0.0139	872/3	0.140
8/3	4.2e-3	44/3	1.8e-4	128	0.0038	292	0.018
4	0.0024	352/3	0.0011	388/3	9.22e-4	880/3	1.106
16/3	2.3e-4	356/3	0.0032	728/3	0.0073	884/3	0.0090
20/3	0.0123	120	0.059	244	5.3e-3	296	0.0677
8	8.5e-4	364/3	0.0344	736/3	0.0167	892/3	0.0023
28/3	1.2e-3	368/3	0.11	740/3	0.0020	896/3	0.0203
32/3	2.2e-4	124	0.038	248	0.0033		

Analysis and calculation done above show that, when motor steadily operated with full load, loss in rotor conducting bar caused by harmonic current is 11.922% of the total losses, also, the proportion of loss caused by each sequence harmonic current in total losses is clearly to get.

243.4 Summary

Combining the methods of finite element analysis and layered solving, aluminum loss of conducting bar is calculated when motor steadily operated. Concerning different bar shape, loss caused by each frequency harmonic current and its proportion in the total loss can be quantitatively attained with layering calculation. Hence, optimizing certain sequence harmonic current is helpful to reduce rotor aluminum loss and motor high frequency loss during motor optimization.

Acknowledgments This work was supported by the National Natural Science Foundation of China (NO. 61074095), and Henan Outstanding Person Plan (NO. 104200510021), and Ministry of Education Scientific Research Foundation for Chinese Overseas Returnees, and Ministry of Education Research Fund for Doctoral Program of Higher Education, and Henan Province Key Project (NO. 092102210359) and Doctoral Fund of Henan Polytechnic University.

References

1. Zhongfeng Z, Guiying Y (2005) Discussion on the method of improving efficiency of small-sized motor. Explos-proof Electr Mach 40:15–17
2. Xinzhen W, Xiangheng W (2003) Calculation of skin effect for double-cage bar of the induction machine. Proceedings of the CSEE 23(3):116–120
3. Xinzhen W, Yuming L, Shixu X, Ming J, Zhou E (1999) Numerical calculation on rotor parameters of single-phase induction motors considering frequency effect. Trans China Electrotech Soc 14(6):15–19

4. Wu X, Wang X, Luo C (2005) Loss calculation for induction machine rotor bar with non-sinusoidal current. Proceedings of the CSEE 25(17):131–135
5. Tang Y (1998) Magnetic field in electric motor. Science Press, Beijing
6. Shikun C (2005) Motor design. China Machine Press, Beijing
7. Fu F, Tang X (2001) Induction motors design handbook. China Machine Press, Beijing

Chapter 244
Experiment of New Radiator in Electronic Device

Shu Xu

Abstract The experimental apparatus is set up to test the heat transfer performance of the new radiator, which mainly includes vacuum-pumping and central processing unit (CPU) tunnel system, heating system and measurement system. The heat dissipation of the CPU chip is imitated by the copper bar which is heated by thermal resistance wire. The experimental results show that the temperature of the CPU chip decrease with growing of cooling air velocity, but the temperature drop tendency becomes gentle when it continues to increase. Higher heat dissipation power results in higher CPU chip temperature, but oversized heat flux density makes the phase changing process worse.

Keywords Experiment · Electronic devices · Radiator

244.1 Introduction

A new radiator heat transfer and flow performance test system based on central processing unit (CPU) chip radiating is established in this paper [1–4]. The system can be a steady and transient heat transfer test. In the system, the effect of different volume, different heating power, different refrigerant filling quantity and different on heat transfer performance and the time of hot fin-plate radiator reaching a steady state of under different conditions are studied through the electric heating of copper analog CPU heating of the chip [5–8].

S. Xu (✉)
School of Mechanical Engineering, Huaihai Institute of Technology, Lianyungang, China
e-mail: gladxushu@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_244,

244.2 New Radiators and its Test System

244.2.1 Design Structure of New Radiator

In this paper surface of a new radiator is designed a rectangular and a new radiator is mainly composed of three parts:

(1) The upper cover plate and the lower cover plate: materials of the upper and lower cover plate are aluminum alloy. They are smooth rectangular plane, and their thickness is 5 mm.
(2) Core of fin: porous straight fin is used and facilitates the filling liquid refrigerant in the core body flow uniformly. It is equivalent to the wick. Fin layer, upper baffles, lower baffles and seals are connected as a whole by the vacuum diffusion welding method, forming a sealing steam chamber with fin support structure like honeycomb; the bearing capacity of refrigerant in the cavity is greatly improved. Refrigerant in the cavity can flow, evaporate and condense.
(3) The straight rib: the role of rib structure is the augmentation of convective. Heat on the surface through the air forced convection is taken away. This improves steam refrigerant condensation speed and makes it reflux to evaporation–condensation cycle (Fig. 244.1).

244.2.2 Test System of New Radiator

The main factors of effects on heat transfer performance of new radiate under certain ambient temperature are: (1) cooling air flowing between fin of the condensation end (wind speed); (2) the heat flux density of the evaporation end; (3) a phase change heat transfer of refrigerant filling quantity; (4) medium of phase change heat transfer.

Test device mainly comprises: (1) vacuum phase change heat transfer working fluid filling system; (2) wind tunnel testing system [1, 2]; (3) the electric heating simulation system; (4) measurement system (see Fig. 244.2).

244.2.3 Filling Process of Working Fluid of Vacuum Phase Change Heat Transfer

The first step of entire test is that the thermal fin radiator steam chamber is pumped into a high vacuum state and this is also the most crucial step. Specific procedures are as follows:

(1) Internal cleaning, degreasing and impurities are done in the thermal fin radiator.
(2) Vacuum diaphragm valve is closed, and the metering of the liquid refrigerant which is well calculated is poured into container.

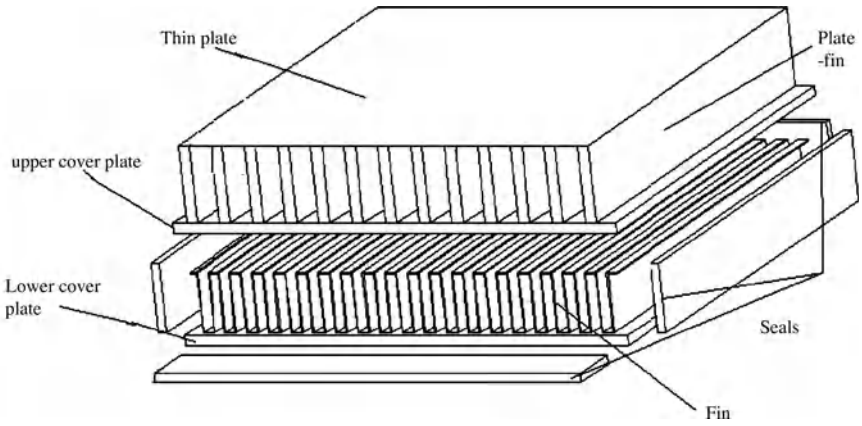


Fig. 244.1 New radiator structure diagram

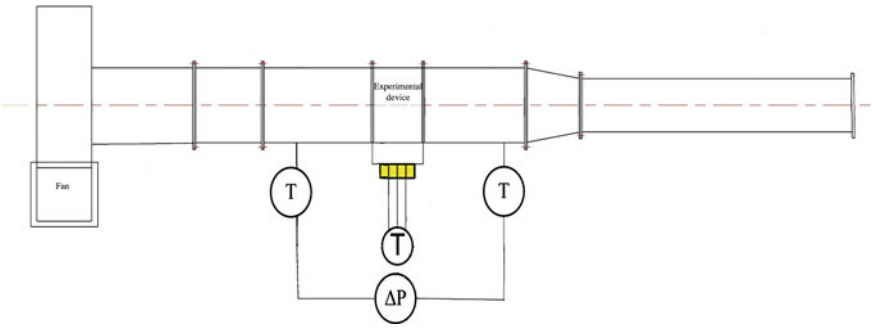


Fig. 244.2 Experimental apparatus for heat transfer performance of new radiator

(3) The vacuum unit is opened, and the vacuum diaphragm valve is opened; the rotary vane vacuum pump to vacuum is used; cooling water and diffusion pump at the bottom of the heater is connected. Oil in diffusion pump is preheated for about an hour and it can work normally.
(4) More than 10 Pa is showed on vacuum gauge display after half of an hour later. The high vacuum valve is opened and pump is used, and the new radiator is pumped vacuum highly at this time; when the vacuum reaches 0.1 Pa, ionization gauge is automatic start to measure high vacuum, until the vacuum degree is about 1.33×10^{-3} Pa.
(5) The thermal fin radiator is heated to 320°C gradually; in the process of heating, heating rate should not be too fast. High vacuum degree inside the new radiator is guaranteed preventing the internal oxidation.
(6) Radiator is decreased gradually to ambient temperature, and the vacuum pump is worked to about 1.33×10^{-3}, vacuum diaphragm valve is shut and

diaphragm valve is opened. Refrigerant is carefully filled, and diaphragm valve B is quickly closed, will heat finned plate radiator is sealed; diffusion pump heater is closed, cooling water is used until the diffusion pump to return to room temperature. The vacuum unit and cooling water are closed.

244.3 Results and Discussion

244.3.1 Transient Heat Transfer Performance of the New Radiator

In Fig. 244.3, every measured temperature points are varied with time under aqueous solution of 15 ml, heating power of 35 W, and maximum wind speed. When the heating is done, each measured temperature of point variation with time. With the heating on, each measured temperature point is increased quickly and then enters to a steady state slowly. Temperature difference between every two temperature points is not large, and this reflect better thermal uniformity in the radiator.

The transient change curve of mean temperature is shown in Fig. 244.4 under aqueous solution of 15 ml, maximum air flow and heating power of 35 and 75 W respectively. With the increasing of heating power, starting time of new radiator is reduced and the stable time is also reduced, but the copper surface temperature also increases accordingly.

244.3.2 Steady Heat Transfer Performance of New Radiator

244.3.2.1 The Effect of Filling Quality of the Working Fluid on Heat Transfer Performance

Figure 244.5 shows that temperature curves of copper surface is in the same degree of vacuum, air flow. Temperature curve of the 15, 20 and 25 ml water is the same basically. In the same heating power condition, the refrigerant liquid rate is increased and the copper surface temperature is made lowerly. The reason may be that the larger refrigerant liquid rate means that liquid level of new radiator steam cavity is higher, and upward steam flows of the evaporation process is under smaller resistance and liquid refrigerant is added more easily. So new radiator has good heat dissipation performance. It can be seen from the above data that the larger filling rate is, the best heat transfer performance of new radiator is.

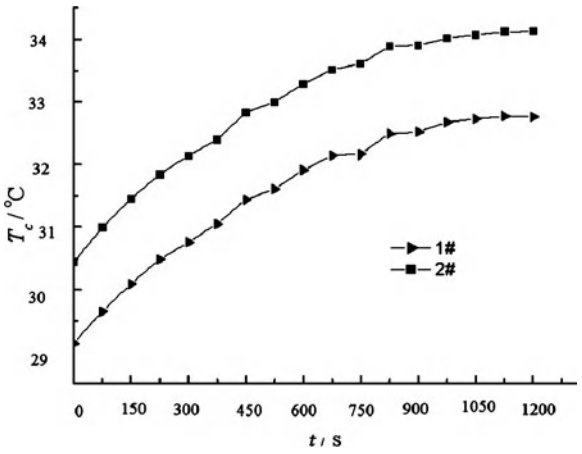


Fig. 244.3 Transient temperature variation curve of different testing points

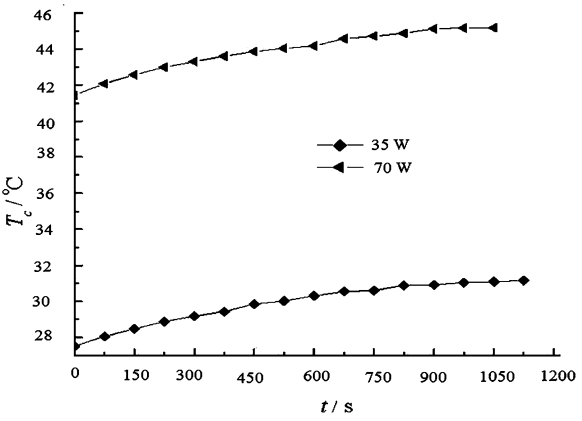


Fig. 244.4 Transient temperature variation curves under different heat flux

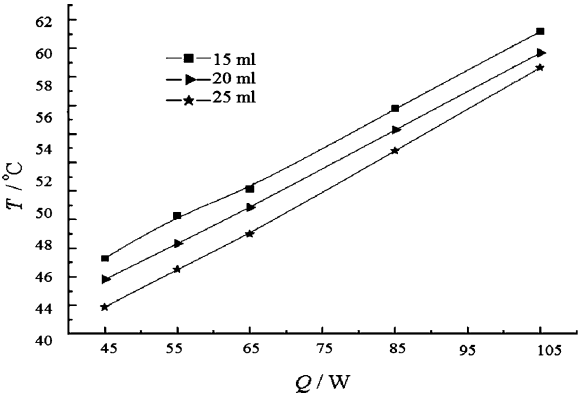


Fig. 244.5 Temperature distribution of copper surface under different filling ratio water

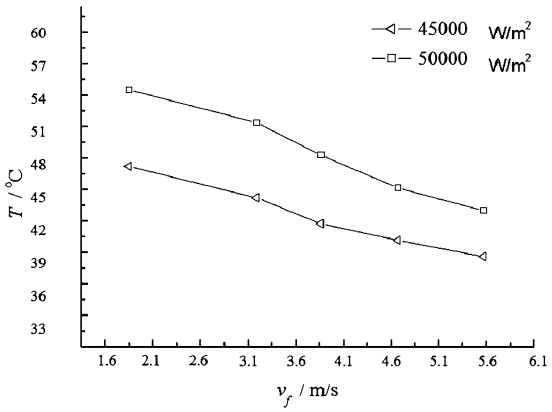


Fig. 244.6 Temperature distribution of copper surface under different cooling air velocity

244.3.2.2 The Effect of Air Flow on Heat Transfer Performance

Figure 244.6 is temperature distribution of a copper surface at different air flow. The test condition is that the vacuum is 2.65×10^{-3} Pa. Its refrigerant filling is water and its filling quantity is 15 ml.

It can be seen from Fig. 244.6 that with the wind speed increasing, the air temperature is reduced; but when it is continued to increase, the change of temperature becomes flat. That is the increased amount of the cooling effect weakens. With air flow increasing, the pressure loss is increased. The loss of energy is in the form of heat, and finally enters the system, the temperature decrease is weakened.

244.4 Summary

In this paper, test device, test results and analysis are systematically introduced the thermal fin radiator heat transfer performance test objectives. Heat Transfer performance test of new radiator for CPU chip, test electric heating rods is used to simulate its calorific capacity. Test for steady state is made. The results are as follows:

(1) Temperature of copper surface decreases with air flow increasing. But when the volume is continued to increase to a certain point, the temperature drop rate will be reduced. This is because the air flow is increased and the pressure loss is also increased and it is in the form of heat into the system eventually.
(2) Compared with water as working fluid the three tests of filling quantity of 15, 25 and 35% respectively. The results showed that heat transfer performance of the new radiator under conditions 35% is best. Because the larger refrigerant charge liquid rate flows upwardly by the smaller resistance, complement of liquid refrigerant is more easily.

References

1. Adami M, Yimer B (1990) Development and evaluation of a planar heat pipe for cooling electronic systems. Chem Eng 10(4):57–74
2. Howard AH, Peterson GP (1995) Investigation of a heat pipe array for convective cooling. J Electron Package 117(9):208–214
3. Basilius A (1987) Heat pipes for cooling high density printed wiring board. In: 6th international heat pipe conference, vol 3(1), pp 531–536
4. Nelson K (1996) Experimental evaluation of micro heat exchangers fabricated by silicon. Natl Heat Transf Conf 3(1):836–839
5. Mafia K, Wang W (1992) Analysis of flow and heat transfer of an asymmetrical flat heat pipe. Int J Heat Mass Transf 35(9):105–112
6. Wang W, Mafia K (1992) Analysis of flow and heat transfer character of asymmetrical flat heat pipe. Int J Heat Mass Transf 135(10):2087–2099
7. Wang Y (1999) An experimental and analytical investigation of the transient characteristics of flat plate heat pipes. Int J Heat Mass Transf 7(10):94–105
8. Siobhan CB, Huang XY (2000) Investigation on transient and steady state performance of a micro heat pipe. Thermophys Heat Transf 14(8):161–169

Chapter 245
Analysis of Characteristics of Permanent Magnet Linear Motor Fed by Sinusoidal and Non-sinusoidal Power Supply

Jikai Si, Shaohua Wang, Xiaozhuo Xu, Haicao Feng and Xudong Wang

Abstract Inverters composed of switch arrays provide non-sinusoidal voltage or current, and the tooth-slot and winding distribution of the motor could lead to a mass of space harmonic components, severely deteriorating the motor performance. This paper adopts the field-circuit coupled adaptive time-stepping finite element method to study the characteristics of PMLSM fed by SPWM voltage source inverter. The solve domain of the field equation is divided into polynomial linear triangular elements, the moving medium is subdivided into layers, and the meshing density in the regions where the magnetic flux density varies acutely higher. The characteristic of having thick air gap is considered in the field-circuit 2D model. The co-simulation using state equation and time-step finite element equation is adopted, the time step of the state equation is smaller than that of the time-step finite element equation.

J. Si (✉) · S. Wang · X. Xu · H. Feng · X. Wang
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China
e-mail: sijikai@hpu.edu.cn

S. Wang
e-mail: wangshaohua@hpu.edu.cn

X. Xu
e-mail: xxz@hpu.edu.cn

H. Feng
e-mail: fhc@hpu.edu.cn

X. Wang
e-mail: wangxd@hpu.edu.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_245,

Keywords Permanent magnet linear motor · SPWM voltage source inverter · Characteristic · Field-circuit coupled · Adaptive time-stepping finite element method

245.1 Introduction

With the rapid development of permanent magnet material, power electronic technology, microelectronic technology, the performance of permanent magnet linear motor tends to be more excellent. Yet inverters composed of switch arrays provide non-sinusoidal voltage or current, and the structural characteristics of the tooth-slot and winding distribution of the motor could lead to a mass of harmonic components, these factors can cause the parametric change of the motor, thrust fluctuation, loud noise, sharp winding temperature rise, severely deteriorating the motor performance.

There are a lot of literatures about the study of characteristics of PMLSM fed by sinusoidal current source, namely the current density distribution of the current source is known, and the finite element method is adopted to solve the equations. In Refs. [1–5], the characteristics of PMLSM fed by sinusoidal current source are studied. In Ref. [6], the time-stepping finite element method is adopted to study the characteristics of PMLSM fed by PWM power source with secondary conductive plane; it only presents the simulation results of the thrust and the current. In Ref. [7], the time-stepping finite element method is adopted to study the characteristics of SVPWM voltage source inverter—PMLSM, and the simulation figures of voltage and current, the simulation result and experimental result of velocity and displacement are presented. Yet the characteristics of SPWM inverter—PMLSM need to be studied further.

When studying the characteristic of voltage source inverter—motor, the current iteration method is often adopted to solve the problem of voltage constraint, which has slow convergence and poor effect, so in this paper, the field-circuit coupled time-stepping finite element method is adopted to have a further study of the characteristics of SPWM voltage source inverter—PMLSM system. The co-simulation has greatly reduced the time for simulation and improved the precision of simulation, and the simulation results are presented at last and validated with the experimental data.

245.1.1 Physical Model of PMLSM

The primary of PMLSM is composed of three-phase windings and the primary iron yoke with rectangular slots, and the secondary is composed of permanent magnets (Neodymium Iron Boron) and magnetic isolation blocks that are placed on the steel plate. The PMLSM studied in this paper is a single side and planar type non-salient pole motor with long secondary and short primary, as shown in Fig. 245.1.

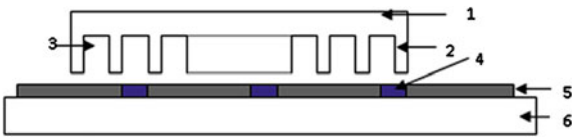


Fig. 245.1 Physical model (Non-salient pole structure). *1* Primary yoke, *2* tooth, *3* slot, *4* magnetic isolation block, *5* permanent magnet, *6* secondary yoke

Table 245.1 PMLSM specifications

	Items	Value
Primary	Phases	3
	Turns	90
	Armature material	Iron
	Pole pitch	39 mm
	Slot pitch	13 mm
	Wire running method	Integral pitch/double layer
	Tooth pitch	13 mm
Secondary	PM material	NdFeB
	PM width	27 mm
	PM height	7 mm
	PM length	120 mm
	PM placement	Surface fix
Air gap	Mechanical	8 mm
Rated thrust		65 kg

In the non-salient pole structure, the magnetization direction of the permanent magnets is concordant with the direction of the air gap flux axis, and has lower pole-to-pole flux leakage and simple craftwork. The specifications of PMLSM are shown in Table 245.1.

245.2 Field-Circuit Coupled Mathematical Model of PMLSM

To take the external circuit fed by SPWM voltage source inverter and the motor end effect into account, this paper adopts field-circuit coupled method to calculate the electromagnetic transient process, solve equation variables of magnetic vector potential and the motor phase current, which are the combination of electromagnetic field time-step finite element equations and three-phase winding circuit equations by electromotive force in the armature windings.

The transient field equation in which A denotes magnetic vector potential is shown as Eq. (245.1) according to Maxwell equations.

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} = -\mu J_s - \mu J_m \tag{245.1}$$

where A—z-axis component of magnetic vector potential, J_s—current density of the primary windings, J_m—equivalent magnetizing surface current density of permanent magnet, μ—the permeability.

In this paper, the model is subdivided into small triangle elements to form a mesh and adopts n-order unit basic function and linear interpolation. After applying the Galerkin method, the governing equations for the analysis model is expressed as follows.

$$[[S] - [C]] \begin{bmatrix} [A] \\ [I] \end{bmatrix} + \frac{\partial}{\partial t} [[T][0]] \begin{bmatrix} [A] \\ [I] \end{bmatrix} = [G] \tag{245.2}$$

where A—unknown node magnetic vector potential, I—current in the windings, S,C,T—coefficient matrix, G—matrix of equivalent magnetization current density.

Maxwell's stress tensor is adopted to calculate PMLSM electromagnetic force, which includes all kinds of harmonics component electromagnetic force. The motor electromagnetic force tangential component is shown in Eq. (245.3).

$$F_{\text{thrust}} = \frac{L_1}{\mu_0} \int_0^{L_2} (-B_x B_y) \mathrm{d}x \tag{245.3}$$

The motor electromagnetic force normal component is shown in Eq. (245.4).

$$F_{\text{normal}} = \frac{L_1}{\mu_0} \int_0^{L_2} \frac{1}{2} (B_x^2 - B_y^2) \mathrm{d}x \tag{245.4}$$

where L_1—winding effective length, L_2—integrating range, B_x—x-axis flux density component in the air gap field, B_y—y-axis flux density component in the air gap field, F_{thrust}—electromagnetic thrust force, F_{normal}—normal electromagnetic force.

Movement equation of PMLSM is shown in Eq. (245.5).

$$F_{\text{thrust}} = m \frac{\mathrm{d}v}{\mathrm{d}t} + F_L \tag{245.5}$$

where m—mass, v—the motor mover velocity, F_L—load force.

245.3 Analysis of Steady-State Characteristics of PMLSM

In order to know the cause of electromagnetic thrust and mover velocity fluctuation of low-velocity PMLSM fed by SPWM-VS, the steady-state characteristics of low-velocity PMLSM are analyzed.

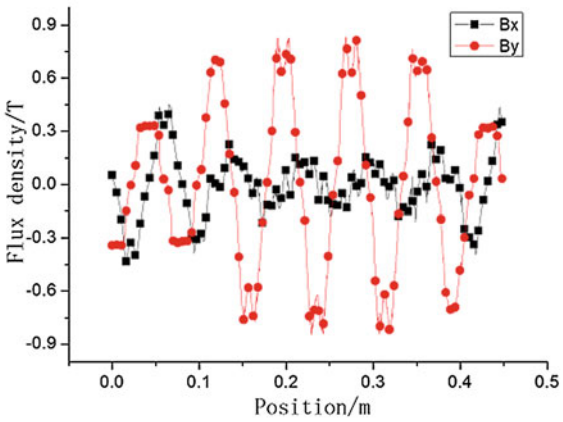


Fig. 245.2 Calculation result of the flux density distribution of SPWM-VI-PMLSM

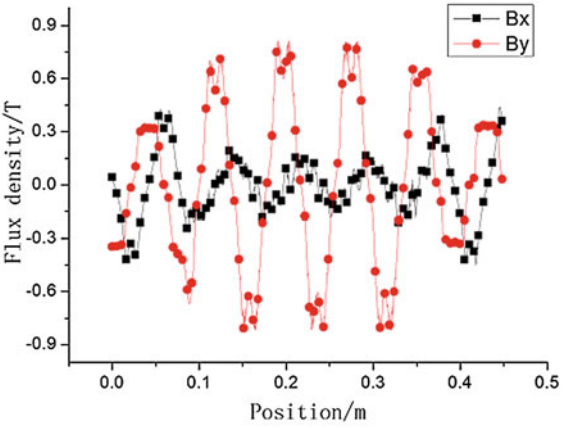


Fig. 245.3 Calculation result of the flux density distribution of S-VS-PMLSM

The magnetic field distribution is the result of the joint action of the primary windings and secondary permanent magnet of PMLSM, the field distribution and variation in different media and its linkage with current decide the electromagnetic parameters and the performance of PMLSM. Figures 245.2 and 245.3 show the result of calculation of magnetic flux density on air-gap centerline of SPWM-VI-PMLSM and S-VS-PMLSM, respectively. B_x and B_y in the figures denote tangential and normal magnetic flux density, respectively.

It can be seen from Fig. 245.4 that the tangential electromagnetic thrust of S-VS-PMLSM under steady state has better sine degree, oscillate around the value of 100.00 N with a period decided by pole pitch, the oscillation amplitude is 267.00 N. The tangential electromagnetic thrust of SPWM-VI-PMLSM under steady-state oscillate around the value of 112.29 N with a period decided by pole pitch, the oscillation amplitude is 393.21 N, and it has various harmonic components. Thus it can be seen that for dampening thrust fluctuation, S-VS-PMLSM has better performance than SPWM-VI-PMLSM.

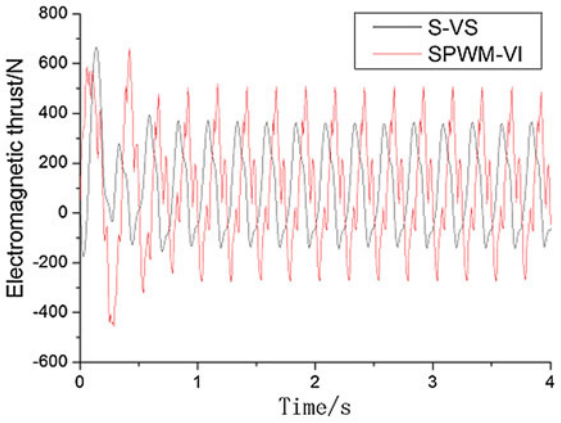


Fig. 245.4 Calculation result of the electromagnetic thrust

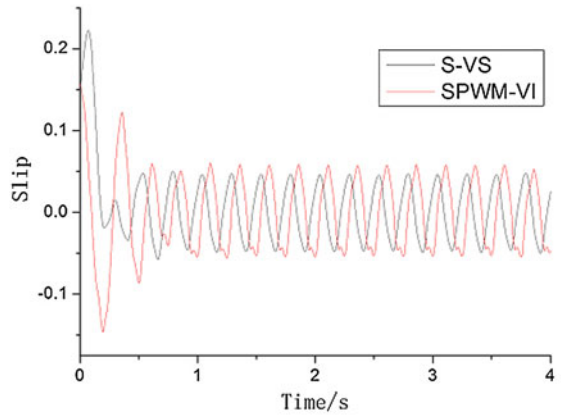


Fig. 245.5 Calculation result of slip

Figure 245.5 shows the calculation result of slip curve of SPWM-VI-PMLSM and S-VS-PMLSM. It can be seen from Fig. 245.5 that S-VS-PMLSM and SPWM-VI-PMLSM have the same periodicity of slip fluctuation, yet the slip value of SPWM-VI-PMLSM is 24.6% larger than that of S-VS-PMLSM.

245.4 Conclusions

For the open-loop scalar and constant flux controlled running mode of PMLSM, to study the characteristics of SPWM voltage source—PMLSM using field-circuit coupled time-stepping method is feasible, and the co-simulation using state equation and field finite element equation is adopted, it has reduced the time for simulation and improved the precision of simulation.

Acknowledgments The authors are greatly indebted to Doctoral Program of Higher Education Research Fund for new teachers funding China Ministry of Education (20104116120001), Natural Science Program of Henan province Education Department in 2011(2011A470003) for its support and henan polytechnic university doctor fund project (B2010-53) for its support.

References

1. Wang X, Yuan S, Jiao L, Wang ZH (2001) 3-D analysis of electromagnetic field and performance in a permanent magnet linear synchronous motor, IEMDC 2001. IEEE Int pp 445–447
2. Sanada M, Morimoto S, Takeda Y (1997) Interior permanent magnet linear synchronous motor for high-performance drives. IEEE Trans Ind Appl 33(4):966–972
3. Jung S-Y, Chun J-S, Jung H-K (2001) Performance evaluation of slotless permanent magnet linear synchronous motor energized by partially excited primary current. IEEE Trans Magn 37(5):3757–3761 (Part 1, Sept 2001)
4. Kwak S-Y, Kim J-K, Jung H-K (2005) Characteristic analysis of multilayer-buried magnet synchronous motor using fixed permeability method. IEEE Trans Energy Convers 20(3):549–555 (Sept 2005)
5. Kang G-H, Hur J, Lee B-K et al (2004) Force characteristic analysis of PMLSMs for magnetic levitation stage based on 3-Dimensional equivalent magnetic circuit network, IAS 2004, pp 2099–2104
6. Jung I-S, Hyun D-S (1999) Dynamic characteristics of PM linear synchronous motor driven by PWM inverter by finite element analysis. IEEE Trans Magn 35(5):3697–3699 (pt 2, Sept 1999)
7. Kwon BI, Woo KI, Kim DJ, Park SC (2000) Finite element analysis for dynamic characteristics of an inverter-fed PMLSM by a new moving mesh technique. IEEE Trans Magn 36(4):1574–1577 (Part I, July, 2000)

Part XVII
Optical, Electrical, Magnetic and Composite Materials

Chapter 246
Research and Simulation of Photovoltaic Cells and its Maximum Power Point Tracking Method

Qi-ying Xu and Yong-sheng Zhu

Abstract The modeling of photovoltaic cells is done according to its mathematical model in this article. Furthermore, a new Maximum Power Point Tracking (MPPT) control method is proposed which applies CV (Constant Voltage)algorithm to adjust the working point of photovoltaic (PV) cells near the MPP (Maximum Power Point) for fast tracking when working point is far from MPP, and uses adapting duty cycle perturb method to optimize the steady state characteristic of MPP for effectively eliminating the output power oscillation under the setting precision. The simulation results verify the correctness of the model and also show that the proposed method tracks the MPP exactly and quickly, improves the energy conversion efficiency, static and dynamic performance of PV system.

Keywords Photovoltaic cells · MPPT · Adaptive duty cycle perturb method

246.1 Introduction

Solar Photovoltaic system is a high quality renewable energy and is of great strategic importance for building a low carbon environment. The output characteristics of photovoltaic cells is obviously nonlinear, only at a certain voltage the cells can output the maximum power when the working point of solar cells reached the highest point of the P–V curve, which called the maximum power point. As the low conversion efficiency, high cost, expensive price, large initial investment and

Q. Xu (✉) · Y. Zhu
School of Electronic and Information Engineering,
Zhongyuan University of Technology, Zhengzhou 450007, Henan, China
e-mail: xqyyl@sina.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_246,

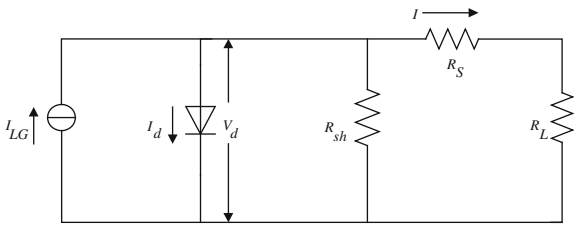


Fig. 246.1 Equivalent circuit of photovoltaic cells

its vulnerable power output to sunlight intensity, temperature and other factors of photovoltaic cells at present, the maximum power point tracking for photovoltaic power generation system is a key technical issue for the effective use of photovoltaic cells [1–5].

The modeling of photovoltaic cells is done according to its mathematical model in this article. And the output characteristics of photovoltaic cells are obtained. On this basis, the constant voltage (CV) with adaptive duty cycle perturb method is proposed. The algorithm stabilizes photovoltaic systems to the maximum power point nearby rapidly by using the CV method, then uses adaptive duty cycle perturb method to fine-tune the system to make the system work at maximum power point fast without oscillation. The algorithm not only has fast response, but also overcomes the shortcomings of power switch in the MPPT which is difficult to choose the adjust volume of the duty cycle (ΔD) circuit.

246.2 Mathematical Model of Photovoltaic Cells

The equivalent circuit of photovoltaic cells is shown in Fig. 246.1.

Generally, R_s and R_{sh} can be ignored when the discussion of the actual equivalent circuit. Simplified equation of photovoltaic cells output characteristics are:

$$I = I_{LG} - I_{\mathrm{OS}}\left\{\exp\left[\frac{q}{\mathrm{AkT}}(V + \mathrm{IR}_S)\right] - 1\right\} \tag{246.1}$$

where, $I_{LG} = [I_{\mathrm{SCR}} + K_1(T - 298)]\frac{\lambda}{1000}$; I, V are the photovoltaic cells output current and voltage. I_{LG} is the photocurrent; I_{OS} is the dark saturated current; K_1 is the short-circuit current temperature coefficient; I_{SCR} is the short-circuit current of photovoltaic cells under standard test conditions; λ is sunlight intensity; A is Diode factor; K is Boltzmann constant; T is kelvin temperature; q is electric charge.

246.3 Maximum Power Point Tracking Algorithm

As the photovoltaic cells maximum power point is a time variable, we can realize the MPPT by using search algorithm which can be divided into two categories: self-optimizing and non-self-optimizing. The so-called non-self-optimizing

algorithm is rarely used in practice. Self-optimizing algorithm, which dose not directly detect the changes in environmental factors, measures the electrical signal to determine the MPPT. The method includes fixed voltage control method, disturbance observe method, admittance increment method and gradient-based variable step admittance increment method [6].

246.3.1 Basic of the Algorithm

As we know, the CV method can not guarantee maximum power output of photovoltaic (PV) cells, but can only guarantee that the system works at maximum power point in the vicinity. To take full advantage of the PV cells power output, when the CV method achieves control objectives, the adaptive duty cycle perturb method will take over around the maximum power point.

Compared with the traditional duty cycle perturb method, when the operating point is far from the maximum power point, the tracking control is achieved by the CV method and the adaptive duty cycle perturb method mainly optimizes the steady state characteristics near the maximum power point. The algorithm keeps the idea of perturb and observe method, constantly disturb the system to find the maximum power point. But the difference is that the amplitude of disturbance changes according to the system operating point which embodies the idea of adaptive, improves response speed, accuracy and stability of the system. After finding the maximum power point, compared with the traditional method's unstablity around the maximum power point, this algorithm can ensure the system works arbitrarily close to maximum power point in the vicinity theoretically.

246.3.2 Principle of Adaptive Duty Cycle Perturb Method

There are still problems in adjustment of step size of the duty cycle D in the duty cycle perturb method. Here the automatic online regulator with a step length λ is introduced to solve the above problems [4, 5]. It can be expresses as (246.2):

$$\lambda(k+1) = M\Delta P/\lambda(k) \tag{246.2}$$

where, $\lambda(k)$ is the adjustment step for the duty cycle D and varies between 0 and 1. $\Delta P = P(k) - P(k-1)$, changes in the size of the power. M is a constant and determines the sensitivity of adaptive adjustment.

When the external environmental factors such as the PV cells temperature, sunlight intensity become suddenly bigger, the general perturb observe method still considers that the interference causes the output power variation is due to the output voltage (or current) to increase or decrease an adjustment step, thus may

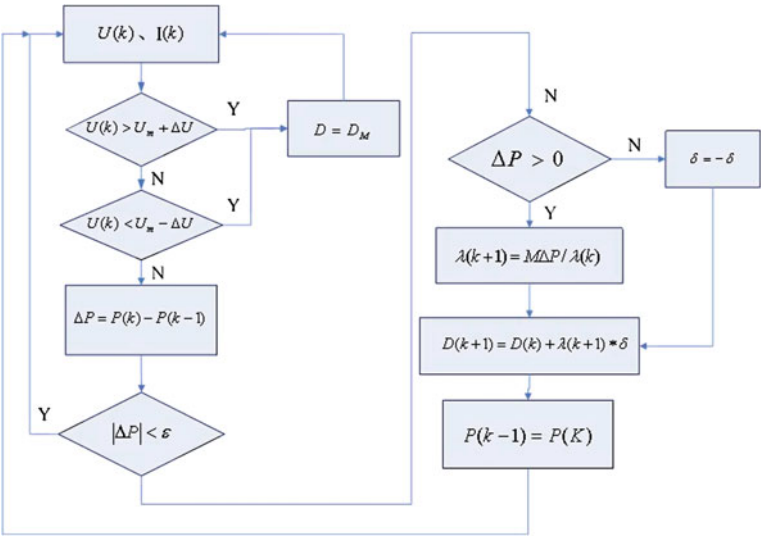


Fig. 246.2 Flowchart of MPPT control by combination method of CTV and ADCP

make the controller away from the maximum power point. The expression (246.2) can solve this confusion question. When $\Delta P/\lambda(k)$ is small, it is shown that the change in output power P was mainly due to adjusting step size of the duty cycle D, then $\lambda(k+1)$ should not change much compared with $\lambda(k)$.When $\Delta P/\lambda(k)$ is large, that changes in power P was mainly due to PV cells surface temperature, sunshine intensity and other external factors. If the maximum power point at this time has a significant shift, then the step $\lambda(k+1)$ become larger, which can make the system fast track to a new maximum power point.

In theory, we can always find a photovoltaic cells maximum power point where $\Delta P = 0$. However, during the actual process, this point can hardly be found. Therefore, the traditional methods will often cause the system working near the maximum power point with oscillations and some power loss. So introducing a parameter ε and determining ΔP, when $\Delta P < \varepsilon$, it is considered that the maximum power point was found. The size of ε can be determined according to different accuracy.

Different from the traditional methods, once the system finds the maximum power point $P_{\max}$, stops perturbing. The reason to do so is that if it continues to disturb, the system will still not work at maximum power point, causing the output of the system instability and reducing system efficiency. Because in the sunshine, the PV cells output changes is very small within a short time, there is no need to disturb all the time. After the cessation of disturbance, monitor the working state, and to make appropriate changes according to different judgments at any time.

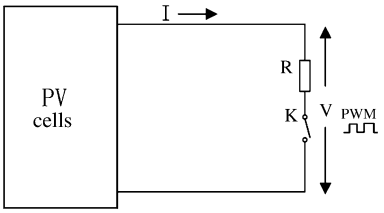


Fig. 246.3 PV system diagram

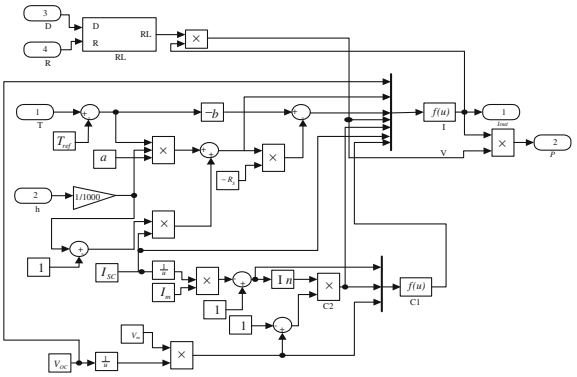


Fig. 246.4 Simulation models of photovoltaic cells

The flow chart of the method is shown in Fig. 246.2. Where D_m is the duty cycle at U_m in CV method. The process repeats until ΔP is smaller than the set of control accuracy where the system works at maximum power point with no disturbance to eliminate oscillation.

246.4 System Simulation

246.4.1 Simulation Model

PV system diagram is shown in Fig. 246.3. Load is a resistor R and a switch K which is controlled by a PWM control signal whose duty cycle is D. When the switch K is disconnected, D = 0%, means open circuit. When the switch K is closed, D = 100%, means the load $R_L = R$. The relationship between the load R_L and duty cycle D can be approximately expressed as $R_L = R/D$. The PV cells output voltage is $V = IR_L = IR/D$. Using the temperature T, sunshine intensity S, fixed load resistance R and the duty cycle D as the inputs, the model of photovoltaic cell structures is built by Matlab/simulink and shown in Fig. 246.4.

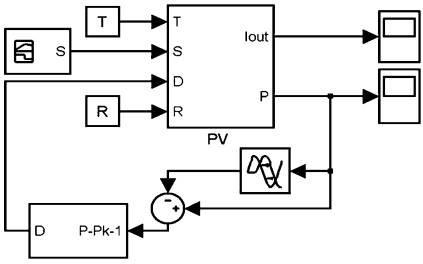


Fig. 246.5 Simulation control model

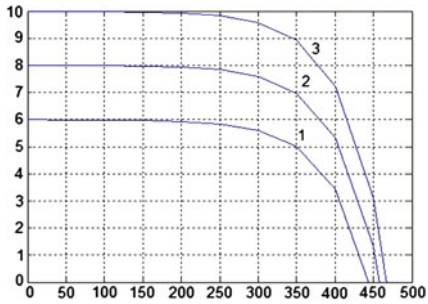


Fig. 246.6 the output I–V of photovoltaic cells

According to the simulation model of photovoltaic cells, the control chart of photovoltaic power generation system is shown in Fig. 246.5.

246.4.2 Simulation Results and Analysis

246.4.2.1 Photovoltaic Cells Output Characteristics

The curves 1, 2, 3 in Fig. 246.6 are the output *I–V* of photovoltaic cells in temperature of 25°C, where, the sunshine intensities are 600, 800 and 1000 W/m^2, respectively.

From Fig. 246.6, in the case of constant temperature, the bigger the sunlight intensity is, the greater the solar cells output current is, while the open circuit voltage is of little change. When the sunshine intensity and temperature is fixed, each curve has a maximum value, which is the PV cells maximum power point. At a constant temperature, the bigger the sunlight intensity is, the greater its power output is.

Figure 246.7 is the output characteristics of photovoltaic cells when the load impedance changes. We can see from the curve, the output power of photovoltaic cells changes with the load. From the above analysis, the greater the switch duty cycle is, the smaller the input impedance becomes. When changing the switch

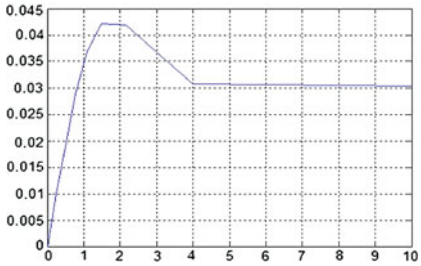


Fig. 246.7 The output characteristics of photovoltaic cells when the load impedance changes

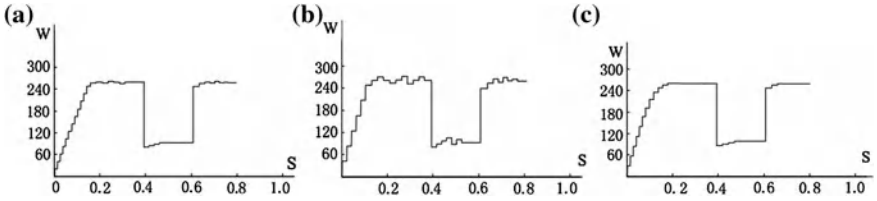


Fig. 246.8 Simulative waveforms of different methods **a** CV method **b** the duty cycle perturb method ($\lambda = 0.05$) **c** the CV method combined with adaptive duty cycle perturb method

duty cycle, making the equivalent input impedance matches the output impedance of the photovoltaic cells, they can output maximum power, which is the theory basis of making use of the duty cycle to achieve MPPT.

246.4.2.2 MPPT Algorithm

In order to verify the proposed method's effectivity, the simulation is carried out at $T = 25°C$, $R = 1.568\ \Omega$. At $t = 0.4$ s and $t = 0.6$ s when the intensity of solar radiation was set from 1000 to 400 w/m^2 and then to 1000 w/m^2. CV method, the duty cycle perturb method and the CV method combined with adaptive duty cycle perturb method are simulated respectively. Figure 246.8 shows the comparison of simulation waveforms of different methods.

It can be seen from Fig. 246.8a, the fixed voltage method can rapidly track the changes in light intensity; In Fig. 246.8b, the duty cycle perturb method can also rapidly track the maximum power point when the light intensity changes, but there is a greater steady-state output power fluctuations. Figure 246.8c shows that the CV method can quickly move the system to the operating point near the maximum power point when sudden changes in light intensity, then adaptive duty cycle perturb method takes over and continuously approaches the maximum power point. After meeting the setting the accuracy, the duty cycle D remains steady and no longer changes. The system stops oscillation and the maximum power point is

found. Comparing Fig. 246.8b and c, the CV method with adaptive duty cycle perturb method has faster tracking speed than that of perturb and observe method. And the duty cycle D remains steady when approaching the maximum power point, when the output power has no fluctuations.

246.5 Conclusion

In this paper, the modeling of photovoltaic cells is done according to its mathematical analysis. And the output characteristics of photovoltaic cells are also obtained. On this basis, this paper presents a new MPPT algorithm-the CV method combined with adaptive duty cycle perturb method. This method not only overcomes the defects in the fixed voltage method which can not accurately track the maximum power point when changes in the external environment, but also solves the conflict about the stability and the response speed in perturb and observe method. Experimental results show that, the photovoltaic cells model can simulate the output characteristics of the real PV system with the cell surface temperature, sunlight intensity and other external environmental factors which has significant nonlinear feature. Furthermore, the proposed MPPT method can quickly and accurately track the PV cell's maximum power point, eliminate the oscillations, improve the output efficiency of photovoltaic cells effectively and make the system have a better dynamic performance.

References

1. Walker G (2001) Evaluating MPPT converter topologies using a Matlab PV model. Elect Electron Eng 21(1):49–56
2. Qiao X, Wu B, Deng Z, You Y (2008) MPPT of photovoltaic generation system using fuzzy/PID control. Electric Power Autom Equip 28:92–95
3. Pandey A, Dasgupta N, Mukerjee AK (2007) A simple single-sensor MPPT solution. IEEE Trans Power Electron 22(2):698–700
4. Li Q, Zhou L, Liu Q, Zhang F, Wu J (2008) Simulative research of MPPT for photovoltaic power system. Electric Power Autom Equip 28:21–25
5. Chen YK, Yang CH, Wu YC (2002) Robust fuzzy controlled photovoltaic power inverter with Taguchi method. IEEE Trans Aeros Electr Syst 38(3):940–954
6. Alippi C, Galperti C An adaptive maximum power point tracker for maximizing solar cell efficiency in wireless sensor nodes. In: Proceeding ISCAS 2006 Kos, Greece, 21–24 May 2006 3722–3725

Chapter 247
Research of MPPT for Solar PV Generation System Based on Independent Voltage Controller

Lingzhi Yi, Hanmei Peng, Genping Wang, Xiaoxue Luo, Jie Liu, Shuhao Wang and Zhezhi Yao

Abstract To overcome the problem of partial shades or the mismatching of photovoltaic (PV) arrays in PV generation system, a method of diode-clamp multilevel inverter to connecting PV array is proposed, and a new pulse-width modulation strategy is adopted, voltage of every PV array can be controlled independently, the maximize system power can be got, the rated voltage of equipment can be reduced, the output voltage distortion can be cut down, so the efficiency of the whole system is improved. Simulation results verify that the feasibility and validity of control strategy.

Keywords Diode-clamped multilevel inverter · PV array · Independent voltage control · Pulse width modulation · MPPT

247.1 Introduction

Multilevel Inverter has many advantages, such a low device withstands voltage, small switching loss, high system efficiency, and good EMI (electro magnetic interference), it can be linked to Voltage power source, so the volume of system is small. Compared to Flying capacitor multilevel inverter and hybrid cascaded multilevel inverter, diode-clamped multilevel inverter has become research

L. Yi (✉) · H. Peng · X. Luo · J. Liu · S. Wang · Z. Yao
College of Information Engineering, Xiangtan University,
Xiangtan Hunan 410005, China
e-mail: ylzwyh@sohu.com

G. Wang
Department of Mechanical and Electricity, Shenzhen Polytechnic,
Guangdong 518055, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_247,

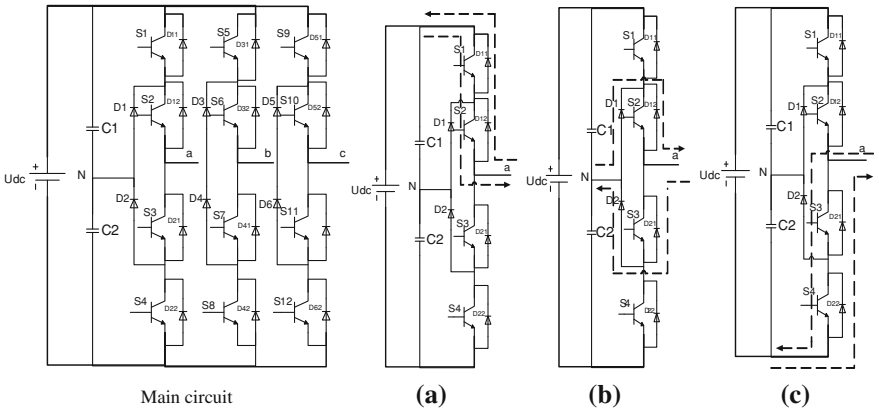


Fig. 247.1 Main circuit and its three valid voltage output state of diode-clamp multilevel inverter

hotspots of power electronics technology to save energy and reduce the cost. Due to the problem of voltage balance control of DC capacitors. To some extent degree, the use of diode-clamped multilevel inverter is restricted. Different DC voltages are generated by different PV arrays, under the condition which these devices withstand voltage, breakthroughs has not been made, multilevel inverter is especially adapted to PV power generation system [1–4].

With high integration and multifunction, mixed PV generation system can be connected to conventional electrical power system or be independent power source to a powered load.

247.2 Diode-Clamp Multilevel Inverter

In a diode-clamp multilevel inverter, when the power switch is turned on, the clamp diode can provide a current channel for the electric circuit, avoiding the capacitor shorted. In Fig. 247.1, each phase bridge arm has four switches in three-level inverter circuit, so totally $2^4 = 16$ states exist. Actually, due to S_{x1} and S_{x3}, S_{x2}, and S_{x4} cannot be switched in the same time, there are only three valid states left [5].

(1) As S_1, S are turned on, S_3, S_4 are turned off (see Fig. 247.1a), when current i_a go to power load from inverter, the direct power will charge the capacitor C_1, phase voltage output $U_a = U_{dc}/2$.

(2) As S_2, S_3 are turned on, S_1, S_4 are turned off (see Fig. 247.1b), when current i_a go to power load from inverter(clamp diode 1, power valve S1), the direct power will charge the capacitor C_1, phase voltage output $U_a = 0$; As the current from power load go to inverter through power valve S_3 and clamp diode D_2, the capacitor C_2 will be charged, phase voltage output $U_a = 0$.

(3) As S_3, S_4 are turned on, S_l, S_2 are turned off (see Fig. 247.1c), if phase current is negative, the direct power will charge the capacitor C_2, the current *will* go through power valves S_3, S_4, if the voltage loss of the power valve can be ignored, phase voltage output $U_a = -U_{dc}/2$; if the current from power load is positive, the current will go through the diodes D_{21} and D_2 which are paralleled with the power valves S_3 and S_4, and the capacitor C_2 will be charged, the phase voltage output $U_a = -U_{dc}/2$.

247.3 SVPWM Technology Based on Independent Voltage Control

In the conventional multilevel inverter strategy, the existence of low voltage oscillation may increase the voltage electrical force, and lead to six order harmonic waves. In order to decrease the voltage fluctuation, besides the independent voltage control algorithm SVPWM given in this paper, the inverter output voltage and grid factor can be controlled, which insure the middle point voltage balance.

247.3.1 Transformation With 3-level and 2-level Electrical Signal Space Vectors

The basic idea of space vector modulation is to determine the inverter's on–off state, and the connected time based on voltage—second balance principle. The geometrical relationship in two-level space vector $A_1 = (1, 0)$, $A_2 = (0.5, h)$, according to the equation $v\ T_s = v_a\ t_a + v_b\ t_b$, in the sampling period T_S, the voltage-second equation can be derived:

$$v_\alpha T_S = t_a + 0.5 t_b \; v_\beta T_S = h t_b \tag{247.1}$$

In above equations, v_a and v_b are elements of v in the $\alpha - \beta$ coordinates system. Assuming each side length of the triangle is 1, $h = \sqrt{3}/2$ in sector S_1, the turn-on time of space vector can be got

$$t_a = T_s\left(v_\alpha - \frac{v_\beta}{2h}\right), \quad T_b = T_s \frac{v_\beta}{h}, \quad t_0 = T_s - t_\alpha - t_b \tag{247.2}$$

In Fig. 247.2, three-level space vectors can be transformed to two-level space vectors. Basis on different switch set in three-level inverter. It can generate many switch vectors. In Figs. 247.2a, c. There are six basic switch vectors (A_0–A_5) in a big sector. Every big sector can be divided into four small triangles (A, B, C, D). All three vectors of small triangle is a switch vector, they are the basic vectors used to Synthesis of space vector. Compare three-level space vectors with

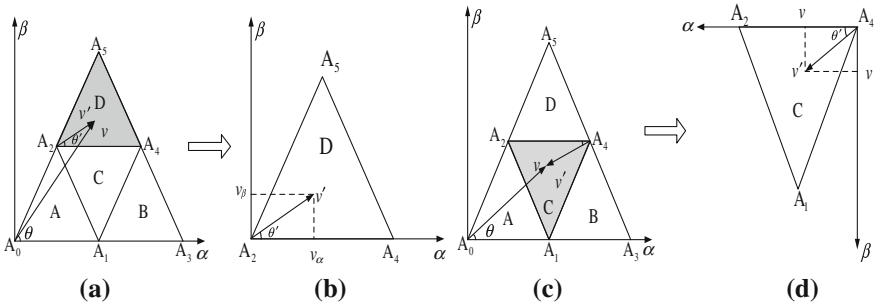


Fig. 247.2 Transform form three-level space vectors to two-level space vectors

two-level space vectors, in small riangles C, A_1, A_2, A_4 are its three vectors. If A_1 is zero vector, it is similar to S_2 sector in two-level space vector. If A_4 is zero vector, it similar to S_4 sector in two-level space vector. In three-level inverter, $|A_0\ A_3| = 2U_{dc}/3$, its magnitude is equal to the side length of small triangle in two-level inverter. The length of $|A_1\ A_3|$ is half of the length of two-level switch vector. To simplify the calculation, suppose the side length of triangle is 1, the reference vector v and its sector is known, the small vector is calculated. See Fig. 247.2a, the amplitude of reference vector v is $|v|$, there is a angle θ between v and the axis α. On the same point, define a small vector v' in $\alpha-\beta$ Coordinate system, see Fig. 247.2b. There is a angle θ' between v' and the axis α. According to the voltage—second balance principle, the approximation small vector v' is equal to the actual vector v. In three-level inverter, the small triangle of the sector can be located, which the reference vector can be determined by voltage–second balance principle, ($V_{\alpha 0}$, $V_{\beta 0}$) can be calculated. The turn-on time of every power switching can be derived and calculated through Figs. 247.2b and d.

247.3.2 Independent Voltage Control of PV Array

PV arrays are often connected in parallel or series in solar PV generation system, the same current must to drive in all series arrays. Compare to centralized inverter, single series inverter can trace the maximum power point better, but the optimization has not been implemented yet. Series PV array consists of 20–30 modules commonly, the total length is 15–25 m. Different I–V characteristics are caused by the shade or dust of PV arrays, the aging of PV Cells, many local MPPT can exist in P–V characteristic of series PV arrays. MPPT algorithm is adopting in PV inverter to raise the output energy. There are many MPPT algorithms, such as hill-climbing algorithm, Perturbation observation algorithm, incremental conductance, Fractional Open-Circuit Voltage control, short circuit control, fuzzy control, and genetic algorithm. In these algorithms, a local MPPT maybe to convergent, sometimes, the global MPPT is replaced by this local MPPT.

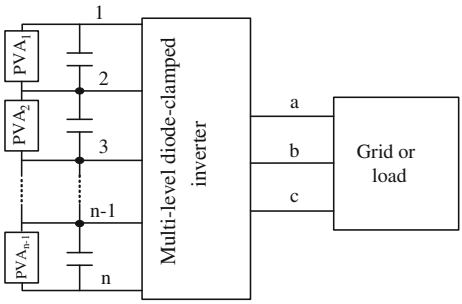


Fig. 247.3 Link between PV array and diode-clamp n-level inverter

Traditional two-level inverter without transformer, it can convert AC to DC, according to notion electric isolate requirement, the output voltage must exceed over the maximum DC link Voltage. In series PV arrays, the same current must flow through all PV arrays, the output voltage of every PV array cannot controlled separately, so the maximum energy cannot be got. By adding an extra circuit, different current can flow through every PV array, the PV array can trace the MPPT. By controlling voltage of every PV array, MPPT of solar PV generation system can be got, but the complexity of the system is also added.

Modulation strategy based on independent voltage controller for diode-clamp multilevel inverter is adopted. Besides the maximum energy form PV array, it also can decrease equipment voltage stress, and raise system efficiency. The link between PV array and diode-clamp n-level inverter can be seen in Fig. 247.3. The principle of independent voltage controller see Fig. 247.4. It mainly consists of DC bus voltage control, balance control, and modulation.

247.3.2.1 DC Bus Voltage Control Based on MPPT

Suppose variables Vpva1,Vpva2,…,Vpvan−1 are detecting voltage (i.e DC capacitance voltage), V*pva1,V*pva2,…,V*pvan−1 are corresponding control voltage which determined by MPPT algorithm. In this MPPT algorithm, the current of PV array is not necessary to consider, two phase current and two phase voltage of grid or load must be detected, and the third phase current (and the third phase voltage) can be counted by sum of three phase current equals zero, the optimal control voltage can be got by Perturbation observation algorithm.

DC bus voltage can be counted by detecting PV arrays voltage

$$V_{n1} = \sum_{k=1}^{n-1} V_{\mathrm{PVAK}} \tag{247.3}$$

A error can be got by comparing the detecting DC bus voltage and corresponding control voltage, through a compensator and a limiter, modulation index m can be got, $m \in (0, 1)$, so the length and angle of three phase output voltage vector can be controlled.

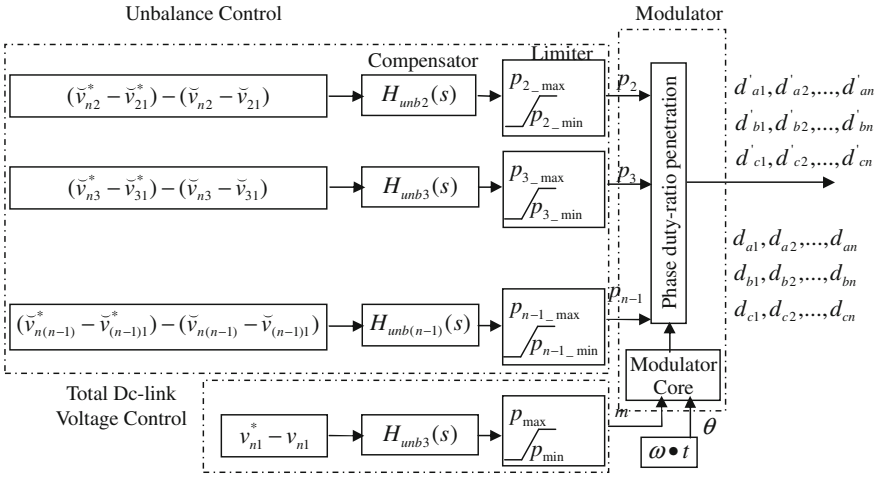


Fig. 247.4 Principle of independent voltage controller

247.3.2.2 Balance Control

To every inner DC bus node j (= 2, 3,…, n−1), DC voltage can be divided into two parts.

$$v_{j1} = \sum_{k=1}^{j-1} v_{\mathrm{PVAK}}, \quad v_{nj} = \sum_{k=j}^{n-1} v_{\mathrm{PVAK}} \tag{247.4}$$

The corresponding average voltage are

$$\bar{v}_{j1} = \frac{\sum_{k=1}^{j-1} v_{\mathrm{PVAK}}}{j-1}, \bar{v}_{nj} = \frac{\sum_{k=j}^{n-1} v_{\mathrm{PVAK}}}{n-j} \tag{247.5}$$

To every inner DC bus node j, a new error can be got by comparing the detecting voltage error and control voltage error, through compensator and limiter, the control parameter Pj can be got. The amplitude of phase current is decided by Pj, whether injection current to node j or extracted current form node j, is decided by whether Pj is positive or negative.

Combining DC bus voltage controller and balance controller, the unit controller can ensure the actual voltage of every PV array equal to corresponding control voltage, make PV generation system work in MMPT. Whether active inverter or passive inverter, all modulation strategy can be used in PV generation system, by adding or reducing control voltage of PV array to implement the valid control.

247.3.2.3 Modulation

Modulation is consisted of modulation module and phase duty-cycle disturb module. Variable phase duty-cycle is generated by modulation module according to the length and angle of three phase output voltage vector. SVPWM strategy is adopted to control voltage of every PV array, even though PV array voltage is inequality, no low frequency obstruction in output voltage.

Phase duty-cycle disturb module can modify phase duty-cycle disturb, to modified control voltage of PV arrays, modified phase duty-cycle are sent to assign unit, assign symmetry to every phase corresponding inner DC bus node, corresponding switch control signal of multilevel inverting in every period can be produced, the output voltage, and the voltage current can be got.

247.4 Simulation

To verify the viability and superiority of proposed control plan, the simulation models of independent PV generation system and PV grid-connected system has been built, some corresponding simulation experiments are executed.

247.4.1 Simulation of Independent PV Generation System Based on Three-Level Inverter

The simulation models of independent PV generation system based on three-level inverter has been built, the inputs of diode-clamp three-level inverter are two PV arrays PV1 and PV2. The main simulation parameters have been set as follows, three giving sinusoidal voltage amplitude is 50*sqrt(2), frequency $f = 50$ Hz, PV array voltage PV1 = PV2 = 400 V, load R = 5 Ω, L = 0.001 H

To give expression to dynamic control property for different PV array voltage, a voltage jump is observed. When $t = 0.03$ s, a voltage jump form (400 V, 400 V) to (430 V, 370 V) occurs, the total voltage is still maintained at 800 V. Two PV array voltage can be controlled independently, different voltage and different current of PV array can be allowed in every PV array, the output performance of system is good, see Fig. 247.5.

247.4.2 Simulation of PV Grid-Connected System Based on Three-Level inverter

Simulation model of PV grid-connected system mainly consists of some modules, such as photovoltaic cell PV module, three phase three-level inverter module, grid module, LC filter module, voltage controller module, current controller module, and SVPWM modulation module Fig. 247.6.

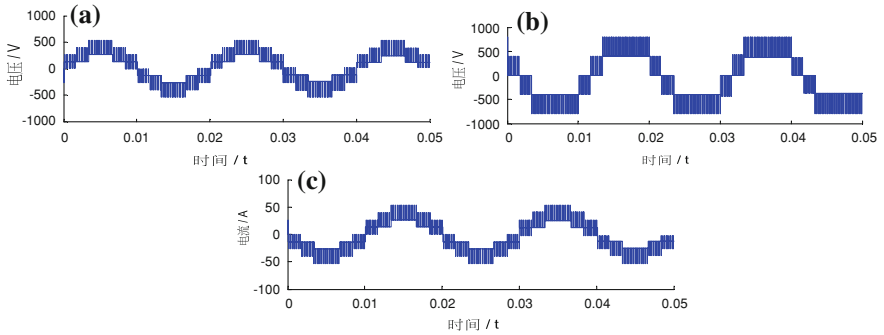


Fig. 247.5 Simulation results of three-level inverter when PV1, PV2 voltage varied. **a** load voltage **b** line voltage of three-level inverter **c** line current of three-level inverter

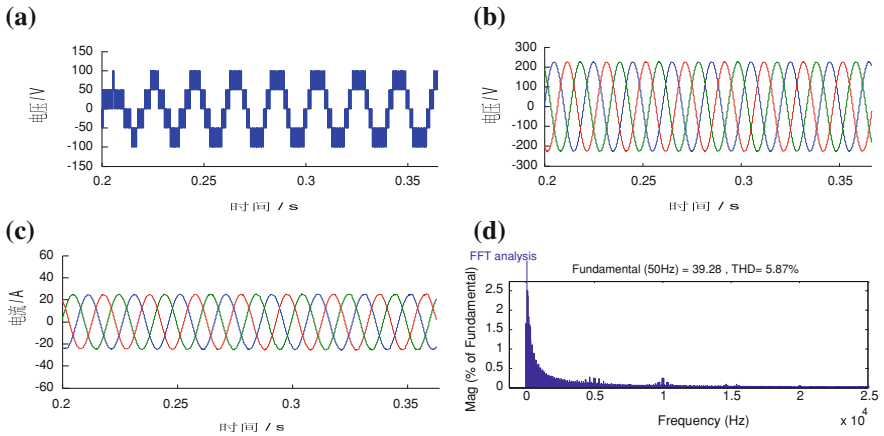


Fig. 247.6 Simulation results of PV grid-connected system based on three-level inverter. **a** Output line voltage. **b** Output voltage after filter. **c** The output current. **d** The spectrogram of Output current

The output current in wave form quality of three-level inverter is better. After harmonic analysis, the harmonic content is low, so it can meet grid requirements. Simulation results have verified the good steady characteristic PV grid-connected system based on diode-clamped three-level inverter.

247.5 Conclusion

A novel control strategy of PV generation system for diode-clamped multilevel inverters based on independent voltage control is presented in this paper. Every PV array voltage can be controlled independently, different current can be allowed

through PV array, so MPPT of system can be achieved. When part shade of PV array or mismatch of PV array, the output power can be increased by adopting independent control strategy. And the rated voltage of equipment is reduced, the power transformer can be saved, and the output harmonic distortion is reduced, and the conversion efficiency is increased.

References

1. Xiao W, Dunford WG, Palmer PR, Capel A (2007) Regulation of photovoltaic voltage. IEEE Trans Ind Electron 54(3):1365–1374
2. Xiao W, Ozog N, Dunford WG (2007) Topology study of photovoltaic interface for maximum power point tracking. IEEE Trans Ind Electron 54(3):1696–1704
3. Kimball J, Krein P (2008) Discrete-time ripple correlation control for maximum power point tracking. IEEE Trans Power Electron 23(5):2353–2362
4. Rodriguez J, Bernet S, Wu B, Pontt JO, Kouro S (2007) Multilevel voltage-source- converter topologies for industrial medium-voltage drives. IEEE Trans Ind Electron 54(6):2930–2945
5. Barkati S, Baghli L, Berkouk EM, Boucherit M (2008) Harmonic elimination in diode-clamped multilevel inverter using evolutionary algorithms. Electric Power Sys Res 78(10): 1736–1746

Chapter 248
Experimental Investigation on the Laser-Induced Breakdown Spectroscopy of Cu Sample

Ribo Ning, Qian Li and Songning Xu

Abstract In order to investigate the mechanism of laser interacting with materials and obtain the parameters related to the laser plasma, a pulsed XeCl laser (308 nm, ~200 mJ) was employed to irradiate Cu targets placed in air. The laser plasma spectra was obtained by grating spectrometer. Under the model of local thermodynamic equilibrium (LTE), the temperature of plasma was deduced to be nearly 1.7×10^4 K, using multi-line slope by relative intensities of characteristic lines. Spectral profile of Cu at 296.117 nm was found to be well represented by a symmetric Lorentz-type curve, and the main cause of broadening mechanism in the plasma was the collision broadening, which was caused by the collision between electron and atom.

Keywords Spectroscopy · Laser-induced breakdown spectroscopy · Multi-line slope · Excitation temperature · Broadening

248.1 Introduction

It is known that intense pulsed laser beams can be focused on a solid material generating ablation and formation of hot plasmas, which are used for several applications in the last few years. Laser-Induced Breakdown Spectroscopy (LIBS)

R. Ning (✉) · Q. Li · S. Xu
The School of Science, Shenyang Ligong University, No.6 Nan Ping Zhong Lu, Hunnan New Zone, Shenyang 110159, China
e-mail: ningribo@163.com

Q. Li
e-mail: liqian198331@163.com

S. Xu
e-mail: xsn_tiger@yahoo.com.cn

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_248,

is a new spectroscopic technique, which using pulsed lasers of high power density irradiate on material surface, causing material ablation, inducing plasma of laser. Through analyzing the laser plasma emission spectrum, the relevant information about the element and content in the samples can be obtained, the interaction mechanism inside the plasma can be investigated. LIBS is an analytical technique, which is able for real-time measurement of elements in solid, liquid and gas, for qualitative or quantitative analysis, for nearly non-destructive analysis. LIBS is also a technique which is simple, fast and can analyze a variety of elements simultaneously, it has unique advantages in the analysis of elements which are difficult to evaporate or difficult to stimulate. In recent years, after the development of the laser source and the instrument performance of spectrum, this technique shows capabilities in many principal fields of applications, such as: foundation research, archeology research, metallurgy mineral analysis and environmental analysis etc. [1–4]. In this paper, plasma was obtained by using XeCl laser to ablate Cu sample, emission temperature of the laser plasma was calculated by measuring the relative intensity of characteristic lines, the line profile of the characteristic lines was discussed.

248.2 Theory Part

Under the action of laser beam, the surface of the sample became hot by the absorption of photons, melting occured, the hot electrons escaped from the surface and formed free electrons. The melted sample which contained atoms, molecules, ions, clusters, particles, etc. expanded rapidly along the normal direction of the solid, formed the fog of plasma. The gas near the solid surface also broke down by the irradiation. The evaporation of the solid began at the time that the front edge of the pulse came to the solid, the fog heating to be ionized, and ultimately formed to be plasma. The process of the formation was shown in Fig. 248.1.

From the point of view in mechanism of interaction between laser and materials, there are two different laser-induced absorption mechanism in the formation of plasma. One is the inverse bremsstrahlung absorption mechanism, that free electrons gain kinetic energy from the laser beam, the ionization and excitation increased by collision with the neutral particles of the excited state or the ground state; the other is the excited particle photo ionization mechanism, especially for UV, the main course is the direct photo ionization of the excited particle in the steam.

Laser-induced plasma is a system with high temperature. In such a system, the material melts into particles, breaks down into molecules or atoms, the collisions between particles make the atoms and molecules ionized to ions, the molecules, atoms and ions can be distributed to various energy levels, the transition from high energy level to low energy level makes the plasma have strong emission spectra.

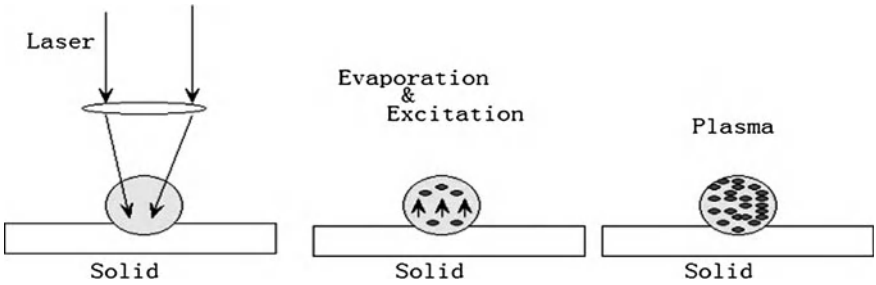


Fig. 248.1 Process of the plasma formation

In the local thermodynamic equilibrium (LTE), the distribution of the particles among the discrete levels of the energy spectrum is described by the Boltzmann formula

$$N_i = N_0 \frac{g_i}{g_0} e^{-\frac{E_i}{kT}} \tag{248.1}$$

where N_i is the number of atoms which is in the excited state, in volume unit; N_0 is the number of atoms which is in the ground state, in volume unit; g_i and g_0 are the statistical weight of the excited state and the ground state; E_i is the excitation energy of the excited state; k is the Boltzmann constant; T is the excitation temperature.

The atoms in the excited state are unstable, return to the ground state quickly, and eradiate spectra. Taking the Boltzmann distribution, partition function, transition probabilities and statistical weight of the excited atoms and ions in the source area, if the particle transit from the excited level i to the level m, then the radiation intensity in unit solid angle is:

$$I = \frac{h\nu_{im}}{4\pi} A_{im} N_i \tag{248.2}$$

where h is Plank's constant; ν_{im} is frequency of the emission lines; A is the atomic transition probability.

Equations (248.1) into (248.2), the emission intensity of a characteristic line, due to the transition $i \rightarrow m$, is linked to the population of the excited level through the following relation:

$$I = \frac{h\nu_{im}}{4\pi} N \frac{gA}{Z} e^{-E_i/kT} \tag{248.3}$$

The Eq. (248.3) can be written as:

$$\lg\left(\frac{I\lambda}{gA}\right) = -\frac{5040}{T} E + C \tag{248.4}$$

where λ is the line wavelength, C is constant.

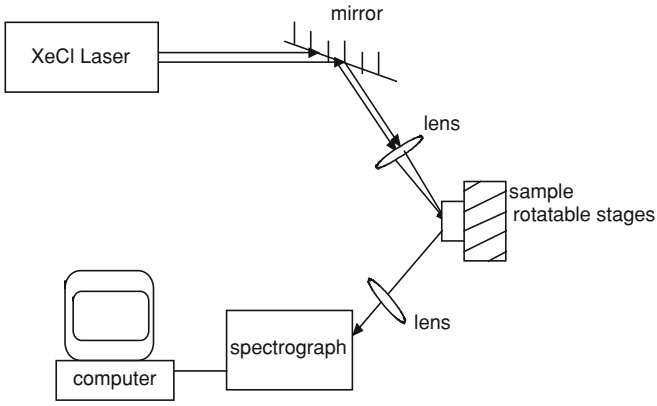


Fig. 248.2 The sketch map of the experimental set

When using multi-lines of one element to measure the excitation temperature, the wavelength λ and gA are known, the spectral relative intensity can be measured, then there are a range of lg(Iλ/gA) value. Constructing the excitation energy E as the abscissa, the value of lg(Iλ/gA) as the vertical. For a LTE plasma, a plot of lg(Iλ/gA) against the energy E for several spectral lines should be a straight line of slope, the slope of the line is equal to -5040/T, then the temperature T can be calculated [5].

248.3 Experimental and Results

248.3.1 Temperature Measurement

The XeCl laser was employed to bombard the Cu sample, the laser-induced plasma formed, the specific experimental set was shown in Fig. 248.2.

The XeCl excimer laser (308 nm, 180 mJ)was employed as a irradiation source, the laser pulse were reflected by the anti-mirror, then focused onto the sample surface by means of a fused silica plano-convex lens (UV grade) of focal length $f = 200$ mm, oblique incidence with angle of 30 degrees, and the laser spot diameter was about 1.5 mm. The elements of sample surface evaporated and stimulated rapidly by laser irradiation, accompanied by a strong crackling sound, then formed into plasma glow. The information of the plasma emission went into the WP1 grating spectrograph (focal length of 8.050 nm, width of slit 0.030 nm, center wavelength of 293.9 nm, slit tilt 4.35 mm) through the lens system. Spectral data processed by the software, the spectra was shown in Fig. 248.3.

The spectrum in the region 240–340 nm wavelength was shown in Fig. 248.3, the spectra was gained by using the XeCl excimer laser (pulse frequency of

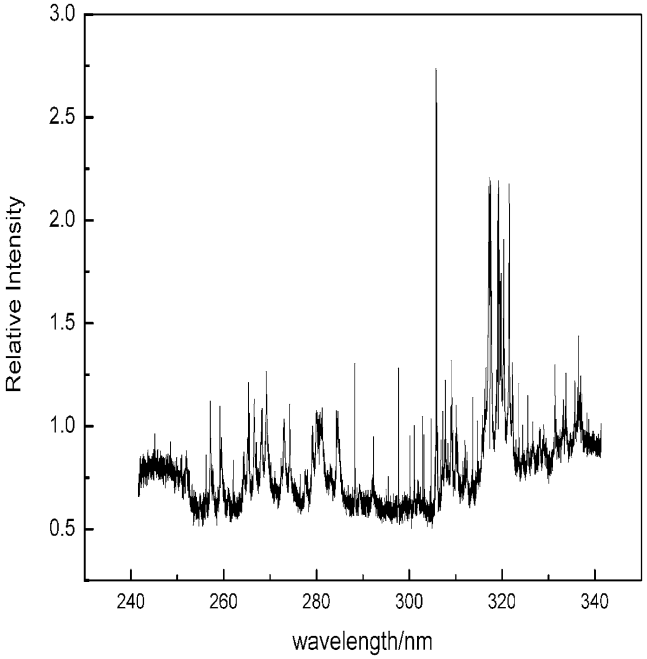


Fig. 248.3 Laser-induced plasma emission spectra of Cu

10 Hz), the sampling time was 30 s. It can be seen from the spectra that the LIBS technique was based on the laser-induced formation of plasma above the sample surface, followed by time-resolved analysis of its emission, where strong lines from excited atoms and ions were present, together with the plasma continuum heat radiation. In the shortwave region, the spectral intensity was weak, and the spectral intensity of the wavelength above 300 nm was much strong. Spectral intensity of four characteristic lines in the vicinity of the laser wavelength 308 nm significantly enhanced affected by the laser beam.

Selected five lines as characteristic line from the Cu once ionized ions lines, the wavelength of five characteristic lines were 2489.7, 2689.3, 2701.0, 2703.2, 2713.5 nm respectively. The data of the characteristic line was in the table below as follows Table 248.1.

The data of the characteristic line was from the CRC handbook of chemistry and physics [6]. Where λ was the wavelength of the characteristic line, g_i and g_k were the statistical weights of the lower (i) and upper (k) states, A was the transition probability.

The proceeded Boltzmann slash figure was drawn by using the relative intensity of the five characteristic lines above, in Fig. 248.4.

Boltzmann plot of the optical spectroscopy analysis gives a coronal plasma temperature. The slope of the Boltzmann fitting line was $k = -0.291531$. The excitation

Table 248.1 Data of the characteristic line of Cu(l)

λ/nm	g_i	g_k	$A/10^8 s^{-1}$
2489.7	5	5	1.5E−02
2689.3	7	7	4.1E−01
2701.0	5	5	6.7E−01
2703.2	3	3	1.2E+00
2713.5	5	5	6.8E−01

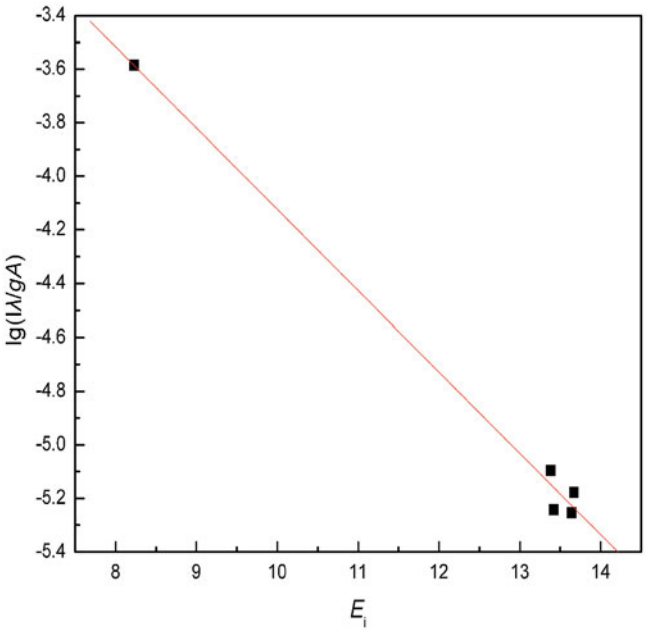


Fig. 248.4 Boltzmann slash figure

temperature of laser-induced breakdown plasma Cu was 17288.02 K, by using the multi-line slope method.

248.3.2 Line-Shape Analysis

The knowledge of the line-shape is important for exact intensity measurement. The shapes of Cu lines from XeCl laser-induced plasmas as a function of time have not yet been investigated. Line profiles are strongly influenced by Stark broadening, which usually results in shifted Lorentz profiles. In the case of pressure broadening the spectral line-shape can be well represented by an asymmetric Lorentz profile,

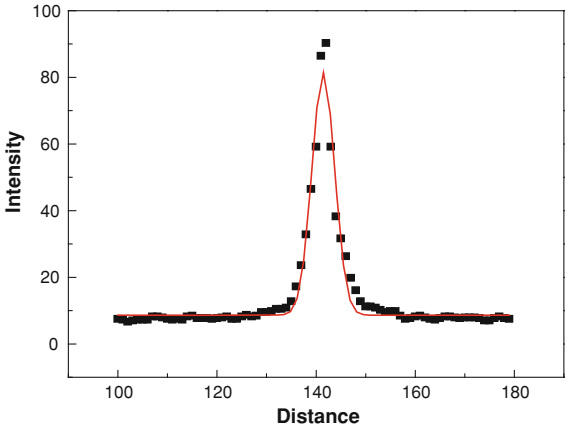


Fig. 248.5 Gauss fitting curve

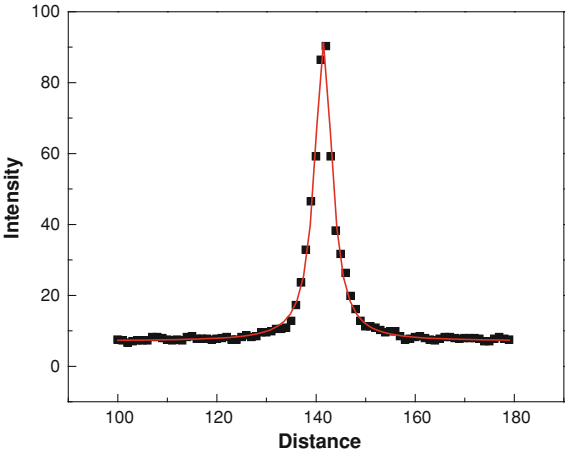


Fig. 248.6 Lorentz fitting curve

while resonance broadening results in line profiles of Lorentz shape with a symmetric core.

In the laser plasma, the contours of the emission spectral lines had a complex relationship with the environment in which the particles were, the electron density and the excitation temperature. The main mechanism of line broadening were Doppler broadening and Stark broadening, The line profile of Doppler broadening was symmetric Gauss line-shape, and the line profile of Stark broadening was basically Lorentz line-shape.

Selected the Cu (296.117 nm) spectrum as characteristic line, drew curve between the relative intensity and distance (the distance was read out as relative distance by using software). The curve fit in Gauss fitting was shown in Fig. 248.5. The curve fit in Lorentz fitting was shown in Fig. 248.6.

It can be seen from Fig. 248.5. that the actual line profile was much different from the Gauss fitting curve, the actual peak was higher than the fitting curve, and the bottom was higher too. The correlation coefficient was 0.96524, then the fitting complexion was poor with Gauss line-shape. From Fig. 248.6 the correlation coefficient was 0.99296, then the profile was basically consistent with Lorentz line-shape.

The results of the comparison showed that, the line slope did not fit the Gauss line-shape, then Doppler broadening corresponding with Gauss line-shape was not the main cause of broadening mechanism in the laser plasma, the Doppler effect can be ignored. The information that the line profile meet with Lorentz line-shape indicated that the main cause of broadening mechanism in the laser plasma was collision broadening, which caused by the collisions between electron and atom.

248.4 Conclusion

Selected the standard samples (Cu) as the research object, chose the XeCl excimer laser as the laser source, the laser-induced breakdown plasma was obtained by spectrometer. The excitation temperature of plasma was 17288 K by the Boltzmann slope method. The line-shape of the characteristic line at 296.117 nm fit Lorentz line-shape, indicated that the main cause of the broadening mechanism was the collision broadening. Exact line intensities can be calculated by line fits which is an important improvement in the technique for quantitative analysis. Further work is in progress.

References

1. Zhao S, Wei YH, Chen JZ et al (2003) Study on the emission spectra of aluminum plasma induced by high energy laser. Spectroscopy Spectral Anal 23:560–562
2. Corsi M, Cristoforetti C, Giuffrida M et al (2005) Archaeometic analysis of ancient copper artefacts by laser-induced breakdown spectroscopy technique. Microchim Acta 152:105–111
3. Yu S-j, Chen J-z, Su H-x et al (2010) Quantitative analysis of manganese, titanium and nickel in nickel-chromium-molybdenum-vanadium-titanium steel by using laser-induced breakdown spectroscopy. Metall Anal 30:11–15
4. Lin Z-x, Chang L, Li J et al (2009) Determination of As in industrial wastewater by laser-induced breakdown spectroscopy. Spectrosc Spectral Anal 29:1675–1677
5. Qiu D-r (2002) Atomic spectral analysis. Fudan University Press, Shangai, pp 32–38
6. CRC Handbook of Chemistry and Physics (2009) CRC Press, Boca Raton p 1575

Chapter 249
Application of Neuron-MOS and Pass Transistor to Voltage-Mode Ternary Logic Circuit

Guoqiang Hang, Yang Yang, Xiaohui Hu and Hongli Zhu

Abstract To realize voltage-mode ternary circuits, a novel design scheme using neuron-MOS transistors and pass transistors is presented. By controlling the voltage of the multiple-input gate, the neuron-MOS literal circuit is firstly designed. Then, the pass transistors used to pass ternary signal are controlled by the outputs of the literal circuit to realize ternary function. The complementarity and duality principles for generation of ternary complementary and dual circuits using pass transistors are also developed. The design results of ternary AND/NAND, OR/NOR, and mod-3 multiplication gates, demonstrate the effectiveness of the proposed scheme. The benefit of the proposed voltage-mode ternary circuits is that they can be fabricated by standard CMOS process with a 2-ploy layer. Besides, they have simple and perfectly symmetrical structure. The effectiveness of the proposed ternary circuits has been validated by HSPICE simulation results with TSMC 0.35 μm 2-ploy 4-metal CMOS technology.

Keywords CMOS circuits · Neuron-MOS transistor · Floating-gate MOS · Multiple-valued logic · Pass transistor

G. Hang (✉) · Y. Yang · X. Hu · H. Zhu
School of Informaton and Electrical Engineering, Zhejiang University City College, Hangzhou 310015, China
e-mail: hanggq@mail.hz.zj.cn

G. Hang
Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_249,

249.1 Introduction

The necessity for multiple-valued logic (MVL) has been pointed out by many research workers. Implementation of MVL using CMOS technology follows two major lines: namely, current-mode and voltage-mode circuits. Although some benefits arise from the current-mode approach, the obvious disadvantage of high power consumption exists due to the constant current flow. Conversely, voltage-mode MVL circuits have the advantage of low power dissipation at the expense of complex fabrication and high cost, since they require a complex process to obtain both depletion and enhancement devices [1–3] or multi-threshold voltages [4, 5]. Among various types of MVL, the ternary logic receives more attention than others because of a lower interconnection cost estimation and a simple electronic circuit implementation method [6]. Double pass-transistor logic (DPL) [7] has a balanced input capacitance. This reduces the dependence of the delay on the input data. The symmetrical circuit configuration results in high logic functionality. The DPL is used to attain a full rail-to-rail swing, and is very attractive for high speed and low power applications. Hang [8] introduced a design method for realizing ternary static circuits using DPL without any modification of the thresholds.

A neuron-MOS transistor [9] is a novel functional transistor with multiple-input gates, and is characterized by a variable threshold voltage achieved by controlling the voltages of the multiple-input gates [9, 10]. This feature is most suitable to use in constructing voltage-mode MVL circuits [9, 11]. Park et al. [12] proposed some quaternary logic gates using DPL with neuron-MOS threshold gate. In this paper, a new static and voltage-mode circuit scheme using neuron-MOS transistors and pass transistors to realize ternary logic gates is presented. The two threshold voltages of ternary circuit are generated using neuron-MOS literal circuit. The proposed voltage-mode ternary circuits can be fabricated by the standard CMOS process with a double-poly layer without the multi-level ion implantation process, which will make the circuit more practically. The proposed ternary circuits are verified by HSPICE simulation results with TSMC 0.35 μm 2-ploy 4-metal CMOS technology and 3.3 V power supply voltage.

249.2 Neuron-MOS Inverter and Ternary Literal Circuit

The complementary neuron-MOS inverter illustrated in Fig. 249.1a consists of an n-channel neuron-MOS and a p-channel neuron-MOS device, and they share the same floating gate. The floating-gate is formed by the first layer of polysilicon, while the multiple-input gates are formed by the second polysilicon layer and they are located over the floating gate. The multiple-input gates are capacitively coupled to the floating gate. The capacitive model is shown in Fig. 249.1b. The floating-gate potential of the complementary neuron-MOS inverter is given by [9]

$$\phi_{\mathrm{F(inv)}} = \frac{\sum_{i=1}^{\mathrm{n}} C_i V_i + C_{0\mathrm{p}} V_{\mathrm{DD}}}{\sum_{i=1}^{\mathrm{n}} C_i + C_{0\mathrm{p}} + C_{0\mathrm{n}}} \tag{249.1}$$

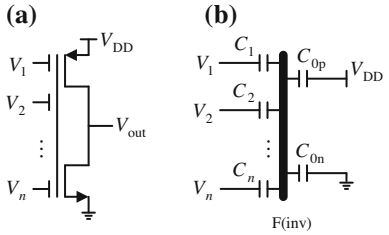


Fig. 249.1 Complementary neuron-MOS inverter
a Circuit diagram.
b Capacitive model.

where $V_i (i = 1, 2, \ldots, n)$ is the input signal voltage, $C_i (i = 1, 2, \ldots, n)$ is the capacitive coupling coefficient between the floating-gate and input gate, C_{0n} and C_{0p} denote the gate oxide capacitances of n- and p-channel neuron-MOS transistors, respectively. One of the most important features of this circuit is clearly expressed in Eq. 249.1, which states that the voltage signals are directly added at the gate level without any power dissipation. This voltage-mode summation is the most attractive feature of the capacitive coupling technique as compared to wired sum of currents in terms of reducing the power consumption [9]. Let the V^*_{inv} be the switching threshold voltage of the inverter as seen from the floating-gate. It is given by

$$V^*_{inv} = \frac{V_{DD} + V^*_{Tp} + \sqrt{\frac{\beta_n}{\beta_p}} V^*_{Tn}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}} \tag{249.2}$$

where V^*_{Tn} and V^*_{Tp} are the threshold voltages of the n- and p-channel neuron-MOS as seen from the floating gate, respectively. In Eq. 249.2, $\frac{\beta_n}{\beta_p} = \frac{\mu_n W_n / L_n}{\mu_p W_p / L_p}$, where μ is the mobility of majority carriers, W and L are the channel width and length of MOS transistor. From (249.1) and (249.2), the inverter threshold as seen from gate 1 is given by

$$V^{(1)}_{inv} = \frac{\sum_{i=1}^{n} Ci + C_{0p} + C_{0n}}{C_1} \left(V^*_{inv} - \frac{C_{0p} V_{DD}}{\sum_{i=1}^{n} Ci + C_{0p} + C_{0n}} \right)$$

$$- \frac{C_2}{C_1} V_2 - \cdots - \frac{C_n}{C_1} V_n \tag{249.3}$$

From Eq. 249.3, if an n-input complementary neuron-MOS inverter is regarded as a single-input inverter, where V_1 is the only signal input gate and the other gates are for threshold control, it is obvious that $V^{(1)}_{inv}$ is determined as a function of control signals $V_i (i = 2, \ldots, n)$. Such potential advantage of variable-threshold can be exploited in construction of the ternary voltage-mode circuits.

A literal circuit is the fundamental element in multiple-valued systems. Most voltage-mode MVL circuits consist of literal circuit. The truth table of ternary

Table 249.1 Truth table of the ternary literal functions

x	${}^{0}x^{0}$	$\overline{{}^{0}x^{0}}$	${}^{2}x^{2}$	$\overline{{}^{2}x^{2}}$
0	2	0	0	2
1	0	2	0	2
2	0	2	2	0

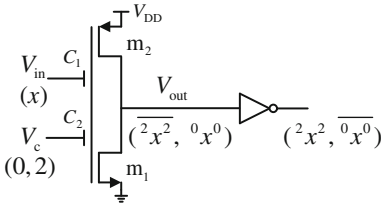


Fig. 249.2 Neuron-MOS literal circuit

literal functions is shown in Table 249.1. The voltage-mode ternary literal circuit using one complementary neuron-MOS inverter (m_1, m_2) with two-input gates, and one conventional CMOS binary inverter, is depicted in Fig. 249.2, where V_{in} (logic value $x \in (0, 1, 2)$) is taken as the input of the literal circuit, and V_c is the control voltage. According to Eq. 249.3, the switching threshold as seen from the input $V_{in}$is given by

$$V_{TC} = \frac{C_{TOT}}{C_1}\left(V_{inv}^{*} - \frac{C_{0p}V_{DD}}{C_{TOT}}\right) - \frac{C_2}{C_1}V_c \tag{249.4}$$

Where $C_{TOT} = C_1 + C_2 + C_{0p} + C_{0n}$. Therefore, V_{TC} of this circuit can be varied by the binary digital control signal V_c. The ternary literal functions ${}^{0}x^{0}$ and $\overline{{}^{2}x^{2}}$ can be implemented, when the logic value of V_c equals 2, and 0, respectively. The literal functions $\overline{{}^{0}x^{0}}$ and ${}^{2}x^{2}$ can be easily realized using only one binary inverter, as shown in Fig. 249.2. The proposed ternary literal circuit gives the ability to construct complex ternary gates.

249.3 Ternary Circuits Design

Based on the theory of transmission switches [4], the following two kinds of variable are introduced in multiple-valued circuits.

(i) Assume that α, β, … are switching variables. Their values are taken as T or F, only representing two kinds of states of MOS transistors, ON and OFF. The basic operations related to switching variables are AND ($\cdot$), OR (+) and NOT ($\sim$), which are expressed as $\alpha \cdot \beta$, $\alpha + \beta$ and $\tilde{\alpha}$ respectively, and are used to describe three physical situations: connections of element in serial, parallel and switching state inverse.

(ii) Assume that x, y, … are signal variables. Taking a ternary signal as an example, its logic value is taken as 0, 1 and 2, which correspond to three different voltage signals, such as 0, 1.65 and 3.3 V in physics. They have a clear meaning of magnitude and can be identified by comparing their magnitude with detection threshold t, $t \in \{0.5,\ 1.5\}$. For example, the two corresponding detection threshold voltages may take 0.825 and 2.475 V, respectively. The basic operations related to signal variables are minimum ($\cap$), maximum ($\cup$) and complement (>) and literal (ixi)operations, etc.

In order to describe the interaction between the on–off states of switching elements and signals, the following four kinds of basic operations can be introduced.

$$x^t = \Delta \begin{cases} \mathrm{T} & (\text{if } x < t) \\ \mathrm{F} & (\text{if } x > t) \end{cases} \tag{249.5}$$

$${}^t x = \Delta \begin{cases} \mathrm{T} & (\text{if } x > t) \\ \mathrm{F} & (\text{if } x < t) \end{cases} \tag{249.6}$$

These two equations represent the low-acting switching characteristic of a pMOS transistor and the high-acting switching characteristic of an nMOS transistor, respectively.

$$s \triangleright \alpha = \Delta \begin{cases} s & (\text{if } \alpha = \mathrm{T}) \\ \phi & (\text{if } \alpha = \mathrm{F}) \end{cases} \tag{249.7}$$

where s is the passed source signal, and α represents the switching state of a transmission switch network. If $\alpha = \mathrm{T}$, the signal s is passed to the output; If $\alpha = \mathrm{F}$, the switch network is off and its output is in the high-impedance state, donated by symbol ϕ. To express the joining of the outputs of two (or more) transmission branches, the following operation can be further defined.

$$s_1 \triangleright \alpha \# s_2 \triangleright \beta \stackrel{\Delta}{=} \begin{cases} s_1 \triangleright \alpha & (\text{if } \beta = \mathrm{F}) \\ s_2 \triangleright \beta & (\text{if } \alpha = \mathrm{F}) \end{cases} \tag{249.8}$$

In Eq. 249.8, the transmission operation $\triangleright$ takes priority over the union operation #. Also, if $s_1 \neq s_2$ and $\alpha = \beta = \mathrm{T}$, a voltage conflict arises between sources s_1 and s_2. This condition is not allowed.

The above theory of transmission switches lays a foundation for realizing the design technique of multiple-valued circuits using neuron-MOS and double pass-transistor at switch level. For example, Fig. 249.3a, b shows the Karnaugh map for a 2-variable ternary AND function and OR function, respectively. According to the coverage shown in Fig. 249.3a, b the following transmission operations can be obtained.

$$x \cap y = 0 \triangleright {}^{0.5}({}^0x^0) \# x\big|_{x=1,2} \triangleright (\overline{{}^2y^2})^{0.5} \# y\big|_{y=0,1} \triangleright [{}^{0.5}(\overline{{}^2y^2}) \cdot {}^{0.5}(\overline{{}^0x^0})] \tag{249.9}$$

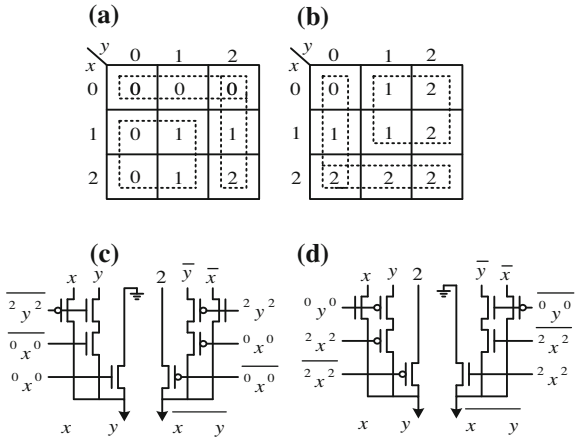


Fig. 249.3 Ternary circuits **a** K-map for ternary AND operation. **b** K-map for ternary OR operation. **c** Ternary AND/NAND gate. **d** Ternary OR/NOR gate.

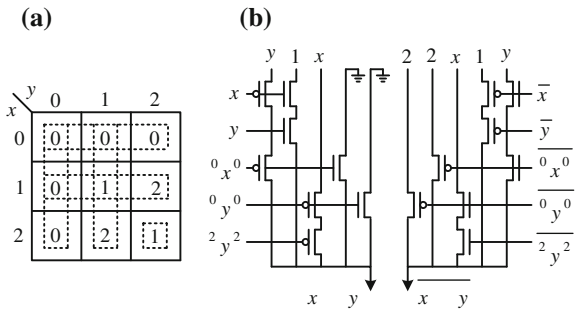


Fig. 249.4 Ternary mod-3 multiplication circuit **a** K-map **b** Circuit realization

$$x \cup y = 2 \triangleright (\overline{{}^{2}x^{2}})^{0.5} \# x|_{x=0,1} \triangleright {}^{0.5}({}^{0}y^{0}) \# y|_{y=1,2} \triangleright [({}^{2}x^{2})^{0.5} \cdot ({}^{0}y^{0})^{0.5}] \quad (249.10)$$

The circuit realizations corresponding to the above two equations are shown in Fig. 249.3c, d respectively, where some input literal signals are yielded by circuit shown in Fig. 249.2. In fact, as AND-OR is the dual logic function, ternary OR circuit can be formed from ternary AND circuit by using the duality principle. The dual logic function in ternary pass circuit can be generated after three-step modifications [8]: (i) swap pMOS and nMOS transistors; (ii) swap logic 2 and logic 0, logic 1 remains unchanged; (iii) swap ${}^{i}x^{i}$ and $\overline{{}^{i}x^{i}}$ $(\overline{{}^{i}x^{\bar{i}}})$, ${}^{i}y^{i}$ and $\overline{{}^{i}y^{i}}$ $(\overline{{}^{i}y^{\bar{i}}})$, etc. And also, using the complementarity principle, the complementary logic function in ternary pass logic can be generated after modifications of the true function: (i) swap pMOS and nMOS transistors; (ii) invert all pass and gate signals.

As another example, Fig. 249.4a gives the Karnaugh map for ternary mod-3 multiplication. According to the coverage shown in Fig. 249.4a, the following transmission operation can be obtained.

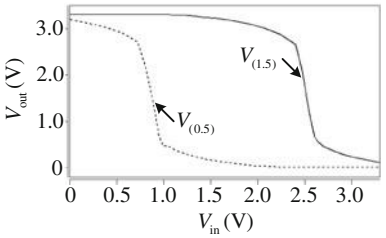


Fig. 249.5 Simulated transfer characteristic of ternary literal circuit

$$x \otimes y = 0 \triangleright [{}^{0.5}({}^{0}y^{0}) + {}^{0.5}({}^{0}x^{0})]\#x|_{x=1,2} \triangleright [({}^{0}y^{0})^{0.5} \cdot ({}^{2}y^{2})^{0.5}] \\ \#y|_{y=2} \triangleright [({}^{0}x^{0})^{0.5} \cdot x^{1.5}]\#1 \triangleright ({}^{1.5}x \cdot {}^{1.5}y) \tag{249.11}$$

Based on Eq. 249.11, the logic function $x \otimes y$ is firstly mapped into circuit realization. Then using the same circuit structure, swapping pMOS and nMOS transistors, and inverting all pass and gate signals, its complementary function $\overline{x \otimes y}$ can be easily obtained as shown in Fig. 249.4b.

As can be seen in Figs. 249.3 and 249.4, they are perfectly symmetrical. This results in a balanced input capacitance and reduces the dependence of the delay time on data. The symmetrical structure makes it easy to realize other ternary complex functions. The behavior of the proposed ternary circuits is straightforward. In the circuit, the ternary logic 2, 1 and 0, are corresponding to voltages V_{DD}, $V_{DD}/2$ and GND, respectively.

249.4 Simulation Results

All the proposed circuits have been simulated using HSPICE and TSMC 0.35μm 2-poly 4-metal CMOS process parameters, with 3.3 V power supply. Figure 249.5 shows the transfer characteristic of the literal circuit given in Fig. 249.2. In simulation, we take $C_1 = 400$ fF and $C_2 = 200$ fF. The aspect rations of transistors are W/L = 5.6/0.7 μm for nMOS, and W/L = 14/0.7 μm for pMOS, respectively. From the simulation results, we obtain $V_{(0.5)} = 0.86$ V, when $V_c = 3.3$ V, and $V_{(1.5)} = 2.51$ V, when $V_c = 0$ V. Therefore, the discrepancy between the simulated and theoretical results is smaller than 5%.

Taking ternary AND/NAND gate shown in Fig. 249.3c as example, the transient simulation results with a 0.1pF load at each output node, is given in Fig. 249.6. In the simulation, the voltages (0, 1.65 V, 3.3 V) are taken to represent logic values (0, 1, 2). From the simulation results, it has been demonstrated that the full logic swing and the maximum possible noise margins are attained due to the usage of both pMOS and nMOS pass transistors. A significant feature of the proposed circuits is the use of the standard CMOS process without any modification of the thresholds.

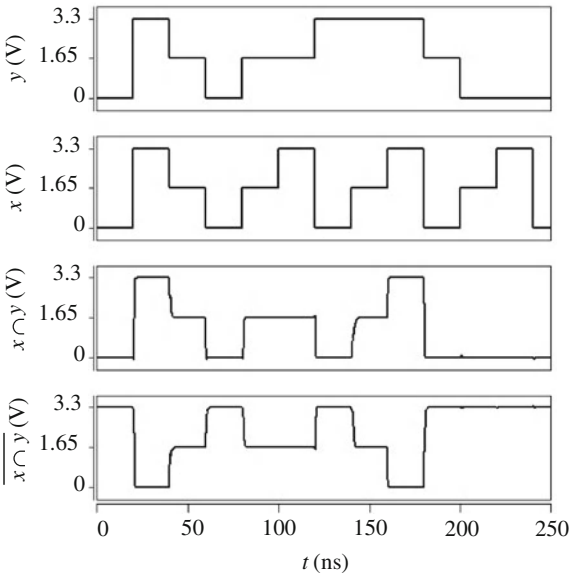


Fig. 249.6 Simulated transient characteristics of ternary AND/NAND gate

249.5 Conclusions

A new kind of static circuit scheme for voltage-mode ternary logic based on neuron-MOS transistors and pass transistors, is proposed. In the proposed design scheme, literal functions are realized by using neuron-MOS transistors. Then, the pass transistors used to pass ternary signal are controlled by the outputs of the literal circuit to realize ternary functions. The way for generation of complementary and dual circuits is also presented. Compared to the traditional voltage-mode MVL circuits, the proposed ternary circuits have several advantages. The circuit can be fabricated by standard CMOS process with a double-ploy layer, without multi-level ion implantation applied in the conventional voltage-mode multiple-valued circuits. As the variable threshold voltage can be achieved easily in neuron-MOS literal circuit, the proposed ternary circuits have simple structure. In addition, they have perfectly symmetrical structure, attain the full logic swing and the maximum possible noise margins. HSPICE simulation results using TSMC 0.35 μm double-poly 4-metal CMOS process parameters have verified the principle of operation and the characteristics of the proposed scheme. Furthermore, the proposed method can be extended to generate larger ternary logic functions and it can also be used to synthesize quaternary logic circuits.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 60971061), the Natural Science Foundation of Zhejiang Province (No. Y106375), and the construct program of the key discipline in Hangzhou

References

1. Heung A, Mouftah HT (1985) Depletion/enhancement CMOS for a low power family of three-valued logic circuits. IEEE J SSC 20(2):609–615
2. Thoidis I, Soudris D, Fernandez J et al (2001) The circuit design of multiple-valued logic voltage-mode adders. In: Proceedings of ISCAS, vol 4, pp 162–165
3. Babu H, Islam M, Ali A et al (2003) A technique for logic design of voltage-mode pass transistor based multi-valued multiple-output logic circuits. In: Proceedings of ISMVL, pp 111–116
4. Wu X, Prosser F (1990) CMOS ternary logic circuits. IEE Proc 137(G1):21–27
5. Philippe JM, Pillement S, Sentieys O (2005) A low-power and high-speed quaternary interconnection link using efficient converters. In: Proceedings of ISCAS, IV, pp 689–692
6. Hurst SL (1984) Multiple-valued logic—its status and its future. IEEE Trans Comput C33(12):1160–1179
7. Suzuki M, Ohkubo N, Shinbo T et al (1993) A 1.5-ns 32-b CMOS ALU in double pass-transistor logic. IEEE J SSC 28(11):1145–1151
8. Hang G (2006) A general method in the synthesis of ternary double pass-transistor circuits. Chinese J Semicond 27(9):1566–1571
9. Shibata T, Ohmi T (1992) A functional MOS transistor featuring gate-level weighted sum and threshold operations. IEEE Trans on Electron Device 39(6):1444–1455
10. Chavan A, MacDonald E, Liu N et al (2008) A novel floating gate circuit family with subthreshold voltage swing for ultra-low power operation. In: Proceedings ISCAS, pp 3354–3357
11. Shen J, Tanno K, Ishizuka O (1999) Down literal circuit with neuron-MOS transistors and its applications. In: Proceedings of ISMVL, pp 180–185
12. Park SJ, Yoon BH, Yoon KS et al (2004) Design of quaternary logic gate using double pass-transistor logic with neuron MOS down literal circuit. In: Proceedings of ISMVL, pp 198–203

Chapter 250
Simulation Analysis of Structure Parameters on Long-Period Fiber Grating

Baishun Su

Abstract Coupling mode theory is adopted in the single-mode step-index optical fiber based on the three-layer modal in the application of long-period fiber grating. This paper is focused on the regularity for the change of effective refractive index core fundamental modes and cladding modes with the wavelength, analysis of the character of the coupling constant between cladding mode in different number of the first order and the core mode by numerical simulation. It can provide theoretical reference for the structure parameters design of the long-period fiber grating.

Keywords Long-period grating · Resonant wavelength · Coupling constant · Effective refractive index

250.1 Introduction

Long-period fiber gratings (LPFG) have been developed and proposed for use innumerous applications in fiber-optic sensor field [1, 2]. LPFG have possess high sensitivities and can be easily configured for parameter measurements, which can be widely used in the components for realizing wavelength dispersion, conversion, modulation, and control of guided wavefronts in optical integrated circuits.

B. Su (✉)
School of computer science and technology, Henan polytechnic university, Jiaozuo 454000, China
e-mail: subaishun@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_250,

Mode coupling between a core mode and a cladding mode occurs in the LPFG at a resonant wavelength λ_v, which satisfies the following phase-matching condition in Eq. 250.1.

$$\beta_0^{co} - \beta_v^{cl} = 2\pi/\Lambda. \tag{250.1}$$

where β_0^{co} and β_v^{cl} are the propagation constants of the fundamental mode and the vth-number cladding mode respectively. Λ is the grating period of the LPFG. The constant of propagation depending on the wavelength, each LPFG couples the fundamental core mode with several cladding modes at resonant wavelengths. For a given LPFG, the grating period determines the cladding modes to which light can be coupled [3–5].

There have been several theoretical studies on the coupling of the core mode to the cladding and radiation. For the coupling to the cladding mode, coupled-mode theory has proved to be a powerful tool for simulation of LPFG structures. We can use the program to perform numerical simulations which are based on solving coupled mode equations that describe the interaction of guided modes by a three layer model [6–9].

250.2 Simulation and Analysis

In this section, we analyze the factor affecting resonant wavelength change. The sketches of the LPFG under the definition of the parameters are given. The gratings are written in a single mode fiber with core radius r1 = 4.15 μm, cladding radius r2 = 62.5 μm, core refractive index n1 = 1.458, cladding refractive index n2 = 1.45, wavelength λ = 1,550 nm. The surrounded by media with different refractive indices.

250.2.1 Solution of the Effective Index of Fundamental Core Mode

A theoretical model is based on scalar approximation analysis to obtain the LP modes of a cylindrical dielectric waveguide. Eigen equation is as follows Eq. 250.2:

$$V\sqrt{1-b}\frac{J_1(V\sqrt{1-b})}{J_0(V\sqrt{1-b})} = V\sqrt{b}\frac{K_1(V\sqrt{b})}{K_0(V\sqrt{b})} \tag{250.2}$$

where $V = (2\pi/\lambda)r_1\sqrt{n_1^2 - n_2^2}$ is the normalized frequency, $b = (n_{eff}^2 - n_2^2)/(n_1^2 - n_2^2)$, b is the normalized effective refractive index, J_0 and J_1 are zero and first order Bessel function of the first kind, K0 and K1 are zero and first order modified Bessel function of the second kind [10].

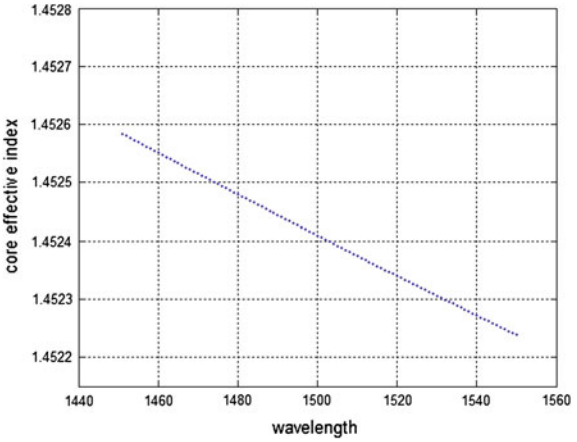


Fig. 250.1 Change of core effective index with resonant wavelength

The change of the effective index of LPFG can be seen in Fig. 250.1. Figure 250.1 indicate that the effective index of the core fundamental modes decreases with the wavelength increase, and effective index change is 4×10^{-4} in the range from 1450 to 1550 nm.

250.2.2 Solution of the Effective Index of Cladding Mode

The solution effective index of cladding mode can be obtained from the following Eq. 250.3.

$$\xi_0 = \xi_0^{'} \tag{250.3}$$

where ξ_0 and $\xi_0^{'}$ are defined by Eqs. 250.4 and 250.5 respectively.

$$\xi_0 = \frac{1}{\sigma_2} \frac{u_2(JK + \frac{\sigma_1\sigma_2 u_{21} u_{32}}{n_2^2 a_1 a_2}) p_l(a_2) - Kq_l(a_2) + Jr(a_2) - \frac{1}{u_2} s_l(a_2)}{-u_2(\frac{u_{32}}{n_2^2 a_2} J - \frac{u_{21}}{n_1^2 a_1} K) p_l(a_2) + \frac{u_{32}}{n_1^2 a_2} q_l(a_2) + \frac{u_{21}}{n_1^2 a_1} r_l(a_2)} \tag{250.4}$$

$$\xi_0^{'} = \sigma_1 \frac{u_2(\frac{u_{32}}{a_2} J - \frac{n_2^2 u_{21}}{n_2^2 a_1} K) p_l(a_2) - \frac{u_{32}}{a_2} q_l(a_2) - \frac{u_{21}}{a_1} r_l(a_2)}{u_2(\frac{u_3^2}{n_2^2} JK + \frac{\sigma_1\sigma_2 u_{21} u_{32}}{n_2^2 a_1 a_2}) p_l(a_2) - \frac{u_3^2}{n_1^2} Kq_l(a_2) + Jr(a_2) - \frac{n_2^2}{n_1^2 u_2} s_l(a_2)} \tag{250.5}$$

where

$$\sigma_1 = iln_{eff}/Z_0, \sigma_2 = iln_{eff} Z_0, Z_0 = \sqrt{\mu_0/\varepsilon_0} = 376.73\,\Omega, u_{21} = 1/u_2^2 - 1/u_1^2,$$
$$u_{32} = 1/w_3^2 + 1/u_2^2, u_1^2 = k_0^2(n_1^2 - n_{eff}^2), u_2^2 = k_0^2(n_2^2 - n_{eff}^2), w_2^2 = k_0^2(n_{eff}^2 - n_3^2),$$
$$J = J_l^{'}(u_1 a_1)/(u_1 J_l(u_1 a_1)), K = K_l^{'}(w_3 a_2)/(w_3 K_l(w_3 a_2)).$$

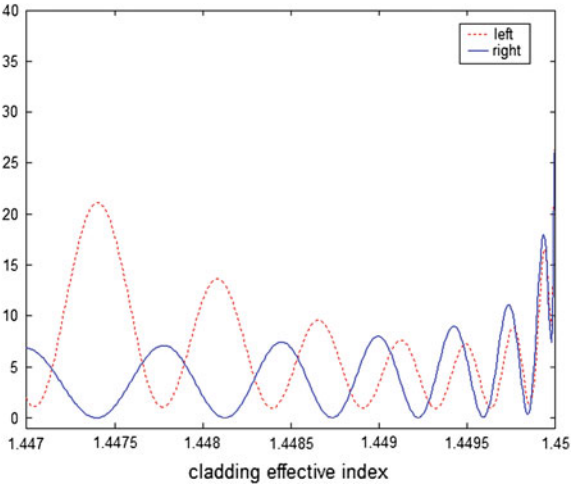


Fig. 250.2 Solution of the cladding effective index by graphic method

$$\begin{aligned} p_l(r) &= J_l(u_2 r)N_l(u_2 a_1) - J_l(u_2 a_1)N_l(u_2 r) \\ q_l(r) &= J_l(u_2 r)N_l^{'}(u_2 a_1) - J_l^{'}(u_2 a_1)N_l(u_2 r) \\ r_l(r) &= J_l^{'}(u_2 r)N_l(u_2 a_1) - J_l(u_2 a_1)N_l^{'}(u_2 r) \\ s_l(r) &= J_l^{'}(u_2 r)N_l^{'}(u_2 a_1) - J_l(u_2 a_1)N_l^{'}(u_2 r) \end{aligned} \tag{250.6}$$

The following parameters of the optical fiber in Eq. 250.6:

(1) The radius of the core a1, the radius of the cladding a2, and the external surrounding refractive index n3.

(2) *J* the Bessel function of first kind, K Modified Bessel function of second kind and *N* Bessel function of second kind.

According to (250.4, 250.5), Eq. (250.3) can be numerical solved by graphic method with matlab. This equation is a transcendental equation. and we can draw the graphics on both sides of the equation with by gradual approaching method. The left and right curve can be seen in Fig. 250.2 (left = ξ_0, right = $\xi_0^{'}$). The crss points values can be calculated accurately.

The relation between first order lower cladding modes and resonant wavelength can be seen in Fig. 250.3. The effective index of the cladding modes also decreases with the wavelength increase as the same as fundamental core modes. The change rates of effective refractive index and the increase according to higher number are bigger than that according to lower number [11, 12].

250.2.3 Solution of the Coupling Constant of Long-Period Grating

The solution of the coupling constant of core mode and cladding mode can be obtained from the following Eq. 250.7.

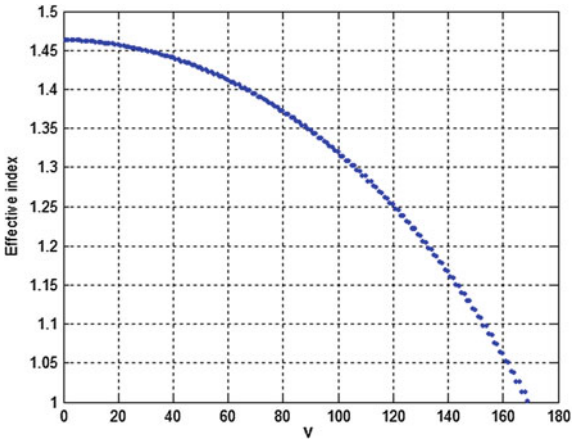


Fig. 250.3 Change of the cladding effective refractive index with different number

$$\kappa_{1v-01}^{cl-co} = \sigma(z)\frac{2\pi}{\lambda}\left(\frac{\pi b}{Z_0 n_2\sqrt{1+2b\Delta}}\right)^{1/2} \times \frac{n_1^2 u_1}{u_1^2 - V^2(1-b)/a_1^2} \times (1+\frac{\sigma_2\zeta_0}{n_1^2})E_{1v}^{cl}\left[u_1 J_1(u_1 a_1) \times \frac{J_0(V\sqrt{1-b})}{J_1(V\sqrt{1-b})} - \frac{V\sqrt{1-b}}{a_1} J_0(u_1 a_1)\right] \quad (250.7)$$

where κ_{1v-01}^{cl-co} is crossing coupling constant, E_{1v}^{cl} is normalization constant and can be defined as the following Eq. 250.8.

$$P = P_{co} + P_{cl} + P_{ex} = \frac{1}{2}\mathrm{Re}\int_0^{2\pi} d\varphi \int_0^{\infty} rdr(E_r^{cl} H_\phi^{cl*} - H_r^{cl*} E_\phi^{cl}) = 1 \quad (250.8)$$

where P_{co}, P_{cl} and P_{ex} are power carried in the core, cladding of the fiber and external region. For a given fiber parameter and a fixed light wavelength, the coupling constant can be calculated with formula (250.7) and (250.8). Its curve can be showed in Fig. 250.4 [13].

As can be seen in Fig. 250.3, the coupling coefficient for even cladding mode and odd cladding mode increased with the number. The coupling between the lowest-number even cladding modes and the core mode is very weak compared to that involving the lowest-number odd cladding modes. The rise rate and increment according to odd cladding mode are bigger than that according to even cladding mode. so even cladding mode can be ignored. It reach the maximum value when number is 25, then the odd mode coupling coefficient begin to reduce while the even mode coupling coefficient continue to increase. the 40 number coupling coefficient reduce to a local minimum value and alternately repeatedly change in the form of gradual increasement and decreasement, difference of the coupling coefficient are not significant for the adjacent mode. They are likely to be coupled with the core mode, which will appear the double peaks structure in transmission spectrum [14].

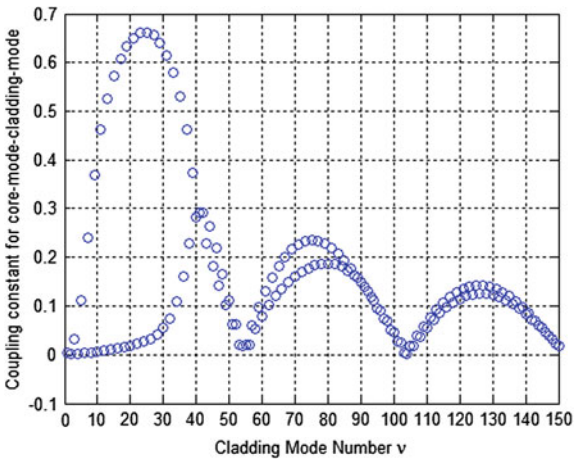


Fig. 250.4 Change of coupling constant with the cladding mode number v

250.3 Conclusion

In this paper, we solve the core effective refractive index numerically by using the weakly guiding approximation. The effective index of the discrete guided cladding mode in LPFGs based on the three-layer optical waveguide modal with the dispersion equation is simulated and analyzed. The numerical results has formed the basis of the research which contribute to further work on the transmission spectra change with the grating period and grating length [15, 16].

References

1. Vergsarker AM, Lemaire PJ, Judkins JB et al (1996) Long-period gratings as band–rejection filters. J Lightwave Technol 14(1):58–65
2. Huang W, Xu C, Chaudhuri SK (1991) Modeling and analysis of fiber-optic mode transducers: single fiber with periodic perturbations. J Lightwave Technol 9(11):1431–1438
3. Erdogan T (1997) Fiber grating spectra. J Lightwave Technol 15(8):1277–1294
4. Erdogan T (1997) Cladding-mode resonances in short- and long-period fiber grating filters. J Opt Soc Am A 14(8):1760–1773
5. Erdogan T (2000) Cladding-mode resonance in short- and long-period fiber grating filter: errata. J Opt Soc Am A 17:2113
6. Tsao C, Payne D, Gambling W (1989) Modal characteristics of the three-layered optical fiber waveguide: a modified approach. J Opt Soc Am A 17(11):555–563
7. Wang L, Erdogan T (2001) Layer peeling algorithm for reconstruction of long-period fibre gratings. Electr Lett 37(3):154–156
8. Song N, Mu J, Huang W-P (2010) Application of the complex coupled-mode theory to optical fiber grating structures. J Lightwave Technol 28(5):761–767
9. Koyamada Y (1999) Numerical analysis of core-mode to radiation-mode coupling in long-period fiber gratings. IEEE Photonics Technol Lett 13(4):308–310
10. Falciai R, Mignani AG, Vannini A (2001) Long period gratings as solution concentration sensors. Sens Actuators B 74(13):74–77

11. Anemogiannis E, Glytsis EN, Gaylord TK (2003) Transmission characteristics of long-period fiber gratings having arbitrary azimuthal/radial refractive index variations. J Lightwave Technol 21(1):218–227
12. Patrick HJ, Kersey AD, Bucholtz F (1998) Analysis of the response of long period fiber gratings to external index of refraction. J Lightwave Technol 16(9):1606–1612
13. Jeong H, Oh K (2003) Theoretical analysis of cladding-mode waveguide dispersion and its effects on the spectra of long-period fiber grating. J Lightwave Technol 21(8):1838–1845
14. Lam CCC, Mandamparambil R, Sun T et al (2009) Optical fiber refractive index sensor for chloride Ion monitoring. IEEE Sens J 9(5):525–532
15. Shu X, Zhang L, Bennion I (2002) Sensitivity characteristics of long-period fiber gratings. J Lightwave Technol 20(2):255–266
16. Villar ID, Matias IR, Arregui FJ (2005) Long-period fiber gratings with overlay of variable refractive index. IEEE Photonics Technol Lett 17(9):1893–1895

Chapter 251
The Simulation of the Principle of MPPT Control of PV Systems Based on BOOST Circuit

Haizhu Yang, Mingzhe Shang and Jie Liu

Abstract In the photovoltaic power generation system, in order to use solar energy maximize, to improve the efficiency of the system, we must consider the maximum power tracking. This paper is about the basic principle and simulation of MPPT based on BOOST circuit impedance matching, by analyzing the basic working principle of BOOST circuit, the principle of achieving maximum power tracking and process, to make the modeling and simulation analysis, simulation results show that the BOOST circuit impedance matching can achieve the maximum power tracking.

Keywords PV system · BOOST circuit · MPPT · Emulation

251.1 Introduction

Energy is the basic driving force of human social development and economic growth, and the material base of human survival. However, human beings enjoy energy with all kinds of interests, but also meet a series of inevitable energy security challenges, such as energy shortage and excessive use of fossil energy cause environmental pollution problems that have a serious threat to human survival and development. These issues will eventually forced people to change the

H. Yang (✉) · J. Liu
Department of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, China
e-mail: yanghaizhu@hpu.edu.cn

M. Shang
College of Electrical Engineering, Dalian University of Technology,
Dalian, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_251,

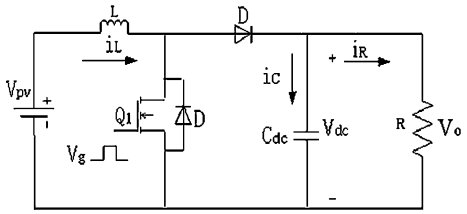


Fig. 251.1 BOOST converter

existing energy structure and rely on solar energy and renewable clean energy sources to solve.

MATLAB/SIMULINK tool can be used in the simulation of complex systems, because of its powerful and quick modular modeling environment, and increasingly attention paid by people. This paper is based on MATLAB simulation environment, first established solar modules model, and then according to the basic principle of BOOST circuit impedance matching, builds the simulation model and simulation analysis, so as to achieve the maximum power tracking capabilities [1].

251.2 The Basic Principle of BOOST Circuit

BOOST circuit is a dc converting circuit, which includes switching tube Q_1, inductance L, diode D and capacity C_{dc}, as shown in Fig. 251.1. According to the inductance current in the cycle beginning whether starts from scratch, whether is successive, there are two kinds of work modes: inductance current discontinuous mode (DCM) and inductance current continuous mode (CCM). Due to the discontinuity of inductive current while the circuit works in the discontinuous state, it means that there is a portion of the electrical energy outputted by solar energy in each cycle is wasted, and the ripple can be bigger, so the circuit parameter selection should satisfy the continuous conductive mode when circuit works [2, 3]. This design makes BOOST circuit work in continuous mode. When it is in inductance current continuous mode, there are two working conditions, as shown in the Figs. 251.2 and 251.3.

Switching tube state 1: Q is conductive, the all input voltage is added in inductance L, and the inductance current increase linearly, when the diode is closed and the load is provided energy by capacity.

Switching tube state 2: Q is off, the diode is positive biased and conductive, and power consumption and the energy stored up by the inductance L are fed to the load and filter capacitance C.

The relationship between input and output is:

$$\frac{V_o}{V_{\text{pv}}} = \frac{1}{1-D} \tag{251.1}$$

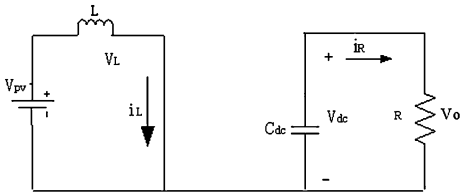


Fig. 251.2 Switch conduction

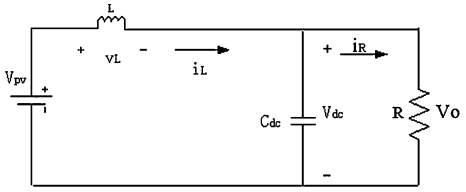


Fig. 251.3 Switch closed

where
V_O is output voltage, namely the load on both ends of the voltage;
V_{pv} is input voltage, namely photovoltaic array on both ends of the voltage;
D is duty cycle.

251.3 BOOST Circuit Implements the Basic Principle of MPPT

Now most methods are based on the established simulation model of solar cells, and establish circuit simulation models through the state space method or add power electronic devices. But this method in determining its characteristics of inside circuit and designing its structural parameters is more difficult, and the time of simulation is longer. This paper is based on analyzing BOOST circuit, and establishes simulation model based on the impedance transformation to realize the maximum power tracking. Relative to the former speaking, this model does not require very precise related parameters and circuit characteristics, and can achieve most power tracing [4].

Photovoltaic array connects BOOST circuit. If the output of BOOST circuit is a resistive load and its conversion efficiency is 100%, namely its input power and output power are equal, also ignore the impedance of BOOST transform circuit itself:

As the type 1 we can know:

$$\frac{V_O}{R_L} = \frac{\left(\frac{V_{pv}}{1-D}\right)^2}{R_L} = \frac{V_{pv}^2}{R_L(1-D)^2} = \frac{V_{pv}}{R_D} \tag{251.2}$$

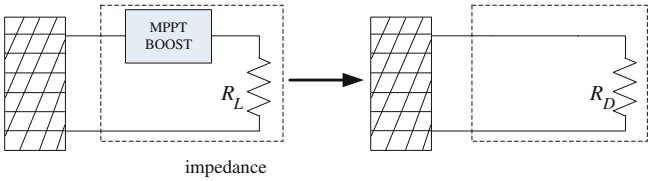


Fig. 251.4 BOOST circuit impedance transformation

That is

$$R_D = R_L(1-D)^2 \tag{251.3}$$

where

D is duty cycle, R_L is load impedance, R_D is equivalent input impedance of BOOST circuit.

As the type 3 we can know: adjusting D can adjust the equivalent input impedance of BOOST circuit, and let it match with the photovoltaic array output impedance to make the output power largest, as shown in the Fig. 251.4.

251.4 Modeling and Simulation of Solar Cells

251.4.1 Mathematical Model of Solar Cell

The solar cells mathematical model which is used for MATLAB modeling in this paper is as follows: under any light intensity $R(\mathrm{w{\cdot}m^{-2}})$ and ambient temperature $Ta(^{\circ}C)$ conditions, the solar cell temperature $Tc(^{\circ}C)$ is [5]:

$$T_c = T_a + R \cdot t_c \tag{251.4}$$

where

R is the illumination intensity on an inclined plane of photovoltaic array,

t_c (deg $\mathrm{m^{-1}{\cdot}m^2}$) is the temperature coefficient of PV array, In reference conditions, I_{sc} is the short-circuit current,

V_{oc} is the open circuit voltage,

I_m, V_m are the maximum power point current and voltage.

When the PV array voltage is V, its corresponding point current I:

$$I = I_{\mathrm{sc}}(1 - C_1(e^{\frac{V}{C_2 V_{\mathrm{oc}}}} - 1)) \tag{251.5}$$

where

$$C_1 = (1 - I_m/I_{\mathrm{sc}})e^{-\frac{V_m}{C_2 V_{oc}}}\ C_2 = (V_m/V_{\mathrm{oc}} - 1)/\ln(1 - I_m/I_{\mathrm{SC}}) \tag{251.6}$$

Considering the changes in light intensity and temperature influence

$$I = I_{\mathrm{sc}}\left(1 - C_1\left(e^{\frac{V-DV}{C_2 V_{\mathrm{oc}}}} - 1\right)\right) + DI \tag{251.7}$$

where

$$DI = A \cdot R/R_{\mathrm{ref}} \cdot DT + (R/R_{\mathrm{ref}} - 1) \cdot I_{\mathrm{sc}} \tag{251.8}$$

$$DV = -B \cdot DT - R_s \cdot DI \tag{251.9}$$

$$DT = T_c - T_{\mathrm{ref}} \tag{251.10}$$

R_{ref}, T_{ref} are temperature reference of illumination intensity and photovoltaic array, generally take for 1 kW/m^2, 25 $°C$;

A is temperature coefficient of current changes, in reference to illumination intensity;

B is temperature coefficient of voltage changes, in reference to illumination intensity;

R_S is photovoltaic module series resistance, by the following type decided

$$R_s = \frac{N}{N_p} R_{s,\,\mathrm{ref}} = \frac{N}{N_p}\left(A_{\mathrm{ref}} \ln\left(1 - \frac{I_{\mathrm{m,\,ref}}}{I_{\mathrm{sc,\,ref}}}\right) - V_{m,\,\mathrm{ref}} + V_{\mathrm{oc,\,ref}}\right)/I_{m,\,\mathrm{ref}} \tag{251.11}$$

$$A_{\mathrm{ref}} = \frac{T_{c,\,\mathrm{ref}}\eta_{V_{\mathrm{OC}}} - V_{\mathrm{oc,ref}} + \varepsilon N}{\frac{\eta_{I_{\mathrm{SC}}} T_{c,\,\mathrm{ref}}}{I_{m,\,\mathrm{ref}}} - 3} \tag{251.12}$$

where

ε is material with energy, $\varepsilon = 1.12\ V$ (silicon);

$I_{m,\,\mathrm{ref}}$, $V_{m,\,\mathrm{ref}}$ are short-circuit current and open-circuit voltage of photovoltaic array, in reference to the conditions

$\eta_{V_{\mathrm{OC}}}$, $\eta_{I_{\mathrm{sc}}}$ are open-circuit voltage and short-circuit temperature coefficient of photovoltaic array, in reference to the conditions

N_s is the united serial number in each module of photovoltaic array

N is the serial number of photovoltaic array module

N_p is the parallel number of photovoltaic array module

$T_{\mathrm{c,\,ref}}$ is the temperature of volt battery, in reference to the conditions, it is generally set to 25°C.

251.4.2 MATLAB Simulation Model for Solar Cells

In the MATLAB environment, based on the mathematical model and using the SIMULINK tool can establish a solar cells simulation model. We have an example of solar cells *NT-R5E3E* and the parameters are: the open-circuit voltage is 44.4 V, the maximum power point voltage is 35.4 V, the short-circuit current is 5.55 A, the

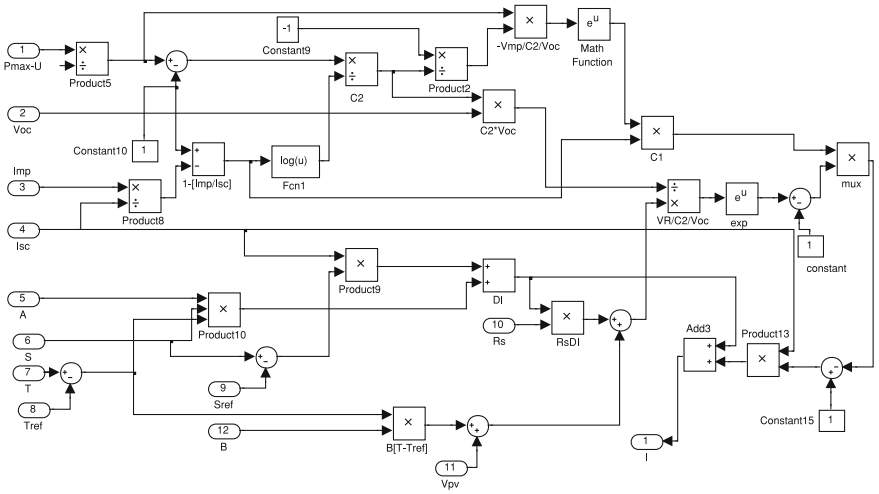


Fig. 251.5 The internal structure of photovoltaic modules

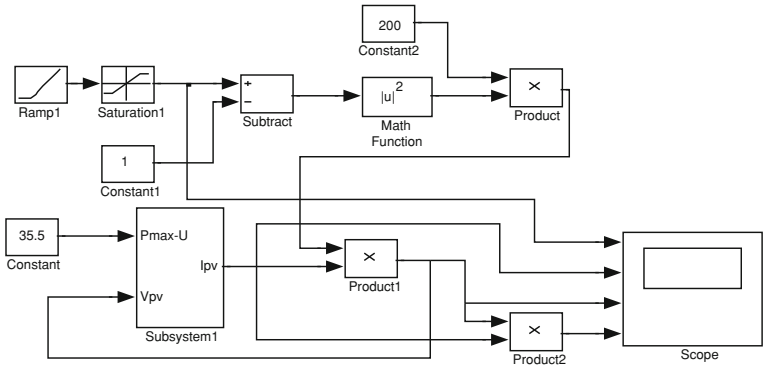


Fig. 251.6 Principles of simulation model of solar MPPT

maximum power point current is 4.95 *A*, the maximum power is 175 *W*. This solar cells model is completely used for the simulation studying of independent and on-grid system [6, 7]. The internal structure of photovoltaic modules is shown in the Fig. 251.5.

251.5 Impedance Matching for MPPT Simulation

In the MATLAB environment, according to the established simulation model of solar cells and the impedance transformation relationship of BOOST circuit, we can set up a simulation model in the SIMULINK, as shown in Fig. 251.6, and the

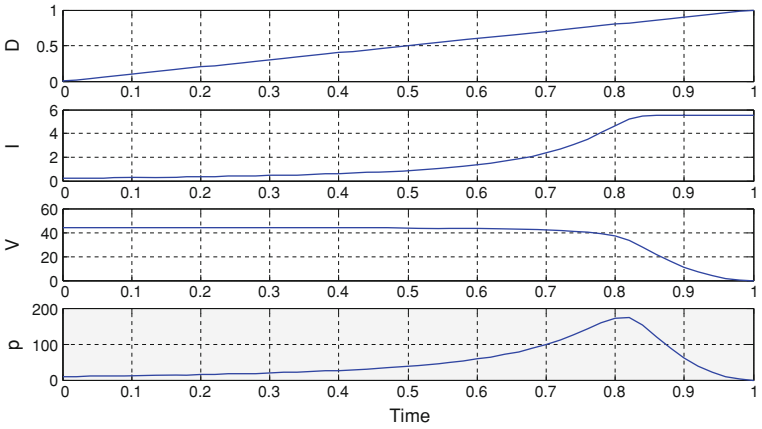


Fig. 251.7 Simulation waveforms for the 200 Ω load

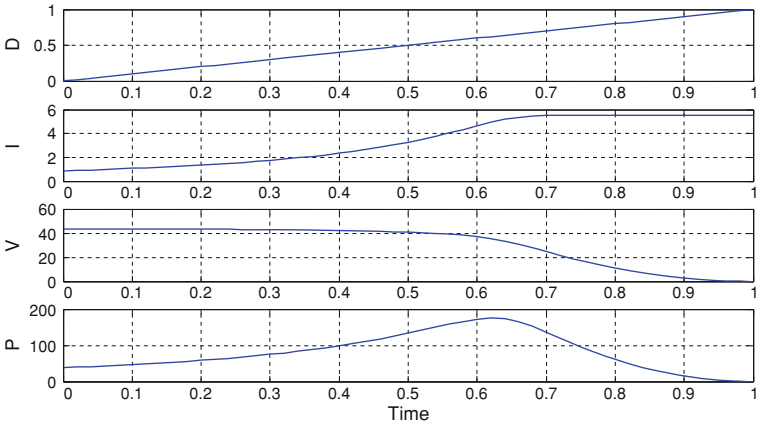


Fig. 251.8 Simulation waveforms for the 80 Ω load

illumination intensity is 1000 W/m^2, the temperature is 25°C and the loads are for 200 and 50 Ω. Adjusting the duty cycle in the range 0–1 and observing the output characteristics, we can get two sets of simulation waveforms shown in Figs. 251.7 and 251.8.

From Figs. 251.7 and 251.8, the load 200 and 50 Ω respectively, along with the D changes, there will be one of the largest solar power, the duty cycle D corresponds to 0.81 and 0.625, respectively, as the type 3 we can obtain the effective resistances were:

$$R_D = R_L(1 - D)^2 = 200(1 - 0.81)^2 = 7.22\ \Omega \tag{251.13}$$

When the load is 200 Ω.

$$R_D = R_L(1-D)^2 = 50(1-0.625)^2 = 7.03\ \Omega \tag{251.14}$$

When the load is 50 Ω.

From the above analysis, the solar panels with different loads by adjusting the duty cycle, the use of BOOST circuit transformation, impedance matching, can make solar cells work at maximum power point.

In the setting work environment, the simulation of the solar panels used to set the maximum power of 175 *W*, corresponding to the maximum power point voltage of 35.4 *V*, then the best match impedance is:

$$R_M = \frac{{U_{\max}}^2}{P_{\max}} = 7.16\,\Omega \tag{251.15}$$

R_M is the best theoretical value matched load under the current working environment, and it is almost equal to the previous calculated equivalent load value of 7.22−7.03 Ω, which means the calculated result is consistent with simulation results and laid the foundation in order to complete for the simulation of MPPT later.

In addition, we need to note that the load is not any value in the simulation process, and BOOST circuit impedance conversion will only make the resistance go small. When the load is less than 7.16 Ω, BOOST through impedance transformation circuit can not achieve maximum power tracking.

251.6 Conclusion

Based on the principle of impedance matching for controlling MPPT of photovoltaic power generation system, this paper has been systematically studied. The basic working principles of BOOST circuit and the most power tracing realizing methods and processes are first analyzed, modeling simulation and analyzing solar cells and the maximum power point tracking, Simulation results show that the impedance matching circuit based on BOOST can realize the maximum power tracking.

Acknowledgments The authors would like to express their gratitude to the managers and teachers of Department of Electrical Engineering and Automation for kindly supporting this research. The research are sponsored by Henan Province key scientific and technological project (082102240008), Doctor Fund of Henan Polytechnic University (648193) and Young Backbone Teachers Fund of Henan Polytechnic University (649093).

References

1. Wu Q, LI X (2010) Analysis of characteristics of solar cells based on Matlab/Simulink[J]. J North China Inst Water Conserv Hydroelectr Power 31:90–93
2. Wang X (2008) The research of stand-alone pv power system and its MPPT[D]. Nanjing University of Aeronautics, Nanjing
3. Chen JM, Chen L (2009) Study of photovoltaic maximum power point tracking techniques[J]. Sci Technol Eng 17:4940–4945
4. Bai LP, Bai S (2009) Design of solar controller with maximum power tracking [J]. Electr Technol 8:104–108
5. AbdulHadi M, Al-Ibrahim AM, Virk GS (2004) Neuro-fuzzy-based solar cell model[J]. Ener Convers 19:619–624
6. Chen JHS (2001) The design and manufacture of maximum power point of solar cells[D]. Tamkang University, Taiwan
7. Mao MQ, Yu SJ, Su JH (2005) MATLAB general simulation model with the photovoltaic array of MPPT function [J]. J Syst Simul 17:1248–1251

Chapter 252
Control and Implementation of Inverters Parallel Operation in Grid-Connected Photovoltaic

Jie Liu, Yongsheng Zhang and Haizhu Yang

Abstract Traditional inverter parallel control schemes have connection among each module. With increase in the number of modules and the increases in the distance among modules, the signals among each module are vulnerable to suffer interference. According to the principle of parallel operation of inverters, this paper analyzes several parallel inverter control schemes, and compares advantages and disadvantages of several common PQ drop methods. Further studies of the control technology of parallel inverters without iconnection besides this paper also introduces a more suitable method for the grid-connected PV systems. This method takes into account the grid-connected line resistance and the impact of harmonics. Meanwhile, it also improves the dynamic response capability of the system. The simulation proves the feasibility of this method.

Keywords PV system · Grid-connected · Parallel inverters · P-Q method

252.1 Introduction

The grid-connected PV system is one of the most hot development direction in PV power system. With the development of society and the demand, there are more and more load equipments that require bigger power capacity, single module

J. Liu (✉)
Department of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo, China
e-mail: liujie@hpu.edu.cn

Y. Zhang · H. Yang
Department of Electrical Engineering and Automation,
Henan Polytechnic University, Jiaozuo, China

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_252,

inverter scalable and reliability get limited, Therefore, to design multi-modules inverters parallel is seeming particularly important before grid power [1], Multiple inverters parallel operation can expand the system capacity, and improve system efficiency and reliability, the common methods of inverter parallel operation are master-slave type and the central limit type, but all have certain restriction. So this paper introduces a kind of inverters parallel operation method without interconnect based on the grid-connected PV system, Through the implicit relationship of modules to realize balanced current, using advanced digital controller, this can not only reduce the size and weight, but also improve analog controller unstable shortcomings [2].

252.2 The Strategy Research of Inverter Parallel Control

252.2.1 The Comparison of Several Inverter Parallel Control Methods

According to having line or not can be divided interconnect control and without interconnect control, the interconnect control mainly has master-slave control, centralized control, scattered logic control, etc. they each have their strengths and weaknesses, the centralized control belongs to the early control mode, it uses a centralized controller to test current deviation, so as to as output voltage compensation quantity, and then eliminate the current imbalance, this method is simple, but the system reliability is low. In the master-slave control every parallel inverter has own controller, through some logical to select one of them as master, the rest as slaves, this method can make capacity expansion be flexible, but reliability low, and easy to introduce noise interference. Each module in scattered logic control has not master-slave, so can reach large redundancy, reliability is high too, this is a kind of fully research method present. Without interconnect control is proposed in recent years, it is mainly to adopt some of the volume adjusting module itself to realize the load power allocation, because only has connection to AC bus, every module has not interconnect, so this kind of control is more complicated, but it completely realize redundancy, each module is independent and the system reliability is high too, so it has good research prospects and application prospect.

252.2.2 The Traditional Strategy of Without Interconnect Inverter Control

First we make a simple analysis to inverter parallel equivalent model, with two grid inverter parallel modules to establish the system's equivalent model as shown

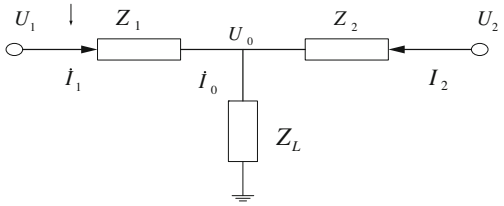


Fig. 252.1 The equivalent model of two inverters parallel

in Fig. 252.1. The U_0 is load voltage, the U_1 and U_2 are separately for first inverter and second inverter's voltage, the I_1 and I_2 are separately for first inverter and second inverter's current. According to the KVL and KCL, we know:

The inverter's output current is:

$$I_1 = \left[\frac{U_1 \sin\phi_1 - j(\cos\phi_1 - U_0)}{X_1}\right]^* \tag{252.1}$$

The inverter's complex power is:

$$S_1 = P_1 + jQ_1 = U_0 I_1 \tag{252.2}$$

Ignoring the line resistance R can be received the active power and reactive power's expressions as shown below:

$$P_1 = \frac{U_1 U_0}{X_1}\sin\phi_1 \tag{252.3}$$

$$Q_1 = \frac{U_0(U_1\cos\phi_1 - U_0)}{X_1} \tag{252.4}$$

Normally the ϕ_1 and ϕ_2 is smaller, according to the sin and cosine laws, can be received: $\sin\phi_1 \approx \phi_1$, $\cos\phi_1 \approx 1$, so can be simplify as:

$$P_1 \approx \frac{U_1 U_0}{X_1}\phi_1 \tag{252.5}$$

$$Q_1 \approx \frac{U_0(U_1 - U_0)}{X_1} \tag{252.6}$$

By the above expressions can we know that active power mainly has relation to the output voltage phase difference. But the reactive power mainly has relation to the output voltage amplitude difference, so can use such a droop characteristics to implement parallel control. This is the basic concept and idea of traditional PQ method, that without interconnect parallel control method is formed by referencing the synchronous generator and voltage drop theory. The traditional PQ method which does not consider load resistor r and harmonic's influence is an ideal theory. Its equation is as below:

$$P = h_1 \times \phi \tag{252.7}$$

$$Q = h_2 \times U_V - Q_0 \tag{252.8}$$

The P and Q are separate for the module's active power and reactive power, the U_V is the mean of the peak value of output voltage, the ϕ is the mean of the voltage phase difference.

The control equations are:

$$f_i = f_0 - m \times P \tag{252.9}$$

$$U_V = U_0 - nQ \tag{252.10}$$

In addition there is another method which is PQD method. It takes into account the influence of output voltage harmonic mean that adjusts output current by reactive power Q, active power P, and harmonic power D, at last to realize balanced current. This method, adjusts the voltage loop gain by testing the harmonic power D, finally completing the harmonic power's average distribution, but using this method can increase the voltage harmonic component. Furthermore, can also make serious implications for grid, so it is not suitable for grid occasion [3].

252.2.3 The Improved Strategy of Without Interconnect Inverter Control

In the actual grid inverter parallel operation, the circuit impedance characteristics changes variably. The circuit resistance cannot be ignored in long distance transmission, so considering that output reactive power Q and active power P jointly effect the output voltage phase and amplitude [4]. that is:

$$\begin{aligned} f_1 &= f_0 - m_1 P_1 + n_1 Q_1 \\ U_1 &= U_0 - m_2 P_1 - n_2 Q_1 \end{aligned} \tag{252.11}$$

The basic relations among these coefficients are m_2 far less than m_1, the line impedance characteristics decides the relationship of n_1 and n_2.

If we still adjust by the traditional PQ method, it may cause the system's positive feedback, so we must consider the effect of the line resistance, through the correction of traditional PQ method, put forward the improved PQ method, taking the first inverter as an example to illustrate, the output active power P and reactive power Q of inverter became:

$$P_1 = \left[\frac{U_0(U_1 R_1 \cos\phi_1 + U_1 X_1 \sin\phi_1 - U_0 R_1)}{R_1^2 + X_1^2}\right] \tag{252.12}$$

$$Q_1 = \left[\frac{U_0(-U_1 X_1 \cos\phi_1 + U_1 R_1 \sin\phi_1 + U_0 X_1)}{R_1^2 + X_1^2}\right] \tag{252.13}$$

Set impedance angles as $\alpha, Z_1 = R_1^2 + X_1^2, \sin\alpha_1 = \frac{X_1}{Z_1}, \cos\alpha_1 = \frac{R_1}{Z_1}$

By type (252.12) and (252.13) simultaneous can be received:

$$\sin\phi_1 = \frac{Z_1(P_1\sin\alpha_1 + Q_1\cos\alpha_1)}{U_1U_0} \tag{252.14}$$

$$U_1\cos\phi_1 - U_0 = \frac{Z_1(P_1\cos\alpha_1 - Q_1\sin\alpha_1)}{U_0} \tag{252.15}$$

Then we use a matrix A to correct P and Q,

$$\begin{bmatrix} P' \\ Q' \end{bmatrix} = A\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} \sin\alpha & \cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} \frac{X}{Z} & \frac{R}{Z} \\ \frac{R}{Z} & -\frac{X}{Z} \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix} \tag{252.16}$$

By type (252.14), (252.15) and (252.16) simultaneous can be received:

$$\sin\phi_1 = \frac{Z_1P_1'}{U_1U_0} \tag{252.17}$$

$$U_1\cos\phi_1 - U_0 = \frac{Z_1Q_1'}{U_0} \tag{252.18}$$

Due to the PQ is corrected, so the corresponding control equations are also changed for:

$$f_1 - f_0 = -k_pP_1' \tag{252.19}$$

$$U_1 - U_0 = -\mathrm{k}_qQ_1' \tag{252.20}$$

When in the inverter modules operation process if appeared to join module, exit module, load switch etc. interference, the system needs have good dynamic response performance, otherwise it might appear concussion phenomenon, so consider to add differential link in algorithm. Using this algorithm is not only considering the system steady performance, but also considering the system dynamic response ability. It also can make the output frequency and voltage more precise and effective [5].

$$\begin{aligned} f_1 &= f_0 - k_pP' - s_1\frac{\mathrm{dP}_1'}{dt} + l_1\frac{\mathrm{dQ}_1'}{dt} \\ U_1 &= U_0 - k_qQ' - s_2\frac{\mathrm{dQ}_1'}{dt} + l_2\frac{\mathrm{dP}_1'}{dt} \end{aligned} \tag{252.21}$$

The basic relations among these coefficients are l_2 far less than s_1, the line impedance characteristics decides the relationship of l_1 and s_2.

252.2.4 Determine the Droop Parameters

At first determine k_p and k_q, second determine four differentiation parameters, because k_p and k_q play a main role in system successful parallel, there are mang

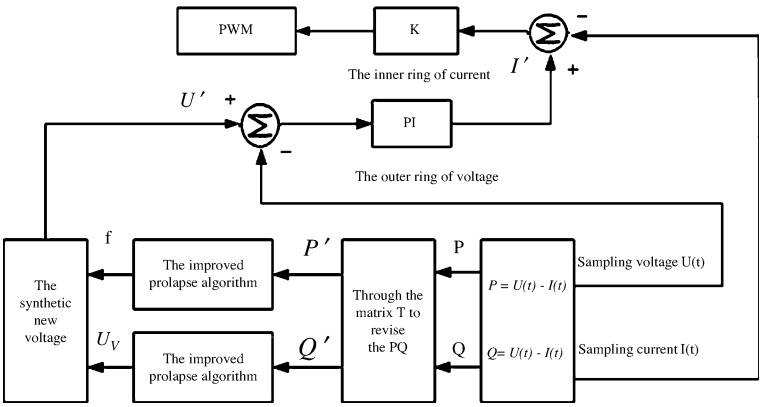


Fig. 252.2 The control diagram of parallel inverters in grid-connected PV

couplings among the differentiation parameters, so it is not easy to derive by formula, only by on-line estimation validation through the simulation. Similarly, we can receive the second's expression:

$$f_2 - f_0 = -k_E \mathrm{P}'_2 \tag{252.22}$$

$$U_2 - U_0 = -\mathrm{k_q} \mathrm{Q}'_2 \tag{252.23}$$

By two type simultaneous to solve equations can be received:

$$-\mathrm{k_p} = \frac{f_1 - f_2}{P'_1 - P'_2} \tag{252.24}$$

$$-\mathrm{k_q} = \frac{U_1 - U_2}{Q'_1 - Q'_2} \tag{252.25}$$

For the rest of four differentiation parameterss_1, s_2, l_1, and l_2, it is difficult to calculate them by theoretical analysis method, so in this paper using on-line estimation validation through the simulation to find appropriate parameters, which can make the system have better dynamic performance.

252.3 To Realize Inverter Parallel System by Simulation

According to the actual demand of improved PQ method introduced in this paper, establish grid-connected PV system by two sets of inverter, the main circuit adopts full bridge structure, using the double closed loop control mode which is constituted by voltage electric ring and current outer ring, the modulation mode is SPWM, the improved system diagram is as Fig. 252.2 [6]:

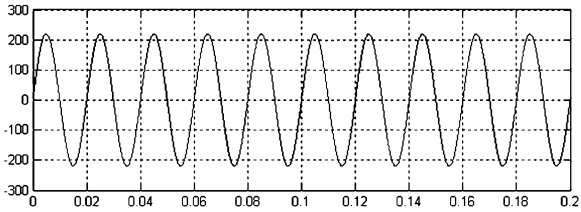


Fig. 252.3 The first inverter's output voltage waveform after parallel operation

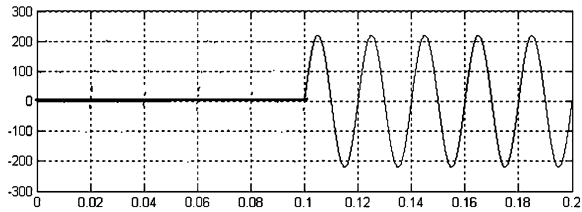


Fig. 252.4 The second inverter's output voltage waveform after first inverter run for several times

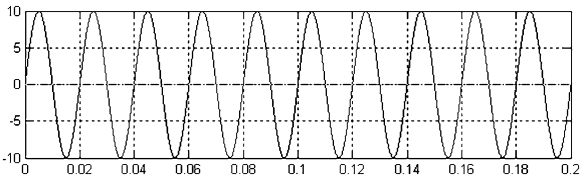


Fig. 252.5 The first inverter's output current waveform after parallel operation

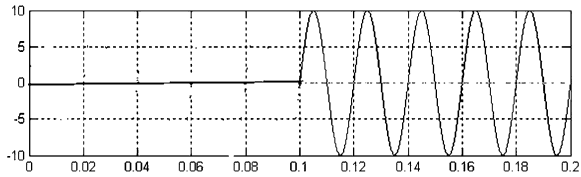


Fig. 252.6 The second inverter's output current waveform after first inverter run for several times

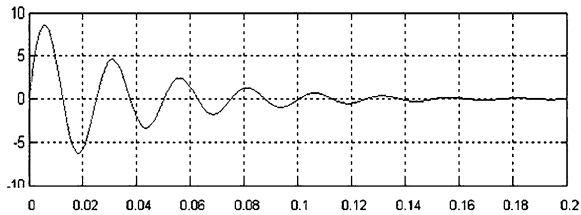


Fig. 252.7 The cycle waveform of system

By Figs. 252.3 and 252.4 we know: when the first inverter runs, the output voltage and current all can reach the desired effect, as the Figs. 252.5 and 252.6 we can know: after the first inverter run for several times. The second inverter merge into the system, this time, the output voltage and current have some distortion, but soon be more stable, as the Fig. 252.7 we can know: the cycle waveform of system is quickly attenuation after the second inverter merge into the system, at last it

soon becomes a straight line, the balanced current effect is satisfied, through the analysis of system simulation results can we know, after using the improved PQ method, the systems output voltage and current become more fit for actual demand, meanwhile, the dynamic response of the system is also very good, the balanced current effect is obvious, in a word, it can realize inverter parallel control of the gird connected.

References

1. Wu CF, Zhang JX et al (2009) Solar photovoltaic grid connected power and lighting system. Science press, Beijing
2. Chen DL (2005) DC–AC inverter technology and Its application . China Machine Press, Beijing
3. Wang B (2009) Research on the technology of parallel inverter based on grid-connected system. HeFei University of Technology, HeFei
4. Li L (2006) research on the technology of parallel inverter based on without interconnect. Wuhan University of Technology
5. SJ LU, Luo YH (2010) The control technology of parallel inverters without control interconnection. Appl Sci Technol 37:53–55
6. Lan X (2005) Analysis and realization of DC–AC inverter parallel system based on active and reactive circulating-current control. Trans China Electrotech Soc 20:8–12

Chapter 253
High Precise Self-Adaptive Digital Illuminometer-Based CPLD

Xiuwu Sui, Jing Lu, Xiaoguang Qi and Dapeng Li

Abstract The paper designs a kind of self-adaptive digital illuminometer of high accuracy. The meter consists of silicon photocell, current to voltage converter, auto gain amplifier with complex programmable logical device, analog value to digital value converter, liquid crystal display. The illuminometer has many advantages of wide measurement range, short response time, high photoelectric conversion efficiency, high resolution, and high portability, so it is very useful in many fields.

Keywords Silicon photocell · A/D conversion · Signal self-adapting · CPLD · LCD

253.1 Introduction

Illumination is significant to our lives, such as on the industry site, airplane track, car lamps and so on. If the illumination is insufficient or extraordinarily strong, it can not only bring vision weakness, but also incur potential safety hazards.

X. Sui (✉) · J. Lu · X. Qi · D. Li
Tianjin Key Laboratory of Advanced Mechantronic Equipment Technology, School of mechanical Engineering, Tianjin Polytechnic University, Tianjin, China
e-mail: allensui@163.com

J. Lu
e-mail: allenui@eyou.com

X. Qi
e-mail: qixiaoguang629@163.com

D. Li
e-mail: ldp513@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_253,

Table 253.1 The comparison of different cosine correctors

Angle shape	0°	10°	30°	50°	60°	70°	80°
Plate	0	−2.5%	−10%	−15%	–	−55%	–
Spherical crown	0	–	−1%	−3%	–	−20%	–
Spherical shell	0	0	0	0	0	0	–
Toric shape	0	–	±2%	–	±7%	–	±25%

The illuminometer is a kind of instrument to measure luminosity, widely used in many fields of civilian life, science research, war industry, traffic, medical care and building industry and so on [1], so it is highly imperative to develop a novel illuminometer with low cost, high accuracy, wide range, and self-adaptive capacity.

253.2 Principle of the Illuminometer

Illuminometer is one of the most common instruments in optics measurement, and is especially designed for test illumination. Illumination is the luminous flux per area expressed in Eq. (253.1), E is the illumination in lux, and φ is the luminous flux in lumen.

$$E = \frac{\mathrm{d}\Phi}{\mathrm{d}A} \tag{253.1}$$

The response of the detector surface to the light from different directions with the same intensity agrees with cosine law, that is to say the response of vertical direction is greatest, but when the angle of incidence equals 90°, the response is zero. When the angle of incidence is greater than 40°, the measuring result is not coincidental to the cosine law, so, in order to eliminate the error caused by the cosine law, a cosine corrector is fixed before the photoelectric detector; the corrector is made of opalescent glass or optical plastics. The correct results of different cosine correctors are listed in Table 253.1 [2].

253.3 Overall Design

The principle of illumination is shown in Fig. 253.1. When the photoelectric detector is lighted, the sensor surface of detector generates micro current, the current is changed to voltage by the I/V converter, then the voltage is converted from analog quantity to digital quantity, and then transferred to the liquid crystal display (LCD) to be displayed.

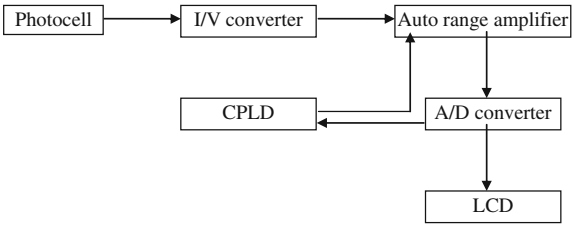


Fig. 253.1 The principle of the illumination

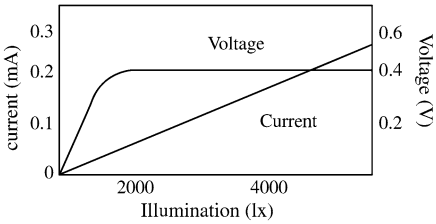


Fig. 253.2 The characteristic of the silicon cell

In order to be adaptive to the illumination measured, that is to say the meter can automatically change the range with the illumination value, the Complex Programmable Logic Device is used to change the amplification times, and then change the measurement range.

253.3.1 Photocell

The photoelectric detector is a silicon cell producing different voltage and current under different lights, the characteristic of the silicon cell is shown in Fig. 253.2.

The silicon cell is used as a current resource controlled by light, where the output current increases as the light increases [3].

253.3.2 I/V Converter

The amplifier is ICL7650, an integrated amplifier of high accuracy, auto-zero, chopped wave, low drift and high input resistant. It consists of two amplifiers, one used for amplification, and the other for drift compensation. The dynamic auto-zero eliminates the inherent misalignment and drift. For the weak signal, ICL7650 is a perfect amplifier, Fig. 253.3 shows the I/V converter.

The resistant is a precise resistance with low temperature coefficient, so the circuit ensures the stability of the current to voltage. A filter of resistance and capacitance is placed to limit the bandwidth, smooth the noise, and then to increase the ratio of signal to noise [4].

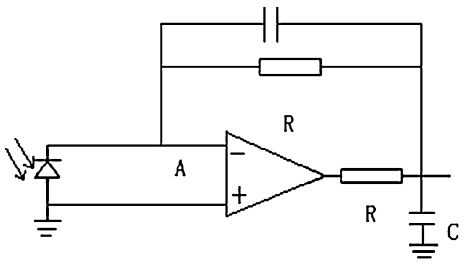


Fig. 253.3 I/V converter

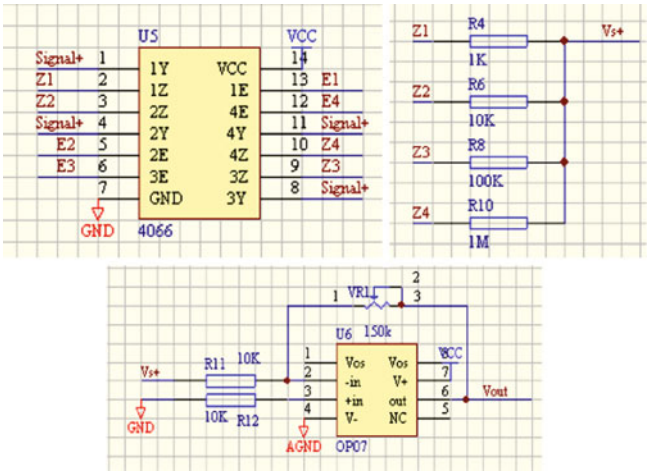


Fig. 253.4 Four channels amplifier with multi-path switch circuit

253.3.3 Auto Range Amplifier

Since the signal from the detector is very weak, the voltage is also very weak after I/V converter. It is difficult to reach high accuracy with the background noise, circuit noise, and components noise; the amplifier must have wide bandwidth, high stability, little gain error, and high ratio of signal to noise [5].

Because the measurement range is very wide, in order to improve the accuracy of the analogy to digit conversion, a four channels amplifier is used to change the measurement range for different illuminations, as shown in Fig. 253.4.

When the signal is small, the higher amplification times is used, and when the signal is larger, the lower amplification times is selected. Selection of the amplification times is realized by multi-path switch 74HC4066.

Table 253.2 CPLD and the amplifier of auto-range

Range no.	CPLD output	Feedback resistant	Amplification times	Illumination range
1	1000	R4 = 1 K	10	0–2 lx
2	0100	R6 = 10 K	100	2–20 lx
3	0010	R8 = 100 K	1000	20–200 lx
4	0001	R10 = 1 M	10000	200–2000 lx

253.3.4 A/D Converter

The A/D converter is ICL7129 from American Intersil Company. The chip is high in the ratio of performance to cost, four and a half bits converter accuracy, suitable for connection to the LCD, and is of high resolution, low power consumption, and easy to be connected with other chips. The chip can directly drive the LCD display, the max display value is ±19999, the input voltage range is 10 μV−2.0000 V, the resolution is 10 μV, equivalent to the 18 bits binary. ICL7129 is a double-integrator analogue to digit converter [6], which is immune to the 50 Hz interference and is able to adjust to zero itself.

253.3.5 Auto Range Switches Circuit with CPLD

The CPLD of EPM7064S is used to control the multi-path switch 74HC4066, the output A, B, C, D of CPLD are connected to the 74HC4066's E1, E2, E3, E4 to control the signal pass to the Z1, Z2, Z3, Z4, separately, the Z1, Z2, Z3, Z4 are connected to the amplifier by the resistance of R4, R6, R8, R10. For example, when E1 is high, 1Y is connected to 1Z, then the signal passes to the amplifier by the R4, the amplification times is 10. The relation between the output of EPM7064S and the range is listed in Table 253.2.

The software of Quartus 6.0 is used to program the CPLD, the circuit of the auto-range program is shown in Fig. 253.5.

In Fig. 253.5, CP is the clock pulse, OR is over range bit connected with A/D converter, UR is deficient range bit connected with A/D converter separately, and A, B, C, D is the output of the CPLD. The hardware simulation results of auto range are shown in Figs. 253.6 and 253.7.

253.3.6 LCD Display

The LCD has advantages of low voltage, low power consumption, small volume, low weighty and is ultra thin.

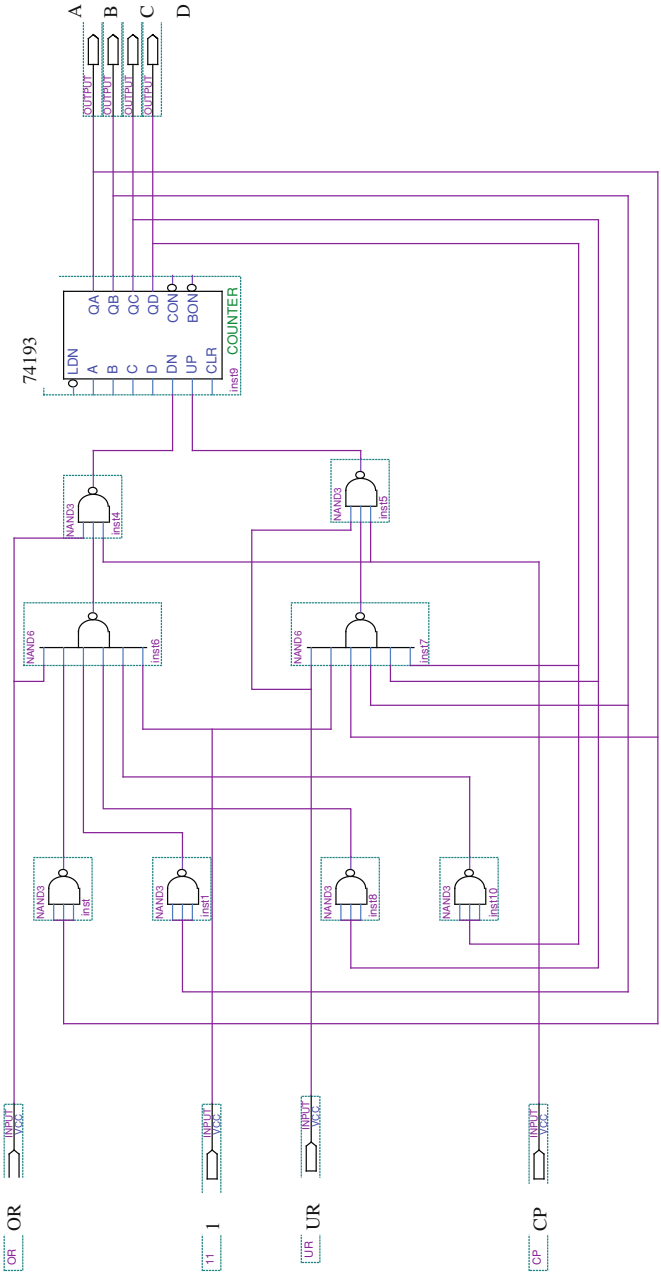


Fig. 253.5 The circuit of the auto range program

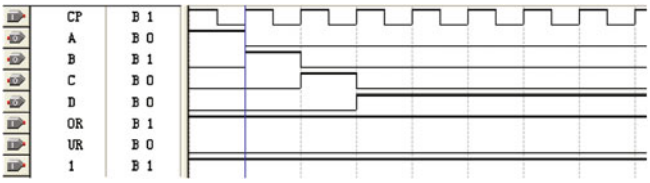


Fig. 253.6 The hardware simulation result of over range

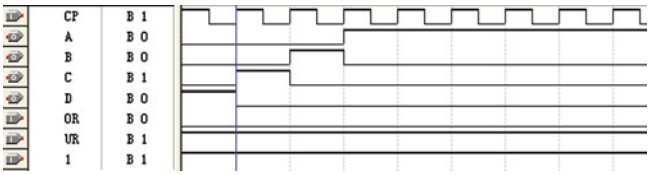


Fig. 253.7 The hardware simulation result of deficient range

Table 253.3 The comparison of the different ranges

Range no.	Range (lx)	Error of fixed range (lx)	Error of auto range (lx)
1	0–2	0.1	0.0001
2	2–20	0.1	0.001
3	20–200	0.1	0.01
4	200–2000	0.1	0.1

253.4 Experiments and Conclusion

This paper carries the experiments by the self-adaptive digital illuminometer. The results show that the illumination can change the range automatically, compared with the traditional illumination of three and a half bits the accuracy of the A/D converter improves an order of magnitude, and compared with illumination of the four and half bits with fixed measurement range, the accuracy of the A/D converter is listed in Table 253.3.

Table 253.3 shows that by the auto range selection, the A/D converter accuracy improves greatly, especially for the small signal. The illumination is high in resolution and accuracy, wide in measurement range, portable type, so it is destined to be used widely.

References

1. Wang QY (2010) The course of comprehensive experiments and design in photoelectric information. Public House of Electrical Industry, Beijing
2. Hou WH (2007) Design of the high accurate illuminometer. Dalian University of Technology, Dalian
3. Xu Y (2002) The illuminometer characteristic and the valuation of the measuring error. Mod Metrol Meas (3):34–36
4. Bai ZS, Bai ZW (2006) The design of a simple illumination measurement circuit. Mod Electric Technol (11):91–92
5. Du MF, Jiang ZJ (2002) The study on the photovoltaic effect of the nonlinear region of the photocell PN junction. Trans Shanghai Technol Univ 24(1):65–67
6. Jiang WB, He LW (2007) The auto-range interface circuit of the common double integrator A/D converter. Instrum Technol 6:59–61

Chapter 254
Centralized Controller and Communication Protocol Design of Remote Meter Reading System

Xiangqun Zhang and Zhongbiao Zhao

Abstract In view of the following problems: the backward technology in remote meter reading, the poor interoperability of communication protocol and the high cost for network operation, the author has proposed a two-layer network construction based on GSM and ZIGBEE and designed the communication protocol of remote meter reading system with consideration of the data characteristics of meter reading system. And then, to ensure the reliability of communication links, the serial communication program is designed by calling, answering and then communicating after repeated experiments. In the end, the designed system is tested under experimental conditions, and the results show that the system had many advantages, such as high reliability, low cost, simple networking, high reliability for communication data and moderate speed of meter reading, and satisfied the design requirements basically.

Keywords Centralized meter reader · Remote meter reading · Handshake · Administration center computer

X. Zhang (✉)
College of Computer Science and Technology,
Xuchang University, Xuchang 461000, China
e-mail: dzzxq18@163.com

Z. Zhao
College of Electrical and Information Engineer,
Xuchang University, Xuchang 461000, China
e-mail: biaogreat@163.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_254,

254.1 Introduction

At present, the method which meter reading personnel read the meter to collect data and charge fees from door to door is widely used in China. Among electric power, gas, water and heat companies, each one reads his own way basically, which not only causes enormous waste of manpower and financial resources, but also brings great difficulties to the construction, analysis and planning of urban pipe network [1]. Although some intelligent technologies in remote meter reading have been applied at present, the following defects [2] are still commonly found: (1) Low technology, decentralized meter reading. Every kind of company sets up its own network, and the interoperability among the networks is very poor. (2) Different networking mode. Each company uses its own networking mode to read meters, such as, the power carrier wave meter reading, the infrared meter reading and RS485 meter reading. (3) Nonstandard communication protocol, poor information interoperability. The communication protocols between the companies are confused and cannot achieve the integration of four meters. (4) High cost. The author has carefully examined the defects of meter reading systems above and put forward a new network structure, combining four (electricity, water, gas, heat) into one. This networking mode has many advantages, such as simple structure, high reliability, low cost [3]. For the particularity of meter reading, the author gives the design plan of centralized meter reader and the implementation method of communication protocol. In addition, this system also has the following functions: water, electricity, gas and heat will be cut off automatically without fees and they can be restored automatically if the fees are renewed. The host computer can send the meter reading instructions from time to time to realize a real-time enquiry of the user's information, with this method, the operating status of the network can be checked and the bad behaviors such as stealing electricity, water, gas and heat can be eliminated.

254.2 Integral Structure of Meter Reading System

The double-layer network (GSM and ZIGBEE) structure is applied in this system [4]. The upper layer uses GSM network, which has been operated in our country for many years with the following characteristics such as mature technology, stable communication and high coverage [5], especially its short message service (SMS) has the two advantages: low cost and high timeliness. There will be the following features: high timeliness, wide network coverage and low cost if SMS data transfer based on GSM network and centralized meter reader is used. The lower layer applies ZIGBEE ad hoc network technology, which effectively reduces the cost of th system's operation and maintenance. The concrete structure is shown in Fig. 254.1.

Administration center computer [6] runs the meter reading software to send the meter reading instruction and other commands to centralized meter reader through GSM network, other commands include remote meter reading batch processing command, electricity, gas and heat will be cut off automatically without fees and

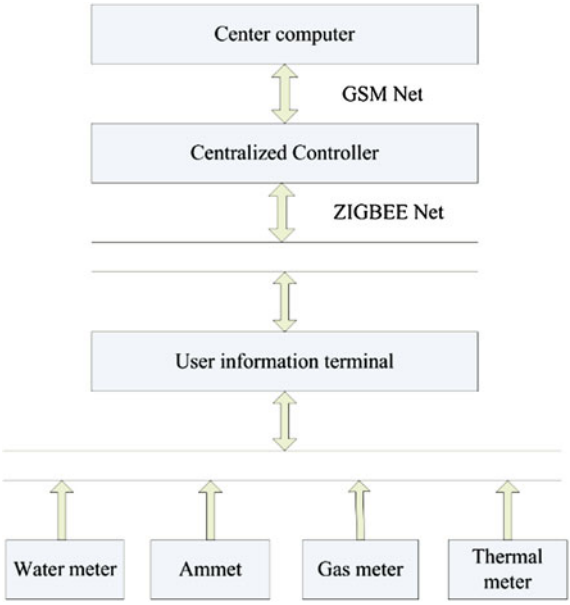


Fig. 254.1 Integral structure of meter reading system

they can be restored automatically if the fees are renewed. Besides, administration center computer can also send meter reading instructions from time to time to check the users' meter information to prevent network fault and eliminate the bad behaviors such as stealing electricity, water, gas and heat. TC35I of Siemens AG is adopted in GSM communication module with the following advantages: simple interface (it is connected to Single Chip Microcomputer through 232 port), high reliability and low cost.

Centralized meter reader connects to user information acquisition system through ZIGBEE wireless ad hoc network technology. The wireless network applies nRF401, a wireless communication chip from Nordic company in Norway with low price and easy network, and its communication frequency does not need to apply and the communication distance can reach 1,500 m in an open area, therefore, 2–3 centralized meter readers can completely satisfy a community's need. Centralized meter reader sends meter reading instruction to user information acquisition system through ZIGBEE wireless network after receiving the command of batch processing meter reading, and user information acquisition system sends the user's information of four meters to centralized meter reader via ZIGBEE wireless network. Centralized meter reader first preserves the received information and then sends them to administration center computer through GSM network; if centralized meter reader receives the user's arrearage information from administration center computer, it will sends the arrearage instruction to user information acquisition system, which can analyze the received instruction and takes appropriate action according to the corresponding instruction.

The main functions of user information acquisition system are as follows: (1) In charge of recording users' information of electricity, water, gas and heat

consumption; (2) Receiving the meter reading and the user's arrearage information from centralized meter reader and taking appropriate actions. For example, the user's four meters information will be sent to a centralized meter reader after receiving the instruction of meter reading; the corresponding arrearage treatment will be taken after receiving and analyzing the user's arrearage instruction; the normal usage condition will be automatically restored after receiving the instruction that the user's fees has already be renewed. Therefore, the reliability of user information acquisition system is more important. In order to improve the reliability, the system adopts an external MAX813 watchdog to monitor the program and uses AT24C04 to preserve four meters' information in time.

254.3 Communication Protocol Design

Remote meter reading system can ensure not only the completion of information collection but also the reliability of information. So the reliability and validity of communication process is the important quality index in the whole meter reading system, which is closely related to whether the system can operate reliably. In order to ensure establishing the reliable links in meter reading system, the serial communication program is designed by calling, answering and then communicating after repeated experiments, and its communication reliability has improved to a large extent. Communication rate: 9,600 b/s; communication method: asynchronous serial method, 1 start bit, 8 data bits, 1 stop bit and 1 parity bit.

254.3.1 Establishment of Communication Links

In order to prevent the error code of frame address by interference, redundancy code has been adopted, that is, centralized meter reader sends 3 sets of call commands in every communication. The call command is made up of 1-word address code and 1-word completion code, specific steps are as follows: if the response code from user information acquisition system can be received at 60 ms interval, the call command is sent continually, and this communication link can be proved to be established after receiving more than 2 responses; if the response from centralized meter reader is not received, which means that communication is not successful. And the next user information acquisition system's address must be called continually in order to prevent deadlock.

The communication reliability is the important quality index in the whole meter reading system, which is closely related to whether the system can operate reliably. This verification method, which takes 2 from 3 call commands, can effectively improve the reliability of communication, but it can also reduce the communication efficiency to a certain extent. This system has respectively designed the communication protocols between user information acquisition

Table 254.1 Call commands of centralized meter reader

Call command	Address code	End code
Case	0000H	E0E0H

Table 254.2 Response commands of user information acquisition system

Response command	Address
Case	0000H

system and centralized meter reader in order to make up for the loss. And the data in short-frame format has improved the communication efficiency effectively. The principle is shown in Table 254.1.

User information acquisition system adopts the receiving method of serial port interrupt. If a set call command from centralized meter reader is received, the judgment must be made first. The majority vote method, which takes 2 from 3 call commands, is applied in the received addresses to confirm whether the address of this certain information acquisition system is called, and if it is called, the instruction will be sent to centralized meter reader and the next operation instruction from centralized meter reader is waited. The principle is shown in Table 254.2.

254.3.2 Communication Frame Format of Centralized Meter Reader

Centralized meter reader adopts the majority vote method to the received address after receiving the response command from user information acquisition system to confirm the right acquisition system's address, and sends the corresponding operation command and data. The operation command consists of two same operation code bytes, operation data, check code and stop code. Communication frame format is shown in Table 254.3.

Address code: address number of user information acquisition system

Operation code: command word, which is sent to centralized meter reader by administration center computer (water outage information with arrearage; restoration of water supply information with the renewed water charges; information of the user's four meters, etc.); Check code: C0H; Stop code: E0E0H

254.3.3 Frame Format of Upload Data of User Information Acquisition System

After receiving the command from upper computer, user information acquisition system will respond to the command from centralized meter reader and send the

Table 254.3 Operation command of centralized meter reader

Operation command	Address code	Operation code	Check code	End code
Case	0000H	EFEFH	C0H	E0E0H

Table 254.4 Upload data of user information acquisition system

Upload data	Address code	Command code	Data	Check code	End code
Case	0000H	00H	1282H	C0H	E0E0H

corresponding data to centralized meter reader, and the data format is as follows. Centralized meter reader uses address code for judging the source of the data in order to conveniently store the corresponding database. The data frame format, which is sent to centralized meter reader by user information acquisition system, is as follows (Table 254.4).

Address code: address number of user information acquisition system; Command code: command of meter reading, command of user's arrearage and command of the user's paid charges; Check code: C0H; Stop code: E0E0H

254.4 Software Design

The communication between centralized meter reader and user information acquisition system adopts master–slave communication. Centralized meter reader sends all sorts of command information such as command of meter reading, information of the user's arrearage and information of the user's paid charges, to user information acquisition system. And the programming flow diagram of centralized meter reader is shown in Fig. 254.2.

After receiving the command from administration center computer, centralized meter reader sends handshake information to user information acquisition system, if the handshake information cannot be successfully sent for 3 times, the next user information acquisition system can be called; if it is successful, centralized meter reader can send operation command word, check code and stop code to user information acquisition system and wait for the information which is sent back by user information acquisition system. User information acquisition system sends its own address, the received command code and the corresponding operation data to centralized meter reader after successfully receiving the command, if the information is correct, centralized meter reader stores the user's information; if the information is wrong, the former process will be repeated.

User information acquisition system analyses and processed the received command, then sends its own address, the received command word, the corresponding data, check code and stop code to centralized meter reader. Centralized meter reader analyses the received address code, command code.

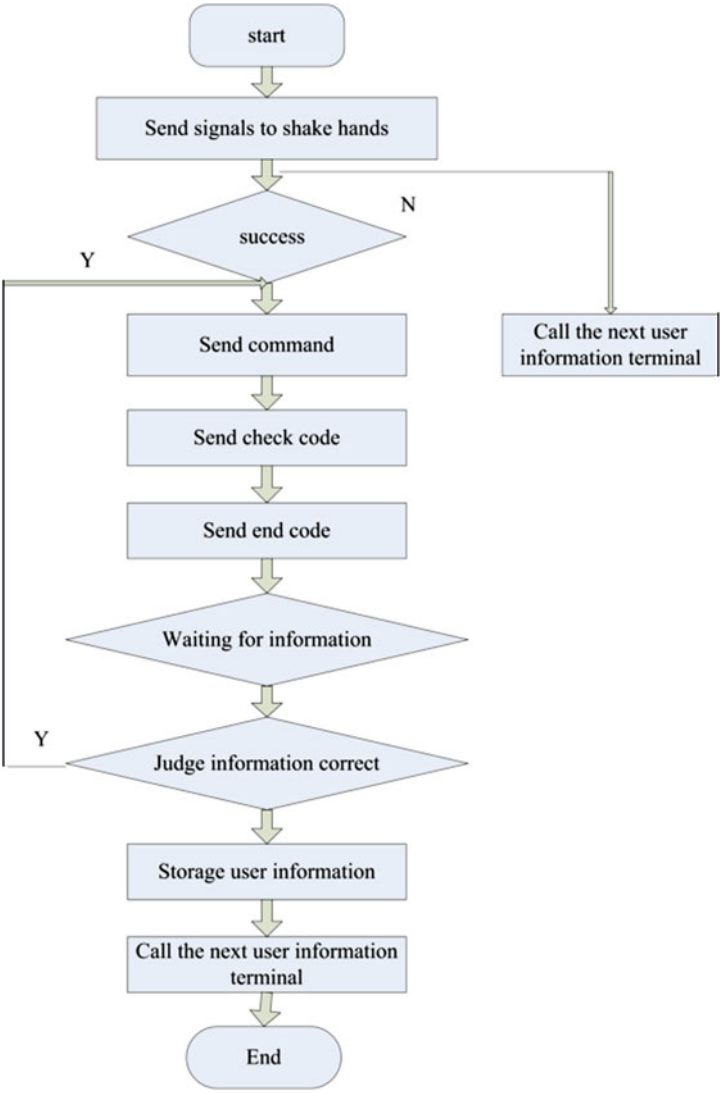


Fig. 254.2 Programming flow diagram of centralized meter reader

User information acquisition system analyses and processed the received command, then sends its own address, the received command word, the corresponding data, check code and stop code to centralized meter reader. Centralized meter reader analyses the received address code, command code, check code and stop code, if they are correct, the acquired data will be stored; if they are wrong, the former process will be repeated. And the flow diagram of software design of user information acquisition system is shown in Fig. 254.3.

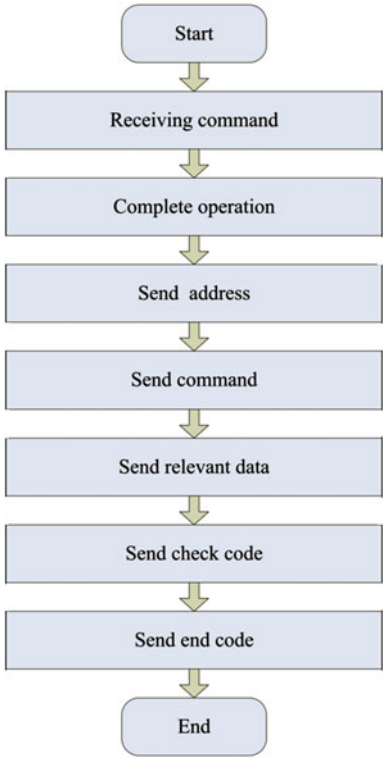


Fig. 254.3 Flow diagram of software design of user information acquisition system

254.5 Conclusion

The double-layer network structure improves the reliability of the system and reduces its operating and maintenance costs. his repeated handshake communication protocol appears to be complicated, but it largely improves the data reliability in the communication process. After long time practice, it shows that meter reading data are mostly short-frame structure and they can be quickly processed, and it can be completely tolerable to complete the communication in the case of 9,600 baud rate.

References

1. Zhao Z, Zhang Y, Gao R (2008) Research on the application of handshaking communication protocol in remote meter reading system. Relay 36(8):145–148
2. Li H (2010) Construction of wireless meter reading system. Sci Technol Innov Her 24:78–82
3. Zhang H, Zhang Y, Zhao H, Du X (2010) Design of wireless meter reading terminal based on GSM module. J North China Inst Water Conserv Hydroelectr Power 31(4):89–92

4. Bai T, Zhang C (2009) Based on the ARM and the ZigBee technology, wireless meter reading system research. Chin High-Tech Enterp (19):35–26
5. Sui H, Wang H, Wang H et al. (2007) Remote meter reading system based on GPRS used in substations.Telecommun Electr Power Sys 28(1):57–61
6. Zhu Y, Chen T, Guo S (2008) Applications of ZigBee technology in wireless meter reading. Telecommun Electric Power Sys 29(8):37–43

Chapter 255
Frames of Subspaces for Banach Spaces

Chunyan Li

Abstract Bessel sequences of subspaces, frames of subspaces and independent frame of subspaces for Banach spaces are introduced and discussed. Frames of subspaces is a generalization of frames of subspaces for Hilbert spaces. Some necessary and sufficient conditions for a Bessel sequence of subspaces to be a frame of subspaces for Banach spaces are given. It is proved that a Bessel sequence of subspaces becomes a frame of subspaces for Banach spaces if and only if its analysis operator is invertible. Lastly, Riesz bases of subspaces for Banach spaces are introduced. It is shown that an independent frame of subspaces is just a Riesz Basis for Banach spaces.

Keywords Frame of subspaces · Riesz basis of subspaces · Banach space

255.1 Introduction

The concept of frames were first introduced by Duffin and Schaeffer [1] in non-harmonic Fourier analysis. Let H be a Hilbert space and I be a countable index set. A family $\{f_i\}_{i\in I}$ in H is called a frame for H if the inequalities

$$A\|f\| \leq \sum_{i\in I} |<f, f_i>|^2 \leq B\|f\|, \quad \forall f \in H$$

C. Li (✉)
Department of Mathematics and Physics,
Chongqing University of Science and Technology,
Chongqing 401331, People's Republic of China
e-mail: lichunyan@126.com

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2_255,

hold for some positive constants A, B. In 1986, Daubechies, Grossman and Meyer [2] applied the theory of frames to wavelet and Gabor transform. After their work, the theory of frames are studied deeply and applied widely to signal processing, image processing, data compressing, sampling theory and so on (see [3–6]). Now, frames are used to mitigate the effect of losses in packet-based communication systems and hence to improve the robustness of data transmission and to design high-rate constellations with full diversity in multiple-antenna code design (see [7–9]).

New methods should be developed to handle these emerging applications of frames. An important problem is that how to build frames "locally" and then piece them together to become frames for the whole space. Many approaches to piecing together families of vectors to get a fame for the whole space have been done over the years. In [10, 11], Fornasier introduced a new approach. Fornasier uses subspaces which are quasi-orthogonal to construct local frames and piece them together to get global frames. In 2004, Casazza and Gitta. Kutyniok [12] introduced the concept of frames of subspaces for Hilbert spaces and formulated a general method for piecing together local frames to get global frames. Let $\{v_i\}_{i\in I}$ be a family of weights, i.e., $v_i > 0$ for all $i \in I$. A family $\{W_i\}_{i\in I}$ of closed subspaces of a Hilbert space H is a frame of subspaces for H with respect to $\{v_i\}_{i\in I}$, if there exist positive constants A, B such that

$$A\|f\| \leq \sum_{i\in I} v_i^2 \|\pi_{W_i} f\|^2 \leq B\|f\|, \quad \forall f \in H,$$

where for every $i \in I$, π_{W_i} is the projection of H onto W_i.

In this paper, we will introduce the concept of frames of subspaces for Banach spaces and discuss some important properties of frames of subspaces for Banach spaces. Frames of subspaces for Banach spaces is a generalization of the concept of frames of subspaces for Hilbert spaces. In addition, the concept of Riesz bases of subspaces for Banach spaces will be introduced. We will characterize an independent frame of subspaces and a Riesz basis of subspaces for Banach spaces by analysis operator.

255.2 Frames of Subspaces for Banach Spaces and Mail Results

255.2.1 Bessel Sequences of Subspaces for Banach Spaces

Throughout this paper, X will be a separable Banach space. Let $\{Y_i\}_{i\in I}$ be a sequence of Banach spaces. Put

$$\oplus_p Y_i = \left\{ \{y_i\}_{i\in I} : \forall i \in I, y_i \in Y_i, \left\|\{y_i\}_{i\in I}\right\|_p < +\infty \right\},$$

where

$$\left\|\{y_i\}_{i\in I}\right\|_p = \left(\sum_{i\in I} \|y_i\|^p\right)^{\frac{1}{p}}, 1 \le p < \infty$$
$$\left\|\{y_i\}_{i\in I}\right\|_\infty = \sup_{i\in I} \|y_i\|.$$

It is easy to check that $(\oplus_p Y_i, \|\cdot\|_p)$ becomes a Banach space over some number field F. From Exercise 4 in III.5 of Ref. [13], we know that the dual space $\left(\oplus_p Y_i\right)^*$ of $\oplus_p Y_i$ is isometrically isomorphic to $\oplus_q Y_i^*$, where $\frac{1}{p} + \frac{1}{q} = 1$. In fact, for every $\left\{y_i^*\right\}_{i\in I} \in \oplus_q Y_i^*$, define a functional $f_{\left\{y_i^*\right\}_{i\in I}}$ on $\oplus_p Y_i$ as follows:

$$f_{\left\{y_i^*\right\}_{i\in I}}(\{y_i\}_{i\in I}) = \sum_{i\in I} y_i^*(y_i), \forall \{y_i\}_{i\in I} \in \oplus_p Y_i.$$

Thus, the following mapping

$$\pi_p : \oplus_q Y_i^* \to \left(\oplus_p Y_i\right)^*, \pi_p(\left\{y_i^*\right\}_{i\in I}) = f_{\left\{y_i^*\right\}_{i\in I}}, \forall \left\{y_i^*\right\}_{i\in I} \in \oplus_q Y_i^*$$

is the isometric linear isomorphism from $\oplus_q Y_i^*$ onto $\left(\oplus_p Y_i\right)^*$.

Let W be a complemented subspace of the Banach space X. We denote by P_W the projection P in X such that $W = \operatorname{ran} P$.

Definition 1 Let $\{v_i\}_{i\in I}$ be a family of weights, i.e., $v_i > 0$, for all $i \in I$. A family $\{W_i\}_{i\in I}$ of complemented subspaces of a Banach space X is a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$, if there exist a constant $B > 0$ such that

$$\left(\sum_{i\in I} v_i^p \|P_{W_i} x\|^p\right)^{\frac{1}{p}} \le B\|x\|, \forall x \in X.$$

Let $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$. Then the analysis operator $T_W^v : X \to \oplus_p W_i$ is given by

$$T_W^v x = \{v_i P_{W_i} x\}_{i\in I}, \forall x \in X.$$

Theorem 1 *Let $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$ and B_W is a bound of $\{W_i\}_{i\in I}$. Then the operator*

$$U_W^v : \oplus_q W_i^* \to X^*, U_W^v(\left\{w_i^*\right\}_{i\in I}) = \sum_{i\in I} v_i P_{W_i}^* w_i^*$$

is well-defined and bounded. Usually, we call U_W^v the synthesis operator for $\{W_i\}_{i\in I}$.

Proof Assume that $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces with respect to $\{v_i\}_{i\in I}$. Let $F(I)$ be the set of all finite subsets of I. For any $\{w_i^*\}_{i\in I}\in \oplus_q W_i^*$, $\forall G \in F(I)$, we obtain that

$$\begin{aligned}\left\|\sum_{i\in G} v_i P_{W_i}^* w_i^*\right\| &= \sup_{\|x\|\le 1}\left|<x, \sum_{i\in G} v_i P_{W_i}^* w_i^*>\right| = \sup_{\|x\|\le 1}\left|\sum_{i\in G} <v_i P_{W_i} x, w_i^*>\right| \\ &\le \sup_{\|x\|\le 1}\sum_{i\in G}\left|<v_i P_{W_i} x, w_i^*>\right| \le \sup_{\|x\|\le 1}\sum_{i\in G}\|v_i P_{W_i} x\|\cdot\left\|w_i^*\right\| \\ &\le \sup_{\|x\|\le 1}\left(\sum_{i\in G}\|v_i P_{W_i} x\|^p\right)^{\frac{1}{p}}\cdot\left(\sum_{i\in G}\left\|w_i^*\right\|^q\right)^{\frac{1}{q}} \\ &\le \sup_{\|x\|\le 1}\left(\sum_{i\in G} v_i^p\|P_{W_i} x\|^p\right)^{\frac{1}{p}}\cdot\left(\sum_{i\in I}\left\|w_i^*\right\|^q\right)^{\frac{1}{q}} \le B_W\left\|\left\{y_i^*\right\}_{i\in I}\right\|\end{aligned}$$

This shows that $\sum_{i\in I} v_i P_{W_i}^* w_i^*$ converges in X^*. Thus, U_W^v is well-defined, linear and bounded.

In fact, for every $f \in \left(\oplus_p W_i\right)^*$, there exist a family $\left\{w_i^*\right\}_{i\in I}\in \oplus_q W_i^*$ such that $\pi_p^{-1} f = \left\{w_i^*\right\}_{i\in I}$. Thus, $\forall x \in X$, we have

$$\begin{aligned}<U_W^v \pi_p^{-1} f, x> &= <U_W^v(\left\{w_i^*\right\}_{i\in I}), x> = <\sum_{i\in I} v_i P_{W_i}^* w_i^*, x> \\ &= \sum_{i\in I} <v_i P_{W_i}^* w_i^*, x> = \sum_{i\in I} <w_i^*, v_i P_{W_i} x> \\ &= <\left\{w_i^*\right\}_{i\in I}, \{v_i P_{W_i} x\}_{i\in I}> = <f, T_W^v x> = <T_W^{v*} f, x>\end{aligned}$$

Hence, we have $U_W^v \pi_p^{-1} = T_W^{v*}$.

255.2.2 Frames of Subspaces and Riesz Bases of Subspaces for Banach Spaces

Definition 2 Let $\{v_i\}_{i\in I}$ be a family of weights, i.e., $v_i > 0$, $\forall i \in I$. A family $\{W_i\}_{i\in I}$ of complemented subspaces of a Banach space X is a frame of subspaces for X with respect to $\{v_i\}_{i\in I}$, if there exist constants $0 < A \le B < \infty$ such that

$$A\|x\| \le \left(\sum_{i\in I} v_i^p\|P_{W_i} x\|^p\right)^{\frac{1}{p}} \le B\|x\|, \forall x \in X$$

We call A and B the frame bounds for the frame of subspaces. If $A = B$, the family $\{W_i\}_{i\in I}$ is said to be tight. In particular, if $A = B = 1$, we say that $\{W_i\}_{i\in I}$ is a Parseval frame of subspaces.

Theorem 2 *Let $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$. Then the following statements are equivalent.*

(1) *$\{W_i\}_{i\in I}$ is a frame of subspaces for X with respect to $\{v_i\}_{i\in I}$.*
(2) *The analysis operator T_W^v of $\{W_i\}_{i\in I}$ is bounded below.*
(3) *The synthesis operator U_W^v of $\{W_i\}_{i\in I}$ is surjective.*

Proof (1) $\Rightarrow$ (2). If $\{W_i\}_{i\in I}$ is a frame of subspaces for X with respect to $\{v_i\}_{i\in I}$ and A, B are the bounds of $\{W_i\}_{i\in I}$, then

$$A\|x\| \le \left(\sum_{i\in I} v_i^p \|P_{W_i}x\|^p\right)^{\frac{1}{p}} \le B\|x\|, \quad \forall x \in X.$$

i.e.,

$$A\|x\| \le \left(\sum_{i\in I} v_i^p \|P_{W_i}x\|^p\right)^{\frac{1}{p}} = \left\|T_W^v x\right\|.$$

Thus, T_W^v is bounded below.

(2) $\Rightarrow$ (1). If the analysis operator T_W^v of $\{W_i\}_{i\in I}$ is bounded below, there is a constant A such that $A\|x\| \le \left\|T_W^v x\right\|, \forall x \in X$, that is,

$$A\|x\| \le \left(\sum_{i\in I} v_i^p \|P_{W_i}x\|^p\right)^{\frac{1}{p}}, \quad \forall x \in X$$

Since $\{W_i\}_{i\in I}$ is a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$, we have

$$\left(\sum_{i\in I} v_i^p \|P_{W_i}x\|^p\right)^{\frac{1}{p}} \le B\|x\|, \quad \forall x \in X,$$

Thus, for any $x \in X$, the following

$$A\|x\| \le \left(\sum_{i\in I} v_i^p \|P_{W_i}x\|^p\right)^{\frac{1}{p}} \le B\|x\|$$

is right.

(2) $\Rightarrow$ (3). If T_W^v is bounded below, then T_W^{v*} is surjective. Since $U_W^v \pi_p^{-1} = T_W^{v*}$, we can know that U_W^v is surjective.

(3) $\Rightarrow$ (2). If U_W^v is an operator of $\oplus_q W_i^*$ onto X^*, then T_W^{v*} is surjective. Hence, T_W^v has a bounded inverse T_W^{v-1} on the range $T_W^v(X)$. For any $x \in X$, we have

$$\|x\| = \left\|T_W^{v-1} T_W^v x\right\| \leq \left\|T_W^{v-1}\right\| \cdot \left\|T_W^v x\right\|$$

i.e. $\left\|T_W^{v-1}\right\|^{-1} \cdot \|x\| \leq \left\|T_W^v x\right\|$. Thus shows that T_W^v is bounded below.

If $\{W_i\}_{i\in I}$ is a frame of subspaces for X with respect to $\{v_i\}_{i\in I}$, then Theorem 2 shows that

$$X^* = \left\{ \sum_{i\in I} v_i P_{W_i}^* w_i^* : \left\{w_i^*\right\}_{i\in I} \in \oplus_q W_i^* \right\}$$

Definition 3 Let $\{v_i\}_{i\in I}$ be a family of weights. A family $\{W_i\}_{i\in I}$ of complemented subspaces of the Banach space X is said to be independent with respect to $\{v_i\}_{i\in I}$ if

$$\sum_{i\in I} v_i P_{W_i}^* w_i^* = 0, \left\{w_i^*\right\}_{i\in I} \in \oplus_q W_i^* \Rightarrow w_i^* = 0, \quad \forall i \in I.$$

Theorem 3 Let $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$. Then $\{W_i\}_{i\in I}$ is an independent frame of subspaces with respect to $\{v_i\}_{i\in I}$ for X if and only if its analysis operator T_W^v is invertible.

Proof Assume that $\{W_i\}_{i\in I}$ is an independent frame of subspaces with respect to $\{v_i\}_{i\in I}$, then its synthesis operator U_W^v is invertible from the definition of frames of subspaces and Theorem 2. Since $U_W^v \pi_p^{-1} = T_W^{v*}$, thus T_W^v is invertible.

Conversely, if T_W^v is invertible, then T_W^v is bounded below. From Theorem 2, we know that $\{W_i\}_{i\in I}$ is a frame of subspaces with respect to $\{v_i\}_{i\in I}$. Suppose that $\{W_i\}_{i\in I}$ is not independent, then there exists a non-zero sequence $\left\{w_i^*\right\}_{i\in I} \in \oplus_q W_i^*$ such that $\sum_{i\in I} v_i P_{W_i}^* w_i^* = 0$. Assume that $w_{i_0}^* \neq 0$, then we have

$$v_{i_0} P_{W_{i_0}}^* w_{i_0}^* = -\sum_{i\neq i_0} v_i P_{W_i}^* w_i^*$$

Since $w_{i_0}^* \neq 0$, there exists a vector $w \in W_{i_0}$ such that $w_{i_0}^*(w) \neq 0$. Denote $w_i = 0, i \neq i_0; w_i = w, i = i_0$, then $\{w\}_{i\in I} \in \oplus_p W_i$. Because T_W^v is invertible, there is a vector x in X such that $T_W^v x = \{w\}_{i\in I}$. Thus, on one hand, we obtain that

$$< v_{i_0} P_{W_{i_0}}^* w_{i_0}^*, x > \; = < w_{i_0}^*, v_{i_0} P_{W_{i_0}} x > \; = < w_{i_0}^*, v_{i_0} w > \; \neq 0$$

On the other hand, we have

$$< -\sum_{i\neq i_0} v_i P^*_{W_i} w^*_i, x > = -\sum_{i\neq i_0} < v_i P^*_{W_i} w^*_i, x > = -\sum_{i\neq i_0} < w^*_i, v_i P_{W_i} x >$$
$$= -\sum_{i\neq i_0} < w^*_i, v_i w_i > = 0$$

This is a contradiction! Therefore, $\{W_i\}_{i\in I}$ is independent.

Definition 4 A family $\{W_i\}_{i\in I}$ of complemented subspaces of X is said to be complete with respect to $\{v_i\}_{i\in I}$ if the space

$$\left\{\sum_{i\in G} v_i P^*_{W_i} w^*_i : \forall i \in G, w^*_i \in W^*_i, G \in F(I)\right\}$$

is dense in X^*.

Definition 5 Let $\{W_i\}_{i\in I}$ be a family of complemented subspaces of X and $\{v_i\}_{i\in I}$ be a family of weights. We call $\{W_i\}_{i\in I}$ a Riesz basis of subspaces for X with respect to $\{v_i\}_{i\in I}$ if $\{W_i\}_{i\in I}$ is complete with respect to $\{v_i\}_{i\in I}$ and there exist positive constants C and D such that

$$C\left\|\left\{w^*_i\right\}_{i\in I}\right\| \leq \left\|\sum_{i\in G} v_i P^*_{W_i} w^*_i\right\| \leq D\left\|\left\{w^*_i\right\}_{i\in I}\right\|, \quad \forall \left\{w^*_i\right\}_{i\in I} \in \oplus_q W^*_i$$

Theorem 4 Let $\{W_i\}_{i\in I}$ be a Bessel sequence of subspaces for X with respect to $\{v_i\}_{i\in I}$. Then $\{W_i\}_{i\in I}$ is an independent frame of subspaces for X with respect to $\{v_i\}_{i\in I}$ if and only if $\{W_i\}_{i\in I}$ a Riesz basis of subspaces for X with respect to $\{v_i\}_{i\in I}$.

Proof Let $\{W_i\}_{i\in I}$ is an independent frame of subspaces for X with respect to $\{v_i\}_{i\in I}$. Then Theorem 3 implies that the analysis operator T^v_W is invertible. Since $U^v_W \pi^{-1}_p = T^{v*}_W$, U^v_W is invertible. Denote by U^{v-1}_W the inverse of U^v_W. For every $\left\{w^*_i\right\}_{i\in I} \in \oplus_q W^*_i$, we have

$$\left\|\left\{w^*_i\right\}_{i\in I}\right\| = \left\|U^{v-1}_W U^v_W \left\{w^*_i\right\}_{i\in I}\right\| \leq \left\|U^{v-1}_W\right\| \cdot \left\|U^v_W \left\{w^*_i\right\}_{i\in I}\right\|$$

i.e.

$$\left\|U^{v-1}_W\right\|^{-1} \cdot \left\|\left\{w^*_i\right\}_{i\in I}\right\| \leq \cdot \left\|U^v_W \left\{w^*_i\right\}_{i\in I}\right\|$$

Let $C = \left\|U^{v-1}_W\right\|^{-1}$ and $D = \left\|U^v_W\right\|$. Then for every $\left\{w^*_i\right\}_{i\in I} \in \oplus_q W^*_i$, the following

$$C\left\|\left\{w_i^*\right\}_{i\in I}\right\| \le \left\|\sum_{i\in G} v_i P_{W_i}^* w_i^*\right\| \le D\left\|\left\{w_i^*\right\}_{i\in I}\right\|$$

is satisfied. From the invertibility of U_W^v, we can know that $\{W_i\}_{i\in I}$ is complete. Hence, $\{W_i\}_{i\in I}$ is a Riesz basis of subspaces for X with respect to $\{v_i\}_{i\in I}$.

Conversely, assume that $\{W_i\}_{i\in I}$ is a Riesz basis of subspaces for X with respect to $\{v_i\}_{i\in I}$. Clearly, the definition of Riesz bases of subspaces implies that the operator U_W^v is bounded below and has closed range. In addition, from the completeness of $\{W_i\}_{i\in I}$, we can know that U_W^v is surjective. Thus, U_W^v is invertible. Therefore, $U_W^v \pi_p^{-1} = T_W^{v*}$ yields that T_W^{v*} is invertible. And so T_W^v is invertible.

References

1. Duffin RJ, schaeffer AC (1952) A class of nonharmonic Fourier serier. Trans Amer Math Soc 72:341–366
2. Daubechies I, Grossmann A, Meyer Y (1986) Painless nonorthogonal expansion. J Math Phys 27:1271–1283
3. Chui CK (1992) An introducion to Wavelets. Academic Press Inc, New York
4. Grochenig K (1991) Describing functions: atomic decompositions versus frames. Monatsh Math 112:1–41
5. Grochenig K (2001) Foundation of time frequency analysis. Birkhäuser, Boston
6. Oswald P (1994) Multilevel finite element approximation: theory and application. Teubner Skr, Numerik
7. Casazza PG, Kovacevic J (2003) Equal-norm tight frames with erasures. Adv Comput Math 18:387–430
8. Goyal VK, Kovacevic J, Kelner JA (2001) Quantized frame expansions with erasures. Appl Comput Harmon Anal 10:203–233
9. Hassibi B, Hochwald B, Shokrollahi A, Sweldens W (2001) Representation theory for high-rate multiple-antenna code design. IEEE Trans Inform Theory 47:2335–2367
10. Fornasier M (2003) Decompositions of Hilbert spaces: local construction of global frames. In: Bojanov B (ed) Proceedings of international conference on constructive function theory, Varna (2002) DARBA, Sofia 275–281
11. Fornasier M (2004) Quasi-orthogonal decompositions of structured frames. J Math Anal Appl 289:180–199
12. Cassaza PG, Kutyniok Gitta (2004) Frames of subspaces. Contemp Math 345:87–113
13. Conway JB (1995) A course in functional analysis. Springer, New York

Author's Biography

Professor Xudong Wang received his Ph.D. degree in electrical engineering at Xi'an Jiaotong University in 2002. His research interest is in linear motor and electrical control. He was the deputy director of the department of electrical engineering, and now serves as the vice dean of the School of Electrical Engineering and Automation at Henan Polytechnic University, China.

Professor Fuzhong Wang graduated from the Wuhan University, China in 1983. He served as deputy director of the department of power and the director of the department of computer applications. His research interests are in the SCM and process control teaching and research. Currently he is the vice dean of the School of Electrical Engineering and Automation at Henan Polytechnic University, China.

Dr. Shaobo Zhong received his bachelors and the master's degrees in mathematics and computer science from Chongqing Normal University, China, in 1995 and 1998 respectively, and earned his Ph.D. in computer science from Chongqing University, China, in 2008. His research interests include machine learning, data mining and web page classification. Now he is an associate professor at Chongqing Normal University in China.

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2,

Author Index

Volume 1

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2,

Author Index

Volume 2

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*,
Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2,

Author Index

Volume 3

X. Wang et al. (eds.), *Electrical, Information Engineering and Mechatronics 2011*, Lecture Notes in Electrical Engineering 138, DOI: 10.1007/978-1-4471-2467-2,