Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning

Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning

Edited by

JOS P. VAN LEEUWEN

Eindhoven University of Technology,
Department of Architecture, Building and Planning,
Eindhoven, The Netherlands

and

HARRY J.P. TIMMERMANS

Eindhoven University of Technology,
Department of Architecture, Building and Planning,
Eindhoven, The Netherlands

KLUWER ACADEMIC PUBLISHERS

NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-2409-6
Print ISBN: 1-4020-2408-8

Created in the United States of America

Visit Springer's eBookstore at: http://ebooks.kluweronline.com
and the Springer Global Website Online at: http://www.springeronline.com

TABLE OF CONTENTS

Simulation and Agent Technology

Design Research and Design Support Systems

PREFACE

As editors of this volume we would like to express our gratitude to the contributing authors who have delivered these highly relevant and inspiring chapters and to the international scientific committee for their help in the review and editing. Special thanks go to our colleagues Mandy van de Sande, Marlyn Aretz, and Leo van Veghel, who were great in organising the conference and took care of everything that made it a pleasant and comfortable event.

Eindhoven, July 2004

Jos van Leeuwen and Harry Timmermans
Conference Chairs

INTERNATIONAL SCIENTIFIC COMMITTEE

INTRODUCTION

The International Conference on Design & Decision Support Systems in Architecture and Urban Planning is organised bi-annually by the Eindhoven University of Technology. This volume contains a selection of papers from the seventh conference that was held at De Ruwenberg Castle in Sint-Michielsgestel, The Netherlands, from 2 to 5 July, 2004.

Traditionally, the DDSS conferences aim to be a platform for both starting and experienced researchers who focus on the development and application of computer support in the areas of urban planning and architectural design. This results in an interesting mix of well-established research projects and first explorations. It also leads to a very valuable cross-over of theories, methods, and technologies for support systems in the two different areas, architecture and urban planning. This volume contains 22 peer reviewed papers from this year's conference.

The chapters in this volume are organised into 5 sections. The first section, on artificial intelligence, deals with three applications of neural networks. The second section is about the role of visualisation in design and decision support. This section includes three chapters on how visualisation can improve public participation in governmental decision-making processes and two chapters on the potential of visualisation and design support systems in architectural design.

The third section deals with simulation technology. Simulation is implemented using agent technology in three chapters, while two other chapters implement micro-simulation techniques and cellular automata respectively. Section four contains six chapters on design research and the development of systems for design support. Finally, the fifth section includes three papers in the area of geographical information systems.

The following paragraphs provide an overview of these five sections.

Applications of Artificial Intelligence

Being able to predict the risks related to soil and rock behaviour is of vital importance in many parts of the world. Statistical methods are generally used but are not sufficient to deal with the complex nature of phenomena such as landslides. *Yesilnacar and Hunter* present a method that utilises a neural network for enhanced predictions of landslide susceptibility. They have been able to demonstrate that their neural network model can deal with the multivariable problem of landslide susceptibility and that it is capable of pointing out the key factors for the risk of landslides.

Optimal planning of facilities, such as schools, is essential when resources are scarce. It requires careful prediction of urban and social development and activities. *Akamine and Rodrigues da Silva* evaluated an application of Artificial Neural Networks in the form of a neural spatial interaction model to address this problem. They discuss the possible reasons why this model behaves differently from the gravity models that are traditionally used in this context.

The chapter written by *Diappi, Bolchi, and Buscema* is about spatial distribution and evolution. They use neural networks (self-organising maps) to simulate the evolution of land use patterns in Milan at two different temporal thresholds between 1980 and 1994 to 2008. The use of neural networks implies that the model can learn trends in spatial evolution.

Visualisation for Design and Decision Support

Governmental organisations are often confronted with the problem of how to make use of a large variety of information resources and how to make this information available to the public. Large scale participation of civilians in governmental decision-making processes regarding the built environment is becoming more feasible with the application of online technologies.

Pettit, Nelson, and Cartwright describe an Australian case study of community participation in land use planning. The case study implemented a host of digital media to allow for well-informed collaborative decision-making. *Barton, Parolin, and Weiley* developed a platform that supports integration and interoperability of many data sources concerning social, financial, and physical aspects of the urban environment. Their Spatial Decision Support Systems proposes a variety of visual interfaces to allow access to this data by both experts and non-experts.

Another interesting chapter about visualization is the one by *Shen and Kawakami*. They discuss the development of an on-line planning support system using WEBGIS in the context of public participation. Especially of interest is the expansion of the possibilities of WEBGIS to 3D *visualization*.

The ongoing development of VR technology has prompted *Göttig, Newton, and Kaufmann* to evaluate and compare a range of contemporary systems in the context of architectural models and abstract data. Aspects such as the quality of display, the quality of the spatial impression, the ease of navigation, and the effect on creativity in design are compared for four different systems. The final conclusion is not overall positive with respect to the capabilities of these systems to support early design tasks.

An interesting approach to design support is offered by *Lee and Qian*. They describe a colour coordinate system that supports the design process by matching designers' feelings, described using adjective words, with colours taken from an indexed colour database. One function of the system is to analyse colour images and match them with adjectives and vice versa, while a second function of the system deals with various kinds of colour harmonies. The latter function supports designers in selecting a complete colour scheme that matches their aesthetic intentions.

Simulation and Agent Technology

This volume contains two chapters about progress in developing the Masque system, a multi-agent framework for generating alternative plans in local land-use planning. It is based on a set of models, formalised expert knowledge, and a different set of decision support tools. In this framework the agents represent land-use experts and initiate the development of plan proposals and request each other to express their claims in order to incrementally draw up these proposals. Thus, some agents take care of the actual planning process, while others are used to represent existing knowledge or simulate likelihood outcomes of plan decisions. *Ma et al* outline how Bayesian belief networks are used to represent expert knowledge as part of this system. This approach is of interest in that it allows the consideration of uncertainty. *Devisch et al* discuss the use of agents to simulate housing choice. The model, still under development, focuses on the process of decision-making as opposed to observed outcomes.

Another development of a multi-agent system is presented by *Bandini, Manzoni, and Vizzari*. Their system takes advantage of elements of cellular automata and defines a model of situated cellular agents. Through this combination, the autonomous agents can operate and interact in a geometrical representation of either a physical or abstract reality. The agents' behaviour can be influenced by their spatial position and relation to adjacent as well as distant agents. The area of application is crowd modelling, where individuals are represented by interacting agents that behave according to their position and signals from elements in their environment. The multi-agent system is tied to a 3D modelling tool to allow for 3D visualisation of the environment and the agents behaviour.

The approach to simulation by *Ballas, Kingston, and Stillwell* is not based on multi-agents, but rather they support the use of general spatial micro-simulation techniques, such as simulated annealing. The chapter discusses and illustrates the potential of their spatial micro simulation-based decision support system for policy analysis. In particular, the system can be used to describe current conditions and issues in neighbourhoods, predict future trends in the composition and health of neighbourhoods and conduct modelling and predictive analysis to measure the likely impact of policy interventions at the local level.

Cellular automata models have recently gained a lot of interest in the urban planning literature. These models are especially relevant to modelling the evolution of some spatial phenomenon. *Ohgai et al* report a state-of-the-art application to the problem of fire spread, providing evidence to the potential of such models in disaster mitigation.

Design Research and Design Support Systems

The ARTHUR project, described by *Penn et al.*, has developed a designers' meeting table that adds digital visualisation to conventional media. Collaborating designers use so-called optical see through augmented reality displays to look at a virtual environment that is blended into the real objects on the table, while still being able to see each other as well. This chapter not only provides a detailed discussion of the system's design and implementation, but also offers a thorough discussion on design theory and the possible impacts of innovative tools on the nature of the design process.

While the importance of images as a resource in architectural design is evident, finding relevant images is still a challenging task. The approach presented by *Kacher, Bignon, and Halin* devises a hierarchy of terms, with given weights, to index images and the architectural elements represented by them. The key for the correct indexing of images is a semantic and unambiguously structured thesaurus that the authors developed to describe the images. Five criteria are used to determine the weight of the relationships between an image and the terms from this thesaurus.

There are two chapters on the application of case-based reasoning. *Bi and Medjdoub* address the problem of how to plan building services in a given space configuration. Their approach offers an interesting combination of case-based reasoning and constraint-based adaptation. With an application for ceiling layout for fan coil systems, they demonstrate how adaptation of retrieved cases can quickly lead to a better design. A strong schematisation of space layouts and sequential constraint solving are instrumental to reducing the complexity of problems.

In his chapter of case-based design, *Lindekens* argues that reduction of context information is essential in design decision-making. Without the ability to reduce the abundance of information that is usually involved in design, designers would not be able to take design decisions. Likewise, the reduction of information makes it easier for others to understand the design rationale. Transformation of information between design steps is the key to understanding the process. Based on this theory, Lindekens proposes a case-based approach to support design and re-design.

One elemental aspect of architecture is the notion of *level*. The chapter by *Bax and Trum* on Domain Theory defines the notion of level in architecture in order to structure both architectural objects and architectural design processes simultaneously in a consistent way. They discuss a number of principles to identify levels and types of levels. This leads to a hierarchy of levels of control, complexity, and composition, which are oriented towards construct, product, and form respectively.

Chang explores the latest developments in the Web-based Architectural Learning Environment that now supports design learning using design puzzles. This development acknowledges the fact that design is not simply about solving design problems as a puzzle, but first of all about creating the puzzle itself. It therefore supports both making the puzzle and solving it. In the WALE environment, which is a multi-user role-playing environment, design students can, in an exploring manner, make and solve interactive design puzzles.

Geographical Information Systems

The final section of this volume contains three chapters on geographical information systems. *Tisma* presents the development of a new tool, Rasterplan, which goes beyond visualisation. Her tool also supports the allocation of future land use. It allows the realization of a quantitative program for future spatial needs for various functions such as housing, green and water areas, working, and recreation. In addition to quantitative calculations, qualitative criteria for location choice can be also expressed in a form of suitability maps or buffers.

Another utilisation of GIS databases is presented by *Rodrigues, Souza, and Mendes* in an extension of the tool ArcView. This extension, called 3DSkyView, is developed to assess and visualise sky view factors of urban environments. The sky view factor, through its thermal and geometrical parameters, is an indication for urban heat islands and can be used to identify problems in the canyons of an urban landscape. With the current implementation of this tool, the authors are able to visualise this factor in the continuum of an entire urban area.

Finally, a detailed study of how geographical information can be used in the analysis of the impact of the urban environment on public safety is shown by *Murakami, Higuchi, and Shibayama*. Through in-depth analyses using a GIS database, they are able to relate criminal incidents to the local urban layout and road network. From these observed relationships, they intend to objectively predict unsafe locations in urban environments on the basis of GIS data.

Applications of Artificial Intelligence

Application of Neural Networks for Landslide Susceptibility Mapping in Turkey

Ertan Yesilnacar and Gary J. Hunter
The University of Melbourne, Australia

Keywords: Landslide Susceptibility Mapping, Neural Networks, Spatial Decision Support Systems.

Abstract: Landslides are a major natural hazard in many areas of the world, and globally they cause hundreds of billions of dollars of damage, and hundreds of thousands of deaths and injuries each year. Landslides are the second most common natural hazard in Turkey, and the Black Sea region of that country is particularly affected. Therefore, landslide susceptibility mapping is one of the important issues for urban and rural planning in Turkey. The reliability of these maps depends mostly on the amount and quality of available data used, as well as the selection of a robust methodology. Although statistical methods generally have been implemented and used for evaluating landslide susceptibility and risk in medium scale studies, they are distribution-based and cannot handle multi-source data that are commonly collected from nature. These drawbacks are responsible for the on-going investigations into slope instability. To overcome these weaknesses, the desired technique must be able to handle multi-type data and its superiority should increase as the dimensionality and/or non-linearity of the problem increases – which is when traditional regression often fails to produce accurate approximations. Although neural networks have some problems with the creation of architectures, processing time, and the negative "black box" syndrome, they still have an advantage over traditional methods in that they can deal with the problem comprehensively and are insensitive to uncertain data and measurement errors. Therefore, it is expected that the application of neural networks will bring new perspectives to the assessment of landslide susceptibility in Turkey. In this paper, the application of neural networks for landslide susceptibility mapping will be examined and their performance as a component of spatial decision support systems will be discussed.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 3-18.

1. INTRODUCTION

Landslides are a major natural hazard in many areas of the world, and globally they cause hundreds of billions of dollars of damage, and hundreds of thousands of deaths and injuries each year. Landslides are the second most common natural hazard in Turkey and the Black Sea region of Turkey is particularly affected. Therefore, landslide susceptibility mapping is one of the important issues for urban and rural planning in Turkey. The reliability of these maps depends mostly on the amount and quality of available data used, as well as on the selection of a robust methodology.

In many circumstances, our fundamental understanding of soil and rock behaviour still falls short of being able to predict how the ground will behave. In most cases, expert judgment plays an important role, and empirical approaches for assessments are widely used. Landslides are complicated processes, mainly because of the many different casual factors involved in the generation of the phenomenon (such as lithology, geological structures, and seismic activities) and geomorphologic features (such as slopes, relative relief, land-use, ground water conditions and climate). This is why a significant number of methods and techniques have been proposed or tested for landslide susceptibility mapping. Broadly speaking, they may be qualitative or quantitative, and direct or indirect (Guzetti et al., 1999). A great variability in scale and mapping procedures exists: in fact, the choice of type and scale of the map depends on many factors, primarily being the requirements of the end user and the ultimate purpose of the application (Varnes, 1974, 1984). Despite all these efforts, no agreement has yet been reached on the techniques and methods for landslide susceptibility mapping (Guzetti et al., 2000).

Although statistical methods generally have been implemented and used for evaluating landslide susceptibility and risk in medium scale studies, they are distribution-based and cannot handle multi-source data that are commonly collected from nature. These drawbacks are responsible for the on-going investigations into slope instability. To overcome these weaknesses, the desired technique must be able to handle multi-type data and its superiority should increase as the dimensionality and/or non-linearity of the problem increases – which is when traditional regression often fails to produce accurate approximations.

Neural Networks recently emerged as computational modeling tools that have found extensive acceptance in many disciplines for modeling complex real-world problems (Basheer and Hajmeer, 2000). Neural networks may be defined as structures comprised of densely interconnected adaptive simple processing elements (called artificial neurons or nodes) that are capable of performing massive parallel computations for data processing and knowledge presentation (Hecth-Nielsen, 1990; Schalkoff, 1997). Many kinds of neural network topology and learning algorithms have been developed as a result of different interconnection strategies. Neural network models can be categorized in terms of two criteria. The first one is based on whether the model employs a supervised or an unsupervised learning strategy. In supervised models input and output information is provided to adjust the weights in such a way that the network can produce the outputs from the given inputs, while in unsupervised models only input information is provided to determine the possible classes in the dataset. The second criterion relates to the directionality of the learning method associated with the network topology. If the information advances from input layer to output layer, the learning method is called "feed-forward". Conversely, if the information proceeds from output layer to input layer, the network is termed "feed-back". The most common neural network topology is Multi-Layer Perceptrons (MLPs) with a feed-forward back-propagation learning algorithm. It is used in about 70% of real-world applications (Werbos, 1995).

Unlike traditional statistical methods, neural networks do not require assumptions about the form or distribution of data to analyze it; hence no prior knowledge of the data is needed (Benediktsson et al., 1990). Neural networks also show the remarkable information processing characteristics of biological systems, such as nonlinearity, high parallelism, robustness, fault- and failure-tolerance, learning capacity, ability to handle imprecise and fuzzy information, and capability to generalize (Jain et al., 1996). Finally, neural network models require less formal statistical training to develop. Hect-Nielsen (1990) reports that neurocomputing is now providing a breath of fresh new air to the 200+ year old gaussian statistical regression.

Sui (1993) successfully integrated MLPs with GIS for the development of suitability analysis, and found that a neural network can make a close approximation of experts' decisions without the explicit expression of experts' knowledge into "if-then" production rules. Wang (1992) used MLPs to strengthen the spatial data modeling capabilities of GIS. Aleotti et al. (1998) and Gomez (2002) applied this approach for landslide susceptibility mapping and found it out-performed traditional statistical methods. Mayoraz and Vulliet (2002) used neural networks for slope movement prediction, while Lee et al. (2003) also used MLPs to determine the weights of landslide factors for susceptibility mapping.

In this paper, the application of neural networks for landslide susceptibility mapping will be examined and their performance as a component of spatial decision support systems will be discussed.

2. NEURAL NETWORKS FOR LANDSLIDE SUSCEPTIBILITY MAPPING

A feed-forward back-propagation (a gradient descent algorithm, also called a generalized delta rule) was used in this study. It is the best known procedure for training neural networks. It is based on searching a performance surface (where error is a function of neural network weights) using gradient descent for point(s) with minimum error. Each iteration in the algorithm constitutes two sweeps: forward activation to produce a solution, and a backward propagation of the computed error to modify the weights. The forward and backward sweeps are performed repeatedly until the neural network solution agrees with the target value within a prespecified tolerance. The learning algorithm provides the needed weight adjustments in the backward sweep (Figure 1).

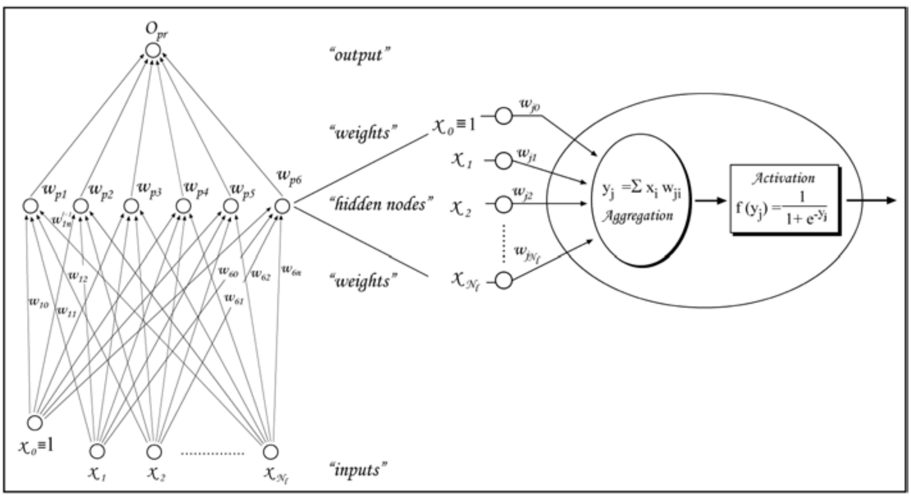


Figure 1. Feed-forward back-propagation neural network

There are important issues relevant to the design and use of feed-forward back-propagation neural networks and they should be addressed before initiation of the network. It is common practice that users design their networks using trial-and-error strategies and employ predefined rates for the learning parameters. It is also a fact that the optimum rates for the parameters and the size of the network required are problem dependent, and should be determined individually for each dataset and network structure. An extensive examination is thus necessitated to enable new users to apply neural network models confidently and successfully. The important design parameters for network generalization are: 1) number of hidden nodes; 2) size of training and test datasets; 3) initial weight range; 4) learning rate;

5) the momentum term; and 6) number of training cycles. Extreme values of these parameters cause some severe effects on training convergence and network generalization (Table 1).

Table 1. Effect of extreme values of design parameters (modified from Basheer and Hajmeer, 2000)

Design Parameter	Too High or Too Large	Too Low or Too Small
Number of Hidden Nodes	Causes overfitting and slow training	Causes underfitting (NN unable to obtain the underlying relation function embedded in the data) but fast training
Size of Training Sets	NN with good recall and generalization	NN unable to fully explain the problem. NN with limited or bad generalization
Size of Test Sets	Ability to confirm NN generalization capability	Inadequate confirmation of NN generalization capability
Initial Weight Range	Solution trapped in a local minimum or in a very flat plateau close to the starting point	May cause better generalization but no guarantee
Learning Rate (η)	Speeds up the convergence however unstable NN (weights) oscillates about the optimal solution	Causes slow training and a greater likelihood of becoming trapped in a local minimum
Momentum Term (μ)	Reduces risk of local minima. Speeds up training. Causes unstable learning	Suppresses effect of momentum leading to increased risk of potential entrapment in local minima. Causes slow training
Number of Training Cycles	Good recall of NN (memorization of data) but causes overfitting in the unseen data	Produces NN that is incapable of representing the data

Traditionally spatial decision support systems rely on statistical approaches in analysing data. As mentioned earlier, neural networks offer a better alternative to statistical methods in certain circumstances therefore they can be used as a component of spatial decision support systems, such as for diagnosis and prognosis of spatial phenomena (Wei et al., 2002; Tchaban et al., 1997). Their outcomes can take critical roles in spatial decision-making and development plans.

3. A CASE STUDY IN TURKEY

A pipeline was constructed in 1997 in order to supply natural gas to the Eregli Steel Factory in Turkey. However, a serious landslide occurred near Hendek (Figure 2). A section of the pipeline was broken due to the landslide with a resulting fire. The fire was extinguished after 3 days and re-routing of the problematic section of the pipeline is still under consideration. The main requirement for re-routing studies is the preparation of sensitive landslide susceptibility mapping at the medium scale. A neural network model was used to produce a landslide susceptibility map by analyzing several intrinsic factors controlling the landslides. Then the accuracy of this map was assessed. According to the assessment result, the susceptibility map will be used as a component of a spatial decision support system to produce a development map for re-routing the pipeline in the future.

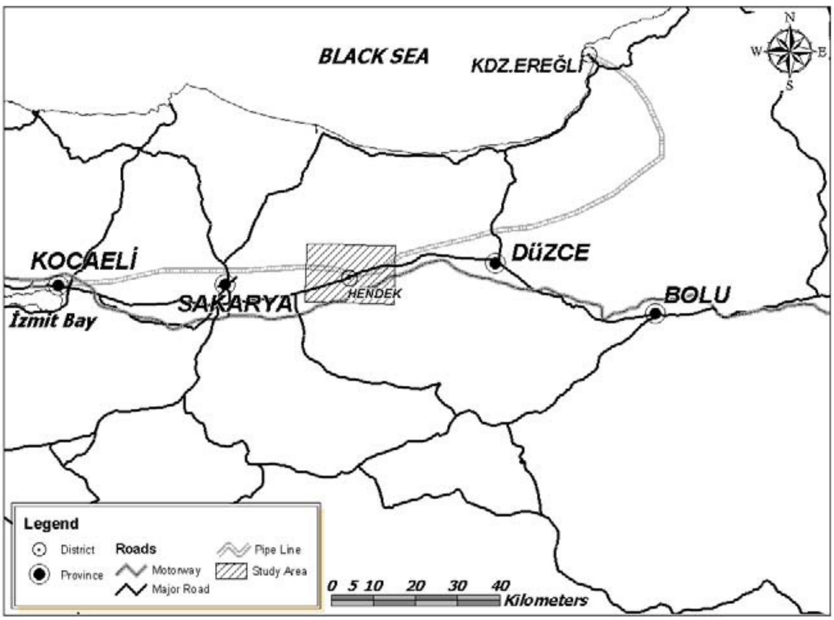


Figure 2. Study Area

The study area covers about 290 km^2. Climatic characteristics of the region resemble the characteristics of both the Marmara and Black Sea regions. In winter weather conditions are generally rainy and mild and in summer, hot. Relative humidity levels are high for every season.

The landslide inventory map is essential for landslide susceptibility mapping because it represents the spatial distribution of mass movements and also includes information about their location, typology and state of activity (Carrara and Meranda, 1976; Hansen, 1984; Wieczorek, 1984; Einstein, 1988; Soeters and Van Westen, 1996). In this study, the landslide inventory map was prepared by Emre et al. (1999) and Duman et al. (2001), however each landslide body was also checked by various field studies. Finally, a total of 112 landslide bodies were extracted and mapped on 1:25 000 scale topographical maps. These bodies were sorted according to their modes of occurrence. This helped in understanding the different triggering factors controlling different slope movement types. Landslides are mainly rotational with a few being translational, flow or complex in type. Flow and complex types were eliminated before the analysis because there were very few numbers of landslides observed as being of these types. Therefore our resultant map shows susceptibility levels of only rotational slides. Watershed basin polylines were taken into account to establish the study area border because it makes the study area more meaningful. The landslide inventory map and final study area border are seen in Figure 3.

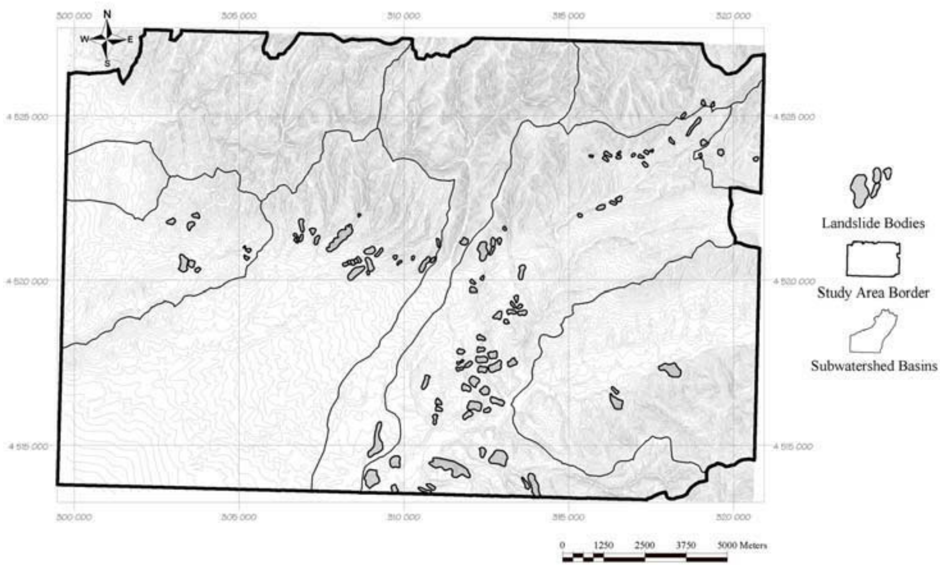


Figure 3. Landslide Inventory Map

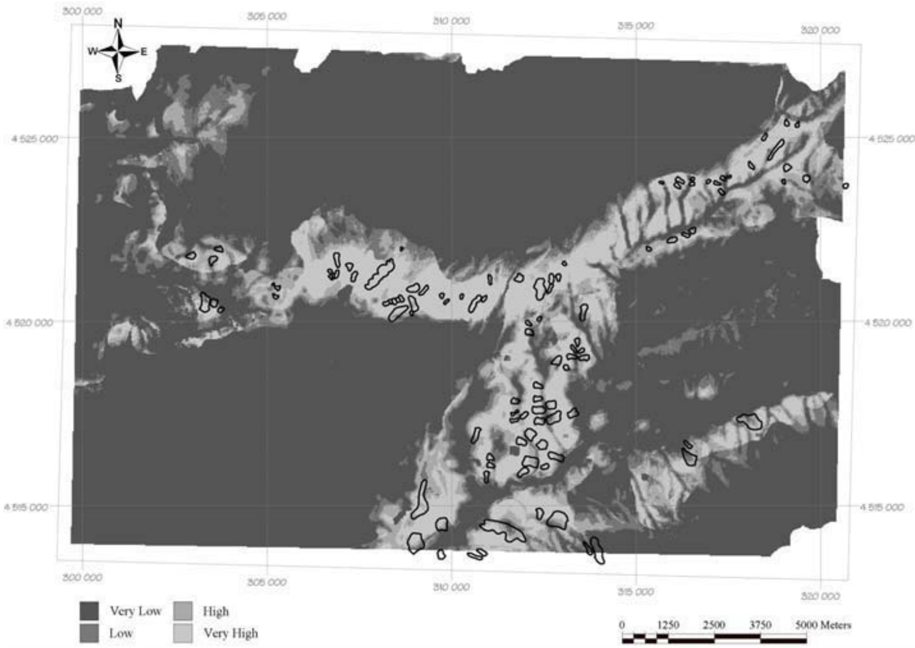
The causes of slope failures are many, complex and sometimes unknown. In statistical approaches, experts have to consider the determination of triggering factors but neural networks have the capability to evaluate the contribution of each factor controlling the instability by looking at training data sets. 19 factors causing slope movements were taken into consideration: Aspect, Distance to Drainage Network, Distance to Fault Planes, Distance to Ridges, Distance to Roads, Drainage Density, Elevation, Fault Density, Geology, Land Cover, Plan Curvature, Profile Curvature, Road Density, Slope, Slope Length, Stream Power, Surface Area Ratio, Topographic Wetness Index, and Subwatershed Basins. All layers were prepared in raster format having 25 meters by 25 meters pixel size. After completing the landslide inventory map preparation and input layers, training data sets were created. For this purpose, a buffer of 100 meters was created around the crown and flanks of the landslide polygons. The aim of this process is to get the pre-failure surface attributes of the landslides. Then a polygon grid having the same extension with the study area and 25-meter pixel size was created. The midpoint of each polygon grid was calculated and then these points were merged with previously created buffer polygons to clip out the ones that fall into these buffer zones. At the last stage, values of input layers were transferred into the database of clipped out points. This technique is called as "*Seed Cell Theory*" (Suzen and Doyuran, 2004). A similar process was followed for the test data set but this time landslide body polygons were used.

"Neurosolutions" software was used in the neural network analyses. Several iterations (a total of 276) were tested to find the best number of hidden nodes and the learning rate-momentum term couple for the problem. Our focus in evaluating the system's performance will be generalization. This refers to the ability of a trained neural network to respond correctly to input not used during the training process. Therefore, we train our models with one partition of the data set and evaluate their performance with a separate partition not used during training. Finally, 13 hidden nodes, with a learning rate of 0.3 and a momentum term of 0.8 with 40000 iterations were found to be the best structure for our case. It produces 82.12% overall accuracy with the test dataset. The weights of neural networks can be followed from input to output, the product of each route giving an indication to the relative importance (or influence) of the chosen input. However, this can only be a coarse measure of an input's effect due to the non-linear nature of NN neuron activation functions. This method can be used to prioritise triggering factors that cause landslides and for this purpose, the formulae proposed by Zhou, 1999 and Lee et al., 2003, were used. The results show eight variables, namely Fault Density, Elevation, Slope, Profile Curvature, Plan Curvature, Distance to Roads, Slope Length and Drainage Density, to have considerable effects on the output (the occurrence of a landslide) (Table 2). These variables have higher influence than the average influence of the total of 19 inputs. Field and desk studies also support the effect of slope, profile curvature and slope length on landslide phenomena. Aspect, Subwatershed Basins and Stream Power, which have relatively lower importance values on the output node of neural networks, are unnecessary factors for landslides occurring in this area. The results confirmed the view before the study that these three layers were unnecessary factors.

Table 2. Relative importance of input nodes in the feed-forward back propagation learning algorithm with respect to output node (Average Relative Importance of Input Nodes is equal to 1)

	Input Node	Relative Importance on the Output Node
Neural Network Feed-Forward Back-Propagation Learning Rule	Fault Density	2.41
	Elevation	2.36
	Slope	1.76
	Profile Curvature	1.73
	Plan Curvature	1.43
	Distance to Roads	1.23
	Slope Length	1.14
	Drainage Density	1.00
	Land Cover	0.88
	Distance to Ridges	0.86
	Geology	0.75
	Road Density	0.59
	Surface Area Ratio	0.59
	Distance to Drainage	0.55
	Topographic Wetness Index	0.53
	Distance to Fault	0.52
	Stream Power	0.29
	Subwatershed Basins	0.23
	Aspect	0.15

The landslide susceptibility map produced from the feed-forward back-propagation learning algorithm is seen in Figure 4. In order to evaluate the model, percentages of very high, high, low and very low susceptibility levels in total area, in landslide bodies and in seed cells were calculated. The preferred map should have a low total percentage of high and very high susceptible areas in the total area and high percentages in the seed cells and landslide bodies. The low and very low susceptibility classes constitute 82.12% of the total area with a corresponding 4.35% of the total seed cells. The rest of the area is classified as high and very high hazard that in 17.88% of the study area with a corresponding 95.65% of the total seed cells. Although the distribution of very high and high susceptible areas is low in the study area, the total percentage of both zones in the seed cells (95.65%) and landslide bodies (80.42%) are quite high. In addition, some critical parts in the study area were also checked by several field visits to assess the sensitivity of the model result.



Susceptibility Levels	PTA	PLB	PSC	PTA/PSC
Very Low	73.24	7.31	0.96	76.29
Low	8.88	12.27	3.39	2.62
High	7.89	24.89	18.85	0.42
Very High	9.99	55.53	76.8	13.0
Total of VH & H	17.88	80.42	95.65	0.55

Figure 4. The landslide susceptibility map of the feed-forward back-propagation learning algorithm and its statistical indicators for the assessment (PTA: percentage in total area; PLB: percentage in landslide bodies; PSC: percentage in seed cells)

Another way to evaluate regression models is by the Relative Operating Characteristics (ROC) curve (Swets, 1988; Pontius and Schneider, 2001). The ROC curve is a method that describes the relation between true and false predictions of the regression model for different cut-off values of the probability. The ROC curve is based on the true-positive proportion and the false-positive proportion (Swets, 1988). The true-positive proportion and the false-proportion will change when different cut-off values are used. The most ideal model shows a curve that has the largest Area Under the Curve (AUC); the AUC theoretically varies from 0.5 to 1.0. If the model does not predict the occurrence of the landslide any better than a random approach the AUC would equal 0.5, like the diagonal line in Figure 5. A ROC curve of 1 shows a perfect prediction. In an attempt to assess the performance of neural networks in modelling landslide susceptibility mapping, the ROC curve was used in this study. For this purpose, a test data set was prepared using randomly selected samples from landslide bodies and safe zones.

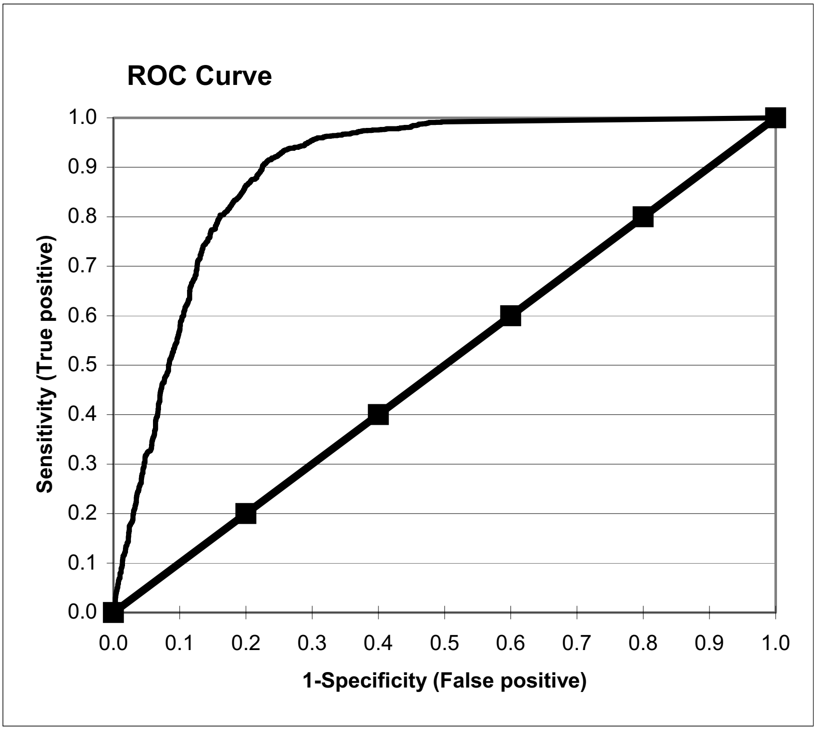


Figure 5. The ROC curve of the Feed-Forward Back-Propagation learning algorithm for "Very High Susceptible Areas", area under the curve (AUC) is 0.89

The ROC curve shows a more rapid increase in the first cut-off values and the area under the curve is 0.89, which is a very good result for prediction purposes.

In the spatial decision support system, determination of problem related thematic maps is important to making appropriate decisions and development plans. In this study, the landslide susceptibility map is a necessity to produce the development plan for the problematic segment of the natural gas pipeline (Figure 6). The evaluation of the study showed that our resultant map is successful for re-routing studies. One of the advantages of using neural network in this study is that user does not need to understand the contribution of probable triggering factors because the model evaluates the relative importance of each input factor by looking at training samples, which decrease the user's role in the selection of necessary factors for landslide susceptibility mapping.

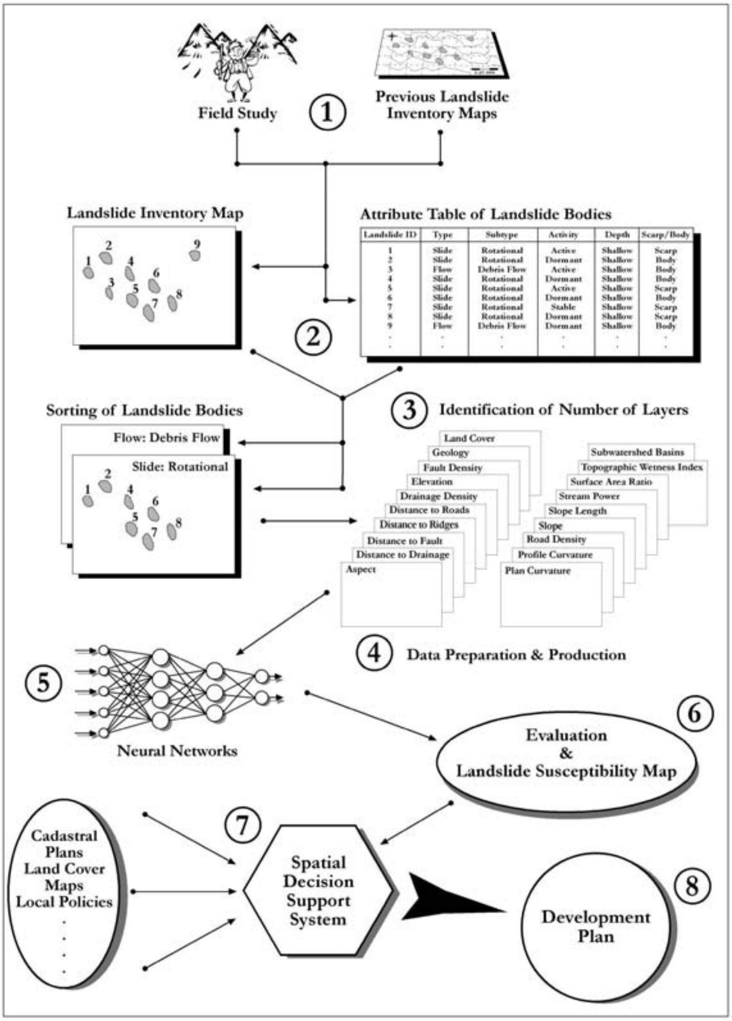


Figure 6. Components of Spatial Decision Support System.

4. DISCUSSION

There are both advantages and disadvantages in using neural networks for spatial analyses.

Input data are always affected by uncertainty and errors which cannot generally be evaluated and controlled. Neural networks show the remarkable information processing characteristics of biological systems, such as fault- and failure-tolerance, learning capacity, ability to handle imprecise and fuzzy information, and capability to generalize (Jain et al, 1996). Neural network models also require less formal statistical training to develop. The independent variables in the neural network usually undergo a nonlinear transformation at each hidden node and output node, and thus a neural network can potentially model much more complex nonlinear relationships than any statistical model can. Sometimes it can be very difficult to extract the most important instability factors in the phenomena and these factors cannot also be observed by the naked eye. However, the relative influence of the inputs (or all probable triggering factors) on the output node can be calculated which in turn prioritises factors that cause landslides. This calculation also makes neural networks a “grey box” rather than a “black-box”.

There are some important issues relevant to the design and use of feed-forward back-propagation neural networks. Model developers need to go through an empirical process of performing sensitivity analyses on training parameters such as learning rates, momentum terms, number of hidden nodes and number of training cycle. Some heuristic approaches exist for predetermining which neural network structure and parameter combinations will perform best for a given problem.

Neural network models can be used as a component of spatial decision support systems. They are powerful tools for diagnosis and prognosis of spatial phenomena. Therefore outcomes of these models can be important for decision makers in spatially difficult and critical cases. In Turkey, urban and rural planning is a critical issue for some regions. Planners and decision makers need to have hazard evaluations or risk maps for making their development plans. At this point neural networks can analyse and produce appropriate background maps for the early planning stages.

On the other hand there are some disadvantages in using neural networks. Neural network model development is a computationally intensive procedure with a standard personal computer hardware and gradient descent algorithms, it can often take hours or days before a network will converge to an optimum learning state with minimum error. Quicker variants of the gradient descent algorithm (second order learning algorithms) have been developed; among them the scaled conjugate gradient is a robust technique, however it still requires quite large amounts of training time.

5. CONCLUSION

Landslides are complex multivariable events, caused by an interaction of factors that are not always completely understood.

The use of neural networks at the medium scale allows the analysis of relationships between the factors determining the areas prone to landslides and their geographical distribution. This paper describes the possibility of the successful application of neural networks, on the issue of landslide susceptibility mapping and the contribution of neural networks to spatial decision support systems.

In this paper, a case study was used to show the practical application of the proposed model. In the Hendek region of Turkey, a section of a natural gas pipeline was broken due to landslides and 19 probable factors causing landslides were taken into hand to prepare the landslide susceptibility map for re-routing studies. Neural networks were used in the analysis stage of the study. Among several trials, the best network architecture produced 82.12% overall accuracy with the test dataset. Although the distribution of very high and high susceptible areas is low (17.88%) in the resultant map, the total percentage of both zones in the seed cells (95.65%) and landslide bodies (80.42%) are quite high. The ROC curve shows also good prediction and the area under the curve is 0.89. Moreover, critical slope faces observed in the field were checked in the resultant map and most of them fall into high or very highly susceptible zones. Neural networks also extracted the relative importance of factors on the phenomena, which are in agreement with expert observations. Based on the results of this study, it is evident that the neural network model has attractive predictive properties and is effective in extracting the contribution of triggering factors to landslides. Concerning the evaluation of the model, the final susceptibility map was compared with observations and records in test areas, and the conclusions were reasonable.

REFERENCES

Aleotti, P., P. Baldelli, G. Polloni and F. Puma, 1998, "Keynote paper: Different approaches to landslide hazard assessment", *Proceedings of the Second Conf Environ Management (ICEM-2)*, Wollongong (Australia), 1, p. 3-10.

Basheer, I. A. and M. Hajmeer, 2000, "Artificial neural networks: fundamentals, computing, design, and application", *Journal of Microbiological Methods*, Vol. 43, p. 3-31.

Benediktsson, J. A., P. H. Swain and O. K.Ersoy, 1990, "Neural network approaches versus statistical methods in classification of multisource remote sensing data", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 4, p. 540-552.

Carrara, A. and L. Meranda, 1976, "Landslide inventory in Northern Calabria, Southern Italy", *Bulletin of Geological Society Am.*, Vol. 87 (8), p. 1153-1162.

Duman, T.Y., Ö. Emre, T. Çan, S. Ates, M. Keçer, T. Erkal, S. Durmaz, A. Dogan, E. Çörekçioglu, A. Göktepe, E. Cicioglu and F. Karakay, 2001, "Turkish landslide inventory mapping project: Methodology and results on Zonguldak quadrangle (1/500 000)", *Fourth International Turkish Geology Symposium.*

Einstein, H. H., 1988, "Special lecture: landslide risk assessment procedure", *Proceedings of 5th International Symposium on Landslides*, Lausanne, 2, p. 1075-1090.

Emre, Ö., S. Ates, M. Keçer, T. Erkal, T. Yilmaz and S. Durmaz, 1999, "Pazarcik-KDZ Eregli dogalgaz ana iletim hatti Pasaköy (Adapazari) – KDZ Eregli arasinin heyelan etüdü", *MTA Report No. 10187* (in Turkish).

Gomez, H. R., 2002, "Modeling landslide potential in the Venezuelan Andes", *PhD Thesis*, The University of Nottingham, School of Geography, 272 p. (unpublished)

Guzzetti, F., M. Cardinalli, P. Reichenbach and A. Carrara, 2000, "Comparing Landslide Maps: A Case Study in the Upper Tiber River Basin", *Environmental Management*, Vol. 25, No. 3, p. 247-263.

Guzzetti, F., A. Carrara, M. Cardinalli, and P. Reichenbach, 1999, "Landslide hazard evaluation: A review of current techniques and their application in a multi-case study, central Italy", *Geomorphology*, Vol. 31, p 181-216.

Hansen, A., 1984, "Landslide hazard analysis", In: Brunsden, D. and Prior, D. B. (Eds), *Slope instability*, John Wiley and Sons, New York, p. 523-602.

Hecth-Nielsen, R., 1990, *Neurocomputing*, Addison-Wesley, Reading, MA.

Jain, A. K., J. Mao and K. M. Mohiuddin, 1996, *Artificial neural networks: a tutorial*, Comput. IEEE March, p. 31-44.

Lee, S., J. H. Ryu and M. J. Lee, 2003, "Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea", *Environmental Geology*, Vol. 44, p. 820-833.

Mayoraz, F. and L. Vulliet, 2002, "Neural networks for Slope Movement Prediction", *The International Journal of Geomechanics,* Vol. 2, Num. 2, p. 153-173.

Pontius, Jr. R.G. and L. C. Schneider, 2001, "Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA", *Agriculture, Ecosystems and Environment*, Vol. 85, p. 239-248.

Schalkoff, R. J., 1997, *Artificial Neural Networks*, McGraw-Hill, New York.

Soeters, R. and C. J. Van Westen, 1996, "Slope instability recognition, analysis, and zonation", in: Turner, A.K. and Shuster, R. L. (Eds), *Landslides Investigation and mitigation*, Transp. Res. Board, Special Report 247, Natural Acad. Press, Washington D.C., p. 129-177.

Sui, D. Z., 1993, "Integrating neural network with GIS for spatial decision making", *The Operational Geographer*, Vol. 11, No. 2, p. 13-20.

Suzen, M.L. and V. Doyuran, 2004, "Data Driven Bivariate Landslide Susceptibility Assessment Using Geographical Information Systems: A Method and Application to Asarsuyu Catchment, Turkey", *Engineering Geology*, Vol 71/3-4, p. 303-321.

Swets, J.A., 1988, "Measuring the Accuracy of Diagnostic Systems", *Science*, Vol. 240, p. 1285-1293.

Tchaban, T., J. P. Griffin, and M. J. Taylor, 1997, The Application of Intelligent Decision Support Systems in Profiling of Retail Outlets. *CiteSeer (Scientific Literature Digital Library)*, http://citeseer.ist.psu.edu/93150.html, 6 pages.

Varnes, D.J. IAEG commission on landslides, 1984, *Landslide hazard zonation: a review of principles and practice*. UNESCO, Paris, No: 3, 63 pages.

Varnes, D. J., 1974, "The logic of geological maps, with reference to their interpretation and use for engineering purposes", *U.S. Geological Survey Professional Paper 837*, 48 pages.

Wang, F. J., 1992, "Incorporating a neural network into GIS for agricultural land suitability analysis", *Proceedings of the GIS/LIS'92 conference*, (1), p. 804-815.

Wei, Y., W. Xu, Y. Fan and H. T. Tasi, 2002, "Artificial neural network based predictive method for flood disaster", *Computers & Industrial Engineering*, Vol. 42, Issues 2-4, p. 383-390.

Werbos, P. J., 1995, "Backpropagation: Basics and new developments", in M.A. Arbib, editor, *Handbook of Brain Theory and Neural Networks*, p. 134-139, MIT Press.

Wieczorek, G. F., 1984, “Preparing a detailed landslide-inventory map for hazard evaluation and reduction”, *Bulletin of the Association of Engineering Geologists*, Vol. 21, No. 3, p. 337-342.

Zhou, W., 1999. “Verification of the Nonparametric Characteristics of Backpropagation Neural Networks for Image Classifications”, *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 37(2), p. 771-779.

An Evaluation of Neural Spatial Interaction Models Based on a Practical Application

Alexandra Akamine and Antônio Nélson Rodrigues da Silva
University of São Paulo, School of Engineering of São Carlos, Brazil

Keywords: Artificial Neural Networks, Spatial Interaction Models, Education Infrastructure.

Abstract: One of the serious problems faced by the Brazilian municipalities is the scarcity of resources for building education infrastructure. This asks for an optimal allocation of the available resources that includes, among other things, a rational spatial arrangement of the supply points (i.e., schools) in order to increase the demand coverage (i.e., students). If it is possible to foresee the regions where the demand is going to be concentrated, it is then possible to plan the location of new facilities and to assess the impact on the future level of service of the entire system. Considering that one of the consequences of the location-allocation process is the distribution of trips from demand points to supply points throughout the city, therefore affecting the overall intraurban accessibility conditions to essential services such as education, there is a strong need of models that planners can rely on to predict the future trip distribution patterns. As a result, the objective of this work was to evaluate the performance of Artificial Neural Networks (ANN) when applied to spatial interaction models, the so-called Neural Spatial Interaction Models. This was done in a practical context, in contrast to the more theoretical works commonly found in literature. The practical application showed that the neural spatial interaction model had different performances when compared to the traditional gravity models. In one case the neural models outperformed the gravity models, while on the other case it was just the opposite. The explanation for this may be in the data or in the ANN model formulation, as discussed in the conclusions.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 19-32.

1. INTRODUCTION

The spatial interaction models were amongst the most studied topics in the field of Transportation Engineering in the second half of the last century. According to Black (1995), research in the area looks essentially for ways to improve the knowledge of the factors influencing trip flows. It also focuses in the development of methods that can help urban and regional planners to forecast the future displacements. There are different approaches for modeling the problem of spatial interaction, such as the intervening opportunities models, and the gravity models. While the use of the former is not very common in practical applications, the latter are largely employed in transportation planning practice and also in theoretical studies. The use of emergent techniques, such as the Artificial Neural Networks (ANN), for modeling the problem has also been tested in the last decade of the 20th century, as can be seen in the works of Openshaw (1993), Black (1995), Fischer, Reismann, et al. (1999), and Fischer and Reismann (2002).

According to Fischer and Reismann (2002), except for their high processing time, the neural models are better than the classical gravity models in terms of general performance. That statement was the starting point of the present study, which is a contribution for a larger project aiming the development of a Spatial Decision Support System for an integrated management of health and education facilities at the local level (Lima, Silva, et al., 2003). The objective of this particular study was to evaluate the performance of Spatial Interaction Models based on Artificial Neural Networks when dealing with different datasets taken from an actual situation. One of the challenges here is to deal with databases that although large are not always necessarily reliable.

This study, which is essentially based on a practical application, was meant to improve the knowledge about temporal and spatial changes of the demand for education infrastructure. That is a key point in the construction of planning scenarios for managing not only the demand but also the supply. In such a way, several alternatives can be tested, such as the reduction of travel distances due to demand relocation or the best locations for opening new educational facilities. Two datasets provided by the Secretary of Education of São Carlos, which is a medium-sized Brazilian city in the state of São Paulo, have been used to test the models performance in practical, real-world conditions. The datasets contain information about children attending day-care centers and elementary schools in two different years, 2000 and 2001.

The data of children from 0 to 6 years-old attending the municipal day-care centers in the year 2000 were initially used for training and validation of the ANN models. Later, these models were used for estimating future trip flows. The estimates were compared to the actual flows observed in 2001 in order to test the generalization capability of the models. The same procedure was carried out with the data of children from 3 to 6 years-old attending the EMEIs, which in Portuguese stands for *Escolas Municipais de Educação Infantil*, or Municipal Schools for Children Education.

The performance of the models was evaluated through a comparison of actual data with the results obtained for the different datasets analyzed with the two modeling approaches: the neural models and the gravity models.

In sections 2 and 3 of this document we discuss some aspects involving the data used in this study and the basic characteristics of the spatial interaction models. The evaluation of the models performance is carried out in section 4, in which the estimates of the gravity models are compared with the estimates of the neural models. Finally, in section 5 are presented the conclusions of the application, followed by the references, in part 6.

2. THE DATA

The data used in the present study show changes in spatial aspects of the demand for municipal educational services in the city of São Carlos throughout two years. The basic information gathered was the home address of the children registered in all public day-care centers and EMEIs and the corresponding locations of these educational facilities in the years 2000 and 2001.

Lima, Naruo, et al. (2001) were able to find the exact location of most children registered in the public educational system of São Carlos in the year 2000 using official data provided by the municipal government. The basic information used in that case was the home address of all students registered in the facilities under analysis. In order to correctly locate the children on a city map, that information was then combined with an address database built by the municipal agency in charge of water distribution and sewage disposal, which has an excellent address recording system based on geographic coordinates of land parcels.

The work started by Lima, Naruo, et al. (2001) was extended for the present study with data of the children attending the educational facilities in 2001. As in the case of the previous work, the data available in the databases used had to be carefully examined before trying to match the records using the address as a common reference. The students' addresses that were not properly typed were then fixed or replaced by the right street names. After that procedure, a Geographic Information System was used to spatially locate the new data. In practical terms, it means that we were able to find each precise address location on a map of the city streets.

That included not only the students but also the facilities they were assigned to, as shown in Figures 1 and 2. While Figure 1 shows the spatial distribution of the children attending day-care centers in the year 2001, Figure 2 displays the information of the demand per EMEI in the year 2001. Later on, GIS tools were also used to calculate the network distances from each demand point to the corresponding day-care center or EMEI.

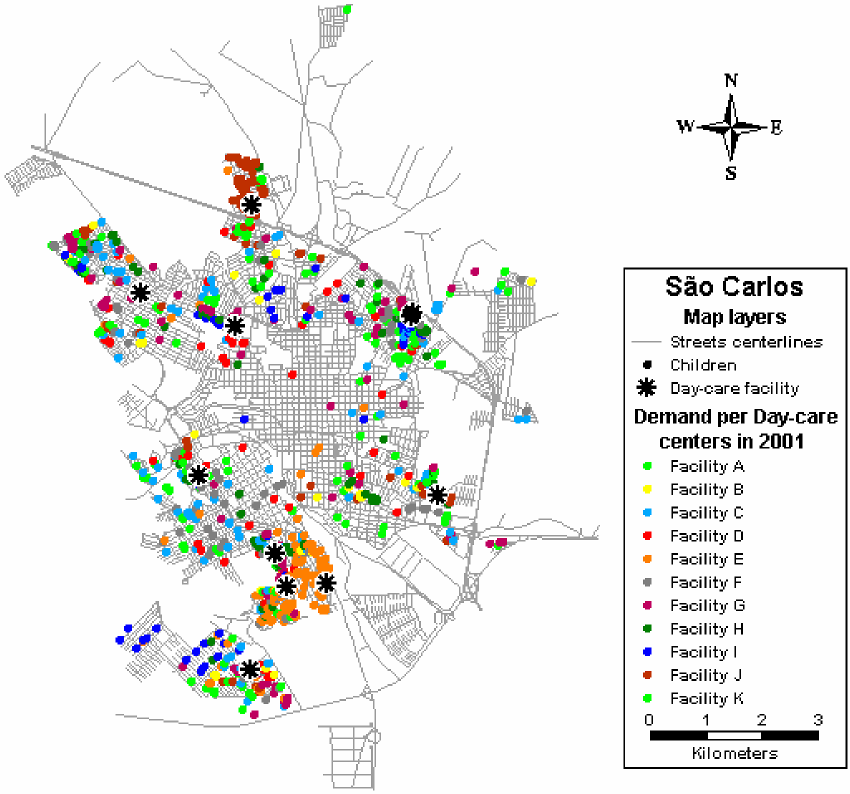


Figure 1. Spatial distribution of public day-care facilities in São Carlos in the year 2001 and respective demand

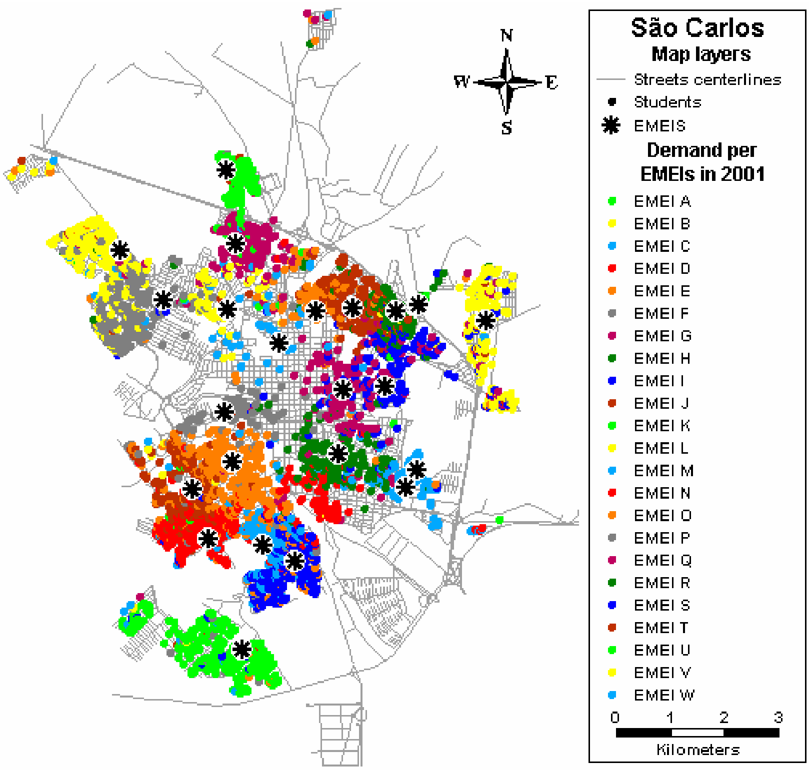


Figure 2. Spatial distribution of EMEIs in São Carlos in the year 2001 and respective demand

3. THE SPATIAL INTERACTION MODELS

Once the demand and supply points were located on the map, the identification of the flows among them was straightforward. The demand points were initially aggregated according to the limits of the 245 census tracts (CT) in which the city has been divided in the year 2000 by the Brazilian census bureau. The centroids of those CT were then assumed as the origin points of all displacements in each particular area. The same procedure was carried out for day-care centers and EMEIs in the years 2000 and 2001. Next, the number of trips attracted to each facility, the total number of trips produced in each CT, and the travel distances from all centroids to all facilities (day-care centers or EMEIs), was obtained. The location of all demand (centroids) and supply points is shown in Figure 3 for the case of day-care centers. An overview of the data is presented in Table 1.

Table 1. Summary of the data applied in the spatial interaction models.

Facilities	Year	Supply points	Demand points	O/D pairs
Day-care centers	2000	10	245	2450
	2001	11	245	2695
EMEIs	2000	22	245	5390
	2001	23	245	5635

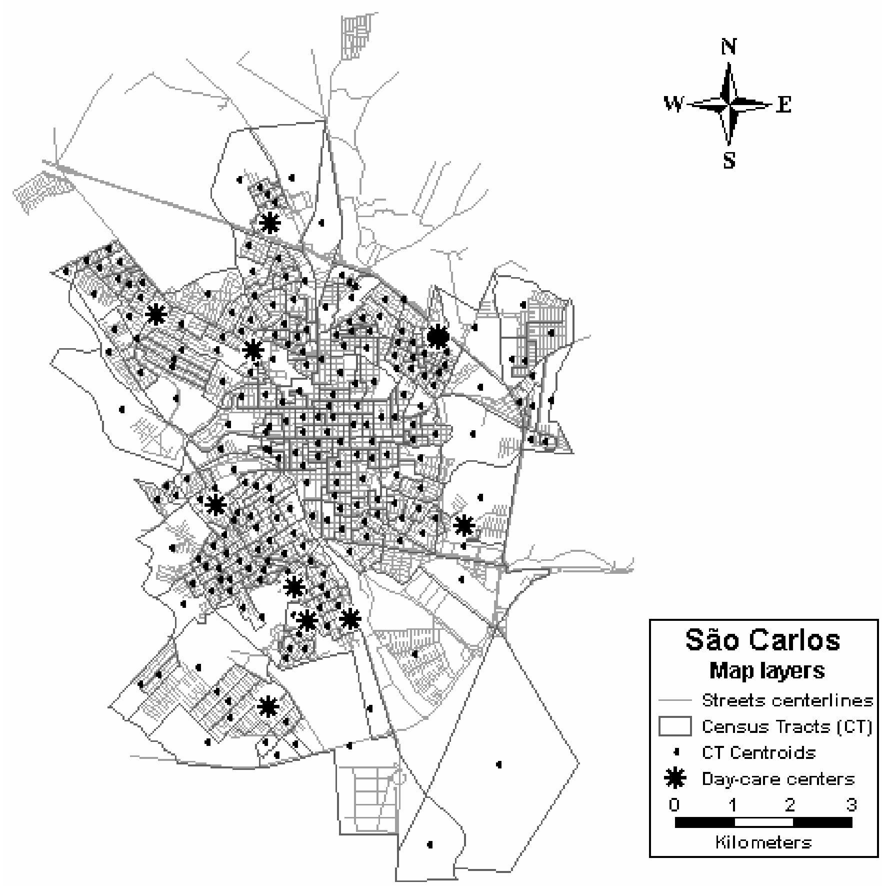


Figure 3. Demand (CT centroids) and supply points (public day-care centers)

The data of the year 2000 were randomly split in three data subsets for each of the two different kinds of facilities. While the first subset, which had fifty percent of the records, was used for training the neural networks, the second and third subsets, both with twenty-five percent of the records, were respectively used for validation and query. The absolute number of records in all subsets is shown in Table 2. Also shown in Table 2 is the total number of records of the year 2001. In order to test the generalization capability of the models, the total number of trips associated to the actual origins and destinations taken from those records were applied as input data in the trained networks for estimating the 2001 flows.

Table 2. Absolute number of records in the subsets used for training, validation, and query in the case of public day-care centers and EMEIs

Facilities	Year	Training	Validation	Query	Total
Day-care centers	2000	1225	613	612	2450
	2001	-	-	2695	2695
EMEIs	2000	2695	1348	1347	5390
	2001	-	-	5635	5635

Before introducing the data in the ANN, they were normalized as shown in Equations (1) and (2). The normalization interval was between 0.1 and 0.9 to avoid problems in the calculation (Equation (3)) of the mean relative error (MRE) if dealing with actual values equal to zero.

For distances: $$dist_{norm.} = \frac{dist}{\max(dist)} \quad (1)$$

For production, attractions, and flows (as in Bocanegra, 2002):

$$Y_i = \left[\frac{(X_i - X_{\min}) * (Y_{\max} - Y_{\min})}{X_{\max} - X_{mín}} \right] + Y_{\min} \quad (2)$$

where: Y_i = normalized value;
X_i = actual value;
X_{min} = minimum actual value;
X_{max} = maximum actual value;
Y_{min} = minimum normalized value (0.1);
Y_{max} = minimum normalized value (0.9).

$$MRE = \frac{1}{n} \sum_{t=1}^{n} \frac{abs(actual_t - estimated_t)}{actual_t} \quad (3)$$

After the pre-processing phase the data were introduced in the software *EasyNN-plus*, which simplifies many of the steps needed for creating simple and efficient neural network models. The software can be used to build Multilayer Perceptron networks with up to three hidden layers, although the developer himself states that most real-world problems can be solved with one or two hidden layers (Wolstenholme, 2002). The *EasyNN-plus* package uses a backpropagation algorithm and a sigmoidal function to build the models. The data needed for training the network can be generated with simple text or spreadsheet software. In addition, the program can either assume values for the learning rate and momentum or let it up to the user. After the models are built, they can be used for estimating output values. The comparison of these values with the actual values making use of a performance measure (such as the MRE) makes possible to select, among the many alternatives built, the ANN model that produce the best estimates of the actual values.

4. EVALUATION OF THE MODELS

In order to evaluate the performance of the neural spatial interaction models, we have calculated the mean relative error (MRE) for each network configuration tested. The results are presented and discussed in this section. Starting with the parameters suggested by the package *EasyNN-plus* we have built networks with one or two hidden layers, with different number of nodes in each of the hidden layers, and with distinct learning rate (L) and momentum (M) values. The schemes shown in Figures 4 and 5 have the values of L and M grouped according to the distinct network topologies tested. In those Figures, the average and standard deviation of MRE values for the three groups of data randomly selected as training and validation subsets are also presented for each network tested, along with the results of the query data in the right-hand boxes. The latter are characterizing the generalization capability of the different ANN models.

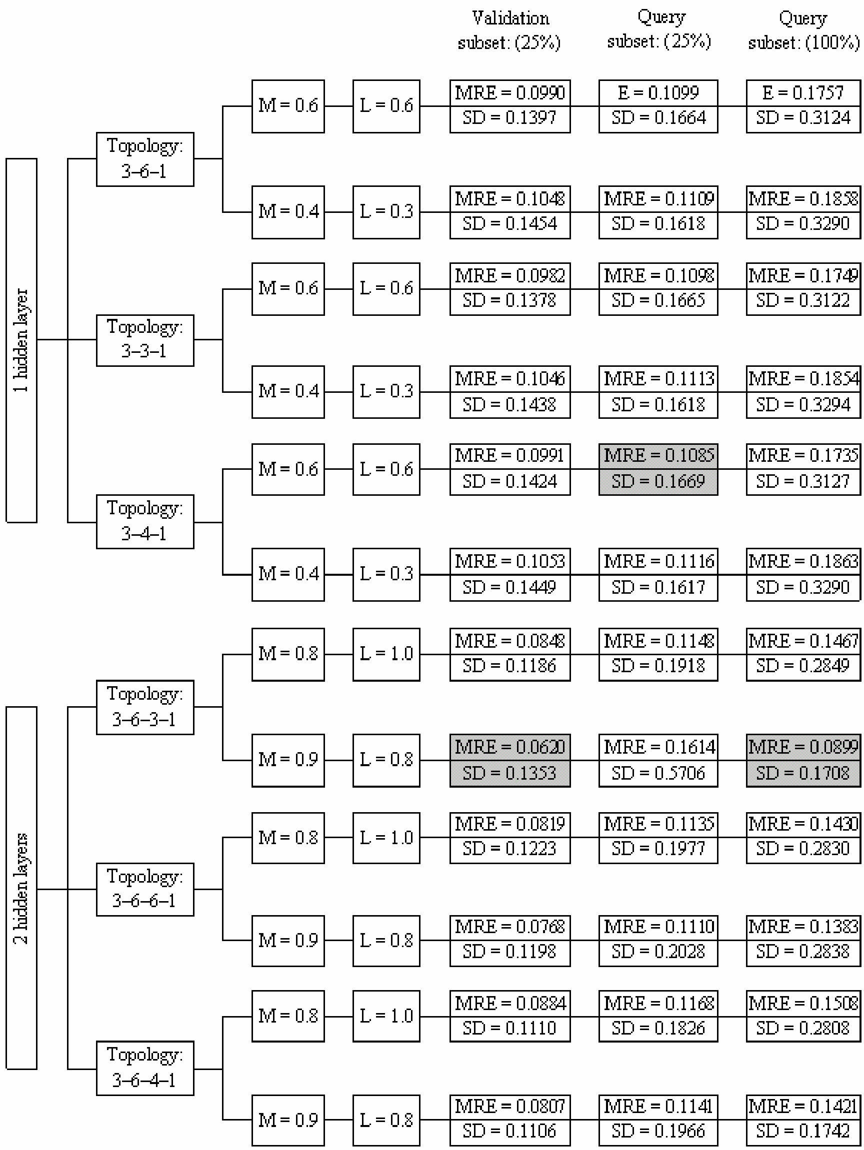


Figure 4. Performance of the neural spatial interaction models when applied to the case of day-care centers

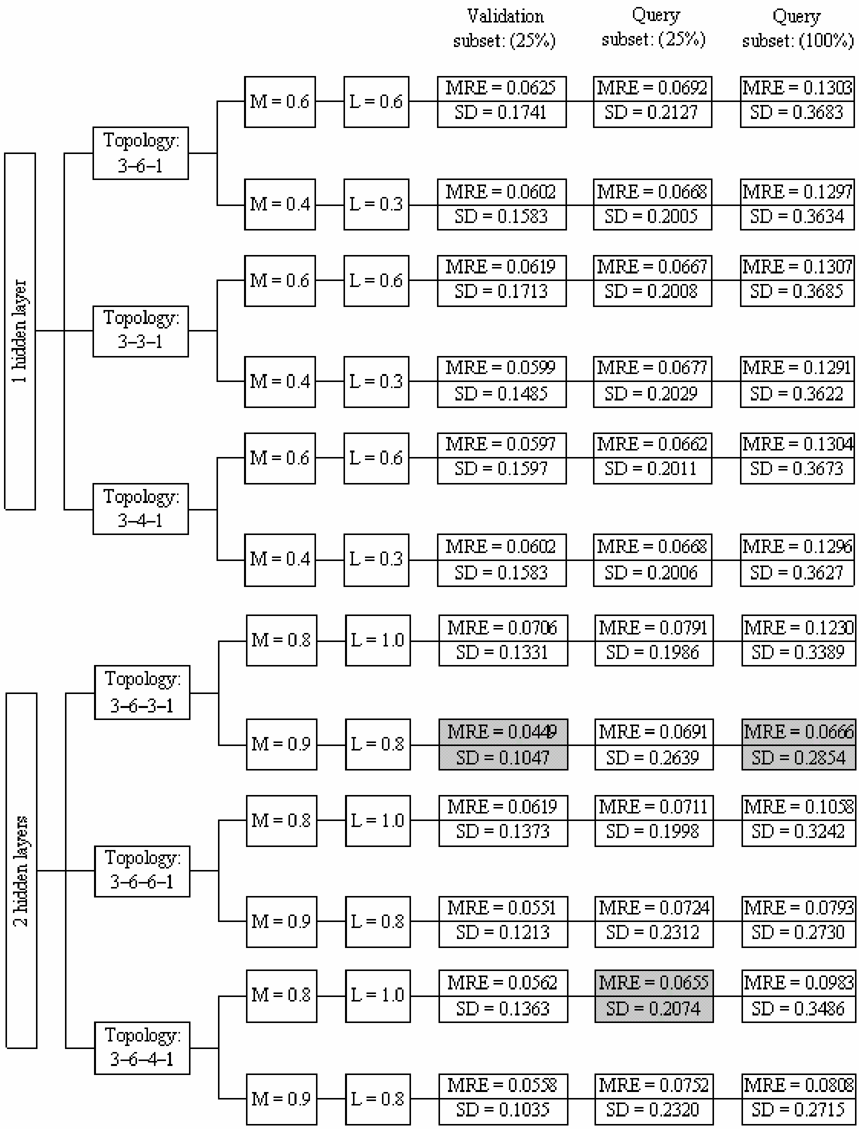


Figure 5. Performance of the neural spatial interaction models when applied to the case of EMEIs

4.1 Analysis of the results obtained with neural spatial interaction models and with gravity models

The data of the year 2000 were also used in both cases (i.e., day-care centers and EMEIs) to calibrate traditional doubly constrained gravity models, which were subsequently used for estimating the flows in the year 2001. Again, the average and standard deviation of MRE values were calculated in order to evaluate the models' performance.

Table 3. Results obtained with the gravity models and with the best neural spatial interaction models

Facilities	Model	MRE	
		Average	SD
Day-care centers	Doubly constrained gravity model	0.1353	0.3263
	Neural spatial interaction model	0.0899	0.1708
EMEIs	Doubly constrained gravity model	0.0456	0.1484
	Neural spatial interaction model	0.0666	0.2854

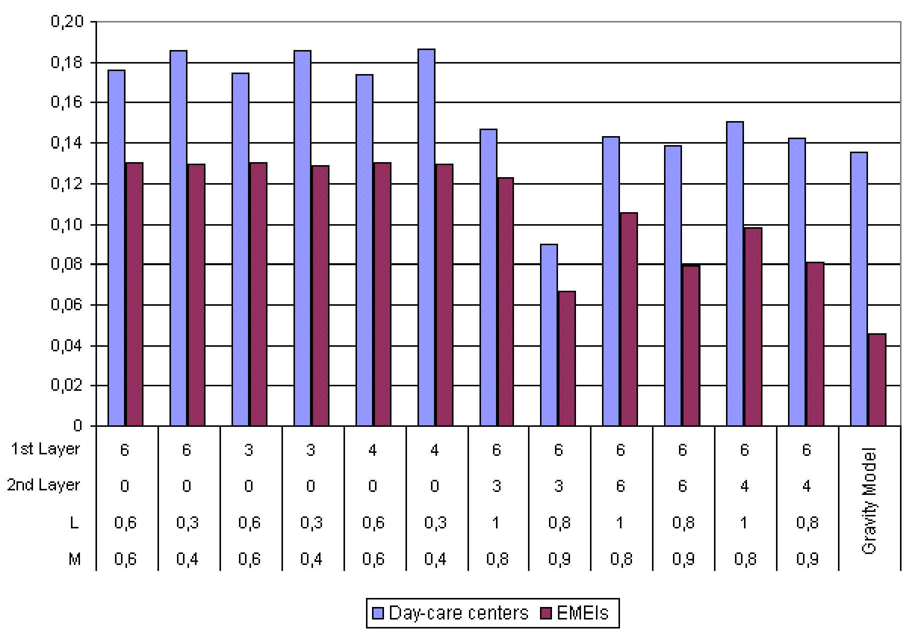


Figure 6. MRE values obtained with the neural spatial interaction models and with the gravity models

They are shown in Table 3, along with the best values obtained for the neural spatial interaction models. A visual comparison of the average MRE values of all models can be done in Figure 6. It is important to highlight that

we were referring in that case to the last column of Figures 4 and 5 (i.e., the 100% query subset), because of the focus of our analysis is on the generalization capability of the models.
The results presented in Table 3 and in Figure 6 showed that one configuration of the neural spatial interaction models had clearly outperformed the gravity model in the case of the day-care centers. In the case of the EMEIs, however, the best results were obtained with the gravity models. It is interesting to notice the fact that in all cases the models were more accurate when estimating the flows to EMEIs than to day-care centers.

5. CONCLUSIONS

The comparative analysis between the results obtained with the neural spatial interaction models and the doubly constrained gravity models for the case of day-care centers showed the superiority of the former, which produced more accurate estimates than the latter, although only in one configuration. In contrast, in the case of EMEIs the gravity models had the best performance. In addition, all models were more accurate in predicting the flows to EMEIs than to day-care centers. The explanation for that may be in the data or in the ANN model formulation, as discussed next.

The spatial distribution of the children in the case of day-care centers was quite irregular, while in the case of EMEIs they were clearly clustered around the facilities they go to, as shown in Figures 1 and 2. The likely explanation for the somehow unexpected spatial distribution pattern of the day-care center attendees (Figure 1) may be in the fact that the facility choice is more strongly influenced by the parents' work location than by their home location. As work location data is not available to the models, which rely only on the total trips produced in the CTs and attracted to the facilities, and on the distances between the points of demand (i.e., home locations grouped in the CT centroids) and supply (i.e., facility locations), this negatively impacts the models' performance. That is not the case of the EMEIs, however, where the students' home location is clearly connected to the school chosen (Figure 2). The regular spatial distribution of the children around the facilities made the prediction task easier for both model types, as indicated by the MRE results shown in Table 3 and Figure 6.

It is interesting to observe that the neural spatial interaction models were remarkably able to capture, although only in one case, the irregular spatial distribution pattern of the children in the case of day-care centers therefore producing better estimates than the gravity models. What is not clear, however, is their performance in the case of the EMEIs. Although better than the performance in the case of the day-care centers it was worse than the performance of the gravity model (Figure 6). Even though, the results of the neural spatial interaction models were not bad. Thus, the case discussed here stressed the promising role that those models can play in practical applications of education, and to a certain extent, health facilities management. The quality of their predictions is crucial in the evaluation of future scenarios of demand and supply spatial distribution aiming the reduction of travel distances obtained either by the demand relocation or by the creation of new facilities.

What seems to be an important conclusion of this study is the fact that certain ANN model configurations can outperform the gravity models. That was once observed in the case of day-care centers despite the irregular distribution of demand points, and it is probably the case with the EMEIs if more ANN configurations were tried. The challenge here lies on developing and applying in real-world conditions efficient methods to select the ANN configuration that better models the problem, such as Genetic Algorithms or the bootstrapping approach suggested by Fischer and Reismann (2002). In addition, given the results obtained in the case of the EMEIs, a point that deserves further investigation is the assumption that those models could be improved by any additional input data that also has influence on the facility selection process, such as the school attractiveness measures used by Almeida and Gonçalves (2001). Considering the difficulty for obtaining this sort of data, from a practical standpoint this may have, however, more costs than benefits. But it is certainly worth investigating. Moreover, the statement that both models performed better with the day-care centers data than they did with the EMEI's data is not necessarily true. It could be that the EMEI's data simply has a smaller variance, and so errors seem smaller when amalgamated using the MRE formula. Therefore, other performance measures should be also tested in the future.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Brazilian agencies CAPES (Post-Graduate Federal Agency), FAPESP (Foundation for the Promotion of Science of the State of São Paulo), and CNPq (Brazilian National Council for Scientific and Technological Development), which have been supporting our efforts for the development of this work in different ways and periods. We also acknowledge the support of the Municipality of São Carlos, which have provided the data used in the present study. Finally, we thank the anonymous referees for their valuable comments.

REFERENCES

Almeida, L.M.W. and M.B. Gonçalves, 2001, "A methodology to incorporate behavioral aspects in trip-distribution models with an application to estimate student flow", *Environment and Planning A.* 33, p.1125-1138.

Black, W., 1995, "Spatial interection modeling using artificial neural networks", *Journal of Transport Geography,* 3(3), p. 159-166.

Bocanegra, C.W.R., 2002, *Alternative procedures to make more effective the application of artificial neural networks in transportation planning*, M.S. Thesis, University of São Paulo, School of Engineering of São Carlos, São Carlos, SP, Brazil.

Fischer, M.M. and M. Reismann, 2002, "A Methodology for Neural Spatial Interaction Modeling", *Geographical Analysis*, 34(3), p. 207-228.

Fischer, M.M., M. Reismann, and K. Hlavácková-Schindler, 1999, "Parameter estimation in neural spatial interaction modelling by a derivative free global optimization method", *International Conference on GeoComputation, 4, Fredericksburg, Virginia, USA*.

Lima, R.S., A.N.R. Silva, and J.F.G. Mendes, 2003, "A SDSS for integrated management of health and education facilities at the local level: challenges and opportunities in a developing country", *International Conference on Computers in Urban Planning and Urban Management, 8, Sendai, Japan*, Proceedings... (CD-ROM). Sendai, Center for Northeast Asian Studies, Tohoku University.

Lima, R.S.; M.K. Naruo, R.J. Rorato, and A.N.R.Silva, 2001, "Influência da desagragação espacial da demanda por educação no cálculo das distâncias de deslocamento em uma cidade média", *Congresso Brasileiro de Cartografia, 20, Conferência Ibero-Americana de SIG, 8, Porto Alegre*, (CD-ROM). Rio de Janeiro, Sociedade Brasileira de Cartografia.

Openshaw, S., 1993, "Modeling spatial interaction using a neural net", in: Fisher and Nijkamp (eds.), *Geographic Information Systems, spatial modeling and policy evaluation, Springer-Verlag, Berlin*, p. 147-164.

Wolstenholme, S., 2002 Neural Network Software for Windows, EasyNN-plus. (http://www.easynn.com in 31/10/2002).

Improved Understanding of Urban Sprawl Using Neural Networks

Lidia Diappi, Paola Bolchim, and Massimo Buscema[1]
Department of Architecture and Planning – Polytechnic of Milan, Italy
[1] Semeion Centre – Rome, Italy

Keywords: Neural Networks, Self-Organizing Maps, Land-Use Dynamics, Supervised Networks.

Abstract: It is widely accepted that the spatial pattern of settlements is a crucial factor affecting quality of life and environmental sustainability, but few recent studies have attempted to examine the phenomenon of sprawl by modelling the process rather than adopting a descriptive approach. The issue was partly addressed by models of land use and transportation which were mainly developed in the UK and US in the 1970s and 1980s, but the major advances were made in the area of modelling transportation, while very little was achieved in the area of spatial and temporal land use. Models of land use and transportation are well-established tools, based on explicit, exogenously-formulated rules within a theoretical framework. The new approaches of artificial intelligence, and in particular, systems involving parallel processing, (Neural Networks, Cellular Automata and Multi-Agent Systems) defined by the expression "Neurocomputing", allow problems to be approached in the reverse, bottom-up, direction by discovering rules, relationships and scenarios from a database. In this article we examine the hypothesis that territorial micro-transformations occur according to a local logic, i.e. according to use, accessibility, the presence of services and conditions of centrality, periphericity or isolation of each territorial "cell" relative to its surroundings. The prediction capabilities of different architectures of supervised Neural networks are implemented to the south Metropolitan area of Milan at two different temporal thresholds and discussed. Starting from data on land use in 1980 and 1994 and by subdividing the area into square cells on an orthogonal grid, the model produces a spatial and functional map of urbanisation in 2008. An implementation of the SOM (Self Organizing Map) processing to the Data Base allows the typologies of transformation to be identified, i.e. the classes of area which are transformed in the same way and which give rise to territorial morphologies; this is an interesting by-product of the approach.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 33-49.

1. INTRODUCTION

Land use dynamics and fragmentation of settlements is a crucial question for planning. In the general framework of sustainability objectives, the policies that control a suitable process of urbanisation increasingly involve a deep knowledge of complex criteria of location chosen by the different agents. Planners realize that it is crucial to understand and provide the best possible explanation for the observed spatial distribution of urban activities.

Principles and technologies of Artificial Intelligence (AI) in general, and of Neural Networks (NN) and Cellular Automata (CA) in particular, offer the potential to increase the knowledge in urban dynamics by multiplying the information capacity of the GIS and by offering a new approach to territorial modelling. Most geocomputation currently deals with models on spatio-temporal dynamics in urban land-use and morphogenesis.

Among them, some applications, mainly based on CA, have opened more promising directions for the goal: Clarke, Hoppen and Gaydos (1997) modelled the historical development of San Francisco area; (Batty, Xie and Sun 1999; Wu 1998) built several urban models and in particular a model on the residential development in the fringe of Buffalo; Portugali, Benenson and Omer (1994; 1997) have focused their research on models of socio-spatial segregation; the many contributions of Engelen, Ulje and White (White and Engelen 2000) have produced several CA based models with integration of several economic theories.

CA appear to be the most attractive and favoured technique for implementing high resolution models of spatial dynamics for a number of reasons:

- They are inherently spatial; their definition on a raster of cells, and on neighbouring relationships are crucial;
- They are simple and computationally efficient;
- They are process models that deal with state changes;
- They are dynamic and can then represent a wide range of situations and processes.

It is worthwhile to note that, in most of the models carried out until now, CA are based on explicit spatial rules which allow the simulation of different dynamic behaviours on the basis of a "trial and error" procedure.

However this condition, the explicit and exogenous formulation of assumptions, represents the greatest limit of this approach, since it reduces the variability of the different territorial contexts on the basis of few theoretical principia (spatial interaction, diffusion processes and so on), inhibiting the discovery and appearance of new features in urban dynamics. Given the complexity and variability of the location behaviours it appears important to learn from reality the most relevant factors affecting the single location with respect to the surrounding conditions.

Recent developments in the natural algorithms, and particularly in Neural Networks, allow the reversal of the approach by learning the rules and behaviours directly from the Data Base, following an inductive bottom-up process.

The aim of this paper is therefore to present an integrated approach on land use dynamics where the transition rules of urban spatial evolution are learnt by NN. The proposed innovation concerns the heart of the CA itself: the growth rules searching and identification.

In the paper the potentialities of NN are experimented with two different architectures: SOM (Self Organizing Maps), (Kohonen 1995) and a set of Supervised NN (Semeion 1998).

SOM permit the investigation of different dynamic behaviors by showing the strengths of the underpinning relationships with the environment. The classification produced by SOM identifies the most relevant clusters of cells for transition rules in quantitative and qualitative terms.

Secondly, for forecasting purposes, a set of Supervised NN is applied to learn the transition rules and to produce a possible future scenario of urbanization.

The case study is the south metropolitan area of Milan, whose extension is approximately 675 Km^2, and which is a rich agricultural area with few historical small centers. The area is under pressure from the spillover, in fragmented residential and manufacturing settlements, of Milan.

The paper is organized as follow:

The second section presents a short overview on NN and their potentialities in urban analysis and forecast; the third section sketches a brief description of the study area and the GIS used. The methodology is explained in the following fourth part which describes the research path.

Section 5 shows the NN SOM implementation results. The implementation of different architectures of Supervised NN is presented in sections 6-8. The input data and the methodology in section 6; the learning and validation phase in section 7 and the results obtained in prediction in section 8.

Some final comments and perspectives on the adopted approach conclude the paper in section 9.

2. NEURAL NETWORKS

With the development of ANN (Artificial NN), which are AI based information processing systems, in recent years new opportunities have emerged to enhance the tools we use to process spatial Data. Their specific advantage lies not only in the enhancement of speed and efficiency in handling urban Data, but specifically in providing a tool to develop new theories and techniques. While the traditional modelling approach is based on explicit a priori rules formulation, through the Neural Networks processing approach, rules are found a posteriori on the base of a learning process of a distributed “unit processing” architecture.

NN structure consists in a set of adaptive processing elements (nodes) and a set of unidirectional data connections (weights).

The most successful applications in territorial Analysis and Planning rely on pattern classification, clustering or categorisation, optimisation (Openshaw and Abrahart 2000; Reggiani 2000; Fischer and Leung 2001), modelling scenic beauty from extracted landscape attributes (Bishop 1994), suitability analysis for development (Sui 1992; Deadman and Gimblett, 1995).

The novelty of our approach lies in the use of NN as a powerful tool for prediction and for building virtual scenarios on urbanisation processes. The results have been achieved through different categories of “training regimen” able to react to different information environment.

The training processes can be divided into three basic categories: monitored training, supervised training, and self-organisation. The monitored training is typical of associative networks, which are NN with essentially a single functional layer that associates one set of vectors x_1, x_2, …x_n with another set of vectors y_1, y_2, …y_n. The primary classification of ANN are into feedforward and recurrent classes.

In a feedforward associative network the x-vector input to the single functional layer of the processing elements leads to the y’-vector output in a single feedforward pass. In a recurrent associative network the output signals of the processing elements of the layer are connected to those same processing elements as input signals. The output vector y’ is ignored until the system converges (that is stabilizes and ceases to change significantly).

Another categorisation of ANN is into autoassociative NN, if y vectors are assumed to be equal to the corresponding x vectors. In a Heteroassociative network $y_i \neq x_i$.

There are many algorithms and procedures to optimize the weight matrix during the learning phase and many algorithms for dynamically querying of the ANN already trained.

In this research we used a Recirculation Neural Network (RCNN) (Hinton and McLelland, 1988). The ANN have been shown to be highly efficient in determining the fuzzy similarities among different Records in any Data Base (DB) and the relationships of gradual solidarity and gradual incompatibility among the different Variables. The ability of ANN to produce *prototypical* generators, to discover *ethnotypologies* and to simulate possible scenarios was already experimented by the authors to investigate the complex structure of urban sustainability in Italian cities (Diappi, Buscema, Ottanà, 1998).

Supervised training implies a regimen in which the NN is supplied with a sequence of examples (X_1, Y_1), (X_2, Y_2)… (X_k, Y_k).. of desirable or correct input/output pairs. As each input X_k is entered into the NN, the "correct output" Y_k is also supplied to the network. In our study the input is given by the territory information at time t and the "correct output" is the corresponding information at time t+1. Once the NN is trained and has learned the rules of transition, it will be able to produce the "desired" land use transformation of the present state of territorial system supplied as Input to the NN.

In self-organizing training, a network modifies itself in response to X Inputs. This category of training is able to obtain a surprisingly high number of information processing capabilities: development of pattern categories based on clustering, estimation of probability density functions, development of continuous topological mapping from Euclidean space to curved manifolds (Hecht-Nielsen 1990). Self-organizing training includes the Self-Organizing Map (SOM), presented in section 5.

SOM is able to develop a continuous topological mapping $f: B \subset R^n \rightarrow C \subset R^m$ by means of self-organization driven by Y examples in C, where B is a rectangular subset of n-dimensional Euclidean space and C is a bounded sub set of m-dimensional Euclidean space, upon which a probability density function ρ (Y) is defined. In the paper their ability to classify has been used to define the prototypical land use change rules in the case study.

3. THE CASE STUDY, THE DATA AND THE GIS

The southern ring of the metropolitan area of Milan presents large extensions of tilled land and natural parks with rare urban centres historically grown on agricultural activities. More recently, in the 70's the area has undergone a rapid urbanization process, principally produced by spill-over effects from the city of Milan.

The scattered and dispersed form of both residential and industrial new settlements is rapidly producing a high land consumption which is compromising the productivity of one of the richer agricultural areas in Europe. The forecast of urban sprawl is therefore a crucial issue which increases the scientific interest to test a new approach in urban modelling.

The available GIS on the area includes the land use coverage only at two temporal thresholds: 1980 and 1994. Even if this is an evident limit, it should be considered that urban sprawl in the area is a quite recent phenomenon whose interpretation and description would be biased if handling a longer sequence of temporal data.

The model uses a regular square grid of 500 m with 2703 cells in total. The land uses taken into consideration are the following: residential, commercial, industrial and "green" or unbuilt land, which in the specific context denotes mainly rural areas.

The construction of the GIS has been based, for each period, on the superposition of two maps: the Lombardy Regional Technical raster Map and the Master Plan Mosaic of the Municipalities of the Province of Milan ("Mosaico Informatizzato degli Strumenti URbanistici Comunali", Provincia di Milano – Settore Pianificazione Territoriale - SIT) showing the different land uses.

The partition of each land use has been executed by ArchView. Then the superposition of the grid has allowed to measure the quantity of land use for each cell.

In this study the information given to the NN has the same structure as a CA. The following information for each cell are supplied to the Neural Network:

- Land use of the cell i at time t (1980);
- Land use of the neighbouring cells at time t (1980);
- Land use of the cell i at time t+1 (1994).

The neighbourhood size is based on the 8 contiguous cells.

The cell and the neighbourhood state are described in terms of *share* for each land use with respect to the total surface of the unit or of the whole neighbourhood. Only the three urbanized functions (residence, industry, commerce) have been processed since the unbuilt cell or neighbouring share is a linear combination of the other three.

4. THE METHODOLOGY

The initial idea was to test the approach in a "toy" example, based on a small urbanisation process produced by a CA evolving with explicitly given rules. The small toy was implemented with different neighbourhood sizes. On the resulting pattern at different time steps an Associative NN has been implemented, with the aim to test to what extent the NN should be able to understand the imposed rules. The results of the experiment were successful: the NN was able to understand the CA rules, and, moreover, the test has produced relevant information on the sensitivity of the NN to the Data (Bolchi, Diappi, Franzini, 2001).

However the same Associative NN, applied to the real Data Base of the south of Milan, produced very poor results. The NN didn't understood many rules of change and the scenario, generated in the querying phase, depicted a quite static situation where even the residential growth, quite numerous and relevant in the period, were much lower than expected.

Then, with an implementation of the SOM (Self Organizing Map), a NN able to classify the pattern in fuzzy clusters of land use dynamics and to produce their prototypical profiles, we tried to find out the rules of change. These profiles, called *codebooks* show the different "average" activation levels of the variables (nodes) of the records in the class, allowing to discover the underpinning relationships among the "average" initial and final state of the group of cells.

Finally, a set of supervised NN had been implemented for forecasting purposes. The approach was reversed: the state at *t* and *t+1* of cell becoming urbanized during the observed time lag represents a "model" which other cells will follow during the time lag *t+1*, *t+2*.

5. THE CLASSIFICATION OF THE LAND USE DYNAMICS WITH SOM

The NN SOM, a powerful tool of classification, have been developed mainly by Kohonen (1995) between 1979 and 1982. As said before SOM are Self Organised NN, where the target is not predefined, but dynamically built up during the learning phase. Their architecture comprises two layers: an input one, acting simply as a buffer, that doesn't modify the data, and an output one, known as Kohonen layer (or matrix), which is formed by units regularly organized in the space and which evolves during the training following a spatial organization process of the data characteristics: the Feature Mapping (Fig. 1*a*). The construction of these maps allows a close examination of the relationships between the items in the training set.

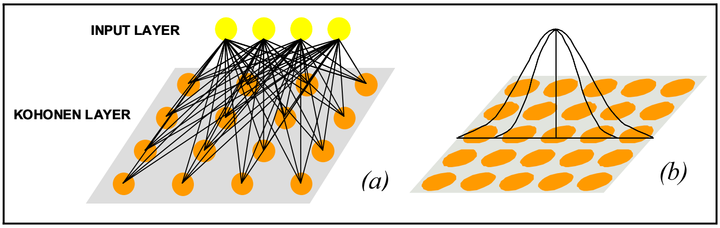


Figure 1. The SOM topology *(a)* and the weight update function *(b)*

When the training phase has calculated the weight matrix, the map shows clusters where each input vector is assigned to the closest codebook in terms of minimum Euclidean distance.

The SOM attitude to “classify” makes it possible to perform a *mapping* with two main peculiarities:

- Clustering: the net performs a logical division of the input space into regions (cluster), associating a point in the N-dimensional input space to the two-dimensional output matrix. In the dimension reduction process the principal components discriminating data are dominant.
- Self-organisation: before the training the weights vectors topology depends only on the initialising criterion: if it is random, weights will be casually organised into their hyper-cube. The learning criterion tends to move the weights vectors towards the input vectors seen during the training. The vector moving affects not only the winner unit vector, but also its neighbourhood according to a decreasing function (fig. 1b).

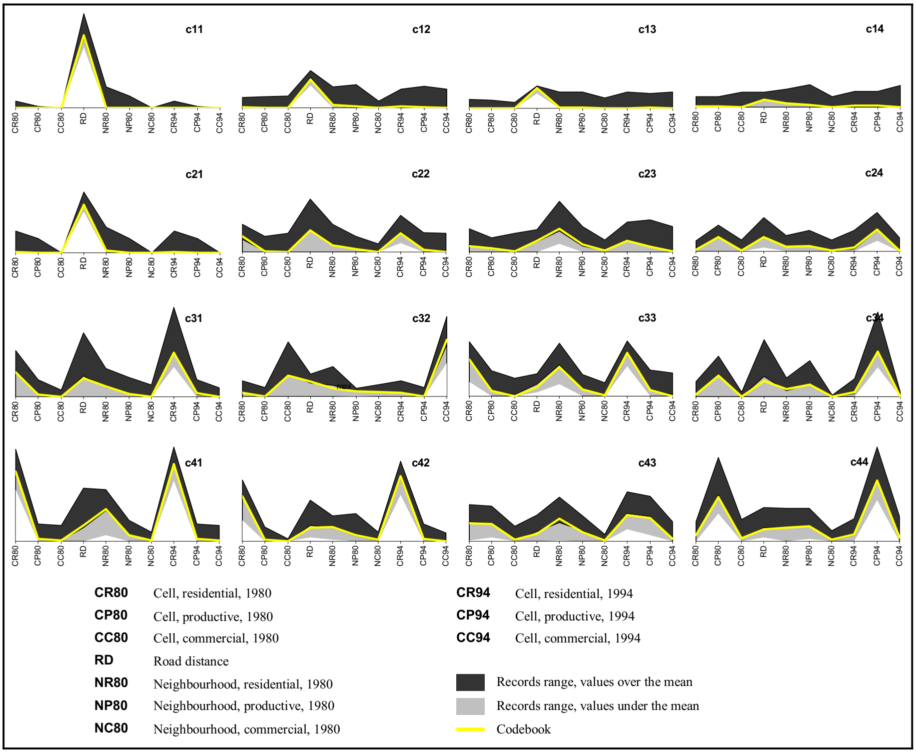


Figure 2. The cluster profiles and the codebooks

SOM NN on square grids of 9,16, 25 nodes have been trained to group the data. The best explaining output was obtained with a 4x4 nodes grid; this result was the best in sufficiently differentiating each group from the other group.

The spatio-temporal classification carried out by SOM has been displayed by:

- cluster profiles and their codebook;
- charts with a colour hatched plot of the zones, based on output units assignment.

Figure 2 shows the clusters and the codebooks; the variables are on the x-axis, and their activation level is on the y-axis. The envelope of the records assigned to each single cluster is charted in grey and black and the codebook is the yellow line.

Figure 3 shows the spatial allocation of the clustered cells; it is crucial to know if cells with similar dynamic behaviour are also spatially contiguous, or are scattered in the territory.

In Fig.2, the first row (C 1-1 ÷ C 1-4) shows groups of stable (in the period) agricultural areas; however moving right along the first row, if no significant different land uses are taking place, the mean distance to the roads drops showing potential "risk" of urbanisation. In Fig. 3, cells of the group C 1-3 and C1-4 are close to urbanised areas.

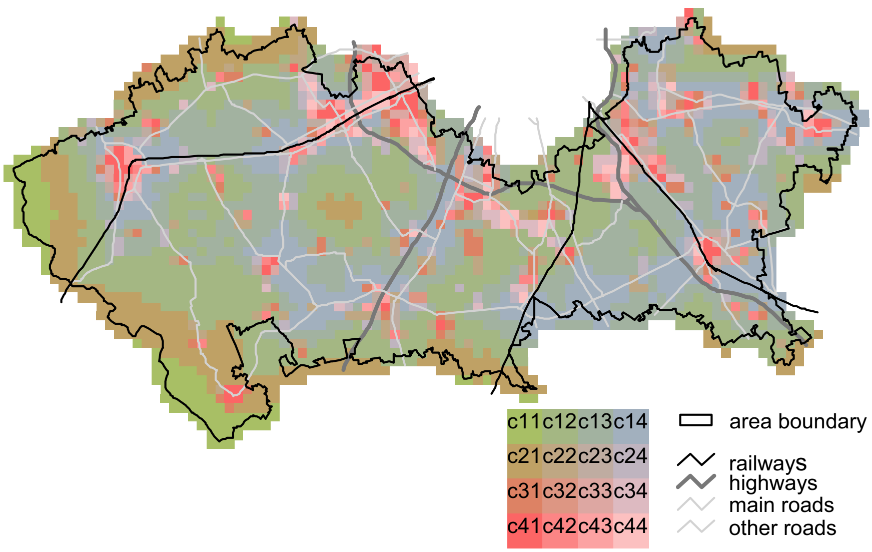


Figure 3. The spatial allocation of the SOM clusters

Shifting from C 1-1 along the column increasing road equipment copes with new residential settlements, which, in C 4-1 infill the consolidated urban centres.

In the fourth column the industrial land use dynamics emerges both in less infrastructured and isolated areas (C 3-4) and near the exiting ones (C 4-4). Looking in the map (Fig. 3) it is worthwhile to note that industrial settlements tend to aggregate spatially, near or far from the urban centres, while road accessibility is not an essential prerequisite for them.

In the fourth row the infilling processes in existing urban areas are represented: from the residential growth in C 4-1 and to the expansion near the existing industrial areas (C 4-4); between the two groups C 4-2 shows peripheral residential growth near industrial areas and C 4-3 classifies the emergence of new linear forms of urbanisation with land use mix along the main roads. The spatial logic of commercial activities is shown in C 3-2 where, as expected, a concentration process near the most important urban centres is taking place.

In conclusion, the adoption of the SOM as a tool to investigate the different dynamics seems fruitful and opens new research directions. Indeed, the different codebooks may be interpreted in a "if then else" approach: given these surrounding conditions at time *t* at time *t+1* the dynamics will change in this way.

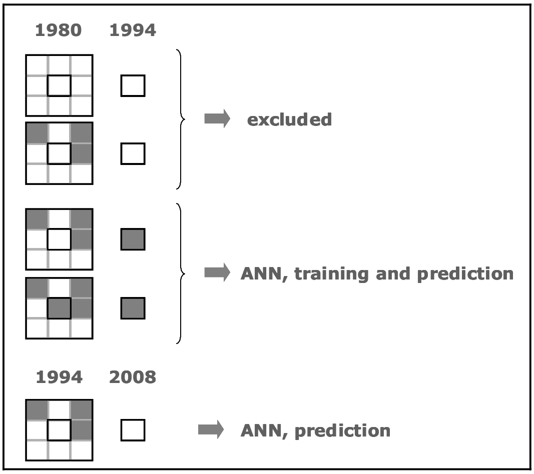


Figure 4. The record set used for the different processing phases

6. THE PREDICTION THROUGH SUPERVISED ANN

The NN which generate prediction on land use dynamics are Supervised (SANN). Trained on the basis of Input/Output examples the SANN is able to reproduce the expected Output starting from the same Input. This means that the network should learn, from the set of cells which change their land use in the time lag considered, the connections between the final state at time *t+1* (the target) and the local and neighbouring conditions at time *t*.

The record supplied to the SANN contains 6 input variables and 3 output ones. The input ones describe the state of the cell and of its neighbourhood at the time *t*, the output variables represent only the cell state at the time *t+1*.

For the implementation of SANN, it is crucial to select "good examples" to feed the network. Therefore the records have been split into three different sets (Fig. 4):

- The first one, composed by 1662 records of "stable green" cells at both the times, has been excluded from NN processing;
- The second one, containing 1041 records for the cells urbanised at one or at both times, has been used for training and prediction;
- The third one, concerning green cells with urbanised neighbourhood at 1994, has been used only for prediction.

Table 1. The different architectures of SANN

Set 1			Set 2		
Topology	Order	Learning Law	Topology	Order	Learning Law
FF		Bm	FF		Bm
FF		Bp	FF		Bp
FF		Sn			
Self	DA	Bp	Self	DA	Bp
			Self	DA	Bm
			Self	SA	Bm
			Self	SA	Bp
Tasm	DA	Bm	Tasm	DA	Bm
Tasm	DA	Bp	Tasm	DA	Bp
Tasm	SA	Bm	Tasm	SA	Bm
Tasm	SA	Bp	Tasm	SA	Bp
Tasm	SA	Cm			
Tasm	SA	Sn			

Learning Law:	Bp = Back Propagation (standard)
	Sn = Sine Net (Semeion)
	Bm = Bi-Modal Network (Semeion)
	Cm = Contractive Map (Semeion)
Topology:	FF = Feed Forward (standard)
	Self = Self Recurrent Network (Semeion)
	Tasm = Temporal Associative Subjective Memory (Semeion)
Order:	DA = Dynamic and Adaptive Recurrency (Semeion)
	SA = Static and Adaptive Recurrency (Semeion)

The split into three sets tries to improve the learning capability of the SANN avoiding the simultaneous presence of records in which the same Input generates different Outputs. Indeed, although one peculiarity of the

SANNs is their ability to deal with fuzzy behaviours, the process of inconsistent patterns should lead to misinterpretations and errors.

The 1041 pattern selected for the experimentation have been randomly divided into two sets (Set 1 and Set 2). Ten different architectures of SANNs (Table 1) have been trained with Set 1 and validated with Set 2. The same SANNs have been also trained with Set 2 and validated with Set 1. In both cases the SANNs performances have been evaluated through statistical functions.

In this way it was possible to evaluate the SANNs prediction capability on the whole 1041 records set.

Finally the average of the 1041 prediction values of the 10 SANNs has been calculated, and again the prediction capability has been tested through statistical functions. Fig. 5 shows a flow diagram of the procedure.

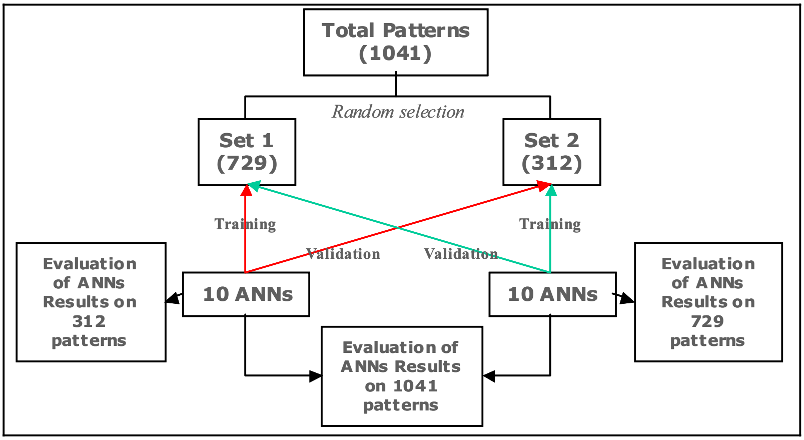


Figure 5. The training and validation procedure for the SANNs

7. LEARNING AND VALIDATION OF THE SANNS

A first evaluation of the results is given by usual statistical functions shown in Table 2.

It should be observed that, from a statistical point of view, the results are quite good for the residential use, a little less so for the industry and not so good for commerce. The difference is probably due to the different sample sizes for the three land uses. Since the recent urbanisation process in the south of Milan concerns mainly residential sprawl, many records are "good examples" for this land use. On the contrary the commercial use, which is the less frequent, gives the worst results. This is shown on the scatter diagram of observed (on the x axis) and calculated values (on the y axis) for each land use (figures 6 a, b and c).

The spatial representation of "errors" allows the evaluation of the spatial logic of SANNs output.

Figure 7 shows the errors concerning the residential land use. Errors are measured in ratio over the whole cell surface.

Table 2. Statistical measures of validation

	Residential	Industrial	Commercial	Average
RMSE	0.06756	0.05543	0.03290	0.09338
Real Error	-0.00262	-0.00938	-0.00550	-0.00583
Relative Error	0.05983	0.04911	0.01867	0.04254
Error Variance	0.11412	0.09735	0.05214	0.08787
NMSE	0.15737	0.22162	0.38873	0.25591
Squared R	0.84310	0.78374	0.62216	0.74967
Linear Corr.	0.91820	0.88529	0.78877	0.86409

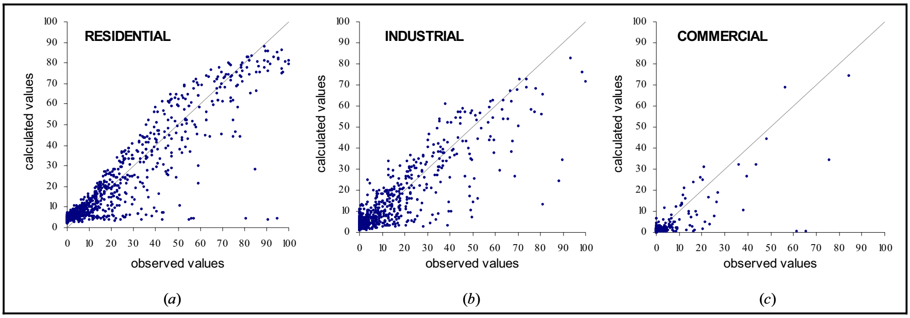


Figure 6. The scatter diagrams of observed and calculated values of land use in each cell

One large underestimation is evident, just in the centre of an agricultural area totally not urbanised and not infrastructured. This is due to an entirely new settlement for affluent people, "Milano 3", which is the result of a negotiation between big investors and the local municipality. Evidently it was impossible for the SANN to predict an event which is totally extraneous to the rules of change discovered by the NN.

Other errors are mainly due to planning constraints, often inhibiting a "natural" growth and forcing the development elsewhere. In fact a limit of the present work is the absence of the information on planning rules, which should allow to capture how planning constraints should play in generating spatial pattern of urbanisation. This is crucial information, which necessarily will be introduced in the future developments of the approach.

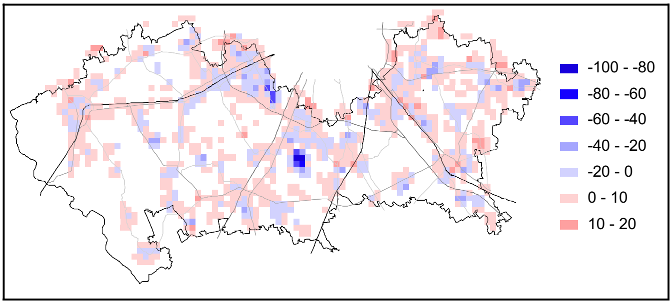


*Figure 7.*The differences between observed and calculated values *(in blue the underestimations, in red the overestimations)*

8. THE PREDICTION CAPABILITIES OF THE SANN

Once the learning and testing phase has been concluded, the averaged weight matrix of the SANNs is processed with the Data set of cells "potentially" in urbanisation in the next time lag (1994-2008). The prediction concerns "green" cells with urbanized neighbourhood in1994.

Figure 8 shows the estimated surfaces for each land use. As expected, the trend is linear, given the availability of only two temporal thresholds.

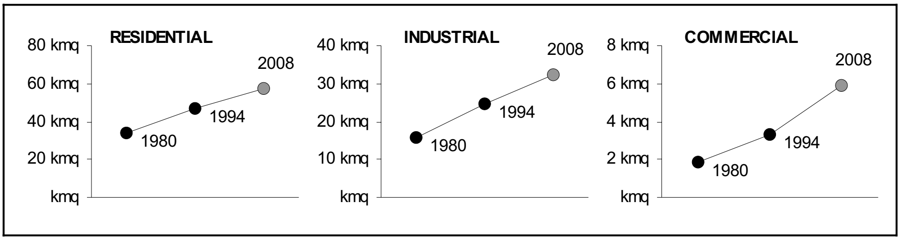


*Figure 8.*The predicted land use growth

The resulted pattern shows a probable scenario (Fig. 9 - Residence prediction) where the prevailing urbanisation process takes place at the boundaries of the urban centres, and, out of the urbanized areas, along the roads. Moreover, the new residence seems to be attracted by the proximity to other activities (industry and commerce). Indeed, this spatial feature characterizes the urban quality of the Italian historical cities and villages and holds particularly in this territory.

New productive settlements will be based mainly around the existing large industrial areas, showing a location criterion mainly driven by agglomeration economies. Such behaviour characterizes also the larger new settlements predicted from the NN for commerce, which, in the considered area, are clustering around new development poles near the highway and far from the urbanised areas.

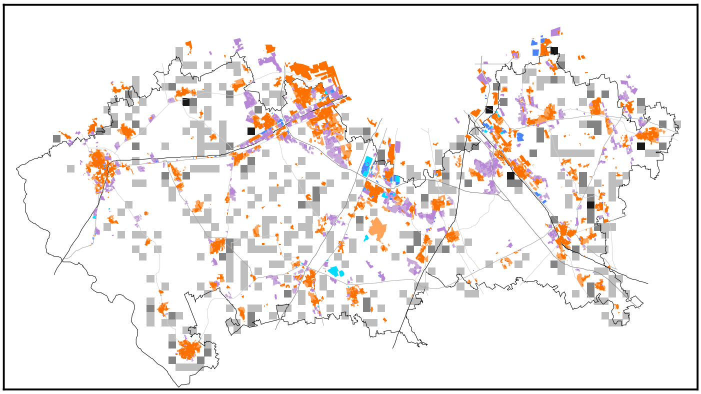


*Figure 9.*The predicted residential growth at 2008 (grey tones)

On the contrary, small and diffused expansions in industry and commerce will take place in and around the existing urban areas, improving the urban mix and showing the existence of urban agglomeration economies.

What is worthwhile noting is that the SANN findings, in terms of spatial patterns, are consistent with the SOM behavioural rules, shown by the codebooks.

To conclude the SANNs process seems able to capture and predict a sound spatial logic for future trends in urbanisation and to drive suitable territorial policies towards the facilitation or the inhibition of the "organic" processes of the considered urban system.

9. CONCLUDING REMARKS

The experiments presented demonstrate that NN is powerful tool for investigation of urban dynamics. The original aim of the paper was to integrate the spatial local logic of CA with the bottom- up knowledge construction of NN in pointing out the rules of change; therefore the information provided is that of a prototypical CA, which is limited to the local and neighbouring land use conditions.

The SOM run was able to show significantly different dynamic behaviours and to clearly distinguish the spatial location pattern of the urban functions considered: compact and urban for commercial activities, compact and peripheral for industry, more scattered and invasive of natural resources for residence, particularly in the last few years.

In each of the produced models, the codebook points out the degree of relevance of each variable in explaining the considered behaviour.

The SANN has produced a suitable scenario of the future urbanisation, even if with the limit of the stability of transition rules, which have been extracted by training in period 1980-1994 and applied in period 1994-2008 for forecasting purposes. But this is a limit due to the availability of Data, not to the method used. NN should be able to learn more complex dynamics if provided with larger temporal data set, unavailable at the moment.

As said before, these quite satisfactory results should be further improved by feeding the networks with other essential information concerning planning constraints, density, morphology, spatial relationships with central functions, political trend of local authorities and so on. Further research with an enlarged Data Base would enrich the knowledge on location dynamics with empirical findings and allow an updating of the concepts and assumptions of urban sprawl phenomenon.

Moreover, the scenario depicted gives an idea of the level of "urbanisation risk" for unbuilt cells, important information for territorial policies in order to drive future settlements towards less spread out spatial pattern.

REFERENCES

Batty M., Y. Xie and Z. Sun, 1999, "Modelling urban dynamics in through GIS based Cellular automata", *Computers, Environment and urban Systems*, 23, p. 205-234.

Bishop I., 1994, "Comparing regression and NN based approach to Modelling of scenic beauty", *WP of Center for GIS and Modelling, University of Melbourne*, Melbourne, Australia.

Bolchi P., L. Diappi and L. Franzini, 2001, "Applicazione di una rete neurale al frame di un automa cellulare", *Proceedings of INPUT 2001, Isole Tremiti, June 27th - 29th 2001.*

Buscema M. (ed.), 1998, "Artificial Neural Networks and Complex Social Systems. I Theory, II Models, III Applications.", Special Issue of *Substance Use and Misuse*, 33.

Diappi L., M. Buscema and M. Ottanà, 1998, "A neural network investigation on the crucial Assets of urban Sustainability", *Substance Use and Misuse Special Issue on Artificial Neural Networks and Social Systems*, 33, p. 793-817.

Clarke K., S. Hoppen and I. Gaydos, 1997, "A self modifying cellular automaton model of historical urbanisation in the San Francisco Bay Area", *Environment and Planning B*, 24, p. 247-261.

Deadman P. and R. Gimblett, 1995, “Merging technologies: linking ANN to GIS for Landscape research and Education”, *WP School of renewable natural resources, University of Arizona*, Tucson, Arizona.
Fischer M.M. and Y. Leung (eds.), 2001, *Geocomputational Modelling*, Springer Verlag, Berlin- Heidelberg.
Hecht-Nielsen R., 1990, *Neurocomputing*, Addison-Wesley Publishing Company, Reading, Mass.
Hinton G.E. and J.L. McLelland, 1988, “Learning representation by recirculation”, *Proceedings of IEEE Conference on Neural Information Processing Systems*, November 1988.
Kohonen T., 1995, *Self-Organising Maps*, Springer Verlag, Berlin.
Openshaw S. and R.J. Abrahart, 2000, *Geocomputation*, Taylor & Francis, London and New York.
Portugali J., I. Benenson and I. Omer, 1997 “Spatial cognitive dissonance and socio-spatial emergence in a self-organizing city”, *Environment and Planning B*, 24, p. 263-285.
Portugali J., I. Benenson and I. Omer, 1994, “Socio-spatial residential dynamics: stability and instability within a self-organizing city”, *Geographical Analysis*, 26, p. 321-340.
Reggiani A., 2000, *Spatial economic science*, Springer Verlag, Berlin and Heidelberg.
Sui D.Z., 1992, “An initial investigation of integrating NN with GIS for spatial Decision Making”, *Proceedings of GIS/LIS 92*, San Jose, California
White R. and G. Engelen, 2000, “High resolution integrated modelling of the spatial dynamics of urban and regional systems”, *Computers, Environment and urban Systems*, 24, p. 383-400.
Wu F., 1998, “An experiment on the generic polycentricity of urban growth in a cellular automatic city”, *Environment and Planning B*, 25, p. 731-752.

Visualisation for Design and Decision Support

Using On-Line Geographical Visualisation Tools to Improve Land Use Decision-Making with a Bottom-Up Community Participatory Approach

Jewell Station Neighbourhood - Case Study

C. Pettit, A. Nelson, and W. Cartwright
RMIT University, Melbourne, Australia

Keywords: Scenario Planning, Geographical Visualisation, GIS, Multimedia, VRML.

Abstract: This paper examines the development of a prototype suite of on-line integrated multimedia-GIS tools to assist in bottom up decision-making. These tools are being developed in the context of scenario planning to enable the community to actively explore different land use options and the implication of government structure and strategic plans. A case study approach is undertaken, focusing on the Jewell Station Neighbourhood, situated in the City of Moreland, Greater Melbourne Region, Australia. The paper documents the first stage of the project, in developing three land use scenarios delivered through a range of technologies including: VRML, HTML, GIS, Pixmaker and Flash. The paper concludes by outlining the future directions of this research that include: the construction of a virtual sandbox, usability testing, and community consultation.

1. INTRODUCTION

The broader context of the research is the trial of on-line spatial planning tools to assist community participation in urban planning and design. The aim of the research is to assess how existing geographical information technology can best be used to improve land use decisions using a bottom-up community participatory approach. The research focuses on: testing the utility of spatial scenario building tools; forming pedagogical insights on the user friendliness of these tools in assisting community members in the on-

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 53-68.

line visualisation of likely scenarios; and integrating community visions into the scenario formulation process.

For decades it has been thought that more effective use of spatial information might assist in the efficient planning and design of urban areas. Since the 1960s, researchers in geography, economics, and planning have been developing a suite of models to improve our understanding of the complex relationships existing between the economic, social, environmental and physical factors influencing urban expansion and intensified developments. However, many decision-makers, planners and community groups have been sceptical of these models and their results. Initially these black box systems based on complex mathematical formula and utilising aspatial data, were not user-friendly and typically top-down in application.

Fortunately, the advent of GIS and IT-proliferated developments and uses since the 1980s has enabled model builders to develop better planning support system (PSS) tools. These tools enable even non-specialist users to better visualise the likely consequences of planned developments by exploring various scenarios for proposed developments. These tools have great potential for supporting joint decision-making by the various stakeholders. Urban planning involves complex and critical trade-offs, for instance associated with sustainable ecological and social practices. PSS tools offer ways to enhance communication, consensual decision-making and conflict resolution at a community level.

The on-line Jewell Station Neighbourhood Project URL: http://www.c-s3.info/jewellwebsite/content/first.htm has been designed to trial spatial planning tools to assist in bottom-up urban planning and design. The project draws on the skill sets of a geospatial scientist with expertise in GIS and spatial planning - Dr Pettit, a social scientist with expertise in community participation - Dr Nelson, and a multimedia cartographer with expertise in web design and usability testing - Associate Professor Cartwright. The research team has been collaborating with urban and strategic planners in Moreland City Council, where the Jewell Station Neighbourhood is situated to assess how on-line geographical visualisation tools can best be used to improve land use decision-making through a bottom-up community participatory approach.

The project is in its first year and has funding for an additional 2-3 years. This paper reports on the development of the first stage of the project, that focussed on the development of a number of on-line GIS and multimedia based planning tools. The on-line website provides access to three scenarios. The first scenario presents the area as it exists today. The second scenario explores the possibility of introducing different buildings and height restrictions to the area. The third scenario shows what the Jewell Station Neighbourhood would look like with different tree planting in the streets and standard shop awnings in the retail area. Each of the scenarios have been developed using a number of different GIS and multimedia tools including: VRML, HTML, ArcGIS, ArcScene, Pixmaker and Flash. Users of the on-line site can hear sounds created from recordings acquired from key nodal points in the area, recordings of noises made by trams, trains, road vehicles and pedestrians.

By providing different representations of spatial information it is hoped that the community will be able to explore different development scenarios and actively participate within the planning process. The next stage of this research is to present the on-line planning toolbox to the Jewell Station Neighbourhood community so that they can comment on the future development of the Jewell Station Neighbourhood. Also, usability testing will be conducted to analyse how the community responds to different geographical and multimedia metaphors: virtual reality, map, sketch and panorama.

2. COLLABORATIVE SCENARIO PLANNING

Since the 1980s there has been a shift in the dominant planning paradigm from *planning for the people* to *planning with the people* (Forester 1999). The change in paradigm has seen public participation in the planning process increase rapidly. Consequently, there has been the need to develop a range of techniques for assisting with public participation processes. To date traditional methods for assisting with public participation in the planning processes have included the dissemination of information at community meetings, with feedback from participants through public forums and focus groups. Today local councils, including Moreland City Council are disseminating such information through their website. However, even with the prolific growth of information and communication technology (ICT) since the late 1980s, until very recently visual tools have not been used much to assist in public participatory planning processes.

Planning support systems (PSS) have been described as geo-information tools used to assist in public and private planning processes (or parts thereof) across a range of spatial scales and within a specific planning context (Stillwell 2003). The term PSS was first used by Harris (1989); Harris and Batty (1993) to describe systems that combined a range of computer-based models and methods. In the 1990s pioneering work in developing and testing collaborative planning support systems was undertaken by Shiffer (1992, 1995a, and 1995b). Since the early 21st century a number of books, reports and journal articles have been published which document more recent developments and use of PSS (EPA 2000; Brail and Klosterman 2001; Geertman 2002; Geertman and Stillwell 2002).

Another form of collaborative geospatial technology, similar to PSS, is participatory GIS (PGIS) (Jankowski and Nyerges 2001; Kingston, Evans et al., 2002; Nyerges and Jankowski 2002). PGIS has been defined as the integration of group communication technology with basic GIS capabilities and is generally considered an enhanced version of GIS called 'participatory GIS' (Nyerges and Jankowski 2002; Malczewski 2004). A third form of technology to assist the planning process has emerged, and is referred to as web-based multimedia-GIS.

An integrated web-based multimedia-GIS approach, as discussed in this paper is considered a loose coupling of multimedia and GIS technologies, rather than PGIS approaches - for example, the multiple criteria decision making (MCDM) GIS approach). This approach is generally considered a tight coupling of technologies (Jankowski 1995). The integration of multimedia and GIS has considerable potential as a geographical visualisation tool for adaptation in land use planning (Malczewski 2004), as the ability to visualise geographical information in 2D and 3D from many perspectives enables communities and decision-makers the ability to explore 'what-if' scenarios (Allen 2001; Klosterman 2001).

3. WEB-BASED MULTIMEDIA-GIS APPROACH

It is important to realise that the functionality contained within proprietary GIS for modelling and visualising is still quite limited (Malczewski 2004). An on-line multimedia-GIS approach is one response to the needs for extending traditional GIS functionality and engaging communities and decision-makers in making collaborative planning decisions. Shunfu (2003) discusses the application of web-based multimedia-GIS approach for enabling users to explore a database on Everglades vegetation, Florida.

This project builds upon previous research undertaken by the authors with respect to web-based multimedia approaches. The work stems from a technique used in the Townsville geoinformation prototype (Cartwright, Williams et al., 2003), which incorporated both multimedia and GIS elements, accessed through the appropriate use of metaphors to explore information landscapes. Our research aims to extend this technique by incorporating some tools/directions in web based planning support systems such as the on-line Wide Bay-Burnett Regional Information System URL: http://www.uq.edu.au/cr-surf/wide_bay/ (Pettit, Shyy et al., 2002) that focussed on engaging regional managers and the on-line Wombat

Community Forest Management Pilot Prototype system URL: http://www.wombatcfm.info/ (Pettit and Nelson 2003) that focussed on engaging the local forest community. Subsequently, the over riding goal of the on-line Jewell Station Neighbourhood Project research has been to work towards developing a plausible web-based multimedia-GIS framework for making collaborative decision about place and space. Figure 1 illustrates the preliminary multimedia-GIS framework and highlights the development of the Jewell Station Neighbourhood to date.

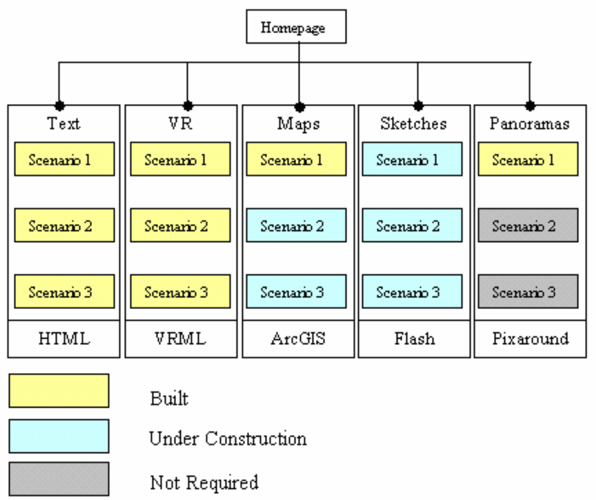


Figure 1. Web-based Multimedia-GIS Framework

3.1 GIS Scene Development

The Moreland Council had limited GIS ortho-rectified aerial photographs. By overlaying existing cadastral boundaries and road layers with the aerial photography within ArcGIS, suitable site maps were created. Fieldwork was undertaken to obtain building height attribute information and to validate building footprints. ArcGIS comprises a number of GIS modules for undertaking various spatial operations. For example, the ArcMap module in a desktop GIS, which has been used to construct and visualise the building footprint data layer created for the Jewell Station Neighbourhood, as illustrated in Figure 2. ArcMap provided a suitable environment for onscreen digitisation of a building footprint dataset and formulating supporting two-dimensional map visualisations.

Figure 2. Building footprint data layers derived from overlaying aerial photography with cadastre boundaries, validated through fieldwork, formulated and presented in ArcMap

3.2 VRML Scene Development

Previous application of VRML to visualise landscape change has been undertaken by Pullar (2001) whose rural case studies focused on water runoff and landform evolution. Pullar (2001) surmises that VRML is a cheap, available, versatile technology for visualisation landscape change however further work is required on dialog interface design.

Figure 3. Building footprint imported into VRML and textured with photographs to produce an interactive three-dimensional representation of the Jewell Station Neighbourhood

Our virtual reality model for the Jewell Station Neighbourhood has been created from exported building coordinates from ArcScene, another ArcGIS module, into VRML format. ArcScene enabled the two-dimensional data building footprint data layer produced in ArcMap to be extruded by building height data (z values), to produce a two and half (arguably three) dimensional data layer. Next, faces and textures have been added to the block faces imported into VRML to contextualise the streetscape.

Finally, sound recordings from street intersections have been encoded into the scene, setting emitter markers to enable the juxtaposition of sound space to create fluid sound dynamics. Figure 3 illustrates the Jewell Station and surrounding area developed primarily using VRML.

3.3 Multimedia Components

Flash multimedia technology is a tool used to create interactive on-line information. Previous research by Cartwright, Williams, et al., (2003) implemented Flash technology to develop an interactive mapping component for a geographical visualisation product for Townsville. Panoramic views have been created using Pixmaker and street sketches have been created in Flash.

Figure 4. Multimedia navigation tools including interactive 360 degree panoramic street intersections and an associated locality orientation map

Three hundred and sixty degree panoramic view capability in a PSS is a useful tool for evaluating sites, buildings and, neighbourhood character (Shiffer, 1995a). Pixmaker was used to stitch together photographs taken at street intersections to create number of interactive 360° panoramic viewpoints. A number of hotspots have been created in Pixmaker to enable the user to undertake an immersive virtual tour of the Jewell Station Neighbourhood.

These multimedia components offer additional mediums for the community to explore the Jewell Station Neighbourhood. Figure 3 shows an example of one of the panoramic views, which has been created using Pixmaker and is linked to a Flash locality map so that users can orientate themselves. The users can navigate through the study area by either clicking the camera icons illustrated in the locality map or by clicking the immersive hotspots found in the panoramic views.

4. JEWELL STATION NEIGHBOURHOOD – STUDY AREA

The Moreland City Council has been conducting consultations on future urban development in the Jewell Station Neighbourhood as part of a broader strategy to engage community in planning scheme reforms in accordance with the Victorian State Government’s Melbourne 2030 Planning Strategy (DOI 2002). Beyond trials of innovative visual and participatory planning techniques, the research deals with social change, collective visions and delegating the responsibilities for implementing planning decisions. While the council has the ‘big picture’ plan for the city, many stakeholders and community members have a narrower focus on their immediate neighbourhoods, the properties they own or their business activities. The key analytical concepts defined by the planners and researchers are ‘perceptions’ and ‘responsibilities’.

4.1 Scenario 1 – Existing Conditions

The Jewell Station Neigbourhood falls within a heritage zone that features small Victorian worker’s cottages and retail businesses in two storey terraces as well as light industry stretched between a railway station and the very busy Sydney Road. It is a diverse urban landscape with richly varied architecture that has been identified for its potential future residential and commercial development. Figure 5 illustrates the base scenario (1) on-line VRML model for a section of Sydney Road.

Figure 5. The Existing Conditions Scenario in Sydney Road, created using VRML

4.2 Scenario 2 – Exploration of Building Heights

The second scenario introduces new residential and office buildings at a variety of heights and step-backs from the main road frontage. This scenario is designed to enable community members to explore what their neighbourhood might look like according to a new Moreland City Council planning amendment involving heights of buildings over three storeys. The VRML geographical visualisation enables the community to compare the existing conditions with the proposed changes to building heights through an interactive environment that allows them to virtually walk around and look up and down their neighbourhood with these specific changes made to it.

Figure 6. Scenario 2: Change in Building Heights in Sydney Road, created using VRML

The council planners are especially interested in this scenario because they are consulting with stakeholders over their responses to recent changes to regulations involving building heights over three storeys. Figure 6 illustrates the building height scenario (2) for the same section of Sydney Road as illustrated in Figure 5.

4.3 Scenario 3 – Exploration of Streetscape vegetation and shop awnings

When the Sydney Road Brunswick Association was consulted in mid 2003 they expressed a distinct interest in the visualisation tools being developed by the researchers. Many shops line Sydney Road and their architecture is no more varied than in their street awnings that provide shade for products in shop fronts and protection from rain for shoppers. The third scenario involves standardising the shop awnings along the street so the traders can more easily discuss changes that might be made to their streetscape in this direction.

Figure 7. Scenario 3: Change in vegetation and shop awning in Sydney Road, created using VRML

The Moreland Landscape Strategy has raised a deal of interest among community members who conflict over the extent and type of vegetation in their streetscapes and especially the benefits of planting native rather than exotic vegetation. Pot plants in public areas, like Sydney Road, have not faired well because of the needs for constant watering and their abuse, for instance, being used as cigarette ashtrays. Figure 7 illustrates the streetscape vegetation and shop awnings scenario (3) for the same section of Sydney Road as illustrated in Figures 5 and 6.

5. WORK IN PROGRESS

The project is a work in progress, nearing completion of phase one which involves building initial on-line multimedia-GIS tools for the community to explore the Jewell Station Neighbourhood.

5.1 Sandbox Approach

Extending a concept first developed by Hudson-Smith and Evans (2003) in the Hackney Building Exploratory Interactive URL: http://www.casa.ucl.ac.uk/research/hackneybuilding.htm, the sandbox approach will ultimately enable community participants to build and submit their own redevelopment scenarios. This is similar to the work by Shen and Kawakami (2003) who have developed a participatory public space design tool using CAD, a VRML browser and a Java applet. In our project Java scripts are embedded within the VRML to create a library of drag and drop objects available for participants to make their own scenarios. For this exercise focus is placed around the parkland located at the western entry to Jewell Station. Figure 8 illustrates the sandbox prototype with the drag and drop object library that will ultimately be used by the community to design their own scenarios.

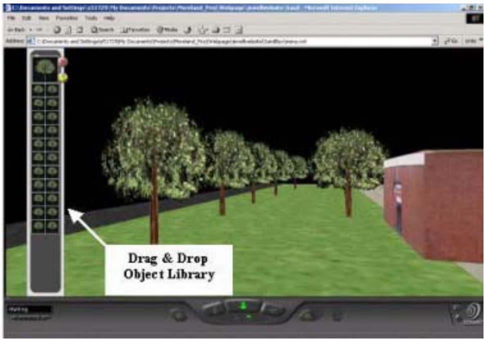


Figure 8. Sandbox VRML Prototype – Community drag and drop scenario formulation tool

5.2 Usability testing

This project is in the process of developing geographical visualisation tools for use by the general community. As the 'user dexterity' and understanding of how best to navigate through the information space is unpredictable it is imperative to assess the success of the user interface. As this can be considered to be a 'different' information access 'type', compared to what has been used before for community consultation, the authors believe that the results from this component of the project will yield results that might leverage better information access in later iterations of this project, as well as future, associated projects. There are interface problems associated with the user not understanding the system and the system not understanding the user. This is especially true when today's users are expecting Startrek's *Holodeck.* Yet users are limited by what is actually being produced, (Sharp and Sharp 1994) and thus will probably be disappointed with what the interface allows them to do. As newer interfaces develop, older interfaces are taken for granted.

Graphics User Interfaces (GUI) and computers enable users to work more effectively. Attention to the GUI of a computer system is just as important to the success of a computer package as are other components. Both the interface and the representational structure of hypermedia are central to the development of more effective and efficient means of transferring knowledge. The combination of effective GUI and powerful computers enables users to work more effectively. Interfaces changed with computer graphics in the 1970s, again in the 1990s using multimedia, followed by 3D interactive multimedia, VR multimedia and now delivering interactive 3D multimedia via the World Wide Web. Nevertheless things like the indiscriminate use of 3D can lead to 'gratuitous dimensionability' (Ingram 1994), which can negate any positive design inputs. Also, an overly literal translation of the 'best' use of any metaphor can make the metaphor useless. The evaluation of the project will assess whether the designed tools are effective.

A formative evaluation method is of use to mapping with new media. It can be used to compare conventional map artefacts with multimedia map use and therefore be able to describe how multimedia maps really should be constructed and used. It involves the collection and analysing of data based on the pedagogic value, the means by which we build up our knowledge of what the medium (multimedia) can do. Formative evaluation requires the practitioner to think in terms of what the user is supposed to get out of a package. The focus of evaluation is directed more towards breaking new ground and defining the unique contributions that multimedia and hypermedia make. This will form the very foundation of the evaluation process.

We will not use a formal evaluation, as such processes deal more closely with the content and performance of the product rather than the concepts behind it. Review by numerous community members will be conducted to receive comments about the concepts behind the design of the product and how such a product could be best delivered as a discrete/distributed unit. Evaluation procedures will observe how users actually use the package and record them ‘speaking their thoughts aloud’ as well as filling in questionnaires.

In the initial stage of the evaluation candidates will be asked to complete a four-part questionnaire that covers the areas of:

- reviewer profile;
- comments on the general use of 3D, in particular the use of VRML;
- navigation through the information space using the standard VRML controls plus the additional controls developed for the product; and
- general comments on the product.

As users become more attuned to using the product, later stages of the evaluation will monitor changes in use of the product as particular users move from novice to expert users, as they become familiar with both the content and the delivery medium.

5.3 Community Consultation

The central objectives of our broader research project focus on developing effective and efficient processes for enhancing community participation in future scenario planning and policy development and implementation at a local level. All community engagement draws inspiration from recent literature and activities involving participatory governance and sustainability research in Australia. Creative, iterative and integrative methods form the backbone of these research techniques and process. The aim is to develop a theoretical framework for enhancing community participation as well as producing practical visualisation tools.

The tools and scenario process developed will be assessed according to their capacity to inform and engage community in consensus-building land use decision making within the City of Moreland. Relevant transport agencies, business and community stakeholders will be targeted through meetings, surveys and forums as well as the dedicated website to select aspects for development in the *what-if* scenarios. There will be interactive screenings of the scenarios in forums of special interest and mixed public groups and a version will be available on-line for individuals to play with and submit their ideas by electronic form to the researchers. The researchers are facilitating open forums for community members and closed stakeholder meetings as well as running dedicated training sessions in a local Internet café and introducing the tools to classrooms in two local secondary schools and one tertiary institution. Community stakeholders will be engaged in the implementation as well as visioning exercises as they participate in meetings that are organised, facilitated, monitored and evaluated by the researchers.

The iterative methodology involves fine-tuning the technical and social processes and revisiting particular groups and/or the whole community for further trials of the on-line spatial planning tools to test developments, modifications and refinements. Various community stakeholders and council strategic planners will be included in evaluations.

6. CONCLUSIONS

We believe like others (Malczewski 2004) that an integrated web-based multimedia-GIS approach provides a platform for building upon traditional approaches by combining a number of modern technologies such as graphics, videos, virtual reality, digital sketches and panoramic displays. By making these tools readily accessible to both community and decision-makers will enable us to assess the extent and ways in which these tools are specifically useful to complement, substitute for and provide innovations in collaborative decision-making.

This research endeavours to develop and test suitable tools for collaborative on-line multi-media GIS decision-making. The success of building such a suite of tools will be the willingness of the community to embrace them to develop and explore their own ideas and ultimately increase their involvement and interest in the future of their urban landscape. Future work on this project will subsequently endeavour to address the research question: *how can web-based multimedia-GIS tools be used to improve land use decisions through a bottom-up community participatory approach?*

ACKNOWLEDGEMENTS

Moreland City Council for supplying structure plans and digital data. RMIT University Research and Innovation Office for research investment funding used to create the Community Spatial Scenario Simulation (C-S3) group http://www.c-s3.info/.

REFERENCES

Allen, E., 2001. “INDEX: Software for Community Indicators”, *Planning Support Systems*, in: R. K. Brail and R. E. Klosterman (eds.). Redlands, ESRI Press, p. 229-261.

Brail, R. K. and R. E. Klosterman, 2001, *Planning Support Systems*, ESRI Press, Redlands.

Cartwright, W. E., B. Williams, C. Pettit, 2003, “Geographical Visualization Using Rich Media and Enhanced GIS”, *MODSIM 2003 - International Congress on Modelling and Simulation*, Townsville, Queensland, Australia, MSSANZ.14-17th July 2003.

DOI, 2002, *Melbourne 2030: Planning for Sustainable Growth*, Melbourne, Victorian Department of Infrastructure: 202.

EPA, 2000, *Projecting Land-Use Change: A Summary of Models for Assessing the Effects of Community Growth and Change on Land-Use Patterns*, United States Environmental Protection Agency, Washington.

Forester, J., 1999, *The deliberative practitioner: encouraging participatory planning processes*, MIT Press, Cambridge.

Geertman, S., 2002, "Participatory planning and GIS: a PSS to bridge the gap", *Environment and Planning B: Planning and Design*, 29,p. 21-35.

Geertman, S. and J. Stillwell, 2002, *Planning Support Systems in Practice*, Springer-Verlag, Berlin.

Harris, B., 1989, "Beyond geographical information systems: Computers and the planning professional", *Journal of American Planning Association*, 55,p. 85-92.

Harris, B. and M. Batty, 1993, "Locational Models, Geographical Information and Planning Support Systems", *Journal of Planning Education and Research*, 12,p. 184-198.

Hudson-Smith, A. and S. Evans, 2003. "Virtual cities: from CAD to 3-D GIS", *Advanced Spatial Analysis*, in: P. Longley and M. Batty (eds.). Redlands, ESRI Press, p. 41-60.

Ingram, F., 1994, "Visual cues: How things look and work", *Proceedings from the Multimedia and Design Conference*, University of Sydney, Sydney, Australia, 26-28th September 1994.

Jankowski, P., 1995, "Integrating Geographical Information Systems and Multiple Criteria Decision-Making Methods", *International Journal of Geographical Information Systems*, 9(3), p. 251-273.

Jankowski, P. and T. Nyerges, 2001, *Geographical Information Systems for Group Decision Making,* Taylor & Francis, London.

Kingston, R., A. Evans, et al., 2002. "Public Participation via On-line Democracy", *Planning Support Systems in Practice*, in: S. Geertman and J. Stillwell (eds.). Berlin, Springer, p. 43-64.

Klosterman, R. E., 2001. "Planning Support Systems: A New Perspective on Computer-aided Planning", *Planning Support Systems*, in: R. K. Brail and R. E. Klosterman (eds.). Redlands, p. 1-25.

Malczewski, J., 2004, "GIS-based land-use suitability analysis: a critical overview", *Progress in Planning* (in press).

Nyerges, T. and P. Jankowski, 2002, "Data-gathering strategies for social-behavioural research about participatory, geographical information system use." *International Journal of Geographical Information Science*, 16(1), p. 1-22.

Pettit, C. and A. Nelson, 2003, "Developing an On-line Public Participatory Planning Support System for Natural Resource Management", *Spatial Sciences 2003: Spatial Knowledge Without Boundaries*, Canberra, Australia, Spatial Sciences Coalition, 22-27th September.

Pettit, C., T. Shyy, et al., 2002. "An On-line Planning Support System to Evaluate Urban and Regional Planning Scenarios", *Planning Support Systems in Practice*, in: S. Geertman and J. Stillwell (eds.). London, Springer-Verlag, p. 331-348.

Pullar, D., 2001, "Using VRML to Visualise Landscape Change and Process", *Cartography*, 31(1), p. 1-9.

Sharp, L. and L. Sharp, 1994, "Interface design aspects for educational multimedia", *Proceeding from the Multimedia and Design Conference*, University of Sydney, Sydney, Australia, 26-28th September 1994.

Shen, Z. and M. Kawakami, 2003, "Study on Development of On-line Cooperative Planning and Design System Using VRML and JAVA", *The 8th International Conference on*

Computers in Urban Planning and Urban Management, Sendai, Japan, CUPUM.27-29th May 2003.

Shiffer, M., 1992, "Towards a collaborative planning system", *Environment and Planning B: Planning and Design,* 19, p. 709-722.

Shiffer, M., 1995a, "Environmental review with hypermedia systems", *Environment and Planning B: Planning and Design*, 22, p. 359-372.

Shiffer, M., 1995b, "interactive multimedia planning support: moving from stand-alone systems to the World Wide Web", *Environment and Planning B: Planning and Design*, 22, p. 649-664.

Shunfu, H., 2003. "Web-based Multimedia GIS: Exploring Interactive Maps and Associated Multimedia Information on the Internet", *Maps and the Internet*, in: M. P. Peterson (eds.). Oxford, Elsevier Science, p. 335-344.

Stillwell, J., 2003, "Planning Support Systems in Practice", *The 8th Conference on Computer in Urban Planning and Urban Management*, Sendai, Japan CUPUM.27-29th May 2003.

A Spatial Decision Support System for the Management of Public Housing

Case Studies and Approaches to Interactive Visualisation

Jack Barton, B. Parolin, and V. Weiley
The University of New South Wales, Australia

Keywords: Spatial Decision Support System, Public Housing, Community Renewal, Security, Open Source, Interoperability, Visualisation, 3D GIS, PPGIS, X3D.

Abstract: This paper is reporting on a research project undertaken jointly between the University of New South Wales (UNSW) and the NSW Department of Housing (DoH) to develop a Spatial Decision Support System (SDSS) to assist planning, management and evaluation in areas of high public housing concentration. In the paper we will describe the development of the SDSS, the specific spatial problems challenging the DoH and the potential for the system to incorporate a range of social, financial and physical data, both internal and from other sources, for interaction and presentation in a three dimensional environment. The prototype SDSS attempts to address the specific challenges of providing better service for clients of the DoH. An information audit and survey has been conducted of the department's resources and needs. Issues identified include the management of high-rise and superlot areas, crime mapping, community interactivity, internal and intergovernmental information sharing, interoperability and maintaining confidentiality and security of data. Interactive 3D visualisation of the model is facilitated by use of the 3map free geospace platform. Use of open source code and open standards such as X3D for 3D graphics interchange allow the project to explore advanced visualisation techniques while ensuring interoperability and data longevity.

1. INTRODUCTION

In recent years the application of management concepts such as efficiency, effectiveness and economy to the assets of government organizations has been occurring at the same time as those managing and

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 69-84.

planning the assets have faced challenges of rationalisation, restructuring, and re-engineering. As part of this process, government organizations are promoting equity, equality and outcomes in terms of their processes and resources.

It is no surprise, therefore, that government organizations have moved into the information era and the development of information management systems to provide accurate information relating to an organization's assets. For example, the New South Wales (NSW) government's Total Asset Management initiative demands accurate information relating to an organization's assets and, in turn, organizations are being held more accountable for the decisions they make concerning their management (New South Wales Government 1993, 2000).

In many government organizations this means that "decision support systems" must be developed to provide information to management. Increasingly these systems incorporate geographic information and are referred to as spatial decision support systems (SDSS) or spatial planning support systems (SPSS) - in recognition that much of the asset, facilities and planning data held by the organization is spatially referenced and can, therefore, be analysed using appropriate geographical information systems (GIS) technology (Naudé, 1999).

In the realm of public housing the mapping of physical, financial and social data accumulated over the life-cycle of particular housing areas may be used to support community renewal strategies as a tool for resource planning, reviewing and predicting changes. Performance and condition indicators may be benchmarked against each other to compare and evaluate strategies.

Spatial decision support systems are increasingly being recognized as a smarter and better way of utilizing information technology for the re-engineering of asset management within government organizations, and for generating productivity and efficiency gains. (Ward, 1997; Clarke, et al, 1997). This is in contrast to previous methods of handling data where information on assets may have been duplicated across different units within the organization, using different standards and formats. Questions from senior management, clients and the community could only be answered after significant efforts in collating information (with reliance on individuals and their specific knowledge of assets and processes).

This paper reports on a research project undertaken jointly between the University of New South Wales (UNSW) and the NSW Department of Housing (DoH) to develop a Spatial Decision Support System (SDSS) to assist planning, management and evaluation in areas of high public housing concentration (Barton et al, 2002). We describe the basic structure of our SDSS (prototype only), the reliance on information technology (GIS) and an application development platform which relies on interactive 3D visualisation facilitated by use of the 3map free geospace platform (Thorne and Weiley, 2003).

2. LITERATURE

A review of the literature on SDSS for public housing indicates a paucity of research on the topic. However, reports on the use of information technology - including GIS - for public housing are only slightly more frequent (Han and Yu, 2001; Yu and Hang, 2001; Fine, 1998) and tend to focus on the domain of facilities management in public housing. These latter studies provide a useful frame of reference for the present study.

The application of GIS for public housing management and maintenance in Singapore (Han and Yu, 2001; Yu and Hang, 2001) is not a SDSS per se but an information system which enables a particular town council and public housing tenants to access various data systems organised into modules. The system is known as EMAPS - a GIS based management information system. Although this system has some of the elements of a SDSS (i.e. GIS, databases) it differs from the definition of a SDSS in that it is not structured for specific decision making purposes (Densham, 1991). It has more of the elements of a Spatial Planning Support System (SPSS) as defined by Naudé (2001).

Han and Yu (2001) state that there were three main objectives for the development of EMAPS: to improve the capability of town councils; to bridge a gap between rapid technology development and slow integration of new technology into facilities management, and to explore the application of GIS to public housing management. EMAPS is based on a series of menus which allow the user access to a range of data from financial, accounts, tenants information, property information, facilities assessments, cyclical works, and telemonitoring systems. Databases are updated regularly and are web enabled. The application of EMAPS is meant to support a managerial function as opposed to a decision-making function, although there is the capability to extend the system to function as a SDSS. The development of EMAPS highlights one of the rationales for using GIS - integration of diverse databases into a common format to support specific tasks.

The report by Fine (1998) is a more specific application of information technology to maintenance management. This application is non-spatial (does not involve GIS) at present and involves database systems to support housing condition surveys, rent analysis and works programs. Of interest is that this system was developed for a tenant based housing management scheme (tenant managed estates) which is allowed under certain conditions in the UK.

The only published example of a decision support system for public housing is that by Conte (1999) for the Bari municipality of the Bari province in Italy. This on-going research project had as its impetus concerns over the lack of sustainable management plans for public housing, socio-economic decline of public housing tenants, severe maintenance problems leading to physical and environmental decay on estates, lack of policies on rehabilitation, and lack of initiatives on self-management of public housing estates. In a broader context there was a need to foster the development of sustainable public housing management plans which recognised the interest and role of the European Union in housing policies of member states, the environmental awareness of public housing tenants, the role of urban rehabilitation as a tool for economic and social equilibrium, and availability of computer technology (Conte, 1999). It is interesting to note that many similar reasons also apply to the rationale for the present study reported upon in this paper.

Of relevance to the present study is that the justification of the decision support system developed by Conte mirrors, in many ways, similar needs and questions posed by the researchers in collaboration with persons at the DoH. In essence the need was for a structured DSS which was practical and useful for the maintenance and management of public housing in the city, and which allowed for collaboration with various local government agencies and participation on the part of tenants. The actual structure of the DSS included, on the one hand, integration of dispersed data and knowledge in the form of systems (housing information, technical legislation, and technical expert systems) for maintenance aspects of management policies. On the other hand, information systems dealing with social aspects, general legislation and policy experts were integrated to support the more strategic aspects of management policies. Elements of public participation occur throughout the DSS, but particularly in developing the housing and expert system modules of the DSS.

According to Conte (1999), the result is an automatic DSS, formed by several modules, which integrates and manages a great deal of data, both quantitatively and qualitatively, and is web enabled. (However, no examples of the working of the system were provided in the paper). The SDSS reported upon in the present paper does not have an identical structure to that developed by Conte, but mirrors the module structure which attempts, among other things, to integrate various information systems into a common platform to support various decision-making and strategic planning outcomes. However, the SDSS we propose differs markedly to that of Conte (1999) and Han and Yu (2001) in its use of interactive 3d visualisation.

3. THE LOCAL SCENE

The New South Wales Department of Housing (DoH) provides and manages public housing for those people who cannot obtain housing through the private sector. The DoH manages 130,000 dwellings in 86,000 buildings valued at around $AU17 billion. (3,000 of these properties are on the heritage register).

Thirty per cent of NSW public housing is in estates with over 100 properties. These were developed from the 50's to the mid 80's during a phase of 'slum clearance' and a push to maximise unit output to overcome a post-war housing shortage. These estates today typically experience inherent design problems, opportunities for anti-social behaviour and poor access to essential services such as health, transport and family support. Furthermore, the concentrations of disadvantaged families are often stigmatised by the broader community. The issues faced by these public housing estates share a similarity to those reported by Conte (1999) in the case of the city of Bari.

To alleviate some of these issues, the Neighbourhood Improvement Program (NIP) was implemented in 1994/5 as a community renewal process to bring estates with the poorest indicators up to modern standards as defined by DoH policies (DoH, 2001). Key strategies include:

- Innovations in housing management.
- Upgrades to dwellings and infrastructure to enhance continuity with the broader context.
- Greater involvement of tenants in estate-level decision making.
- Involving other social housing providers in managing properties on estates and improving the provision of services from other agencies.
- Targeted sales of assets to break up heavy concentrations of DoH ownership and enhance socio-economic and demographic diversity.

However, on a practical note, these key strategies have not previously been linked to a DSS or SDSS such that the implementation of strategies remains uncoordinated, use of relevant data is not coordinated, data is at present not spatially linked to GIS (although some exceptions are present), and visualisation of data is restricted to two dimensions.

3.1 Case Study Areas

Two study areas have been chosen for our project: Redfern/Waterloo and Penrith/Cranebrook. The sites have been chosen as case studies for their diversity of building types, as both have high densities of public housing and are undergoing community renewal programs.

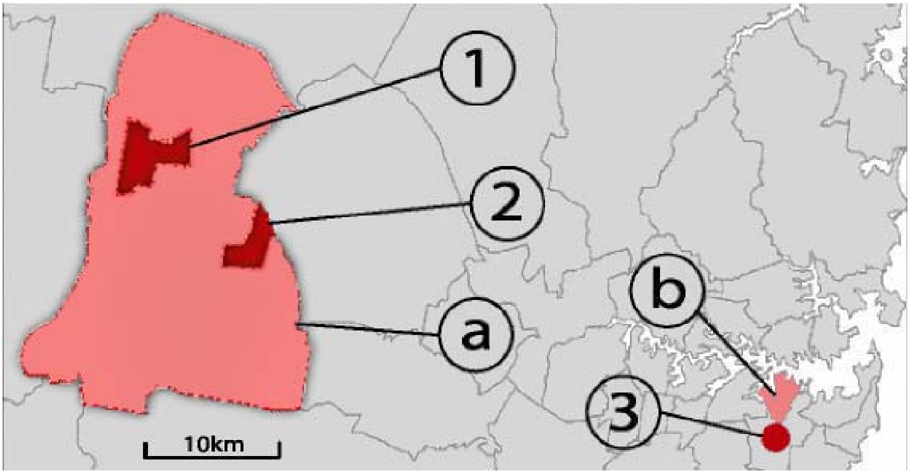


Figure 1. Both Cranebrook (1) and St Marys (2) are located within the Penrith LGA (a). The Redfern/Waterloo housing area (3) is just south of the Sydney CBD (b).
(Source of base map: Land and Property Information NSW)

The Redfern/Waterloo social housing area, located only 3km from the Sydney CBD, consists of a mix of housing types including high-rise apartment blocks, walk-up flats, terrace and modern infill housing and specialist aged housing. The area has an ongoing community renewal process instigated in 1995/6. Issues raised by the community became the catalyst for the Community Renewal Strategy (CRS) for the estate and focused on the lack of security, problems with lack of ownership of common space, the quality of the dwellings, particularly bedsits and the level of unemployment on the estate (DoH, 2001).

Public housing in the broader Penrith Local Government Area (LGA), located some 50km from the Sydney CBD, consists of three-story apartments, townhouses, cottages and 'Radburn' style detached housing. The DoH Special Projects Unit is undertaking an Integrated Strategic Planning Process based on GIS for the Penrith LGA, and has already established an operational Geographic Information System (GIS). Two focus areas have been identified within the Penrith LGA: Cranebrook and St Marys. Cranebrook was the last broadacre estate developed in 1979 by the then Housing Commission and has been involved in an ongoing community renewal strategy (DoH, 2000). The St Marys area will be subject to an urban improvement strategy to realize market potential and up-zoning.

The key problems at Penrith/Cranebrook stem from American style planning principles (Radburn principles) which created separated car and pedestrian networks, houses backing onto the road, facing shared open areas, with unclear delineation of public and private space and isolated pedestrian networks which presented many opportunities for crime. Currently, the DoH is undergoing a broad process of 'De-Radburnisation'; reversing the orientation of houses on the estates, removing access ways and hence opportunities for crime and increasing areas of private, controllable outdoor space for tenants (DoH, 2000).

Another problem with the estate model was the creation of 'superlots'. At the time of their creation, the housing authority was exempt from local government planning requirements. The tenure of the land parcels is illustrated on the Digital Cadastral Database (DCDB) by large, unconsolidated expanses. Large swathes of properties were built on these superlots. This makes it difficult for traditional data systems to present information, as properties have no spatial locator within the lot. Subdivision of these lots is costly and is only validated on areas intended to be sold and divided into smaller tenable lots. Furthermore, approval for subdivision has to be made by local councils. The DoH has to approach local councils individually, and often faces lengthy negotiations and sometimes obstructive conflicts in the process (Barton, 2003).

In a managerial and operational context, the DoH must locate the properties themselves so the information is useful internally, eg. for coordinating maintenance teams, and for external agencies, especially emergency services. The problem of locating individual properties within superlots also has a vertical component for high-rise estates, and inversely some low density occupancies extend over more than one lot. To resolve these issues the DoH needs to employ more sophisticated tools with 3d locators and a spatial interface in order to view these 3d subsets within superlots (Barton 2003).

Consultations with various stakeholders and individuals within DoH suggest to us an SDSS structure which must be capable of the above decisions, but which should also allow use of more detailed datasets from across a diverse range of providers (private and public) that are able to be linked and geo-referenced; have the ability to view data in a spatiotemporal and 3d context; the ability to share knowledge and improve continuity; and the ability for estate tenants to interface with management, particularly in relation to community renewal strategies.

4. THE STRUCTURE OF THE SDSS

The SDSS reported in this paper has been designed to incorporate a range of social, financial and physical data, both internally and from other sources. Further, it should be interoperable with internal systems and facilitate intergovernmental information sharing (Nairn, 2000). Users should be allowed to operate at a range of scales in a 2d/3d environment (Levy, 99) and be allowed to record, update, manage and represent knowledge by means of a knowledge base and expert component (Goel, 99). Finally, the system must be operable by executives, experts and non-experts whilst maintaining the confidentiality and security of data (Barton, 2003).

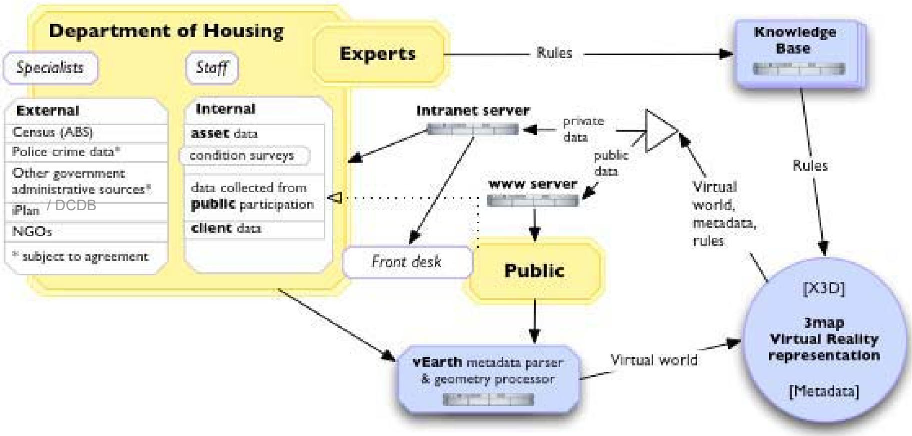


Figure 2. Structure of SDSS

The input data for the SDSS may incorporate various information systems available from within the DoH, or from external sources, such as national census data (socio-economic and housing related), electronic planning information systems (iPlan) and the DoH asset and client information management systems and potentially other data systems provided by a range of agencies (eg. transport, police, social welfare). A characteristic of these data systems, as currently used within DoH, is absence of a common spatially referenced platform. For example, the asset management information system ascribes building condition survey information to properties located on lots (cadastre), but the system is not able to be linked to the DCDB or to the census for GIS storage, retrieval and analysis functions. The presence of individual standalone information systems is legacy of an environment in which decision-making at a management and/or strategic level reflected isolation as opposed to whole-of-government approaches.

In addition, the availability of more specialised information from external sources, which was already spatially referenced and portable across different GIS platforms, further highlighted the problem of a lack of an integrated 'corporate' GIS within DoH which could facilitate a range of decisions. The model of 'compartmentalised' decision-making based on one information system (eg. maintenance management using the asset management system) did not allow interoperability or visualisation of the data. Until recently, most visualisation of the data was non-spatial and restricted to 2d CAD. Our SDSS model supports integration, interoperability and advanced visualisation to facilitate decision-making.

4.1 Three-Dimensional Visualisation Techniques

Three dimensional visualisation techniques are the platform through which the SDSS is visualised and web enabled (Figure 2). These techniques may usefully show the patterns, forms and character of the case study areas (Levy, 1999), as well as representing non-physical attributes. Some data has important 3-dimensional spatial attributes that need to be visualised (e.g. digital cadastral data and building attribute data). The extra spatial dimension of 3D, and the extra temporal dimension that can be shown through animation and explored through interaction, gives more room for layering extra information dimensions affording comparative data visualisation (Tufte, 1990).

There are several possible platform choices available when designing a 3D representation of data. Various 3D engines exist which are closely tied to the multitude of available file formats. However these engines tend to be highly optimised for their purposes (e.g. games), constraining the size and type of space that can be represented, as well as the types of interaction that can occur in the space. Also, not all such engines can operate over the Web.

A more flexible option is to use a language designed for general-purpose 3D worlds, such as Inventor, Shockwave 3D, MPEG-4, BiFS, VRML or X3D. These languages are designed without such specific optimisations in mind; some are open standards, which can support multiple runtime engines implementing different optimisations. This comes with a speed penalty when compared to game engines, but corresponds with greater flexibility in the kind of environments and interactions that can be implemented.

4.1.1 X3D for 3D Geographic Representation

X3D is the ISO standard for interactive 3D which has been developed by the Web3D Consortium (2003) as the successor to VRML97. It is the standard chosen to visualise our SDSS. X3D is a general-purpose virtual reality language that affords interactive 3D visualisation. Large environments can be supported through level-of-detail (LOD) nodes. It is intended for Web delivery, which allows for much larger spatial models to be managed on distributed servers but still be viewed on limited clients. It is back-compatible with VRML97, the most widespread general-purpose VR language and file format. Almost all tools capable of exporting 3D support VRML, including the tools in use at DoH.

X3D is standard, back-compatible and XML based, thereby making it straightforward to write parsers and editors. Multiple implementations exist which are royalty-free; and all modern operating systems are supported. It is scriptable through ECMAScript and Java (through the Scene Authoring Interface), supports embedded metadata and is optimised for online delivery.

Figure 3. Incidents of Victimisation in a High-Rise apartment. (Base model source: Barton et al 2003)

Figure 3 depicts an example of visualising fictitious incidents of victimisation in a high-rise residential apartment block.

The model is accessing several spatially referenced datasets: the X3D building data itself links to the DCDB and a table of reported crime incidents. Sigils represent different categories of victimisation. The system is accessing the database of incidents, extracting relevant categories and displaying the result in 3d through a web browser.

Patterns are revealed in three dimensions that may not be apparent in a traditional 2d system, for instance; vertically stacked zones that may be public or private, solid or void, roof top, ground level or subterranean.

It should be noted that problems with X3D include the fact that single-precision data representation gives insufficient dynamic range to support global datasets. This is fixed in the geospatial component of X3D with the addition of a GeoOrigin node to create local single-precision zones; however there is no currently available implementation of the geospatial X3D component. The Web3D Consortium is supporting the development of Xj3D, an open-source Java-based X3D implementation, which is on the verge of implementing the geospatial component. The availability of open source code makes these problems addressable.

4.2 Metadata standards

Interoperability with spatial data providers will be crucial in broadening the scope and longevity of the data and information systems which underpin our SDSS. The emergence of international standards for geospatial metadata facilitates this goal. 3Map uses Open GIS Consortium (OGC) standards wherever possible (Ping, 2003). Key OGC specifications are OGC Feature Geometry (ISO 19107:2003 - Spatial Schema) and OGC Metadata (ISO 19115:2003 - Metadata). The baseline for discovery metadata is the mandatory set of OGC Metadata elements. The joint Australia/New Zealand government Spatial Information Council (ANZLIC) has been working with ISO/TC 211 on ISO 19115, and the ANZLIC Metadata Guidelines are designed to comply with that standard.

Figure 4 shows an application of meta-tagged X3D data in a community renewal context. Once again, several databases are being interrogated: X3D building volumes, the DCDB, a digital terrain model and a layer dedicated to comments invited from the public. The community may visualise the area through web-enabled access and attach different comments to specific geographic locations, reflecting their perceived quality of neighbourhood facilities: neighbourly or neglected locations, valued or feared areas.

Qualitative data are captured upon upload and structured with metadata. A time and date stamp could be shown, coupled with a numeric value to register perceived quality. In turn this input is added to the community knowledge base. Personnel involved in formulating community renewal strategies may then access and interoperate with this data, for instance analysing change in community perceptions of the area over time.

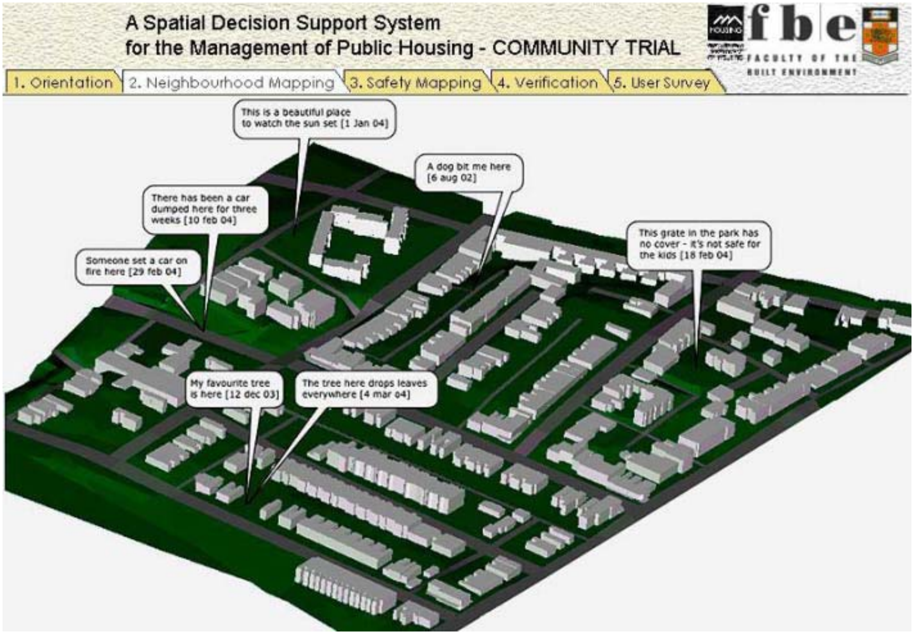


Figure 4. Example of cumulative community knowledge displayed in an urban renewal area (Base model source: Barton et al 2003. Enhanced by authors using dummy data)

4.3 Interaction design

The use of dynamic queries for visual information seeking (Schneiderman, 1994) is an effective method of visualising and interacting with metadata, particularly for geospatial datasets. Experts can apply their knowledge of a specific subject area to find relevant metadata elements and represent them with an annotated key.

This key can be presented to the user as a configurable slider, affording direct manipulation (Schneiderman, 1997) of metadata within the spatial context of the virtual reality representation.

Metadata on the server, when combined with user actions will effect the presentation of that data to the client. In Figure 5, a user is visualising compliance with a maintenance survey. The figure shows building footprints corresponding to the DCDB where building maintenance compliance is graphed in 3d in response to user configurable sliders. Sliders may vary attributes such as colour.

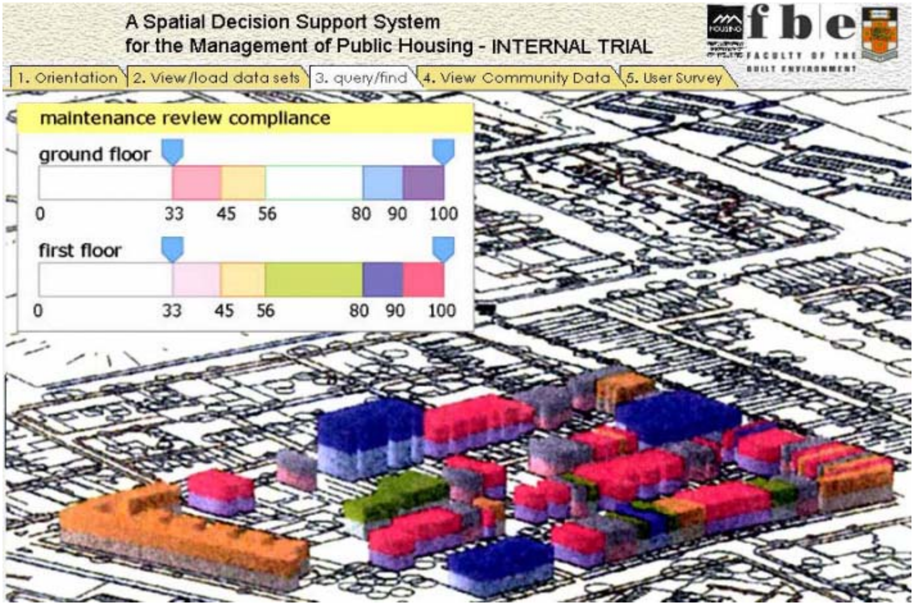


Figure 5. Envisaged use of sliders to identify compliance with user defined criteria
(Source of base image: Walia, 2003)

4.4 Security

The DoH database contains data which cannot be published due to privacy concerns. There will be various levels of access to data in-house, as well as selective public access. One issue is that even revealing the data source is a privacy issue. If the DoH database only publishes data on their assets then everything they publish is identifiable as public housing. One solution is to publish broader areas including non-DoH assets, without identifying which objects are DoH assets.

Privacy is related to detail. Publishing of street-level data is acceptable for public access, but details of assets including their interiors are private. The model has the ability to zoom from a broad view - regional, national or global - down to a specific asset and its attributes complete with power points, smoke detectors etc. Access to lower levels of the LOD tree is restricted to bona-fide users.

Server-side exclusion should be employed in order to maintain security of confidential data. One can locate the server with public data on the public internet (figure 2), and maintain the server with private data on a private intranet. This provides physical security; users must be on the premises and connected to a private network in order to access the data. Remote access to data may be facilitated by HTTPS (HyperText Transmission Protocol, Secure) requiring the end-user to log into the server. This approach is compatible with open standards, and is resistant to reverse-engineering.

4.5 Experts and Knowledge Management

Experts are a key component to an operational SDSS (Goel, 1999). It is envisaged that a knowledge management system form an integral part of the model, to record, update, manage and represent knowledge by means of a cumulative knowledge base. This will assist in improving continuity of knowledge within the department and add further value to their existing information. Integration of an inference mechanism provides a rule based environment adding constraints to guide the decision maker and informing users how to approach specific problems, especially if it is not the first time within the organisation the problem has been encountered.

5. CONCLUSION

This paper has described the development of an SDSS for public housing which is relevant to a government department (DoH) which is undergoing rapid change in its organisational structure to better facilitate management and strategic development of public housing in the state of NSW. The specific problems faced by DoH decision makers in dealing with our case study areas focus on community renewal strategies, crime prevention, asset management and maintenance issues, socio-economic status in estates, and issues of property boundaries and ownership.

Our SDSS model, which is in its prototype phase, is designed to assist decision-makers explore alternatives, constraints and benefits to complex and interacting problems through the avenue of integrated spatially based data and information systems which are manipulated and visualised in a 3D environment. We have described a structure which is entirely built around X3D, but also enables more traditional 2D views of spatial data.

A further strength of our system is its ability to enhance possibilities for public participation - especially in community renewal, asset maintenance, and crime prevention issues – and, more generally, for the sustainable management of public housing estates. Public housing tenants will be able to access the web enabled 3D model and to directly contribute inputs to the above issues.

The next stage of our research is to move from a prototype SDSS to a more formally approved system which has been tested by DoH personnel and public housing tenants to ensure adequacy in support of management decision making and strategic planning. We envisage that our SDSS will evolve to incorporate other modules and/or information systems, and will benefit from further development of the knowledge system (crucial given future restructuring and re-organisation of DoH).

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the NSW Department of Housing for funding this project, the support of staff in the Central Sydney and Western Sydney Regional offices, and to the support from the DoH steering committee formed for this project. UNSW students assisted in the construction of the initial model for the Redfern/Waterloo estate and critical inputs were given by Dr. Bruce Judd.

REFERENCES

Australia/New Zealand government Spatial Information Council (ANZLIC), 2003, http://www.anzlic.org.au/

Barton, J., Judd, B. and BENV2411 Students, 2002, Urban Modelling Winter Elective, Faculty of the Built Environment, University of New South Wales, http://www.fbe.unsw.edu.au/courses/benv/2411/

Barton, J., 2003, "Report 2: Survey of Department of Housing Staff", Completed as part of the masters by research thesis: *A SDSS for Public Housing,* Faculty of the Built Environment, University of New South Wales.

Clarke, G., Dean, L., Segal, J., 1997, "*Improvements in Property Asset Maintenance and Facilities Management Through the use if Better Technology"*, Presented at AURISA97 – The 25th Annual Conference of the Australasian Urban and Regional Information Systems Association held at Christchurch, New Zealand, 19-21 November 1997

Conte, E., 1999, "A Decision Support System For A Sustainable Management Of Public Housing", Presented at *The International Conference in Computers in Urban Planning and Urban Management* (CUPUM), Venice, 8 - 11 September, 1999.

Densham, P.J., 1991, "Spatial Decision Support Systems". In *Geographical Information Systems: Principals and Applications*, edited by D.J. Maguire, M.F. Goodchild and D.W. Rhind, 403-12. New York: John Wiley & Sons, Inc.

DoH, 2000, *Factsheet: Community Renewal in Cranebrook,* NSW Department of Housing Communication and Media Team, Sydney.

DoH, 2001, *Community Renewal: Transforming Estates into Communities - Partnership & Participation,* NSW Department of Housing, Sydney.

Fine, B., 1998, "Application of Information Technology to Housing Management", *Facilities*, Volume 16, Number 11, pp.306-313, November 1998, MCB University Press

Goel, R. K., 1999, *Suggested Framework (Along With Prototype) for Realising Spatial Decision Support Systems* (SDSS) GISdevelopment.net

Han, S.S. and Yu, S.M., 2001, "GIS for Public Housing and Maintenance: EMAPS in Tanjong Pagar – West Coast Town Council", Singapore, *7th International Computers in Urban Planning and Urban Management Conference*, University of Hawaii at Monoa Honolulu, Hawaii USA.

iPlan, last updated 3rd December 2003, http://www.iplan.nsw.gov.au/industrylands/

Levy, R.M., 1999, "*The Role Of Computer Visualization in Design Review*", Presented at *The International Conference in Computers in Urban Planning and Urban Management* (CUPUM), Venice, 8 - 11 September, 1999.

Nairn, A.D., 2000, "Australia's Developing GIS Infrastructure - Achievements And Challenges From A Federal Perspective". Presented at the *5th International Seminar on GIS,* Seoul, Korea, 28-29 September, 2000

Naudé, A.H., 1999, "The Essential Elements of a Spatial Planning Support System". Presented at *The International Conference in Computers in Urban Planning and Urban Management* (CUPUM) Venice, 8 - 11 September, 1999.

New South Wales Government 1993, 2000, *Total Asset Management Manual*, New South Wales Government Printing Office, Sydney.

Ping Interactive Broadband pty ltd, 2003, *Ping: 3map the 3D Mapping Platform*, http://www.ping.com.au/3map/

Schneiderman, B., 1994, *Dynamic Queries for Visual Information Seeking,* IEEE Computer Society Press, Los Alamitos, CA, USA.

Schneiderman, B., 1997, *Direct manipulation for comprehensible, predictable and controllable user interfaces,* IEEE Computer Society Press, Los Alamitos, CA, USA.

Thorne, C. and Weiley, V. 2003, "*Earth's Avatar: the Web Augmented Virtual Earth"* Proceedings of *The SIGGRAPH 2003 Conference on Web Graphics: in Conjunction with the 30th Annual Conference on Computer Graphics and Interactive Techniques,* ACM Press New York, NY, USA.

Tufte, E. R, 1990, *Envisioning Information*, Graphics Press, Cheshire, CT, USA.

Walia, J., 2003, *Integrating Geographic Information Systems (GIS) in Asset Management for NSW Public Housing* Faculty of the Built Environment, University of New South Wales.

Ward, K., 1997, "School Assets Management System", Presented at *AURISA97 – The 25th Annual Conference of the Australian Urban and Regional Information Systems Association*, Christchurch, New Zealand, 19 - 21 November 1997.

Web3d Consortium, 2003, Extensible 3D (X3D), International Draft Standards: http://www.web3d.org/fs_specifications.htm

Yu, M.S. and Hang, S.S., 2001, "Information Systems in Public Housing: Case Study of a Town Council in Singapore" *Facilities*, Volume 19, Number 5/6, pp.233-259, November 2001, MCB University Press.

Visualization of Usable Building Space According to Planning Permission Ordinances for Public Participation in District Plan in Japan

Z.J. Shen and M. Kawakami
Faculty of Engineering, Kanazawa University, Japan

Keywords: WEBGIS, JAVA3d, Planning restrictions, Public participation.

Abstract: The district plan in Japan is designed and implemented through public participation, which is based on the zone restrictions of land use system. The usable space of a building can be generated according to the zone restrictions implemented in a district plan, which include the planning controls of high altitude, oblique line and so on. Residents can choose control items as their preference in order to control the urban physical environment of the district. In this paper, we discuss how to generate the virtual world according to the data set of GIS including planning control items and how it can work for net participation. In the future, we would like to conduct an Internet social experiment with the cooperation of a city government and residents in order to support the district plan in a local city.

1. INTRODUCTION

There are on-line planning support systems that have been developed to enable planning practice to take place using a public participation workshop. The expansion of possibilities of WEBGIS, the 3d visualization (Michael B. et al., 2001) of GIS data is taken into account. The possibility has been described for modelling an urban environment using web based mapping and virtual reality technologies (Doyle, S. et al., 1998), and the method of computer simulated visualizations of a built environment has been discussed for better public participation in the design process (Daniel Bulmer, 2002). There are also many others reports regarding the visualization of an urban environment.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 85-98.

A building volume that can be legally permitted be constructed on a land parcel according to the planning law system is defined in terms of usable space of building (Kobayashi, F. et al., 1999). For generating such a building volume employing a computer system where parameters include land conditions and shape parameters are necessary. These attributes are relative to the respective parcel. The land condition is defined in a term that describes the size, vertexes coordination data of parcel polygons and the surrounding road situation, which can be saved in a GIS database. In addition, shape parameter is defined in a term that describes optional restriction values according to the designated legal requirements. The shape parameters consist of an index regarding length, width and height and an oblique line that defines the maximum shape of a construction space on a particular land parcel.

For visualization of the usable space of a building, land conditions of the parcel polygons that are saved in the GIS database should be combined with the shape parameters in order to generate a 3d world. In addition, a 3d simulation tool such as JAVA3D and VRML and so on is necessary. Visualization of the usable space of a building is very useful in order to understand what kind of building volume can be constructed according to planning permission ordinances. In public participation, most of the participants are not specialists in the field of urban planning and design so visualization using a computer system is certainly helpful for participants to understand the building volume.

The users of such a support system are considered to be administrative officers and citizens in a local city. They use this system to confirm what kinds of building volumes can be constructed according to planning permission ordinances. Inherently, users can choose planning restrictions as their preference in order to control the urban physical environment of the district. Therefore, this system should prepare as an easy and simple operative interface. Furthermore, users should moderate the shape parameters in order to consider the possible building volume according to land conditions of a parcel, the interface of moderating shape parameters should be prepared after inputting land conditions from a GIS database. In terms of possibilities of system utilization without restrictions of time or space, a WEB system is considered to be the optimum solution in our current project for public participation in a district plan, which is supported by the Kanazawa city government in Japan. The shape parameters can be obtained with form objects designed in HTML, and participants can choose the parameters as their preference.

In our project, a prototype system has been developed according to the above description. In a community where a district plan is discussed or decided, this system can be used as a visual tool in the office of administrative counselling division for responding to the citizens' counsel on requirements regarding the planning restrictions. This system can also allow citizens to understand what kind of building volume can be permitted on their private land parcel. Meanwhile, citizens can also access an official web page to confirm the planning restrictions and usable space of building volume in land parcels. Therefore, the functionalities offered by the system include opening planning restrictions using WEBGIS on respective land parcels and visualization of the usable space of building based-on WEBGIS using JAVA3D.In brief, this paper discusses how to visualize usable space of buildings according to various types of restrictions designated in the plan permission ordinances.

2. PLANNING PERMISSION ORDINANCES

The district plan in Japan is decided and implemented through public participation, which is based on zone restrictions of the land use system of Japan. The details of the planning restrictions of district plan are extracted from the zone restrictions.

These details of planning restrictions for the usable space of buildings are actually land use control measures in a horizontal direction and vertical direction. In the horizontal direction, the backspace of the building line from the road borderline is the main control measure. In the vertical direction, the optional oblique line restrictions including the oblique line from the front road and the oblique line from an adjacent parcel are the main control measures on a land parcel.

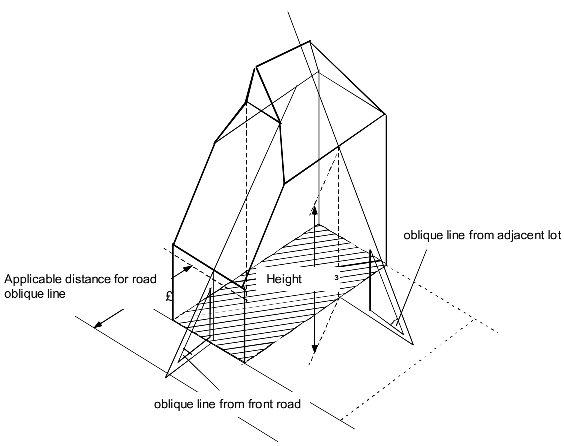


Figure 1. Usable space of building according to plan permission ordinances

Table 1. Rate of oblique line from front road in different zonings

Zoning	Floor area ratio	Applicable distance	Rate of oblique line from front road
The 1^{st} and 2^{nd} low-rise exclusive residential district, The 1^{st} and 2^{nd} mid-high rise exclusive residential district, the 1^{st} and 2^{nd} residential district, quasi-residential district	Less than 200%	20m	1.25
	More than 200%, less than 300%	25m	
	More than 300%, less than 400%	30m	
	More than 400%	35m	
Neighborhood commercial district, commercial district	Less than 400%	20m	1.5
	More than 400%, less than 600%	25m	
	More than 600%, less than 800%	30m	
	More than 800%, less than 1000%	35m	
	More than 1000%, less than 1100%	40m	
	More than 1100%, less than 1200%	45m	
	More than 1200%	50m	
Semi industrial district, industrial district, exclusive industrial district	Less than 200%	20m	
	More than 200%, less than 300%	25m	
	More than 300%, less than 400%	30m	
	More than 400%	35m	
No zoning district	Less than 200%	20m	
	More than 200%, less than 300%	25m	
	More than 300%	30m	

Oblique line from the front road is a set of restrictions, which are relative to the floor area ratio, applicable distance and rate of oblique line from the front road shown in *Table 1*. These restrictions are different in different zonings. An applicable distance is the legally required distance from the front road to an applicable position of the oblique line restriction (See *Figure 1*). The angle of the oblique line from the front road is from the counter side of the front road, which is defined in *Table 1* and has respective rate values in the respective zoning type.

Another plan restriction is the oblique line from an adjacent parcel designated in a vertical direction and it also has different values according to the different zoning types as shown in *Table 2*.

Both of the oblique line restrictions are applicable for controlling the height of buildings and the angle of oblique planes on the foreside of the usable space of buildings.

Table 2. Rate of oblique line from adjacent parcel in different zonings

Zoning	Height restriction and Rate of oblique line from Adjacent lot
The 1st and 2nd low-rise exclusive residential district, The 1st and 2nd mid-high rise exclusive residential district, the 1st and 2nd residential district, quasi-residential district	H<1.25X+20m（or,H<2.5X+31m）
Neighborhood commercial district, commercial district Semi industrial district, industrial district, exclusive industrial district	H<2.5X+31m（or,no ordinances）
No zoning district	H<1.25X+20m（or,H<2.5X+31m）

There are also other plan restrictions such as an oblique line from the parcel of north side, building coverage, floor area ratio, and restriction due to shadow and height restriction in planning system of Japan.

3. USABLE SPACE OF A BUILDING ACCORDING TO PLAN PERMISSION ORDINANCES

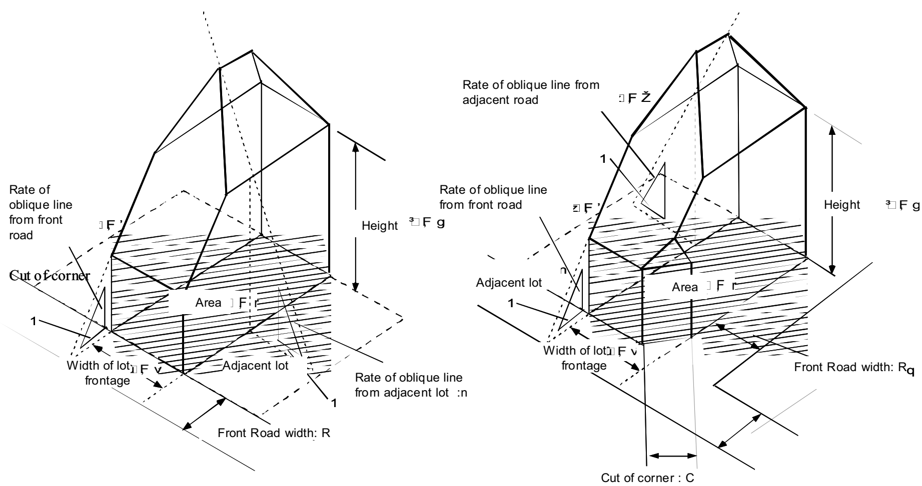
In this session, we discuss how to visualize the usable space of a building according to these planning restrictions.

A geometric model of the usable space of a building can be generated according to the zone restrictions implemented in the district plan, which include planning restrictions of high altitude, oblique line and so on. In this paper, zoning types and oblique lines from the front road and an adjacent parcel are firstly considered to be conditions for establishing a usable space of a building on a certain regular square parcel of land for developing our current support system. In addition, the current system is only applicable for the land lots in case of the 1st and 2nd low-rise exclusive residential district and 2nd mid-high rise exclusive residential district. However, other restrictions such as an oblique line from the parcel of the north side, building coverage, floor area ratio, and restriction due to shadow will be considered in the further research.

As mentioned above, the parameters that are necessary for generating a usable space of a building on certain land parcels can be divided into two parts as shown in *Table 3*. These parameters include land conditions and shape parameters. The part 1 is land conditions and the part 2 is shape parameters.

Table 3. Parameters of usable space of building

Parameters of the usable space of building	
Part 1	Land conditions
S	Lot area Scale
W	Width of lot frontage
Af	Floor Area
Y	Zoning type
D	Direction of adjacent road
R	Front road width
c	Cut of corner
Part 2	Shape parameters
r	Rate of oblique line from front road
z	Height limitation from front road
n	Rate of oblique line from adjacent road
H	Height of oblique line from adjacent road



Example with one front road Example with two front roads on a corner

Figure 2. Shape parameters on usable space of a building

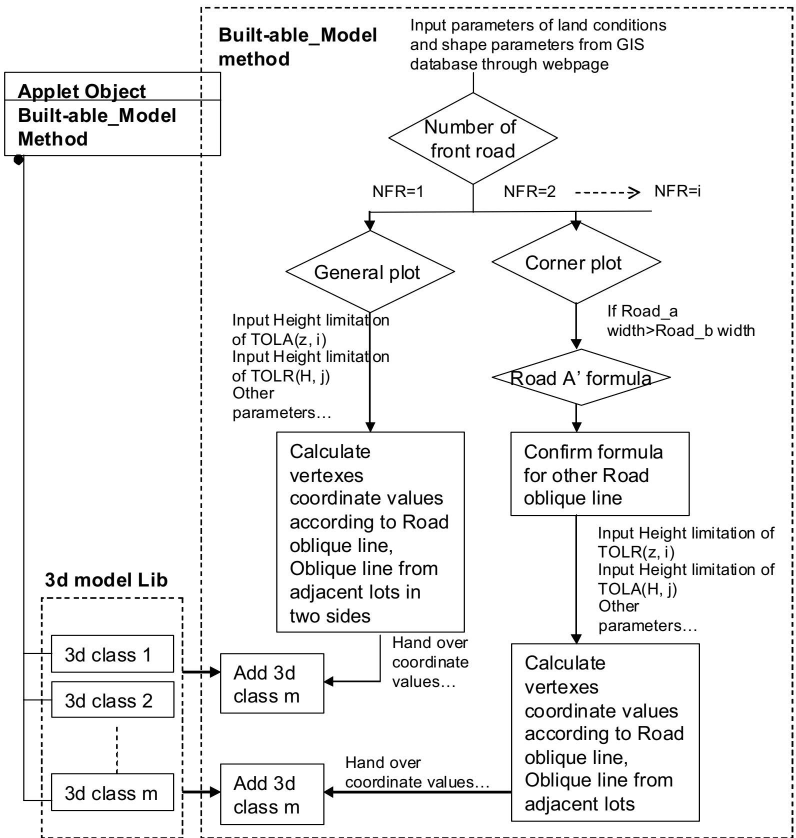
Two examples that have been generated by these parameters are shown in *Figure 2*. One of them is built up on a land parcel with one front road; another is built up on a land parcel on a corner with two front roads.

The usable space of a building generated by land conditions and shape parameters consists of many diversified types in terms of shape and volume, which has been discussed in detail in CUPUM’99 (F. Kobayashi et al., 1999). The reason why there are so many diversified usable spaces of buildings is that the width, length and other conditions of a parcel size decide the order of oblique line restrictions from roads or adjacent land lots, and that different direction of building ridge makes roof forms very diversified and complicated.

Table 4. Proposition variable used in propositional types

Typical types of shape of usable space of a building		
Proposition variable		Description
NFR(i)		Number of front roads
TOL		Types of oblique line
	TOLR(z, i)	Road oblique line (Height limitation from number i road)
	TOLA(H, j)	Oblique line from adjacent lots (Height limitation from number j adjacent lot)

Abridged description flow of generation of usable space of building



Even though there are many types of usable spaces of buildings according to the number of front roads and optional oblique line restrictions, these types can be drastically reduced if using a Boolean operation of a 3d solid shape. Actually, if the types of usable spaces of buildings are classified with proposition variables shown in *Table 4*, the usable spaces of buildings can be divided into several basic types in this case. The proposition variables are only NFR(n), TOLR(z, i) and TOLA(H, j). NFR (n) is number of front roads, TOLR(z, i) and TOLA(H, j) are proposition variables of oblique lines. H and z are height limitations applicable in oblique line, number i is road number and number j is number of adjacent land lot.

The process of generation of usable space of a building is shown in the flow graph embedded in *Table 4*, which is a method procedure in Java applet class object. The java applet input parameters including land conditions and shape parameters are from a database saved as attribution data in WEBGIS through a web browser. In the next step, the parameter of the numerical quantities of roads around the land parcel can be used to state how many oblique lines of the road will be applicable on the parcel, and the left sides of the parcel should be adapted to the oblique line of adjacent lots. Furthermore, the vertexes values of usable space of a building can be calculated according to the shape parameters and land conditions. Finally, these coordinate values can be handed over to class objects in a 3d model library that are used to be parts of the applet object. The applet object is a Java3d object combined with the objects from the 3d model library is a virtual world for expression of usable space of a building.

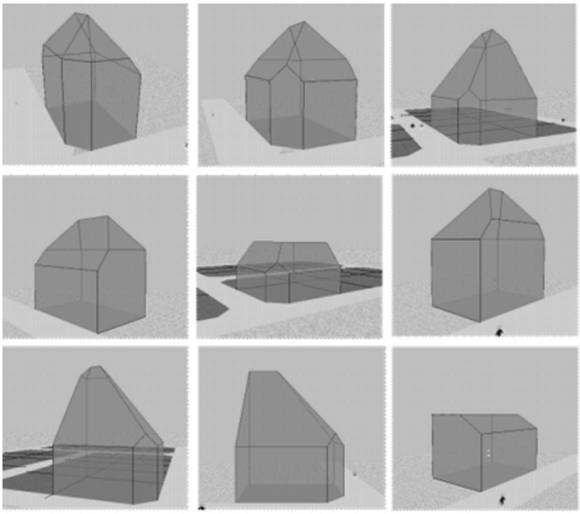


Figure 3. Auto generated types of usable space of a building according to plan ordinances

In *Table 4*, the flow graph is drawn as an abridged description of the generation of usable space of a building, in which the Java3d applet is the main class used to input land conditions and shape parameters from a GIS database. Built-able_model method of this applet object calculates the vertexes coordinate values according to respective planning restrictions, and hand over these values to the 3d model library. Class objects of the 3d model library use the coordinate values to create a 3d virtual world in Java3d applet. Typical types of shapes of usable space of a building and some of them are shown in *Figure 3*.

4. SYSTEM FRAMEWORK FOR VISUALIZATION OF USABLE SPACE OF A BUILDING

As mentioned in the introduction, a visualization tool is necessary to represent the usable space of a building in an urban environment. For developing such a visualization tool, a system framework is discussed in this part of the paper.

Our system framework is a suggestion that integrates WEBGIS with a 3d presentation tool. In our system proposal, we use Jshape as WEBGIS software and Java3d as the 3d visualization tool which are integrated on an httpd server designed for net participation, as shown as *Figure 4*.

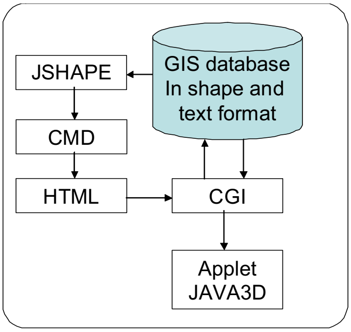


Figure 4. Development tools in the system

For generation of an image of usable space of a building, this system inputs coordinate information of the parcels and their attribution data from WEBGIS. These land conditions information can be handed over to form objects on the homepage, and next to be integrated as Java's parameters in html tag with planning restrictions in JAVA3d for the generation of a 3d virtual world.

As shown in *Figure 5*, shape parameters are prepared in the form of an object in html and land condition items can be input from a GIS database through the CMD script of JSHAPE. For the last step of 3d visualization, the parameters of land conditions and shape parameters can be input from a form object as Java's parameters into a class of applet using a JAVA3d program. In the applet of Java, built-able_model method can be used to generate a usable space of a building and building_model method can be used to generate a building model for reference.

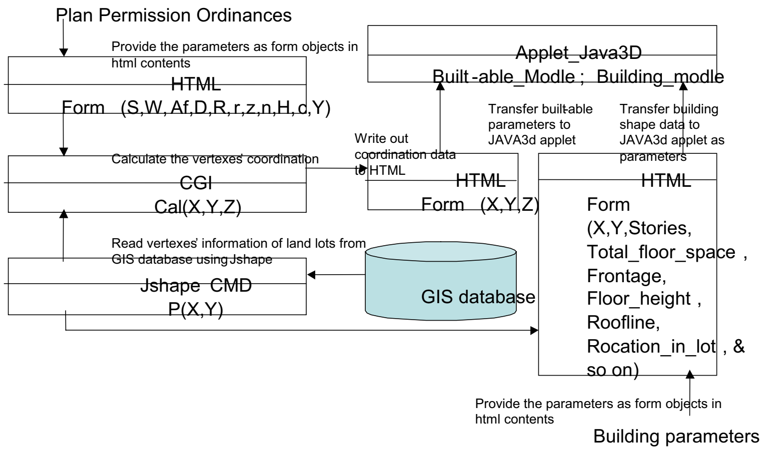


Figure 5. Plan permission ordinances and 3d visualization

5. POSSIBILITIES IN DISTRICT PLAN IN THE FUTURE

As mentioned above, if there are no visualization tools to represent the usable space of a building in an urban environment, the alternative district plans are difficult for local residents to imagine. Consequently, they have to make a planning decision without a concrete image of the future of their community. For making more consensual planning decisions, a visualization tool is necessary to represent the future image of the urban environment proposed in alternative plans. The 3d visualization tool can be used as a kind of visualized method for making planning decisions in consensus meetings for alternative district plans. However, at this stage, only single usable spaces of a building can be represented and the visualization of all the community will be the next aim of our project.

The users of this system are considered to be the administrative officers and citizens from a local city who use this system to confirm what kinds of building shape and volume can be constructed according to planning permission ordinances. Inherently, users can choose their preference planning restrictions in order to control the urban physical environment of the district. In this part of the paper, we discuss the user-friendliness of our project. This system should be prepared as an easy and simple operative interface.

5.1 Land Conditions Input from a GIS Database Using Jshape

Figure 6 shows a map created in ARCGIS and saved in shape file. Jshape (http://www.jshape.com) can be used to open the shape data to the Internet environment as shown in *Figure 7*. The land conditions and shape parameters relative to each land parcel are saved as attribution data of parcel polygons.

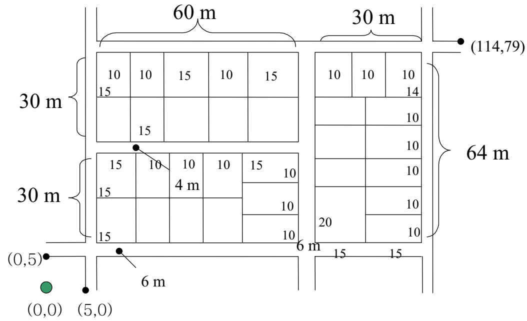


Figure 6. A map in WEBGIS(Jshape)

5.2 Visualization of Planning Restrictions as Usable Space of a Building by JAVA3d

Figure 7 shows the WEBGIS interface for collecting attribution data from a GIS database. The mapview component and map control are arranged in the above part in *Figure 7* in order to show map and legend, which is a customized interface of Jshape. In the section below, a list of attribution data that links to a polygon will be shown if the user clicks on the polygon using the mouse in the mapview component.

If users click the show button in *Figure 7*, a set of homepage contents for inputting shape parameters of planning restrictions can be launched for the next operation. After the process of inputting shape parameters, a button which is designed to launch the Java3d applet can create the usable space of a building which is compatible with the land conditions and shape parameters. *Figure 8* shows the usable space of a building based on the land conditions and shape parameters which are relative to the identified land parcel. Users can walk through the 3d world by moving the mouse and are able to change viewpoints by pushing the buttons on the left of the site shown in *Figure 8*.

The button designed on the below left of *Figure 8* is used to launch a tool to model a building on the land parcel which is shown in the same applet.

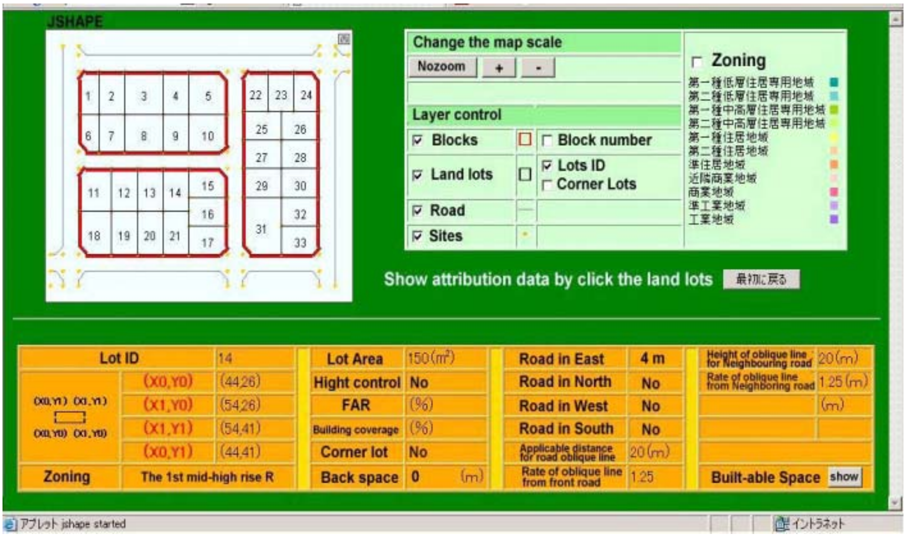


Figure 7. WEBGIS and attribution data of land lots

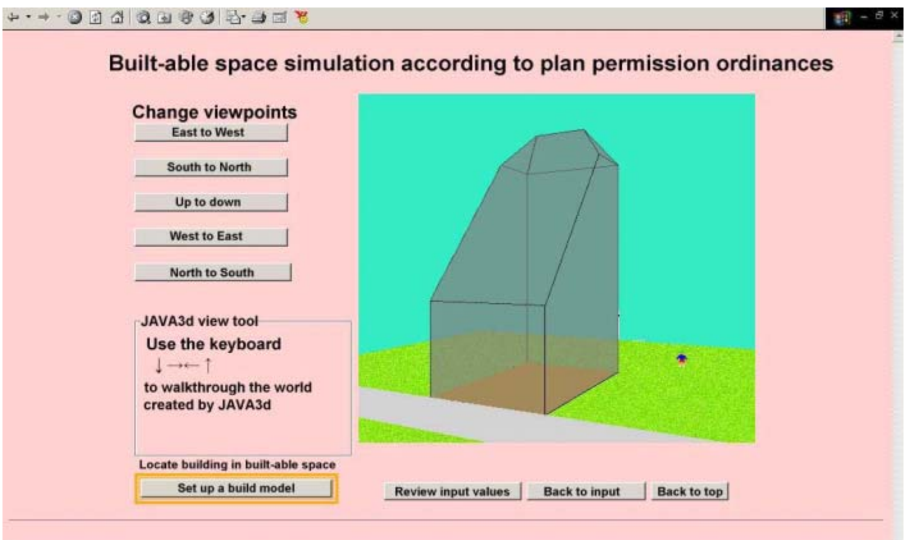


Figure 8. Simulation of the usable space of a building by JAVA3d

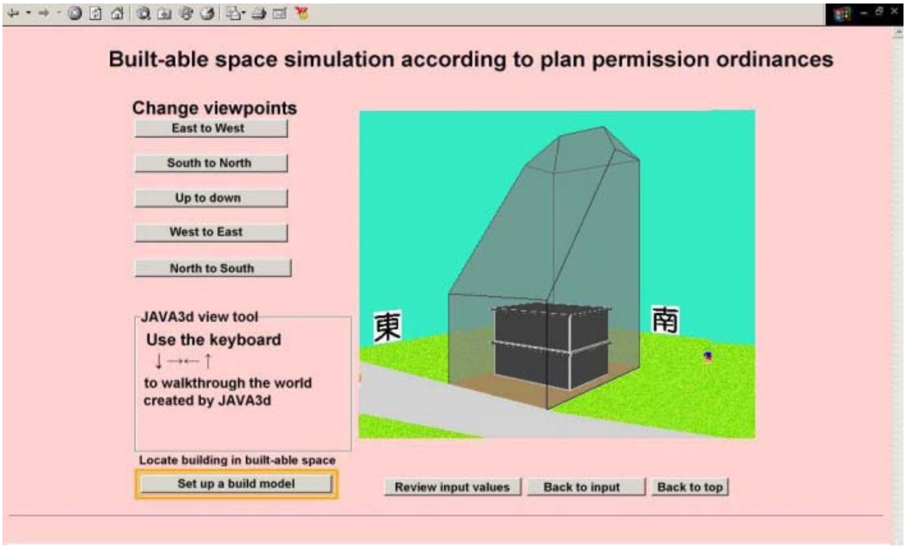


Figure 9. Simulation of the usable space of a building and building model using JAVA3d

5.3 Simulation of Building Model

The user can set up a building model by pushing the button in *Figure 9*. The necessary parameters regarding the building model can be input in the form of objects and transferred to Java3d to establish a building model that is relative to the usable space of the building generated in *Figure 8*. These parameters include coordinate information of relative land parcel, stories, total floor area, frontage, floor height and roofline, which are the basic pieces of information required for the building model.

This function will be useful for users to understand that what kinds of buildings are compatible with the usable space of a building in plan restrictions. Because the building model function is still too simple to depict a usual building, further research for providing an effective representation is still on the way.

6. CONCLUSIONS

For generating usable space of a building using a computer system, parameters include land conditions of a parcel, shape parameters of planning restrictions. Land conditions implicate vertexes coordination data of parcel polygons and the road situation of the surrounding parcels, which can be saved in a GIS database. In addition, the shape parameter indicates optional values for visualization of the usable space of a building according to legal requirements, which can be combined with land conditions in system.

The visualization of usable space of a building according to planning permission ordinances can be auto generated using WEBGIS and JAVA3d. A system framework has been suggested in this paper. The usable space of building generation has been discussed in the case of the 1st and 2nd low-rise exclusive residential district and 2nd mid-high rise exclusive residential district.

In the future, further research regarding the building model and other zoning types should be continued, and a virtual image of the total community within a region is also necessary for obtaining consensus for a district plan. In the future, we would like to conduct an Internet social experiment with the cooperation of a city government and residents in order to support the district plan in a local city.

ACKNOWLEDGEMENT

This project is supported by the Kanazawa city government.

REFERENCES

Daniel Bulmer, Victoria University of Manchester, UK. 2002, "How can computer simulated visualizations of the built environment facilitate better public participation in the design process?" *On-line design Journal*, http://www.onlineplanning.org/

Doyle, S., M. Dodge and A. Smith, 1998, "The potential of web based mapping and virtual reality technologies for modelling urban environment", *Computers, Environments and Urban Systems* Vol.22(1), pp.37-55.

Kobayashi, F., M. Kawakami and H. Nogawa, 1999, "Modelling of usable space for building and evaluation of regulation easement on building form", *Proceedings of CUPUM'99* on CD-ROM.

Michael B., David, C., Steve E., Scott F., Mordechai H., Stefan K., Naru S., Andy S., and Paul M.T. 2001, "Visualizing the city: Communicating urban design to designers and design makers", USA: ESRI Press, Redlands, California, pp.405-443.

Sidjanin, P. 2000, "A conceptual model of the design tool for analysis and visual quality control of urban environment", In H. Timmermans, & B. de Vries, *Proceeding of DDSS '00, Design & Decision Support Systems in Urban Design,* Nijkerk, The Netherlands, pp.229-240.

A Comparison of 3D Visualization Technologies and their User Interfaces with Data Specific to Architecture

Roland Göttig, Joanna Newton, and Stefan Kaufmann
Technische Universität München, Germany

Keywords: 3D-Systems, Virtual Reality, Powerwall, Holobench, HMD, CAVE, User Interfaces, Visual Display Qualities, Design Process.

Abstract: Contemporary advanced virtual reality systems use different stereoscopic 3D visualization technologies. In this study, VR systems from one projection wall up to VR systems with six projection walls have been evaluated. Besides the optical properties tested with architectural 3D models, the user interfaces have been analyzed with reference to exact and intuitive control abilities. Additionally, the workflow of an early architectural design process with CAAD generated 3D models and VR visualization techniques was analyzed. It turns out that current VR systems exhibit shortcomings in visual and spatial representations, as well as tools for an early design process.

1. INTRODUCTION

A computer simulated, immersive, and interactive environment can be set up in various different configurations ranging from desktop virtual reality (VR), through partial and fully immersive virtual environments to augmented reality and telepresence. Desktop virtual reality is a low-cost solution, usually non-immersive, with a conventional computer monitor as the output device for a rendered three-dimensional environment. Partial-immersive systems are normally projection-based, such as so-called powerwalls or workbenches. They display a depth effect in the field of vision that is covered by the projection area. Thus, they offer an intensified sense of presence, as well as greater appreciation of scale than desktop installations. Nevertheless, the most compelling experience of virtual reality

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 99-111.

is conveyed by fully immersive VR systems, known as head mounted displays ‘HMDs’ or cave automatic virtual environments ‘CAVEs’ (Types of VR Systems, 2003). Immersive virtual reality utilizes specific display technologies and tracking devices in order to create virtual presence, a sensation that “Is experienced by a person when sensory information generated only by and within a computer compels a feeling of being present in an environment other than the one the person is actually in.” (Sheridan, 1992) The interaction devices of these VR systems vary from combinations of typical 2D mice and keyboards over so-called spacemice up to 3D pointers and data gloves (*Figure 1*).

Figure 1. CAVE application and data glove

New developments in telecommunication and computerized media technology of the late twentieth century had an enormous influence on the process of drawing and drafting, and on the dependent design professions. Computer aided architectural design (CAAD) enabled designers to gain more control over the realization of their proposals, as the computer was a fast and accurate medium for three-dimensional modeling, and thus, a valuable method for representing detailed propositions. Hence, computerization has reorganized and redefined the nature of architectural design and even that of the architectural office itself (Campbell, 1994). Nevertheless, determining strategies for architectural design is a complicated procedure, realizing them being even more difficult. Different solutions were proposed to cope with this situation, mostly constituting an aid for the architect in one way or another, but not dealing with the principal questions and intrinsic problems of CAAD. As a result, the early design phase is still dominated by the traditional design tools and doesn't show any attempt of shifting towards a more computerized design process (Kurmann, 1998). At the early conceptual design phase, the architect’s work is essentially based on ambiguity and vagueness. In contrast, commands in CAD systems always expect exact values.

The difficult use of computer interfaces compared to the intuitive use of a pen comprises another vital disadvantage in CAD systems. Additionally, the selection of new commands and the input of exact values distracts from the design task. Furthermore, most commands in CAD systems are based on precise dimensioning, which almost freezes the creativity of an inexperienced user. The fixed scale is an additional problematic issue, as it is missing in computer drawings, and thus complicates the judgment process in respect to the verification of the functional and aesthetic correctness of a design (Kulinski, 2003).

2. PROJECT FRAMEWORK

A major advantage of computer based visualization of 3D models is that it offers the possibility to enhance the imagination, comprehension, and evaluation of models or concepts, which are otherwise difficult to capture. It provides a highly informative overview, strengthened by the interaction with the simulation, which in turn enables the conceptualization of relations that are not apparent from a less dynamic representation (Weiss, Jessel, 1998).

Another major advantage of virtual reality is the increase of engagement that can be reached, as the control of the virtual environment can be put down to natural movements of the user's hand or head (Smets et al., 1994). Since the interface is directly linked to the natural way of examining real environments, it can be intuitively used for the exploration of virtual environments. The users are able to look and move around as in reality, which minimizes the danger of misconception inherent in the bird's eye view perspectives that scale models and computer generated images often give (Henry, 1992).

For architects drawing sketches and building physical models seems to be the natural way to start a new design process. At that point the use of any computer based technology often leads to misunderstanding or even refusal among traditionally orientated architects. The reason for that is most likely the exactness of virtual reality models – they look finished at a period of time when typically most changes in an architectural design take place. Difficulties in giving hints or corrections appear, because basic knowledge in 3D modeling and interaction with the VR system is needed.

3. PROJECT DESCRIPTION

To date, a great number of virtual reality research has been carried out, which explored many problematic issues concerning the technology. But while most studies were focused on special technical issues, or were rather general, reviewing a wide range of application areas, little work was done to compare these systems in respect to their state-of-art in the architectural domain. It is also important to note that the illusion of Virtual Reality is based on a great deal of unspoken beliefs, which make it necessary to critically review the potential and relevance of the use of virtual reality within the architectural domain.

In this study four categories of VR systems have been examined, in order to provide a cross-section of contemporary systems: A ‘powerwall’, a ‘holobench’, a ‘head-mounted display’ and a ‘cave automatic virtual environment’.

Figure 2. (left) Powerwall set-up at the Technische Universität München (TU Munich), (right) Holobench set-up at the Leibniz Rechenzentrum Munich

The powerwall set-up at the TU Munich is an individual construction of three connected 1.4GHz PCs with GeForce3 graphic cards and a spacemouse as control interface. Two beamers equipped with polarization filters project the computer generated pictures on a special screen, which preserves the polarization direction. To get a true three dimensional impression, an observer needs polarization glasses that select the pictures for each eye (*Figure 2 (left)*).

The holobench set-up at the Leibniz Rechnezentrum in Munich was acquired as a complete package, which utilizes a SGI Onyx2 Workstation (4 processors) with an InfiniteReality2 graphic card (1 graphic pipe). It is equipped with two projection screens. The beamers alternately produce one picture for the right and one picture for the left eye at a frequency of 48Hz. So-called shutter glasses darken alternately the view through the glasses on each side likewise at 48Hz. Here, a pointer device is used as control and interaction interface (*Figure 2 (right)*).

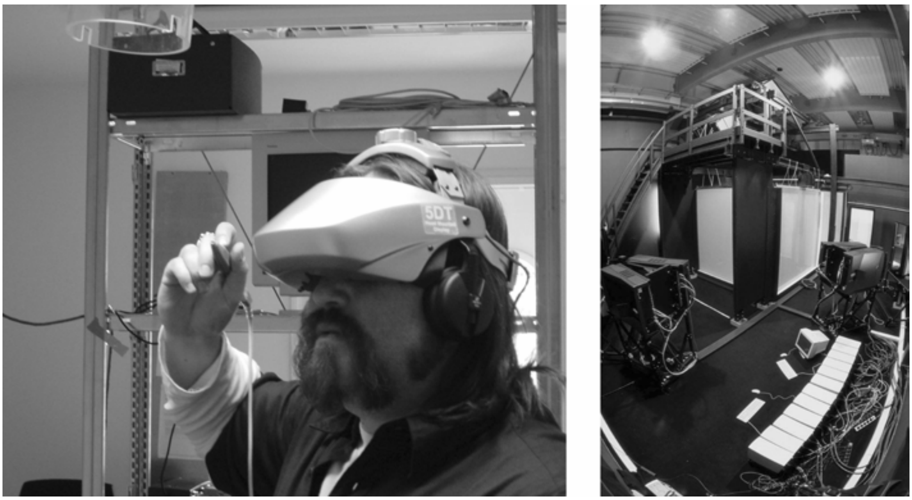


Figure 3. (left) HMD set-up at the Bauhaus Universität Weimar, (right) CAVE set-up at the IAO Stuttgart of the Fraunhofer Gesellschaft

The HMD set-up at the Bauhaus University in Weimar is also an individual construction, based on a 1.9GHz Dual Athlon PC with a GeForceDDR graphic card. Furthermore, its visual device consists of two small monitors close to the eyes and a lens system for an optical correction. The same pointer device as in the holobench set-up is used (*Figure 3 (left)*).

At the IAO in Stuttgart the most sophisticated VR system examined in this study was build up. It was bought as a complete package including a SGI Onyx3 Workstation (8 processors) with an InfiniteReality3 graphic card (6 graphic pipes). The observer of a virtual scene is located in a room of 2.9m x 2.9m x 2.9m size with six projection walls around. Beamers with alternating pictures for each eye and shutter glasses are used. The control and interaction device was developed by the IAO and is non-commercial. It is similar to the pointer devices of the holobench and HMD set-ups (*Figure 3 (right)*).

The first part of the examinations deals with the basic properties of the VR systems. They can be outlined as follows: Visual display qualities (VDQ), spatial impression, and navigation. Therefore, three predefined models of varying complexity have been used. Details about these models are shown in *Table 1*. The simplest model shows a spatial design based on a number of tilted planes with cut-in holes. The second model could be described as semi-complex. It shows two buildings with a higher level of detail, such as visible construction elements and staircases, while the surroundings were kept rather abstract. The third model offers a complex scene of an exhibition hangar. It comprises a very detailed depiction of steel construction elements as well as people, trees and air-planes.

Table 1. Predefined models, general information

	Abstract Model	Semi-complex Model	Complex Model
Polygon Number	1 295	23 696	226 760
Lights	1	3	5
Textures	-	-	18
File Size (VRML 1)	55 kB	1 043 kB	17 414 kB

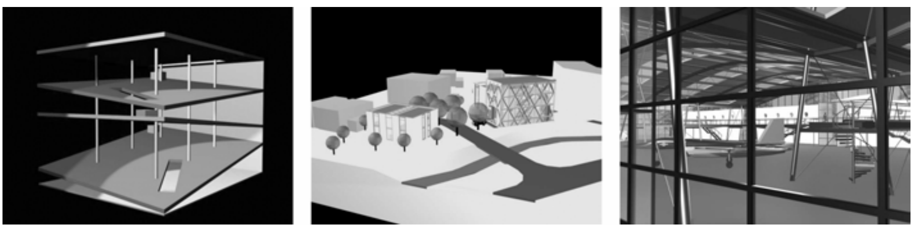


Figure 4. (left) Abstract model, (middle) Semi-complex model, (left) Complex model

Due to different software capabilities in importing 3D data, VRML 1 and VRML 2 were chosen, because all systems were able to read at least one of these formats. Moreover, import filters for file formats with a wider range of possibilities in defining surfaces, lights or animations, like 3DS or MAX, are missing in the examined VR systems.

The goal of the second part of the study is to reveal the impact of different VR systems on the creativity of a designer and the way he communicates his proposal. In addition, the aim is to describe a possibility of using these systems for the conceptual design phase, in contrast to the traditional design approach. The subject of that task was set in relation to the abstract architectural model used in the first part. It was the result of a competition for the rearrangement of the main interior space of the German pavilion at the architectural biennale in Venice in 2002. The theme was formulated as ‘Next’, addressing the next generation of architects, who should analyze the obvious, the essence of architecture, and formulate a thematic interpretation (*Figure 5*).

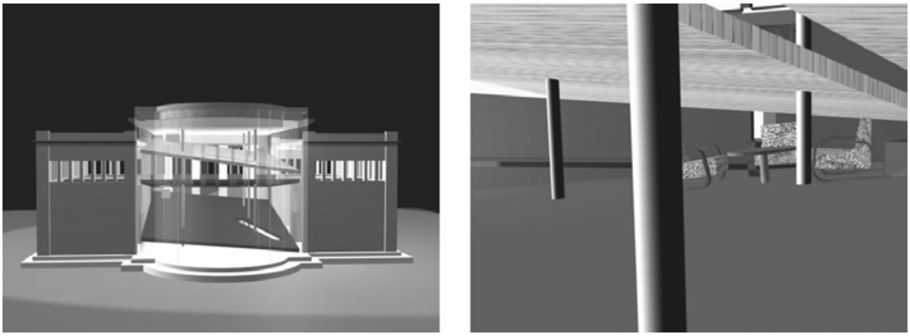


Figure 5. (both images) Experimental design for the main interior space of the German pavilion in Venice

The tests have been carried out with questionnaires. For that task, sixteen probationers were randomly assembled in teams of two. Firstly, the VR systems were examined on issues like spatial impression, visual display qualities and ease of use of the 3D interaction devices. Then the probationers had to develop a design for the main space of the pavilion as described above, starting only with CAD and VR systems, and then only with sketches, cardboard and glue.

4. SYSTEM COMPARISON AND EVALUATION

High quality visual representations are crucial for a realistic experience of virtual environments. They should guarantee the perception of stereoscopic imagery and be capable to track the varying positions of the head, thereby constantly updating the images and revealing the performed movements in the virtual reality. Basically, the viewer should receive colored visual stimuli of acceptable resolution, an adequate brightness as well as motion representations of high quality (Gobetti, Scateni, 1998). Unfortunately, the reality turns out to be different as shown in *Figure 6*. One main reason for the poor estimations might be a too simple light and surface representation. Because of the need to calculate new images from new viewpoints in a fast manner, extensive render methods are not used. *Table 2* illustrates an overview of information relating to the visual display qualities inherent in the tested VR systems.

Table 2. Visual display qualities, basic technical data

	Powerwall	Holobench	HMD	CAVE
Resolution	1024 x 768	1280 x 1024	800 x 600	1024 x 1024
Projection Screens	1	2	2	6
Display Size	2.4m x 1.8m	1.8m x 1.1m	0.90m x 0.67m (optical corrected)	2.9m x 2.9m
Refresh Rate	LCD	48Hz	LCD	60Hz
Color Depth	32 bit	24 bit	32 bit	24 bit

Moreover, a major factor that accounts for display quality and thus for visualization of computer graphics is the rate at which those can be produced, also known as frame rate. As the virtual world is interactive, and therefore the position of the user and the viewed parts are unknown, the computer images can't be created in advance, but have to be produced in real-time as they are being shown to the observer. The frame rate is expressed in terms of frames per second (fps), with higher values leading to smoother graphics and lower values leading to images that appear jerky. For smooth simulations at least 24 (the frame rate of movies) or better 30 fps need to be displayed (Burdea, Coiffet, 1994). Due to viewpoint changes of average speed, a range of frame rates was determined with the VR systems depending on the complexity of the models (*Table 3*).

Table 3. Frame rates depending on the predefined architectural models

	Powerwall	Holobench	HMD	CAVE
Abstract Model	38-70 fps	15-48 fps	8-16 fps	27-36 fps
Semi-complex M.	29-70 fps	6-16fps	5-8 fps	19-25 fps
Complex Model	15-51 fps	< 1fps	not displayed	7-13 fps

At least two of the four tested VR systems generate satisfactory frame rates (powerwall and CAVE set-up) when abstract or semi-complex models were displayed. However, none of the systems fulfill the requirements of a fluent real-time presentation with a complex architectural presentation model.

The first impressions of the probationers being 'in' the virtual world varied from confusion to fascination. Typically, after a short period of time - less than five minutes - they got used to the new environment and reported about their spatial impressions. The highest values that indicate an acceptable immersion into the virtual world were evaluated for the holobench and CAVE set-ups (*Figure 6*).

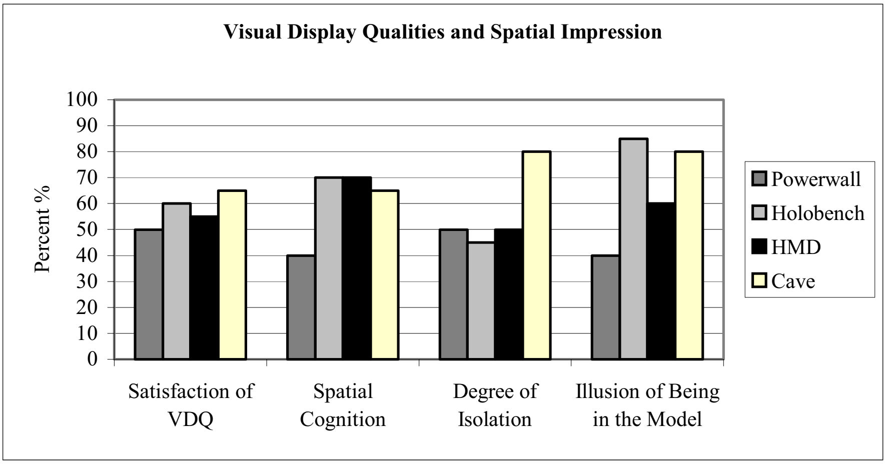


Figure 6. Visual display qualities and spatial impressions of the tested VR systems

Today, the interaction with a computer mostly occurs in a two-dimensional environment, including mouse and keyboard inputs. In contrast, all tested control interfaces of the VR systems are 3D devices with six degrees of freedom. The input device of the powerwall set-up is a spacemouse, it shows similarities with common 2D computer mice. However, besides lateral movements pushes, pulls, and twits can be performed. For the probationers it turned out to be quite difficult to learn how to utilize this device efficiently. More trained persons do not seem to face these kinds of difficulties. Even more, this piece of equipment was elected as the best 3D input device in a magazine for professional CAD users in 2003 (CADCAM report, 2003).

The holobench and the HMD set-ups were equipped with the same stylus devices. The user may hold them in one hand, and points to the direction he wants to look at. For movements the button on top of it must be pressed. Unfortunately, backward movements were not supported by the HMD set-up. To reach a starting point, one has to go forward and then turn around. The interaction device of the CAVE set-up works similarly to the stylus system. It is a self-developed tool of the IAO in Stuttgart called 'hornet' without limitations in any direction. This mechanism turned out to be the most manageable one, as documented in *Figure 8*. Beside issues concerning the movements, some problems, such as to keep the plane, and the ease of collision with objects in the virtual world were evaluated.

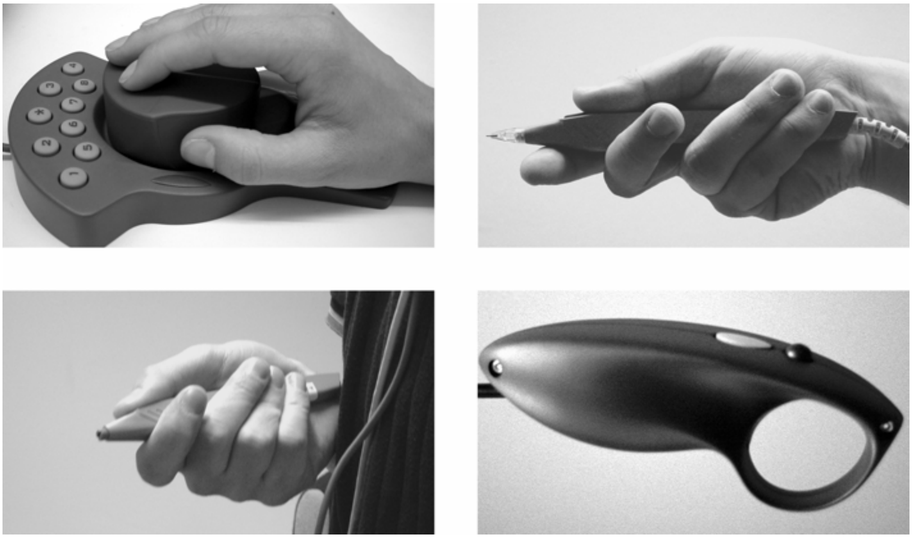


Figure 7. Interaction devices: (top left) Spacemouse (top right) Stylus, holobench set-up (bottom left) Stylus, HMD set-up (bottom right) Hornet

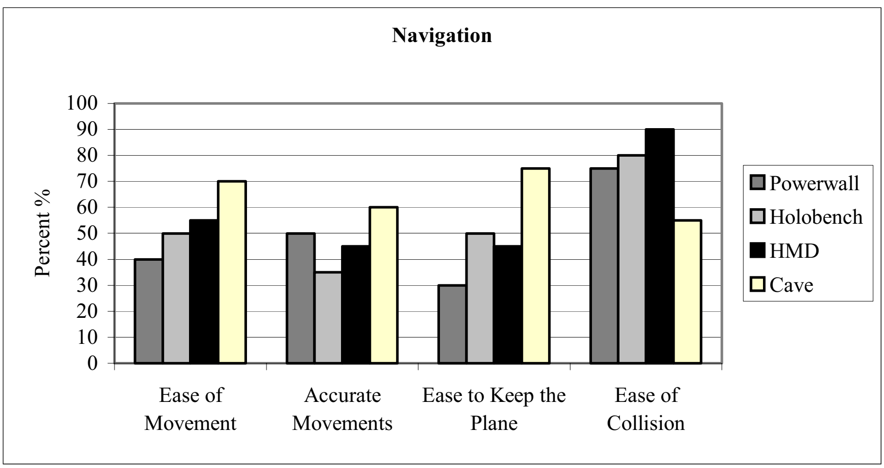


Figure 8. Navigation in the VR systems

Finally, the test persons had to compare their traditional way to design with the use of CAD and VR systems. Therefore, it was necessary to choose a simple task which made it possible to develop a spatial design in a short period of time with just a few architectural building parts like pillars, walls, and ceilings. The probationers were told to find their own solution for the main hall of the German pavilion in Venice as shown in *Figure 5*. It was essential that they had to begin with a complete new design to guarantee a comparison of the early architectural design process in a traditional way with sketches and physical models and with CAD-models and VR systems.

It was thought that the spatial impression of a VR system, and the ability to walk through a virtual space with dimensions similar to reality, could be very helpful for that early phase. The probationers had to keep in mind that their traditional way of designing was set on a level of 100%. Yet all of the questionnaires exhibited values below 100%, indicating that none of the test persons characterized the use of CAD and VR systems superior to the traditional tools (*Figure 9*). Mainly, because design on CAD systems with exact drawings instead of the fast and abstract way of sketching was reported to be baffling. But also time-consuming necessities like importing data from the CAD computer system in the VR system hindered an overall fluent process.

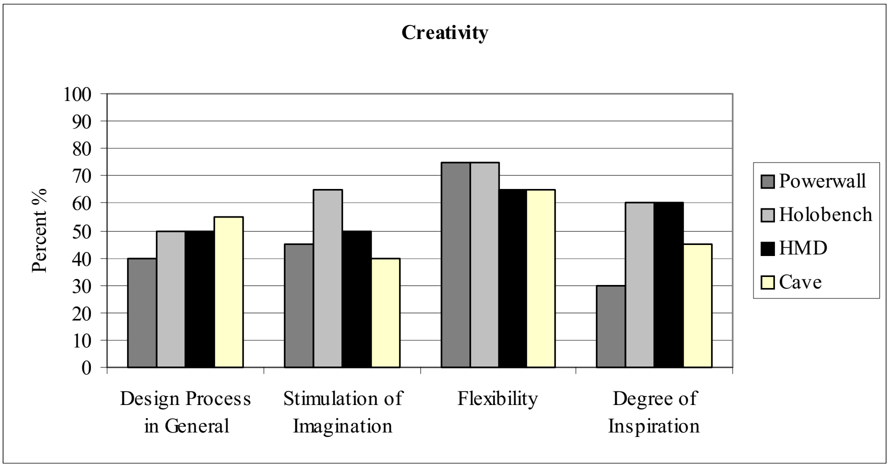


Figure 9. Creativity using VR systems

However, some positive aspects must be annotated. The most evident advantage of a VR system as design tool is the possibility to move inside a virtual world and experience its spatial impression. Furthermore, simple changes of a basic design, like variations in size or the repetition of several building parts, can be performed more easily than with cardboard models. Additionally, other observers than architects e.g. stress analysts or building owners, may find visualization in a true scale very helpful.

5. CONCLUSION

Contemporary VR systems have the capability to display semi-complex architectural models in an adequate manner. In contrast, presentation models, which display more complex details, typically used for rendered images or animation movies, cannot be presented with sufficient frame rates. Another disadvantage is the reduced quality of the displayed images in terms of natural lighting and surface representation. Furthermore, navigation of the user within the virtual world is a critical issue. Some new devices like the 'hornet' turned out to be easily and intuitively utilizable, especially in combination with a fast VR system. Otherwise disorientation could affect an observer.

However, Moore's law is still unbroken, which describes that approximately every 18 months the computer capabilities double (Moore, 1965). Thus, acceptable frame rates even for complex models will be realized in the near future. Even more, the display quality will improve: Modern non-stereoscopic but interactive virtual scenes, as shown in *Figure 10*, allow smooth navigations even with highly detailed models and complex light simulations.

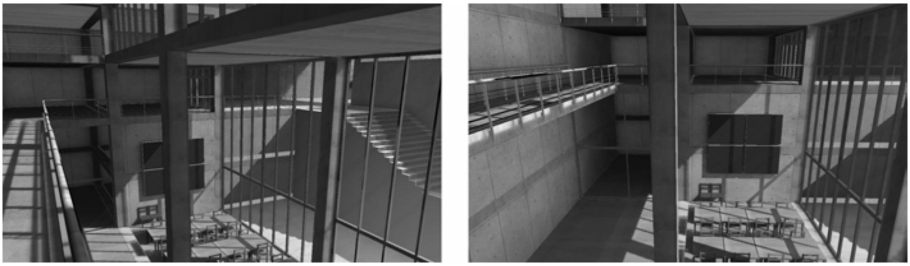


Figure 10. Highly detailed VR scene (Osmosis, 2003)

For an architectural design process the low speed of the entire procedure, as well as the difficulties in translating spatial and conceptual thoughts in precise CAAD models, leads to a restricted usability of the VR systems. Skilled designers do not seem to benefit from true-scale impressions of planned objects. Due to their education they are familiar with spatial imagination, and often do not need this kind of presentation. Nevertheless, the possibility to quickly alter CAAD models was reported to be quite positive, especially for making changes after a basic concept was found.

Several major improvements could be introduced by the integration of enhanced interaction devices, such as digitizer tablets with pressure-sensitive pens for the basic design, and easy to use software interfaces for a better 3D data exchange. After that, a workflow with a higher flexibility could be established.

REFERENCES

Burdea, G., P. Coiffet, 1994, *Virtual Reality Technology*, John Wiley & Sons, New York, USA.

CADCAM report, 2003, "Innovator 2003", *CADCAM report*, 22(7), p. 12-19.

Campbell, D., 1994, "Vers Une Architecture Virtuelle", University of Washington. Seattle, *http://www.hitl.washington.edu/people/dace/portfoli/*

Gobbetti, E., R. Scateni, 1998. "Virtual Reality: Past, Present and Future." *Virtual Environments in Clinical Psychology and Neuro-science*. Amsterdam. Netherlands *http://www.crs4.it/vic/data/papers/*

Henry, D., 1992, "Spatial Perception in Virtual Environments:Evaluating an Architectural Application.", *http://www.cs.utexas.edu/users/graphics/seminar99/vr_displays/*

Kulinski, J., 2003, "An Inspiring Method of Teaching CAAD Programs." *http://www.arch.usyd.edu.au/~jarek/research/*

Kürcherer, K., 2003, "Angereicherte Wirklichkeit, Augmented Reality applications at the industry ", *c't*, 16, p. 80-82.

Kurmann, D., 1998, "Sculptor – How to design space?", in: Sasada, et al. (eds) Proceedings of CAADRIA 1998, Osaka, Japan, p. 317-325

Moore, G.E., 1965, "Cramming more components onto integrated circuits", *Electronics*, 38(8).

Osmosis, 2003, "Examples", *http://www.osmosis.com.au*

Sheridan, T., 1992. "Musings on Telepresence and Virtual Presence." *Presence: Teleoperators and Virtual Environment*, 1(1), p. 120–125

Smets, G., et al., 1994, "Designing in Virtual Reality: Implementing Perception-Action Coupling with Affordances.", in: Feiner, et al. (eds) Proceedings of VRST 1994, Singapore, p. 97-110.

Types of VR Systems, 2003, Advisory Group of Computer Graphics, *http://www.agocg.ac.uk/reports/virtual/37/chapter2*

Sheridan, T., 1992. "Musings on Telepresence and Virtual Presence." *Presence: Teleoperators and Virtual Environment*, 1(1), p. 120–125

Weiss, P., A. Jessel, 1998, "Virtual Reality Applications to Work.", *WORK*, 11(3), p. 277-293, *http://www.utoronto.ca/atrc/rd/library/papers/weiss.html*

Color Your Feeling

Design Support System for Color Coordination

Ji-Hyun Lee and Wei Qian
Graduate School of Computational Design, National Yunlin University of Science & Technology, Taiwan (R.O.C.)

Keywords: Design Support System, Sensibility Ergonomics, Color Coordination, Color Image.

Abstract: Color selection plays a vitally important role in creating impressions of individuals or companies because colors have sensibility aspects and relate to some images or associations. Based on both the theory of color harmony and the sensibility ergonomics, some quantitative and systematic researches on the color image have been developed. In this paper, we suggest a color coordinate system that supports the color analysis and the color harmony functions using color images, which can be captured by corresponding adjective words. We focus on a system prototype for interior design domain to exemplify our concepts in this paper, even though this system can be applied for all design domains.

1. INTRODUCTION

Much of our perception of physical world includes our identifying objects by vision, and colors are typically essential to our vision. Because colors have perceiver-dependent properties, it seems to be considered as subjectivism. However, the concept of color harmony has to be removed from the realm of subjective attitude into that of objective principle (Itten, 1970). For example, most people will think red color represents hot feeling and blue color represents cold feeling. This fact reveals that most people have common sense for color, and colors also have objective essences. Moreover, based on both the theory of color harmony evolved from the nineteenth century and the sensibility ergonomics, some quantitative and

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 113-125.

systematic researches on the color image have been developed in many objective ways.

From 1990s the rapid spread of color monitor as an output device and advancement of digital technology allow designers to move paper-based traditional research to computer-aided design (CAD) area. Current software related to color provides color palette function to select desired colors easily. Some of the simple color picker software are built into the operation systems. Sometimes software provides different color palette system such as RGB, HSB, CMYK, etc. However, the major issue in handling color is the huge amount of data, as there might be more than ten million kinds of colors. When designers want to coordinate some colors with one or another color, they should deal with delicate variations in a lot of data so that the reasons often cause that the result of color coordination depends on designers' training and ability or push designers for referring and comparing with hundreds of colors and spending a lot of time. Therefore, just providing traditional color picker is not intelligent enough to provide designers with faster or more efficient color coordination.

This paper proposes a color coordination system that supports designers by matching their sensibilities with a group of colors in the color coordinating process. The system integrates color analysis sub system with color harmony sub system to select the type of color harmony or inspect exhaustive alternatives to find out the best match with the designers' intention.

This paper has the following tasks to satisfy these objectives:

1. Get a color image of a given picture from the color analysis knowledge base and the color image database.
2. Generate a set of colors matched with the color image.
3. Select a color from the set of colors and a type of color harmony.
4. Generate harmonious colors from the color harmony knowledge base.
5. Display the color coordination to check if the result is satisfied or not.

In this paper, we focus on an interior design system prototype to exemplify our concepts in this paper, albeit this system can be applied for all kinds of design products.

2. HUMAN-COMPUTER SYSTEMS DESIGN USING COGNITIVE MODEL

Among the phases of information presented from computer to users, the phase related to color or shape is users' **cognitive perception** (Ito and Nakakoji, 1996). *Cognitive models* are those models which represent the human being's sensory and cerebral processing system, his characteristics and limitations related to the elements of that system, and the outcome of such processes (Hasdogan, 1996). Cognitive models in HCI point of view is **task-based** models because these models were developed based on the emphasis on people interacting with computers and the dominant information processing paradigm of cognitive psychology (Benyon, 2004).

To develop effective human-computer systems, there are several levels of abstractions at which we want to describe the system in HCI point of view; the organizational level, the workplace level and the operational level (Benyon and Imaz, 1999):

- **Organizational level**: to consider the design of working practices and the impact of new technologies on organizations
- **Workplace level**: to design for the activities which users will undertake, at the structures and artifacts which enable people to pursue the goal
- **Operational level**: to look at how knowledge is constructed, represented or how it is distributed through the system, or consider the usability of system and the cognitive and emotional demands that systems make on people

We describe below each level of cognitive models in greater detail.

3. ORGANIZATIONAL LEVEL

As mentioned in Chapter 2, cognitive models represent the human being's sensory. The designer's sensibilities are represented using color images, which can be captured by corresponding adjective words.

3.1 Sensibility Ergonomics

Sensibility ergonomics can be defined as follows; it is a kind of engineering integrated emotional agreement into the human-computer interface (HCI) design. In the United States, the term **human factors engineering** is often used, while the term **Kansei engineering** is used in Japan. The technology of sensibility ergonomics can be used to understand the relationship between the factors of designers' feeling and the design elements. Because of the fuzzy characteristic of an image, using color language is the most convenient method to select the color, which can be agreed by most of people (Nagumo, 2000). As mentioned in the previous section, most people have a common image for a group of colors, therefore, we use color images to represent the colors.

3.2 Color Image

Nagumo (2000) classifies 160 color images captured by adjective words, each of which has 9 to 24 colors. He uses x-y coordinate system to categorize thousands of colors using adjectives. The vertical axis represents *time* and the horizontal axis represents *energy*. Figure 1 illustrates the image scale for color.

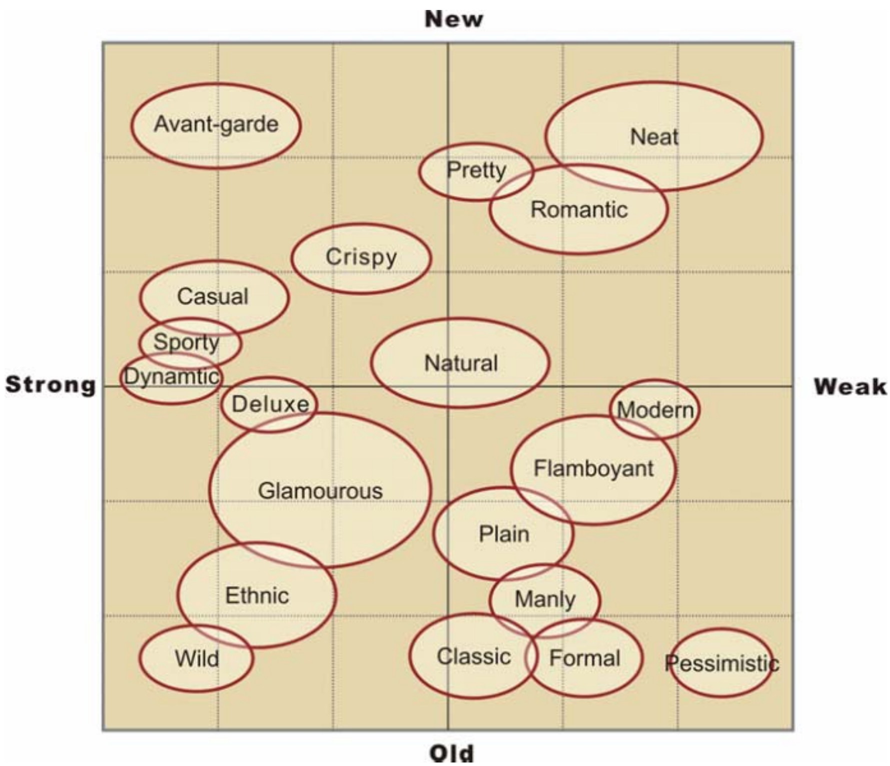


Figure 1. The image scale for color (redrawn from Nagumo (2000))

Using color image to coordinate color, designers can get the result more precisely and quickly. However, there are several drawbacks when we use the paper-based data form, even though we are able to use a specific color image with corresponding colors. The major problem with this is that the data is arranged sequentially.

A second drawback of using books is their static nature. Once they are printed, data cannot be inserted or updated. It is also impossible to combine or merge data from heterogeneous sources. Finally, because CMYK for printing and RGB for monitor take two different color spaces, the colors not in the intersection area cannot be convertible so that the designers may not able to get the same color to the monitor *what they see* from the book. For these reasons, developing a digital color image system offers a promising application for the design domain because it can overcome some limitations of traditional methods of information storage and exchange in the design industry.

Although it is possible that one viewer's perception of color may be different from another's, experimental evidence suggests that the relationships between colors are, in many respects, universal, and thus relatively free from individual and cultural influences (Jacobson and Bender, 1996). This fact can be utilized to build a general architecture for adding guidance to interactive systems.

3.3 Color Harmony

Color is not beautiful or ugly when it is stand-alone. It starts to be meaningful only if color is placed next to each other. Most of the theory of color harmony is based on color wheel. According to Itten (1970), we organize 11 types of color harmony, which are listed in Figure 2.

Color harmony is influenced by adjacent background. Because of the interaction among foreground and background colors, each color looks different. Therefore, color harmony system should consider the visual context. For example, the system can keep visual identity by unified background. However, if the foreground color is very close to the background color, the unified background can be changed to the appropriated one.

4. WORKPLACE LEVEL

This level deals with two issues; design for the activities which users will address and construct the structures and artifacts which enable people to pursue the goal. From those points of view, we try to understand the general activities of color coordination probed by previous works and point out advantages and drawbacks in those works to develop our innovative solution. The outline of the proposed system is shown in the next step.

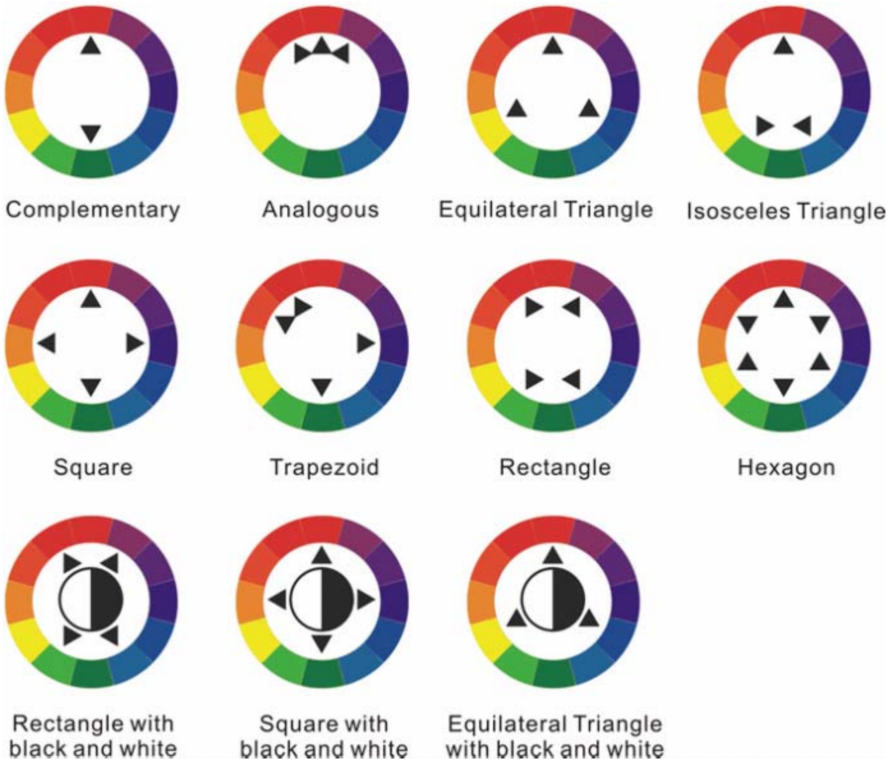


Figure 2. The eleven types of color harmony

4.1 Related Researches and Software

Several researches or Commercial-Off-The-Shelf (COTS) software are in the past 5 years.

Hsiao (2000) applies digital information to computer aided design to execute consultative glossaries from consumers' perceptional demands in color planning. The objectives are to make designers and consumers closer and to achieve color aided marketing. The result of consultative system is a color set according to measures of fuzziness under different consultative condition. However, the system is lacking of color harmony function so that users are hard to make color harmony among a set of suggested colors corresponding a color image.

Tokumaru et al. (2002) propose a system that automatically composes color schemes, which are in harmony with a color inputted in the system and correspondent to user's image. They make fuzzy rules to evaluate harmony of color scheme. In the proposed system, effective judgments of color harmony and color image are executed and suitable color schemes can be acquired by the system. However, the color schemes lack of considering the size, because specific color in different area has different color effect. Moreover, the system does not apply the color scheme to the target objects directly so that the designer still has to do several trial-and-errors to apply and compare each color combination.

Ueki and Azuma (2003) develop methods and system that support a Web page designer to decide appropriate color combination for a specified image. After that, they apply their system to support background color coordination for Web page by the use of the analysis. If a designer uses this system, the designer can know the combination of colors by the intended Web page impression. The designer does not need to have sensibility or knowledge for colors. However, the system can only deal with background of a simple type of web page layout so that it is still not enough to handle realistic web page color coordination. In addition, the system also lacks of color harmony functions.

Through Internet we gathered several color harmony applications like Color Wheel Expert, ColorImpact, Color Wheel Pro, etc. to compare their advantages and disadvantages. Some basic color harmony types based on the color harmony theory are given, but all software provides the function of user defined color harmony to make their system flexible and extensible. Most of the software use the default templates to apply their color harmony types, because it is relatively easy to make the harmonious system fit to the chosen shape, size, and colors. However, using fixed templates makes their systems quite limited. If a designer has totally different kinds of design from any of the default templates –unfortunately, those situations happen very often–, those systems will not be much helpful.

Based on the related researches and software described above, a beneficial and enhanced system is conceived. First, analyzing color information is needed to acquire designer's concept beforehand. Second, providing color harmony function is useful to make color coordination. Finally, if the system can provide a function to get pictures from outside, it will be more flexible; only dealing with fixed templates or just support recommends colors are not satisfied.

4.2 Framework for the Proposed System

The outline of the proposed system is shown in Figure 3. The system is composed of two sub-systems, "Color Analysis Sub System" and "Color Harmony Sub System". Designers can run these two sub-systems independently or combine them together.

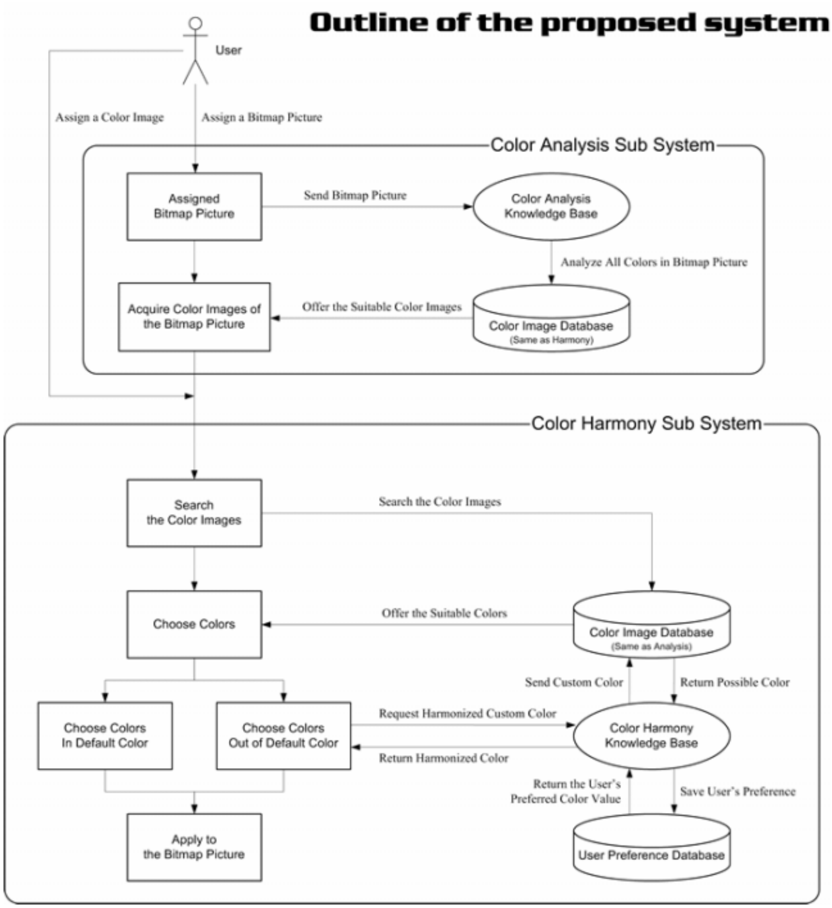


Figure 3. Outline of the proposed system

4.2.1 Color Image Database

The central piece of the system is a Color Image Database (CID), which is organized into appropriate adjective categories. Each category has 3 to 11 adjectives to describe typical feelings of a set of colors. From Nagumo (2000) we extract 160 color images grouped by 23 adjective categories and save them into the CID. On each color image 9 to 24 colors are allocated. This color image database can be customized to represent designers' specific definitions to achieve an expandable feature.

4.2.2 Color Analysis Knowledge Base

The main function of Color Analysis Knowledge Base (CAKB) breaks down a full picture into pixels and calculates RGB color value for each pixel. Each color can be ranked and proportioned using percentages according to the number of pixels contained on the picture. Taken up to 70% of the total area (100%) colors from the top are compared with the data in CID. CAKB uses *CIE 2000 color different formula* to calculate the similarity between two colors and to match one color image to one or several others. The user can decide the tolerance threshold.

4.2.3 Color Harmony Knowledge Base

Eleven types of color harmonic methods are saved in the Color Harmony Knowledge Base (CHKB). Given a specific color, the designer can select a type of color harmony or exhaust all the 11 methods. The system then will find the best matches for the designer. CHKB provides color harmony functions, which not only outputs individual harmonious but also calculates and weighs their areas to get the better solutions. CHKB also has a color adaptation function to meet the designer's specific favour. This function depends on the User Preference Database (UPD) which will be discussed in the next section.

4.2.4 User Preference Database

The capability to allow color preference makes the system more efficient. After a designer has decided a color combination and has applied to the picture, UPD records the personal color choice. When the designer uses the system next time, the system checks the record in UPD and calculates personal offset between CID and UPD. The UPD becomes more powerful over time as designers solve more problems and thus add more color solutions to the database to be re-used when needed.

5. OPERATIONAL LEVEL: IMPLEMENTATION

To understand the application at the operational level, we can consider the interface of system in cognitive artifacts. Among user interface styles "what you see is what you get" (WYSIWYG) technique can always tell what final result will be and be more intuitive, especially applying to the color coordination. The prototype has used an Object-Oriented Software Engineering (OOSE) methodology, which governs analysis, design, implementation, and test phases. It also has been implemented by Java, an object-oriented language, to fully take advantage of accessibility through multiple platforms and capability to connect the system in any place.

Designers can start with a picture captured from any particular domain. In our example shown in Figure 4, we obtain a picture from interior design domain.

Figure 4. Main screen of the proposed system

5.1 Color Analysis Sub System

The acquired picture is dispatched to Color Analysis Knowledge Base (CAKB) to investigate colors in the picture. The CAKB analyzes colors according to the gross area (100%), calculates the percentages to be taken by each color and displays in the screen maintaining the ratios (Figure 5). The designers can create their own color image word to save it to Color Image Database (CID) or get the suggested color image word from the CID. If the designer wants to get the adjective from the system, he/she sends the colors analyzed in CAKB to CID to match with the specific color image for describing the group of colors. CID offers the best matching color image to the designer can decide whether to accept the result. If the designer accepts the result, all colors related to the color image will be listed in the screen, and the designer would also be able to get the RGB value for each of the colors by clicking on it.

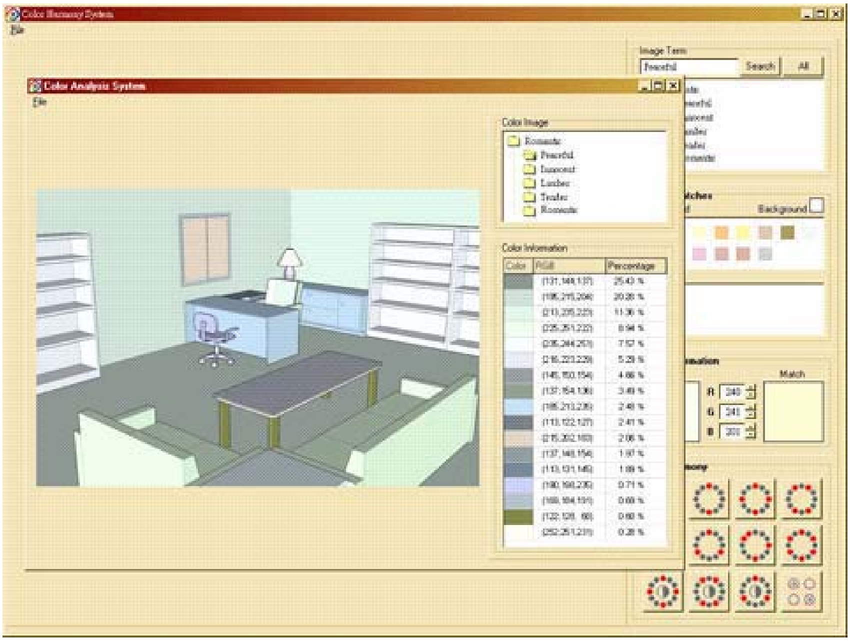


Figure 5. The system with color analysis wizard

5.2 Color Harmony Sub System

The designers can use this Color Harmony Sub System (CHSS) either continuing from the Color Analysis Sub System (CASS) or starting this system from first principles. In either case, a color image would be given into this sub system. Similar to CASS, the system extracts all the colors from CID that match to the color image and displays them on the screen. The designer selects a desired color from the color list to harmonize the initial image. The designer can choose to experiment with only one type of color harmony among the eleven classes or exhaust all the options provided by our system. The CHSS also provides another two advanced functions. One is that the system can make recommendation to the designer based on his/her past selections or preferences from User Preference Database (UPD). The other is that the system can suggest different set of harmonious colors areas, i.e., there are some commonly accepted rules, for example, some vivid colors cannot take up a large portion of an image, etc.

One example is shown in Figure 6 where the system suggested alternatives are displayed. The designer can select a desired alternative and apply the colors to the picture.

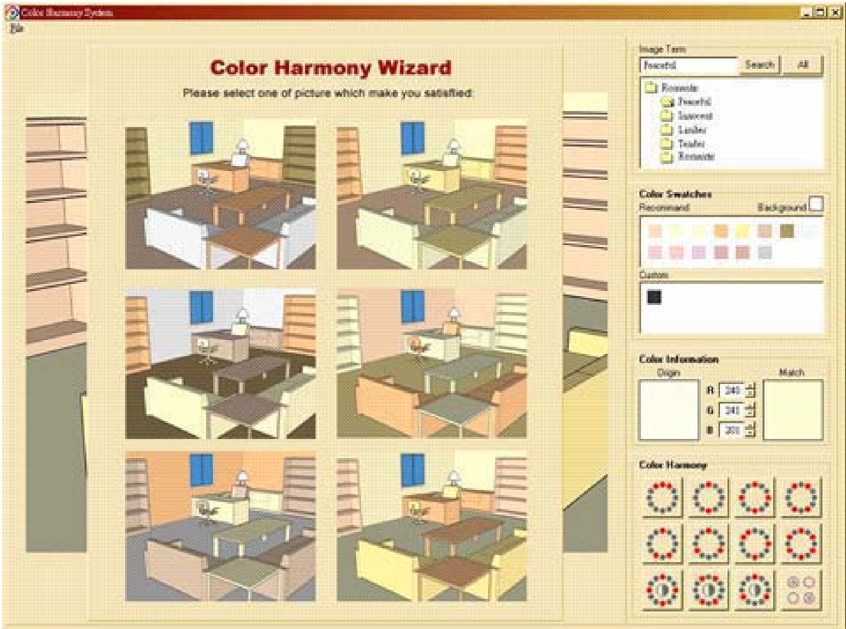


Figure 6. The system with color harmony wizard

6. CONCLUSION

This paper proposed a color coordination system that supports the color analysis function and the color harmony function by using color images. We create a new system which is not only a “smart e-palette”, which is able to overcome traditional rough combinations of RGB values, but also “purpose-oriented”, in that it can be used with two knowledge bases and databases. Given this system, the designers are no longer required to possess expert level color knowledge to make ideal color. The system increases design efficiency and accuracy in expressing your *feelings*.

REFERENCES

Benyon, D., 2004, "Representations in Human-Computer Systems Development", in press, http://www.dcs.napier.ac.uk/~dbenyon/publ.html

Benyon, D. and M. Imaz, 1999, Metaphors and Models: Conceptual Foundations of Representations for Interactive System Design, Human-Computer Interaction **14**: 159-189.

Cabarga, L., 2001, *The designer's Guide to Global Color Combinations: 750 Color Formulas in CMYK and RGB from Around the World*, How Design Books, Cincinnati.

Hasdogan, G., 1996, "The role of user models in product design for assessment of user needs", *Design Studies* **17**(1), p. 19-33.

Hsiao, S.-W. and H.-Y. Ou, 2000, "A Study on the Application of the Spatial Color Image to Design", Master Thesis, Department of Industrial Design, National Cheng Kung University, Taiwan. (in Chinese)

Jacobson, N. and W. Bender, 1996, "Color as a determined communication", *IBM Systems Journal* **35**(3), p. 526-538.

Image Research Institute Inc., 2003, *Web Color Design*, DrMaster Press Co., Ltd., Taipei. (in Chinese)

Itten, J., 1970, *The Elements of Color*, Van Nostrand Reinhold, New York.

Komachiya A., 1991, *Polyphony of Colour and Sensibility*, Keisoshobo Publisher, Tokyo. (in Japanese)

Ito, M. and K. Nakakoji, 1996, "Impact of Culture on User Interface Design", in *International User Interfaces* (Galdo, E. and J. Nieksen eds.), John Wiley & Sons, Inc., New York, p. 105-126.

Nagumo, H., 2000, *Color Image Chart*, Long See International Book Co Ltd., Taipei. (in Chinese)

Tokumaru, M., N. Muranaka, and S. Imanishi, 2002, "Color Design Support Considering Color Harmony", in *Proceedings of the Eleventh IEEE International Conference on Fuzzy Systems*, *Hawaii, USA*, IEEE, p. 378-383.

Ueki, H. and M. Azuma, 2003, "Background Color Coordination Support System for Web Page Design", in *Proceedings of the Second IEEE International Conference on Cognitive Informatics (ICCI '03)*.

Web Design, http://www.webdesignref.com/chapters/13/ch13-15.htm

Simulation and Agent Technology

Using Bayesian Decision Networks for Knowledge Representation under Conditions of Uncertainty in Multi-Agent Land Use Simulation Models

Linda Ma, Theo Arentze, Aloys Borgers, and Harry Timmermans
Urban Planning Group, Eindhoven University of Technology, The Netherlands

Keywords: Land Suitability Analysis, Multi-Agents, Knowledge Representation, Bayesian Decision Networks.

Abstract: Land suitability analysis typically involves the assessment of the suitability of land units without knowing the future spatial distribution of land use. Traditional planning techniques have used "algebraic equations" to express land suitability as a weighted function of suitability scores across multiple criteria. However, the existing multi-criteria evaluation methods do not systematically account for uncertainty about the land use in adjacent and other cells. This paper proposes an alternative approach to land suitability analysis that does address the problem of uncertainty. In particular, Bayesian decision networks are suggested as a means of knowledge representation for agents in a multi-agent land use simulation system. Bayesian decision networks model the uncertainty in terms of probabilities specified in the network representing the expertise of specialists with respect to specific land uses. This paper discusses the approach and illustrates its use in the context of a retail agent.

1. INTRODUCTION

1.1 Background

The decision where to locate a particular land use is a key decision problem for urban planners and other location decision makers alike. The location decision problem involves the assessment of the suitability of a

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 129-144.

particular piece of land (cell, zone) for a particular land use in light of the characteristics of the cell itself and the spatial distribution of a series of land uses. Often, this decision making process is a multi-dimensional and multi-disciplinary activity embracing social, economic, political, and technical factors (Kim et al., 1990). This very nature makes the problem complex, uncertain and subjective (see Yeh et al., 1999).

The most difficult aspect of any location decision problem facing an urban planner/decision maker is uncertainty. Many types of uncertainty can be distinguished, caused by various reasons such as a lack of data, imperfect information, and uncertain future developments. This paper focuses on one type of uncertainty that decision markers are facing: the uncertain spatial distribution of competitive and synergetic land uses.

Over the past decades, land suitability analysis, defined as "finding an appropriate use for the land unit that has the suitability for that desired use" (see Fabos, 1985; Hossain, 1989), has been the focus of much research. The tools that have emerged assist planners to systematically evaluate the suitability of a particular piece of land for different land uses. Typically "algebraic equations" have been used to express land suitability as a weighted function of suitability scores. The widely used Multi-Criteria Evaluation (MCE) techniques are based on a weighted sum of evaluation scores (*i.e.,* Fedra et al., 1990; Carver, 1991; Pereira et al., 1993; Jankowski, 1995; Malczweski, 1996; Lin et al., 1997). The set of criteria is usually divided into criteria that pertain to characteristics of the land unit itself, and to criteria, depicting the accessibility to other land uses in adjacent or more distant cells.

The future land use of other cells is often not or only partly known and hence the suitability of any cell has to be judged under conditions of uncertainty. Decision makers will assess the suitability of any cell partly based on their beliefs about the spatial distribution of land use in the area of interest. Unfortunately, traditional land suitability methods do not consider this type of uncertainty. As an alterative, this paper therefore suggests the use of Bayesian decision networks in land-use decision-making to take into account the problem of uncertainty. The approach described in this paper is part of the development of a wider multi-agent planning support system, called Masque (A Multi Agent System for Supporting the Quest for Urban Excellence) that is currently developed by the authors and their co-workers.

1.2 Structure of this paper

This paper consists of five sections. After this introductive section, the second section shortly presents the Masque framework. The third section introduces Bayesian decision networks. It is followed by a section that describes how these decision networks can be applied in land suitability analysis, explores the approach and provides an illustrative example. Finally, the fifth section summarizes the major conclusions and discusses future research activities.

2. MASQUE FRAMEWORK

Masque is a research program and multi-agent planning support system that aims at supporting decisions related to complex, uncertain and subjective urban planning problems in a user-friendly environment.

This section first presents the components of agents defined in the MASQUE framework. Then, it sketches the functions they need to perform in the MASQUE framework.

2.1 Components of the multi-agent systems in MASQUE framework

Agents have been defined in various ways in the literature (see Maes, 1994; Wooldridge et al., 1995; Dijkstra et al., 2000). This paper defines an "agent" as an entity that has its own knowledge and goals to perform certain tasks. In Masque, a set of agents works together to achieve a set of goals, given an uncertain environment. The framework of the system consists of data structures and protocols to co-ordinate the actions of the agents.

Each individual agent in the system has a set of problem-solving skills and experiences. Four types of agents are distinguished: 1) facilitation agents; 2) interface agents; 3) tool agents; 4) domain agents. The domain agents represent of particular land-uses and are identified based on a commonly used classification of land-use in Dutch planning practice (*i.e.,* the IMRO model). These land uses are business, housing, transportation, recreation, landscape, technical infrastructure, hydraulic construction and service (see Ravi, 2000; Saarloos *et al*, 2001).

2.2 Function of agents

Each agent is autonomous and able to provide (updated) information and to perform different kinds of tasks. The system uses a raster-based representation of the study area. Each agent uses his knowledge (*i.e.,* regulations, requirements, criteria, attributes, constraints, decision rules, techniques, etc.) to develop beliefs about the likely distribution of land uses across space. Based on these expectations, each agent determines his preferences and expresses his claims (*i.e.*, the cells he wants for his land use) and passes this information to other agents. In a cyclic procedure, the initiator of a plan proposal (which can be any one of the domain agents or the planning agent) processes the claims and makes allocation decisions until a plan is fully determined.

A key problem in developing the system therefore is how to represent the domain knowledge that the agents are assumed to use in assessing the suitability of any particular cell in light of their beliefs about the likely (future) distribution of land use in the planning area. In the present paper, we focus on a knowledge representation method (based on Bayesian Decision Networks) for agents to formulate land-use claims.

3. KNOWLEDGE REPRESENTATION IN BAYESIAN DECISION NETWORKS

Knowledge representation refers to the way in which knowledge of experts is modelled. As discussed, for the envisioned planning support system, the knowledge representation format should be able to deal with uncertainty and allow probabilistic reasoning based on beliefs under conditions of uncertainty. Therefore, agent or expert knowledge is represented using Bayesian Decision Networks (*e.g.,* Neapolitan 1990; Ames 2002). In this section, we discuss this means of knowledge representation in which uncertainty is modelled explicitly.

3.1 Structure of Bayesian decision networks

Bayesian decision networks can model uncertainty and provide a framework for representing cause and effect relationships between variables in a decision problem. A Bayesian decision network consists of three types of nodes: decision nodes, nature nodes and utility nodes (see Figure 1). Decision nodes represent the variables on which agents can make a decision. A network may have one or multiple decision nodes. In case of multiple decision nodes, modelers have to decide one the sequential order of these decision nodes. Nature nodes represent the variables over which the decision maker has no control. The outcomes of these nature nodes are typically uncertain to decision makers because nature decides on the value of outcome variables. Utility nodes represent the utility values, reflecting agents' preferences for the possible states of the system being planned. In general, arrows in the network represent cause-effect relationships between nodes. In the envisioned planning support system, they can be interpreted as reasoning relationships and are detailed in a conditional probability table (CPT) in case of a nature node and a conditional utility table (CUT) in case of a utility node. Both decision variables (decision nodes) and nature variables (nature nodes) are represented in terms of an exhaustive set of mutually exclusive states that represent the possible values/outcomes of the variables. This means that if the variables are continuous, they should be discretized.

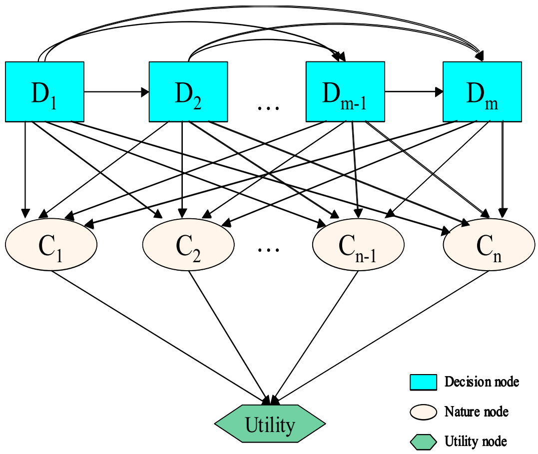


Figure 1. General structure of Bayesian decision networks

3.2 Expected utility

Once the structure of the decision network is defined and the Conditional Probability Tables (CPTs) and Conditional Utility Tables (CUTs) are specified, standard algorithms can be used to reason and determine the expected utility of each decision option. In case of a single decision node, the expected utility of a decision option is calculated as the sum of the products of probability and utility across the possible outcomes of the decision option. In case of multiple decision nodes, the expected utility of a decision option is defined as the expected utility of that decision option under the condition of the best decisions on the next decision variables. To evaluate decision options when multiple decision variables are involved, the following procedure is applied:

1) Calculate the expected utilities of the options of the first decision node;
2) Choose the decision option that maximizes the expected utility;
3) Enter the decision that has been made in the decision networks.
4) Repeat the procedure for the next decision node;
5) Continue this procedure until the last decision is made.

3.3 Conditional Probability Table

A condition Probability Table (CPT) quantitatively defines how an outcome variable (nature node) is impacted by its parent nodes. More specifically, it specifies the probability distribution across the states of the outcome variables for each combination of states of the parent nodes. The rows of the conditional probability table represent all possible combinations of states of the parent nodes and the probability distribution across the states of the nature node under concern. The rows can be interpreted as probabilistic rules, reflecting for example expert knowledge, to determine the state of the nature variables.

Given

1) A nature variable C with a set of states: $\{c^1, c^2, \ldots, c^k, \ldots, c^m\}$ and parent node D_i with states: $\{d_i^1, d_i^2, \ldots, d_i^{n_i}\}$; n_i is the number of states for the parent node D_i.

The probabilities of the states of the nature node can be calculated as:

$$P(c^k) = \sum_S P(c^k \mid S) \qquad (1)$$

where S is the set of all combinations of states of parent node D_i.

The conditional probability table (CPT) represents domain knowledge that can be provided by knowledgeable experts, observational data, results of complex model simulations, written documents or combinations of these knowledge sources depending on the variable under concern.

3.4 Conditional Utility Table

Conditional utility table (CUT) has a similar structure as the conditional probability table (CPT) of a nature node, *i.e.,* the rows of the table represent all possible combinations of states of the parent nodes. However, rather than a probability distribution across all possible combinations of states, the table indicates the utility value of each combination of states of parent nodes. The scale used to measure utility is free to choose, as long as it is metric (*i.e.*, interval scale) and the direction is positive (*i.e.*, a higher value indicates a higher preference).

Given

1) Any set of parent nodes

- C_1 has a set of states: $\{c_1^1, c_1^2, \ldots, c_1^{m_1}\}$;

- C_2 has a set of states: $\{c_2^1, c_2^2, \ldots, c_2^{m_2}\}$;
- …
- C_j has a set of states: $\{c_j^1, c_j^2, \ldots, c_j^{m_j}\}$; m_j is the number of states for the nature variable C_j.

The conditional utility table defines utilities as a set of values:

$$U(c_1^1, c_2^1, \ldots, c_j^1),\ U(c_1^1, c_2^1, \ldots, c_j^2), \ldots, U(c_1^{m_1}, c_2^{m_2}, \ldots, c_j^{m_j}) \qquad (2)$$

3.5 Order of decisions, non-forgetting links and feasible utility nodes

In the case of multiple decision nodes in a network, the user should define the order in which decisions are made by drawing the arrows from one decision node to another decision node. An arrow is drawn between any two decision nodes D_i and D_j if the decision on D_i is known the moment the decision on D_j is made. Therefore, arrows between decision nodes are called non-forgetting links; they do not present causal relationships but rather define the order in which the decisions are made.

In real decision problems it often occurs that two (or more) options from different decision variables are incompatible with each other. Bayesian decision networks cannot directly represent such incompatibility relationships. A possible way to represent incompatibility indirectly is to add a utility node for each pair of decision variables between which incompatible relations exist. In the conditional utility table of such a utility node, any arbitrary large negative value is specified for each incompatible combination of decision options, whereas the utility for all other compatible combinations is set to zero. Thus, this principle ensures that the expected utility values will be strongly negative for the incompatible option combinations, making sure that the incompatible decision options will not be chosen.

3.6 Knowledge representation and uncertainty

Bayesian decision networks can represent expert knowledge qualitatively and quantitatively. The type of knowledge involved depends on the decision problem itself and the type of nodes. In case of a decision node, agents should know the position of the decision in the sequential order of decisions. In case of a nature node, the conditional probability distributions reflect an agent's assessment of the likelihood of outcomes of the variables under each possible combination of decision options, or in general, the influence of parent variables on outcomes. In case of a utility node, agents specify the conditional utility tables to indicate their preferences regarding outcomes in terms of the nature variables.

Bayesian decision networks can model uncertainty in terms of the probability distribution across the states of the nature variables. The degree of uncertainty is reflected by the uniformity of the probability distributions across the states of the outcome variable. The more equal the probabilities are distributed across the possible outcomes, the higher the degree of uncertainty. Uncertainty is inversely proportional to predictability. A perfect uniform distribution (*i.e.,* equal probabilities across alternatives) implies maximum uncertainty, complete unpredictability (random choice), while a deterministic distribution (*i.e.,* 0-100 distribution of hard evidence) means no uncertainty and a deterministic prediction of the outcomes.

4. ILLUSTRATION

To illustrate how a decision network can be used to model expertise and uncertainty in land-use decision problems, we consider a hypothetical medium-sized future district as an example (see Figure 2).

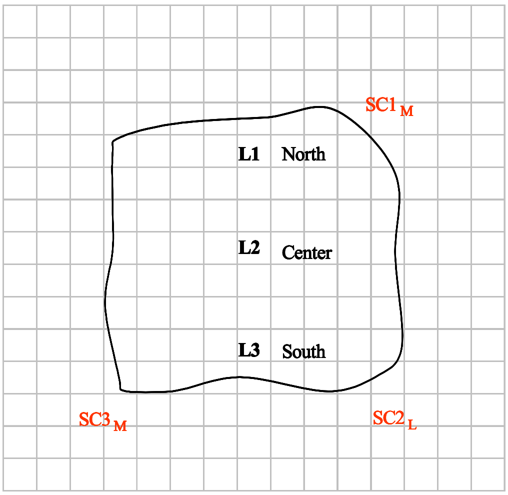


Figure 2. A hypothetical study

We use a raster of cells [i,j] to represent the area. In the present example, the area is divided, on a high level, into three larger regions referred to as: North (L1), Center (L2) and South (L3). In the area, approximately 5000 new houses and all necessary facilities will be located. As an example, we further consider a retail agent that has decided to allocate a maximum of two new shopping centers (a large one denoted as SC_L, a small one denoted as SC_S) in the plan area after he conducted some relevant studies (*i.e.,* market analysis). The decision problems for the retail agent then are: 1) how many (none, one or two) new shopping center(s) should be located in this area? 2) If the number of new shopping centers is not less than one, where should the shopping center(s) be located in the plan area? To reach these decisions, we assume that the retail agent conducts a suitability analysis at two levels: first at the regional level (North, Center and South); second at the cell-based level within that suitable region(s) identified at the regional level.

There are four existing shopping centers in the neighborhood of the plan area (see Figure 2), labeled as medium shopping center one ($SC1_M$), large shopping center two ($SC2_L$) and medium shopping center three ($SC3_M$). The agent is supposed to be aware of this. However, the allocation of land uses in the plan area is not known with certainty until final allocation decisions have been made. The agent only has particular beliefs of land-uses in adjacent cells (and all other relevant cells across the plan area) are represented as probabilities.

4.1 Simplified decision network for the location decision of new shopping centers

Figure 3 shows a simplified decision network to decide where to locate the maximum two shopping centers. This simplified network serves to explain how a Bayesian decision network can be specified to represent knowledge under conditions of uncertainty. It also illustrates how a decision network can be used to address two sequentially related decisions and how to deal with infeasible combinations of shopping center locations. For example, it is not feasible to allocate a small shopping center to the same location where a large shopping center has been allocated. In this simplified decision network, an extra utility node labeled 'feasible' is used to represent these incompatible combinations.

The second layer in Figure 3 shows the nature nodes relevant to this problem. It shows that the location decision is assumed influenced by:

1) total minimum travel distance to the nearest shopping center (C_1) and the sum of ratios of the supply of all shopping centers to distance summed across all housing cells (C_2). Both influence consumers' satisfaction;
2) profit of future retailers in the planning scheme (C_3) and profit of other existing retailers outside the plan area (C_4). They have effects on retailers' satisfaction;
3) efficient use of facilities and space (C_5) and total travel distance to all shopping centers (C_6). These variables influence the satisfaction at the level of the community.

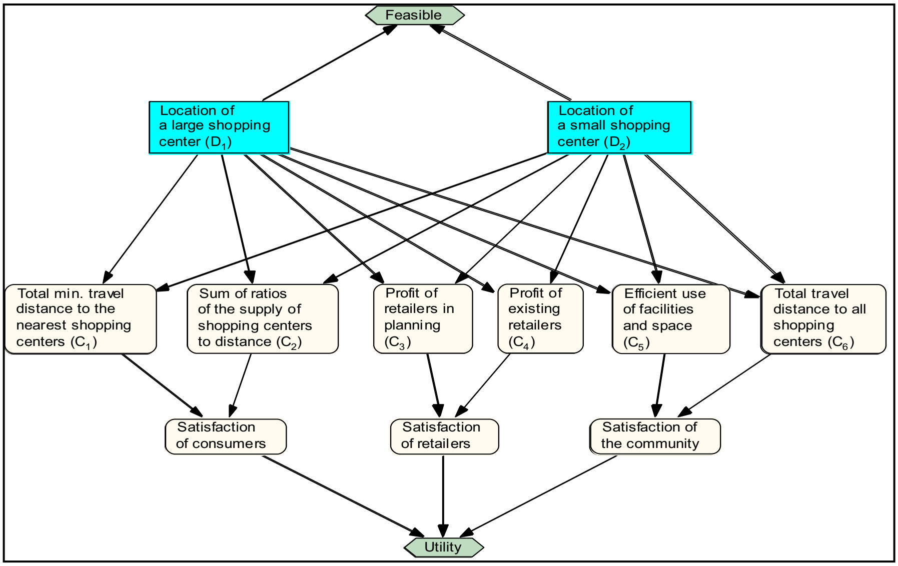


Figure 3. A simplified example of a location decision networks for shopping centers

Each decision variable has four decision options 'North, Center, South and none'. Each of nature variables in this network has three states. The variable C_1 has a set of states 'short, average and long' defined as distance (in meter) intervals: [0~500), [500, 1000), [1000, infinity] respectively. The variable C_2 has states 'high, average, low'. Variables C_3 and C_4 each have states 'high, normal, low'. Variable C_5 has states 'high, normal, low' and variable C_6 has states 'short, average, long'.

The third layer in Figure 3 shows the variables 'satisfaction of consumers', 'satisfaction of retailers' and 'satisfaction of the community'. Each of them has states 'good, normal, poor'. In the following section, we will focus on one nature variable C_1 'total minimum travel distance to the nearest shopping center' to explain how to specify the probabilities for all possible combinations of its parent nodes in a conditional probability table.

4.2 Approach to define the conditional probability table

The retail agent is assumed to make location decisions first at the regional level and then at the cell level within the chosen region(s). The crucial step is how the agent can specify the probabilities in the CPT of for example the nature variables C_1 and C_2 at the regional level and at the cell-based level respectively. Specifying the CPTs requires computing the values of the nature variables, C_i, for each combination of location choices. To determine the values of the variables with certainty, the allocation of the residential land-use must be known, whereas in the planning stage we know only a probability of a residential use for each cell. One possible method would be to use the probabilities as weights in calculations of the measures. However, in that way the notion of uncertainty at the cell level would be lost. To solve this problem, we propose a sampling method to specify the probabilities in conditional probability tables (CPTs).

4.2.1 At the regional level

For the combination of North-North (SC_L in the North and SC_s in the North), the retail agent repeatedly draws a sample from all possible configurations of housing allocations in the plan area by drawing a decision for each cell whether or not the residential land use will be allocated to that cell based on the known probability of a residential land use for that cell. Then the total minimum travel distance from any housing cell to the nearest shopping center is then calculated.

The retail agent repeats to select random samples of configurations of housing allocations, say *n* times, and calculates C_1 for each sample. In doing so, he obtains *n* different values for C_1. Then, he categorizes these results into the distance intervals defined for C_1, e.g.: [0~500), [500, 1000) and [1000, infinity] for three states: short, average and long, respectively. The number of times a value falls in each of these categories can then be counted. Finally the probabilities in the CPT of variable C_1 for the combination of North-North are determined as the number of counts (frequency) in each interval divided by the total number of simulations/samples. This procedure is repeated for each of the remaining fifteen combinations of possible location decisions.

Next, in a similar vein, the agent specifies the CPTs of other variables C_2, C_3, C_4, C_5 and C_6. Note that the CPTs of the satisfaction variables do not require such a sampling procedure as they do not directly depend on land use allocations. Finally the agent specifies the CUT of the utility variable for all state combinations of three variables: ‘satisfaction of consumers’, ‘satisfaction of retailers’, and ‘satisfaction of the community’. After the CPTs for all other variables and the CUT have been specified, the agent “reasons” to derive the expected utility for the decision variables D_1 and D_2 (at the regional level).

There are several incompatible combinations of location of the two shopping centers. To handle these incompatible combinations, a feasible utility node (see Figure 3) is used to make these incompatible combinations infeasible (not to be chosen).

4.2.2 At the cell-based level

To find out where (which cell(s) in that region) the new shopping center(s) should be located, the agent assesses the suitability of each cell, *i.e.*, cells in the North for SC_L and cells in the South for SC_s.

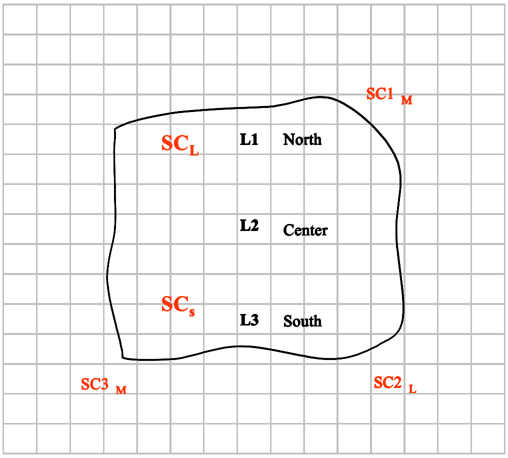


Figure 4. A hypothetical study area

In this example, there is a maximum of 702 combinations of cells for two shopping centers at the cell level (26 cells in the North and 27 cells in the South, see Figure 4). However, this number might be reduced because only cells not closer than a certain distance from any existing shopping center are feasible for the location of the new shopping centers. For each combination of cells, the agent uses the same procedure as explained at the regional level to specify the probabilities in the CPT's and in the conditional utility table. Based on these beliefs, the agent can reason the expected utility of any options of decision variables D_1 and D_2 and decide where to locate the two shopping centers at the cell level.

4.3 Application

This section only analyzes the partial decision network (see Figure 5) discussed in Section 4.0 to show how the technique is applied to a decision problem. The 'feasible' node represents the incompatible location decisions, which are indicated by a large negative penalty. The probability distribution in Table 1 indicates the agent's beliefs of the probabilities of the states of the travel distance variable for each possible combination of the location decisions. The degree of uniformity of the probability distribution indicates the degree of uncertainty. The uncertainty is highest for combinations

"North-South" and "South-North", while lowest for combination "None-None".

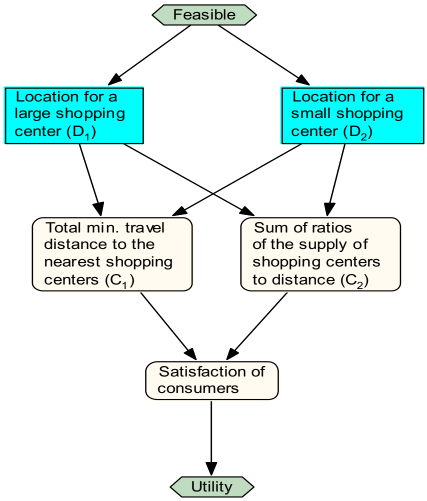


Figure 5. Example of analysis

The degree of uncertainty in Table 2 is rather low because the agent has high confidential beliefs. We assume the agent's preferences are portrayed in Table 3 on a 0-500 scale, where the highest utility receives a value of five hundred and the lowest utility receives a value of zero. These values can be determined from experts' experience or questionnaires.

Table 1 CPT of the total minimum travel distance to the nearest shopping center (%)

Location of a large shopping center	Location of a small shopping center	Total minimum travel distance		
		Short	Average	Long
North	North	8	33	59
	Center	12	38	50
	South	37	36	27
	None	8	33	59
Center	North	12	38	50
	Center	5	30	65
	South	12	42	46
	None	5	30	65
South	North	37	36	27
	Center	12	42	46
	South	8	32	60
	None	8	32	60
None	North	8	33	59
	Center	5	30	65
	South	8	32	60
	None	2	15	83

Table 2 CPT of the sum of ratios of supply of shopping centers to distance (%)

Location of a large shopping center	Location of a small shopping center	Sum of ratios of supply of shopping centers to distance		
		High	Average	Low
	North	8	62	30
North	Center	7	68	25
	South	7	73	20
	None	6	42	52
	North	3	77	20
Center	Center	3	80	17
	South	5	77	18
	None	3	68	28
	North	7	58	35
South	Center	7	53	40
	South	9	48	43
	None	5	42	53
	North	1	2	97
None	Center	2	2	96
	South	2	5	93
	None	1	2	97

Table 3 Conditional utility table

Satisfaction of consumers	Utility
Good	460
Normal	200
Poor	30

Table 4 Expected utility

The large shopping center		The small shopping center	
Decision options	Expected utility	Decision options	Expected utility
	-816	North	-816
	205	Center	205
North	**266**	South	**266**
	160	None	160
	207	North	207
	-821	Center	-821
Center	216	South	216
	167	None	167
	249	North	249
	195	Center	195
South	-833	South	-833
	154	None	154
	110	North	110
None	97	Center	97
	111	South	111
	-934	None	-934

Given these values, the expected utilities derived from the network at the regional level (see Table 4) indicate that both shopping centers should be planned for the planning area. The maximum expected utility value '266' is generated for the location combination of 'North-South'. Thus the best option for location of the large shopping center is the North while the best option for locating the small shopping center is the South. The expected utility associated with this combination of locations is 266. To decide on locations of the two shopping centers at the cell level, a similar process is followed.

In addition, Table 4 also shows that the incompatible combinations of locations are considered infeasible as reflected by their large negative expected utility.

5. CONCLUSIONS AND FURTHER RESEARCH

This paper argued that most traditional methods of land suitability analysis do not consider the inherent uncertainty in the future spatial distribution of land uses, that is typically assumed to influence the accessibility and hence the suitability of a site. To overcome this limitation, this paper proposes an alternative approach to represent the knowledge and reasoning of multiple agents under conditions of uncertainty. More specifically, the use of Bayesian Decision Networks is advocated. Uncertainty is modelled in terms of conditional probability tables, which represent the impact of particular variables on decision options. The use of this approach is illustrated in the context of a retail location decision example. A sampling method is proposed to maintain the concept of uncertainty across a sequence of decisions. The results of the simple example indicate the potential of this alternative approach.

Further research is required to develop an appropriate support system. First, the knowledge structure should be elicited for every envisioned agent in Masque. This involves the construction of the network, and the derivation of the conditional probability tables and conditional utility table, unless the latter is seen as part of the use of the system where users can specify these tables. Second, the size of the tables can become very large as the complexity of the network increases. For such complex networks, models and standard algorithms that can compute the probability of all possible combinations of parent nodes in the decision network formulated and explored. We plan to report on such development in the near future.

REFERENCES

Ames, D., 2002, *Bayesian decision networks for watershed management*, PhD dissertation, Utah State University, Logan, Utah, USA, pp.31-74.

Carver, S.J, 1991, “Integrating multi-criteria evaluation with geographic information systems”, *International Journal of Geographical Information Systems 5*, p.321-339.

Dijkstra, J., H. Timmermans and B. de Vries, 2000, “Towards a multi-agent system for visualizing simulated behavior within the built environment”, *Proceedings of the 5th international conference on Design and Decision Support Systems in urban planning.* Nijkerk, the Netherlands, August, 22~25, 2000.

Fedra, K. and R. Reitsma, 1990, “Decision support and geographic information systems”, in: Scholten and Stillwell (eds.) *Geographic information systems for urban and Regional Planning*, Kluwer, Dordrecht, the Netherlands, p.177-188.

Fabos, J., 1985, “Land-use planning from global to local challenge”, Chapman and Hall, London, 1985.

Hossain, H., 1989, “Use of computers in urban planning and management”. *Proceedings of the international conference on computers in urban planning and urban management* (CUPUM), Hong Kong, August 22-25, 1989, pp427~438.

Jankowsk, P., 1995, “Integrating geographic information systems and multi-criteria decision-making methods”, *International Journal of Geographical Information Systems 9*, p.251-273.

Kim, T., L. Wiggins and J. Wright (eds), 1990, *Expert Systems: applications to urban planning*, Berlin: Springer, 1990.

Lin, H., Q. Wan, X. Li, J. Chan and Y. Kong, 1997, “GIS-based multi-criteria evaluation for investment environments”, *Environment and Planning B: Planning and Design 24*, p. 403-414.

Maes, P., 1994, “Agents that reduce work and information overloaded”, *Communications of the ACM*, 37(7), 1994, pp. 31-40.

Malczweski, J., 1996, “A GIS-based approach to multi-criteria group decision-making”, *International Journal of Geographical Information Systems 10*, p.955-997.

Malczewski, J. ,1999, “GIS and Multi-criteria decision analysis”, Wiley, London.

Neapolitan, R.E.,1990, “Probabilistic Reasoning in Expert Systems: Theory and Algorithms”. John Wiley & Sons, New York.

Pereira J. M.C. and L. Duckstein, 1993, "A multi-criteria decision-making approach to GIS-based land suitability evaluation”, *International Journal of Geographical Information Systems 7*, p.407-424.

Ravi, 2000, Information-model for environment planning, IMRO (in Dutch). Ravi publication 00-06, 2000. Amersfoort, The Netherlands.

Saarloos, D., T. Arentze, A. Borgers and H. Timmermans, H., 2001, “Introducing Multi-Agents in an integrated GIS/VR system for supporting urban planning and design decisions”, *Proceedings (CD-Rom) of 7th international conference on computers in urban planning and urban management (CUPUM)*, Hawaii, USA. July 18-20, 2001.

Wooldridge, M. J. and N.R. Jennings, 1995, “Intelligent agents: theory and practice”, *Knowledge Engineering Review*, 10(2), 1995, pp. 115-152.

Yeh, A. and X. Shi, 1999, “Applying case-based reasoning to urban planning: a new planning support system tool”. *Environment and Planning B: Planning and Design*, 26(1), pp. 101-115.

Towards a Generic Multi-Agent Engine for the Simulation of Spatial Behavioural Processes

MASQUE / SwarmCity

O.T.J. Devisch, H.J.P. Timmermans, T.A. Arentze, and A.W.J. Borgers
Eindhoven University of Technology, The Netherlands

Keywords: Micro Simulation, Multi-Agent Systems, Spatial Simulation Models, Scenario Analysis, Heuristic Decision-Making, Location-Choice Behaviour, Decision-Analysis Trees, Residential Mobility, Lifetime Utility.

Abstract: SwarmCity is being developed as a micro-simulation model, simulating the location-choice behaviour of a population of households, retailers, firms, developers, etc. reacting to an urban plan. The focus of SwarmCity lies –in a first phase- on the decision-making procedures of households, conceptualised as a series of three processes: awakening, search and choice. The methodology used to implement these processes makes use of life-time utility and decision-analysis trees. The final model should work as a scenario-analysis tool, allowing planners, developers, retailers and municipalities to test intervention-proposals, to evaluate legislations, to measure the attractiveness of services, to quantify residential mobility, etc.

This paper illustrates the state of the art in household location-choice modelling and introduces a first attempt in developing a conceptual framework.

1. INTRODUCTION

The design of an *urban plan* depends on, among other things, expected location-choice decisions of the future plan-population. The problem is that this population is, in most cases, unknown and if known, the possible reaction of this population to the plan is not. Each future resident, from households, over retail companies, to firms, developers, service-providers, etc., chooses locations based on particular preferences and constraints. This makes it very hard for the planner to argument his design decisions. Scholars

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 145-160.

now claim that location-choice behaviour displays regularities: for example: households changing location mostly stay within the same area (Clark, 2003), households in relatively large units are less mobile than households in smaller dwellings (Dieleman, 2001), rising income increases mobility rates of owner-occupiers, whereas it decreases mobility rates of renters (Van Der Vlist, 2001). These regularities make it possible *to model the location-choice behaviour* of the future plan-population and, in that way, provide the planner with a tool to test design decisions. It is the *ambition* of SwarmCity to develop such a model, simulating the reactions of a population to an urban plan.

Section 2 explores the state of the art in location-choice modelling. Section 3 defines the focus and relevance of the SwarmCity model. Section 4 & 5 introduce the underlying theory and research methodology. Section 6 proposes how to implement such a model and section 7 ends with some concluding remarks.

2. STATE OF THE ART

Location-choice models describe the behaviour of an individual (be it a household, firm, company, etc.) searching for a location to settle. Modelling urban phenomena, like location choice, originated somewhere in the end of the fifties in North America in reaction to the increased use of the car (Batty, 1976). A first type of models, for this reason, mainly focused on transportation: for example: calculating the number of trips between a set of destinations. Later, a second type of models focused on land-use allocation processes (Alonso, 1970): for example: predicting the spatial distribution of programs around a city-centre based on economical factors. Both type of models later integrated into land-use-transportation models. “The main engine of the generic land-use–transportation model has traditionally been the spatial interaction model or variants thereof” (Torrens, 2000). *Spatial interaction models* approach the city as a system of interacting aggregates in a continuous equilibrium condition. This (balanced) system can be represented by a set of mathematical equations: for example: predicting migration- and job-streams between different parts of a city, etc. This type of models is *static*, analysing the structure of an urban system at one moment in time.

In reality, equilibrium conditions do not exist; an urban system needs time to adjust to change, making a city inherently *dynamic* (Batty, 1976). Dynamic urban models therefore no longer approach the city as a collection of interacting aggregates but as interacting individuals. The focus shifts from analysing the structure of the total urban system to simulating the behaviour of single individuals. These individuals are driven by goals and desires and make decisions based on incomplete knowledge. From these micro-interactions, macro level behaviour emerges on the scale of the system (Torrens, 2000). Two modelling techniques were developed to incorporate these dynamics into land-use-transportation systems: Cellular Automata and Multi-Agent Models. "In *cellular automata*, space is represented as a uniform lattice of cells with local states, subject to a uniform set of rules, which drives the behaviour of the system. These rules compute the state of a particular cell as a function of the previous state and the states of the adjacent cells" (Dijkstra, Timmermans, 1999). *Multi-Agent systems* go a step further and link the rules directly to the individual (or cell), no longer to the system as a whole. Each modelled individual now interacts following a personal behaviour. This makes a multi-agent system an enormously powerful tool to model location-choice behaviour and, in turn, explains the boom, during the last 10 years, of agent-based land-use-transportation models: for example: UrbanSim (Waddell, 2002), OBEUS[1] (Aronovich, 2001), SprawlSim (Torrens, 2003), ILUTE[2] (Miller, Salvini, 2003), ILUMASS[3] (Wegener, 2002), etc. For a state-of-the-art overview on land-use-transportation models, see Berger (Berger, Manson and Parker, 2001).

ILUTE, for example, is currently being developed by Miller & Salvini at the University of Toronto. Implemented, ILUTE should work as a tool to analyse "a broad range of transportation, housing and other urban policies" (Miller, 2003). ILUTE is a dynamic model: a series of mathematical equations and decision rules is applied sequentially through iteration. This sequence simulates, among others things, the migration behaviour of households and firms and daily travel behaviour.

Most of these land-use-transportation models are very ambitious in their intentions: combining market interactions, activity scheduling, (non)residential migration, etc. This ambition might explain why these models remain mainly *conceptual*. To date, implemented agent-based location-choice models do not exist. Analysing the conceptual framework of the different models reveals, for the same reason, that only rather crude behaviour is incorporated: for example: individuals will always choose for the alternative that guaranties them maximum profit or utility, given a set of constraints and decision rules. Plus, preferences and behaviour are assumed fixed over time, meaning that modelled individuals do not learn based on experience.

[1] OBEUS is the acronym of "Object-Based Environment for Urban Simulations".
[2] ILUTE is the acronym of "Integrated Land-Use, Transportation, Environment".
[3] ILUMASS is the acronym of "Integrated Land-Use Modeling and Transportation System Simulation".

The intention of SwarmCity should therefore be twofold: on the one hand to develop an implemented dynamic urban model and on the other hand to incorporate a more complex behaviour. This implies, for example, that decisions are no longer solely driven by immediate profit or utility but might anticipate future events, that individuals make decisions based on limited information, that individuals do not necessarily behave totally rational, that decisions depend on how people search for information, etc. Behaviour is not fixed, but can be adjusted by the individual, based on his experience. To allow for this *complexity*, the models driving the behaviour of individuals will not be based on mathematical equations but on heuristic *if-then-else* rules.

3. FOCUS OF SWARMCITY

Swarm refers to the phenomenon that objects, interacting without the intervention of a regulating super-object, nevertheless seem to – unconsciously- generate an ordering logic. *Swarm/City* takes as point of origin that a city can be interpreted as such a self-organizing object emerging out of the interactions of a population of individuals.

The aim of SwarmCity is to model the spatial behaviour of these individuals. The focus lies on defining and implementing a *generic model*, simulating the decision procedures underlying this behaviour. All assumptions are based on existing literature on location-choice behaviour of firms, households, retail, etc. It should nevertheless be possible in future research, to calibrate the generic model for specific cases.

“The urban system is presented as a mechanism for resolving conflict between various groups who require land for their various purposes” (Lowry, 1968 in Batty, 1976). Groups could be households, retailers, firms, service-providers, etc., all searching for the ideal place to live, to open a shop, to start a business, etc.

In the European context, service-providers like schools and hospitals are mostly planned by the government. Households, retailers and firms each follow a particular location-choice logic. For most firms and retail companies, this logic is mainly rational, market-driven, whereas for households this logic is much more *emotional*, reflecting the life-style of the household. This makes that households are more *flexible* and at the same time more *out-of-control*.

SwarmCity chooses to focus on the location-choice behaviour of *households*: where do they typically locate? (How) do they influence each-others choice? What factors do they take into consideration? The behaviour of all other groups, like retailers, firms, developers, etc. is limited to static decision-rules. The structure of the model will allow that their behaviour can be elaborated in future research.

The *input* of the model is an *urban plan*, defining the spatial allocation of housing types, neighbourhood character, density, amount of green, etc. The model then simulates the immigration, emigration and internal mobility of households. This process is known as *residential mobility* (Dieleman, Mulder, 2002). The *output* is a series of development scenarios, tables and graphs at different moments in time.

The *relevance* of the model lies in the possibility to implement changes in the urban plan and/or behaviour of the individuals while instantly being able to observe the reactions of a future plan-population to these changes. The user can in that way experiment with different spatial and behavioural scenarios. This might help the user to evaluate his decisions and/or convince others of these decisions. Spatial scenarios could, for example, be used to evaluate physical planning interventions, alternative legislations, plausible plan-layouts, etc. Whereas behavioural scenarios could, for example, help to test the robustness of the plan, the sensitivity of the population for certain elements of the plan, the appropriateness of concepts like target-groups and life-styles (Nio, 2002), etc. The model will be tested on *Meerhoven*, a new VINEX[4] location west of Eindhoven.

SwarmCity is a component of a planning support system called MASQUE[5]. MASQUE is currently being developed as a tool supporting urban planners to generate *and* evaluate local plans. The intention of MASQUE is to model the negotiation process of the different experts involved in the making of a local plan in order to generate a multitude of alternative plans, approved by these experts. SwarmCity can then be used to evaluate these plans by simulating the reaction of the future population.

4. CONCEPTS & THEORY

As mentioned in the previous section, SwarmCity will, in a first phase, only focus on the location-choice behaviour of households. This behaviour can be summarized as a series of 3 processes: (1) a household decides it is willing to move, (2) searches for available properties and (3) chooses a property. This series can be conceptualised as: *awakening*, *search*, and *choice* (Bettman, 1979 in Goetgeluk, 1997). Each of the 3 processes involves

[4] VINEX is the acronym of "Vierde Nota Ruimtelijke Ordening Extra". This are plans for new urban expansions, approved by the Dutch Government.

[5] MASQUE is the acronym of "Multi-Agent System supporting the Quest for Urban Excellence".

making decisions, choosing strategies and is subject to more or less *regular* behaviour (see figure 1).

(1) AWAKENING

Each household *demands* a number of things from a property: for example: a minimum size, a level of comfort, a particular location, etc. Over time, certain *events* might change these demands: for example: a house might become too small because of an extra child.

As a result the household is no longer satisfied and decides something has to change. It is *woken up*.

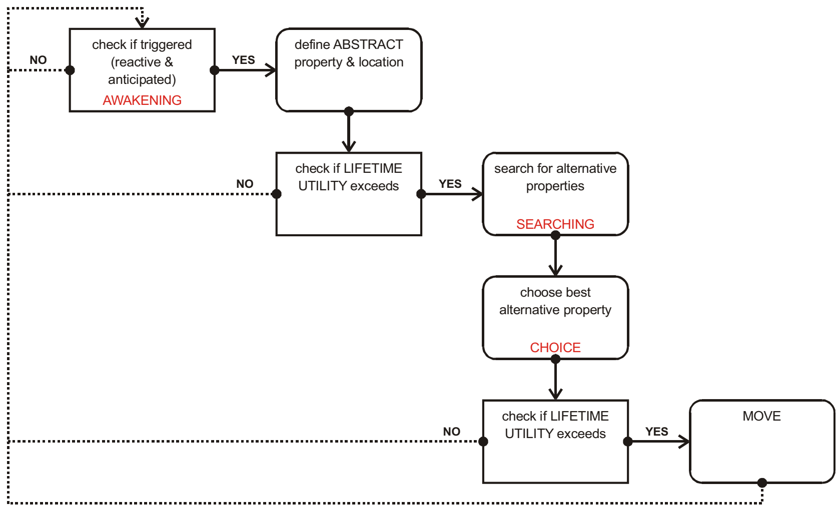


Figure 1: Conceptual framework illustrating awakening, search and choice

Events are not the only awakening-triggers; a household might also start to feel uneasy out of *uncertainty* concerning the appropriateness of his current situation: for example: the household wants to know if there exist properties that better suit its demands. Events might trigger the household to move, whereas uncertainty might trigger the household to look for information.

Once triggered the household has to decide which strategy to follow: will it move to another property, will it extend the current property or will it adjust its requirements? If it decides to move, what will it then move to: a villa with a big garden in a villa quarter or a bigger apartment in the same block?

The answer depends, among other things, on the type of trigger. A first distinction can be made between *actual* versus *anticipated* triggers: a household might wake up because something changed at that moment: for example: one of the household members gets fired, or because it expects something to happen in the future: for example: a child. Secondly, a trigger can be *voluntary* or *involuntary*: for example: because it would like to try something new or because the renting contract ends. Thirdly, a trigger can be a *sudden* event or a *gradual* process: for example: a household might stumble across their dream-house without even having the intention to move, or, the household might decide to move in 5 years out of old-age. Fourthly, a trigger can be related to the current property or to another property. These triggers are known as *push-motives* and *pull-factors*: push-motives push a household out of their current property: for example: a house that becomes too big, whereas pull-factors attract a household to a new property: for example: a cheaper rent. A final distinction is related to *scale*: triggers can be related to the micro, the metropolitan, the national and the international level (Dieleman, 2001): for example: age, tenure, household composition, etc. belong to the micro scale, turnover rate of housing stock, total number of housing, level of urbanization, etc. belong to the metropolitan level, rates of inflation, mortgage interest, etc. belong to the national level and housing policies, variations in wealth, etc. belong to the international level (Dieleman, 2001).

The sensitivity for a certain trigger depends on the profile and preferences of the household. Change in household composition is for example a trigger that makes most young households move from rental to owning (Clark, 2003), whereas a good housing-market might stimulate owners to move to bigger and better located properties (Clark, 2003).

(2) SEARCHING

Searching implies collecting information. A household can be more or less sure of the content of that information. This can be formalized by drawing a probability distribution of all the messages the piece of information could hold. The probability assigned to each message indicates how convinced the household is that this message fits to that piece of information. The more horizontal this distribution, the more uncertain the household is. The *entropy* of this distribution is a measure for this uncertainty. Searching then implies reducing the number of possible outcomes or minimizing the information-entropy. The point where this uncertainty becomes acceptable depends on the particular household.

Searching can be done in different ways. The household, for example, has to decide where to search, how to search, how long it will search, which selection criteria it will take into consideration, etc. Searching costs time and money. The search-radius, -mode, -demands, etc. therefore mostly depend on the *budget* of the household, the *urgency* of the search and the *expected* *information gain*: for example: a household that urgently needs to move will have a different searching behaviour than a household that is just curious about what the market has to offer.

Collected information is stored in a *choice-set*, a list of options meeting the criteria set by the household. The size of this choice-set depends firstly, on the budget and urgency, secondly, on the available information and thirdly, on the capabilities of the household to store information. This last factor is known as *mental effort*. Research, for example, indicates that most households mainly search in their immediate environment and within their current housing market (Clark, 2003) and that households without children have lower demands concerning housing than households with children (Dieleman, 2003).

This can be translated in two types of searching: *querying* versus *exploring*. Querying implies that the searcher has a clear objective, while exploring is used when this objective is not so clear: the more urgent the search, the less explorative. Independent of the objective, a household can search in three ways: (1) through *interaction* with his environment: for example: reading newspapers or driving around, (2) through *communication*: for example: using social networks or consulting estate agents, and (3) through relying on *experience*: for example: a household might only recognize something as an opportunity if it saw it before.

(3) CHOICE

Choice implies evaluating and selecting. The household has to decide upon the type and number of evaluation criteria and the relative importance of these criteria. This might require negotiation and/or group decisions.

What the household does is comparing the utility he expects to derive from each alternative in the choice-set. The alternative with the highest level *wins*. This is not necessarily the best alternative available at that moment in time because the household might have searched in the *wrong* areas or might have misinterpreted important pieces of information. This is known as *limited information condition*: each individual only has access to a limited amount of information.

The final choice depends on the size of the choice-set, the urgency of the choice and the preferences of the household. A household might for example be *risk-aversive* (familiarity seeking) or *risk seeking* (novelty seeking). The result of a choice is a loss or gain in utility or information. A risk-aversive household assigns a bigger weight to expected loss than to expected gain. A risk-seeking household does the opposite.

In reality, search and choice are sometimes part of one and the same process: a household finding a property matching its preferences, might immediately decide to buy this property, because if it waits too long, someone else might buy the property. In this case, there is no choice-phase.

The opposite also occurs: searching without choosing. This might happen when the household is unsatisfied with the search result. It might then decide to *adjust* its demands, preferences and/or search-mode and start searching again. In this case, the household goes more than once through the awakening / search / choice series.

5. METHODOLOGY

In each of the awakening / search / choice processes the household has to make decisions: whether it will move or stay, where and how it will search, which alternative to choose, etc. This requires a standard against which the household can weigh choice-options. This standard is *utility*. In many cases this can be expressed in terms of money (Arentze & Borgers, 2002). A person using an object derives an amount of utility from that object. The more suited this object is to the demands of the user, the higher that amount. Utility thus depends on the preferences of the user. Another implication is that utility can change over time because the demands of the user regarding the object can change over time. The utility derived from using an object thus has to be measured over a period of time, related to a particular user. This is known as *lifetime utility*.

Applied to residential mobility, lifetime utility can be used as a *measurement tool* indicating whether it is better to move or to stay. What the household basically does is continuously comparing the utility derived from staying in his current property with the utility he expects to derive from moving to another property. In SwarmCity, this *other* property is referred to as *abstract property* representing the property the household thinks of moving to if possible/necessary. Abstract refers to the fact that this property only exists in the mind of the household and is thus not a physical property. Besides an abstract property, the household also has an *abstract location* in mind, representing the location where the abstract property should preferably be situated. Depending on the household, the abstract property and location can be more or less detailed: from a Mediterranean villa with two bathrooms and a big garden located in a particular neighbourhood to, just a villa somewhere in the suburbs. The move from current to abstract property can be seen as a step in the *housing career* of the household. “A housing career is the way people change their housing as they progress through the life course” (Abramsson, 2000). Such a step could be upwards: for example: from renting to owning, or downwards: for example: from a house with a garden to a house without a garden. The abstract property can change over time parallel with the household, without necessarily always becoming real.

A household derives utility, on the one hand, from living in the current property, and, on the other hand, from performing activities: for example: going on vacation, meeting friends, doing sports, etc. The *property-utility* depends on the budget of the household, the household composition, the value of the property, the condition of the property and the satisfaction with the neighbourhood. For the current property the utility can be calculated on the basis of direct experience, whereas for the abstract property this is calculated on the basis of the *beliefs* the household has of these factors. The *activity-utility* depends on the budget of the household and on the value of the property. A decreasing budget, for example, has no direct influence on the property-utility but limits the money left for travelling, recreation, shopping, etc. The decrease might be so dramatic that the total utility derived from the current property goes under the utility the household expects to derive from the abstract property. The household then decides it is willing to move and starts searching.

Some of the location-choice decisions are interdependent: for example: the decision to search in a certain area depends on the housing type the household is looking for. These interdependencies can be represented in *decision-analysis trees*. Such a tree consists of nodes and branches; *nodes* call for a decision, whereas *branches* point at the alternatives to choose from. The simplest tree has only one node and two branches, representing a decision with two options: for example: a household having to choose between two properties. The household assigns a utility and a probability value to each of these options. The utility indicates the amount of utility the household expects to derive from that property and the probability value indicates how sure the household is about these expectations. A lower probability results in a lower utility. The option with the highest utility is selected.

In most cases, a decision tree has more than one node. For example, in the case of a household changing location, there are *three nodes* (see figure 2) implying three decisions: first the household has to decide whether to move or to stay. Secondly, if the outcome is to move, it has to search for locations and thirdly, it has to search for properties. The first decision now depends on the second, which in turn depends on the third: a household will only decide to move when the utility he expects to derive from the abstract location and property exceeds the utility derived from the current property. Once this is the case, the household starts searching, building up the decision tree. First, locations are chosen, then properties for each location. Once the searching is over, the household has to choose. He starts by calculating the utility he expects to derive from each property. Then, he selects, for each location, the property with the highest utility. In the next step he has to calculate the utility he expects to derive from each location, each time adding the utility of the selected property. The combination of location and property with the highest total expected utility results in the final choice. A consequence of this interdependency might, for example, be that a household prefers an average villa in a super neighbourhood to a super villa in a bad neighbourhood.

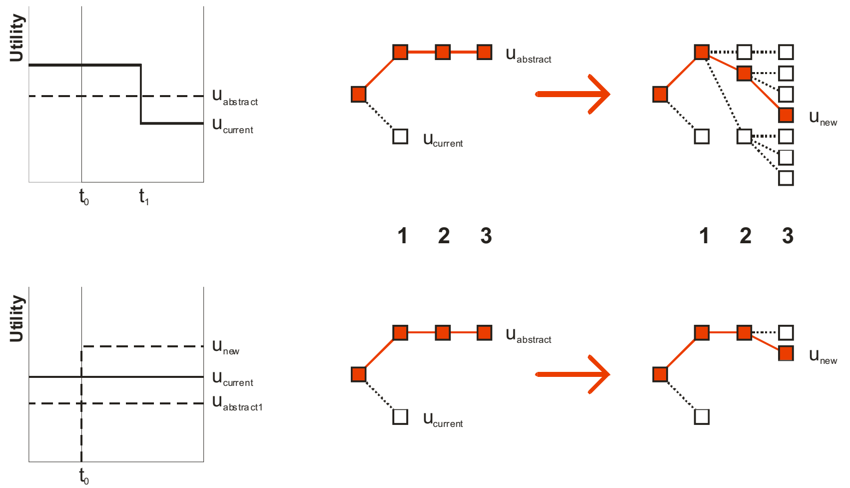


Figure 2: Lifetime utility & decision-analysis trees are used as a tool to visualize scenarios: for example: a household anticipating a change in family composition (top), or a household suddenly stumbling across the ideal property without having the intention to move (bottom). $U_{current}$ represents the utility derived from the current property, $U_{abstract}$ represents the utility the household expects to derive from the abstract property, U_{new} represents the utility derived from the new property, t_0 represents the present and t_1 represents the future.

To conclude, the combination of lifetime utility and decision-analysis trees makes it possible to model both *reactive and anticipating behaviour* and offers a powerful tool to *visualize different scenarios* of awaking, search and choice: for example: a household suddenly stumbling over its dream-property or a household expecting a child (see figure 2). The methodology is generic enough to also apply for location-choice behaviour of retailers, firms, etc.

6. THE MICRO SIMULATION SYSTEM

As mentioned in section 2, the SwarmCity model is a *micro-simulation* of households searching for the ideal place to live. "Micro-simulation models aim at reproducing human behaviour at the individual level, i.e. how individuals choose between options following their perceptions, preferences and habits subject to constraints, such as uncertainty, lack of information and limits in disposable time and money" (Moeckel, 2002).

Multi-agent technology is used as the tool to implement this micro-simulation model. A multi-agent system "consists of a set of agents which together achieve a set of tasks or goals in a largely undetermined environment" (Timmermans, 1999). In SwarmCity, agents represent both persons and objects: for example: households, individuals, properties, neighbourhoods, etc. Each agent has *attributes* representing the characteristic features of this agent: for example: the attributes of a household might be budget, composition, number of children, etc. (see figure 3). The agents representing persons also have *methods*, representing the behaviour of this person: for example: behaviour might be driven by family-motives, profit, etc.

As mentioned in section 4, all these agents demand certain things from their current property. Once these demands no longer hold or are no longer satisfied, they might decide to move. On their search for alternative properties, the agents might interact with other agents. The model consists of three *basic agents*: actors, institutions and an environment. *Actors* can be individuals, households, retailers, firms, service providers and developers. *Institutions* can impose regulations, award subsidies, define average prices and indexes, etc. As mentioned in section 3, the study area considered as a case is Meerhoven, a new VINEX location west of Eindhoven. This study area makes up the *environment* and is represented by 4 *layers*: a zone-, a neighbourhood-, a cell- and a property-layer (see figure 3). The *zones* and *neighbourhoods* are defined by the planner. Each zone has a certain land-use and density. Each neighbourhood has a certain character, building-type, amount of green, etc. Neighbourhoods can vary widely in scale. Therefore, a cell-layer is added. Each *cell* counts 50 by 50 meters and is mainly used to calculate distances.

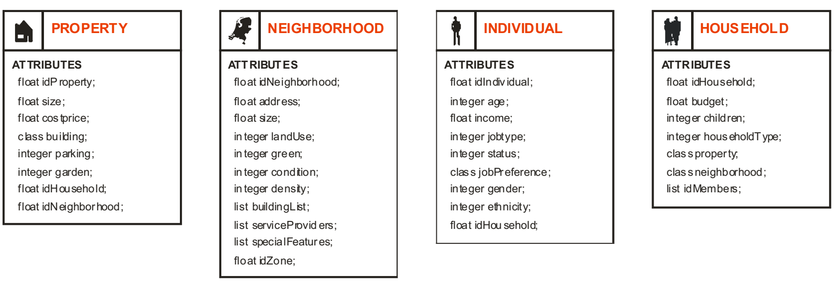


Figure 3. Agent attributes

The properties represent the plots the households can buy. Each *property* has a certain building-type and size. The zone- and neighbourhood-layer are vector-based, whereas the cell layer is grid-based. The property layer is a non-physical layer.

The simulation starts with a planner making an *urban plan*. This plan defines the future neighbourhoods of a given city-part. The finished plan is then inserted as a GIS-file in the SwarmCity model. Now the simulation can start. Each time-step (representing one month), new households enter the plan searching for a property to settle.

Searching starts with selecting neighbourhoods. For each selected neighbourhood, the households then read newspapers and cut out housing-ads that meet their demands. Once a household decides it collected enough information, it contacts the developers selling the selected properties. If the property is not sold yet and if the household decides that the price of the property is acceptable, it will buy the property and put its current property for-sale, if not, it keeps on searching. The price is defined by the developer, based on a general market analysis. Whether or not a household will accept the property depends, among other things, on the budget, the expected utility, the urgency of the move, etc.

From the moment a household buys a property it will each time-step evaluate if that property still meets its demands: for example: the neighbourhood features might change over time because of new developments. If the demands are no longer met, the household will start searching again. What the user sees of this scenario is a change in residential-mobility-density, continuously visualizing the amount of households moving in and out of a cell.

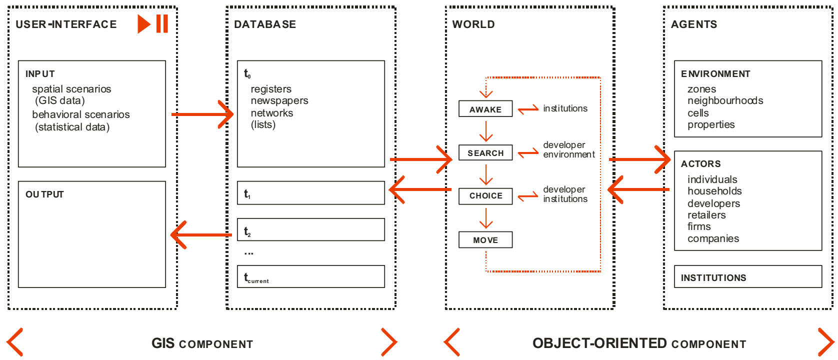


Figure 4. Structure of the implemented model

At this moment, only a limited test-model is implemented. There are, for example, only three actors: individuals, households and developers, and there are no institutions yet. All these actors are randomly generated. In a future version, they will be synthesized through a *Monte Carlo* micro-simulation of a sample of representative households derived from existing surveys and statistical databases like WBO[6], CBS[7], the database of the research group, etc. The test-model is written in the *object-oriented language* C++ Builder with MapObjects as an integrated *GIS-component* (see figure 4).

7. CONCLUSIONS

A review of the current modelling literature learns, firstly, that implemented micro-simulations of location-choice behaviour do not exist yet and, secondly, that the conceptual frameworks underlying existing models only incorporate very simple choice-behaviour. SwarmCity will address both points and can therefore be considered *innovative* for two reasons: firstly, it will be an implemented micro-simulation model and secondly, it will incorporate flexible and dynamic behaviour. The implementation will make use of multi-agent technology. This technology makes it possible to assign a unique behaviour to each household, adjustable over time.

The final model should work as a planning support tool allowing the user to evaluate urban plans. The focus lies on *scenario analysis*, not on prediction. The user will be able to change both spatial and agent settings and can in that way experiment with both *planning- and behaviour-scenarios*. The user could, for example, be a planner trying out different plan-layouts, plan-compositions and/or plan-implementation-strategies; a retailer evaluating if it is suitable to open a new outlet given this type of neighbourhood or checking out what the population composition distribution would be over 10 years; a municipality testing the effect of a new legislation or trying out alternative locations for a particular service; or, a developer trying to get an insight in the willingness of households to adjust their residential-behaviour.

At this moment the model is only partly implemented. *Future development* will, in a first phase, elaborate on the location-choice behaviour of households. In a second phase, the location-choice behaviour of retail will be incorporated. And finally, SwarmCity will be extended with an existing activity-based model AURORA[8] (Joh, 2004), a model simulating the daily

[6] WBO is the acronym of "Woningbehoefte Onderzoek" or "Housing Demand Survey".

[7] CBS is the acronym of "Centraal Bureau voor Statistiek" or "Central Bureau of Statistics".

[8] AURORA is the acronym of "Agent for Utility-driven Rescheduling Of Routinized Activities".

activities of households. This extension might be in the form of a parallel research.

The final model will be *validated*, firstly, by comparing observed model-phenomena with existing *location-choice literature*, secondly, by *simulating historical events* and thirdly, by performing *robustness-tests*. Location-choice literature provides an extensive amount of data on, for example, length of property occupation related to household types, residential movement patterns, choice-behaviour of households, etc. These data will be used to synthesize the model population. Historical events refer, for example, to the developments of VINEX locations. These locations are quite recent and the underlying developments are therefore very well documented. This makes it possible to test if a simulation based on a particular VINEX-plan would evolve parallel with the documented developments. One way to test the robustness is by confronting the model population with 'strange' phenomena and registering if their reactions are 'normal': for example: would households still move to a villa if this villa is surrounded by apartment buildings. With each of these evaluations, the model will grow more valid.

REFERENCES

Abramsson, M., L.-E. Borgegaard and U. Fransson, 2000, "Housing careers – some empirical evidence of a complex concept", *ENHR 2000 conference, Gavle, 26-30 June 2000.*

Alonso, W., 1970, *Location and land use. Toward a general theory of land use*, Harvard University Press, Cambridge, Massachusetts.

Arentze, T.A. and A.W.J. Borgers, 2002, "Evaluation and Decision course", *DDSS Master Program*.

Aronovich, S., I. Benenson, and S. Noam, 2001, "OBEUS: Object-Based Environment for Urban Simulations", *University of Tel Aviv*.

Batty, M., 1976, *Urban modelling. Algorithms, calibrations, predictions*. Cambridge University Press

Berger, T., S.M. Manson, and C.P. Parker, 2001, "Agent-Based Models of Land-Use and Land-Cover Change".

Clark, W.A.V. and Y. Huang, 2003, "The life course and residential mobility in British housing markets", *Environment and Planning A,* 35, p. 323-339.

Dieleman, F. M., 2001, "Modelling residential mobility; a review of recent trends in research", *Journal of Housing and the Built Environment,* 16, p. 249-265.

Dieleman, F.M. and C.H. Mulder, 2002, "The geography of residential choice", in: Aragones, J.I., G. Francescato and T. Garling (eds.) *Residential Environments. Choice, satisfaction and behavior*, Bergin & Garvey, London, p. 35-54

Dijkstra, J. and H..J.P. Timmermans, 1999, "Towards a multi-agent system for network decision analysis".

Goetgeluk, R. W., 1997, *Bomen over wonen*, Thesis publication, Utrecht University.

Hooimeijer, P. and A. Oskamp, 1999, "Advances in the microsimulation of demographic behavior", in: Van Wissen and Dykstra (eds.) *Population Issues: an interdisciplinary focus*, Plenum Press, New York, p. 229-263.

Joh, C.H., 2004, *Measuring and predicting adaptation in multidimensional activity-travel patterns*, Thesis publication, Technical University of Eindhoven.

Lu, M., 1998, "Analysing migration decisionmaking: relationships between residential satisfaction, mobility intentions, and moving behaviour", *Environment and Planning A*, 30, p. 1473-1495.

Miller, J. E. and P. Salvini, 2003, "ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems", *10th International Conference on Travel Behaviour Research, Lucerne, August 10-15*.

Moeckel, R., C. Schurmann, and M. Wegener, 2002, "Microsimulation of Urban Land Use", *42nd European Congress of the Regional Science Association, Dortmund, August 27-31*.

Nio, I., 2002, "Van levenswijzen tot lifestyles. Over de bruikbaarheid van leefstijlen voor de ruimtelijke ordening", *Stedebouw & Ruimtelijke Ordening*.

Timmermans, H.J.P., 1999, "MASQUE, a project proposal".

Torrens, P., 2000, "How land-use-transportation models work", *A CASA working paper*.

Torrens, P., 2003, "SprawlSim", http://www.sprawlsim.org/

Van Der Vlist, A.J., C. Gorter, P. Nijkamp, and P. Rietveld, 2001, "Residential Mobility and Local Housing Market Differences", *Tinbergen Institute Discussion Paper*.

Waddell, P., 2002, "UrbanSim: Modelling Urban Development for Land Use, Transportation and Environmental Planning", *Preprint of an article that appeared in the Journal of the American Planning Association*, 68(3), p. 297-314.

Wegener, M., 1998, "The IRPUD Model", http://irpud.raumplanung.uni-dortmund.de/irpud/pro/mod/mod_e.htm.

Crowd Modeling and Simulation

Towards 3D Visualization

Stefania Bandini, Sara Manzoni, and Giuseppe Vizzari

Department of Computer Science, Systems and Communications, University of Milan-Bicocca, Italy

Keywords: Multi-Agent Systems, 3D modelling, Simulation.

Abstract: The paper introduces a Multi Agent Systems (MAS) approach to crowd modelling and simulation, based on the Situated Cellular Agents (SCA) model. This is a special class of Multilayered Multi Agent Situated System (MMASS), exploiting basic elements of Cellular Automata. In particular SCA model provides an explicit spatial representation and the definition of adjacency geometries, but also a concept of autonomous agent, provided with an internal architecture, an individual state and behaviour. The latter provides different means of space-mediated interaction among agents: synchronous, between adjacent agents, and asynchronous among at-a-distance entities. Heterogeneous entities may be modelled through the specification of different agent types, defining different behaviours and perceptive capabilities. After a brief description of the model, its application to simple crowd behaviours will be given, and an application providing the integration of a bidimensional simulator based on this model and a 3D modelling application (3D Studio) will also be described. The adoption of this kind of system allows the specification and simulation of an architectural design with reference to the behaviour of entities that will act in it. The system is also able to easily produce a realistic visualization of the simulation, in order to facilitate the evaluation of the design and the communication with involved decision-makers. In fact, while experts often require only abstract and analytical results deriving from a quantitative analysis of simulation results, other people involved in the decision-making process related to the design may be helped by qualitative aspects better represented by other forms of graphical visualization.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 161-175.

1. INTRODUCTION

Designing different kinds of environmental structures, at different detail levels, from the corridors or emergency exits of a building to the whole transportation system on urban or regional scale, requires some kind of simulation system, in order to evaluate strategies and designs before their actual implementation. There are different approaches to simulation, based on various theoretical models, ranging from analytical ones (Helbing [1991]) to those based on Cellular Automata (Wolfram [1986]). Rather than tackling simulation scenarios in a global way, defining centralized solution mechanisms that manage different aspects of the modelled system, they can be suitably reformulated in terms of local interacting entities, which try to achieve their own goals (explicitly specified or implicitly emerging from their collective behaviour), by means of coordination or competition schemes. The solution, global system behaviour, emerges as an effect of agents' individual local behavior (Ferber and Drogoul [1992]). In this framework, approaches based on Multi Agent System (MAS) principles (Ferber [1991]) that propose to focus on interaction aspects of agent groups and crowds have been defined. These works have shown that intelligent group behavior and solution to complex problems can be obtained as the result of interactions between agents characterized by a simple internal model (Drogoul [1995]). A MAS could thus represent a mean of modelling those self-organizing systems that carry out the planning operation to obtain the solution of the design problem.

MASs could even be useful in another kind of iterative design process, centered on simulation (Caneparo and Robiglio [2003]). The latter could provide a first phase in which the designer makes some preliminary choices about the environment, the entities that inhabit it and their behaviour, then a cycle of simulations is performed and the design can be evaluated (by the expert, maybe assisted by some semi-automatic mechanism provided by the system itself). If necessary the design can be then suitably modified by the designer, possibly supported by the system. Then this simulation, evaluation and adaptation cycle could be iterated until the design produces results that are considered acceptable.

The acceptance of the design is generally based on some kind of quantitative analysis of the simulation results performed by experts, but especially in this area there are often a many stakeholders involved in the decision-making process that are not able to understand this kind of information. Where tables, graphs and analytical approaches may not be effective, a realistic visualization of simulation dynamics can integrate them and allow non-experts to better understand effects of the design.

The aim of this paper is to describe the Situated Cellular Automata (SCA) model, a particular class of Multilayered Multi Agent Situated Systems (MMASS Bandini et al. [2002]), that has been designed for situations in which an explicit representation of the spatial structure of the environment is a crucial factor. This structure can be regular or irregular and agents' behaviour is strongly influenced by their position, as it is determined as a consequence of synchronous interaction with other adjacent entities (i.e. *reaction*) or according to the perception of signals asynchronously emitted by at-a-distance agents (i.e. *field diffusion*). Synchronous reaction can represent a direct cooperation between neighbours, and operates in a way that is very similar to CA's transition rule. Remote interaction, implemented through the field emission-diffusion-perception mechanism represents instead a mean of modelling the concept of locality. The latter is generally obtained in CA-based approaches through extensions of the basic model (see, e.g., long-range neighborhood extensions in White and Engelen [1997] or Evans [2003]). The model allows thus to represent heterogeneity in the spatial structure, in agent behaviour and interaction.

The following Section briefly describes the SCA model, focusing on agent interaction, while Section 3 describes how to exploit it in order to represent crowd behaviours. Later two sample applications will be shown, respectively in indoor and outdoor situations. Conclusions and future developments will end the paper.

2. SCA AT A GLANCE

The Situated Cellular Agents (SCA) model defines MAS that are situated in environments whose structure is defined as an undirected graph of sites. This $Space$ may represent an abstraction of a physical space, but also a conceptual one, for instance the adjacency relations in “logical” domains, such as the space of collaborations between roles in a business unit or the fact that different entities share an interest with reference to some kind of topic.

A SCA agent is defined by the 3-tuple $\langle s, p, \tau \rangle$ where $s \in \Sigma_\tau$ denotes the current *agent state* and can assume one of the values specified by its type, $p \in P$ is the site of the $Space$ where the agent is situated, and τ is the *agent type*. The latter allows the definition of heterogeneous agents (i.e. agents of different type), and the related perceptive capabilities and behavioural specification. An agent type τ is thus defined by the 3-tuple $\langle \Sigma_\tau, Perception_\tau, Action_\tau \rangle$, where: Σ_τ defines the set of states that agents of type τ can assume and $Perception_\tau : \Sigma_\tau \rightarrow \{N \times W_{f_1}\} \times \ldots \times \{N \times W_{f_{|F|}}\}$ is a function associating to each agent state a vector of pairs representing the receptiveness coefficient and sensitivity thresholds for that kind of field. This function defines the perceptive capabilities for the related agent type; field diffusion and perception mechanism will be better described in Section 2.1. $Action_{\tau}$ represents instead the behavioural specification for agents of type τ , and will be further described in Section 2.2.

Agent may interact in two different ways, both dependant on the environment and on the position of the involved entities. In particular two or more agents may interact if they are adjacent in their $Space$ they agree to synchronously change their state with a *reaction* operation. A field emission-diffusion-perception mechanism allows instead communication among non adjacent agents. An agent emits a field that is, it generates a signal defining its parameters (i.e. intensity value, diffusion function, and so on), and this signal propagates throughout the space according to its diffusion function. The $Perception_{\tau}$ function, characterizing each agent type τ , defines the other side of an asynchronous interaction among agents: that is, the possible reception of signals conveyed through a field diffused in the environment, if the sensitivity of the agent to the field is such that it can perceive it. This means that a field can be neglected by an agent of type τ if its value at the site where the agent is situated is less than the agent sensitivity threshold computed by the $Perception_{\tau}$ function.

Field perception constitutes a fundamental aspect of the perception-deliberation-action mechanism that specifies MMASS *agent behavior*. This mechanism describes agents as characterized by a set of possible actions, and a mechanism for the selection of the action to be undertaken based on the internal state and the position of the agents themselves. The set of possible actions (i.e. $Action_{\tau}$) specifies whether and how agents of type τ change their state and/or position, how they interact with other agents, and how neighboring agents can influence them. In the following more details of agent interaction and behaviour will be given. The suitable definitions field types, field sources and triggered behaviours for various agent types may allow to model effects of *attraction* and *repulsion* guiding agent movement in the environment. In this way it is possible to obtain agent movement without the need of providing each agent with an internal representation of the environment on which search algorithms and pathfinding techniques can be applied. Moreover, as agents are field sources, dynamic aspects of the environment such as their density in specific areas (i.e. crowding) may be modelled and considered by agents in the choices related to their movement (e.g. avoid or seek groups).

2.1 At-a-distance agent interaction

Each SCA agent is provided with a set of sensors that allow its interaction with the environment and other agents. At the same time, agents are influenced by the perception of those signals and can assume the role of source of given fields acting within the spatial structure.

Each field is characterized by the *set of values* that it can assume during its propagation throughout the space, a *diffusion* function, and field *comparison* and field *composition* functions that define field manipulation. Formally, a field $f \in F$ is defined by $\langle W_f, Diffusion_f, Compare_f, Compose_f \rangle$ where W_f denotes the set of values that the field can assume; $Diffusion_f : P \times W_f \times P \rightarrow (W_f)^+$ is the diffusion function of the field computing the value of a field on a given space site taking into account in which site and with which value it has been generated (since the structure of a $Space$ is generally not regular and paths of different length can connect each pair of sites, $Diffusion_f$ may return a number of values depending on the number of paths connecting the source site with each other site). $Compose_f : (W_f)^+ \rightarrow W_f$ indicates how field values have to be combined (for instance, in order to obtain the unique value of a field type at a site), and $Compare_f : W_f \times W_f \rightarrow \{true, false\}$ is the function that compares field values. For instance, in order to verify whether an agent can perceive a field, its value on the site it is placed on (modified according to the sensitivity coefficient of the agent type) and the agent sensitivity threshold are compared by this function (more details on agent perception will be given in the following subsection).

An agent interaction mechanism, regardless the specific kind of involved sense (e.g. sight, hearing capabilities of a man) can be defined through the specification of:

field sources that can correspond to agents (that may model objects). For instance, fields can be emitted by agents to indicate their availability to fulfill given tasks (e.g. a door, a guide in a museum, a phone booth);

functions to define diffusion, specifying how field values have to be modulated (e.g. when an agent moves far from a group of agents its view must be reduced and it is no more visible when he goes out from a room);

field sensors and *perception functions* associated to each agent, allowing the representation of different and dynamic agent perceptive abilities, that can be dependent on agent state, goals, and context (e.g. agent sensitivity to the presence of a fire exit must be higher in an emergency situation, but its sight may be impaired by the smoke).

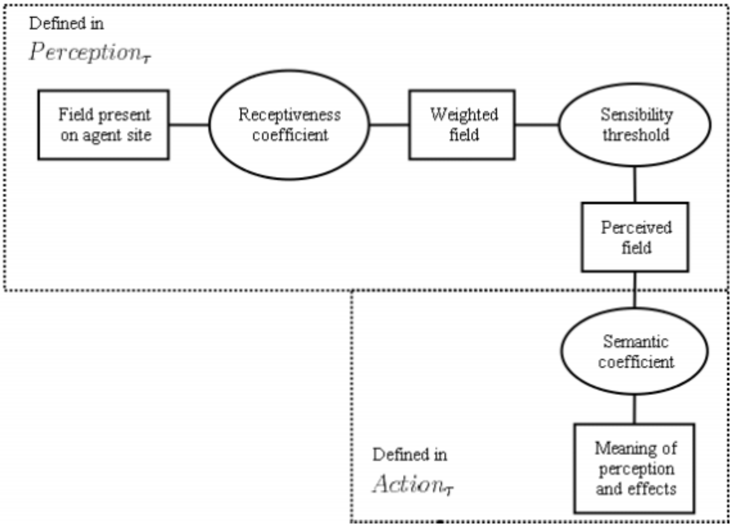


Figure 1. Bidimensional abstraction of a 3D environment

2.2 Agent behaviour

SCA agents are entities that are situated in an environment and that, according to their state, perceive their local environment (*perception*) and select an operation from a set of possible actions (*deliberation*) in order to move, modify their state, and interact with other agents (*action*).

A formal specification of agent perception mechanism and function can be found in (Bandini et al. [2002]). Field perception for each agent type and each field can be defined through the specification of the set of *fields* that agents are able to perceive (i.e. maximal agent capabilities, e.g. deaf agents can not perceive sound fields); a *sensitivity threshold* that indicates, according to agent state, the minimum field value that the agent is able to perceive (e.g. when they are involved in a conversation, agents are less sensitive to surrounding noises); a *sensitivity coefficient* that modulates field values according to agent state (e.g. when an agent is in a hurry, it is more sensitive to exit and elevator signs). The perception mechanism is summarized in *Figure 1*: the first part is related to the physical possibility to perceive a signal in a certain situation, while the second one refers to the semantic value that the agent assigns to the perception in the current circumstances. With reference to the previous example, related to the perception of a human agent in a fire emergency situation, the first part of the mechanism is used to specify that smoke reduces agent possibility to perceive visual signals, and this can be modelled as a change in agent state upon the perception of smoke. The reaction to this perception may also make the agent more sensitive to signals related to emergency exits (i.e. rise the related sensitivity coefficient). The global effect can be fine tuned through a suitable definition of coefficients and thresholds, but the model can represent physical impairment (or improvement) related to perception and increased attention related to semantic aspects.

Agent behaviour can be specified using the L*MASS (Bandini et al. [2001]) language that defines the following primitives:

$reaction(s, a_{p_1}, ..., a_{p_n}, s')$: this primitive defined for agent a situated in the site p allows it to synchronously interact with agents $a_{p_1}, ..., a_{p_n}$ situated in $p_1, ..., p_n$ adjacent to p, that have agreed to take part in the interaction; the effect of this interaction is the change of its state from s to s';

$emit(s, f, p)$: the *emit* primitive allows an agent to start the diffusion of field f on p, that is the site it is placed on;

$trigger(s, f_i, s')$: this primitive specifies that an agent must change its state from s to s' when it perceives a field f_i;

$transport(p, f_i, q)$: the *transport* primitive allows to define agent movement from site p to site q (between whom an adjacency relation must be present) upon reception of field f_i.

For all these primitive, additional conditions (i.e. sort of guards on the execution of the related operation) on agent state, perceived fields and adjacent sites can be specified. In some cases these two parameters can be insufficient, as they are just related to a single site, therefore these conditions can include the intensity of fields present in adjacent sites. For instance, in order to specify a *transport* operation, this is necessary to model the behaviour of an agent wishing to move to the adjacent site with the highest intensity of a certain field. In order to specify that a certain agent of type τ_a can attract agents of type τ_b one must respectively define a field type $F_{a \leftarrow b}$, specifying required parameters, insert a specific *emit* action in $Action_{\tau_a}$ and a *transport* operation in $Action_{\tau_b}$ indicating that the related agent of type τ_b should move towards the adjacent site with the highest value for field type $F_{a \leftarrow b}$. More precisely the *transport* action would be specified as follows:

action: transport(p, $f_{a \leftarrow b}$*, q)*
condit: position(p), empty(q), near(p,q), perceive($f_{a \leftarrow b}$*),best(q)*
effect: position(q), empty(p)

where *p* and *q* are sites, *position(p)* specifies that the related agent is placed in *p*, *empty(q)* indicates that site *q* is not occupied by other agents, *near(p,q)* specifies that the arguments are adjacent sites (i.e. connected by an edge in the spatial structure) and *perceive(* $f_{a\leftarrow b}$*)* indicates that the agent is able to perceive a field $f_{a\leftarrow b} \in F_{a\leftarrow b}$. The additional condition *best(q)* is verified when for all sites *r* adjacent to *p* and currently empty, the intensity of field type $F_{a\leftarrow b}$ is lower or equal than is site *q*. Repulsion requires the same operations, with a difference in the *transport* action, whose destination is the site with the lowest value for the repelling field type. More complex conditions for the transport operation can cause interesting effects, such as an agent that keeps at a certain distance from the source of a specific field type, following thus its movement but avoiding contact. The definition of different field sources and types (or, equivalently, the inclusion of an indication of the related source in the information related to fields) allows the definition of different way-points, intermediate goals in a script that specifies a path in the environment. An agent may be perceptive to the first field type and move towards its source. When the perceived field intensity reaches a certain level (i.e. when the distance has reduced under a certain degree) the agent may change its goal, becoming perceptive to the field emitted by the next way-point. In order to obtain more complex behaviours, for instance related to agents interests, goals, and more autonomous behaviours requires a more composite field definition, that should encapsulate more information than their simple intensity, and different agent actions. A formal description of the introduced modelling elements can be found in Bandini et al. [2004a].

3. CROWD MODELLING AND SCA

As previously specified in Section 2, the SCA model provides a discrete representation of the environment, an abstraction of the actual spatial structure in which the simulation takes place. In order to specify a situation exploiting the concepts defined by the model, the first step is to describe the simulation scenario in terms of a discrete and possibly irregular network of nodes. *Figure 2* shows the 3D representation of a museum room and the related abstraction with a grid structure. Black squares are occupied by walls, grey ones represents artworks, while agents are represented with black circles. The decision on the granularity of the tessellation depends on the features of the scenario and especially on the goal of the simulation: for instance, if the main aim is to evaluate the design of a corridor in an evacuation situation, the spatial abstraction should reflect the actual dimensions of a human body and its space occupancy.

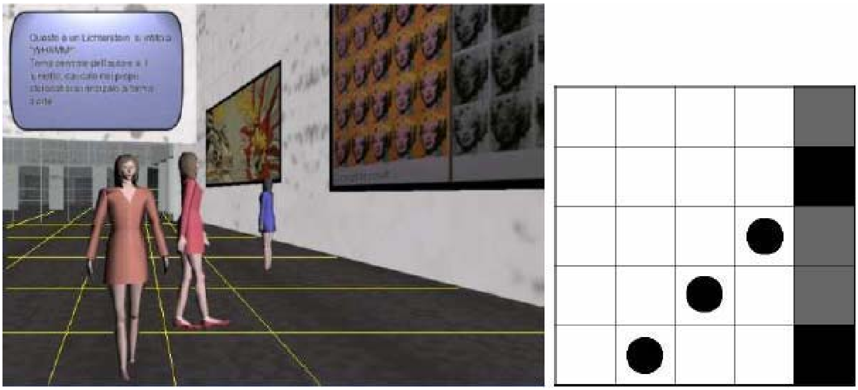


Figure 2. Bidimensional abstraction of a 3D environment

The second step is to describe the behaviour of the different entities placed in the environment, in order to obtain a realistic system dynamic. With reference to the same figure, we could model artworks and doors as sources of fields that can have an attractive effect on agents placed in the environment, according to their own internal state and goals. In fact agents may be perceptive to fields emitted by artworks placed at a distance lower than a specified threshold, and be attracted for a specific time interval, but only if they still have not already observed it. They may also be perceptive to doors and passageways, with a different priority. For instance they may decide to move from one room to another (according to a specific order among rooms), following the field emitted by the related door, after they have observed every artwork present in the room.

The placement of field sources must be specified with considering the diffusion function related to field types: in order to obtain a realistic behaviour of agents they should not be able to perceive the presence of an artwork if they are not able to see it. This consideration indicates that the diffusion of some signals can exploit the actual 3D model of the environment for the diffusion of specific field types, while for other ones (e.g. audible signals) the bidimensional abstraction can be enough to obtain believable system behaviour.

The definition of fields and diffusion functions, with reference to the representation of the environment, is just one side of behaviour modelling. The other is related to the specification of the reaction to the perception of a certain signal, in the form of actions undertaken by agents. Different entities may react in a completely different way to the perception of the same field, and even the same agent can perform different actions according to its own state. For instance, in order to specify that an agent should follow a specific path, the related way-points could be associated to field sources. The agent could be sensitive to the field emitted by the closest way-point and moving towards it, becoming sensitive to the next one once its distance becomes lower than a certain threshold. Moreover the way-points could be exploited by different agent types related to different paths. In other words the order of the points to visit could be defined in the behaviour specification related to a specific agent type, and sources could be just relevant points indicating their presence through the emission of a presence field.

In the same way effects of attraction and repulsion can be defined in order to fine tune the behaviour of various entities roaming in the environment according to its infrastructure. Anyway agent movement can also be based on the behaviour of other active entities. In other words agents may be at the same time affected by fields but also sources of signals affecting other entities. For instance, a specific agent may be the source of a presence field that is considered attractive by a specific agent type. In this way crowds can concentrate around a leader and follow him/her in a procession (see *Figure 3*). In a similar way lanes and queues can be obtained specifying that every agent is only sensitive to the signals emitted by the preceding one, and having leaders that guide the crowd, following specified paths.

Figure 3. A screenshot of the 3D application showing the formation of a procession

4. SAMPLE SIMULATIONS

One of the applications developed to implement SCA based simulations exploits a simulator based on a bidimensional spatial structure representation and an existing commercial 3D modelling instrument (3D Studio MAX). The simulator has been developed as experimentation and exploitation of a long term project for a platform for MMASS based simulations (Bandini et al. [2004b]). This software is based on the Java platform and its goal is to implement basic elements and mechanisms of the MMASS model in order to allow a user to rapidly use these components to build a simulation.

This simulator produces results that can undergo a quantitative analysis whose results can be easily understood by experts of the application area. In different situations it can be useful, for sake of communication with non-experts, to obtain a more effective visualization of simulation dynamics. To do so, the bidimensional simulator produces a log-file provided with a fixed-record structure, in which every record is related to a node of the spatial structure or the position of an agent with reference to this structure. Initially, the simulator prints the structure of the environment, then the starting position of each agent. For every iteration of the simulation the new position of every agent is also printed. This file is parsed by a 3D Studio Max script which generates a plane and walls related to the spatial structure, nodes related to sites, and bipeds related to agents. Splines are then generated starting from the discrete positions assumed by various agents, and represent bipeds' movement. This process introduces modifications to trajectories defined by the bidimensional simulator whose sense is to give a more realistic movement to agents' avatars. This application was tested in an abstract indoor situation, related to an evacuation scenario (see *Figure 4*). Agents placed in this room are attracted by the door, which is source of a field that generates a gradient that can be ‘climbed’ by the entities placed in the room that are able to perceive it.

The perceptive capabilities, and especially agent behaviour, is object of current developments in collaboration with sociologists and psychologists: in fact it is possible to model psychological data and knowledge in order to define specific fields (e.g. presence signals that can model crowding) and *transport* actions for various agent types. For instance very lucid agents could favor sites where there are signs of an exit (whatever it is) and the crowd density is low. For instance this can be obtained introducing a presence field, which is emitted by every agent, whose *Compose* function simply sums the intensity of signals emitted by various agents. Other agents instead may head towards the closest exit, completely disregarding overcrowding situations or simply stand still. These considerations can be suitably represented by conditions specified for every *transport* operation in the behavioural specification of an agent type.

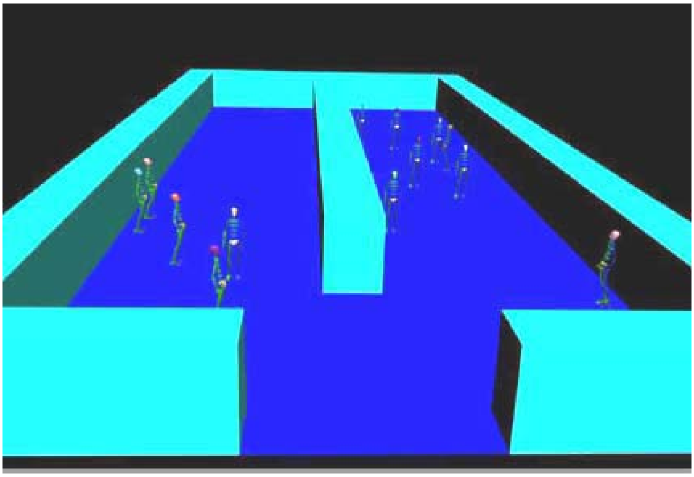


Figure 4. A screenshot of the animation produced by the 3D modelling tool in an indoor scenario

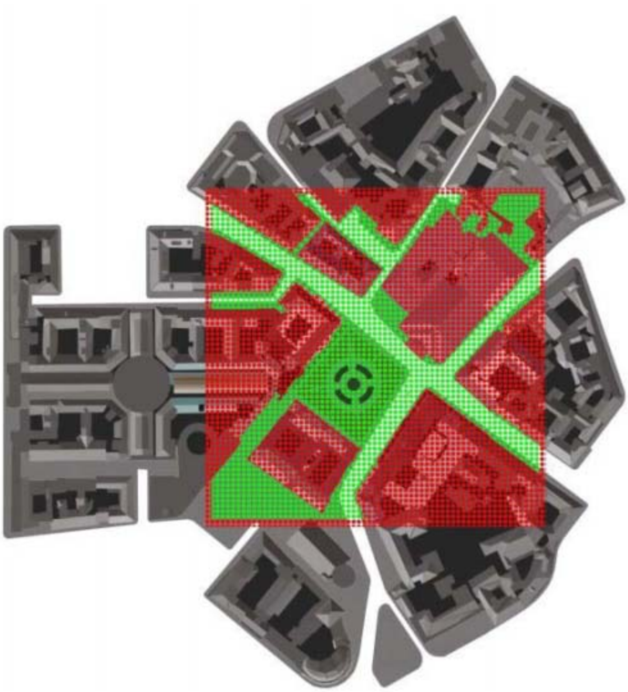


Figure 5. A sketch of the bidimensional spatial structure editor

The interaction mechanism between the simulator and the 3D modelling tool is currently being modified in order to allow an easy integration with existing models of the environment. By doing so, it will be possible to draw the bidimensional abstraction of the space directly on images obtained by the 3D modelling tool.*Figure 5* shows a top view of a Scala Square in Milan, and a lattice of nodes has been generated in order to define a spatial structure suitable for adoption in a SCA model.

Figure 6. A screenshot of the animation produced by the 3D modelling tool in an outdoor scenario (the 3D model of Scala Square appears courtesy of Geosim systems)

Some nodes are related to portions of space that can be occupied by agents, while other ones are already occupied by static objects. This instrument allows the definition of a graph structure, inserting and deleting nodes and edges, starting from scratch or modifying predefined structures like regular lattices. The development of this kind tool for the specification of the environment is currently on-going and will allow an easy adoption of these models and instruments even in a realistic outdoor scenario.

As a preliminary result, some sample animations related to the area of Scala Square were produced: a screenshot of these animations is shown in *Figure 6*[1]. Agents related to passers-by, heading for different ways out of the square,

[1] The 3D model of Scala Square appears courtesy of Geosim systems.

only walk on “legal” spaces (i.e. sidewalks and zebra crossings), avoiding each others and, if needed, waiting for other agents to move out of their way. Their trajectory, as previously specified, is actually smoothened by a script that parses the simulation log and produces agents’ trajectories in the 3D modelling tool.

5. CONCLUSION AND FUTURE DEVELOPMENTS

This paper has presented the application of the SCA model to crowd modelling, simulation and visualization. The model provides the possibility to explicitly define a spatial structure of the environment in which the simulation takes place. Relevant objects are modelled as sources of fields, signals that diffuse in the environment and can be perceived by agents. The reaction to the perception of these fields is defined by agent type, which also specifies perceptive capabilities with reference to their state. The interaction model defined by the SCA model provides the possibility to obtain agents able to act and interact in an environment according to their *context*, in terms of spatial relationships among agents and local perception of signals. In this way it is possible to obtain agents that are more autonomous, meaning by that that they do not need an internal representation of the environment, on which traditional search algorithms or pathfinding techniques must be applied, but are instead guided by signals perceived in the environment. Moreover dynamical aspects of the environment can be modelled (e.g. the crowding of an area) through the field emission-diffusion-perception mechanism, and agents can adapt their behaviour to these additional contextual conditions. For instance agents may thus try to avoid or seek crowded areas, according to the behavioural specification that has been adopted for the simulation.

Currently the SCA model is being applied to simulation supporting localization (Bandini et al. [2003]), design of environments, but it is also being considered as an instrument for urban and environmental planning. The applications that were described in this paper provide the integration of a simulator and a 3D modelling tool, but the design of an integrated simulator and 3D engine for the development of real-time dynamic (and possibly interactive) applications exploiting elements of the model is currently under-way.

REFERECENCES

Bandini, S., Manzoni, S., Pavesi, G. and Simone, C., 2001: A language for situated multi-agent systems, In Esposito, F., editor, *AI*IA 2001: Advances in Artificial Intelligence*, volume 2175 of Lecture Notes in Computer Science, Springer–Verlag, p. 249–254.

Bandini, S., Manzoni, S., Simone, C., 2002: Dealing with space in multi-agent system: a model for situated MAS. In Castelfranchi, C., Johnson, L., eds.: *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2002)*, ACM Press, p. 1183–1190.

Bandini, S., S. Manzoni, and G. Vizzari, 2003, MMASS Approach to Localization Problems. In: *Proceedings of IAT 2003 -- The 2003 IEEE/WIC International Conference on Intelligent Agent Technology,* IEEE Computer Society, p. 357–363.

Bandini, S., S. Manzoni, and G. Vizzari, 2004a, Situated cellular agents: a model to simulate crowding dynamics. *IEICE Transactions on Information and Systems: Special Section on Cellular Automata,* E87-D(3), p. 669–676.

Bandini, S., S. Manzoni, and G. Vizzari, 2004b – to appear, Towards a specification and execution environment for simulations based on MMASS: Managing at–a–distance interaction. In: *Fourth International Symposium From Agent Theory to Agent Implementation*.

Caneparo, L. and M. Robiglio, 2003, Urbanlab: Agent–based simulation of urban and regional dynamics. In: *Digital Design – Research and Practice – Proceedings of the 10th International Conference on Computer Aided Architectural Design Futures*. Kluwer Academic Publisher.

Drogoul, A, 1995, When ants play chess (or can strategies emerge from tactical behaviors?). In *Proceedings of 5th European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW93)*, volume 957 of Lecture Notes in Computer Science, Springer–Verlag, p. 13–27.

Evans, K M., 2003, Larger than Life: threshold-range scaling of Life's coherent structures, *Physica D*, 183(1-2), p. 45–67.

Ferber, J, 1991. *Multi–Agent Systems*. Addison–Wesley, Harlow, UK.

Ferber, J. and A. Drogoul, 1992, Using Reactive Multi-Agent Systems in Simulation and Problem Solving, chapter *Distributed Artificial Intelligence: Theory and Praxis*, Kluwer Academic Publishers, p. 53–80.

Helbing, D, 1991, A mathematical model for the behavior of pedestrians. *Behavioral Science*, 36, p. 298–310.

White, R. and G. Engelen, 1997, Cellular Automata as the Basis of Integrated Dynamic Regional Modelling*, Environment and Planning B: Planning and Design*, 24(2), p. 235–246.

Wolfram, S., 1986, *Theory and Applications of Cellular Automata*. World Scientific, Singapore.

Using a Spatial Microsimulation Decision Support System for Policy Scenario Analysis

Dimitris Ballas[1], Richard Kingston[2], and John Stillwell[3]
[1] *Department of Geography, University of Sheffield, England*
[2] *School of Planning and Landscape, University of Manchester, England*
[3] *School of Geography, University of Leeds, England*

Keywords: Spatial Microsimulation, Spatial Decision Support Systems, Geotools.

Abstract: This paper discusses the potential of a spatial microsimulation-based decision support system for policy analysis. The system can be used to describe current conditions and issues in neighbourhoods, predict future trends in the composition and health of neighbourhoods and conduct modelling and predictive analysis to measure the likely impact of policy interventions at the local level. A large dynamic spatial micro-simulation model is being constructed for the population of Leeds (approximately 715,000 individuals) based on spatial microsimulation techniques in conjunction with a range of data, including 2001 Census data for Output Areas and sample data from the British Household Panel Survey. The project has three main aims as follows: (i) to develop a static microsimulation model to describe current conditions in Leeds; (ii) to enable the performance of 'What if?' analysis on a range of policy scenarios; and (iii) to develop a dynamic microsimulation model to predict future conditions in Leeds under different policy scenarios. The paper reports progress in meeting the above aims and outlines the associated difficulties and data issues. One of the significant advantages of the spatial microsimulation approach adopted by this project is that it enables the user to query any combination of variables that is deemed desirable for policy analysis. The paper will illustrate the software tool being developed in the context of this project that is capable of carrying out queries of this type and of mapping their results. The decision support tool is being developed to support policy-makers concerned with urban regeneration and neighbourhood renewal.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 177-191.

1. INTRODUCTION

This paper presents a spatial microsimulation-based decision support system. Traditionally, confidentiality concerns have been the main reason why demographic and socio-economic data on individuals, despite being collected from censuses and surveys, have not been available for researchers. Spatial microsimulation is a methodology that attempts to estimate the demographic and socio-economic characteristics of human behaviour of individual people or households (Clarke, 1996). Ballas and Clarke (2003) have recently provided a detailed review of the development of the methodology from Orcutt's original work in the 1950s (Orcutt, 1957).

In this paper we present a spatial decision support system for Leeds City Council. The system is based on a microsimulation model which is capable of constructing a list of 715 thousand individuals within households along with their associated attributes for any point in time, past or future. There are various different ways of calibrating the model but the results are particularly valuable because they combine data from different sources to provide estimates of the probabilities that individuals or households will have particular characteristics and thus create new population cross-classifications unavailable from published sources. So, for example, it becomes possible to identify individuals with the characteristics of being aged 18, a lone parent, unemployed and living in private accommodation in an area prone to high levels of crime. Alternatively, households can be identified in the outer suburbs that contain five persons and have a head of household who is a professional working in another city and earning over £50,000 per year. Once the long list of individuals and their attributes has been simulated, the individuals and households (and the attributes which they possess) can be aggregated to any geographical scale which is deemed appropriate such as output areas, wards or postal sectors for example, or more specific areas designed for policy implementation such as regeneration areas.

The system presented in this paper utilises spatial microsimulation techniques to provide a spatial decision support tool for local council officers. In particular, the system can be used to describe current conditions and issues in neighbourhoods, predict future trends in the composition and health of neighbourhoods and conduct modelling and predictive analysis to measure the likely impact of policy interventions at the local level. The overall aims of the project underpinning the system presented here has are as follows: (i) to develop a static microsimulation model to describe current conditions in Leeds; (ii) to enable the performance of '*What if?*' analysis on a range of policy scenarios; and (iii) to develop a dynamic microsimulation model to predict future conditions in Leeds under different policy scenarios.

This paper reports progress in meeting the above aims and outlines the associated difficulties and data issues. In particular, section 2 briefly describes the data used and the combinatorial optimisation spatial microsimulation technique that has been adopted. Section 3 describes the spatial decision support Micro-MaPPAS system, section 4 presents some policy relevant simulation outputs and section 5 offers some concluding comments.

2. DATA AND METHODS

The spatial decision support system presented in this paper utilises a spatial microsimulation model which links a wide range of data sets, including 2001 Census data for Output Areas and sample data from the British Household Panel Survey. The framework for the spatial decision support system presented in this paper is illustrated in Figure 1.

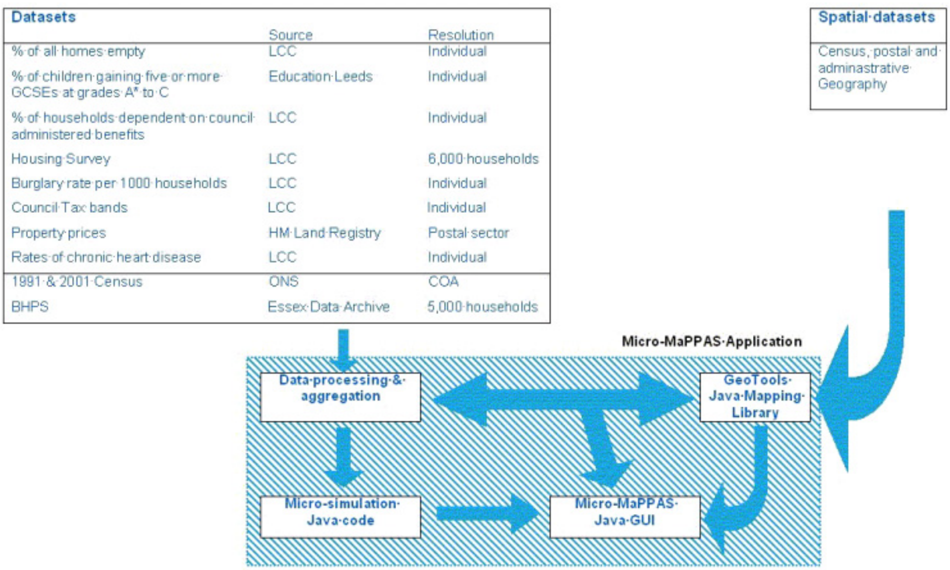


Figure 1. Data sets and system architecture of the Micro-MAPPAS project

The data sets described in figure 1 are linked on the basis of a spatial microsimulation model which implements a combinatorial optimisation approach to generating spatially disaggregated population and household microdata sets at output area (OA) level for the metropolitan district of Leeds as defined in 2001. In particular, the modelling exercise involves the construction of micro-level population using existing 2001 Census Key Statistics (KS) tables for population and household characteristics of all

2,439 OAs in Leeds together with sample data from the British Household Panel Survey (BHPS), a national microdata set of household characteristics. The BHPS contains data on different variables (such as income) for households and their occupants that can be used to derive estimates of 'new' variables for OAs. The technique we have adopted is known as 'simulated annealing' and is distinguished from other methods such as iterative proportional fitting (Ballas, 2001; Norman, 1999; Ballas and Clarke, 2000). Simulated annealing involves reweighting the microdata sample from the BHPS so that it fits OA data for Leeds from the Census. In the first instance, the BHPS microdata set has been reweighted to estimate its parent population at the micro-spatial scale. The BHPS provides a detailed record for a sample of households and all of their occupants. The reweighting method can enable the sampling of this universe of records to find the set of household records that best matches the population described in the KS tables for each OA.

The actual procedure works as follows. First, a series of Census (KS) tables that describe the small area of interest must be selected. The next step is to identify the records of the BHPS microdata that best match these tables. However, there are a vast number of possible sets of households that can be drawn from the BHPS sample. Clearly, it would be impractical to exhaustively consider all possible sets so this is where the simulated annealing[1] (Ballas *et al.,* 1999; 2003) is used to find a set that fits the target tables well. The Micro-MaPPAS simulation model builds on a previous computer software known as SimLeeds (Ballas, 2001) and uses the 10th wave of the BHPS to provide a detailed record for a sample of households and all of their occupants. Ballas *et al.* (2004) describe the combinatorial optimisation method in some detail with the use of illustrative examples. The remainder of this paper describes the various modules of the spatial decision support system and provides some policy relevant simulation outputs.

[1] Annealing is a physical process in which a solid material is first melted in a heat bath by increasing the temperature to a maximum value at which point all particles of the solid have high energies and the freedom to randomly arrange themselves in the liquid phase. This is then followed by a cooling phase, in which the temperature of the heat bath is slowly lowered. The particles of the material attempt to arrange themselves in a low energy state during the cooling phase. When the maximum temperature is sufficiently high and the cooling is carried out sufficiently slowly then all the particles of the material eventually arrange themselves in a state of high density and minimum energy (Kirkpatrick et al. 1983; Dowsland, 1993; Pham and Karaboga, 2000; Van Laarhoven and Aarts, 1987). In geography, simulated annealing has been applied in various contexts for different problems (see for instance Alvanides, 2000; Ballas, 2001; Openshaw and Rao, 1995; Openshaw and Schmidt, 1996; Williamson *et al.*, 1998).

3. THE MICRO-MAPPAS SYSTEM

The Micro-MaPPAS software is written in the Java programming language version 1.4, which means that it can be installed and operated on any computer system and platform. A default set of simulations generated for OAs is loaded when the system is booted up. The structure of the Micro-MaPPAS system is illustrated in Figure 1. Through the graphical user interface (GUI), the user has access to various modules: the model controller, the model diagnostics, the data analyser, the mapping controller and the scenario builders.

First, the *model controller* module allows the user to set the parameters for a small area simulation. The user can set the temperature, the number of model iterations and the number of restarts (Figure 2) and also apply weights to input tables (Figure 3) using slider bars. A simulation can take several hours to run if results are required for all OAs in Leeds and if relatively high temperature and large number of iterations or restarts are selected. It is also possible to run the simulation model for OAs contained within individual wards or community areas.

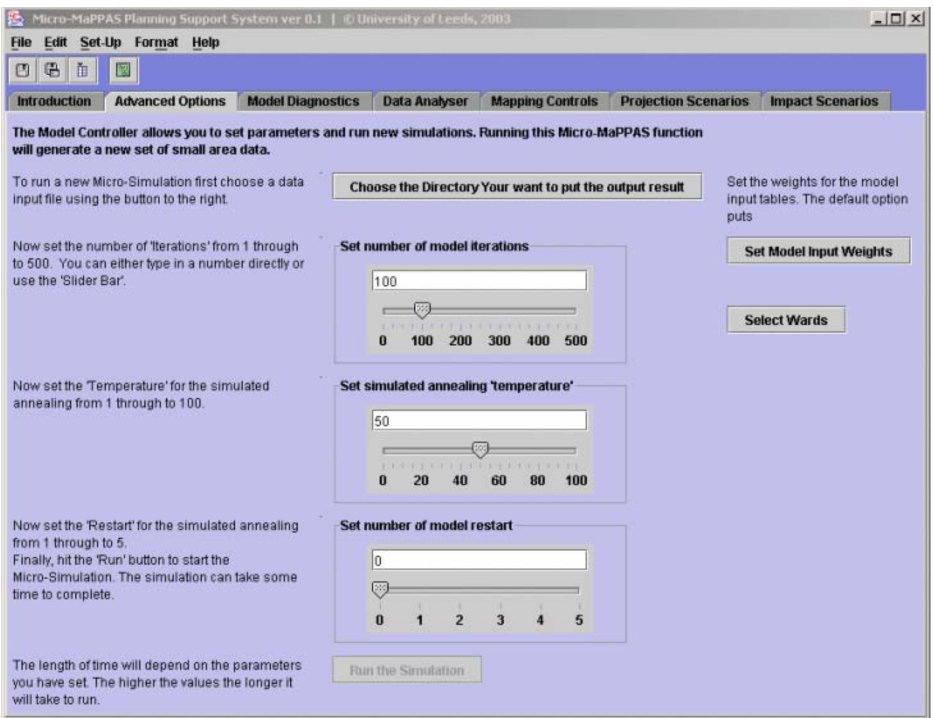


Figure 2. Model controller interface

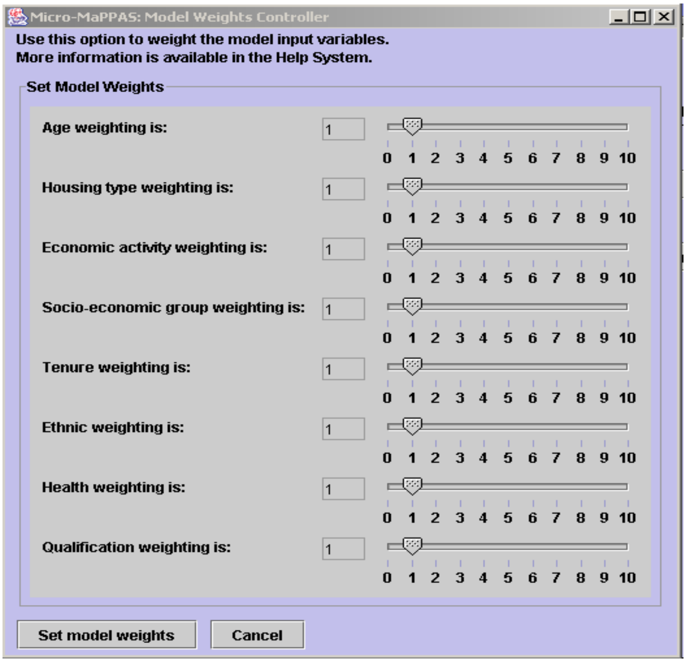
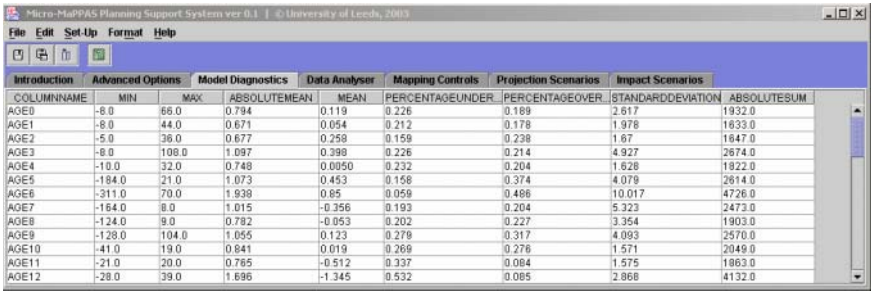


Figure 3. Weights controller interface

The *model diagnostics* module provides details of the accuracy of the microsimulation. It compares the simulated data for OAs with the actual census variables and produces a set of basic statistics (minimum, maximum, absolute mean, mean, standard deviation) and also the percentages of values that have been over-predicted and under-predicted. Each simulation generated has a corresponding model diagnostics table and Figure 4 illustrates the statistics associated with the single year of age variables from 0 to 12. The mean error is lowest (closest to zero) for those aged 10 and highest for those aged 12, for example.



COLUMNNAME	MIN	MAX	ABSOLUTEMEAN	MEAN	PERCENTAGEUNDER	PERCENTAGEOVER	STANDARDDEVIATION	ABSOLUTESUM
AGE0	-8.0	66.0	0.794	0.119	0.226	0.189	2.617	1932.0
AGE1	-8.0	44.0	0.671	0.054	0.212	0.178	1.978	1633.0
AGE2	-5.0	36.0	0.677	0.258	0.159	0.238	1.67	1647.0
AGE3	-8.0	108.0	1.097	0.398	0.226	0.214	4.927	2674.0
AGE4	-10.0	32.0	0.748	0.0050	0.232	0.204	1.628	1822.0
AGE5	-184.0	21.0	1.073	0.453	0.158	0.374	4.079	2614.0
AGE6	-311.0	70.0	1.938	0.85	0.059	0.486	10.017	4726.0
AGE7	-164.0	8.0	1.015	-0.356	0.193	0.204	5.323	2473.0
AGE8	-124.0	9.0	0.782	-0.053	0.202	0.227	3.354	1903.0
AGE9	-128.0	104.0	1.055	0.123	0.279	0.317	4.093	2570.0
AGE10	-41.0	19.0	0.841	0.019	0.269	0.276	1.571	2049.0
AGE11	-21.0	20.0	0.765	-0.512	0.337	0.084	1.575	1863.0
AGE12	-28.0	39.0	1.696	-1.345	0.532	0.085	2.868	4132.0

Figure 4. Model diagnostics table for simulated data for OAs

It should be noted that the model generates a simulated set of data of individuals but these are never visible through the interface. The simulated data are aggregated to OAs and the *data analyser* module provides a table view of this information (Figure 5). Below the table are a number of buttons that enable the user to run queries on the OA data to select the information required but also to aggregate the data to another spatial scale if required. The query builder interface is shown in Figure 5 and, once a query has been constructed, the results are returned to a table in the bottom half of the data analyser window. As an example, a query is undertaken which selects the number of individuals in each ward aged between 20 and 30 whose household income is between £20,000 and £30,000. The results of the query are shown in the bottom section of the data analyser and these can then be mapped.

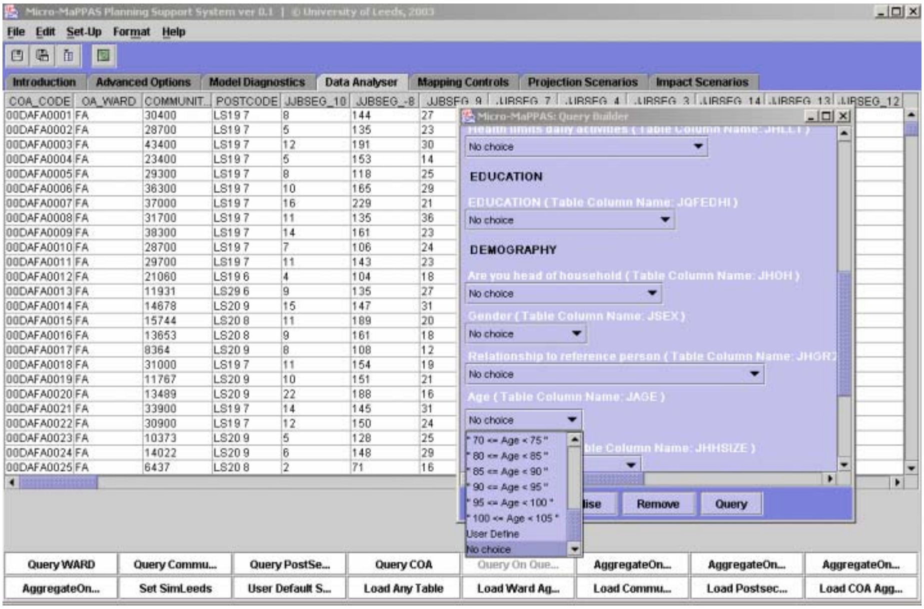


Figure 5. Data analyser and query interface with age drop down menu in use

Further, the *mapping controls* module allows the user to select a variable from a query and map the results at any of the geographical scales of OA, community area, ward or postal sector. Figure 6 illustrates the mapping of the query relating to persons aged 20-30 with incomes of £20-30,000. Mapping functions include panning and zooming and symbology editing.

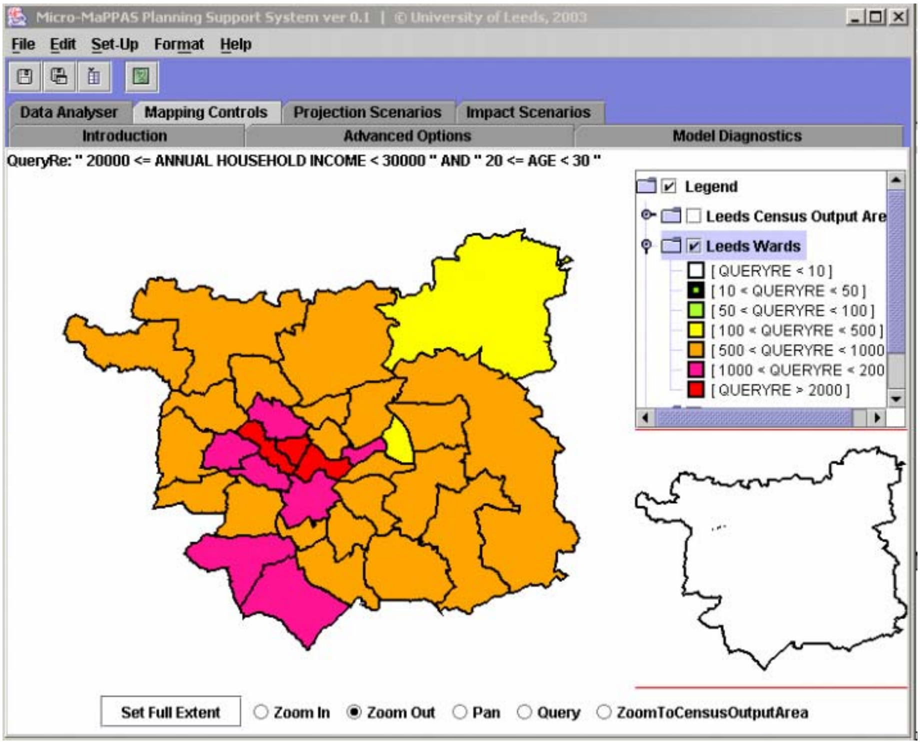


Figure 6: The results of the query as shown in the mapping controller

Finally, one of our immediate priorities is to render the system capable of designing and running simulations for future scenarios based upon different assumptions about population change in the future, and also to undertake some evaluation of 'what if' scenarios. Thus, we are currently developing two additional modules: *projection scenarios* and *impact scenarios* modules.

4. SIMULATION OUTPUTS AND POLICY RELEVANCE

The MicroMaPPAS system has been designed to create sets of simulations for 2001 as the base period. Three examples of sub-group distributions generated by the MicroMaPPAS query are presented by way of illustration. These examples are all produced from the simulation that gives equal weight to all the tables used in the simulation. The first two queries provide more detailed information about the locations of disadvantaged groups. Figure 7 show the distribution of female lone parents with dependent children, where the household income is less than £10,000 per year. The query has been carried out at the scale of community areas and the distribution demonstrates the concentration of this particular subgroup in areas to the east of the city centre (Chapeltown, Burmantofts, Harehills and Seacroft South) and to the south (Beeston Hill, Belle Isle and Middelton).

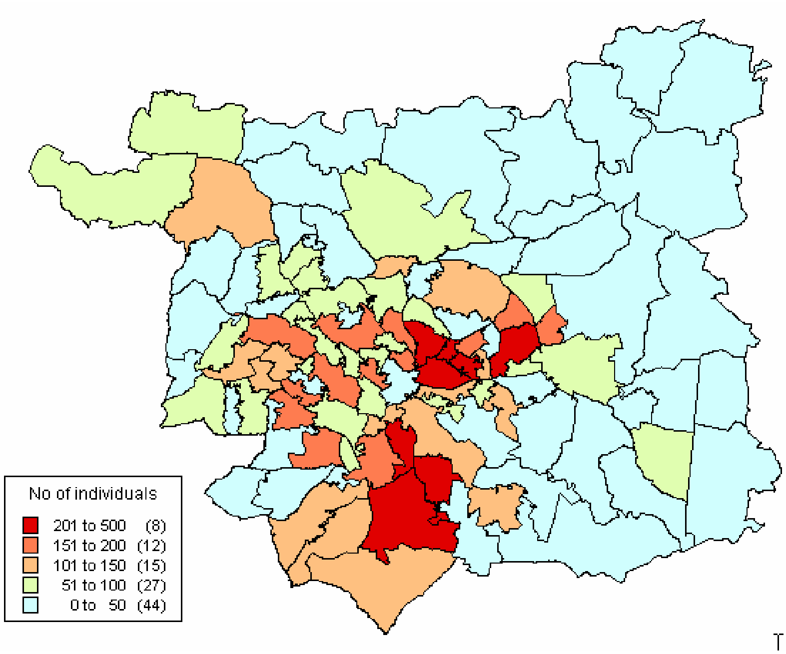


Figure 7. Simulated distribution of female lone parents with dependent children in households with low income by community area, 2001

The second query identifies children aged 0 to 15 in households where income is less than £10,000. In Figure 8, the largest numbers are again found to the east of the city centre in community areas like Halton Moor, Harehills, Gipton North, as well as Seacroft South, with pockets also in Cottingley and Belle Isle in the south

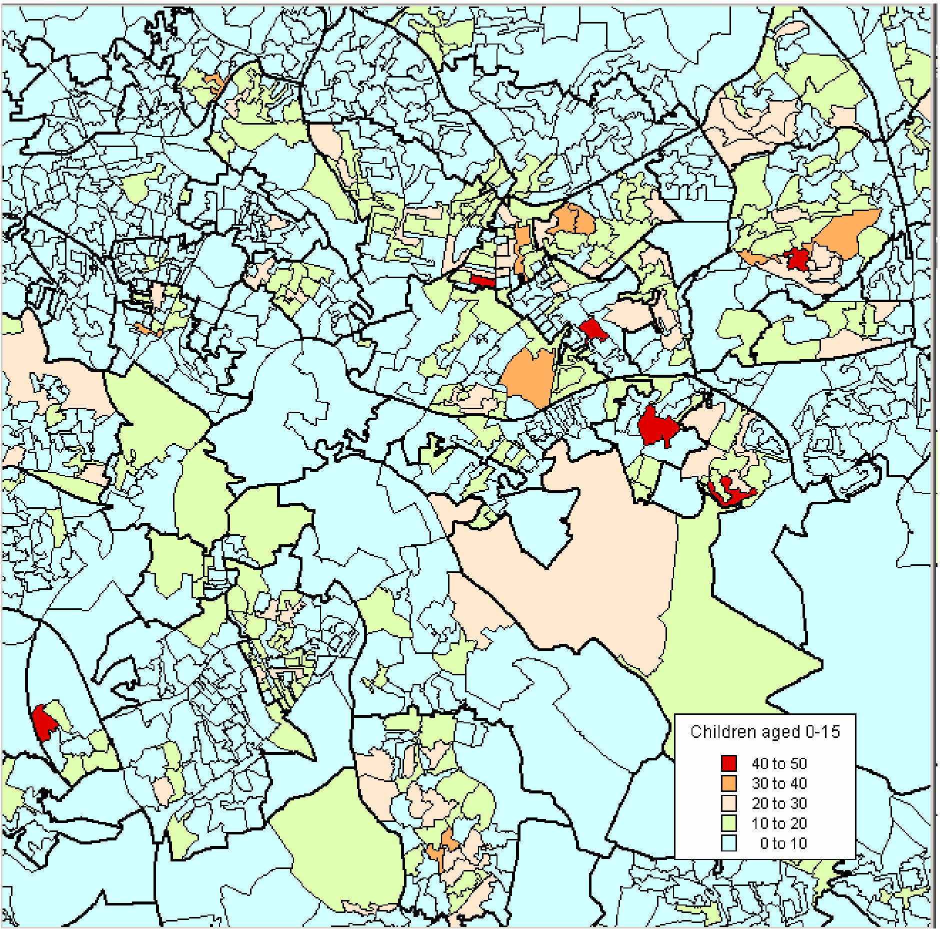


Figure 8. Simulated distribution of children in households with low income by output area, 2001

The third example involves individuals at the other end of the social spectrum. Figure 9 shows the distribution by output area across the whole of the metropolitan district of those individuals living in two person households with no dependent children and with a household income in excess of £50,000. These are the so-called DINKies and their spatial incidence is observed across the north of the district, particularly in the outer suburbs of Harewood and Arthington and Pool in the north and Calverley in the west, but also in pockets in the south.

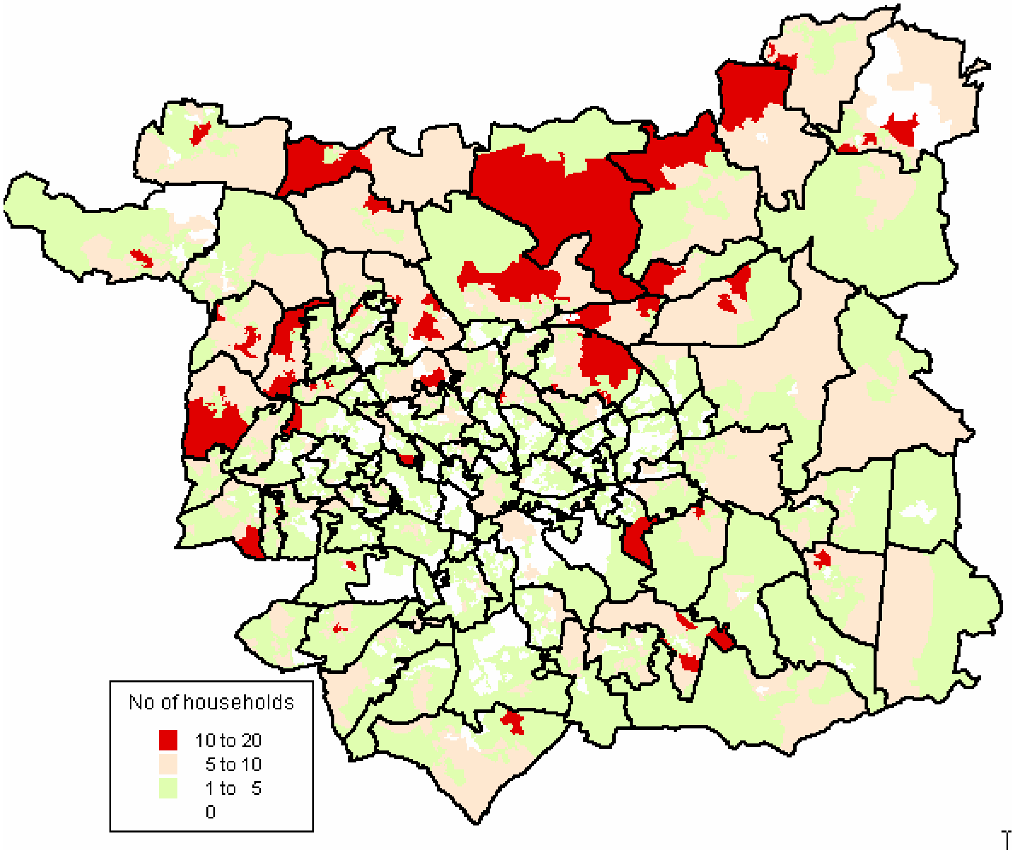


Figure 9. Simulated distribution of two person households without dependent children with high household income, 2001

In addition to modelling the population in 2001, microsimulation can be used to predict the characteristics of the Leeds population in future years. The method we adopt is to build a crude projection model for aggregate ward populations and then to use these totals to constrain more detailed projections generated by the MicroMaPPAS model. In other words, we apply the MicroMaPPAS reweighting methodology to readjust the weights of BHPS households so that they fit small area constraint data in any selected year.

It is important to recognise that, currently, there is no provision of small area projections by a single agency in Britain. The ONS is responsible for creating sub-national population projections for England but these are only produced for local authority areas and on a relatively infrequent basis. A comparison of the projected population for Leeds in 2001 from the latest (1996-based) set of ONS sub-national projections (ONS, 1999) with the mid-year estimate for 2001 indicates an over-projection of nearly 16,000 persons overall. Figure 10 shows the differences between projected counts and estimates for males, females and persons; most of the over-projection has occurred for males and females (to a lesser extent) aged 35 to 54 and these differences offset under-projections of females in their 20s.

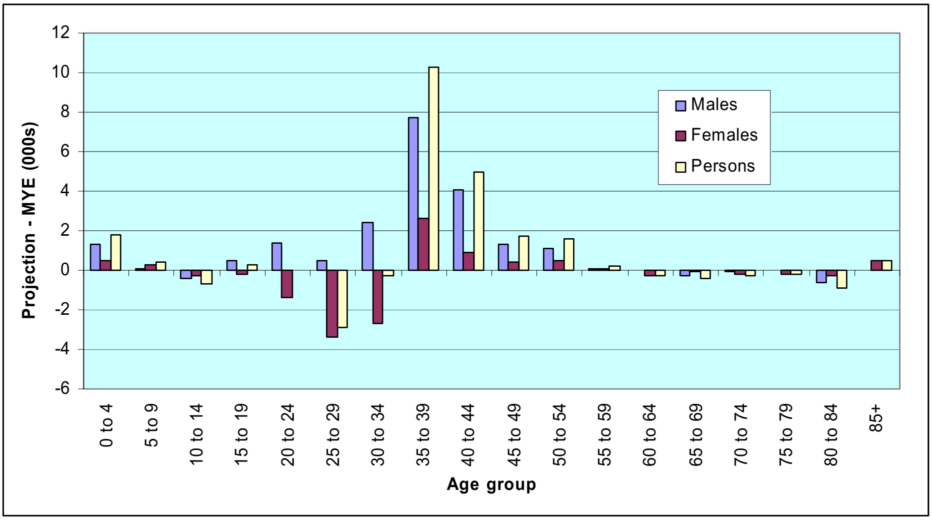


Figure 10. A comparison of 1996-based projections and mid-year estimates for Leeds in 2001

The lessons we take from the comparison are to recognize that even short term projections for relatively large areas may be error-prone and it might not be wise to use the 1996-based district projections as constraints. Consequently, the first step in our methodology is to project the numbers of households and individuals for every year into 2021 by ward within the district independently. In order to do this, we make the simple assumption that the annual rate of change between 1991 and 2001 will continue until 2021. Therefore, we calculate for each ward the annual rates of change between 1991 and 2001 for households and individuals and apply these to successive years after 2001 to give ward-based population totals up to 2021.

The projection of the disaggregated counts of the population in future years according to marital status, socio-economic group, number of cars owned, *et cetera*, is undertaken by applying annual rates of change between 1991 and 2001 in exactly the same way as with the aggregate populations. The projection of these counts is necessary given that these variables are used as constraints in the simulated annealing household reweighting procedure. So, for example, having projected the car ownership characteristics by household, the next step involves calculating the proportions of all car ownership categories in each ward. These proportions are then applied to the projected numbers of households by ward in each future year. In this way, we ensure that the sum of all cars by household categories adds up to the aggregate household projection. The same method can be applied for all other household (e.g tenure) and individual (e.g. ethnic group) variables.

The above discussion implicitly assumes that the 1991 Census recorded accurately the populations living in Leeds wards. However, over 2% of the population was missed overall in 1991 and this underenumeration did not occur uniformly across all areas or age-sex groups (ONS, 2003). Further, the 1991 Census did not record the number of students, which is quite large in some electoral wards such as Headingley. In order to deal with these problems, the following strategies were adopted. To tackle the problem of the undercount in 1991, the ward populations in 1991 can be readjusted on the basis of alternative assumptions on the extent of the undercount. For instance, if it is assumed that the Leeds population in 1991 was underestimated by 4%, the 1991 population numbers can be increased by this rate and the projection procedure described above is applied using the annual rates of change recomputed on this basis. A reasonable solution to the problem of not counting the students in the 1991 Census is to estimate their numbers on the basis of 2001 proportions. For instance, according to the 2001 Census, students in Headingley comprise 54.5% of the total population and this proportion can be added to the published 1991 population total.

5. CONCLUDING COMMENTS

It has long been argued that spatial microsimulation has a great potential for socio-economic impact assessment (see for instance, Ballas and Clarke, 2001a) and for the geographical analysis of the impacts of social policies (Ballas and Clarke, 2001b; Ballas *et al.*, 2003). In this paper we demonstrated this potential further by presenting the first attempt to link spatial microsimulation modelling frameworks to Spatial Decision Support Systems (SDSS). In particular, we presented Micro-MaPPAS, which is a system that added spatial decision making capabilities to a spatial microsimulation model. We believe that systems such as Micro-MaPPAS can play a very important role in the on-going debates on the role of potential of new technologies to promote local democracy and electronic decision-making. This paper addressed the relevant data issues and technical aspects of the linkage of spatial microsimulation modelling frameworks to SDSS and described the capabilities of the Micro-MaPPAS system. We also presented some simulation outputs which demonstrate the policy relevance of the system. We believe that Micro-MaPPAS has a great potential for local policy analysis and may have wider implications for local governance procedures. It also demonstrates the prospects of policy simulation models for the enhancement of local decision making and democracy.

Further, it can be argued that systems such as Micro-MaPPAS developed in JAVA, which is a platform independent programming language, can be put on the World Wide Web and linked to Virtual Decision-Making Environments (VDMEs). The latter are Internet World Wide Web based systems that allow the general public to explore 'real world' problems and become more involved in the public participation processes of the planning system (Kingston *et al.,* 2000). Systems such as Micro-MaPPAS can potentially be used not only to provide information on the possible consequences and the local multiplier effects of major policy changes but also to inform the general public about these and to enhance, in this way, the public participation in policy making procedures (Ballas, Kingston and Stillwell, 2003).

ACKNOWLEDGEMENTS

The research presented in this paper was funded by Leeds City Council. All Census and digitised boundary data used in this paper are Crown Copyright. The BHPS data were obtained from the UK Data Archive (University of Essex). The authors would like to thank Jianhui Jin at the School of Geography, University of Leeds for his contribution to the development of the computer code for the Micro-MaPPAS system.

REFERENCES

Alvanides S, 2000, Zone design methods for application in human Geography, *Unpublished PhD Thesis*, School of Geography, University of Leeds, Leeds.

Ballas, D. 2001, A spatial microsimulation approach to local labour market policy analysis, *Unpublished PhD Thesis*, School of Geography, University of Leeds, Leeds.

Ballas, D., Clarke, G.P. 2000, GIS and microsimulation for local labour market policy analysis*, Computers, Environment and Urban Systems, 24*, 305-330.

Ballas, D. and Clarke, G. P. 2001a, Towards local implications of major job transformations in the city: a spatial microsimulation approach, *Geographical Analysis* 33, 291-311.

Ballas, D., Clarke, G.P. 2001b, Modelling the local impacts of national social policies: a spatial microsimulation approach, *Environment and Planning C: Government and Policy, 19,* 587 – 606.

Ballas, D., Clarke, G.P. 2003, Microsimulation and Regional Science: 30 years of spatial microsimulation of populations, paper presented at the *50th Annual North American Meeting of the Regional Science Association International*, Philadelphia, USA, 19-22 November 2003.

Ballas, D., Clarke, G. P. and Turton, I. 1999, Exploring microsimulation methodologies for the estimation of household attributes, paper presented at the *4th International Conference on GeoComputation*, Fredericksburg, Virginia, USA, 25-28 July.

Ballas, D., Clarke, G. P. and Turton, I. 2003, A spatial microsimulation model for social policy micro-spatial analysis, in Boots, B., Okabe, A. and Thomas, R. (eds.) *Modelling Geographical Systems: Statistical and Computational Applications*, Kluwer, Dordrecht, pp. 143-168.

Ballas, D, Kingston R, Stillwell, J, 2003, Public participation in local policy-making with the use of GIS-based microsimulation, poster presented at the *2nd annual Public Participation GIS Conference*, Portland State University, Portland, USA, 20-22 July 2003

Clarke, G.P. 1996, Microsimulation: an introduction, in G.P. Clarke (ed.) *Microsimulation for Urban and Regional Policy Analysis,* Pion, London, 1-9

Dowsland, K. 1993, Simulated annealing, in Reeves, C. (ed) *Modern Heuristic Techniques for Combinatorial Problems*, Blackwell, Oxford.

Kingston, R., Carver, S., Evans, A. and Turton, I. 2000, Web-Based Public Participation Geographical Information Systems: an aid to Local Environmental Decision-Making, *Computers, Environment and Urban Systems*, 24, 109-125

Kirkpatrick, S, Gelatt, C. D. Jr. and Vecchi, M. P. 1983, Optimization by simulated annealing, *Science*, 220: 671-680.

Norman, P. 1999, *Putting Iterative Proportional Fitting on the Researcher's Desk*, working paper 99/03, School of Geography, University of Leeds, Leeds LS2 9JT

Openshaw, S. and Rao, L. 1995, Algorithms for reengineering 1991 Census geography, *Environment and Planning A*, 27: 425-446.

Openshaw, S. and Schmidt, J. 1996, Parallel simulated annealing and genetic algorithms for re-engineering zoning systems, *Geographical Systems*, 3: 201-220.

Orcutt, G.H. 1957, A new type of socio-economic system, *The review of economics and statistics*, *39*, 116-123

Pham, D.T. and Karaboga, D. 2000, *Intelligent Optimisation Techniques: Genetic Algorithms, Tabu Search, Simulated Annealing and Neural Networks*, Springer, London.

Van Laarhoven, P.J.M. and Aarts, E.H.L. 1987, *Simulated Annealing: Theory and Applications*, Kluwer, Dordrecht.

Williamson, P., Birkin, M. and Rees, P. 1998, The estimation of population microdata by using data from small area statistics and samples of anonymised records, *Environment and Planning A*, 30: 785-816.

Cellular Automata Modeling For Fire Spreading As a Tool to Aid Community-Based Planning for Disaster Mitigation

A. Ohgai, Y. Gohnai, S. Ikaruga[1], M. Murakami[2], and K. Watanabe[3]
Toyohashi University of Technology, Toyohashi, Japan
[1] Yamaguchi University, Yamaguchi, Japan
[2] Kogakuin University, Tokyo, Japan
[3] The University of Tokushima, Tokushima, Japan

Keywords: Community-Based Planning for Disaster Mitigation, Fire Spreading, Fire Fighting Activity, Cellular Automata Modeling, Planning Support Tool.

Abstract: As a tool to support collaboration in community-based planning for disaster mitigation in Japanese old wooden built-up areas, we attempt to develop a fire spreading simulation model incorporated a fire fighting activity using Cellular Automata (CA). The proposed model can deal with the process of fire spreading in a building that traditional models can not represent. Whether or not fire can spread is based on a stochastic calculation process to reproduce uncertain fire spreading. The errors caused by the stochastic factor are analyzed by carrying out simulation two or more times under the same condition. Moreover, the reproductivity of the model is examined by comparing simulation results with actual fire records.

1. INTRODUCTION

There are a lot of old wooden built-up areas in Japan. When the Hanshin-Awaji (Kobe) earthquake occurred in 1997, those areas were badly damaged by building collapse and fire spreading. Compared to other areas, some of the badly damaged areas could be reconstructed rapidly due to the existence of community-based planning practices in such areas. This situation clearly emphasizes the importance of a planning strategy that collaborates with residents to achieve effective disaster mitigation. In general, the planning

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 193-209.

process requires the enhancement of residents' tendencies to participate in the planning process through their heightened awareness of the disaster risk and the need for improvement of the local environment. This enhancement leads to increased and improved collaboration in planning between residents and planning authorities. In order to increase disaster mitigation, an alternative plan based on residents' consensus is required. Due to the difficulty of achieving this enhanced state, as well as in achieving consensus among participants, consensus-building tools are required. In this regard the development of planning support tools in Japan has received widespread attention.

In old wooden built-up areas, the damage caused by the spread of fire may increase significantly. Therefore, it is important to improve fire prevention performance by limiting the spread of fire as much as possible.

Through the comparison of virtual conditions of fire spreading (speed, direction, range, etc.) based on the assumption of certain fire prevention measures (widening of road, establishment of disaster prevention plaza, etc.) district improvement can be visually illustrated to participants and compared to existing circumstances of the specific districts. Participants are then able to confirm for themselves the effect of proposed improvement projects on limiting the spread of fire. Additionally, this could also serve as and effective source of information to educate residents on the hazard of fire. To achieve this, a model that can visually and dynamically simulates the spread of fire based on various assumptions about improvement projects is required.

The previous model focused the reproductively regarding fire spreading. In other words, there is no useful model to serve as a tool for collaboration in community-based planning for disaster mitigation.

From the above-mentioned background, we attempt to develop a fire spreading simulation model, which can plainly offer residents the information on the effect of increase in disaster mitigation performance by improvement of the local environment as well as enlighten them on the disaster risk and the need of the improvement.

2. FIRE SPREADING MODEL AND CA

CA can reproduce a complicated phenomenon by setting up simple rules into lattice space. Many researches by use of this technology have been advancing in some fields; such as transportation, economy, and chemistry. In the field of city planning, CA is used to understand dynamic and complicated urban phenomenon and to predict urbanization and land use change (Batty and Xie, 1994; White and Engelen, 1993; White and Engelen, 1997; Li and Yeh, 2000).

Fire spreading in built-up areas is not natural but a physical phenomenon. In the phenomenon, various factors about circumstances of built-up areas are intricately related. But, by using the technique of CA, the phenomenon must be reproduced by constructing some appropriate rules. Though simulation models of forest fire have been studied using CA (Resnick, 1995), there is little city fire spreading simulation model using CA, but Yamada, Takizawa, et al (1999) and Xie, Sakamoto, et al (2001) are trying to develop the model using CA.

On the other hand, many researchers have been studying about city fire spreading simulation model and a lot of good results have been found in Japan. Because these researches basically use 'building' as a unit of output about whether or not burning occurs in simulation, it is difficult for these models to realistically express the fire spreading process. For example, they cannot deal with gradually spreading in a building with large area. Moreover, they may not provide enough the information on the likely effect of increase in disaster mitigation performance by improvement of the local environment. In this research, by using the results of previous researches regarding fire spreading simulation model and the characteristics of CA, we attempt to develop a model that can simulate fire spreading by using a smaller unit than the ones of the previous researches, which can visually show detailed fire spreading process in built-up areas with wooden buildings.

3. FIRE SPREADING MODELLING

This chapter explains a fire spreading modelling by use of CA, parameters used in the model and the process of simulation.

3.1 Cell size

Naturally, if the size of cell that is a basic unit of fire spreading model is smaller, more detailed spreading simulation becomes possible. However, when the size of cell is too small, the explosive increase in the calculation time and the volume of data occurs. As the purpose of a simulation model proposed here is to show the appearance of fire spreading, the size of cell needs not be small so much. Then, in this research, the size of cell is set at 3 by 3 meters using the following equation representing spread speed to leeward proposed by Hamada (Jirou and Kobayashi, 1997).

$$D = 1.15(5 + 0.5v) \qquad (1)$$

where v is a wind velocity meter per second [m/sec] and D is the limit of distance which fire can spread [m].

3.2 Attribute of cell

From previous researches on fire spreading models in built-up areas, the following three factors of fire spreading are defined: building condition, weather condition and characteristics of built-up area (Murosaki, Ohnishi, et al, 1984). The building condition is composed of building material, floor area, opening area and height that represent the characteristics of building. These attributes relate to the possibility of building burning. The weather condition is composed of wind velocity and direction. These relates to the direction, the range, and the possibility of spreading. The characteristics of built-up area consist of vacant land, pitch of building, and road. These attributes are used as factors relating to the spreading control.

3.3 Neighbourhood

Based on the idea of the limit of distance that fire can spread, which has been brought out by previous researches (Jirou and Kobayashi, 1997), the neighbourhood type of the model is set as shown in *Figure 1*, changing according to wind velocity. Cell *kl* is a causing one of the fire spreading and cell *ij* is a received one in the neighbourhood.

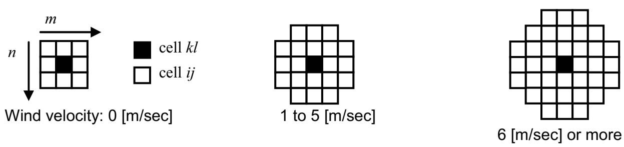


Figure 1. Neighbourhood used in the model

3.4 State of cell

Transition of cell state is set as shown in *Table 1*. The cell with state [1] is not burning yet, but has the possibility of burning. Once the cell catches fire, the state of the cell changes to state [2]. And then, according to the elapsed time after outbreak of fire, the state of the cell changes from state [2] to [4]. But the cell with state [0] is unburnable and never changes during the simulation period.

Table 1. State of a cell

Expression	Explanation
[0]	n_{mn} = 0: Unburnable; Having nothing burning.
[1]	n_{mn} = 1: No burning yet; Having the possibility of burning.
[2]	n_{mn} = 2: Catching fire; Just catching fire, but having no ability to cause fire spreading.
[3]	n_{mn} = 3: On fire; Having the ability to cause fire spreading.
[4]	n_{mn} = 4: Extinction

3.5 Fire spreading probability

The fire spreading judgment, that is whether or not cell *ij* changes from state [1] to state [2], is done by using the fire spreading probability of cell *ij* and the stochastic calculation process in section 3.6. The probability is given by the fire spreading judgment index F_{ij}. This is calculated for all cells *ij* within the neighbourhood of cell *kl* and defined by equation (2).

$$F_{ij} = \alpha \cdot (S_{ij} \cdot P_{ij}) \cdot W_{ij}{}^{\beta} \cdot p(t_{ckl}) \qquad (2)$$

where, S_{ij} is a building structure parameter, P_{ij} is the ratio of the area occupied by wooden or fire preventive wooden building in cell *ij*, W_{ij} is a parameter decided by wind velocity and direction and $p(t_{ckl})$ is the ability of cell *kl* causing fire spreading. These values range from 0 to 1. α is a coefficient to tune the degree of slowdown in spreading caused by the condition except wind velocity ($0 < \alpha \leq 1$), and β is a coefficient to adjust the range and direction of spreading, that is $0 < \beta$.

S_{ij} and P_{ij} are parameters representing combustibility of cell *ij* itself. The value of S_{ij} is given as follows; wooden=1.0, fire preventive wooden=0.6, and fireproof=0.0. P_{ij} plays the role of controlling combustibility by adjusting the value of S_{ij} according to the area of wooden or fire preventive wooden building in cell *ij*. The parameter W_{ij} relates to spreading speed and direction, the value of which is set subject to the distance and direction from cell *kl* to cell *ij*. As shown in *Figure 1*, the number of cells *ij* that is an object of fire spreading judgment increases as wind velocity becomes higher. When wind velocity externally given in the model is 0 to 1 [m/sec], the parameter value of cell *ij* adjacent to cell *kl* is set at 0.5. Other cells are set at 0.3. When wind velocity is more than 1[m/sec], the parameter values variously change as influenced by wind velocity and wind direction. For example, as shown *Figure 2*, in the case of west wind, the value of cell *ij* on the east side of cell *kl* is set at higher than the west side. Moreover, as wind velocity is higher, cell *ij* adjacent to cell *kl* takes a higher value and other cells take a lower value subject to the distance from cell *kl*. Consequently, the possibility of fire spreading becomes higher on the leeward than windward and in the higher wind velocity.

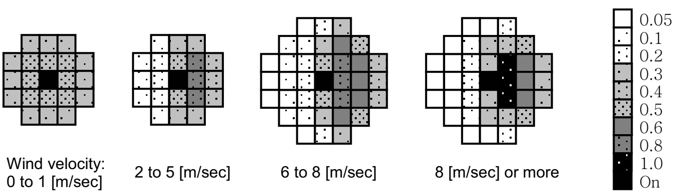


Figure 2. W_{ij} in the case of a west wind

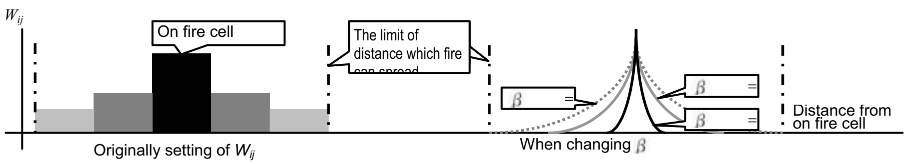


Figure 3. Image of W_{ij} changing by tuning β

In general, the actual area devastated by a fire per unit time varies significantly depending on the whole of condition related to the spreading. For instance, in case of fires at Hanshin-Awaji earthquake, combustion and fire spreading slowed down because the building collapse restricted the supply of air. As a result, the areas devastated by the fires per unit time were smaller than the past fire records. In order to consider the situations such as the above that can not be represented by parameters used in F_{ij}, α is incorporated.

W_{ij} is set so as to decrease linearly subject to the distance from cell *kl* to cell *ij* as shown in *Figure 2*. Thus, by tuning the value of β, a sharp or gradual decrease in the value of W_{ij} is expressible as shown in *Figure 3*. In other words, the coefficient β plays the role of controlling the possibility and range of spreading.

The parameter $p(t_{ckl})$ is set by equation (3) using the indoor temperature standard curve of wooden house at fire as referenced to the literature (Wakamatsu, 1978).

$$p(t_{ckl}) = \begin{cases} \dfrac{4.0}{t_2 - t_1} \cdot t_{ckl} + \dfrac{0.2 \cdot t_2 - 4.2 \cdot t_1}{t_2 - t_1} & \left[t_1 \le t_{ckl} \le \dfrac{t_2 - t_1}{5} + t_1 \right] \\ \dfrac{5}{4 \cdot (t_2 - t_1)} \cdot (-t_{ckl} + t_2) & \left[\dfrac{t_2 - t_1}{5} + t_1 \le t_{ckl} \le t_2 \right] \end{cases} \tag{3}$$

The time t_1 until cell *kl* becomes having the ability to cause fire spreading after catching fire is set at 2 [min]. The reason is below. Horiuchi (Yasuno and Nanba, 1999) proposed the equation as follows.

$$t_x = (3 + 3a/8 + 8d/D)/(1 + 0.1v) \tag{4}$$

where, *a* is a length of average of a side in building [m], *d* is a pitch of building [m] and *D* is the limit of distance which fire can spread [m].

This equation shows the time until an adjacent house to the point of fire origin ignites after catching fire in the origin (Yasuno and Nanba, 1999). In this research, the time t_x was considered as the time until a building becomes having the ability to cause fire spreading after catching fire. But it is by building and thus can not be used in the model simulating by cell. Then, the examination of the value of the time t_1 used in the model was carried out by calculating t_x changing the values of d and v but fixed a at 3 meters. This leads to setting t_1 at 2 minutes.

The time t_2 until cell *kl* burns out after catching fire is defined as follows.

$$t_y = (w/5.5)/(A_w\sqrt{H}/A_f) \quad (5)$$

where, w is a fire load [kg/m^2], A_w is a opening area [m^2], A_f is a floor area [m^2] and H is a height [m].

This equation approximately represents the continuation time of fire as assuming wooden building and constant combustion speed (Sugawara and Naruse, 1997). In this research, we consider the value calculated by this equation as the time until a building burns out after catching fire. However, as t_y is the value by building, unrealistic results can be outputted when carrying out a simulation using the value, for example all cells within a building with large area remain on fire more than a likely burning out time. The reason is because a floor area A_f critically influences the value of t_y. Then, t_2 is set at 10 minutes as a result of calculating t_y by considering a cell as a building of 9 square meters.

The value of W_{ij} depends on the position of cell *ij* within the neighbourhood and a given wind velocity and direction. $p(t_{ckl})$ depends on an elapsed time after outbreak of fire in cell *kl*. Therefore, F_{ij} is not constant in cell *ij* but changes according to W_{ij} and $p(t_{ckl})$.

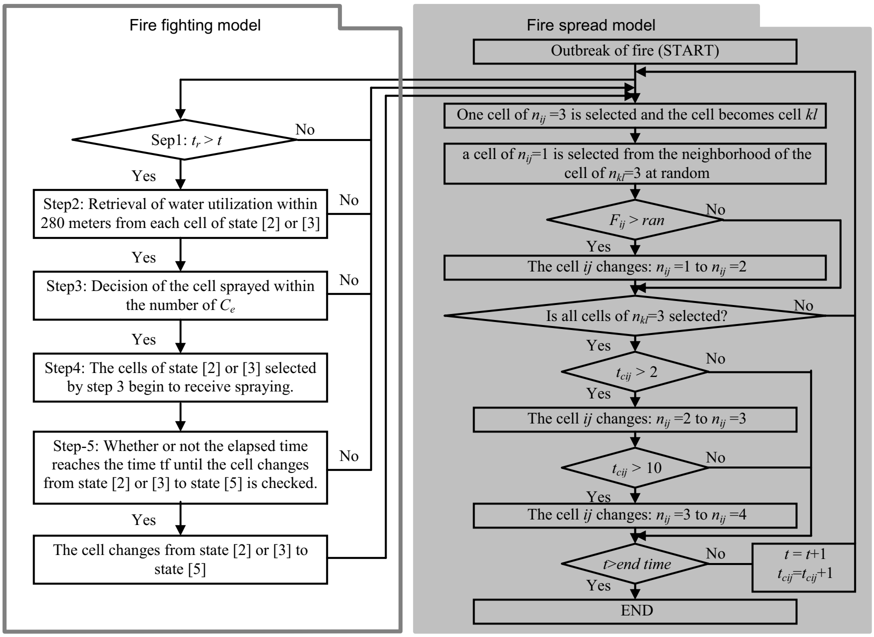
3.6 Process of the fire spreading model

Figure 4 shows the flow of a fire spreading model proposed here in addition to a fire fighting model described later.

- Step 1: First of all, one cell *kl* of n_{kl} =3 is selected.
- Step 2: One cell *ij* of n_{ij} =1 is selected at random from the neighbourhood defined by a wind direction and a wind velocity externally given when simulation starting.
- Step 3: Fire spreading judgment index F_{ij} of a selected cell *ij* is calculated. If F_{ij} is satisfied with the following requirements, the sate of cell *ij* becomes n_{ij} =2 (Catching fire).

$$\textit{If}\ \ n_{ij}(t) = 1\ \ \textit{and}\ \ F_{ij} > \textit{ran}\ \ \textit{then}\ \ n_{ij}(t+1) = 2 \quad (6)$$

where n_{ij} (t) is the state of cell *ij* at simulation time t and *ran* is the random number which takes from 0 to 1.



t : The simulation time [min], t_{cij} : The elapsed time of after outbreak of fire [min], *end time* : The end time of simulation [min], F_{ij}: The fire spread judgment index of cell *ij*, *ran*: The random number given within the range from the maximum to the minimum of *Fij* value, t_r: The time until the nearest fire brigade rushes to the fire site and then begins to spray water [min], C_e: The number of cells of which the fire can be extinguished with a water utilization [cell]

Figure 4. The flow of the model

Note that t_{cij} is an elapsed time after outbreak of fire in cell *ij*. The time count of t_{cij} starts when the state of cell *ij* changes from state [1] to [2], that is satisfied with equation (6). Step 2 and 3 are iterated for all cells *ij* within the neighbourhood of cell *kl*. Further, the process regarding cell *kl* is executed for all cells of state [3].

– Step 4: The state of cell *ij* after catching fire changes according to the elapsed time after outbreak of fire t_{cij}. In other words, if the value of t_{cij} takes more than t_1=2[min], cell *ij* changes from state [2] to [3]. Namely, n_{ij} =3. Moreover, if the value of t_{cij} takes more than t_2=10[min], cell *ij* changes from state [3] to [4]. Namely, n_{ij} =4.
– Step 5: The simulation ends when the time *t* [min] becomes the end time externally given in the model. The time *t* is counted from the simulation start to the end. When the time *t* does not reach *end time*, 1 is added to t_{cij} and *t*, and then the process returns to Step 1.

4. FIRE SPREADING SIMULATION IN JAPANESE HISTORICAL BUILT-UP AREA

4.1 Subject area

The subject area is located at Futagawa district, the east part of Toyohashi city, Aichi Prefecture, that is a typical historical built-up area with Futagawa-Honjin museum. The size of the area is 400 meters at the north-south and west-east direction (see *Figure 5*). In this area, there are old stores called Nishikomaya, Higasikomaya, and Kurebayashi-Syohyu. And there is a road that was called the Tohkai highway in the Edo era. But, old wooden houses are densely built up and there are many roads of which the width is not so wide with 4-5 meters in this district.

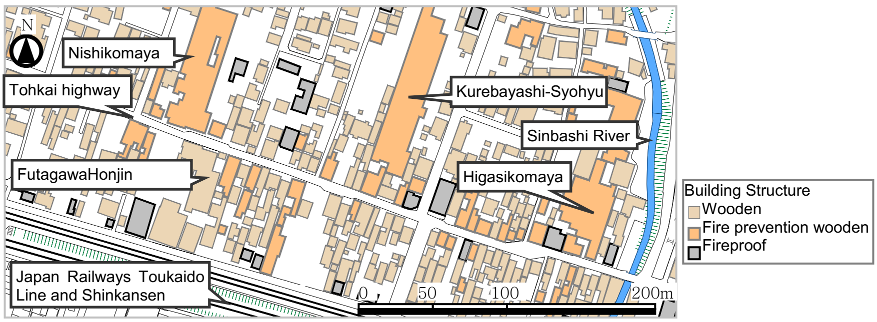


Figure 5. Subject area

4.2 Data arrangement

The date about building attributes was first arranged by building in the subject area. In order to carry out the simulation using a CA model developed here, it is necessary to identify the data by building with cell. The method is as follows. First, the layer of grid cell (3 by 3 meters) is overlaied on the city map representing buildings, roads and land use. Second, The attributes of building in a cell were defined on the supposition that the type of building occupying the largest area among all types in the cell is the type of the cell used in the model. But, the building data is not set to the cell where the building does not exist. Weather condition is given at the start time of a simulation.

4.3 Fire spreading simulation

Figure 6. Two simulation results by the model

The model developed here was applied to the subject area. The probabilistic calculation process is built into this model. Therefore, even if simulation is executed two times at most by a setting of parameters, weather condition and the point of fire origin, one result is different from another. Then, a simulation was carried out one hundred times supposing the point of fire origin as shown in *Figure 6*, α =1 β=1, wind velocity of 3 [m/sec] and the west wind. The simulation end time was set at 180 minutes. The calculation time of the simulation for one hundred is about 26 minutes by the personal computer of CPU: Pentium-M 1.4GHz and Memory: 1GB. The number of cells with each state in each simulation time is outputted. Two simulation results are visually shown in *Figure 6*.

This figure indicates that fire spreading has the tendency to extend to the leeward side of the origin of a fire. Each range of spreading after 180 minutes was about 200 meters on the leeward side. But there is the little difference in detailed process of spreading. Moreover, two examples are different in the number of cells with each state.

5. ANALYSIS OF SIMULATION RESULTS

5.1 The errors analysis

As mentioned above, simulations by the same condition are even different in the results of spreading. Therefore, in the scene of collaboration in community-based planning for disaster mitigation, it may be difficult to examine the effect of improvement in local environment by using one simulation result of this model. However, as you know, from the characteristics of the model with a probabilistic calculation process, the number of cells with each state is converged to a certain value when infinitely iterating simulation by the same condition. Accordingly, it must be more suitable to use a probabilistic expression as a way of offering the information to the participant. Then, first of all, all of the numbers of extinction cells acquired by one hundred simulation samples are prepared. Second, the mean and the standard deviation in each of 20 cases of 5, 10, … , 100 samples were calculated.

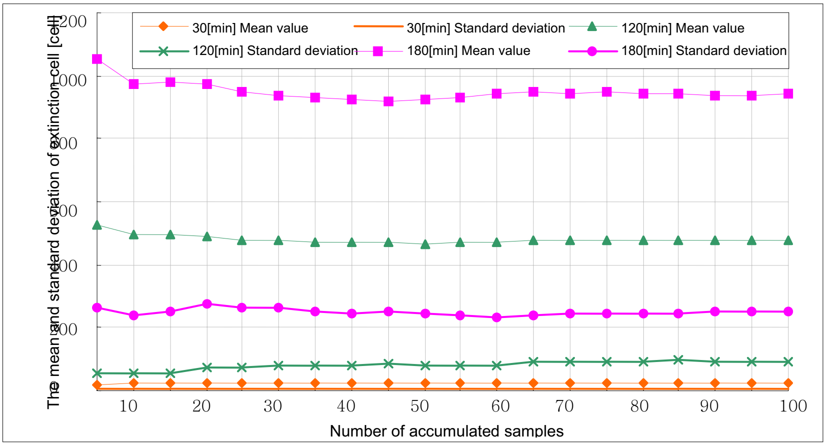


Figure 7. The mean and the standard deviation by simulation results

The following can be confirmed from *Figure 7*. There is increase or decrease in the mean and the standard deviation at each simulation time according to the increase of samples until 30 samples. But, in 50 or more samples, the amplitude of the mean and standard deviation becomes smaller than less than 50 samples. As a result, it is found that the number of extinction cells becomes stable in 50 samples at most. Therefore, hereafter in the analysis of simulation result, 50 samples are used.

5.2 Visualizing probability of cells changing

Figure 8 shows the probability that cells change to extinction at the simulation time of 180 minutes. This result is based on the 50 simulations at α =1, β=1 and wind velocity of 0 [m/sec]. From this figure, it is confirmed that the cells that burn out with the probability of more than 60 percent exists linearly in the southeast direction from the point of fire origin.

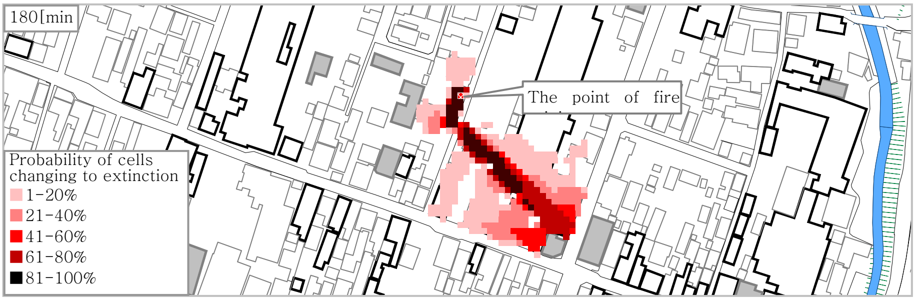


Figure 8. Probability of cells changing to extinction

In the scene of collaboration in the planning, for example using such a visual presentation of probable result must be useful for participants to discuss some concrete improvement projects such as widening of road, rebuilding from wooden to fireproof.

5.3 Comparison between actual fire records and simulation results

Figure 9 shows the relation between the number of extinction cells and the spreading time based on the following three kinds of actual records adding to simulation results; First type of actual record is one at Nagata district of Kobe City in Hanshin-Awaji earthquake (Matsubara and Suzuki, 1996). The district is an old wooden built-up area comparatively similar to the subject area. The gray and thick line is based on regression analysis using those records. Second type expressed with a triangle mark, is calculated by applying the data of wind velocity and fire spreading speed of Hanshin-Awaji earthquake to an oval model of Hamada (Jirou and Kobayashi, 1997). Third type is calculated from the record of fire in the subject area, which had occurred in April, 2003. The blue line is based on the means of 50 samples in each simulation time at α=1 β=1. On the other hand, the red line is based on the means at α=0.7 β=1.

From *figure 9*, it is found that the number of extinction cells of the first type of record is about 70 percent of the estimated number of those cells in simulation at α=1 β=1. This reduction can be considered as the slowdown of the spreading because of the restriction of air supply by building collapse. The simulation result at α =0.7 β =1 can approximately reproduce the situation of Hanshin-Awaji earthquake. Therefore, we can see that α plays the role of expressing the uncertain situation like the above. The number of extinction cells based on the records in the subject area (third type) is less than that of others because of a fire fighting.

As there is the difference in the circumstance of district, weather condition, situation of the earthquake damage, etc., the appearance of spreading actually has a wide range. According to the record of past conflagration, the fire spreading speed is with the wide range of 100 to 1,300[m/h] but depending on the velocity of the wind (Jirou and Kobayashi, 1997). In other words, the spreading scale considerably depends on the circumstances. Therefore, it seems to be able to explain the error of each simulation result and the variety of range of fire spreading to some degree caused by stochastic calculation process used in the model.

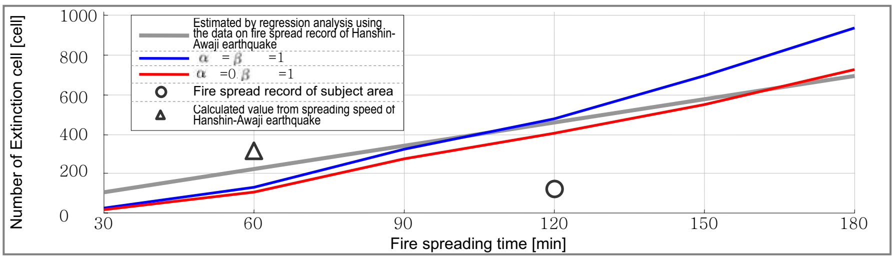


Figure 9. Comparison between the actual fire record and simulation results

6. FIRE FIGHTING SUBMODEL CONCEPT AND STRUCTURE

In chapters 4 and 5, we constructed a fire spreading model, analyzed simulation results of the non-interference model and carried out calibration. But, in the collaboration, a discussion must be necessarily done that how the fire spreading changes by fire fighting activity besides non-interference spreading. Therefore, to make the model more useful as a support tool, the incorporation of a fire fighting activity into our model proposed in previous chapter is significant.

In this research, the following fire fighting process was modelled; a fire brigade rushes to the site, the brigade sprays water on burning buildings, and the fire of the building is extinguished.

To support the collaboration, offering the information on the spreading interference by the fire fighting activity seems to be important. Then, state [5] (=cell of which the fire is extinguished not naturally but by fire fighting activity) is defined additionally. The following describes the concept and the parameter of a fire fighting model proposed here.

6.1 The time until the fire brigade rushes to the fire site

In this sub-model, the time t_r [min] is assumed to be the time until the nearest fire brigade rushes to the fire site and then begins to spray water. The value of this parameter can be optionally set and externally given in the model. A non-interference fire can be then reproduced by setting t_r at the time longer than the simulation time t. Introducing t_r enables the model to provide participants with whether or not the fire brigade can put down the spreading when a certain time t_r.

6.2 The number of cells of which the fire can be extinguished with water utilization for fire fighting

The number of cells of which the fire can be extinguished with water utilization, C_e, is given by equation (7). This is based on the assumption that one nozzle continues spraying water to a cell of state [2] and [3] until extinction of the cell.

$$C_e = \left\{ V_w /(C_w \cdot t_f) \right\} \cdot R_w \tag{7}$$

where, V_w is the capacity of a water utilization for fire fighting [m^3], C_w is the water volume that a nozzle consumes every a minute [m^3/min], t_f is the time until the cell sprayed changes from state [2] or [3] to state [5] [min] and R_w is the range of a nozzle spraying water [cell]. Referring to the previous research (Sekizawa and Endou, 2003) and hearing the fire department of local government in Japan, the values of C_w, t_f and R_w are set at 0.5, 20 and 3 respectively. The value of V_w can be optionally set and externally given within the range of 40 to 100 [m^3] which is based on the Japanese standard of establishing the water utilization. Therefore, the number of cells of which the fire can be extinguished by fire fighting depends on only the capacity and number of water utilizations located at the subject area in the model.

6.3 Process of the fire fighting sub-model

The flow of the fire fighting sub-model is constructed based on the above mentioned and the previous research (Sekizawa and Endou, 2003). The flow of the model is shown in *Figure 4*.

– Step 1: Whether or not the time t in the simulation reaches the time t_r until the fire brigade rushes to the fire site is checked.
  The following steps are iterated for all cells.

– Step 2: If the time t reaches the time t_r, water utilizations within 280 meters from each cell of state [2] or [3] are retrieved. If not so, the following steps are skipped.
– Step 3: The cell sprayed is decided within the number of C_e, subject to the priority of spraying to the cell decreasing according to the distance from the cells to a water utilization retrieved in step 2.
– Step 4: The cells of state [2] or [3] selected by step 3 begin to receive spraying. The elapsed time after spraying by the cell starts to be counted at the same time.
– Step 5: Whether or not the elapsed time reaches the time t_f until the cell changes from state [2] or [3] to state [5] is checked, and then the state of the cell changes when reaching the time t_f.

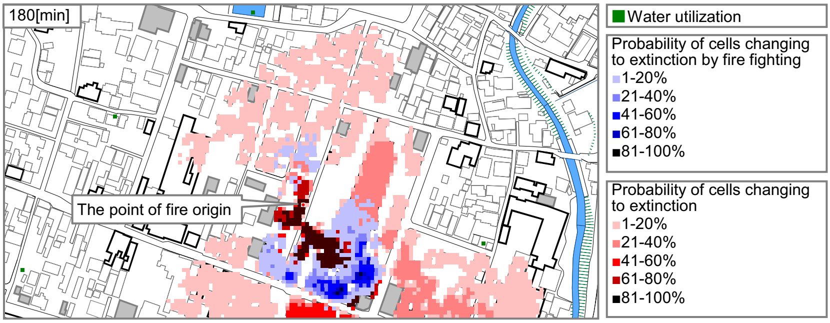


Figure 10. Probability of changing to natural extinction and extinction by fire fighting

7. SIMULATION RESULTS OF THE MODEL INCORPORATED FIRE FIGHTING ACTIVITY

Here, to analyze the behaviour of the fire fighting sub-model, two types of simulation with fire fighting activity (t_r =5 and t_r =60) were carried out 50 times respectively. The value of V_w of 4 water utilizations is set at 40 [m^3]. The location of these utilizations is shown in *Figure 10*. Other parameters and conditions externally given is the same as ones used in chapter 4.

There was no cell that changes from state [2] or [3] to state [4] in all of 50 simulations at t_r =5. As a result, we can see that the fire brigade rushing at 5 minutes can absolutely extinguish the fire in all cells burning. This result seems to be reasonable compared to actual fire fighting generally stated.

Figure 10 visually shows two types of the probability of cells in 50 simulation results at t_r =60. First type is one that cells change to state [4] and second is that cells change to state [5]. The followings can be seen from this figure. Cells changing to state [4] at the probability of 80 percent or more are distributed around the point of fire origin. On the other hand, there are cells changing to state [5] at the probability of 80 percent or more in the outside. From this, it is found that fire spreading is obstructed to some degree by the fire fighting.

8. CONCLUSION

In this paper we have attempted to develop a fire spreading simulation model using CA in order to provide the participants with the useful information in collaboration for the community-based planning for disaster mitigation. The conclusion is as follows.

Using CA enables the model to visually represent the fire spreading in detail. The proposed model can deal with the process of fire spreading in a building that traditional models can not represent. The actual fire spreading varies dramatically according to the circumstance of district, weather condition, situation of the earthquake damage, etc. The fire spreading judgement of the model can be done considering those conditions by using operational parameters. Although the full investigation of parameters by sensitivity analysis still remains, it was shown that simulation results can reproduce the actual fire spreading records approximately.

Though the model is based on a stochastic calculation process to reproduce uncertain fire spreading, we are sure that the use of the probable expression makes it useful for examining the effect of improvement in collaboration. Visual presentation of simulation results shown in this paper may help participants discuss about the improvement projects and enlighten them on the disaster risk and the needs of improvement. However, computational time takes over 10 minutes in order to get a stable simulation result in the present model. Shortening the calculation time is significant in order to enhance the usefulness of the model for really aiding the collaboration. Moreover, in our future work, we are going to develop a support tool of community-based planning for disaster mitigation, devising a visual and effective display of fire spreading process to participants, and incorporating the revised model into GIS.

REFERENCE

Batty M and Y Xie, 1994, "Form cells to cities", *Environment and Planning B, Planning and Design*, 21(5), p. 531-548.

Jirou K, K Kobayashi, 1997, "Large area fire", in: Fire Institute of Japan (eds.) *Fire Handbook third edition*, Kyoritsu Publication Co., Ltd., Tokyo, p. 508-573 (in Japanese).

Matubara M and K Suzuki, 1996, "Actual conditions of fire", in: Fire Institute of Japan (eds.) *Inspection report concerning a fire in southern part of Hyogo Prefecture earthquake in 1995*, Tokyo, p. 32-39 (in Japanese).

Murosaki Y, K Ohnishi and K Matsumoto, 1984, "Research on spreading factor of an urban expansion fire", *Journal of The City Planning Institute of Japan*, p. 373-378 (in Japanese).

Resnick M, 1995, *Turtles, Termites, and Traffic Jams -Explorations in Massively Parallel Microworlds-*, A Bradford Book, MIT Press

Sekizawa A, S Endou, 2003, "information system of supporting fire fighting based on real-time fire spreading forecast", in: Zama (eds.) *Research report concerning development of fire fighting activity support system*, National Research Institute of Fire and Disaster, Tokyo, p. 133-198 (in Japanese).

Sugawara S, T Naruse, 1997, "Building fire", in: Fire Institute of Japan (eds.) *Fire Handbook third edition*, Kyoritsu Publication Co., Ltd., Tokyo, p. 368-447 (in Japanese).

Wakamatsu T, 1978, "The risk of fire", in: Editorial committee of architectural outline (eds.) *Architecture outline No. 21 theory of architectural fire prevention*, Syokokusya, Tokyo, p. 279-298 (in Japanese).

White R and G Engelen, 1993, "Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land-use patterns", *Environment and Planning A*, 25(8), p. 1175-1193.

White R and G Engelen, 1997, "Cellular automata as the basis of integrated dynamic regional modeling", *Environment and Planning B*, Planning Design 24(2), p. 235-246.

Xie L and A G Yeh, 2000, "Modelling sustainable urban development by the integration of constrained cellular automata and GIS", *International Journal of Geographical Information Science*, 14 (2), p. 131-152.

Xie M, N Sakamoto, K Fujita and H Ogura, 2001, "Application of cellular automata to the city fire simulation" *Memoirs of Fukui National College of Technology*, 35, p. 19-24 (in Japanese).

Yamada A, A Takizawa, H Kawamura and M Tani, 1999, "The city fire simulation using the cell automaton", *Journal of Architecture Institute of Japan Tokai Chapter*, p. 657-660 (in Japanese).

Yasuno K, Y Nanba, 1999, "City fire prevention design", in: Horiuchi (eds.) *New edition architectural fire prevention*, Asakurasyoten, Tokyo, p. 178-235 (in Japanese).

Design Research and Design Support Systems

Augmented Reality Meeting Table: a Novel Multi-User Interface for Architectural Design

A. Penn, C. Mottram, A. Fatah gen. Schieck,
M. Wittkämper[1], M. Störring[2], O. Romell[3], A. Strothmann[4], and F. Aish[5]
The Bartlett Graduate School, University College London, UK
[1]Fraunhofer FIT, [2]Aalborg University, [3]SaabAvionics, [4]Linie4 Architekten, [5]Foster and Partners

Keywords: Design Collaboration, Tangible Interface, Gesture, Agent Simulation, Augmented Reality.

Abstract: Immersive virtual environments have received widespread attention as providing possible replacements for the media and systems that designers traditionally use, as well as, more generally, in providing support for collaborative work. Relatively little attention has been given to date however to the problem of how to merge immersive virtual environments into real world work settings, and so to add to the media at the disposal of the designer and the design team, rather than to replace it. In this paper we report on a research project in which optical see-through augmented reality displays have been developed together with prototype decision support software for architectural and urban design. We suggest that a critical characteristic of multi user augmented reality is its ability to generate visualisations from a first person perspective in which the scale of rendition of the design model follows many of the conventions that designers are used to. Different scales of model appear to allow designers to focus on different aspects of the design under consideration. Augmenting the scene with simulations of pedestrian movement appears to assist both in scale recognition, and in moving from a first person to a third person understanding of the design. This research project is funded by the European Commission IST program (IST-2000-28559).

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 213-231.

1. INTRODUCTION

1.1 The co-evolution of design theory and computing

The way in which designers design has been the subject of considerable investigation and debate over the last 50 years. Much of this debate has been closely associated with the development of design computing which by its nature has required us to understand and define exactly what we mean by design. In this paper we report on the ARTHUR project, the latest in a long line of research projects aimed at developing computing for design, which perhaps helps throw further light on the nature of design itself.

In the early years of architectural research, the contention was that formalised procedures or 'design methods' would bring rationality to the design process. In this view the methods of operational research, developed with such success to provide logistic support for the military during the war, could be brought to bear on design which until then had been largely seen as an intuitive (and therefore unscientific) activity (Rittel and Webber, 1973). Herbert Simon championed the cause during the 1950s (Simon, 1969) and Christopher Alexander during the 1960s, and by the time of the Oxford Conference on Design Methods in 1963 this view was well established. Although it was noted that design methods might inhibit imagination and creativity, it was felt that by integrating brainstorming into the process even this problem could be overcome (Jones, 1963). However, only a decade later this view had come to be comprehensively challenged from a number of directions.

Alan Colquhoun noted the centrality of typology and precedent in design thinking (Colquhoun, 1967). As often as not design proceeded by reference to past 'good solutions' to a problem, or through an analysis of the range of solution types that might be applicable to a specific problem. Meanwhile Bill Hillier developed a critique of design method based on Popper's analysis of scientific method (Hillier, Musgrove, et al., 1972; Hillier and Leaman, 1974). He proposed that knowledge of the effects of local design moves on global outcomes was the specifically architectural knowledge that was needed in design. By the 1980s even those who had been the originators of the Design Methods movement had completely reversed their positions: *"in the seventies I reacted against design methods. I dislike the machine language, the behaviourism, the continual attempt to fix the whole of life into a logical framework."* (Jones, 1977); *"Jacobson: In what areas should future work centre in design methodology? Alexander: I think I should be consistent here. I would say forget it, forget the whole thing."* (Alexander, 1971). Analysis of design processes evolved radically during this period through a recognition of the inadequacy of the linear process characterised in the early days of Design Methods (eg. Archer, 1979) and the realisation that design processes were both fuzzy and iterative in nature (Checkland, 1981; Checkland and Scholes, 1990). In this analysis it was clear that the path followed by the design process was not from a well defined problem to an optimal solution, but as often as not the reverse: through an exploration of possible solutions to a better definition of the brief or design problem at hand (Rittel and Webber, 1973). Herbert Simon noted that architectural design was characterised, like chess, by problems that are "well structured in the small but ill structured in the large" (Simon, 1973). That is, whilst in detail it is often possible to give a well constrained definition of the design problem at hand, it is often impossible to do so for a whole design.

Attention had been drawn to the important distinction between design as a procedure and as a cognitive process by Hillier at UCL (Hillier, Musgrove, et al., 1972), but an understanding of the latter was developed through the analysis of design activity by Donald Schoen and his colleagues at MIT. Their key contribution was to formalise from a theoretical standpoint the notion of 'knowledge in action'. This is that class of knowledge that is captured not in words, but in actions or 'know-how'. For Schoen, design, and especially the interaction between the designer and the plastic media with which they work during design – sketches and maquettes – is largely learned and practiced through action and reflection. The sketch is a concretisation of an abstract idea of a building, which can then be reflected on, and it is this dialogue which characterises an important dimension of the design process (Schoen, 1983).

Schoen's analysis goes further, however, to discuss the education of the 'reflective practitioner'. He notes that knowledge in action must be developed through action itself, and he proposes that the design studio and the design jury are typical means to accomplish this. It is by presenting a design to colleagues and tutors that the implicit understanding of what design entails is built up (Schoen, 1987). This sets out design education and practice (along with most other forms of professional practice) as a primarily social and collaborative activity. The work of Tom Allen on innovation in engineering design is relevant in this context (Allen, 1984). Allen's studies of large defence engineering projects found that despite the hopes of the brainstorming enthusiasts, innovative ideas in design projects often arose through the import of ideas from outside the group. Innovation, it seems, is substantially a product of collaboration, and the most productive collaborations involve individuals who bring radically different mindsets to the problem (Penn, Desyllas, et al., 1999).

Others have noted that design practice is characterised by a dichotomy around the design-meeting table, between the 'form generators' and the 'analyst critics' (Hillier, 1998). In this situation design proposals (generated by the ‘designer’) are reviewed by a range of experts: engineers, cost consultants, contractors and client or user representatives. Here the key capacity of the designer is to listen to the critique, to internalise its consequences and to use this knowledge to help refine the problem definition and to narrow the range of solution types that need to be developed. The process has been described as akin to solving Rubik’s cube. Since each domain that must be resolved is independent, but all interact through the proposed built form, as a move is made to resolve one issue it creates problems for another domain. It is for this reason that design requires both a broad range of expertise around the table and an iterative process (Penn, Treleavan, et al., 1996). However, it is also clear that collaboration in design is itself a reflective activity in Schoen’s terms. The process of listening to domain experts and internalising the dynamics of different form/function relationships is itself learned through action and largely an implicit skill. Habraken and Gross investigated these collaborative skills through game play (Habraken and Gross, 1988), and found that experience of collaboration in games improved the effectiveness of team based design work. In a more recent analysis of historical design projects, Chengzhi Peng reviewed the importance of different representational media in design collaboration, including the use of overlays, scores and shared modelling spaces (Peng, 1994).

By the 1990s the true complexity of architectural design processes was well established and the research field had divided to address the range of specific sub themes that seemed to form part of the whole. The field was characterised, it seemed, by a plethora of dichotomies: should research focus on the process or the product? Was design bottom up or top down? Intuitive or rational? Art or science? Individually or socially constructed? Emergent or rule governed? Which comes first, the problem or the solution?

Answers to these binary oppositions were generally ‘both’, perhaps presaged most eloquently by Robert Venturi. Architectural design, it seemed, was consistent in its need to span and work from both ends of any axis, and creativity thrived on the contradictions involved in this (Venturi, 1966).

Among the key components of the research field was the investigation of part-whole relationships both in analytic mode (eg. Space syntax studies – Hillier and Hanson, 1984; Hillier, 1996) and in generative mode (eg. Shape grammars – Stiney, March, 1981). Both of these approaches (and others akin to them) lent themselves to computation, and in fact required computation for their execution. At the same time, in architectural practice, the development of computing to assist in the organisation and management of production information had taken over the way that this aspect of the process was organised. The effect of Computer Aided Design (CAD) in the office was to bring into clear focus the distinction between drawing (or sketching) as a cognitive activity supporting design, and drawing as method of communication of production information (Lawson, 1990). The latter is now almost entirely computer based in the architectural profession, but the former remains firmly rooted in the media of the plastic arts (including of course, digital media).

Developments in computing for architectural design since the 90s have tended to follow one of a number of main lines: three dimensional modelling for visualisation; parametric and scripted modelling; analysis and simulation of functional performance for concurrent engineering; case based or precedent databases; and, evolutionary or optimised form generation. Of these, one of the great hopes of the 90s was that improvements in visualisation, and Virtual Reality (VR) in particular would help the 'intuitive' aspect of the creative process. By allowing designers to inhabit and visualise their emerging design, now possible at 1:1 scale through fully immersive VR, it was hoped that computing could be brought to the service of design rather than just production information.

These hopes for fully immersive VR in architectural and engineering design have not yet been borne out in reality for a number of reasons, including cost and the unfeasibility of working in a completely immersive environment. Full immersion has therefore remained of greater academic than practical interest. More recently, the focus of research has shifted again, towards lower cost systems and the merging or augmentation of the real work setting with virtual media.

Amongst recently developed technologies, Augmented Reality (AR) offers a number of attractive properties which may help support design processes as we are coming to understand them. AR as a field has developed rapidly over the past ten years (Azuma, 1997). AR systems and technologies originated in the head-up display technologies developed for defence (Wanstall, 1989) and were introduced into a broader engineering context in 1992, when they were developed to support industrial manufacturing processes by Boeing (Caudell and Mizell, 1992). Since 1996 AR and its applications have been investigated for various surgical applications such as needle biopsy and minimally invasive treatment, where, by combining Computer Tomography, Magnetic Resonance and Ultrasound Imaging, with the real scene, doctors are provided with information superimposed on the patient (Weghorst, 1997; Fuchs, Livingston, et al., 1998).

In the architectural and construction field it has been suggested that an AR system might give users "X-ray vision" inside a building, allowing them to see where the pipes, electric ducting, and structural supports are inside walls and above ceilings (Feiner, Webster, et al., 1995; Webster, Feiner, et al., 1996) or generate 3D models of pipelines in a factory and register them with the user's view (Navab, Bascle, et al., 1999). In all of these applications the main thrust of the research is to use AR techniques to augment physical objects in the environment with additional information or views of their internal, hidden structure or anatomy. The feasibility of optical see-through (as opposed to video see-through) AR for these applications is severely limited by current wide area head tracking technologies needed if the virtual and models are to be properly registered with respect to the real environment (Azuma, Baillot, et al., 2001). To date these applications have remained laboratory based, however the development of computer vision systems for head tracking using head mounted cameras have been shown to allow viable frame rates outside the laboratory (Argyros, Trahanias, et al., 1998).

A number of current and recent research projects are directly relevant to the augmented reality approach presented in this paper:

BUILD-IT (Rauterberg, Fjeld, et al., 1997) a multi-user planning tool using a 2D projection on a table top. Additionally, a video camera is used to track manipulations of a small, specialized brick that can be used as a "universal interaction handler". A second, vertical projection screen behind the table provides a 3D view of the virtual scene.

MagicMeeting (Regenbrecht, Wagner, et al., 2002) supports product review meetings by augmenting a real meeting location. Virtual 3D models are loaded into the environment from desktop applications or Personal Digital Assistants (PDAs). The MagicMeeting system explores several interaction techniques, including the MagicBook metaphor (Billinghurst, Campbell, et al., 2000), annotations and a clipping plane tool. The 3D models are linked to real placeholder objects to create a tangible interface (Ishii and Ulmer, 1997).

MARE (Grasset and Gascuel, 2002) is a multi-user augmented reality system designed to be used in any application domain. The table space is divided into two parts: a personal area for the private real objects (pen, PDA, cup) and virtual ones (private menus); and a shared interactive space.

MIXDesign (Dias, Santos, et al., 2002) provides a Mixed-Reality system oriented towards tasks in Architectural Design. The system explores tangible interfaces using AR Toolkit patterns on a paddle and gestures.

TILES (Poupyrev, Tan, et al., 2000) is a collaborative tangible AR that allows two-handed interactions. The interaction techniques are aimed at the application for rapid prototyping and evaluation of aircraft instrument panels. The spatial interaction is in 3D allowing the user to manipulate virtual data on any working surface.

URP (Underkoffler and Ishii, 1999) is an application for urban planning and design. The infrastructure allows physical architectural models placed on a projected table surface to cast shadows accurate for any time of day; to throw reflections off glass facade surfaces; to visualise a simple 2D

Computational Fluid Dynamics (CFD) analysis of wind flow. This is accomplished using *I/O Bulb* techniques to attach projected forms to physical architectural models.

In these projects four main concerns are notable:

- the use of tangible interfaces to engage both hand and eye in response to the cognitive importance of action in design;
- the merging of environments with visualisations of functional performance;
- a catholic approach to mixing of different conventional, digital, VR and AR media;
- the support of collaboration through sharing of a single representation on a 'table-top' or large screen visualisation.

This is the context within which the ARTHUR project described in this paper is set. By using optical see through Augmented Reality (AR) displays the ARTHUR system aims to put virtual 3D models on the designer's meeting table alongside conventional media. In the next section we describe the ARTHUR system in detail. We then review early findings from the system in use, including qualitative feedback from our first user tests, before finally drawing conclusions about the light these new media throw on our understanding of the nature of the design process.

2. ARTHUR

Imagine the scene. A design meeting consists of a group of people seated around the meeting table. On the table are various drawings, a polystyrene sketch model, photographs, perhaps some material samples, and inevitably a roll of layout paper and some felt tip pens. As people discuss the scheme they use the drawings and model to illustrate their point. From time to time someone will sketch something – a possible solution to a problem. Often they will do this alone as the conversation proceeds before tabling it to help others understand their proposition. Sometimes the sketching is done in front of the group where the act of drawing is itself an aid to communication. The poly model becomes the focus of attention, a block is moved and another cut to change its shape. One of the consultants around the table raises a point and illustrates it with reference to a spreadsheet calculation on her laptop. However, there is a key concern. Something that can only be resolved through a simulation based on the new plan. Tasks are allotted and decisions postponed until this can be done.

Now imagine that everyone has lightweight clear glasses; these have a pencil camera on each side and are connected to a computer. The glasses are augmented reality displays – the kind that the military use to give head-up display information to fighter pilots – the cameras track the scene and allow a 3D computer model generated from the CAD system to be kept still on the table as the user moves around it. The cameras also recognise placeholders on the table and the user's hand gestures. These are used to interact with the model. Moving a placeholder or making a gesture changes the model. As the model is changed there is an instantaneous link to the consultant's spreadsheet and its results are recalculated. The results of analysis and simulation can be rendered into the model on the table, and so long as calculation time allows, these can be updated as the model is manipulated. This is the ARTHUR concept (Broll, Störring, et al., 2003). It is not a replacement for traditional media, but an attempt to bring the computer – as an interactive and creative medium in its own right – onto the meeting table.

2.1 System architecture

The ARTHUR system consists of four main components. An optical see-through AR Head-Mounted Display (HMD) prototype has been developed by Saab Avionics specifically for the project. This incorporates head mounted stereo cameras to provide Computer Vision (CV) based object tracking of 'placeholder objects' and pointing devices for use in interaction with the virtual media, as well as finger tracking and gesture recognition. The CV system is also used to track the users head position and so to keep the virtual model stable in the real environment from the user's point of view. An integrated software framework provides for visualisation, distributed communication for multiple users and CAD integration through an Application Programming Interface (API). A graphical language (GRAIL) has been developed to allow users to configure the relationships between inputs and outputs within the system, including through external applications. Each of these components is described in detail below.

2.2 AR display

The ARTHUR head-mounted display is a transparent high-resolution binocular 3D HMD. The image is generated on two independent full color 1280x1024 pixels (SXGA) Liquid Crystal on Silicon micro displays. The ARTHUR HMD has excellent image quality, high brightness and contrast. The image can be shown superimposed on the environment with up to 35 % see-through or fully immersed. The ARTHUR HMD is designed for a 46-degree diagonal 100% stereo overlap field of view. A 50% overlap can also be used, giving a 54-degree horizontal or 60-degree diagonal field of view. It features a patented optical design that combines a wide field of view with high transparency see-through and a patent pending head fitting system with easy adjustments, low weight and eyeglass compatibility for most users.

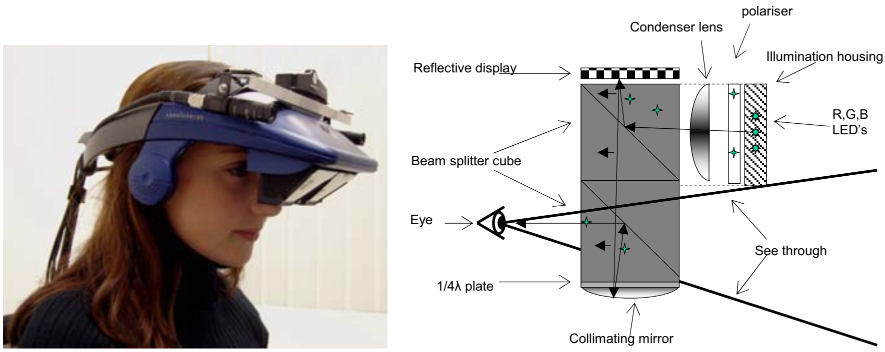


Figure 1. The optical see-through HMD with mounted cameras and tracker unit

2.3 Computer Vision

Computer Vision (CV) techniques take input from stereo Head-Mounted Cameras (HMCs) and a fixed camera, and are used to track the movements of real world items (placeholder objects and two hand held pointers) and to recognise hand gestures. The computer vision based approach allows users to interact without cables or sensors connected to the objects they manipulate to create an interface. The hope here is that interaction will be more natural and intuitive.

2.3.1 User Interface Devices: Placeholder and Pointer Tracking

The user interface devices are dedicated objects that are tracked by the computer vision system using colour and shape information. There are two types of device: placeholder objects (PHOs) and wand-like pointers.

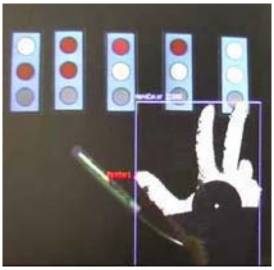
Placeholder objects are tracked in the table plane, in two translational and one rotational degrees-of-freedom giving 3DOF information. They are of a convenient size to be grasped and moved by users. More than ten placeholders may be used concurrently. The users may take a placeholder object, associate it with any virtual object and move this virtual object by moving the placeholder object, thereby creating a direct manipulation interface. Alternatively, any one of the 3DOF can be associated with any object attribute for any virtual object. Thus, for example, a PHO's orientation could be used to set a virtual object's height or colour.

The pointers are tracked in 5DOF – all except roll – by the head mounted cameras. This has the advantage that the pointer is always tracked when it is in the user's field of view. The pointers have three buttons for functionalities such as pick or select. Users may select and manipulate the shape of virtual objects with a pointer or use it to navigate in virtual menus.

2.3.2 Gesture Recognition

Two types of gestures can be used, static command gestures and 5DOF fingertip tracking. Both gesture types are tracked by the head-mounted cameras and the CV system.

A set of five static command gestures have been implemented. The number of fingers shown to a head-mounted camera is interpreted as a gesture. These command gestures can be associated with a change of mode in the user interface. For instance, they can be used to display or hide a pop-up menu, or to execute functions such as cut-copy-paste. The 3D position of the users index fingertip and the direction their finger is pointing is also tracked using the head mounted cameras. Moving the thumb generates a 'click' event. Fingertip tracking may be used to draw a line in space or to select items in pop-up menus, select objects and execute actions.



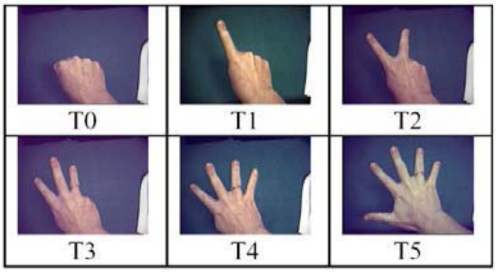


Figure 2. Left: 3 DOF PHOs and a 5 DOF pointer. Right: Static command gestures

2.3.3 Computer Vision Based Head Tracking

In an AR system, fast, accurate, and jitter-free tracking of the user's head pose (position and orientation) is crucial if the virtual objects in the display are to appear stable in the real environment. In ARTHUR the head pose was originally tracked using a commercial tracking system. In order to reduce jitter the head pose is now tracked using the HMCs and the two sources of information are fused. The head tracking via the HMCs is accomplished using the positions of the PHOs found using the fixed camera system.

2.4 AR Framework Components

The AR framework is a C++ API, which provides the basis for bringing all the individual ARTHUR components together. It consists of three main elements:

2.4.1 3D-Visualization

The 3D-visualisation component renders geometric objects in 3D using head mounted displays or other output devices. Rendering can be either stereoscopic (quad-buffered or dual screen) or monoscopic. Augmentation within ARTHUR is usually achieved using see-through augmentation, i.e. the virtual image is superimposed optically on the real environment. The 3D-visualisation component however, additionally supports video augmentation, where the virtual image is superimposed on a video image from a head-mounted camera. This may be used for screen-based presentations or larger projections to show people not wearing an HMD, what is currently visible to those participants. The scene contents displayed by the visualisation component are based on Virtual Reality Modelling Language (VRML) 97. In the overall system, one visualisation component exists for each individual user, allowing rendering to be performed locally. In order to achieve this, the scene graphs are replicated among the individual visualisation components and kept synchronised by distributing changes to geometry.

2.4.2 Distribution and Communication

The AR framework provides the distribution and communication mechanisms needed to connect the input devices such as head tracking, computer vision input (placeholder tracking, pointer and finger tracking, and gesture recognition) to other system components (e.g. the 3D stereo visualisation components). This allows the visualisation components to adapt their current viewing position and orientation to the tracking input. Additionally this distribution mechanism is also used to keep the scene graphs of multiple users synchronised (i.e. upon changes to one local scene graph, these changes are immediately distributed to all other replicated scene graphs). Synchronised clocks between all PCs involved are required in order to resolve ambiguities. The general communication mechanism used within the AR framework is Common Object Request Broker Architecture (CORBA). A universal sequencer mechanism provides an overall virtual time management for all components depending on virtual time.

2.5 Graphical AR Interface Language

In ARTHUR a high level graphical interface gives non-programmer users the ability to fully configure their interactions with the system. This makes use of a Graphical Reality Augmentation Interface Language (GRAIL), which provides an interface combining both 2D applications and 3D manipulation and visualisation tools in the 3D space.

The GRAIL application sits on top of the AR framework and acts as a ‘tool building tool’ to define the properties and characteristics of the overall interaction environment. In this sense it performs as a Graphical User Interface (GUI) builder, but for AR environments and interactions.

Users can create a range of tools by graphically defining the relationship between the 3DOF PHOs and 5DOF pointers and command gestures and the virtual objects (defined as VRML nodes) or system modes. The relationships within a GRAIL configuration can easily be changed and reassigned, providing the ARTHUR system with a flexible interface that supports a rich type of interaction with the system and the virtual objects. Essentially, we provide this flexibility in an attempt to define ARTHUR as a creative or plastic ‘medium’ in which users are free to define their own forms of interaction between virtual and real objects and their own gestures and actions.

GRAIL also provides the user with the ability to link the ARTHUR system to external applications, such as pedestrian and environmental analysis applications, using scripting commands as an alternative to using the full API. The use of scripting commands allows these links to be made by the user at run time without the need to recompile.

2.6 The user interface approach

Although desktop software applications now have a well-defined set of graphical user interface concepts to work with, no common user interface paradigms have yet been defined for immersive 3D or AR systems.

A starting point for the design of the ARTHUR user interface was to use the architectural desktop as a metaphor, to define the characteristics of the interaction environment. The critical question however, was the degree to which an augmented reality interface to design should follow the conventions that have been developed for interaction within the architectural desktop environment, bearing in mind that the conventional architectural desktop is designed for individual interaction and the efficient management of production information rather than to support the cognitive processes of design. In view of this our approach to interface design has been through the development of a series of usage ‘scenarios’ in collaboration with end user partners. These scenarios are then implemented using the GRAIL interface and amended in response to users’ feedback.

Figure 3. Collaboration within ARTHUR environment

2.7 Designing with a live movement simulation

Architecture differs from the plastic arts – painting and sculpture – in that in addition to its aesthetic, cultural and structural roles it is expected to perform functionally. That is, to provide shelter by modifying the environment and to support specific social functions by ordering spatial arrangement to generate or control patterns of contact between different groups of people. It is through the latter that architecture functions socially, and it is this that leads to a hospital being differently planned to an office or a prison. It has been proposed that the social function of both architecture and urban form derives from the way that the configuration of space affects the way that people are brought into contact as they move around. It is this ‘construction of the interface’ between different categories of people that characterises the social function of architecture at its most basic (Hillier, and Hanson, 1984; Penn, Desyllas, et al., 1999).

In the development of the ARTHUR system it would clearly be impossible to develop the full range of analytic methods used even on a relatively simple building project. However it was also clear that the ability to subject a design to a simulation of functional performance was a critical component of round table collaborative design. It is for this reason that we chose to investigate the use of a pedestrian simulation to bring the ARTHUR model on the designer’s meeting table to life, however, this is but one example of the kinds of analysis that could be incorporated into the system.

In order to simulate realistic movement patterns and rates we have developed agent simulations in which an agent’s forward facing field of view is sampled, and an immediate next step chosen from that field. In this we have relied on previous research that has found that visual field based movement of this sort provides a good correlate with observed aggregate movement behaviour in urban environments and buildings (Penn and Turner, 2002; Turner and Penn, 2002). In this case, however, in order to allow the environment's geometry to be varied in real time while the simulation is running, our algorithms are based on sampling the current environment rather than pre-computation of visual fields.

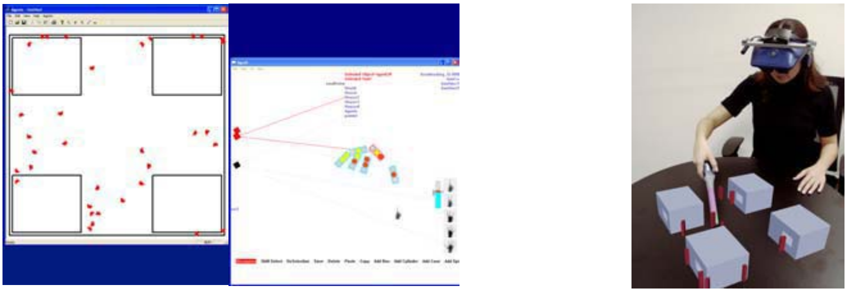


Figure 4. Left: Linking the agent software in GRAIL. Right: Adding agents using the pointer

3. SYSTEM EVALUATION AND DISCUSSION

3.1 Early results of user studies

All the main components of the ARTHUR system have now been developed and integrated, and early user evaluations of each component and of the whole system have begun. To date, three main types of user evaluation observations have been conducted. First, single user tests were conducted with subjects from our application partner organisations to evaluate usability for simple manipulations using simple geometric transforms (translation, rotation, scaling) with different interaction techniques using the pointer to navigate in virtual menus or to select and manipulate the shape of virtual objects. Here the protocol was use of the system followed by debriefing discussions. These observations must be considered only as preliminary and a part of an early prototyping development process since the subjects had been intimately involved in the specification of the system, however, they proved to be very useful in helping us to develop appropriate interaction techniques. The results of these tests were directly incorporated in the development of the system.

The second form of evaluation was through formal user tests and associated observations and questionnaires. An experiment was designed to assess the impact of the ARTHUR environment on the level and quality of collaboration between users by evaluating the way users perform a simple task. Two teams of collaborators were first given a short period of training and time to get used to the system, and then were asked to collaborate to design a simple structure. The interaction techniques we tested were restricted to adding 3D objects to the common space using two types of gestures; a two-finger gesture to create an object (a simple cuboid) and a one-finger gesture to locate it. In this way the objects could be aligned and stacked to form a structure. The form of collaboration was observed and annotated by a researcher, and notes were taken of the form of interaction between the subjects, of difficulties experienced with the interface and of their comments on the system while using it. Following the session, the participants were debriefed in both a structured discussion and using a questionnaire.

The third form of user evaluation took place as a part of a public demonstration of the ARTHUR system at the Fair for Information Technology and Telecommunications (CeBit 2004) in Hanover. A large number of people participated in either a single user or in a multi-user environment depending on their preference. These users were all first time users with no previous knowledge of the system. A brief description of the system and instructions were given by a demonstrator, and then they were invited to use the system. The interaction techniques consisted of simple object creation and manipulations using simple geometric transforms (translation, rotation, scaling). Placeholder objects and pointers were used to interact with virtual menus or to select, move and manipulate the shape of virtual objects. All participants were observed, and notes were made of the difficulties they experienced and assumptions they made about the user interactions within the system. Of these subjects, ten who showed particular interest in the system were asked to provide detailed feedback in a form of a questionnaire. The main observations we report here are therefore qualitative and derived from the discussions and observations described above.

System evaluation at such an early stage in development and integration was surrounded by technical glitches and inconsistencies, and it was clear that these hampered a user's ability to evaluate the system. Having said that, we observed that with simple instructions the users were able to use gestures, pointers and placeholders effectively. All subjects found the system to be both enjoyable and to offer a potential for collaborative design. However, subjects pointed out that the use of gesture was not natural "I am used to having something in my hand while designing: a pen or a mouse". Subjects who used the gestures effectively were able to do this only after a period of training. This is unsurprising given the command gestures comprise a language to be learned.

Generally, it appeared difficult to understand the virtual objects as integrated in the physical environment. This might be expected since physical laws e.g. gravity and inertia, are not applied. Subjects were not always aware of the new objects that were added by another collaborator, especially when the scene was cluttered, “It was confusing to see a box suddenly appear floating”.

It was difficult to understand distances, and subjects, especially those with CAD experience, pointed out that a common view of a situation might help collaboration “to have a view that the other user can see”. This was unexpected. It appears to be partly a result of the restrictions on a user’s movement in their environment imposed by the system, the weight of the HMD and its cables, and the need to keep the PHOs that sit on the table in the field of view in order to allow head tracking. Ideally, users should be able to move their heads freely and wander around the table and observe the design from different viewpoints. In addition, the appearance of the virtual objects as always superimposed on real objects, especially the user’s hands, may have led to a loss of the feeling of the virtual as existing within the real world. In conclusion, users found that the properties of the interaction space, viewed through a relatively narrow field of view provided by the headset, are currently closer to those of an immersive VR than an augmented real space “I was unaware of the real environment, immersed and engaged in the task until somebody (the collaborator) talked to me”. In this situation speech appeared to be crucial to subjects’ perception of the social activity and to the collaboration itself.

When users are faced with simulated pedestrian movement two key factors seem to emerge. The first is that the simulation lends a sense of scale to the model through size and speed of movement of the simulation agents. The second is that it seems possible to empathise with the simulation agents, and so to ‘inhabit’ the virtual model at its reduced scale. Games are played by moving buildings to affect movement patterns and these games actively involve both designers and simulation agents.

3.2 Conclusions: social organisation and play

Every new design medium brings to light insights into the design process. The advent of CAD made it clear that the cognitive dimension of design was separate to the communication of production information, and perhaps one of the most interesting aspects of the ARTHUR development lies in what it seems to tell us about the way that designers collaborate. There is clearly an asymmetry around the designers’ meeting table between the activities of form generation and those of critical analysis. As a system, the intention was that ARTHUR should support both aspects of design interaction. However, we have found that collaborative design is complex: form generation as an activity seems to be substantially an individual one whilst analytic criticism is a group effort. Users find it extremely hard to even think about how it might be possible to really collaborate in real time in the act of form generation itself.

At the same time, the new AR medium seems to hold out a prospect for quite different ways of working. Two different kinds of behaviour were observable in our user tests of collaboration. In the first, one member of the team would take charge of the process, and direct actions “you do this, and I’ll do that”. This form of organised teamwork is familiar in design teams, and is a normal feature of most architects’ offices. In the second, collaborators began to play games. Here for instance they would place blocks to channel the movement of the simulated pedestrians, in much the same way that children collaborate in damming streams.

We believe that creating architectural forms and working on a task collaboratively became a game that users enjoyed and consequently this increased their level of collaboration: “it was great to see the other user and talk, and shape the design space together”. Playing such games, particularly where they involve simulations of functional performance, may also promote learning. This is supported by Habraken. “The actual making of decisions about forms in space - had a strong and inevitable social dimension and as such was influenced by the way in which involved parties interacted” (Habraken and Gross, 1988). Our findings to date support this, and suggest that if technical and human factors barriers can be overcome, augmentation of the real world by digital interactive media may lead to new forms of genuinely collaborative form generation. One consequence of this is that design education may need to begin to learn some lessons from time based performance media about how to train designers to collaborate and play.

ACKNOWLEDGMENT

Extra contributions from I. Lindt, C. Yuan, W. Broll, T. B. Moeslund, Y.Liu, E. Granum, H. Hildmann, S. Preikschat, and H. Whitehead.

REFERENCES

Alexander, C., 1971, “The State of the Art in Design Methodology”, (replies to questions by M. Jcobson), DMG Newsletter, p. 3-7.

Allen, T., 1984, *Managing the Flow of Technology*, MIT Press.

Archer, L., 1979,” Whatever became of Design Methodology”, *Design Studies*,1 (1), p.17-18.

Argyros, A. A., Trahanias, P.E., and Orphanoudakis, S.C., 1998, Robust Regression for the Detection of Independent 3D Motion by a Binocular Observer**,** *Real-Time Imaging*, 4(2), p.125-141.

Azuma, R.T., 1997, “A Survey of Augmented Reality”, *Presence: Teleoperators and Virtual Environments*, 6(4), p. 355-385.

Azuma, R.T., Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, 2001, “Recent Advances in Augmented Reality”, *IEEE Computer Graphics and Applications*, 21 (6). P.34-47.

Billinghurst, M., S. Campbell, D. Hendrickson, W. Chinthammit, I. Poupyrev, K. Takahashi, and H. Kato, 2000, Magic Book: Exploring Transitions in Collaborative AR Interfaces, *SIGGRAPH 2000 Emerging Technologies Proposal*, ACM SIGGRAPH 2000.

Broll, W., M. Störring, and C. Mottram, 2003, “The Augmented Round Table - a new Interface to Urban Planning and Architectural Design”, *INTERACT 2003*, p.1103-1104.

Caudel, T.P., and D. Mizell, 1992, “Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes”, *Proceedings of the Hawaii International Conference on Systems Sciences, Maui, Hawaii*, IEEE Press, p. 659-669.

Checkland, P., 1981, *Systems Thinking, Systems Practice*, John Wiley and Sons, Chichester.

Checkland, P., and J. Scholes, 1990, *Soft Systems Methodology in Action*, John Wiley and Sons, Chichester.

Colquhoun, A., 1967, “Typology and Design Method”, *Arena*, 83, p.11-14.

Dias, J.M.S., P. Santos, and N. Diniz, 2002, “Tangible Interaction for Conceptual Architectural Design”, *ART 02*.

Feiner, S., A. Webster, T. Krueger, B. MacIntyre, and E. Keller, 1995, “Architectural Anatomy”, *Presence: Teleoperators and Virtual Environments*, 4(3), p. 318-325.

Fuchs, J.F.H, M. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, J. Crawford, P. Rademacher, S., Drake and A. Meyer, 1998, “Augmented Reality Visualization for Laparoscopic Surgery”, *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 98)*, Springer-Verlag, Heidelberg, Germany, p. 934-943.

Grasset, R., and J.D. Gascuel, 2002, “MARE: Multiuser Augmented Reality Environment on table setup”, *SIGGRAPH, Conference Abstracts and Applications, Computer Graphics Annual Conference Series 2002*, p. 213.

Habraken, N.J., and M.D. Gross, 1988, “Concept Design Games”, *Design Studies*, 9(3), p.181

Hillier B., J. Musgrove, and P. O'Sullivan, 1972, "Knowledge and Design", *Environmental Design Research and Practice*, University of California Press, Los Angeles.

Hillier, B., and A. Leaman, 1974, “How is design possible?”, *Journal of Architectural Research*, 3(1), p. 4-11.

Hillier, B., and J. Hanson, 1984, *The Social Logic of Space*, Cambridge University Press, Cambridge.

Hillier, B., 1996, *Space is the Machine*, Cambridge University Press, Cambridge.

Hillier, B., 1998, “A Note on the intuiting of Form: Three Issues in the Theory of Design”, *Environment and Planning B: Planning and Design*, 25, p. 37-40.

Ishii, H., B. Ullmer, 1997, “Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms”, *CHI’97*, Atlanta, Georgia, p.234-241.

Jones, C., 1963, “A Method of Systematic Design”, *Conference on Design Methods*, Pergamon, Oxford.

Jones, C., 1977, “How My Thoughts about Design have Changed During the Years”, *Design Methods and Theories*, 11(1), p.50-62.

Lawson, B., 1990, *How Designers Think: the design process demystified*, 2nd edition, Butterworth, London.

Navab, N., B. Bascle, M. Appel, and E. Cubillo, 1999, “Scene Augmentation via the Fusion of Industrial Drawings and Uncalibrated Images with a View to Markerless Calibration”, *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99),* IEEE CS Press, Los Alamitos, California, p. 125-133.

Peng, C., 1994, "Designers in Teamworking", *ITRI-94-6, University of Brighton.*

Penn, A., P. Treleavan, B. Hillier, L. Bull, R. Conroy, N. Dalton, L. Dekker, C. Mottram, and A. Turner, 1996, "Intelligent Architecture: Pangea, an Intelligent Workbench for Architectural Sketch Design.", *The Intelligent Systems Integration Programme ES96 Papers* , SGES Publications, Oxford, p. 127-139.

Penn, A., J. Desyllas, and L. Vaughan, 1999, "The Space of Innovation: Interaction and Communication in the Work Environment", *Environment and Planning B: Planning and Design*, 26 (2), Pion, London, p. 193-218.

Penn, A., A. Turner, 2002, "Space Syntax based Agent Simulation", *Pedestrian and Evacuation Dynamics*, Springer-Verlag, p. 99-114.

Poupyrev, I., D.S. Tan, M. Billinghurst, H. Kato, H. Regenbrecht, and N. Tetsutani, 2001, "Tiles: A Mixed Reality Authoring Interface", *INTERACT 2001 Conference on Human Computer Interaction,* Tokyo, Japan.

Rauterberg, M., M. Fjeld, H. Krueger, M. Bichsel, U. Leonhardt, and M. Meier, 1997, "BUILD-IT: a Video-based Interaction Technique of a Planning Tool for Construction and Design", *Proceedings of work With Display Units - WWDU'97*, Tokyo, p. 175-176.

Regenbrecht, H.T., M.T. Wagner, and G. Baratoff, 2002, "MagicMeeting: A Collaborative Tangible Augmented Reality System", *Virtual Reality*, 6, p.151-166.

Rittel, H., and M. Webber, 1973, "Dilemmas in a General Theory of Planning", *Policy Sciences*, 4, Elsevier, Amsterdam, p. 155-169.

Schoen, D.A., 1983, *The Reflective Practitioner: How Professionals think in Action*, Harper Collins, New York.

Schoen, D.A., 1987, *Educating the Reflective Practitioner*, Jossey-Bass, San Francisco.

Simon, H., 1969, The *Sciences of the Artificial*, MA: The MIT Press, Cambridge, Mass.

Simon, H., 1973, "The Structure of ill-structured Problems", *Artificial Intelligence*, 4, p.181-202.

Stiny, G., and L., March, 1981,"Design Machines", *Environment and Planning B: Planning and Design*, 8, Pion, London, p. 213-238.

Turner, A., and A. Penn, 2002, "Encoding Natural Movement as an agent-based System: An Investigation into Human Pedestrian Behaviour in the Built Environment", *Environment and Plan B: Planning and Design,* 29 (4), p. 473-490.

Underkoffler, J. and H., Ishii, 1999, " Urp: A Luminous-Tangible Workbench for Urban Planning and Design. ", *Proceedings of Conference on Human Factors in Computing Systems (CHI99),* ACM Press, p. 386-393.

Venturi, R., 1966, *Complexity and Contradiction in Architecture*, The Museum of Modern Art, New York.

Wanstall, B., 1989, "HUD on the Head for Combat Pilots", *Interavia*, 44, p. 334-338.

Webster, A., S. Feiner, B. MacIntyre, W. Massie, and T. Krueger, 1996, "Augmented Reality in Architectural Construction, Inspection and Renovation", *Proceedings of the ASCE 3rd Congress on Computing in Civil Engineering,* Anaheim, CA, p. 913-919.

Weghorst, S., 1997, "Augmented Reality and Parkinson's Disease", *Communications of the ACM*, 40(8), p. 47-48.

A Method to Index Images in the Wooden Architecture Domain

Terms hierarchy and weight given to terms

S. Kacher, J.-C. Bignon, and G. Halin
MAP-CRAI – Architecture and landscape Modeling - Research Center in Architecture and Engineering, Architecture School of Nancy, France

Keywords: Weight Given to Terms, Image Describers, Semantic Description, Visual Criteria.

Abstract: Architectural design is a domain where using pictures (e.g., drawing, photographs, …) is essential because the nature of the information transmitted by photographic image is often easier to interpret. The fact is that an image requires less interpretation than a text. The information transmitted by image (element shape, colour, light, ...) is already "put in shape" and so can be more easily integrated into the design process. This paper presents a way to index more efficiently an image database of the wooden architecture domain. Images in our databases illustrate real architectural elements. This work aims to analyse the representation of the real element illustrated by images. The analysis will allow us to identify some criteria related to the visual features of each image. The identified criteria will be used in a discriminating way to associate a weight with an indexation term describing its representation illustrated by an image. The importance of that representation (according to what is seen at first) is evaluated depending on graphic rules which correspond to the graphic properties of the representation of the element in each image.

1. INTRODUCTION

Design process requires research into ideas and documentation to help the designer in his task (Heylighen, 2000). The design process groups two distinguished parts where image use is essential. The first one concerns the problem formulation (image allows the user to express his design question

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 233-245.

and enables him to advance in his problem formulation). The second one concerns the problem's solution (images illustrate several potential solutions to the design's problem) (Lebahar, 1997).

When the designer has a design problem, he tends to transpose the elements appearing in his mind directly into the images that he visualises (Denis, 1982). This transposition happens as a virtual simulation with virtual objects. This kind of reasoning is based on the information transmitted by an image, on the possibility to develop it and to reuse it later. More precisely, it is based on the principle of permutation between the elements that appear in mental images and those represented by real images. Martine Joly (Joly, 1993) defines this principle as a way of segmentation which aims to identify the various components of an image. The identification is possible by locating the "autonomous" elements illustrated by the image. Once these elements are recognised, this principle allows the viewer to identify and replace them by other elements situated in his mental images. To realise these mental operations (Denis, 1989), the designer should have in mind other similar elements which are capable of being substituted which but are absent from the visualised images.

The objective of this research work is to better use the help that an image can give to the designer during a design process. More precisely, the proposition is not to suggest help to the designers who want to index their own image database. It is an indexation method that will be used in an information resource centre to offer an information service to designers. For this, in this paper we present a set of hypothesis and an experiment to define the best way to construct our image database. At first we will present the importance of image in the design process. Then we will show our proposition to index in a better way our image databases. And finally we will present an experiment which aims to validate our proposition.

1.1 The research system

In order to help the designer to find solutions to his design problem, an interactive and progressive research system (Bignon, Halin, et al, 2000) by image was developed by the MAP-CRAI[1]. Within the framework of this research, our work consists of defining a structured vocabulary, "a thesaurus" (Aitchison and Gilbrichrist, 1987), to describe architectural elements illustrated by the image databases. That defined vocabulary will be inserted into the system developed by the MAP-CRAI in order to better meet to the user's needs. The distinctive features of our research engine are divided in two parts. The first one concerns the assignation of a weighted

[1] "MAP-CRAI" Architecture and landscape Modeling - Research Center in Architecture and Engineering, Architecture School of Nancy.

value to every thesaurus term, which has been used for indexing images. The second one concerns a research process using images. For each image presented by the system and visualised by a user, the user can choose, reject or not give an opinion. A method of relevance feedback is used to propose new images for his query. The indexing document is represented by a weighted vector of thesaurus terms. A vectorial matching model is then used between the query and the indexing document. The results of this matching will be given as an ordered list of images representing the user's choices.

2. SEMANTIC IMAGE DESCRIPTION

One of the limits of the image is the "wrong semantic interpretation" which happens when the receiver interprets iconic information in a different way from that wanted by the transmitter. To reduce this wrong interpretation, and to make our database interpretable by the system proposed, we have decided to describe images in a semantic and unambiguous and structured language, "a thesaurus" (Kacher, Bignon, and Halin, 2003).

2.1 The proposed language

This stage of the work is based on a database containing about 1000 images. It is important to remember that for more efficiency, the corpus of the selected images illustrates architectural works belonging to a particular domain of architecture which is the wood construction field. This limitation allows us to add another dimension to the description language (Cabré, 1999) which is the type of "material used". The vocabulary is structured in three hierarchical levels (Rosch, 1977) and is divided into 4 classes *(Figure 1)*.

2.1.1 The architectural element

Indicate every physical part of a whole architectural work which has an essential or a particular function (post, beam, window, ...).

2.1.2 The material

Includes every wood material and its by-products (species, glued-laminated).

2.1.3 The products

Include any component aimed at protecting and decorating wooden elements (fungicide, impregnation).

2.1.4 The type of architectural realisation

Includes the name of the category to which the work element illustrated belongs (school, single-family dwelling).

Figure 1. The proposed language

2.2 Defining weights for image indexing

For each image in the database, a weight is associated to each thesaurus term. It is important to make clear that each term of the thesaurus corresponds to the "name" of the real architectural element, which is illustrated by an image. Weight values are given according to the importance of that illustration. This importance is evaluated depending on graphic criteria which correspond to the graphic features of the representation of the element (Bignon, et al, 2000) on each image and are defined as follows:

2.2.1 The area occupied on the image

Depends on whether the illustration of the real architectural element occupies a large area in the image or not (*Table 2)*. It is important to remember that images belonging to our database illustrate real objects put in situation. Then, we can say that the representation of the element is in prominent position if it occupies a larger visual area than the other elements which surround it. The visual area depends on the shot distance and on the size of the photographed element.

Table 2. The area occupied

	In which of these images is the roof the most important?	

2.2.2 The likeness with its archetype

Depends on the likeness of the illustration of the real architectural element to the ideal model shared by professionals belonging to the same domain. The fact is that if the representation of the element keeps the structural and spatial properties of the real objects so allowing the viewer to identify the element appearing in the image, the element will be easier to recognise (Reed, 1999).

Table 3. The archetype likeness

	Do these elements look like their archetypes?	

2.2.3 The contrast with the image background

Depends on the capacity of the illustration of the real architectural element to emerge from the rest of the image. The element will be in a conspicuous position if the representation of the element contrasts strongly with the rest of the image (colour, light, ...).

Table 4. The contrast

	Do these elements contrast with their image background?	

2.2.4 The focus

Depends on the position of the illustration of the real architectural element in the image. If it occupies the centre of the image (diagonal junction), the element should be more obvious than the rest of the elements illustrated.

Table 5. The focus

	Are the windows more obvious in the left image than in the right one?	

2.2.5 The completeness

Every representation or illustration of a real architectural element shows only a part of this element. This graphic criteria "completeness" depends on the fact that the part of the illustration of the real object represents the semantic features (Reed, 1999) that allow a viewer to identify the represented element.

Table 6. The completeness

	Is the element more easily identifiable in the right image or in the left one?	

3. THE EXPERIMENT

This experiment concerns the validation of the weights associated with the elements recognised. Weights assigned to the indexation terms according to these criteria will be used by the system as a basis of calculation for the relevance feedback. In order to validate our hypotheses, an experiment has been carried out.

- The first hypothesis concerns the fact that a relation exists between the rank and the value associated to each criterion.
- The second hypotheses concerns the fact that a relationship exists between the rank associated with each term and the kind of the architectural element to which belong the term.

At the moment, all these criteria are at the same level of importance. The second hypothesis concerns the fact that all the criteria do not have the same value.

3.1 The subjects selected for the experiment

The subjects who participated to this experiment are divided in two groups: (1) Architects in the wood construction domain (2) architects researchers.

3.2 The protocol of the experiment

The experiment was performed in several stages aiming to fill in the table illustrating in the (*figure 2)*:

Stage (1)

This first step is performed by presenting to the subjects a series of images illustrating concrete architectural elements on sheets of paper. Below every image a table with 7 columns. One of these columns includes the list of terms describing the architectural elements illustrated by images. This list is classified in alphabetical order.

Stage (2)

This second step is performed by asking subjects to classify in decreasing order the list of terms according to what they consider important in each image or not. Then they associate the rank "1" with the term describing the most important architectural element in each image.

Stage (3)

Image

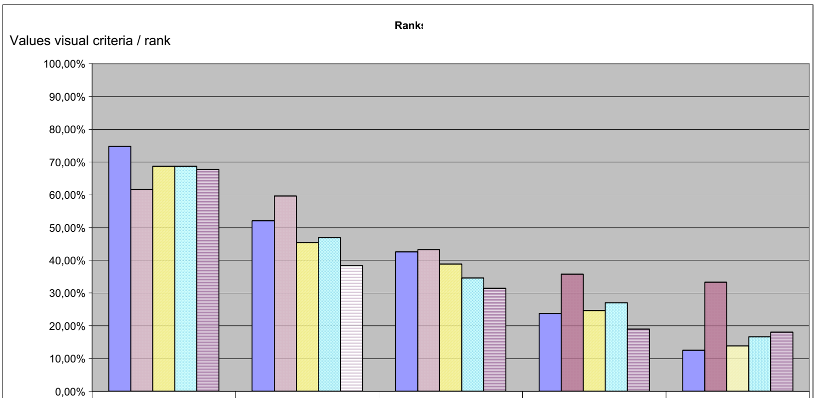
Rank	Element	The occupied area	The archetype likeness	The contrast	The focus	The completeness
2	Cladding	x	x		x	x
5	Outside flooring	x	x	x		
1	Outside shutter	x	x		x	
4	Outside terrace	x	x			
3	Window		x	x	x	x
Stage (2)	**Stage (1)**	**Stage (3)**				

Figure 2. The experiment protocol

This third step is also the last one. It was performed by asking the subjects, for every term in the presented list, to tick the shared cell on the table only if the graphic criterion is filled in by the representation of the element in the image. For example in the figure 2, a subject classify the “cladding” term at the rank 2 and selects the occupied area, the archetype likeness and the focus criteria with this term.

3.3 The results

This experiment showed that the subjects put in relation the number of the selected criterion with the importance of the rank. In the figure 3, we showed the rates associated by the subjects of the experiment to every visual criterion according to the rank. The results allow us to point the most important visual criterion for each rank. Then, for the first rank the subjects selected more often the “occupied area” criterion. For the other ranks the most important visual criteria which have been selected is the “archetype likeness”.



	Rank N°1	Rank N°2	Rank N°3	Rank N°4	Rank N°5
Surface	74,75%	52,02%	42,59%	23,81%	12,50%
Archétype	61,62%	59,60%	43,21%	35,71%	33,33%
Contrast	68,69%	45,45%	38,89%	24,60%	13,89%
Focus	68,69%	46,97%	34,57%	26,98%	16,67%
completeness	67,68%	38,38%	31,48%	19,05%	18,06%

Figure 3. Rates of graphic criteria with ranks

The subjects also had the opportunity to criticise or to add missing terms or missing graphic criteria. For some of them, they asked to be added as graphic criterion, “ the repetitiveness of the element “. Actually, they considered that when an illustrated element was repeated it became important in the image even if it did not necessarily fill all other criteria.

We also analysed the relation between each kind of architectural elements and the visual criteria. For this, we classified all the terms describing images in the experiment according to their geometric features. For example we classified floors and walls as “planar elements” and posts or beams as “linear elements (Ching, 1996). Then we defined 4 families:

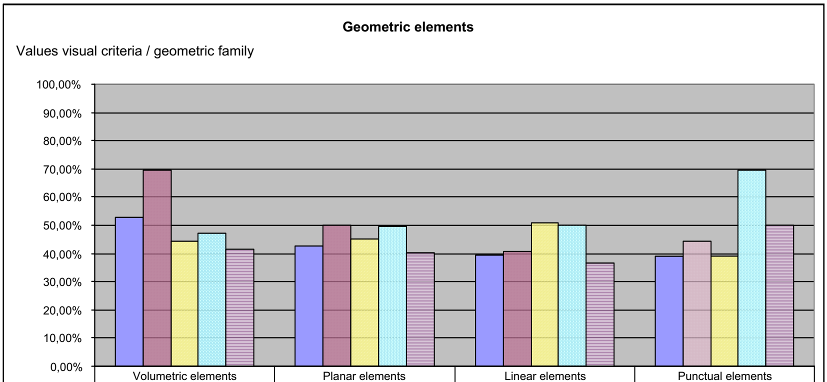
- Punctual elements: are seen as a singular element in a group. They could be assimilated to a building product such a furnishing element or a junction point between two linear elements (table, chair, spigot joint, ...).
- Linear elements: are principally defined by a length, a direction and a position, these include post, beam, ...
- Planar elements: are defined principally by an area (length and width) and also an orientation and a position. We can identify 3 kinds of planar elements related to their position in space. The first one is the “overhead plane” which forms the upper enclosing surface of a space. The second one is called the “wall planes”, they possess principally a vertical orientation and they are also used to shape and enclose architectural space. The third one is called the “base planes” which constitute the ground plane that serves as the lower enclosing surface of an architectural space.
- Volumetric elements: all volumes can be analysed as the sum of point elements, line or edge elements, and plane or surface elements. All these elements and more specifically the surface elements, define the limits of a volume.

The experiment allows us to identify the visual criteria which are strongly related to each family. As a result we obtained figure 4:

- For volumetric elements, the visual criterion which has been selected very often is the archetype likeness.
- For planar elements, the visual criteria which have been selected very often are the archetype likeness mixed with the focus criteria.
- For linear elements, the visual criteria which have been selected very often are contrast mixed with focus.
- For punctual elements, the visual criterion which has been selected very often is the focus.

3.3.1 The results analyse

The results obtained by this experiment allow us to validate some hypotheses such as the fact that the more an important rank is associated with a describing term, more subjects select a greatest number of visual criteria. As another result, we can conclude that, for the subjects of the experiment the most important criteria are the occupied area in the image and the archetype likeness.



	Volumetric elements	Planar elements	Linear elements	Punctual elements
Surface	0,527777778	0,42650463	0,395061728	0,388888889
Archetype	0,694444444	0,501446759	0,407407407	0,444444444
Contrast	0,444444444	0,450810185	0,50617284	0,388888889
Focus	0,472222222	0,49537037	0,5	0,694444444
Completeness	0,416666667	0,40162037	0,364197531	0,5

Figure 4. Rates of geometric elements with visual criteria

This means that if an illustrated element is recognise because it looks like the ideal model of the viewer and if it also occupies a large area, it is will be more important than the rest of the illustrated elements.

We also obtained with this experiment other results related to the geometric features of the elements. It appears that for volumetric elements the most important visual criterion is the archetype likeness. This means that if a volumetric element looks like the viewer's archetype, the volumetric element becomes very important to him.

For the planar elements we obtained the best rates also for the archetype likeness. For the linear elements we obtained the best rates for the contrast and the focus criteria. This means that because of their lack of surface, the linear elements needed to be in contrast and in focus to be visually important to a viewer. And finally, for punctual elements the most important criteria is the focus. Then, if a punctual element is situated in the image centre, it will be more obvious than the rest of the elements illustrated by the image.

4. CONCLUSION

In this article, we have presented an experiment from which the results obtained will be defined as indexation rules to improve image indexing in our database. These results will be implemented and integrated with the research system tool by image, in order to better index images. The objective is to build a useful database in which a designer will search for and find the relevant solution to his design problem.

At first, the achieved results allow us to identify the most important graphic criteria to index images. The fact was that, before the test they were 5 criteria to each of which we associated the same weight and the sum of their weights was associated with the term used for the image indexation. In a second way this experiment allows us to carry out the future indexation with a new graphic criterion, which is "the repetitiveness of the element ". In a third way, the experiment showed which visual criteria are related to each geometric family of elements.

Consequently, this experiment allows us to propose an indexation method to further image indexing:

- Firstly we must identify the element represented by the image that will be described.
- Secondly for each selected element, its representation will be compared to the results obtained by this experiment. According to the results obtained though this experiment, the system proposes rates to weight the indexed terms. For example, if an image illustrates a floor and if the illustration fill in the contrast criterion, the value associated to the term will be "0,45".
- Thirdly, the person who will index the images will decide if she/he will validate or refuse the rates proposed by the system during the indexation process.

At the moment, the statistical conditions were satisfied, more subjects have been tested. Finally, another experiment will be undertaken with professional people to validate the fact that design activity could be assisted significantly thanks to images illustrating wooden references elements.

REFERENCES

Aitchison, J and Gilbrichrist, A., 1987, *Thesaurus construction: a practical manual*, Second Edition, London.

Bignon, J.C., Halin, G., Nakapan, W., 2000, "Building Product Information Search by Images," Proceedings of the *5th International Conference in Design and Decision Support Systems in Architecture*, Nijkerk, The Netherlands, p. 47-61.

Cabré,M.T., 1999, *Terminology: Theory, Methods and Applications*, ed. Juan C. Sager, Amsterdam, Netherlands.

Ching, F.D.K., 1996, *Architecture: Form, Space and Order*, second edition, VNR, USA.

Dauzats, M., 1994, *Le thésaurus de l'image : étude des langages documentaires pour l'audiovisuel*, Paris.

Denis, M., 1989, *Image et Congnition*, PUF, Paris.

Denis, M., 1982, “On figurative components of mental representations,” In: F. Klix, J. Hoffmann, &E. van der Meer (Eds.), *cognitive research in psychology*, Amsterdam.

Joly Martine, 1993, *Introduction à l’anayse de l’image*, Nathan Université, Paris.

Heylighen, Ann, 2000, *In Case of Architectural Design*, critique and praise of case-based design in architecture, Katholieke Universiteit Leuven.

S.Kacher J-C bignon and G.Halin, 2003, “Image indexing vocabulary in architecture: Taxonomic hierarchy and categorisation,” *EIA9: E-Activities and Intelligent Support in Design and the Built Environment, 9th EuropIA International Conference*, 8-10 October 2003, Istanbul, Turkey, p 109-118.

Kattnig, C., 2002, *Gestion et diffusion d’un fonds d’images*, Paris.

Lebahar, J.-Ch., 1997, *Le dessin d’architecte, simulation graphique et réduction d’incertitude*. Parenthèse, Marseille.

Reed, S. K., 1999, *COGNITION théories et applications*, Paris.

Rosch, E., 1977, “Human categorization,” In: Warren (Ed.), *Studies in cross-cultural Psychology*, volume 1, Academic press, New-York.

Hybrid Approach to Solve Space Planning Problems in Building Services

G. Bi and B. Medjdoub
School of Built Environment, University of Nottingham, UK

Keywords: Case-Based Reasoning, Constraint Satisfaction Problem, Ceiling Voids Layout, Complex Geometry, Large Problem.

Abstract: In this paper an object-based CAD programming is used to take advantage of standardization to handle the schematic design, sizing, layout for services in a building ceiling void. From the specification of the building 3D model, our software proceeds through different steps; from the determination of the standard number and size of fan coils to the generation of 3D solutions. In order to deal with more complex geometry and larger problems, we have used a hybrid approach: Case Based Reasoning (CBR) within Constraint Satisfaction Problem (CSP) approaches. In practice, engineers in building services use previous solutions and adapt them to new problems. CBR mirrors this practical approach and does help us to deal with increasingly complex geometry effectively, and meanwhile CSP has been used for layout adaptation. The results have shown that it is possible to define and implement standard solutions to produce designs comparable with current practice. The benchmarking exercise has underlined many advantages and made some suggestions for further development. This project is funded by The Engineering and Physical Sciences Research Council (EPSRC) in UK.

1. INTRODUCTION

Standardization is widely recognized as a key element in reducing design time, cutting construction costs and ensuring efficient design solutions. The previous project "Building Services Standard Solutions implemented in CAD" (B. Medjdoub, P. Richens and N. Barnard, 2003) has shown that it is

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 247-261.

possible to define and implement standard solutions to produce designs comparable with the practice. The previous project dealt with middle size problems within simple geometry using constraint-programming technique to implement the design rules and the solution generation algorithms.

In this project, our approach uses standard solutions in conjunction with IT and extends the work of the previous one for fan coils to make the solutions useable for more complex geometry. This will bring benefits in terms of increasing the range of applications for which the solutions can be used (with consequent reductions in design time etc.) and to improve its usability. The extension to deal with more complex geometry will be based on Case Based Reasoning (CBR) (B. Dave, G. Schmitt et al., 1994) (M.H. Sqalli, L. Purvis et al., 1999) (A. Aamodt and E. Plaza, 2000) as well as Constraint Satisfaction Problem (CSP) approach. Very often, design engineers retrieve similar solutions from case based and adapt them to new problems. CBR mirrors this practical approach and will help us to deal with increasingly complex geometry effectively. In this paper we do not deal with pipe routing.

The objectives of the project are:

- Implement the standard solutions in an industry standard CAD system using advanced techniques in Artificial Intelligence combining CBR with CSP.
- Develop an interactive and friendly user interface, simply to use with a high-level modification of the 2D and 3D solutions.
- Test and evaluate the standard solutions against conventional ones in a benchmarking exercise with our industrial partners.

2. FAN COIL STANDARD SOLUTIONS

The standard solutions were developed in consultation with practicing engineers. They are essentially a collection of rules that address the selection, sizing and location of services equipment in the ceiling void. The rules are in a number of different forms including:

- Schematics
- Written rules
- Geometric layouts

The ceiling void solution is based on a four-pipe-fan-coil system. Fresh air is provided by a central air-handling unit. Air supply to the space is via slot diffusers in the perimeter zones and the square diffusers internally.

Whereas with plant rooms, the main integration issues are between different services elements, for ceiling voids the emphasis is more on integration with non-services elements including beams, ceiling tiles and core areas. It is firstly necessary to define these before the suited services solution can be generated.

In practice, once the non-services elements have been defined, the services are located approximately in order from the most to least geometrically constrained. For the structural solution illustrated in Figures 1, the distribution runs are located next as these are constrained to pass through the beam holes. These are then followed in sequence by; diffusers selected to provide the required air flow and located to integrate with the lighting; fan coils selected to meet the zone load requirements and located to ensure maintenance access from below; ductwork from fan coils to diffusers; distribution ductwork (for air flow) and pipe work (for cooling water) to the fan coil units (where not routed through beams and columns); condensate runs from the fan coils to column droppers.

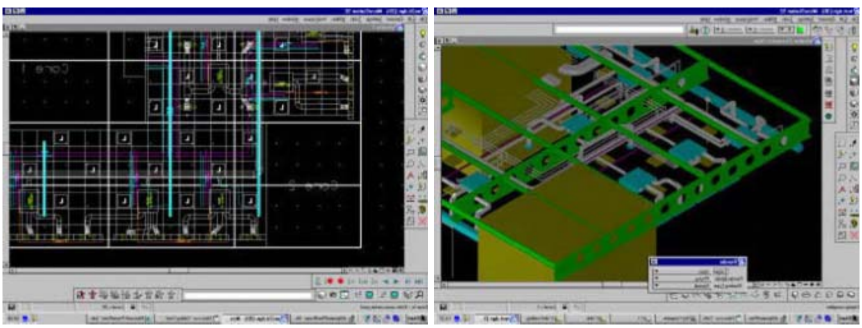


Figure 1. 2D (left) and 3D (right) ceiling layouts

The rules and solutions documented in the following table (Table 1) are for a four pipe ceiling mounted fan coil system as the HVCA (DW144, 1998) standard solution.

Table 1. Four-pipe ceiling mounted fan coil system as the HVCA standard

Ceiling voids	Type	Rules
General	Locations	The equipment location sequence is generally as follows: • Luminaries • Diffusers • Fan coil units • Pipe work • Ductwork

Ceiling voids	Type	Rules
		Equipment (i.e. diffusers and lights) should not be located in adjacent tiles unless there is a reasonable margin around equipment.
Fan coil units	Selection	Units selected from manufacturers standard range using chilled water for cooling and LPHW for heating as required. Standard spigot plenum connections used for supply - blanked if not all required.
	Sizing	Following variables defined for selecting size of fan coil units: • Total air volume • Fresh air volume • Cooling required • Heating required • External pressure drop • Noise level If largest unit not big enough to serve zone, then treat with 2 of the same type, then 3 of the same type, and so on.
	Location	Heating and cooling pipe work connections staggered to assist coordination. Locate fan coils on riser side of zone to minimize condensate and other pipe work runs and air flow direction changes from the risers through unit to the diffusers (see Figure 2). Locate 50-100 mm below slab to allow for slab inconsistencies and slope towards drain exit. Limit distance to diffusers to keep pressure drop down. Avoid grouping intakes as this can cause acoustic problems through reinforcement.
Diffusers	Selection	Perimeter zones - slot diffusers selected from manufacture standard range. Internal zones - square diffusers selected from manufacture standard range.
	Sizing	Use manufacturers sizing algorithms to meet throw, pressure drop and noise requirements.
	Location	Located square diffusers to nearest possible location to suit lighting layout. Locate slot diffusers along perimeter.
Ductwork	Selection	Local ductwork from fan coils to diffusers circular. Distribution ductwork circular up to 200 mm, flat oval above to limit depth requirement. Use standard ISO/DW144 ranges (DW144, 1998). Use fittings as defined by BSRIA (1995) Standard Details project.
	Sizing	Use CIBSE Guide (1986). Local ductwork from fan coils to diffusers sized as plenum spigot connection subject to a maximum velocity of 3 m/s. 5 m/s maximum velocity limit for ductwork distribution to fan coils.

Ceiling voids	Type	Rules
	Location	Run duct and pipe work headers out from riser and tap off to fan coil units. Route ductwork through cellular beams if possible, otherwise run below beams. Route ductwork down centre of area to be served. Branch supply ductwork local to risers to facilitate crossovers. Terminate extract stub duct with attenuator as “bell mouth”.
Pipe work	Selection	Use fittings as defined by BSRIA Standard Details project.
	Sizing	For LPHW and CHW pipe work use CIBSE Guide C “steel pipe” (bigger and higher k factors). Base sizing on 200 Pa/m and never exceed 250 Pa/m. Size condensate at 20 mm from units, 40 mm for 2 or more units, and 50 mm in risers.
	Location	Try to run in pairs (side by side) for F & R but not essential. Reverse return arrangements preferred where possible. Share commissioning sets if adjacent units similar (cost and commissioning benefits) running pipe work with a self-balancing “T”. Route through cellular beams if possible, otherwise run below beams.

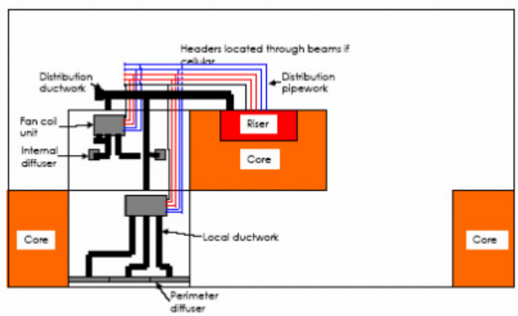


Figure 2. Ceiling void layouts - fan coils

3. OBJECT MODELS OF CEILING VOIDS SOLUTION

The object model holds three main classes of objects representing the ceiling voids space geometry, equipment and pipe work. Each class is characterised by attributes and constraints.

3.1 Space

The space class contains two sub-classes: the zone class and the structural element class (e.g. column and beam). The zone geometry can vary from a simple rectangle to an ellipse or a curved shape with known obstructions, doors and external walls. The attributes of a zone depends on its shape, if the zone is a polygon it will be defined by a set of points *(X[n], Y[n], Z[n])* which represent the polygon vertexes, where *n* equals to the number of all vertexes. The internal structure is characterised by a reference point, its length and its width.

3.2 Equipment

The equipment class includes the fan coil and the diffuser classes. These classes are characterised by a reference point (X, Y, Z), a length, a width, a Height and an orientation attribute defined as a constraint discrete variable defined over the domain *{0°, 90°}* (see Figure 3, left).

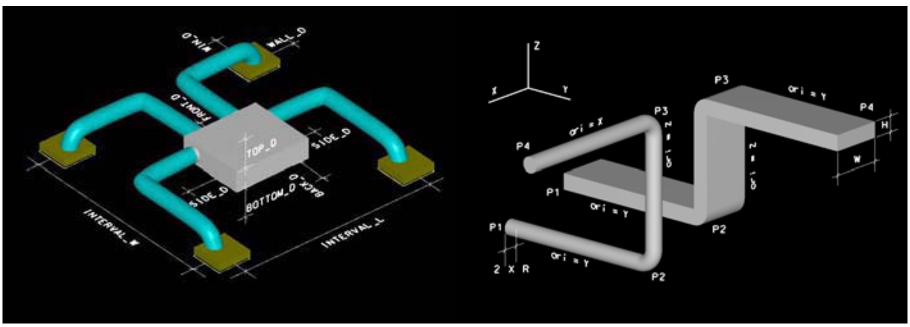


Figure 3. (left) The size and distance attributes of fan coils and diffusers. (right) Geometrical representation of pipe with set of 4 points (P1-P4) corresponding to 3 segments, 2 bends

3.3 Pipe

The pipe class is defined by a set of points *(X[n], Y[n], Z[n])* and a radius *R*. Each pair of successive points defines a segment of the pipe, each segment has an orientation variable defined over the domain *{0°, 90°, 180°, 270°}*. A class constraint is defined to ensure that two successive segments have different orientations. We consider the number of bends equal to the number of pipe segments minus one (see Figure 3, right).

4. CEILING VOID SPACE ALLOCATION PROCESS

The generation of a solution will follow four main steps:

1) Definition of the fan coil number and type.
2) Floor zoning.
3) Case retrieving.
4) Constraint-based case adaptation.

4.1 Definition of the fan coil number and type

Buildings have different functions like offices, shopping malls, hospitals, restaurants etc, and each one need a particular cooling and heating loads to serve the indoor environments. So only after defining the floor function, we could deduce the fan coil number and type. For example, if the floor is for an office use, the perimeter (window side) cooling load is 100w/m^2, and the setting is $6 \times 4.5\ \text{m}^2$ per fan coil. For the internal area (area far from the perimeter), the cooling load is $50\ \text{w/m}^2$, and the setting is $50\ \text{m}^2$ per fan coil[1]. To find out the appropriate fan coil type, we also need the necessary environment data like air density, specify heat capacity, cool temperature difference etc. All these data are saved in the application database, thus through the user interface the engineers have just to choose the floor function to generate automatically the appropriate fan coil type and number.

4.2 Floor zoning

In the case of a building within a complex shape, the user will through the user interface divide the building floors in zones. As indicated in Figure 4 the zones are simple geometric primitives (e.g. circle, rectangle, triangle, trapezoid and ellipse).

The reason for floor zoning is because the shapes of large buildings are very complicated, which are composed of different geometrical primitives. Therefore, it becomes impossible to find solutions directly from this compound shape. To make it possible to solve the problem, our approach is dividing the complex shape into the basic geometrical primitive zones, such as rectangle, triangle, trapezoid, ellipse etc. These zones are linked together with joint sides, and the unlinked sides are perimeter sides. The fan coils and diffusers can be set within perimeter (window) sides and the internal areas of different zones; finally pipes will pass through the joint sides to connect them. Meanwhile, engineers are required to define the locations of columns,

[1] Data provided by Faber and Maunsell.

beams and internal structures (i.e. store rooms, lifts, stair rooms etc) where the pipes and fan coils need try to avoid going through (non-overlapping layout). For the locations of columns and beams definition, engineers are required to select all reference lines (row and column base lines) which cross the columns.

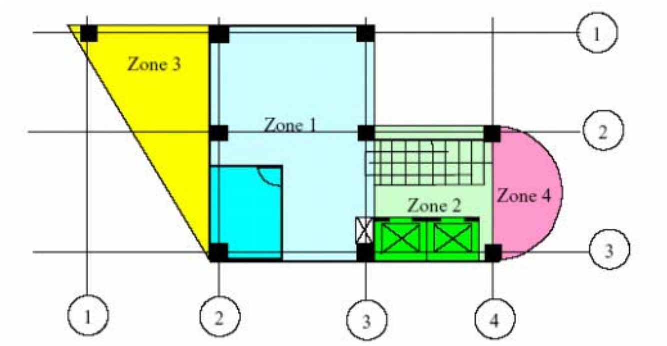


Figure 4. The floor is divided into 4 zones, which are assigned different colours

4.3 Case retrieving

After having defined the zones and their attributes, the next step consists on retrieving the most similar case from the case library. The solutions of fan coil systems for the basic geometrical floor shapes (rectangle, triangle, trapezoid, circle and ellipse etc) are stored in the case library, where the match between the floor zones and the retrieved similar cases is based on a descriptive index of the shape, which approach is called Case Based Reasoning (CBR).

The matching information we need to know includes the zone shape (i.e. rectangle, triangle, ellipse etc), side number of polygon, position of joint and window side, shape obliquity etc. The following table (Table 2) has listed the key information to retrieve the case. Then, from the case retrieving a first incomplete solution (see Figure 5) is generated and will need further adaptation to make it usable. The retrieved case contains all necessary topology and setting approaches. And therefore the case must be a program. While system finds the similar case, the case program attached will be retrieved and run to generate the first solution. The case program will set fan coils arrangement using grid topology and CSP approach, which can set the locations of fan coils and diffusers to avoid overlapped with internal structures, and also avoid setting outside the boundary of grid and perimeter.

Grid topology approach represents that the zone is divided into a number of grids, each grid will be assigned to one fan coil and no fan coil can be out of its grid boundary. As we known, the floor has been divided into zones; we need to set fan coils and diffusers in each zone. An easy and quick way to set them is using grid topology, where the zone is divided into grids, and the sizes of grids are different between perimeter and internal area. The size of grid will depend on the requirements of user. For example, if the floor is for office use, then the perimeter grid area is $L \times W$: $6 \times 4.5\ m^2$ and internal grid area is $50\ m^2$. System will automatically calculate out the length L and width W of internal grid based on the size of the zone.

Table 2. Key information to retrieve the case

Key information	Explanation
Shape type	What shape the zone is (i.e. rectangle, trapezoid, triangle, ellipse, circle, arc, curved etc).
Side number	How many sides that the shape has if shape is a polygon.
Horizontal side number	How many horizontal sides the shape has if shape is a polygon.
Vertical side number	How many vertical sides the shape has if shape is a polygon.
Top vertex number	How many top vertexes with same Y value the shape has if shape is a polygon.
Bottom vertex number	How many bottom vertexes with same Y value the shape has if shape is a polygon.
Primary radius and second radius	Check whether primary radius equal to second radius or not if shape is an ellipse, circle or arc.
Joint side number & position	How many joint sides the shape has, and where are their location. If shape is a polygon, joint side should be at perimeter. If shape is an arc or curved, joint side should be at the line perimeter.
Window side number & position	How many window sides the shape has, and where are them. If shape is a polygon, window side should be at perimeter. If shape is an ellipse or circle, window side should be at elliptical perimeter. If shape is an arc or curved, window side should be at arc and curve side, or at line side.
Shape obliquity	What the obliquity the shape is, the range is 0° to 90°.
Fan coil arrangement	What the arrangement the fan coils are (i.e. horizontal arrangement).

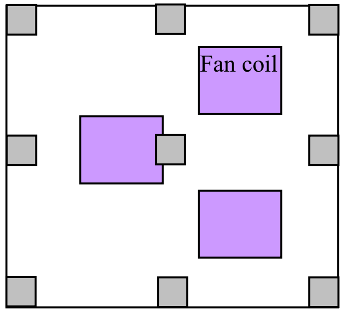


Figure 5. Incomplete solution generated by the case retrieving, one fan coil overlap with a column (this is a small example, and the floor shape is not derived from *Figure 4*)

4.4 Constraint-based case adaptation

The adaptation process is based on constraint programming techniques. The method used is by substitution, where we replace some parts of the old solution that does not fit the current situation requirements. Thus from the incomplete solution from the case retrieving step, the system will identify the incoherent parts of the solution, as indicated in Figure 6, where one fan coil overlap with a column. Next this incoherency is solved using simple constraints (e.g. inclusion constraint, non-overlapping constraint, dimension constraint) to make it coherent. The advantage of this approach is that each incoherency is solved separately, which decrease drastically the complexity of the problem.

4.4.1 Dimension constraints

Dimension constraints assign a minimal or a maximal value to the object constrained variables. This constraint is expressed by equality or inequality, i.e. length *L*, width *W* and height *H* of fan coil.

4.4.2 Inclusion constraints

Inclusion constraints represent the object must be within the space, and do not overlap with other spaces. These type of constraints are especially important for finding the locations of fan coils and diffusers, where they must be located within their local grid areas, and do not overlap with other grids.

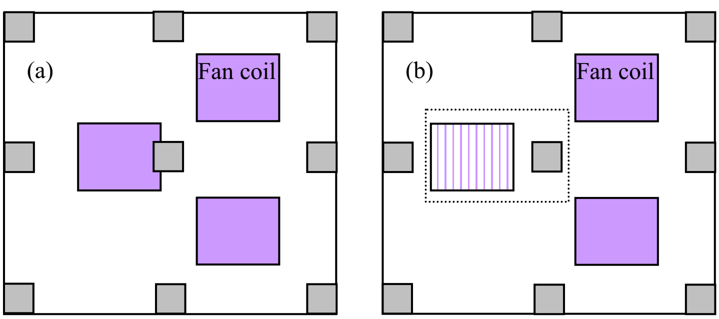


Figure 6. (a) Incomplete solution, (b) solution after local adaptation within the use of a non-overlapping constraint between the fan coil and the column (this is a small example, and the floor shape is not derived from *Figure 4*)

4.4.3 Non-overlapping constraints

Non-overlapping constraints represent the fact that two objects cannot overlap each other; it is automatically applied to all pairs of objects in our application (i.e. fan coil to structures, fan coil to fan coil etc). Non-overlapping constraints introduce a new non-overlapping variable with 4 elements *{North, East, South, & West}*. These orientation variables are represented in order to avoid making two objects overlapped.

5. IMPLEMENTATION & BENCHMARKING

This application is developed in JMDL (Java Modelling Language) as embedded in Microstation/J and JSolver the constraint programming system. Microsoft Access, the relational database application is used for the case library. Finally we use Microstation/J to layout the object model and the 3D rendering.

Our research has made explicitly progress, and has completed some parts of the work which can layout the fan coils and diffusers within polygon zones. Our completed parts include user interface, CBR and CSP programming. The user interface is composed of several panels and the implementation includes (see Demo in Figure 7 and 8) the data input, zoning process and equipments (fan coils and diffusers) generation. The main steps that user need to input are:

Step 1: Use place shape tool (i.e. place block, place polygon etc) to enclose the zones and internal structures in the floor.

Step 2: Define the floor information which includes floor drawing unit (i.e. millimetre or meter); floor type (i.e. office, shop, hospital etc); floor noise requirement NR; floor height, ceiling voids height and z value of floor bottom. And system will find out the type of fan coils to be used in the case based from the information given.

Step 3: Define the diffuser information which includes slot diffuser size, square diffuser size, interval of diffusers, and minimum distance from diffuser to wall and to window. If user thinks there would be necessary to modify the attributes of environment (i.e. perimeter and internal area sizes, area load rate, air density etc) to match the case situation, he/she can go to advanced panel to change them.

Step 4: Define the architectural coordinate lines (the lines locate the columns) which are including row and column lines. The aim to define these lines is to locate the positions of columns and beams.

Step 5: Define the zone by selecting the shape which enclosed the zone, and system will automatically generate the zone information (i.e. zone area, zone shape and zone position). Also, the user needs to define the window and joint sides of the zone by selecting the target side. Because zone shape could be many and varied (not only the rectangle), so to make user understand which side he/she want to target, a small scale drawing of the zone top view is displayed, and each corner of the zone is numbered (i.e. a, b, c …).

Step 6: Define the internal structures and the riser which user wishes fan coils topology are non-overlapped with them. Such structures like lifts, stairs, and some function rooms. The riser position is also need to be defined to give the start position of pipe routing.

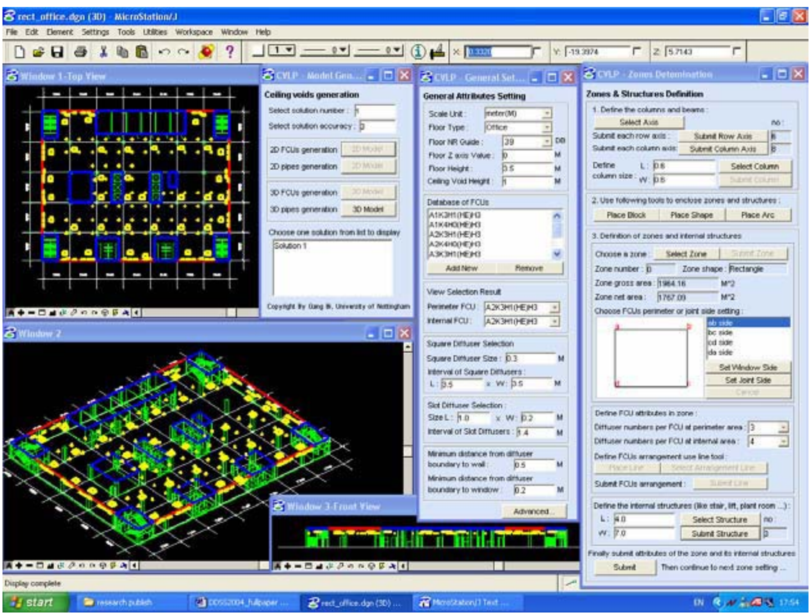


Figure 7. User interface panels to fan coil layout

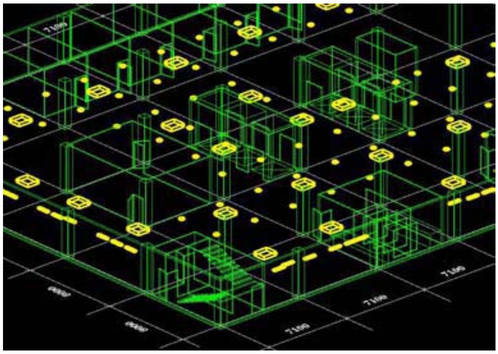


Figure 8. 3D fan coils and diffusers layout solution displayed on Microstation/J

The results have shown that it is possible to define and implement standard solutions to produce designs comparable with current practice. This benchmarking exercise has underlined many advantages and made some suggestions for further development.

The main advantages are:
1) The system deals with complex floor shapes.
2) The retrieving of the similar case is done in reasonable time.
3) The constraint-based adaptation approach is done sequentially which decrease the complexity of the problem.
4) The output from the solutions such as the 3D data model is beneficial to other parties in the supply chain.

The main improvements needed:
1) To deal with curved shapes.
2) To enrich the case library.
3) To develop a retain solution mechanism.

6. CONCLUSION

This approach has shown the potential to significantly reduce design costs by reducing design time, improve the quality of the solution and produce additional benefits elsewhere in the supply chain. On the computational part, the integration of CBR and CSP approaches did achieve a synergy, which produces results that could not be obtained if each mode were operating individually. Further developments are being done and concern mainly the case library enrichment and the complex problem of pipe routing.

Pipe routing includes the pipe routing for fan coils to their diffusers, and fan coils connection with riser and main pipe. Pipe routing will cover ventilation ductworks and cooling water pipes. To improve the system automation and reduce modification time, system can update the pipe routes after user modified the locations of fan coils and diffusers by hand. The principle of pipe work design is connecting the fan coil (X_m, Y_m, Z_m) and the pipe node (X_n, Y_0, Z_0) with the shortest distance D and the fewest bended corners $\sum_r (X, Y, Z)$, where *{minimize(n, D), 1 < n < m}*. Meanwhile, pipe route must be non-overlapped with fan coils, diffusers, columns, stairs and lifts etc. Therefore the general rules to create the pipe routing among the fan coils, pipes, and riser are:

1) The pipe routing must have the shortest pipe distance, minimizing the length of each pipe using the "branch and bound" algorithm.
2) The joints and bending points of pipe routing must be minimized.
3) The segments of pipe routing are normally horizontal, vertical or Z-depth direction.
4) All segments of pipe routing must be non-overlap with other objects (i.e. pipes, fan coils, diffusers, columns, stairs and lifts etc).

The idea to have a compromise between full automation and interactivity gives to the designer full control of the design while assisting him to solve complex problems automatically. This compromise is the main difference with the aforementioned approaches in facilities layout and pipe routing. In our application, several aspects were discussed which include case based reasoning approach, constraint satisfaction approach, algorithm and rules of layout solution enumeration.

The implemented user interface has shown an interactive system for ceiling voids layout design. Easy to use with high level modification of fan coils layout topology, pipe work position and pipe sizing. Output from the solutions such as the 3D visualisation is beneficial to other parties in the supply chain.

ACKNOWLEDGEMENTS

This project is funded by the EPSRC[2] in the UK. There are three industrial partners including Faber and Maunsell, Bentley Systems and Biddle Air Systems. The authors would like to thank Mr Nick Barnard from Faber and Maunsell and Mr Mike Price from Biddle Air Systems for their valuable help and comments about various aspects of this research project.

[2] The Engineering and Physical Sciences Research Council

REFERENCES

Aamodt, A. and E. Plaza, 2000, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", *AI Com – Artificial Intelligence Communications*, Vol. 7, No. 1.

Dave, B., G. Schmitt, B. Faltings, and I. Smith, 1994, "Case based design in architecture", *Artificial Intelligence in Design – AID' 94, Kluwer Academic*, p. 145-162.

Medjdoub, B. and B. Yannou, 2000, "Separating topology and geometry in space planning", *Computer Aided Design*, 32(1), p. 39-61.

Medjdoub, B. and B. Yannou, 2001, "Dynamic space ordering at a topological level in space planning", *Artificial Intelligence in Engineering*, 15(2001), Elsevier Science Ltd, p. 47-60.

Medjdoub, B., P. Richens, and N. Barnard, 2003 "Generation of Variational Standard Plant Room Solutions", *Automation in Construction*, 12(2), Elsevier Science Ltd, p. 155-166.

CIBSE Guide C, Reference Data, 1986, *"Section C4 Flow of Fluids in Pipes and Ducts"*.

BSRIA report, 1995, *Rules of thumb*, Technical Note TN17/95, BSRIA.

DW 144, 1998, *Specification for sheet metal ductwork – Low, medium and high pressure/velocity air systems*, HVCA.

Sqalli, M.H., L. Purvis, and E.C. Freuder, 1999, "Survey of Applications Integrating Constraint Satisfaction and Case-Based Reasoning", *PACLP99: International Conference and Exhibition on the Practical Application of Constraint Technologies and Logical Programming*.

Reduction Mechanisms Explored in Architectural Re-Design

Improving the understanding of precedents in a CBD tool

Jonas Lindekens
Vrije Universiteit Brussel, Belgium

Keywords: Architectural Re-Design, Design Process, Case-Based Design, Reduction Mechanisms.

Abstract: Observation of the design process of an architect shows that, while building up an argumentation for taking a design decision, different mechanisms of data transformation are used. The paper argues that this transformation is a key element in understanding architectural design processes. A theoretical description of these mechanisms forms the framework to discus a sequence of design decisions derived from a real-world design situation. After outlining how this can be implemented in a case-based design supporting tool, the paper concludes with a discussion of advantages and downsides the use of the tool might entail.

1. INTRODUCTION

The research reported on here is part of a broader research project on architectural re-design. The aim of this project is to understand how architects value existing buildings, and how their value judgement guides the re-design process. Additionally we try to use this understanding to develop a case-based design tool that can support (novice) architects when dealing with re-design. This paper investigates how design information can be represented in this tool so that users can easily understand it and adopt it in their own designs.

Two arguments direct our plea in this paper. First of all, in architectural design – and more specifically in re-design – *reduction* of context information plays a key role in the advancement of a design process. In new

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 263-278.

building design a large body of information is formed gradually during the process, as more design dimensions are tackled. In case of re-design a considerable part of this information is already present at the start of the process: the existing building embodies all dimensions of a mature design. This observation has steered us to concentrate on re-design for this investigation. Without the ability to reduce in an appropriate way the richness of the information involved, a possible information overload would make it very difficult for an architect to take any design step.

Secondly, we argue that similar problems arise when third parties try to understand a finished design. Too much information is present (in the artefact itself or in its representations) to easily reconstruct the design decisions. To explain a specific design decision, reducing the overloaded representations can be an adequate means to shed light on the underlying motives of a design.

Before illustrating these two arguments in sections 2 and 3, section 1 first discusses the intended mechanisms to clarify how transforming information can elucidate meaning.

2. TRANSFORMATION MECHANISMS

In their book ‘Design and analysis’, Leupen et al. explore ‘the diversity of analytic methods used by architects, designers, urban planners and landscape architects to understand the structure and principles of the built environment’ (Leupen, Grafe, et. al., 1997). By analysing historical and contemporary cases, an overview is created of mechanisms which provide an insight in designs made by others. In different chapters covering different dimensions (context, typology, construction, use and composition), it is shown how a specific dimension that influences the design can be identified and represented. In a final chapter the drawing techniques, called *transformation mechanisms*, are summarised.

The aim of *reduction* is to reveal the structure of a design. Several kinds of structures may be revealed: morphological structure, typological structure, spatial structure, technical structure, … In order to reveal the morphological structure of a building e.g., all information in the drawing is omitted which has no connection with the form. Similar reductions are made for other structures. Sometimes, however, references are maintained to allow for recognition of the original building. *Addition* is used to bring in information that is not graphically visible. This is often done after the drawing has been reduced. *Disassembly* comes down to drawing the object as if it were taken apart, which allows to reveal the relationships between separate elements of the object. Here reduction and addition can be added to make clear a specific dimension.

These techniques are not only proposed by Leupen et al.: Ching (1979), Baker (1989) and Unwin (2003) amongst others use the drawing as a means for analysis. Most often, these operations are used as ways to analyse an existing project. Leupen et al. indicate, though, that this analysis becomes only interesting when the abstract architectural types, morphological means and composition rules can be used again in a different context. In other words, analysing an existing building is only valuable when the knowledge achieved can be applied in one's own design practise. In this sense, re-design constitutes a special case. We can assume that architects use the same mechanisms to translate the existing building data. The analysis of the existing building offers information that can be valued and used in the new design. The only difference, however, is that it is not applied in a different context, but in the same location, even in the same building. Therefore a re-design project has to react to and interact with the existing building, which means that the analysed information cannot be used in the same way.

Before illustrating the use of transformation processes with analytical purposes, its use during the conceptual stage of a re-design process is exemplified in the following section.

3. INFORMATION IN THE RE-DESIGN PROCESS

The examples used henceforth are extracts from a protocol generated in the context of earlier experiments. They were originally conducted for testing the prototype of an on-line design case base called 'DYNAMO' (Heylighen and Neuckermans, 2000) and have been re-analysed in the context of this research (Lindekens, Heylighen, et al., 2003a; 2003b). Four architects – two junior and two senior designers – were invited to develop a concept for the reorganization of and extension to an architecture school, which is located in a 16th century castle (*Figure 1*). The protocol used here applies to one of the senior designers. Although not a 'world-famous' architect, he can be considered an expert designer in that his work has attracted major design awards and/or won important competitions.

The design task consisted of the reorganization and optimization of the West wing of the castle (design studios, lecture rooms, secretariat and photocopy room) and its extension with a reception hall, material museum and exhibition room. Small scale plans of all floors and pictures of the exterior of the building were provided in the experiment room. The architect knows the building quite well because he is currently design teacher in the same school.

Apart from having access to DYNAMO, the architects could go about the task as they preferred. However, two restrictions resulted from the method used: they were asked to 'think aloud' and the design session was limited to two hours. During these sessions all actions of the designers were audio- and videotaped. At the end, the drawings and notes were collected and chronologically numbered. After the session, the tapes were transcribed and the resulting text was divided into different segments.

Figure 1. Exterior view of the castle (photo: Paul Van Aerschot)

3.1 Lines of thought in the design process

The resulting protocol clearly fell apart in coherent fragments. Each of them forms a logical sequence of thoughts, a *line of thought*. A line of thought is defined as what Suwa and Tversky call a *dependency chunk:* 'a sequence of conceptually interrelated design thoughts, each of which was evoked in relation to preceding thoughts in the chunk' (Suwa and Tversky, 1997). We purposely did not adopt the name *dependency chunk* here because we also want to allude to the *parallel lines of thought* as defined by Lawson (Lawson, 1993). The notion of having several ways to tackle a design – as is elegantly described by this term – is appealing to us.

In what follows, two lines of thought are described to illustrate some of the ways in which the architect treats the design information in this protocol. One of the first things the architect does is analysing the ground floor plan

(*Figure 2*). He is looking for measurements on the drawing, which apparently add meaning. Because they are missing, he adds these measurements to the drawing. Except for these pragmatic dimensions, he also adds some circles and arrows. The circles point to the windows that give to the inner court of the castle, while the arrows indicate possible communication these openings might offer between the inside and the outside spaces.

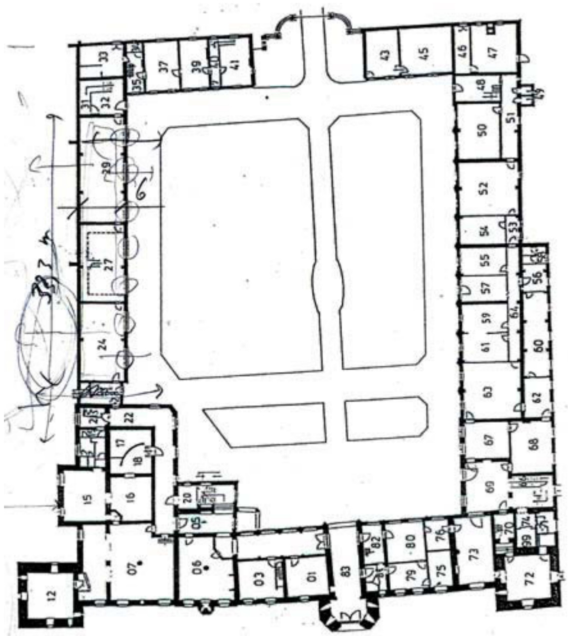


Figure 2. First notes on the given ground floor plan

Some time later, but still at the beginning of the design session, the architect wants to determine the character of the new part of the building. Although he is not looking at any picture or drawing of the castle, he makes the following consideration:

Table 1. Character of the building (own translation from Dutch)

0:16:41	since the castle is a clear entity…
0:16:44	consisting of … of course… several wings and euhm elements
0:16:49	… but still has a certain identity
0:16:51	It is important that that part which will be added
0:16:54	gets its own identity
0:16:56	and definitely does not become an annex to the castle
0:16:59	It should not become one with the castle
0:17:01	It should have a certain autonomy
0:17:02	… a certain contrast

0:17:05	So it can be a clear entity
0:17:07	It does not have to be contrasting
0:17:08	It could still be brick euhm
0:17:13	In the end it is also a space which has a lot to do with the functions in the castle…

It seems that the architect has a clear image of the existing building. He knows the building well, so probably he has even different images in mind when he talks about the building. Figure 1 shows one image which – in our opinion – describes best what the architect is discussing.

In the first line, he filters out one element from the rich source of information in his head: 'the castle is a clear entity'. He does not explain what this is, so probably this is obvious for him. He continues: 'consisting of course of several wings and euhm elements… but still has a certain identity'. It seems as if the fact of having several wings weakens the fact of being a strong entity, but it is still strong enough to use as an argument. So probably, for the architect 'being a strong entity' means creating a self-contained, independent image. Here this image is mainly created by the scale, the uniform use of materials, the static, symmetrical construction and the gothic style of the building. In the rest of this fragment, the architect considers the material for the new part. Out of the rich body of information in his mind, he abstracts only a very limited number of elements, and uses these to take his first design steps.

Table 2. Circulation between the different building parts (own translation from Dutch)

0:43:54	actually an important circulation euhm connection is still between…
0:44:03	the elbow of the mill here
0:44:05	Because here the important steps are situated
0:44:08	also the shortcut towards auditorium 'de molen' [the mill]
0:44:12	So this is an important circulation
0:44:17	we the stairs over here euhm
0:44:20	will be restored
0:44:21	the former stairs are continued till the ground floor
0:44:25	so that a better connection is created that euhm
0:44:30	that corridor with the gothic arches euhm towards euhm the stone staircase
0:44:37	so we will make sure that this actually is the most important
0:44:41	the most important route
0:44:43	maybe here somewhere
0:44:44	we will place a building somewhere
0:44:45	thus that there is a kind of division
0:44:49	Between a zone
0:44:50	a calm place to sit
0:44:53	euhm where also euhm in the afternoon euhm a sunny elbow can be created
0:45:01	so I think that we have to situate the workshops at that side
0:45:06	euhm such that of course the sash windows that euhm outside space…
0:45:13	and that it is not disturbed by people walking by
0:45:16	and always disturb the workshops or euhm
0:45:22	so this will be more the passage with the parvis

Later on in the session, he is eager to determine the relationships between the existing and the possible new part of the building.

During this fragment the architect is almost constantly drawing. He is adding new information to the drawing. In contrast to the first fragment, he is not drawing directly on the existing plan, but on tracing paper he placed on top of it. On this paper (*Figure 3*), he first traces the major outlines of the existing building, leaving out almost all detail. Then he starts adding symbols for connections, trees, volumes, pavements, … All are drawn in a very sketchy way, with multiple lines on top of each other. This contrasts with another drawing he creates later on, where he actually redraws the same plan be it in more accurate lines.

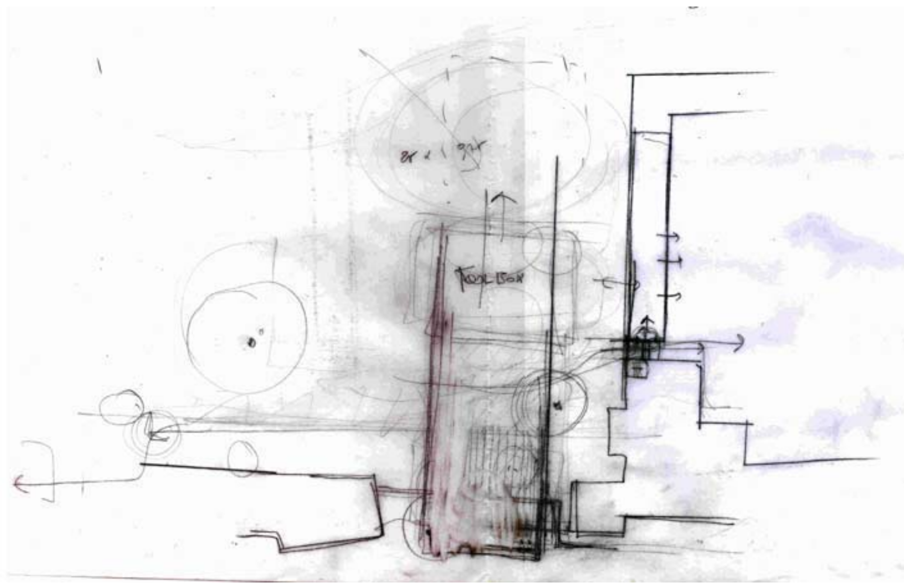


Figure 3. Active drawing phase during the design

3.2 Transformation during design

In the examples given above the architect tries to 'read' specific data from the building. He tries to understand how the building is like in several ways. He derives several kinds of information from the building (dimensions, visual relations, functional relations, meaning, spatial organisation) and uses several sources to do so. In this way, he develops a story that becomes richer as more angles are used, more layers are stacked, or more lines of thought are considered.

Being able to extract these data from the given sources seems to be important because several design decisions could be made based on the information retrieved. Now, how can the actions described in these fragments be interpreted? For an inventive process, Goldschmidt contends, 'what is required is an ability to use the representational act to reason with on the fly' (Goldschmidt, 1999). This could explain what we have described above: the architect uses different techniques to reason with (and in these he applies reduction, addition and disassembly), each with a specific outcome.

First he uses an existing drawing and 'reads' new information in it by sketching directly on the plans (addition). The visual links are not present in the drawing, but can be derived from it. He is also pointing to elements that are there, the windows e.g., but to which he attaches a specific value.

A second technique mainly makes use of imagery. He actually 'reads' his own mind in search for design clues. This imagery seems to be very powerful, because the architect continues like this for a long time. According to Lawson and Loke (1997), this can be explained by the fact that words have a degree of uncertainty, which is why they are evocative and direct towards more creative design processes. The meaning of this fragment is difficult to grasp by third parties though, it is not entirely clear what he means exactly by e.g. 'entity' or 'identity'. Goldschmidt claims that sketching serves as an extension of imagery. The fact that he only sorts out some elements for use in design decisions could be explained by his preoccupations (Goldschmidt, 1999) or gazing point (Taura et. al., 2002): because of his experience he is automatically driven to his own preferences.

In a third technique he is tracing the plan. This way a reduction of the plan is made, bringing forward what is important to the architect at that time. It enables him to 'read' only what is essential and omit disturbing side-information. Then he starts adding multiple rudimentary, dense, ambiguous, lines to this drawing. Goldschmidt refers to this as representations that are produced, evaluated, transformed, modified, refined and replaced by others. Suwa (1997) suggests that the architect sees 'unanticipated relations and features that suggest ways to refine and revise ideas, using the cycle: sketch, inspect and revise'. Both explanations direct towards the notion of uncertainty as creative means, as was the case with the previous technique. Schön (1983) talks about the 'backtalk' of sketches: one reads more information off the sketch than was invested in its making (Goldschmidt, 1997). Lawson (1997) adds that 'the drawing does not explain the rules by which it should be interpreted', which is an opportunity for creativity of the expert designer, but can be a problem for the novice.

Later in the design session, the architect redraws the drawing shown above. Lines are more clear now, and unambiguous. At this stage the architect has taken decisions which he wants to verify and optimise in this drawing.

4. INFORMATION IN THE ANALYSIS PROCESS

In the examples given above, we were able to trace the design intention of the architect in a detailed way. In many cases though, it is impossible to observe a subject this intensively. Architects are not always willing to cooperate in an experiment as this one. In practise a design takes much more than two hours and for use over long periods of time these methods are too time-consuming. A case-based system as is proposed here[1] is only as valuable as the contents of its case base. Using these operations as a means to almost 'reconstruct' the architect's design intentions is a way to overcome the problem of observation and ease the accumulation of cases. As a basis, the 'raw' material showing the existing situation and the result after re-design can be used. A systemised way of doing this can help gaining insight and building a knowledge base for re-design.

By way of example, this section will reconstruct one line of thought, based on a case study. To this end, we observed the re-design of a medieval watermill by noA-architects, a young architecture office in Brussels. The design is developed for a national architecture competition among four selected architecture offices in Belgium. The offices were asked to develop a concept for the re-use of the watermill. They had to make a proposal for a new function in the building and show how they would treat its refurbishment. Participatory observation was chosen to minimise the interference of the experiment – and the presence of the observer – with the actual design task. The office consists of three partners, at that time working with three collaborators. The latter only participated in the final phase of the competition, merely helping with the presentation of the project. For several reasons, a competition design was chosen for observation. The first phase of a competition typically concentrates on the conceptual design, which is our field of interest. Due to the fixed deadlines a competition imposes, the design phase is limited in time, in this case taking about six weeks, which makes it more convenient for research purposes. As pointed out above, it is not possible to audio- and videotape all design sessions for such a long period. Nevertheless it has been possible to keep track of the main design decisions and their interconnections. Another advantage we see in selecting a competition design is that, in order to be clear to a jury, the proposed concepts should be unmistakably represented and are therefore better documented than when designing for a client[2].

[1]See also the description in section 4.

[2]This notion has been confirmed in a conversation with Johan Anrijs (51N4E space producers, architecture office, Brussels), who indicates that for each design task the office makes a different kind of presentation, suited to the respective target audience. In doing so, they noticed that competition designs demand a greater effort compared to designs for

4.1 Line of thought as an analysis

Figure 4 explains one line of thought, which considers the spatial configuration in the new setting of the building. The first drawing (1) displays the current section of the building. The light coloured part burned down and was demolished. In the second drawing (2) only the spatial piling of similar block shaped spaces is indicated inside the circumscribed volume. The cross section shows how these spaces interconnect via a series of doorways along the east façade. This spatial arrangement is a quality of the building that is maintained in the re-design. The third drawing (3) shows how the machinery – which was scattered throughout the whole building – physically linked all these spaces by means of driving systems transferring energy from the waterwheels to the machines, transportation systems for raw and grinded material, cleaning systems etc. Most of these machines have already disappeared, but traces of them cutting through walls and floors are still visible in the building. In the drawing the architects made this is shown in a symbolic way: it was not important for them to capture the real places of these mechanisms. The symbolic value of the lively activity that must have been present in the building, was what they found intriguing. The building acted as a real machine in all its extremities. This notion is translated and shown in the fourth drawing (4). Here the real machine activity has disappeared, but is replaced by human activity in the public spaces. The architects want the public to experience the whole building, up to its extremities. Therefore they place most public functions on a diagonal through the section of the building. By doing so, they hope to reinstate activity – meaning human presence – till the extremities of the watermill. As a consequence, the upper left and lower right spaces can accommodate more private functions, which are necessary to develop a realistic building program. The last drawing (5) shows the different spaces with the proposed functions. New vertical circulations are added to make a diagonal circulation through the building possible.

Drawing 4 and its accompanying explanation already indicate some of the architects' design intentions, which would have been impossible without the intermediate abstracted drawings (2-4). Of course not all clarifications for the design decisions are given in this line of thought. Other lines of thought will give more information on the exact location of the different functions. It is obvious that the sizes of the available spaces and the space required for each function played a determining role too, while possibilities to add staircases determined whether the diagrammatic intent could be realised physically (3).

---

individual clients. The latter need less graphical material as this can always be supplemented with verbal explanation.

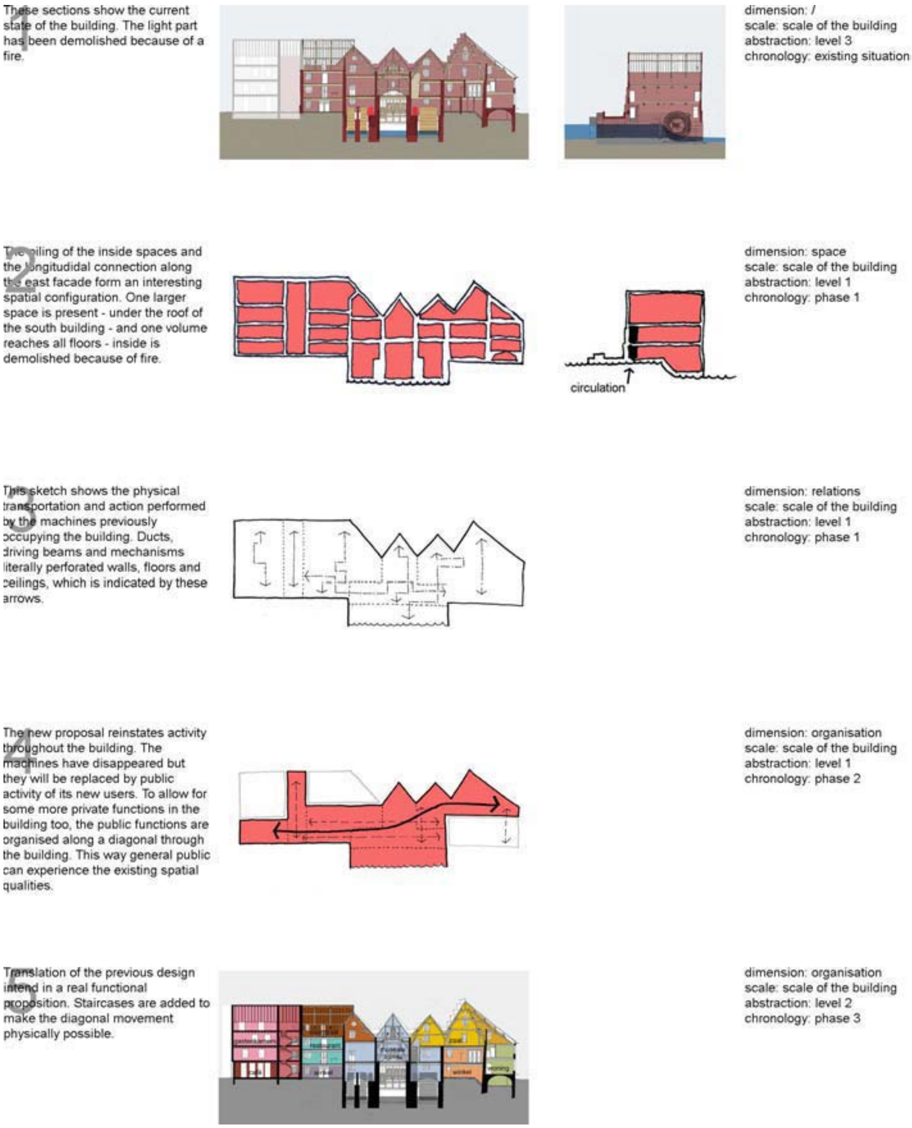


Figure 4. Line of thought regarding the spatial organisation

4.2 Transformations in analysis

How are the operations from the previous section used here? The first drawings (1) are based on an existing survey plan and show internal spaces and divisions. These drawings are already an abstraction of reality: textures, fissures, real colours, light, … are not represented. Still too much information can be read in this drawing to clearly make a single point: beams, roof timber, floor construction, windows, lock mechanism, … are all elements that distract from the spatial configuration. That is why we added the second drawings, which make abstraction of the elements unnecessary for the design intent under consideration (reduction) (2). The longitudinal section shows only the spatial piling of separate block shaped spaces inside the circumscribed volume. The cross section adds connection doorways from one space to the other. These elements can easily be read in the first drawings (1). Experienced architects do not need the second drawings to come to the same conclusions. Novices, though, might not see the wood for the trees and therefore not immediately recognise the spatial configuration. After seeing the second drawings, they will learn to 'read' these elements in the first drawings. The next drawings (3) are not obvious for experts nor novices. Based on the drawings above, it is just not possible to infer this information. The described cuts in the walls, floors and ceilings were observed during visits of the watermill. They were not present in any drawing, and if they had been, it is improbable that they would have triggered the same connection. This drawing actually adds information to the sections in (1) and is very helpful to understand the next design decision (4). Here a design intent is represented (addition). Because the building cannot be given back its historical function, the designers hope at least to reinstate a lively atmosphere – be it in a completely different sense. At the time this drawing is proposed, they do not know yet how to realise this intent physically. As described by Lawson (1997), their abstraction at this point prevents them from getting stuck in practical obstructions before they can develop a clear idea. The last drawing shows that, eventually, it was no problem finding a solution to realise this intent. Although this drawing is again less abstract, the use of colours (addition) to indicate the functions make the reader focus on the program, as suggested in section 2.

4.3 Case-based logic

Previous discussion shows how the proposed analysis and representation mechanisms can elicit understanding of design thoughts. The way in which information is reduced in this example can be enlightening for novices, who do not always recognize a clear logic when investigating (only) the design results. We believe that understanding the logic behind a design is more helpful for one's own design process than describing the formal / organisational / spatial / technical / … appearance of a project. Especially students tend to 'copy' the latter from examples without being aware of their meaning and/or their underlying motivation.[3]

We also believe that understanding the existing building is an essential condition to make a valuable re-design. We do not claim that drastic alterations to the existing situation are to be avoided, but they should be made in a conscious way.

Here the advantage of cases comes into play. Seeing how other designers interpret existing building information, what they consider as valuable and how it is treated, might be helpful to make value judgments in design. The real process information in this case – and in most architectural (re-)design cases – is not available, so we try to reconstruct it in some way. We propose to use reduction mechanisms, derived from real design processes, because we believe they can construct explanatory lines of thought (which eventually might be discussed with the architects to check their relevance). But even when it is not the exact logic of the real design process, offering a hypothetical logic can also be insightful. In fact, this is what happens on a larger scale in architectural theory all the time. When discussing a project, critics cannot – and probably should not – always check whether what they write is also what the architects think.

Due to space restrictions only one line of thought has been illustrated in this paper. The discussion indicates that it interferes with other lines, creating a more substantial logic for defining the spatial organisation of this building. Apart from designing the spatial organisation, the architects address many more issues, which in turn can be untwined in even more lines of thought. We think of the determination of the materials for the new wing and new elements, the composition of the reconstructed façades, the choice of building structure, the solution to the stability problem of the south façade, the circulation towards and around the building, the visual links with its surroundings, etc. Lawson talks about parallel lines of thought (1997). Several starting points are chosen from where design ideas are generated. At a certain point, these ideas collide – and then they have to be reconsidered – or they accord – and then they can strengthen each other. The next paragraph will briefly discuss how the re-design supporting tool allows for capturing more lines of thought and reveals the logic underlying the design process.

[3] This has also been acknowledged by Paul Van Aerschot, who states in an interview (Heylighen, 2000) "I don't blame them [the students] for adopting things, but for adopting them without question, without reflecting". It is exactly this reflection we try to trigger in this re-design tool.

5. RE-DESIGN SUPPORTING TOOL

So far we have explained how we can reveal the knowledge embedded in a line of thought. It is our aim to make this knowledge easily accessible to student and novice designers. Therefore we have chosen to develop an on-line computer tool[4]. We will only provide a short description of the tool here and focus on how it supports the framework discussed above. A more detailed description of the tool can be found in (Lindekens and Heylighen, 2004).

The tool can be called a case-based design (CBD) tool to the extent that it is conceived as a collection of cases. The aim is to provide more information than solely the design result represented by *raw material,* as e.g. in magazines. Except for storing general data about the building, each case therefore contains detailed information on a re-designed building or structure, the re-design process behind it, and the aims and strategies of the architect(s) involved. This information is organised along the lines of thought described above. These lines of thought are collected in a database and visualised through an interface as shown in the example given (*Figure 4*). Keywords are added to the information embedded in a line of thought. Apart from browsing through the case collection, users can thus conduct a directed search to select only these lines that are of interest to their present design task. This way they can compare how different architects tackled similar problems, and decide in a more informed way as to how they will act themselves. If they would like to see how a new case – which is not available in the case base yet – can be compared with those already present, they can use the techniques illustrated in other cases and apply them to their own case, after which they can submit the data to the case base.

6. DISCUSSION

The paper introduced a theory of transformation mechanisms and described how they can help designers. Next a real-world design phase indicates how these mechanisms are used in the design process. These observations formed the basis to exemplify how transformation mechanisms can be used in an analytical context opposed to the design context. Finally, which benefits and shortcomings can this proposal entail? At some point in the protocol the available pictures are investigated by the architect.

[4] The current prototype of this tool is based on DYNAMO, a Dynamic Architectural Memory Online, developed at the CADlab of the K.U.Leuven. We will investigate how the tool can be further developed within DYNAMO, and how DYNAMO and this re-design supporting tool may interact later on.

Goldschmidt considers this less advantageous than sketching, because images fade away rather quickly. This might be taken to suggest that taking pictures of a building does not serve much when nothing is done with them afterwards. Just looking at the pictures, without analysing them through sketching is likely to offer less design clues than when sketching is involved. Sketching enhances seeing things and reacting to them, which strengthens our belief in the relevance of this re-design tool.

At the same time, an argument against this tool might be that, 'seeing' an analysis made by others might not offer students the same learning experience as sketching the analysis themselves. Goldschmidt clearly endorses this concern by stating 'consulting self-generated displays is, for the most part, cognitively more economical than seeking useful information in other displays' (Goldschmidt, 1997). In view of this, the tool invites users to create self-generated displays, by adding new cases to the case base. On the other hand, the tool invites to use more hard-lined drawings, which according to Goldschmidt 'no longer have the same potentials for harbouring unexpected clues'. But this is also what we are looking for: during analysis it is our first aim to provide the user with the design intent of the designer in order to understand the exact meaning of the design. No real backtalk is needed here.

One needs to distinguish between drawings made for oneself, as in the protocols, and those made for sharing: the former are a working tool, should trigger new info, the latter should be unambiguous. To some extent, the analysis of a building can be considered a design too: first self-generated displays will be created, which are then formalised by means of the transformation processes described, before submission to the case base. It is this sequence that will provoke most of the learning skills.

Suwa argues that architects have remarkably longer dependency chunks - or, in our terms, lines of thought - than students, and these chunks or lines are more interrelated. This indicates that once architects shift their focus of attention, they think more deeply about the topic, what she explains by the ability of architects to " 'read-off' more different types of information from their sketches" (Suwa, 1997). It is exactly this ability the tool aims to support: showing what dimensions architects use in their design, implicitly suggesting students to try it out as well.

Developing different levels of detail in the analyses makes browsing the cases interesting for both students and experts. The idea of zooming in and out can be used for this purpose. That way novices can zoom in more when they need more explanation, experts have enough with a more general analysis. A working prototype of the proposed tool will enable us to test our findings with both novices and experts, hopefully advancing our project.

ACKNOWLEDGEMENTS

Jonas Lindekens is Research Fellow of the Fund for Scientific Research Flanders – FWO. Marijke Mollaert and Herman Neuckermans (K.U.Leuven) are supervisor respectively 2nd supervisor of this research.

We would like to thank Ann Heylighen for conducting the design experiments and for commenting this paper. We would also like to thank the designers for collaborating in the experiments and noA-architects for allowing the participatory observation of the competition design.

REFERENCES

Baker, G.H., 1989, *Design strategies in architecture: an approach to the analysis of form*, Van Nostrand Reinhold (International), London.

Ching, F.D.K., 1979, *Architecture: Form, Space and Order*, Van Nostrand Reinhold Co., London.

Goldschmidt, G., 1999, "The Backtalk of Self-Generated Sketches", in: Gero and Tversky (eds.) *Visual and Spatial Reasoning in Design, Cambridge, MA, Sydney, Australia, Key Centre of Design Computing and Cognition,* University of Sydney, p. 163-184.

Heylighen, A. and H. Neuckermans, 2000, "DYNAMO—Dynamic Architectural Memory Online", *Journal of Educational Technology and Society*, 3(2), p. 86-95.

Heylighen, A., 2000, In case of architectural design: Critique and praise of Case-Based Design in architecture, K.U.Leuven, Leuven.

Lawson, B., 1993, "Parallel lines of thought", *Languages of design,* 1(4), p. 357-366.

Lawson, B., and S. M. Loke, 1997, "Computers, words and pictures", *Design Studies,* 18(2), p. 171-184.

Leupen, B., C. Grafe, N. Körnig, M. Lampe, and P. de Zeeuw, 1997, *Design and analysis*, 010 Publishers, Rotterdam.

Lindekens, J., A. Heylighen and H. Neuckermans, 2003a, "Re-designerly ways of knowing, Towards a deeper understanding of architectural re-design", in: *Proceedings of the 5thEuropean Academy of Design Conference, Barcelona (Spain), April 28-30.* (in print)

Lindekens, J., A. Heylighen and H. Neuckermans, 2003b, "Understanding architectural redesign", in: Aouad and Ruddock (eds.), *Proceedings of the 3rd International Postgraduate Research Conference in the Built and Human Environment, ESAI - Lisbon (Portugal), April 3-4, 2003,* University of Salford (UK), 2003, p. 671-681.

Lindekens, J. and A. Heylighen, 2004, "Re-Using Re-Design Knowledge", in: *7th International Conference on Design & Decision Support Systems in Architecture and Urban Planning*, 2 - 5 July 2004 (in print).

Neuckermans, H., 1994, *Ontwerpmethodiek*, course text, K.U.Leuven, Leuven.

Schön, D.A., 1983, *The reflective practitioner: how professionals think in action*, Temple Smith, London.

Suwa, M. and B. Tversky, 1997, "What do architects and students perceive in their design sketches? A protocol analysis", *Design Studies*, 18(4), p.385-403.

Taura, T., T. Yoshimi and T. Ikai, 2002, "Study of gazing points in design situation - A proposal and practice of an analytical method based on the explanation of design activities", *Design Studies*, 23(2), p. 165-185.

Unwin, S., 2003, *Analysing Architecture*, Routledge, London.

On the Notion of Level in Architecture

Developments in Domain Theory

M.F.Th. Bax and H.M.G.J. Trum
Eindhoven University of Technology, The Netherlands

Keywords: Levels, Hierarchy, Architecture, Composition, Complexity, Control.

Abstract: The notion of Level (in a scale of Levels) is probably the most authentic notion in Architecture. Already in the work of Vitruvius the notion is implicitly present in the triad 'ordinatio – symmetria – eurythmia'. In more recent times, the notion always appears in relation with hierarchical organization as a means of control of quality. However used in drawings and in architectural discourse, the term lacks precision; there are many types of level like abstraction, specification, dependency, resolution levels etc., but no operational definition can be found as a notion that structures architectural objects and design processes simultaneously in a consistent way. Defining this notion of Level is the purpose of this paper. An example of application in an architectural decision-making process completes the paper.

1. INTRODUCTION

This paper aims at establishing a definition of the notion of Level, which 1) is dual in nature, i.e. it applies to both the articulation of objects as products of a process, and to the articulation of processes in which these objects are created; 2) is operational in the sense that it may articulate architectural concepts both as a notion and as an image, and 3) applies to research, design as well as to education processes.

Before going into the method used in order to arrive at a definition, it is appropriate to reflect on the nature and the general application of the notion of Level in architectural design, inspired by applications in some other fields. In general, the arrangement of an object in terms of levels entails positioning the object in a range of levels (a scale), as well as articulating the object itself in levels, based on particular properties of the object. These

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 279-292.

properties are denoted by the term ‘stratification’, a manifestation of a category of thinking of architects, scientists and philosophers.

The notion of level in an architectural object always functions in interplay with the notions of Phases en Criteria. Therefore the relevance of the definition of level we want to achieve can only be demonstrated after the moment that also these notions are dealt with as subjects of separate papers. We will treat the subject of levels as independently as possible, but because of the very nature of architecture, it is sheer impossible to avoid some overlap.

In science the notion of level is an indispensable tool to abstract and represent the complex reality into reduced and simplified models, which may, because of their predictive value, be deployed again in order to bring about a new reality. In political parlance for instance, the notion of level is an essential means to communicate with other people in all nuances, varying from coarse to meticulous, either in an open or mandatory way (open: by use of general terminology in order to attain agreement on legislation; mandatory: by use of specific terminology in order to enforce execution of legislation).

In Architecture the notion of level is clarified by the more familiar notion of Scale. This notion has a twofold disposition: numerical (Scale 1:100) and intuitive (Harmony).

As to make the first type operational, e.g. for designing, insight in the nature of the second type is needed. This second type of Scale is subject of the Proportion Principle, as known from Vitruvius and Van der Laan’s interpretation.

By means of the Vitruvian notions of ordinatio, eurythmia and symmetria (Morgan, 1960), the concept of proportion ingeniously conditions the accomplishment of a proper balance of dimensions within and between the parts by which the building is composed (Van der Laan, 1967). This is achieved by establishing an arrangement of super- and subordination between those parts and by adopting units of measurement by which the dimensions of the parts can be measured in a countable way. Application of this principle in Architecture, based on human perception, establishes scale in the sense of Harmony. The units of measurement again correspond to numerical scales that enable an orderly image and perception of the building. The sense of scale and the practical application of scale are brought together through perception and application of proper units of measurement.

The units and the scale, on which they can be represented, determine classes of elements that determine the content of spatial levels.

As any building, because of its super- and subordinated arrangements of parts, covers several classes of elements with their own units, several levels are simultaneously active in the perception of a building.

This simultaneous perceptible presence of several levels in an object beside the presence of an object in a scale of levels is denoted as the feature of stratification.

2. METHOD

In accordance with the paper's objective the notion of Level will be adequately defined through a number of steps:

1. Carrying out an analysis of linguistic aspects of the term Level.
2. Discerning different principles for identification of levels.
3. Discerning and defining types of levels.
4. Formulating a preliminary layered (working) definition of the Notion of Architectural Level.

Presenting an example of application.

3. ANALYSIS

A quick glance in a dictionary makes clear that 'Level' refers to 'Hierarchy' and 'Hierarchy' refers to 'Level'.

At first the term Hierarchy is considered. It is interesting to note that 'Hierarchy' refers to a principle of classification in a process of ordering as well as to the product of such a process.

Stratification or level-based ordering may be considered as 'hierarchy-as-product' (of a process of ordering, guided by a hierarchical principle).

The term 'Hierarchy' stems from the world of administration and government (of church and state). A scale of levels is the manifestation of a hierarchical principle of ordering, aimed at efficient control, policy-making and management of organizations. So, the leading principle in discerning levels must be found in a *principle of control*.

The scope of this principle should be interpreted widely and concerns both 'control' in the sense of governing institutions (leading, administering churches, states and other organizations), as well as controlling (commanding, regulating, having in hand, keeping in check) processes and projects in almost any field of human activity.

Etymologically the term Hierarchy refers to a principle of stratification, in which for instance decisions on a higher level are 'sanctified', i.e. inviolable and thereby unchangeable, irreversible and definitive for people and institutions on a lower level in the hierarchy.

Historically the term Hierarchy is used to indicate stratification primarily in social and political systems (of e.g. church and state) and secondarily to indicate and describe forms of corresponding stratification (according to societal conventions) in physical and other systems: perceived as being organized in an ascending or descending order.

Practically the term Hierarchy is used to distinguish between main issues and side issues, between things to which major or minor importance is attached, or matters that are considered to be of a higher of lower degree of permanence, e.g. things that to have be maintained or changed in a design process, depending on the situation at hand.

Hierarchical ordering can be found in both natural and artificial systems. A tree may be analyzed and described as a hierarchically organized system. On the highest level the tree is considered a complete overall object, on a lower level as a more detailed system of constituting elements: roots, trunk and branches and on an even lower level, as a system of many more smaller parts; leaves, flowers, fruits, root-hairs, bark, etc. These types of elements may again be considered subsystems, containing their own parts on a lower level, etc. In a hierarchical description the complete system is depicted on every level, but with a decreasing or increasing degree of generalization or specification between the levels. The contents of the levels can be strictly separated and the relations between the levels precisely determined. A company is hierarchically organized, however the number of levels varies, depending on the type of organization (flat or multi-layered).

Hierarchical organization occurs in sensory perceptible *concrete* systems as well as in mental *abstract* systems. An urban plan may be described in a similar way as the above example of the tree, every element being equally concrete, but a city may also be described on various levels of concreteness, as is usual in the presentation of a plan on different scales. On a certain level a city may be described as a composition of blocks and on a more detailed level as a set of dwellings that together constitute blocks again. Moreover, on every level of concreteness (or abstraction) a hierarchical ordering is conceivable, the elements of which having the same degree of concreteness.

As to *designing* an effective organization, both types of hierarchy are relevant: the principles governing the relations between *elements* within a certain level of concreteness are also valid for the relation between *sets of elements* on different levels of concreteness. All these principles may be classified under the common denominator of Control.

This analysis leads to basic assumptions about levels on which the following distinguishing characteristics of the next section are based, namely: Levels can be distinguished by themselves and from each other; multiple levels can exist simultaneously; interactions (of any kind) can be defined between levels; it is possible to transfer information from one level to another.

4. PRINCIPLES OF LEVEL IDENTIFICATION

In literature, through observation of architectural practice, education and by authors' reflection six principles were found that are related and directed to the identification of levels in architectural objects. In order to structure the principles the object may be considered in first instance an Object as Form ('Percept' or 'Gestalt'), in second instance an Object as Product (Form & Process) and in third instance an Object as Construct (Product & Criteria), which hierarchically and cumulatively ordered threesome correspond with the three modalities of self-regulating organizations. In this section these three object modalities enable an ordering of the identification principles that will be applied for distinguishing types of Level.

4.1 Principle of Discreteness

Identification of levels by discerning elements, corresponding with perceptible qualitative articulations of the Object (as a Form).

Before going further into the nature and issues of hierarchically ordered levels, it should be noticed that the hierarchical type of levels occurs in a concatenated chain (or *scale*) of levels. This is a specification of the general (Webster) notion of level, which also includes linear scales on which levels only differ from each other quantitatively and gradually. This study however, is focused on types of levels with qualitative differences, corresponding with a distinction in classes and types.

The essence of this kind of series is (like in the example of the tree), that each level has its own quality and thus its own identity. So, the transition from one level to another entails a *quality shift*. This notion of level is comparable with the way the notion is interpreted in dialectical discourses, where a conflict on a certain level is made solvable by conversion to a higher level of abstraction. For instance, if the dimensions of a girder, in order to be strong enough, have to be so large that the height of the remaining underpass is too small, this conflict may be solved on a higher, more abstract level, where the problem of spanning may also be solved by means of an arch instead of a girder.

In a mental and artificial way, levels enable the introduction of discontinuities in complex entities, e.g. in a (whether or not imaginary) space, being continuous by nature. Herewith the space can be made *discrete* for the purpose of observation, perception and intervention.

In a concatenation of levels of this type, a (situation on a) given level presumes the presence and working of a (situation on a) lower level. The chain of levels is principally unlimited and open at both sides; for that reason they are indicated as Open Hierarchical Systems (Koestler, 1976).

For examining a certain level and to gain insight in the working of the complete system of levels, it is possible and useful to constrain a study field by selecting repeatedly three adjacent levels and to focus on the middle one, positioned between its two neighboring levels.

4.2 Principle of Recursivity

Identification of levels by discerning similar *patterns* of elements in situations returning on other levels of the Object (as a Form).

In many hierarchically ordered systems similar patterns of ordering may be perceived on every level. For example, in the nerves of each leaf the branching pattern of the tree recurs, and in the patio-house the repeated pattern of the patio on an urban level recurs as a piazza or forum. In ancient China the pattern of walled territories was applied from the simple hedged house, via the walled (Forbidden) city in the city, up till the circumvallated town in the country and even on the level of the walled country in the surrounding world.

Application of such a principle contributes to coherence and harmony in the whole system. It is also a means to define different levels against each other; occurrence of Recursivity is an initial indication that we are dealing with several levels.

4.3 Principle of Restricted Complexity

Identification of levels by discerning situations (composed of elements) with a similar *numerical complexity* on each particular level of the Object (as a Product of a process).

A noteworthy feature of hierarchically ordered systems is that *situations* occurring on different levels are characterized by a similar degree of complexity (measured in numbers and quantities of types of elements of which they are composed). Levels in musical or architectural systems for instance, may be characterized by an equal number of elements on each level: 3, 7 (octave) or 12 (they often are 'holy' numbers).

This characteristic is especially important in artificial systems (being the result of human decisions) because it is sensible to assume that the capability of persons and institutions to make decisions is comparable on every level. For this reason it is important that the number of (types of) elements, that has to be decided about (with regard to application, position, dimension, etc.), should be comparably large on each level.

However, this does not alter the fact that the complexities of the *elements* on each level may vary considerably. The total complexity, i.e. that of the situation in addition to that of the constituting elements is not constant of course, but here a regulated complexity exists.

This principle implies that for aesthetic (perception-based) reasons the numbers of elements per level are subject to limitation: too many elements result in chaos, too small an amount in deprivation.

This principle is a powerful means for identifying levels and for stratifying a complex entity into a hierarchically organized system of levels.

According to parlance every higher level in a hierarchical organization is characterized by an increased complexity of its *elements*.

This particular principle is closely related to the principle of Economy of Thought, and is applied on efficiency of mental processes.

4.4 Principle of Double Role

Identification of levels by discerning *mechanisms* that cause the transition of a situation into an element (of a lower level), and vice-versa, of the Object (as a Product of a process).

If elements of a higher level are more complex than elements of a lower level, then (in this vision), this higher degree of complexity can only be traced back to a larger number of (types of) parts (elements again) that constitute the complex element.

The complex *element* itself may be considered a *situation*, in which again elements occur (as its constituting parts) that already have obtained their specific positions, dimensions, pattern of mutual relations, etc. So, in this view one and the same entity fulfils a double role: both as an element and as a situation. Obviously two kinds of elements exist: namely the kind of elements that may become a situation and the kind of elements of which the situation is put together. In order to discern these two kinds of elements, they may be assigned to two adjacent levels in a hierarchical order. The kind of element that becomes a situation belongs to a higher level and the kind of element that constitutes the situation belongs to a lower level. Thus, this principle enables discrimination between hierarchically ordered levels, as known from e.g. a means-and-ends hierarchy.

For example, a decision-maker on a high level should be well aware that decisions about elements belonging to his ‘own’ level not only determine ‘his’ particular elements, but that simultaneously and implicitly decisions are made that influence and condition situations on a lower level.

4.5 Principle of Open Form

Identification of levels by discerning (structural) *rules* in elements, which allow for the generation of equivalent functional variants in the context of a lower level, and vice-versa, of the Object (as a mental Construct resulting from judgement).

If an element of a higher level becomes a situation on a lower level (according to principle 4.4), and if furthermore only *one* situation is involved, then this situation is already implied by the way the hierarchical system is arranged. In the most extreme case, a decided situation on the highest level completely determines (by implication) all situations on all lower levels. Such cases concern closed systems, for which a distinction in levels is merely useful for a systematic description of the object's stratified anatomy as the existing result of a completed design process.

In contrast, the principle of Open Form is particularly relevant for e.g. phased and participatory decision-making processes, but it is also of general importance for gaining insight in the nature of hierarchical systems in general. In participatory decision-making processes levels should be determined in such a way, that they not only correspond with a hierarchical ordering of spatial/material objects, but also with a hierarchical ordering of participants (individuals and groups) in corresponding process phases and in the functional criteria applied by them.

In decision-making processes, deciding participants and parties on every level and in every phase need some freedom in order to be able to contribute to the process in accordance with their own interests and needs. The transition from a higher to a lower level therefore should allow full play and provide elbowroom for parties appearing on that particular level in that particular process phase. In terms of hierarchically ordered levels this means that an element (of a situation already decided about on a higher level) should not generate just one, but several situations on the next lower level, which still may be elaborated in different ways, i.e. in *variant solutions*. Though the element of the higher level already has obtained some form, it should be an unfinished, open type of form.

As this form (just like the closed form) is intended to ensure a certain quality, yet to be worked out in future decisions, the open form inherently consists of two groups of properties: *material* properties (its dimensions, position, material, etc.) and *regulative* properties. The regulating properties set out the perspective for possible future developments of the material properties, however without determining their final form. At the beginning of a process the regulative properties of the element will be substantial, at the end of the process the material properties will be dominant, but there will always remain a regulative constituent, which often implicitly determines its usability.

The property of elements and situations by which they contain rules for their use and development is denoted (in conformance with usage in Structuralism) as a transformational or *structural property* (Bax, 1976).

4.6 Principle of Controlled Change

Identification of levels by discerning incremental purposeful (functional) *transformations* of situations from a high-level Model to a low-level Plan, or vice-versa, of the Object (as a mental Construct resulting from judgement).

The term Hierarchy may still be used in its etymological meaning: a system is hierarchic if changes of subjects on lower levels do not affect subjects on higher levels, i.e. they remain untouched. However, the opposite is not the case: changes of elements on higher levels imply changes of elements on lower levels.

In a hierarchically ordered system decisions are carried out in a one-way direction: top down (to this the term owed its pejorative connotation in the seventies). Yet, exchange of information (in architectural modeling and planning processes) requires and is carried out in two-way traffic (for this reason the more neutral term Stratification is often preferred to Hierarchy). Such an asymmetric interdependence of transformations and controlled change of elements is typical for the arrangement of elements in levels in a hierarchically ordered system. An ordered arrangement of parts in levels of an ordered whole is that of an *organization*.

These properties concerning controlled elaboration of transformations within a level-ordered system are characteristics of a principle underlying any arrangement in levels: the principle of controlled change.

5. TYPES OF LEVEL

All above-mentioned principles contribute to the formulation of the notion of level. It appears to be possible to determine types of level by means of the number of principles being relevant for those specific types. The principles in the preceding treatise are arranged in a sequence of increasing complexity, in which each next principle is based on and contains the previous one.

Three types of level are discerned: firstly by applying principles 1 and 2 (Form-oriented) Levels of Composition, then 1 through 4 (Product-oriented)

Levels of Complexity and after that 1 through 6 (Construct-oriented) Levels of Control, in an order of increasing complexity.

Moreover, it may be informative to point out that the principles appear in pairs: the even-numbered principles perform an instrumental role, whereas the odd-numbered are more involved with content.

Levels of Composition are a type of level in which an ‘Object as Form’ may be characterized by elements and situations based on principles of Discreteness (qualitative differences) and Recursivity (similar patterns), representing the composition of the object. In practice Levels of Composition are also denoted as Aggregation levels, Containment levels, Distribution levels and Framework levels. This type of level plays a role in spontaneous observation of a situation at the beginning of a process, yet without the subject’s attention for the way this situation arose. The subject perceives the entity as (outward) *Form.*

Levels of Complexity are a type of level in which an ‘Object as Product’ may be characterized by elements and situations according to Levels of Composition and additionally according to principles of Restrained Complexity and Double Role (the mechanism determining the transition of elements to situations on a lower level and vice versa). In practice, levels of Complexity are also indicated as levels of Abstraction and Concreteness, levels of Specification and Generalization, levels of Scale and Resolution, etc. This kind of level plays a role in the subject’s action aimed at changing the situation, i.e. to adapt it to the requirements of the environment. The subject now perceives the entity as *Product* of a process

Levels of Control are a type of level in which an ‘Object as Construct’ may be characterized by elements and situations according to Levels of Complexity and additionally according to principles of Open Form (with structural rules) and Controlled Change (gradual transformation). In practice Levels of Control are decision-making levels. The Form and the Product are now perceived as *Construct*, i.e. a product in relation with the reasons, arguments and considerations that eventually resulted in the origination of the product. The term Construct also matches the Anglo-Saxon ‘Control’ which includes more than merely supervision of the process. The term accommodates so-called levels of Subsidiarity, tooled up according to the homonymous principle and as applied in ecclesiastic law and more recently in the European administration. According to this principle decision-making is assigned to the lowest level of competence.

The distinguished types of level again are hierarchically related. The level of Control presupposes the presence and operationality of the level of Complexity, which again presupposes the presence of the level of Composition. Moreover, as a consequence of the way they are defined, the level of Control comprises both lower levels. Thus, the levels themselves may be ordered in levels. Therefore the level of Control is the highest level in the hierarchy of levels, in spite of the fact that it is positioned on the lowest row in Figure 1.

Principles	Levels	Reference
1. Discreteness	Level of Composition	Form-oriented
2. Recursivity		
3. Restricted Complexity	Level of Complexity	Product-oriented
4. Double Role		
5. Open Form	Level of Control	Construct-oriented
6. Controlled Change		

Figure 1. Hierarchy of Levels

6. DEFINITION OF ARCHITECTURAL LEVEL

For an advanced study of levels and hierarchy in the field of Architecture only the highest type of level, the Level of Control, is relevant. Herewith no information gets lost, because this level comprises the contents of both lower levels. Qua name and contents this type of level also joins with the leading principle of Hierarchy, viz. Control en thereby with the (work) definition of Architecture: Control of Complexity by means of Composition (Form). For this reason Levels of Control are also denoted Architectural Levels.

Though the other types of level contribute to the definition of the Level of Control or Architectural Level, they are especially intended to position the commonly used kinds of level, in relation to the ones discussed in this study. For example, the widely used Level of Scale is a level of Complexity.

An Architectural Level is a particular articulation of an Object (next to other articulations like Phases and Criteria) on a certain position in a Scale of Levels, which structures the organization of the Object in fields of perception, conceptualization and decision-making: every level being defined initially by classes of spatial elements and situations of the type of Level of Composition, secondly by the type of Level of Complexity and thirdly by the type of Level of Control.

These three instances of the Notion of Level correspond with the degrees of profundity of consideration of an architectural object, in which the first instance is a condition for the second, etc.

The instances correspond with the phases of an imaginary process. The first one concerns an open-minded observation, the second one enables closer investigation of the complexity of the object that came into being as an answer to a principally discordant request from the environment and that may be explained by the third one as the result of application of (structural) rules in a modeling and planning process.

7. APPLICATION

The powerful interdisciplinary properties of the notion of Level (of Control) become apparent by its application in structuring objects, which are subject of disciplinary fields and have to be integrated in a multi-disciplinary project. Architectural objects, even the simplest ones, are subject of such a complex type of process. An Architectural Object may be considered the result of a synthesis of three types of objects, which may be distinguished similarly as the types of Levels, viz. as Form, Product and Construct (oriented) Objects, and are articulated in Levels as any other object. In order to integrate these three types of object it is necessary to tune the levels of those objects to one another. This means that the levels of the Form Object also articulate levels in the parties and participants who produce a Product Object in the variety of phases of a process and moreover articulate levels in the rules and criteria that lead to a Construct Object.

In the sixties of the last century John Habraken (Habraken, 1962, 1972) and the Foundation for Architectural Research (SAR) introduced a new kind of architectural object in the field of social housing that was denoted as a Support. Although theoretical work, as presented in this paper, resulted 40 years after the introduction of the notion of Supports, it appears to fit perfectly well. The innovative feature of the Support object is that Form, Product and Construct aspects are conceptually integrated in a complex, but nevertheless easy manageable concept in order to make complex processes of decision-making simpler and open for participation by all involved parties.

A Support is a type of Object of such *general* characteristics that it opens to *specification* by dwellers. Beside that, the object is positioned on a scale of Levels, in which a Support is an *element* of a higher level Tissue, and a *situation* open for Infill on a lower level. As a consequence a Support is an object with *structural properties* that allows for a controlled transformation according Principle 5 and 6 of Section 4. From a conceptual point of view a support is a Construct: a Support has Form properties (zones and margins and sectors) and beside that Product properties in so far it addresses specific parties in the design process, and more specifically it contains rules, which control the generation of variant plans.

A Support functions not only in a mediating role between the various Levels of an Object as Construct, but functions also in a mediating role between the Object as Form and the Object as Product. The Object as Form is represented in its Levels of Block Building and House Building, and the Object as Product is represented in its Levels of Block Dwellers and House Dwellers.

This twofold mediating position of a Support Plan is rendered below in Fig. 2. On a higher and lower level a Tissue Plan and an Infill Plan are distinguished respectively. The schema of Fig. 2 is an adapted version of a scheme developed by John Carp (Carp, 1979).

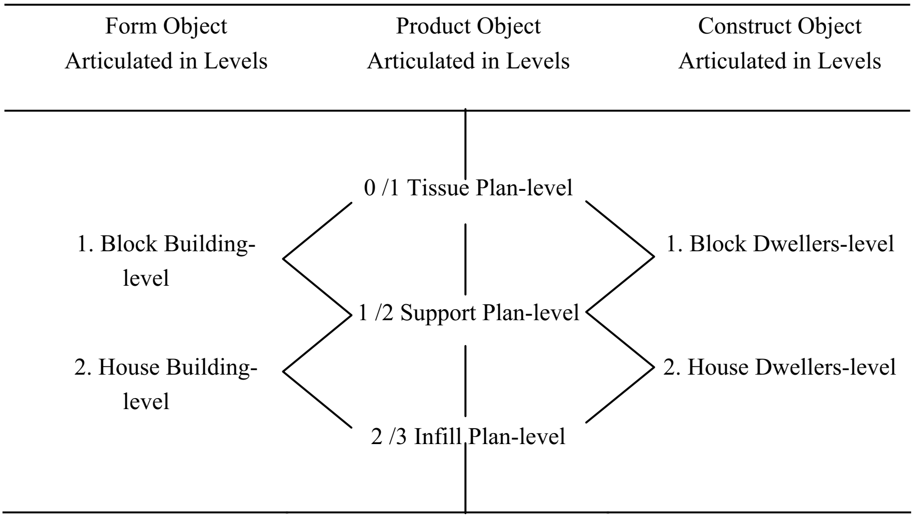


*Figure 2.*Levels of decision making in a Support-structured process

The scheme represents principles for decision-making in the field of spatial organisation. The scheme may be extended to higher levels of decision making – even to a national level – and to lower levels, but always following the same pattern as indicated by the arrows in the scheme. In the process of tuning of the various Levels of Form, Product and Construct objects, the notion of Support appeared to be a crucial instrument.

A Construct object, in general, is a determining factor for the demarcation and articulation of a territory, because, by its hierarchical and cumulative position, it integrates and harmonises the Form- and Product objects it encompasses. Such a territory is denoted a Domain.

A Domain is a complex articulation of an object that regards the object as Form, Product and Construct. In a more general terminology a Domain regards the Material, Procedural and Organic capacities of an object.

A Support is a first specific example of the application of the general notion of Domain, a subject that inspired the theoretical work of the authors since a long period of time. The study into Domains as the most complex units for the articulation of architectural objects and related process in a project are subject of so-called Domain Theory. This paper is a presentation of developments in this theory; hence the sub-title of this paper.

The paper fits in a series of papers, which aim at the establishment of a connection between Domain Theory and the pragmatistic philosophy of Ch. S. Peirce, one of them previously presented in the DDSS 2000 conference (Bax, Trum, et al., 2000).

Beside this example of application, numerous examples may be given. They are all specifically situated in the field of multidisciplinary design in the form of concurrent design, integral design, participatory design, or whatever other name may be given to this kind of activity. Designing processes is the subject of a separate discipline, of which the Postgraduate Designers' Course 'Architectural Design Management Systems' at TUE is a good example.

An annotated document on the topic of this paper in Dutch language (30 pages) is available via h.m.g.j.trum@bwk.tue.nl

REFERENCES

Bax, M.F.Th., 1976, *Meten met twee maten: de ontwikkeling van een maatstelsel als kader voor een besluitvormingsproces op het gebied van ruimtelijk ordenen*, Proefschrift, Technische Hogeschool Eindhoven, 290 p.

Bax, M.F.Th., H.M.G.J. Trum, and D. Nauta jr, 2000, Implications of the philosophy of Ch.S. Peirce for interdisciplinary design: developments in domain theory, *Design and Decision Support Systems in Architecture*, Proc. 5th Int. DDSS Conf., Nijkerk, The Netherlands, pp.25-46.

Carp, J.C., 1979, *Inspraak en Zeggenschap: een preadvies over ruimtelijke ordening*, Stichting Architecten Research, Eindhoven, 214 p.

Habraken, N.J., *De Dragers en de Mensen*, Scheltema en Holkema, Amsterdam, 1962 English edition: Habraken, N.J., 1972, *Supports: an alternative to mass housing*, The Architectural Press, London, and Praeger, New York.

Koestler A., 1976, *The ghost in the machine*, Hutchinson, London, 384 p.

Laan, Dom H. Van der, 1967, *Het Plastische Getal*, E.J. Brill, Leiden, 142p.

Morgan, M. H., 1960, *Vitruvius: The ten books on architecture*, Dover, New York, 331 p.

Supporting Design Learning with Design Puzzles

Some Observations of On-line Learning with Design Puzzles

Teng-Wen Chang
National Chiao Tung University, Taiwan

Keywords: Design Puzzles, Design Collage, Puzzle-Making, Andragogy, Game Play

Abstract: The design process is a puzzle-solving process. Two groups of researches that share many similarities with Puzzle-solving design process are the process of game-playing and playful learning. The main argument is using the “playing” characteristics to amplify and explore the learning process, furthermore the design process. In addition, puzzles imply playful exploration that utilizes the characteristics of “playing a game” as “solving a puzzle”. Puzzle making and puzzle solving provides an incremental exploration mechanism that is more intuitive for design learning. For understanding and realizing puzzles in design learning, this research is divided into two stages of researches—manual design puzzles and interactive design puzzles. By analysing the outcome from manual design puzzles, this research proposes a framework called (interactive) “design puzzles”. The conceptual and implementation framework of this view of design is elaborated in this paper as well as a particular design puzzle called *puzzle collage* is described as the realization of design puzzles.

1. INTRODUCTION

The design process is a complex human behaviour. One common practice is to apply a metaphor to explain such a complex behaviour. It is called a *model of design*. Two well-known models of design are *design-as-search* (Woodbury, 1991) and puzzle-making (Archea, 1987). By decomposing design problems into a set of hierarchical sub-problems, design as search models design process as an iterative process of searching for design solutions in related to the design problems. While design-as-search shows an effective mechanism, puzzle-making argues that designers cannot solve a

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 293-307.

design problem unless they can define the design problems within the solving process—design process is about making the puzzles rather than solving the puzzles. Design as puzzle-making depicts different aspects of design (as puzzles). These two models of design are compliment to each other in a way to frame a design process (puzzle solving and puzzle-making). Instead of emphasizing on either puzzle making or puzzle solving, this research intends to develop a design supporting system that integrates both concepts as a model of design.

Two groups of researches that share many similarities with Puzzle-solving design process are the process of game-playing and playful learning. The main argument for the process of playful learning is by using the playing characteristics to amplify and explore the learning process, furthermore the design process. In addition, puzzles imply playful exploration that utilizes the characteristics of "playing a game" as "solving a puzzle". Therefore, within the scope of the game-playing researches, people tend to create a learning platform entirely within game environment such as video games or networked 3D virtual game environment.

Briefly speaking, solving puzzles requires a clear and deterministic goal, and also permits the possibility for creative and alternative searches (Akin and Akin, 1998). This allows users to explore the alternative solutions and be motivated in their learning process. Therefore, another group of researches is online design learning. A model of learning is called *andragogy* that provides a conceptual model for describing how to motivate the learners in various self-constructive situations. Three groups of relevant researches show different concepts of playing puzzles in the interest of this paper. The relations are shown in Figure 1.

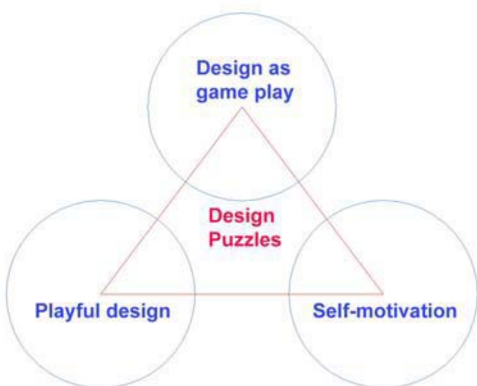


Figure 1. The relation between design as game play, playful design, and self-motivation

Insights of each group of researches in relation to this research will be described more in details in the following sections.

1.1 Design as a playful learning process

Through *playing*, one of metaphors for design learning, this specific group of researches implies technical and methodological refinement over a standard design exploration process. Such as (Radford, 1997) provides a simple mechanism for exploring the form-making process. Further developed by (Woodbury, Shannon, et al., 2001; Woodbury, Wyeld, et al., 2001), while using the name of "game", these researches tend to focus on the conceptual level of play (Klugman and Smilansky, 1990). Every game specified in these papers is using the concept of playing games at its conceptual model level by applying the metaphor of "playing games" on either teaching design or modelling design learning. The exploration of design-as-search is transmitted into a playful experience. While restricting the transmission between different games, the characteristics of playful experience is unleashed that paves the route towards a puzzle-making and puzzle-solving experience.

What we learn from this group of research is its capability for exploring the design concepts following the characteristics of playful learning. Such will allow designer to think more freely and creatively, meanwhile, the puzzle-exploration process might invoke.

1.2 Using games as a design supporting tools

Design games are realized in several instances such as (Chien, 2002; Woodbury, Wyeld, et al., 2001). The results are promising and effective. To their extent, this research group goes even further into analogising the real design experiences of game playing. Such as (Lee and Wai, 2003) using game engine as a development environment for cooperative design; and (Lehtinen, 2002) using 3D video game environment for teaching and learning design. Another effective approach is to use gaming environment for design presentation as well as interaction (Sallkachat and Choutgrajank, 2003). Among those, one thing in common is that "game" is taken not just conceptually but also literally. By putting design into a well-defined game environment, the limitation as well as its significances of games can apply onto design problems in our interests.

As for the purpose of puzzles, this group of researches provides an important insight that games all have their own strength and features that can be mapped into our puzzle systems. In this case, studying the characteristics of puzzles will give us the benefits to stress the concept of puzzles back into design. Further, our system might provide the features of puzzles that all people can be familiar without re-inventing the wheel.

However, the most important issues for design learning—self-motivation and exploration strategy are hard to achieve within game-only context. We need to gain more insights from learning theory in the next group of research.

1.3 Learning design with self-motivation

From Design learning to a model of learning, we look at a group of theory that can provide significant insights for realizing our puzzle-exploration system. There are many learning theories, especially online learning, that will provide both cognitive level and computational mechanism of design learning. However, for the nature of puzzle exploration process, we tend to search for a learning theory that will provide self-conscious or motivation for interactive guidance of exploration process.

In stead of the learning theory that is based on observation of children learning, an adult learning theory called *andragogy* (Knowles, 1990) focuses on motivating adult learning. Motivation ends up is an essential device for guiding the interactive behaviours needed for puzzle exploration process.

While adjusting andragogy into design learning, (Chang, 2002) outlines the conceptual learning framework for our learning model. Andragogy enforces the adult learning process with six characteristics—(1) The need to know, (2) Learner self-concept, (3) Role of learners' experience, (4) Readiness to learn, (5) Orientation to learning and (6) Motivation. Each characteristic represents certain aspect of learning with specific design constraints, thus, the puzzle-solving and puzzle-making processes in this paper.

Main strength of andragogy is to enforce the concept of self-directed learning. To say it explicitly, andragogy provides a puzzle-exploration mechanism that presents an environment for motivating designers to learn and design from their peers as well as themselves.

1.4 Towards a puzzle-base learning environment

By looking at playing games at metaphorical or system levels, playful learning environment shows some important insights for the development of our design supporting system. Therefore, we will provide conceptual structures of puzzle-as-metaphor as well as the analysis of puzzle games as our methodology towards the implementation. By integrating the concepts of andragogy into puzzle-based design learning system, the self-constructive process is then mapped as a puzzle-making process. Each individual interaction provides enough information for guiding the exploration as well as creating new puzzles. Furthermore, the motivation for solving a puzzle should provide the mechanism for the components of puzzles.

The development of design puzzles is divided into two stages and described in the following sections. First stage: manual design puzzles is to describe the accomplishment of development in WALE and its experimental outcomes for manual design puzzles. Stage two: interactive design puzzles is the further development built on the top of stage one. A workable and interactive design puzzles system is then specified and implemented based on these two stages.

2. STAGE 1: MANUAL DESIGN PUZZLES

Further understanding on the behaviours of design puzzles and learning behaviours, we conduct a preliminary experiment: manual design puzzles, using an acting role-interplayed system (*WALE*) (Chang and Huang, 2001; Chang and Huang, 2002). WALE (Web-based Architectural Learning Environment), is generated through a CAD subject in a positive virtual learning space. WALE is based on a game-playing learning environment for students to interact motivated and to evolve the design potential of individual (seen Figure 2). The study of WALE is facilitated with CAD tools and developed to help participants to explore possible design alternatives by acting multi-roles in the process of learning design. Several useful characteristics of digital architectural subjects on web are learned from various experiences (Radford, 1997; Woodbury and Chang, 1997). There are five primary characteristics of the web-based learning aspects are concluded that proposed acting role model of WALE are built on top of. They are 1) Asynchrony; 2) Dynamic Interaction/immediate feedback; 3) Multi-role interplay, including One-one, one-many, many-one and many-many relationships; 4) Recording/ history; 5) Game playing.

Despite that synchrony is a useful feature of web; asynchrony characteristic of design behaviour provides a more flexible design outcomes by shifting the design ideas input on web. Web offers an enormous virtual space where information is accessed and exchanged freely and fast. Thus, the features of dynamic interaction/immediate feedback are derived. The concept of multi-role proposed for students is to offer an opportunity of acting different proposition to extend their perspective views of design on web where anything is possible. The features of recording the design process (Huang, Chang, et al., 2000) and playing as a reflective action are two crucial aspects toward web-based learning.

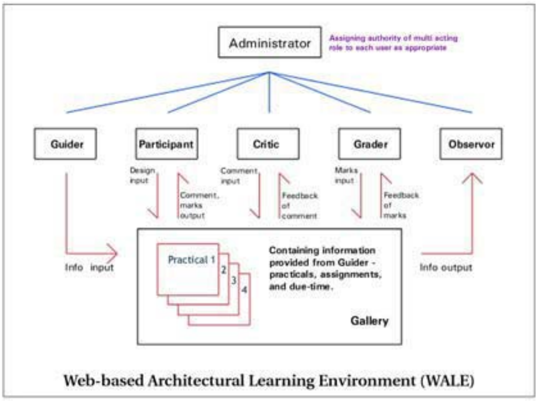


Figure 2. The role-interplay process of WALE

2.1 The experiment

On the top of WALE, we have developed the first prototype of design puzzles manually as a pilot study for understanding the behaviours of puzzle making and puzzle solving on the behalf of digital media learning. The duration of the experiment is three semesters and with a group of 40-50 participants. The backgrounds of participants are mostly artists with design training. Only few have computer experiences before participating this experiment. Therefore, the objective of design puzzles at this stage is to guide and motivate participants with the digital media. In addition, each puzzle is divided into two parts: a puzzle with design media concept and a puzzle-solution evaluation. As the nature of design, the evaluation of puzzle solution has to somehow depend on the critic (the human) involvement. Three evaluation criteria applied to the puzzle are: design, technology and skill. Different criterion represents different aspects of puzzles in a particular design goal related to the learning motivation. Several online techniques such as page-hits, preferences and grading are used for quantifying the evaluation outcome.

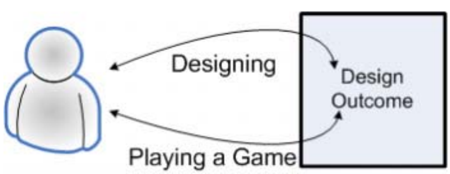


Figure 3. Playing a game and designing a design all have similar behaviour

The experimental steps are 1) using a more conventional method of curriculum at the first semester as an introduction to digital technology. 2) For second semester, WALE is used and applying with our design puzzles for a set of predefined goals within the acting role framework. 3) For third semester, the curriculum is changed to be professional skill oriented. Consequently, design puzzles developed for the third semester will require a more pedagogical approach along with andragogical methodology proposed in this research.

2.2 Some observations and limitations

As a teaching outcome, some design puzzles are successful for some participants but fail in others. With comparison of the puzzle-solution proposed by participants, we gain some observations and drawbacks of this experiment. There are five main limitations of this version of design puzzles such as the implementation platform, the domain knowledge of puzzle, the duration of experiment time, the number of participants and the will of participants.

As WALE is developed as a general Web-based learning environment, the puzzle-making and puzzle-solving facilities have to move around by using other mechanism such as digital media exploration. This is to say that puzzle in this stage is represented as a metaphorical experience rather than implementing the mechanism of design puzzles. Furthermore, solving a puzzle requires some knowledge of that particular puzzle itself. Thus a general mechanism for designing puzzles might not be feasible and useful. Also, as such experiment needs a long time to expose these puzzles for participants and allow participants to refine them. The duration time of puzzle exploration becomes more critical and inefficient in some cases. In addition, while solving puzzle participants need to either know the context or to understand the puzzle goal with instincts with time limitation. Consequently, a large number of participants decrease the communication time among participants as well as their will in solving the puzzles. As well, in our experiment, the process of puzzle solving and making is only invoked by the personal communication among participants that cannot be shared with others.

As the first pilot study, the puzzles are intended to be simpler and the hints for solving the puzzles have to be constructive and effective (Figure 4). This ends up to be an important factor for a successful puzzle. Another observation is the media. While using media as a representation, each successful puzzle should represent certain view of digital media. Several instances show that an interesting or inspiring view of digital media will motivate the participants to explore further on their own learning. The duration time of each puzzle solving process is set for about 10 hours. The minimum is 3 hour when average time is about 8-9 hours. When an interesting and challenging puzzle has been explained, participants are more willing to spend time to explore the possibility to solve the problem than an obvious solution.

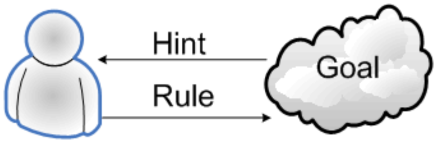


Figure 4. Relation between hints and rules

Another insight is that more successful puzzles have more inefficient or "fun" rules that will allow many alternatives or exploration to be made. For example, a design puzzle is about topological relations among geometry. The hint is *push*/*pull*. The goal is a space without space. Participants are getting into it right after using a few steps of commands instruction. Another important observation is that puzzle making provides an incremental exploration mechanism that is more intuitive for designers during the learning process. While learning affects are varied, the outcome of applying puzzle as a metaphor shows its strength in exploring and refining possibilities. The process of doing such exploration is showing promising results that motivate a further development of design puzzles in this research.

3. STAGE TWO: INTERACTIVE DESIGN PUZZLES

With the observations gained from stage one, manual design puzzle experiment, we then further specify a design supporting system by using puzzle-exploration metaphors. We called it (interactive) "design puzzles". According to the nature of learning behaviour and rules of puzzles, the design stage in this research is taken place in an early phrase of design. With design puzzles, learning design and making design can be modelled and studied with a specific view of design. Consequently, by learning from puzzle-games, several components and their mechanism are described in the following sections as well as the characteristics of design puzzles.

3.1 Puzzle Games

Puzzle-games (Bates, 2002) are a special class of games that will require player's logical reasoning capability to overcome the obstacle (the puzzle) in order to complete the game. Such as "finding a missing information piece" will require players to know the background of situation and reasoning it with the hints in order to solve the puzzle. According to (Bates, 2002), there are 19 types of puzzles such as mazes, criddles and cryptogram/word puzzles. In addition, such as "building puzzles" that players need to build the "missing pieces" in order to solve the puzzles can be applied on design domain. Each individual puzzle has its strength and twists that will provide some informative hints for players as well as traps. Players need to use the given hints with a set of rules in order to solve the puzzles within some particular situation or environment.

In mapping to design problems, especially puzzle-exploration, we identified six types of design puzzles according to the 19 types of puzzle games. There are 1) the geometric composition puzzles that compose 2D or 3D geometric components under the interaction with environment such as 3D puzzles; 2) design collage that using images as a symbolical meaning of design and the collage represents the outcome of design process; 3) information puzzles such as fractals or design space navigation; 4) role-play simulation with users playing different roles within different timeframe to simulate the puzzle-exploration process; 5) pattern recognition (jigsaw puzzle) using visual similarity and pattern recognition represents a simple but power design puzzles; 6) maze that can be viewed as a design space navigation process.

Each puzzle represents different design representation and aspect of puzzle exploration—puzzle-solving and puzzle-making process. To certain extend, several characteristics among those puzzle games are unleashed that will be described in the next section.

3.2 Characteristics of Design Puzzles

While studying online learning in design process, this research is applying the metaphor of "puzzle" from games and other resources onto developing an ongoing design supporting system called the "Design Puzzles". Design puzzles while supporting the self-constructive learning pattern of design is an essential device for design by exploring the possibility of design rules. By exploring design with modifying the goals as well as generating new design with generative rules, a puzzle exploration process can then be supported by the machinery we have developed. This goes to keywords exploration or re-definition as image collage (shown in the example). One common characteristic of above is to keep player motivated and focus with intuitive components. With the observation of puzzle game as well as preliminary experiment at WALE, the characteristics shared among them are elaborated further as followed:

1. A goal has to be clear but indirect. The goal of a good puzzle cannot be a binary choice and requires a bit of mind-twist. However, puzzle goals have to be clearly addressed and relevant to the given hints.
2. Puzzle exploration has to be self-constructive but with the potential to be shared by others online. Even puzzles are possible to be explored or solved by the cooperative work of multiple users; the self-constructive puzzle exploration is the key for learning design.
3. Hints can only provide partial information for the environment of puzzles but is enough for further exploration. Hints are the triggers for creating alternatives and decision strategic control.
4. Hints have to be intuitive and direct for shorten the duration time of puzzle solving process.
5. Rules have to be simple and manipulated directly with several entities described in hints. Rules are needed to be simple to avoid complicate firing sequence that will make the puzzle uncontrollable. In addition, rules are based on the representation of puzzles that will indeed be a set of components provided by the hints.
6. Puzzle making process is an interactive process with both rules and puzzles data. This is to say that a puzzle will not be completed unless the interacting with the users. In this case, an automatic process is rather not desired since the outcome of puzzle is dynamically made when users encounter the puzzle hints.

As seen from the characteristics described above, three key issues for an effective puzzle are the need to have a clear defined goal, some components and a mechanism to manipulate them in solving the puzzles. The clear defined goal is the design issue that will be elaborated here. The components and the mechanism are the issues we will address in the following sections.

3.3 Components of Design Puzzles

A design puzzle in our representation is comprised of *hints*, *puzzle goals* and *puzzle rules*. Each represents partial information that will direct the understanding how a puzzle can be built. Hints provide a description of situation that a puzzle is located in addition to the information players need to solve the puzzles. Puzzle goals determine how the puzzle can be satisfied or evaluated. Puzzle rules then describe how the piece of information provided by hints can be manipulated in order to satisfy the goals. In addition, with informational point of view, the outcome (the solution to the puzzles) as well as the exploration process of any participated players should be recorded for a further examination or exploration.

We have developed a system called *design puzzle* that utilizes the functionality of puzzles on the behalf of design, for this view of design described above. Design puzzles while comprised of “hint”, “puzzle rules” and “puzzle goal”, provides a way to describe our teaching capability. Consequently, an experimental system is conducted according the theory of design puzzles described here.

4. IMPLEMENTATION AND AN EXAMPLE

With the conceptual level of design puzzles, we then describe our implementation—design collage as an experimental system for computerizing the puzzle exploration behaviours described above. The reasons for choosing collage as our first implementation of design puzzle is its simplicity in both metaphorical and implementation levels. With such simplicity, we can then focus and further develop the systematic issues and the interactive mechanism of design puzzles.

4.1 Implementing a puzzle-making environment

For realizing the design puzzles, a system called “design puzzle zone” has developed in this research. The system is comprised of three system blocks—puzzle-making facilities, puzzle rules and puzzle server. The relations among these system blocks are shown in Figure 5.

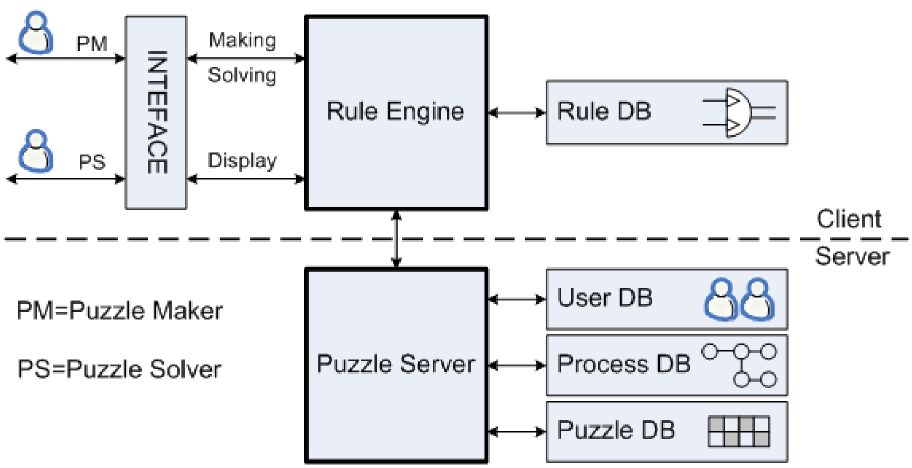


Figure 5. The system components of design puzzles

Briefly, the puzzle server is developed using LAMP (Linux, Apache, MySQL and PHP) approach with communication with a rule-based inference engine for resolving and editing the puzzle-rules created by the users at the client side. The interface of puzzle-exploration is built on top of Flash/ActionScript with a Web-based interface for its rich text capability and integration with a server. Figure 5 has clearly demonstrated this system concept. The details of each system block are described as followed.

Puzzle-making facilities comprise several tool-smiths allowing designers interactively create their own puzzles, control the exploration strategy and display the content of design puzzles. This facility will allow an extensible design puzzles to be made in respect to designers expectation and design needs. Puzzle rules and a puzzle server are another two main components of design puzzle zone. Puzzle rules provide the mechanism for simple and intrinsic rule behaviours to be made into puzzles. A puzzle server then records all the transaction between different puzzles and their exploration process according to users, process and puzzles database. Furthermore, a puzzle server will provide an access to share the learning experiences—puzzle-exploration process with others within the server domain. Or, server implementation allows the puzzles done within one server can be further searchable on the net.

For the experimental purpose, we choose representation of design collage as our design puzzle in this paper for its simplicity and design-oriented concept. While the representation and transaction between different puzzle solving stages might be different, the mechanism for rules, server and puzzle-making facilities can remain the same.

4.2 A prototype—design collage

For realizing a design puzzles system, we develop a prototype called "design collage" (an example output is shown in Figure 6) that is an implementation derived from our stage two. With designed for invoking design concepts, design collages analyse the conceptual development stages with a group of design students using puzzle-exploration metaphor. Design collage while utilizes a special view of design—image-based collage, presents a view to retrieve and incorporate the puzzle making and puzzle solving functionality.

The learning process of design collage is through converting the intention of a design sketch provided by users to generate an image collage over that intention. In Figure 6, the left side is a sketch provided by users with symbol-network attached to it, the system then generates an initial puzzle collage according to the symbolic network and keywords on the left side.

These two pair presents the puzzle-hint in our design puzzle system. Similar to jigsaw or maze puzzles, design collage as well as its concept intention—sketches and their symbolic meanings are all parts of design puzzles that users have made. By interactive defining the symbolic means as the hints for solving the puzzles, several production rules and algorithms are applied for the generative process. Further, several production rules of design collage are applied and controlled by the users for exploring the puzzle alternatives. The outcome of design puzzles is shown in the right side as the outcome of puzzle-solving/making process. Every strategic moves as well as outcome are recorded in the server with a version tag for further exploration later.

Figure 6. An example outcome from design collage

While refining the design puzzles as well as the puzzle rules users made, they would be able to redefine and further explore the puzzle-goal and its potentials. This shows the usages of design puzzles in another learning aspect. In addition, the server will be recording the puzzle exploration behaviours that can be shared among different users on that server. The puzzle rules are implemented as a set of rules that will incorporate with users in terms of generating alternatives. By using web-based search engine, puzzle collages will share with many users their collection of images as well as using web as a huge image database.

5. CONCLUSION

People are using puzzle solving and puzzle making for teaching design for both methodological and metaphorical levels. This idea is simple and effective shown in many researches, namely design games and playful learning environment. Within this scope, this particular research applies the idea to an extent for developing a workable design system that helps exploring design more effective and motivated. Several observations are unleashed by building up a puzzle-solving curriculum on a Web-based environment. They are then led to the implementation of design puzzles specified in this paper.

As observing the outcome from experiment, design puzzles sure have their niches in terms of creative thinking as well as problem solving. With the key of learning is at the content not the facility. This is to say any good metaphor will not help the learning unless it is in the hold of an inspiring teacher. However, while taking the advantage of Web-based services, self-motivation and game-playing environment can then be achieved with just a bit of help. This will be done with a lot of efforts while using different media. In addition, the process of puzzle and its exploration process cannot be studied without the help of client/server model.

Furthermore, developing a puzzle-exploration system surely helps addressing the functionality and behaviours of using puzzle solving as design exploration metaphor. Hopefully, with this tool, we might be able to understand the strength and power of design puzzles. In a long-term research project, design puzzles described in this paper come with a set of requirements and limitation that will help develop a further exploration on the puzzle-solving process that (certainly) assist design and enhance the potential of design capability of individual.

REFERENCE

Akin, O. and C. Akin. 1998, “On the process of creativity in puzzles, inventions, and designs”, *Automation in Construction*, 7(2-3), p. 123-138.

Archea, J. 1987, *Puzzle-Making: What Architects Do When No One is Looking*, Wiley-Interscience, New York, p. 37-52.

Bates, B. 2002, *Game Design: The Art and Business of Creating Games*, Premier Press.

Chang, T.-W. 2002, “Pedagogy or Andragogy: a role-interplaying approach for digital media learning”, in: Luther, Dawson, Ham, Moore, Rollo and Treloar (eds.) *36th Annual Conference of the Australian and New Zealand Architectural Science Association (ANZAScA) 2002 Conference: Modern Practice of Architectural Science: from pedagogy to andragogy*, School of Architecture and Building, Deakin University, Geelong, Victoria, Australia., p. 69-76.

Chang, T.-W. and J.H. Huang. 2001, “The acting roles model of web-based learning - Web-based Architectural Learning Envirnoment (WALE)”, *CAADRIA 2001*, Sydney(Asia), p. 323-327.

Chang, T.-W. and J.H.-L. Huang. 2002, “A pilot study of role-interplay in Web-based learning environment”, *Educational Media International*, 39(1), p. 75-85.

Chien, S.-F. 2002, “Design Gaming, Designing Games - Learning Design through Game Playing and Game Making”, *20th eCAADe*, Warsaw(Poland), p. 28-33.

Huang, J.H., T.-W. Chang and A. Radford. 2000, "A Derivation Graph of Computer Models for the design process on the Web", in: Tan, Tan and Wong (eds.) *Proceedings of the fifth Conference on Computer Aided Architectural Design Research in Asia*, School of Architecture, Centre for Advanced Studies in Architecture, National University of Singapore, Singapore, p. 307-316.
Klugman, E. and S. Smilansky. 1990, *Children's Play and Learning: Perspectives and Policy Implications (Early Childhood Education Series)*, Teachers College Press.
Knowles, M.S. 1990, *The Adult Learner: a neglected species*, Gulf Publishing, Houston.
Lee, K. and A. Wai. 2003, "Game-engine-based Cooperative Design Agent", *SIGraDi*, Rosario Argentina, p.
Lehtinen, S. 2002, "Visualization and teaching with state-of-the-art 3D game technologies", *eCAADe*, Warsaw (Poland), p. 538-541.
Radford, A. 1997, "Games and Learning about Form in Architecture", *eCAADe*, Vienna (Austria), p.
Sallkachat, R. and A. Choutgrajank. 2003, "Collaborative Design: Integrating Game Engine to Online Interactive Presentation Tool", *Proceedings of 8th CAADRIA Conference*, Bangkok Thailand, p. 295-306.
Woodbury, R.F. 1991, "Searching for Designs:Paradigm and Practice", *Building and Environment*, 26, Pergamon Press, p. 61-73.
Woodbury, R.F. and T.-W. Chang. 1997, "Using the WWW for design teaching", in: Liu (ed.) *the 2nd annual conference of Computer Aided Architectural Design Research in Asia (CAADRIA) 97*, National Chiao Tung University, Hsinchu, Taiwan, 2, p. 465-474.
Woodbury, R.F., S.J. Shannon and T.D. Sterk. 2001, "What works in a design game? Supported by student reactions to being made to play", *CAADRIA*, Sydney, p. 411-420.
Woodbury, R.F., T.G. Wyeld, S.J. Shannon, I.W. Roberts, A. Radford, M. Burry, H. Skates, J. Ham and S. Datta. 2001, "The Summer Games", *19th eCAADe*, Helsinki (Finland), p. 293-297.

Geographical Information Systems

A new computer supported design tool: RasterPlan

Alexandra Tisma
Netherlands Institute for Spatial Research, The Hague, The Netherlands

Keywords: Map Making, Computational Tools, Interactive Decision Making.

Abstract: By developing framework plans on a large scale such as a country or a region, planners use economic prognoses which show future needs for space for new spatial developments expressed in units of surfaces. Planners and designers make drawings and sketches to show where those new areas will be situated, but they do not really measure the surfaces of spaces they created. This often leads to incorrect images which can be wrongly understood by actors in decision making. To avoid this problem the Netherlands Institute for Spatial Research developed mapmaking software called RasterPlan. The purpose of this tool is to make design decisions quantitatively checkable and transparent. The maps which result from the RasterPlan are precise in geographical positioning and defining the surfaces of designed areas. RasterPlan allows realization of a quantitative program for future spatial needs for various functions such as housing, green and water areas working, and recreation. In addition to quantitative calculations, qualitative criteria for location choice can be also expressed in a form of suitability maps or buffers. This paper presents an experimental implementation of RasterPlan for the design of the future development of the Province North Brabant.

1. INTRODUCTION

Creating maps is an essential activity in spatial planning. Often maps show predicted future developments visualized in the form of sketches or vague drawings. According to Monmonier (1996) as a "scale model" map represents selective, incomplete picture of reality, where symbols used are bigger and thicker than the characteristics they represent.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 311-324.

The power of these images is actually underestimated because they can have large influence on political decisions and public opinion. Vast majority of maps users trust the maps because, “as with many things beyond their full understanding, they readily entrust mapmaking to a priesthood of technically competent designers and drafters working for government agencies and commercial firms….Map users seldom, if ever, question these authorities, and they often fail to appreciate the map’s power as a tool of deliberate falsification or subtle propaganda” (Monmonier, 1996).

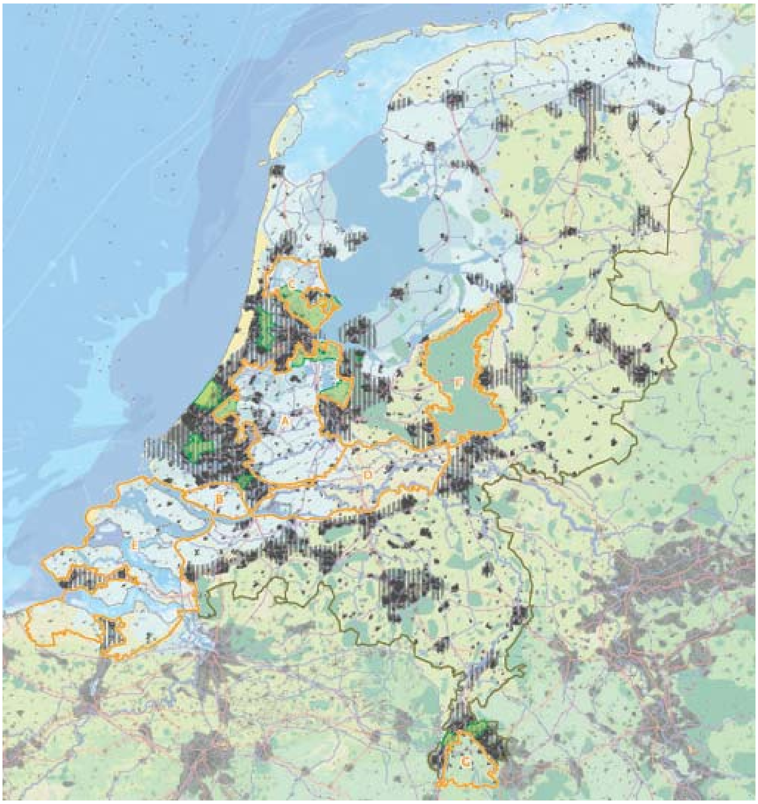


Figure 1. The Map of Urbanization Areas and National Landscapes from the Fifth Act of National Planning

An example of such a map is the image of areas reserved for urbanization and national landscapes from the Fifth Act of Spatial Planning of the Netherlands. In this image, vertical grey lines show the areas which can be committed to urbanization. As the lines are dark and thick, they give the impression that large amount of the country will be completely built up, which is not what the text of the Act explains. Actually, the grey striped areas are only search zones within which some smaller areas will be developed as urban extensions. So, urbanization zones are only a framework for planning that will be later worked out by provincial or regional planning authorities.

As many participants in a planning process do not read the whole text of the plan, they assume that the images represent real plans, and do not realize that these images are just sketches. A sketch is never meant to be precise—when images have to show exact amounts of space for future developments, other design techniques should be employed.

With the creation of Geographic Information Systems (GIS) techniques, many designers thought that GIS would be the solution to this problem. But although GIS is a powerful tool for spatial data collection and analyses, it is not frequently utilized in the design process. While the reasons for this are not the subject of this paper, we would like to mention the most important one—conceptually, GIS is a database and not a designing tool. The skills of a GIS specialist are different then those of a designer, and for that reason designers needs extra training to be able to draw in GIS. The strict logic of drawing in GIS is not designed to accommodate designers' need to create sketches when needed. For that reason, the flow of ideas is disturbed, and the process is shifted from conceptual to technical aspects, which all together require more time and efforts than usual design techniques.

In the Netherlands Institute for Spatial Research, we often deal with predictions of future spatial developments expressed in quantitative needs for extension of housing, working, nature, water areas, or infrastructure. Therefore, we needed a simple tool which can be easily and quickly used to design realistic scenarios or alternative plans for future spatial developments. In order to accomplish this, the RasterPlan software was developed and implemented for an experimental design of a plan for the Province of North Brabant. The design of the plan was also conducted with two other methods Land Use Scanner (Ruimtescanner of The National Institute of Public Health and the Environment - RVIM) and GIS based design method AGORA, developed by designers of our Institute. The three methods were then evaluated and compared. The emphasis of this paper is though on characteristics and potentials of the RasterPlan software.

2. CONTEXT OF THE RASTERPLAN AND ITS RELATION TO OTHER SIMILAR TOOLS

Since the 1960s, there have been concerted efforts to make possible robust representation of spatial problems embodied in functional models of the spatial system. In its most extreme form, this perspective assumes that spatial problems can be represented within functional models, that formal processes enable such problems to be consistently resolved, and their solutions tested using such models (Batty et al., 1998). As a result, a large number of functional models have been developed throughout the world. Those models in many cases use automatic methods, such as cellular automata, logit or agent based models, to allocate the functions. One of the most known Dutch models which uses logit model for allocation is Land Use Scanner. In this research RasterPlan will be compared with Land Use Scanner.

RasterPlan is though not a functional model. It belongs to the other group of tools which will in this text be called Computer Supported Design Methods (CSDM).

The basic difference between CSDM from functional models is that CSDM allocation is not automated. A designer or a group of designers decide where to allocate which function.

The aim of allocation models is to formalize planning process, which gives an impression that planning is conducted in a quasi-scientific manner. Unlike allocation models, CSDM enable users/actors themselves to design and choose the strategy they want when developing an area. However, this action of design is supported by quantitative and qualitative information. In that sense CSDM differ from other design techniques and are closer to functional models.

There are many CSDM systems developed nowadays; they differ in their goals, and they are applied in different settings. Some of the goals are:

- to make design transparent and controllable,
- to predict or control spatial developments,
- to manage spatial information needed for the design process,
- to develop plan alternatives,
- to evaluate the impacts of plan alternatives,
- to support design decisions,
- to support actors decision making, and
- to promote and stimulate collaborative planning process.

The web site www.PlaceMatters.com gives an overview of currently used systems which are partly allocation models and partly CSDM tools. After a quick scan of CSDM systems and in consultation with world wide planning support systems experts, it is came out that the most elaborate and the most popular ones are Community Viz, Urban Sim, Smart Growth Index (with Paint the Town as its known implementation), and PlaceIt. However, because all of these systems are either model based or GIS based and therefore not completely usable for the purposes of our Institute, we decided to develop our own tool.

3. RASTERPLAN SOFTWARE

RasterPlan is software which has been developed by the Netherlands Institute for Spatial Research in cooperation with the firm "Digital Architects." Unlike other types of software used for land use planning or allocation, RasterPlan does not rely on any kind of computational model such as cellular automata or similar algorithms. Instead, a designer or a planner decides where and in to which extent to allocate functions. RasterPlan is programmed in Delphi language as a standalone application. This paper presents the first version of the software which was developed during the past year and has been used in one case study.

Technically, RasterPlan combines some basic GIS operations with simple drawing possibilities similar to those of Photoshop. It consists of three parts:

- a drawing section where raster images can be created,
- a calculations section which calculates land use changes, and
- a section that includes some simple GIS functions such as intersections and buffers.

The interface of RasterPlan looks partly like Photoshop (figure 2) and function on similar way.

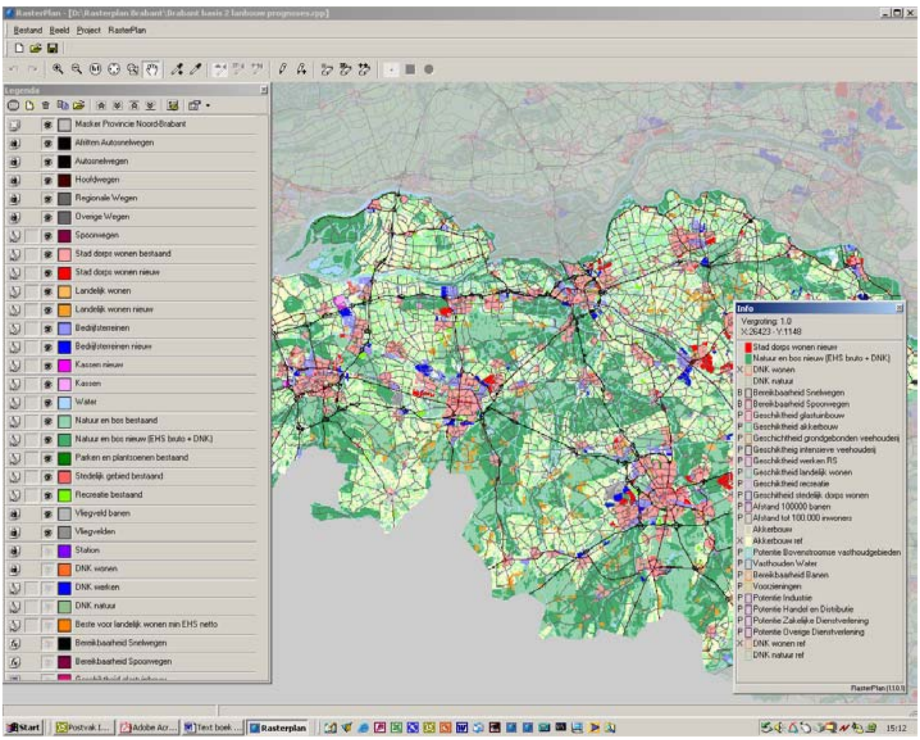


Figure 2. Interface of RasterPlan

RasterPlan uses different kinds of maps and numeric data as input information. This data can originate from GIS and other software applications which can produce bitmap files. The maps are used as overlay information for spatial analyses, but suitability maps from functional models can also be imported and used as a reference for allocation.

As the name says, RasterPlan is a raster based drawing tool. The cell of the raster can be of different size, which is defined at the start up of a project. For the regional or state scale the proper size is 100 by 100 meters, 1ha. Cells of the map are the basis for all the calculations.

RasterPlan combines drawing and calculation functions so that the sketch created can be quantitatively checked. The final product of this process is a map that shows the amounts and spatial distribution of new functions and a table that shows the changes in land use caused by this allocation. Such a map can be exported as bitmaps and used in other applications, while a table can be exported as an excel file.

As RasterPlan involves some features of GIS, it is also possible to perform qualitative analyses important for the choice of location for certain functions. For example, intersections and buffers functions can be used to produce usability maps or to find hotspots. RasterPlan is designed for making large-scale sketches, mostly on regional or above regional scale.

3.1 Implementation of RasterPlan on a Case Study of Province of North Brabant

Together with other two techniques – Land Use Scanner and AGORA, Raster Plan was used for the design of spatial development plan for the Province North Brabant. As a starting point for the design all three methods used a scenario study developed by Netherlands Institute for Spatial Research, Free University of Amsterdam (VU), Institute for Agricultural Economics (LEI-DLO), and The National Institute of Public Health and the Environment (RIVM). The scenario study was used to define quantitative and qualitative requirements for the spatial development of the Province.

There were two scenarios designed within the study. The scenario called "Individualistic World" was used for this experimental design. That scenario describes future developments in terms of quantitative program for new areas for housing, work, recreation, and glasshouses, which will replace agricultural land in the province. Next to the quantitative assignment, the scenario describes the main lines of spatial development. So, according to authors of the scenario, it is economic development which is the most important for the future of the Province, and everything that can contribute to economic progress is stimulated. The spatial consequence of such an approach is that more low density housing areas are needed. These areas are in attractive natural surroundings, and the environment is only protected where it has high values. Agricultural production is represented by large and robust enterprises.

Table 1 shows the quantitative assignment used for the design with RasterPlan, Land Use Scanner and AGORA.

Table 1. Quantitative assignment for the spatial development of the Province of North Brabant

Spatial function	Number of hectares
Housing in urban areas	5780
Housing in rural areas	6898
Working areas	10272
Gashouses	99
Nature	105173
Recreation	3978
Agriculture	-127366

As one of the aims of the research was to compare RasterPlan with Land Use Scanner, for the analyses and location choice a number of usability maps originating form the Land Use Scanner were used as analyses layaers showing where is good and where not to allocate certain functions. Genarally, RasterPlan is not dependent on data or usability maps of Land Use Scanner. It is possible to conduct own analyses by using GIS functions in RasterPlan. It was in this case more for the practical reasons and time limits that we have chosen to import the maps from Land Use Scanner.

Next to that the New Map of the Netherlands and the map of Ecological Infrastructure of the Netherlands were used to allocate new housing and new nature areas.

Figure 3 shows the methodology of use of RasterPlan for the design of the experimental plan. Figures 4 and 5 show the existing map and the planned situation of the Province North Brabant which are the result of this process.

In this plan all the hectares required in the quantitative assignment were realized. As a consequence the huge amount of agricultural land was converted to housing and nature areas.

The results of this experimental design are not meant as a definitive plan for the Province because design conducted with RasterPlan was in this case more oriented towards methodological aspects then to the content. Nevertheless, this exercise produced useful insights into a possible spatial development of the province if such a scenario as described above was to come true.

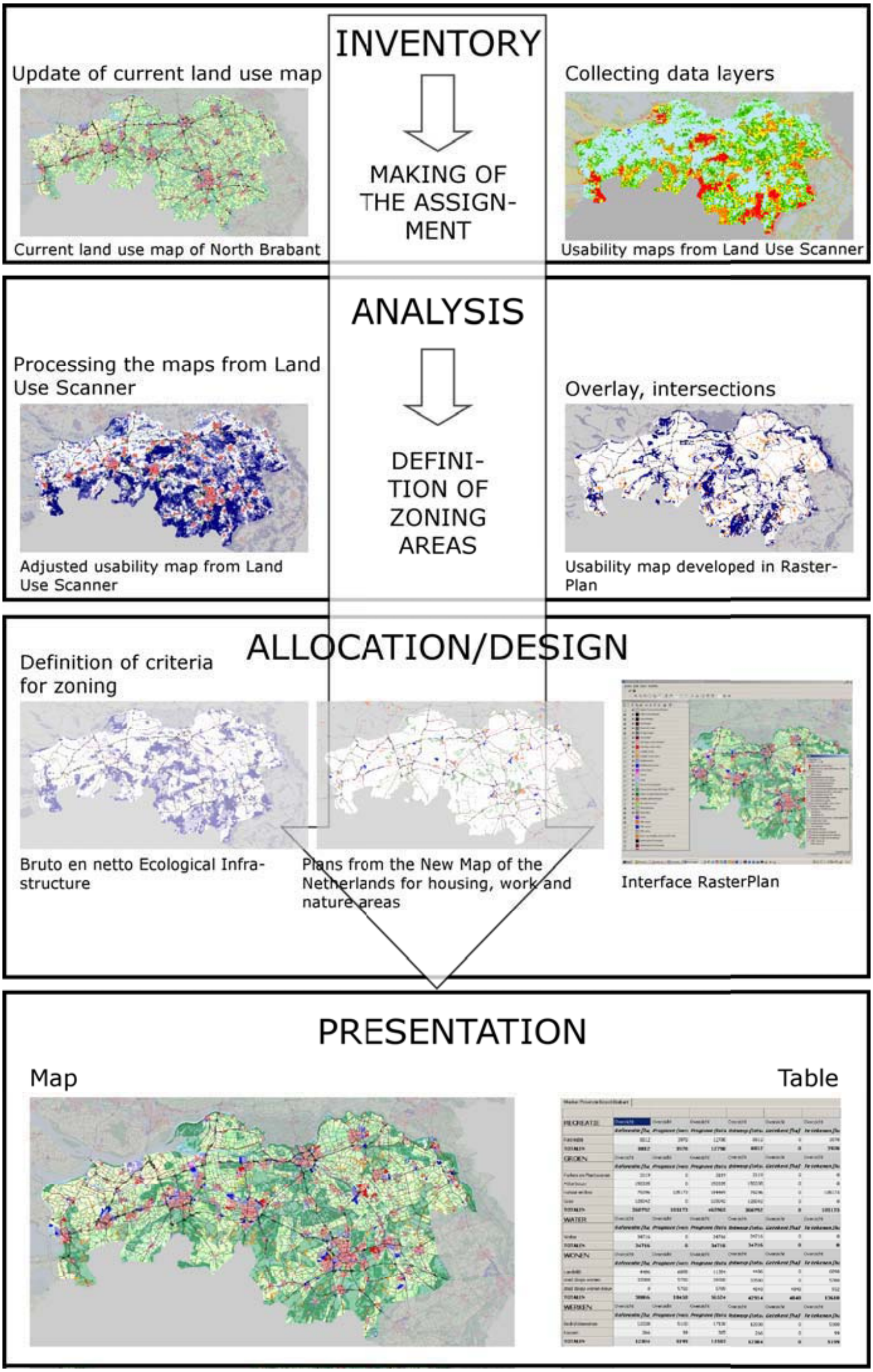


Figure 3. The methodology for the development of the spatial plan for the Province North Brabant

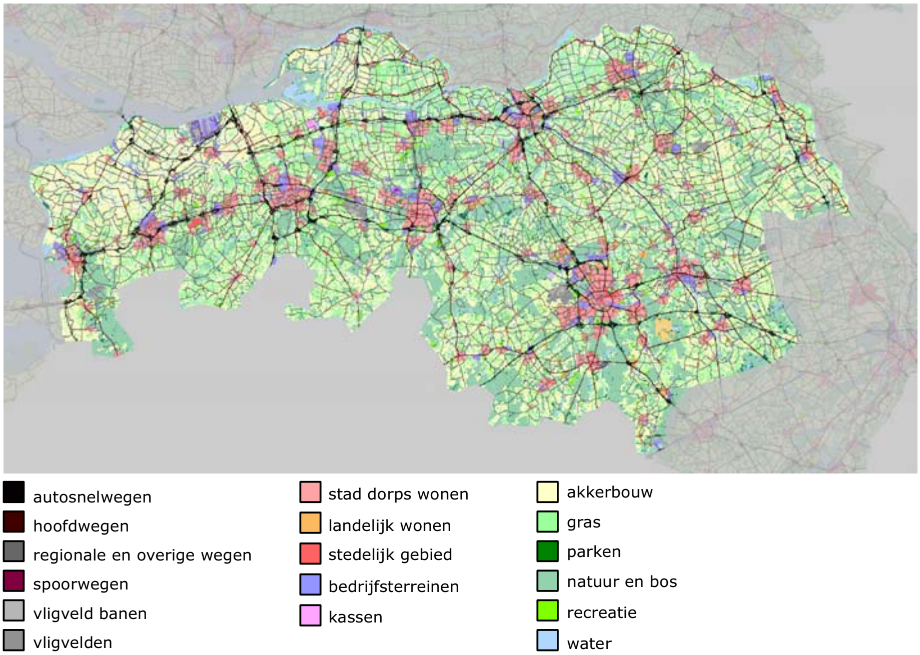


Figure 4. Map of the current land use in the Province of North Brabant made in RasterPlan

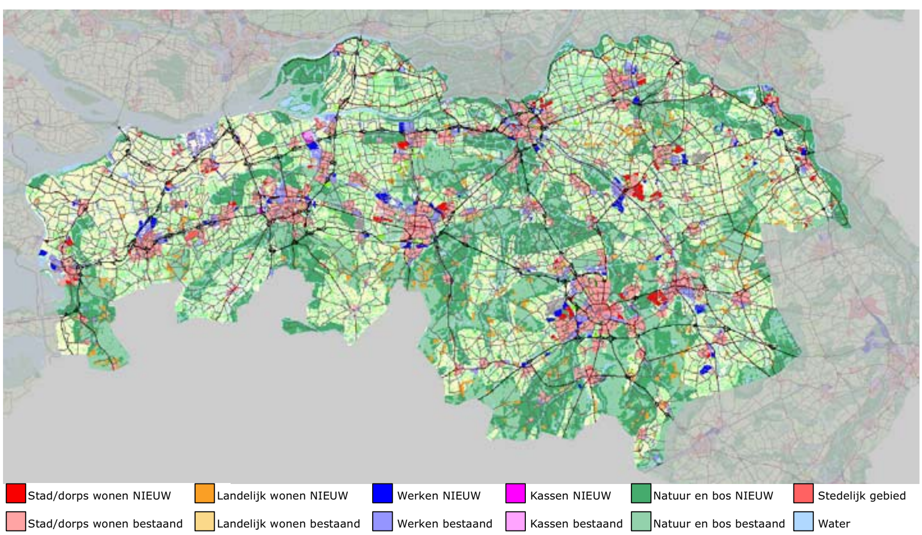


Figure 5. The result of the designing process with RasterPlan: The image of the spatial development of the Province of North Brabant till the year 2030. The grid cell is 100x100m; lighter colors represent current and darker colors future land use. To produce this plan the usability maps from Land Use Scanner were combined with national and local policy plans.

3.2 RasterPlan versus Land Use Scanner and AGORA

In this chapter we will describe other two methods which were also used for the allocation of the land use in the Province North Brabant: a functional model - the Land Use Scanner, and a design method named AGORA. Thereafter RasterPlan will be compared with these two methods.

The two methods were chosen because they are the most known and used in the planning practice in the Netherlands.

The Land Use Scanner is a known method for land allocation invented by The National Institute of Public Health and the Environment (RIVM). According to its developers, the firm Object Vision, the Land Use Scanner (Ruimtescanner in Dutch) is a GIS based information system, calculating scenarios for future land use. A spatial allocation model, taken into account the current land use, the attractiveness for future land use and regional claims, is used to predict multiple views on future land use. The Land Use Scanner has been used for various spatial planning projects. It was amongst others applied for simulating the consequences of different planning perspectives and assessing the changes in land use caused by a possible new airport location (http://www.objectvision.nl). Figure 6 shows the result of allocation which was conducted by Land Use Scanner.

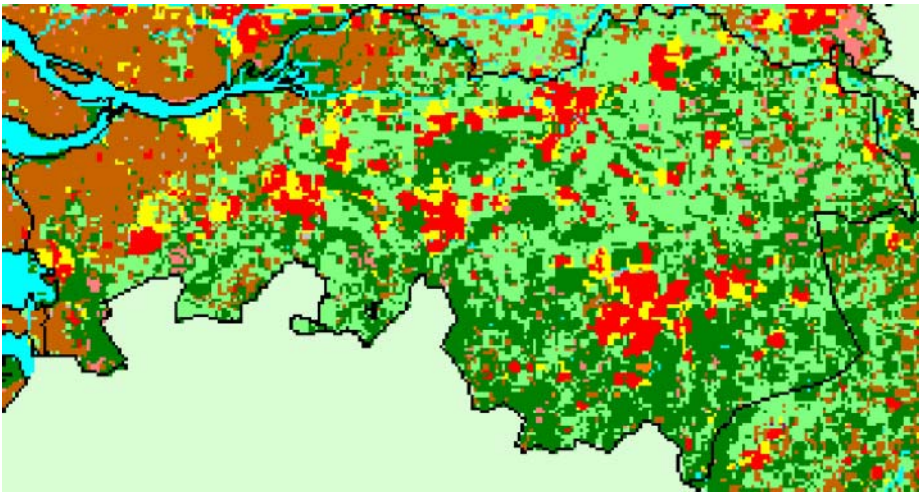


Figure 6. Future development of the Province North Brabant as a result of allocation by Land Use Scanner. Grid cell is 500x500m, only future land use is shown. The same usability maps were used for allocation as in RasterPlan

AGORA stands for Analytic GIS-supported Spatial Allocation (Analytische GIS-Ondersteunde Ruimtelijke Allocatie). Basically it is a design technique which uses GIS for analyses but also for drawing of the future scenarios. In this case the ArcMap software was used to produce experimental plan for the Province North Brabant. Figure 7 shows the "plan" developed by two young urban designers by the means of ArcMap.

The table 2 shows similarities and differences between the three methods used. The table focuses on the most important criteria which were used to compare the three methods. There were also other criteria used but because of the lack of space they won't be considered in this article.

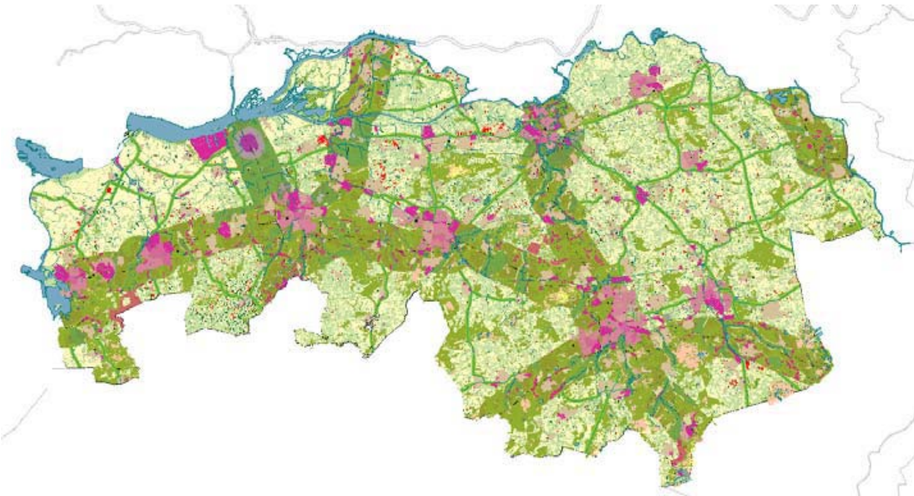


Figure 7. The plan for the Province North Brabant as a product of AGORA method. The vector drawing was produced in ArcView. The allocation is provoking and based on imagination rather than on realistic analyses.

The three methods have a lot in common. The following similarities were found:

- The final products of all three methods are not predictions or plans.
- All three methods had as a goal to translate quantitative assignment of the scenario to plausible spatial images, and all three methods succeeded in that.
- The process of making of the maps was comparable because all methods went thorough the usual design steps: definition of spatial requirements, analyses, allocation, presentation and evaluation.

There are also considerable differences between the three methods. Those are:

- Land Use Scanner is an automatic method where creativity can be expressed only in the beginning of the analyses – by definition of usability criteria. By RasterPlan and AGORA creativity of the designer is present through almost whole process and employed in analyses, allocation and presentation phase.

- RasterPlan and AGORA can easily adjust to new and unpredicted situations, to new legend units and different spatial scales. Land Use Scanner is rather rigid in that sense.
- Allocation process of Land Use Scanner is better controllable and reproducible than the design processes in RasterPlan and AGORA.

Generally said Land Use Scanner attempts to simulate the process of change on more realistic manner, AGORA tends to create provoking solutions, while RasterPlan can achieve both.

Table 2. Comparison of RasterPlan with Land Use Scanner and AGORA

Criteria	**Land Use Scanner**	**RasterPlan**	**AGORA**
Grid	500x500m	100x100m	vector
Cell	conains more functions	one function per cell, possible to pile functions	
Legend	limited number of units, defined beforehand	unlimited and flexible number of units	unlimited and flexible number of units
Spatial requirements	defined by prognoses/scenario's	defined by prognoses/scenario's	defined by prognoses/scenario's
Usability maps	produced by model, input criteria defined by user	taken from Land Use Scanner and combined with own analyses	own analyses
Current land use	basis for allocation, but not visible in the final product	basis for design and calculation of land use changes, visible in the final product	basis for design and calculation of land use changes, visible in the final product
Future land use	defined by allocation model	defined by designers in RasterPlan	defined by designers in ArcMap
Product	map and a table of future land use, various evaluation maps	map and a table of future land use	map and a table of future land use

3.3 Strong and Weak Points of RasterPlan

These initial experiments with RasterPlan highlighted many of its good points, as well as some weak points. Most of the weak points are of technical nature and occured because the software was developed a very short time. So although we knew that it is possible to make everything we missed we had not enough time to do it. As this is the very first prototype, its further developments will easily eliminate these weaknesses.

The strong points of RasterPlan are:

- Design can be quantitatively tested.
- Images, which are the product of design, are realistic.
- Alternative scenarios or hypotheses for spatial development can be easily created and tested in RasterPlan.
- It can be used individually of in workshop settings.

- No knowledge of GIS is needed to operate RasterPlan.
- It is user friendly and easy to learn.

Weak points of RasterPlan are:

- If the grid cell is small and the area large, drawing takes a lot of time. To improve this additional drawing tools (polygon and paint bucket) can be added to the software.
- The resulting images are geo-referenced and precise, comparable with topographic map. If input data are not accurate, mistakes are very easily seen in the final design. Therefore RasterPlan needs good underlying maps of existing land use, and it requires good knowledge about the area in concern.
- RasterPlan is not proper for vague sketch planning.
- Software is in the development stage so it still needs some technical improvements.

4. CONCLUSIONS

RasterPlan is a Computer Supported Design Method which makes the design process transparent and quantitatively checkable. The maps which result from the design process are precise in geographical positioning and in surfaces of designed allocation areas. RasterPlan allows realization of a quantitative program for future spatial needs for various functions on regional or state level, such as housing, working, recreation, green and water areas as well as networks of infrastructure and connections. Next to quantitative calculations, qualitative criteria for allocation can be expressed too. For that purpose RasterPlan uses some basic GIS operation.

RasterPlan is geared towards professional designers and planners; it is user friendly and compatible with GIS systems. It can be used either individually or in group settings. The tool can be actively involved in direct planning process where different actors can express their preferences and immediately bring them on the map.

Finally, RasterPlan is not meant for all design purposes, and has no intention of replacing existing allocation models or sketch techniques. Rather, its purpose is to enrich the assortment of tools which can be used either alone or in combination with other tools, such as Land Use Scanner and AGORA, to enhance the planners' performance and products.

REFERENCES

Batty, M, Dodge, M., Jiang, B. and Smith, *A*., 1998, *GIS and Urban Design*. http://www.casa.ucl.ac.uk/urbandesifinal.pdf

Monmonier, M., 1996, *How to lie with maps*. Chicago: The University of Chicago Press.

Scholten, H.J., Velde, R.J. van de, Borsboom van Beurden, J.A.M., 2001, *Ruimtescanner: Informatiesysteem voor de lange termijnverkenning van ruimtegebruik*. Vrije Universiteit van Amsterdam.

Enhancing 3DSkyView Extension Performance

A Multi-Observer Determination of Sky View Factors

Daniel S. Rodrigues[1], Léa C.L. Souza[2], and José F.G. Mendes[1]

[1] *University of Minho – Department of Civil Engineering, Braga, Portugal*

[2] *State University of São Paulo – Department of Architecture, Urbanism and Landscape, Bauru, São Paulo, Brazil*

Keywords: Sky View Factor, Urban Geometry, GIS Extension, Urban Heat Island.

Abstract: This paper presents a second version of the 3DSkyView extension. The purpose of that extension was to implement a calculation algorithm for assessment and visualization of sky view factors (SVF) by means of tools available in a Geographical Information System (GIS). The sky view factor is a thermal and geometric parameter pointed out in the specialized literature as one of the main causes of urban heat islands. A 3D-GIS is a powerful tool for reaching the goal of this research because it allows the storage, treatment and analysis of tri-dimensional urban data, in addition to a high level of flexibility for incorporating calculation algorithms. The objective in the 3DSkyView extension is to optimize the determination of that factor, not only reducing its demanding calculation and graphical representation time, but also generating a simplified tool for replacing expensive photographic equipment usually applied on this matter. Enhancing functions of *ArcView GIS 3.2*, the first version of that extension showed a very good performance allowing the automatic delineation and determination of SVF. That performance was although limited to a single observer point. The simulation of SVF for several view points in urban canyons was only possible by applying the extension as many times as the number of observers considered. Therefore, this second version was now developed in order to allow simultaneous determination of SVF for many view points. In addition, the 3DSkyView new interface is more flexible, in a way that the user may choose the kind of output wanted (graphical and/or tabular). With this new feature it is then easier to create a continuous SVF map for an entire area.

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 325-340.

1. INTRODUCTION

1.1 About the general goals of the tool

The approach of this paper suggests the use of a GIS environment for simulating obstructions that urban canyons can cause to the sky vault. Nowadays, the use of Geographical Information Systems (GIS) as a tool to understand and analyze urban areas is wide spread. Based on a technology that allows spatial and non-spatial data storage, analysis and treatment, GIS are able to optimize calculations and tasks, while reducing decision-making time. Therefore, the potential of GIS is here explored, showing its potentiality to help not only environment specialists, but also urban architects in deciding shapes and configurations for healthier cities.

In order to achieve this purpose, a tool named 3DSkyView was developed as an extension of a three-dimensional GIS, promoting the calculation and visualization of sky view factors (SVF). The 3DSkyView was conceived in Avenue scripting language in an *ArcView GIS 3.2* software with its *3D Analyst* extension switched on (all ESRI - Environmental Systems Research Institute products).

The first version of that extension (Souza, Rodrigues, et al., 2003) showed a very good performance allowing the automatic delineation and determination of SVF. That was, although, limited to a single observer point. The simulation of SVF for several view points in urban canyons was only possible by applying the extension as many times as the number of observers considered. Enhancing the performance of that tool, in this paper a second release of it is presented. This second version allows simultaneous determination of SVF for many view points. In addition, the 3DSkyView new interface is more flexible, in a way that the user may choose the kind of output wanted (graphical or tabular). With these new features it is then easier to create a continuous SVF map for an entire area.

1.2 Defining Sky View Factors (SVF)

The SVF represents an estimation of the visible area of the sky from an Earth viewpoint, being defined as the ratio between the total amount of radiation received from a plane surface and that received from the whole radiant environment. It is thus a dimensionless parameterization of the quantity of visible sky at a location. In this way the sky area results from the limits of urban canyons generated by the tri-dimensional characteristics of urban elements and their mutual relationships. As once studied by Steyn (1980), Oke (1981), Johnson and Watson (1984), Bärring, Mattsson, et al. (1985), Souza (1996), Ratti and Richens (1999), Chapman (2000), and

Chapman, Thornes, et al. (2001), the SVF is one of the main causes of the urban heat island phenomenon, therefore required as a parameter for modelling it.

As the sky usually presents lower temperatures than the Earth surface, it has an important role on the energy balance. In the process of the Earth heating loss and its consequent temperature reduction, the sky is an element that receives the long wave radiation from Earth surface. Therefore, the urban radiation loss has a straight relationship with the obstruction buildings or any other urban element can cause to the sky, when considering an Earth viewing point. Long waves are not only trapped by the warm urban surfaces during the day, but also released into the cold sky at night. So, the geometry of urban surfaces influences the radiation exchange between the Earth and the sky.

There are many methods of estimating SVF values, including mathematical models, fisheye-lens photographs analysis, image processing, diagrams or graphical determination. The calculation is, however, not straightforward and these methods are usually time demanding. In addition, the main problem of these methods is the delineation of the sky from buildings in the graphic representation. This delineation is often a task that has to be done by hand. In this matter, the work of Chapman (2000) must be remarked, since it develops a technique to enable direct calculation from a digital fish-eye image, by delineating sky pixels from the non-pixels in the image.

A more simplified method was although developed and automatized by Souza, Rodrigues, et al. (2003) in the 3DSkyView extension, whose principles are presented in the next section.

2. PRINCIPLES OF THE 3DSKYVIEW EXTENSION

The issue of SVF lies on an identification of angular dimensions between the observer and the urban element obstructions caused to the sky vault. These angles allow the urban canyon to be projected in a bi-dimensional plane, in a process where the stereographic projection is very useful. The stereographic projection of an urban canyon is an azimuthal projection, in which points of urban elements are projected to the sky vault surface (which is a hemispherical surface) and then transferred to the equatorial plane of the same sphere. This transference is possible by the union of each point on the upper sphere surface to the Nadir vanishing point, as shown in Figure 1. In this way any point on the sphere is projected into the circle representing the sky vault on the plane projection.

In order to estimate the SVF value, the sphere can be homogeneously divided and its parts projected stereographically onto the equatorial plane, creating a stereonet (Figure 2). By overlaying this stereonet on the equatorial plane projection of the obstructions, their parts (i.e., sky and obstruction areas) can be compared to the total area of the whole sky, determining their ratio (i.e., the SVF).

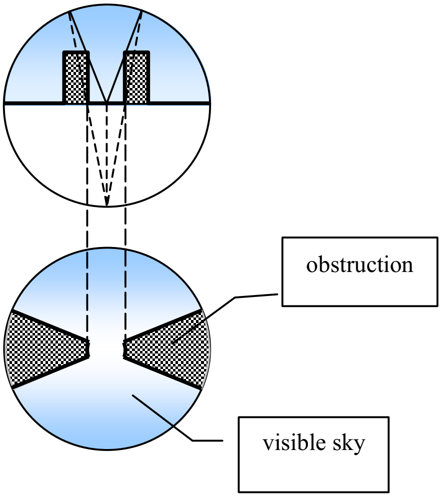


Figure 1. Stereographic Projection

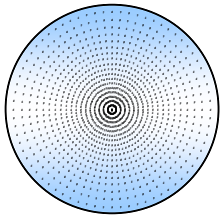


Figure 2. Stereonet

This method of calculation is automatic in the 3DSkyView extension. In practical terms, the aim of the 3DSkyView is to identify a new coordinate system for the tri-dimensional urban elements, so they could be represented in a stereographic projection on a bi-dimensional plane, in this way allowing the calculation of the SVF parameter. In the 3DSkyView extension the viewing point position is movable for all three dimensions and it can be fixed inside the urban canyon level with its focus point centred at the urban canyon level. This new coordinate system of a stereographic projection refers to the tri-dimensional relationships in the canyon. There are three important angles in the canyon determining the scene, as it is shown in Figure 3.

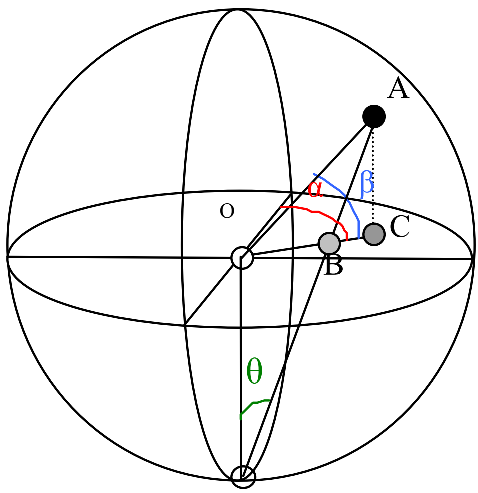


Figure 3. Important angles in the stereographic projection determination

The first one is the horizontal angle α created between the viewer North-South axes, on viewer horizontal plane, and the point of interest. The second one is the vertical angle β between the viewer plane and the point of interest. And the third one is the Nadir vanishing point angle θ between the vertical plane that contains the Nadir point and the line projected from the point of interest to the vanishing point. Considering that the viewer is in a movable position, and regarding the particularities that α should always be related to the vertical plane that contains the viewer (point O in Figure 1) and that β should always be related to the viewer horizontal plane, those angles are comparable to the azimuth and altitude angles that can be easily determined. The angle θ can be calculated by Equation 1, as it belongs to an isosceles triangle.

$$\theta = \frac{90 - \beta}{2} \qquad (1)$$

The new coordinates can then be expressed by Equations 2 and 3, which define the new coordinate system on a stereographic projection, where r is the radius adopted for the projection. Here, the α angle was submitted to an adjustment in order to have the same origin of the trigonometric relationships. This is done because α was calculated based on the North side corresponding to 0º, while the same angle for trigonometric calculation corresponds to East side. This rotation is the reason for the subtraction of α value from 90º in Equations 2 and 3.

$$x = \cos(90 - \alpha).r.\tan\theta \tag{2}$$

$$y = \sin(90 - \alpha).r.\tan\theta \tag{3}$$

With the new coordinates of the points of interest it is possible to have the stereographic projection by plotting them on the horizontal plane in *ArcView* GIS. The determination of SVF is then just a question of spatial manipulation of layers by overlaying a stereonet of equal radius on the stereographic projection of the scene. The value of SVF is calculated by Equation 4, where q is the visible area of the sky and Q is the total area of the sky defined by the area of the circle applied on the stereographic projection.

$$\varphi = \frac{q}{Q} \tag{4}$$

The simulation process of the 3DSkyView follows the steps described below:

- Based on the input themes containing the viewer point and urban elements polygons, the XY coordinates of the observer and of the vertices of the polygons are identified;
- According to the observer coordinates, the XY coordinates of the polygons are transformed into a stereographic projection. As a side product, they are also transformed into an orthographic projection;
- The polygons vertices on new coordinates are linked, depending on their original characteristics, shaping a 2D plan of the scene;
- The boundaries resulting from the new projection system are the limits of two new themes for each projection: one represents the obstruction caused to the sky and the other represents the visible sky;
- By applying GIS tools, a netpoint of the whole sky stereonet is compared to each one of these new themes, allowing the calculation of their areas and therefore the sky view factor;
- A scene simulating a projection of the urban canyon on the hemisphere is presented in a 3D environment.

As one can draw from the steps above, shapefiles containing polygons, which represent the buildings in urban areas, are required for the operation to be successful. These files can be either imported from CAD and any other compatible extension accepted by the *ArcView GIS 3.2*, or also generated in that GIS environment.

In Figure 4, the user interface of the first 3DSkyView version is presented.

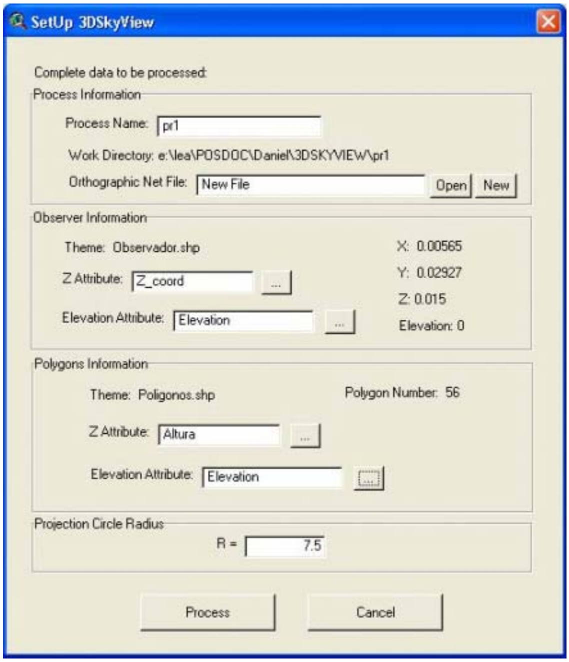


Figure 4. User interface of the 3DSkyView first version

3. 3DSKYVIEW RELEASE 2

3.1 Description and Options

Here the potentiality of the second 3DSkyView version is highlighted, demonstrating its advantages in relation to the first version. The principles of that original version in determining SVF and presented in the previous section are exactly the same. However, the ability of applying the algorithm in a simultaneous and automatic way for multi-observer points is unique. This is the main feature of version 2. The development of this new capability was the aim of the whole process now implemented. Furthermore, in version 2 the simulation time has also been taken care of.

The new users' interface presented in Figure 5 has six input data groups so that the process can be started. They are the process information, orthographic net data, information about the observer points, polygons information, radius of the circle projection, and the desired outputs.

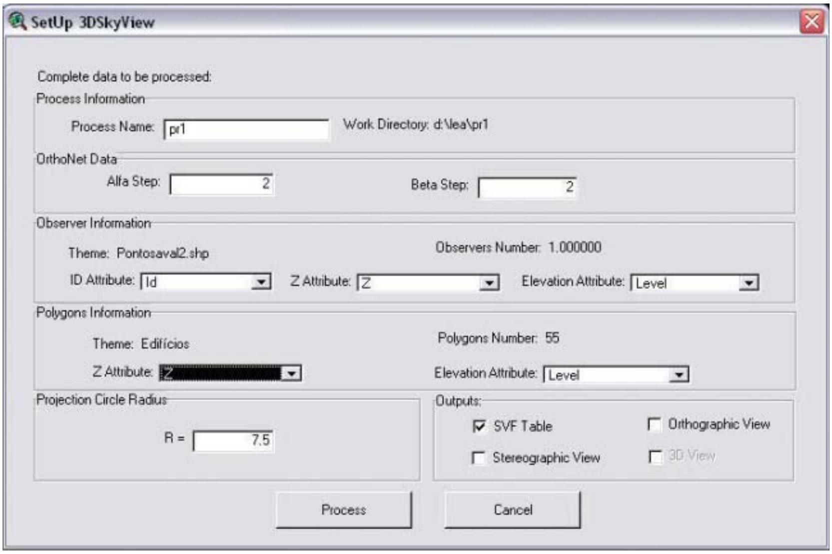


Figure 5. User Interface of 3DSkyView version 2

The process information refers to the name the user designates the process. The directory, where the resulting data (tables, shapefiles) will be stored, appears automatically next to that field.

For the generation of the skynet (the orthonet) created in this calculation, the user must supply the information about the increment of alpha and beta angles. The values of this increment have influence on the 3D model resolution. The lower the increment, the better the resolution shapes.

As the user should select both, the theme (layer) containing the observer points and the buildings (polygons) theme before running the routine, the software identifies the number of observers and the name of those themes. The observer information group requires the user identification of input data fields in the table of contents of the observers theme. Therefore, the fields on the input data table must include a unique identifier for each observer point. Also the height of the buildings and their elevations (contour line levels) are fields that ought to be available on that table of contents. The identifier will be then associated to the resulting data. For the polygons information group, these requirements are the same, except for the identifier field, which is useless here.

The projection circle radius group allows the user to choose a radius for the graphical representation of the SVF in stereographic and orthographic projections.

At last, the outputs group highlights the flexibility of this new version in allowing the user to make choices of outputs. That means that the user can optimize the time of simulation getting only the outputs of interest. (Note: the interface of 3DSkyView version 2 presented in Figure 6 does not yet allow the generation of the 3D model as a result because this option is still under construction).

These input data, followed by the click on the *Process* button, starts the process without any further user intervention.

The user' *outputs* selection is one of the advantages of version 2. It makes possible, for example, to use the extension only to get tabular results, without the graphical outputs. If the user selects only the *SVF Table* option that will create a table as shown in Figure 6. By means of the identifier number, its structure adds to each observer point the values of the sky area, the obstructed area (*CanyonArea*), and the relative area of visible sky (*SVF*).

It is then important to highlight that the simulation time is dependent on the number of observers in the input data. This happens because the final values are stored point by point in the table. This facet assures the availability of partial results, even if there is any sudden interruption on the process.

ObsID	SkyArea	CanyonArea	VisiSky	SVF
3236	353.25000	57.68043	295.56957	0.83671
3294	353.25000	97.82636	255.42364	0.72307
3342	353.25000	43.94293	309.30707	0.87560
2586	353.25000	97.82636	255.42364	0.72307
2595	353.25000	115.06223	238.18777	0.67428

Figure 6. Resulting table with sky view factor values

In comparison to the first 3DSkyView version the other results are basically the same. Only for demonstration purposes, a stereographic projection is presented in Figure 7, in which the visible sky area (Stsky2586.shp) and the obstructed area (Stcanon2586.shp) can be observed. In addition, Figure 8 shows the stereographically projected points that generated both themes of Figure 7.

In other words, this extension now called 3DSkyView2 allows the determination of urban geometry by calculating and representing sky view factors simultaneously viewed from multi-observer points. Thus, as the outputs are already stored in a GIS software, these data can be handled and plotted to create new databases and maps.

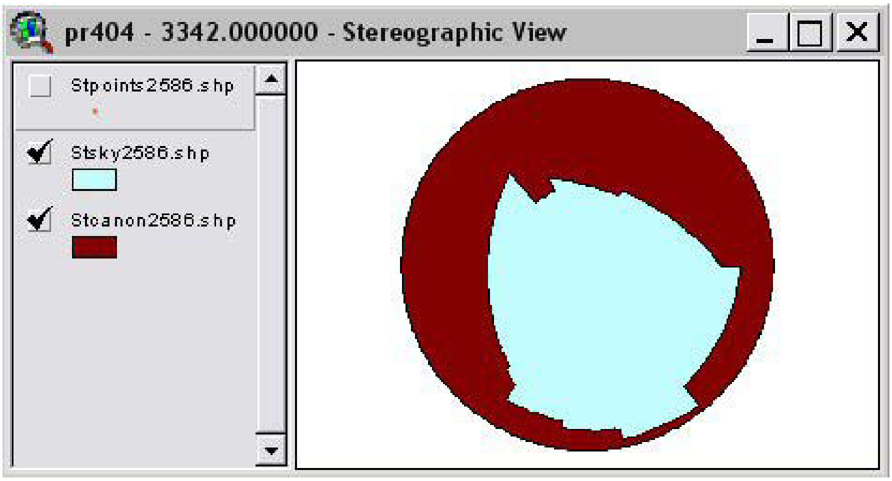


Figure 7. Example of Stereographic Projection

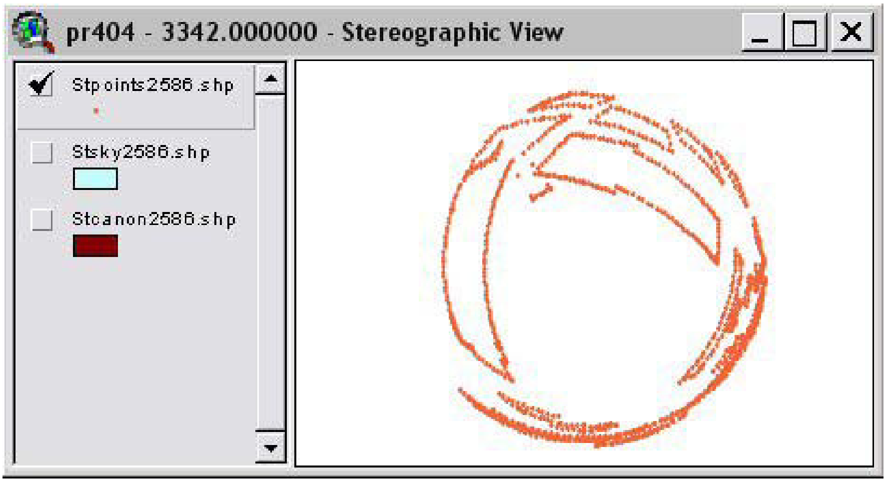


Figure 8. Points stereographically projected that allowed the delineation showed on Figure 7

4. APPLYING AND GETTING RESULTS

An application of the extension has been carried out in order to validate the main feature of 3DSkyView2, which is the multi-observer simultaneous input. This application has taken into account the University of Minho – Campus de Gualtar as the study area. That Campus lays on a peripheral area of the city of Braga, between the east side of the city and the former village of Gualtar. It occupies an area of twelve hectares. The community of the Campus has about 13000 users, with 12000 students, 800 lecturers and 300 staff employees. The buildings support academic activities, congregating Schools and Institutes, three Classroom Complexes and several buildings for services, such as the Library, the Computational Center, the Academic Services, the Sports Complex, and so on. (see Figure 9)

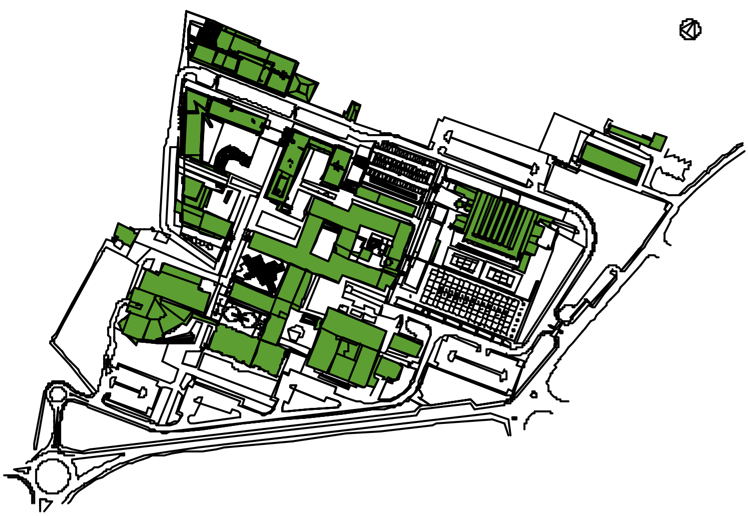


Figure 9. Plan of the University of Minho in Braga – the Gualtar Campus

A campus database was prepared for the simulation process. Two themes (layers) were required as input data, one containing attributes of the observers, and the other containing the attributes of the buildings polygons.

The first theme has been developed by means of a square net with an increment of 5 m, so that a representative cover of observers could be generated. Observer points have then been created on the nodes of this square net. Each node corresponded to an observer point. Finally, all points on this net but outside the Campus area have been removed, as well as those coincident to existing buildings or with no relevance for the study. It means that remote zones and the boundaries are out of the range of this study. This preparation steps ended up with an observer net of 3502 points, as can be seen in Figure 10. The heights of these observers were constant and corresponded to the pedestrian level of 1.50 m, with their elevation varying according to the correspondent contour line level of their position.

For the polygons representing the buildings, attributes of elevation (ground level) and height have been collected and/or estimated in field. This investigation has brought up the fact that a same building can present different heights. Thus, to have a simulation as nearest to reality as possible, the buildings theme may contain a multi-height building stored as several distinct polygons of different heights.

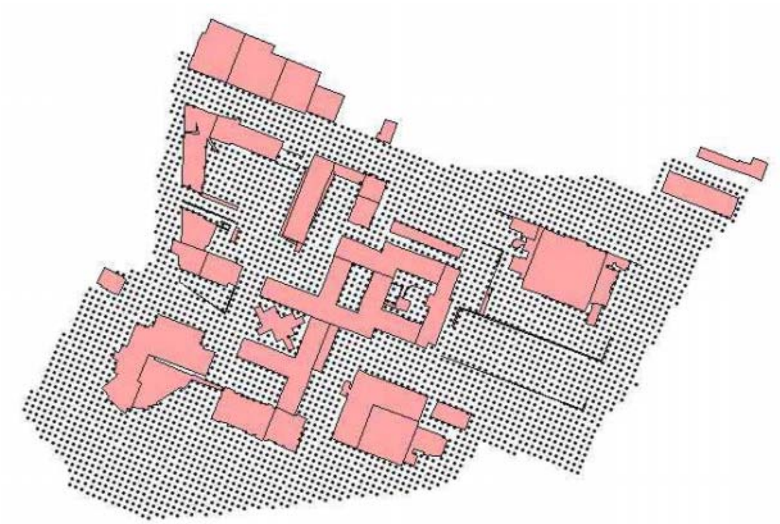


Figure 10. Buildings and Observer points in the Gualtar Campus

Due to the large number of observer points (i.e., 3502) and considering the aim of testing the main facet of version 2, just the SVF table option of the output field was checked in this application. In this way, the routine skips the graphical outputs, which usually represents a significant time-consumption in this method. Nonetheless, for a computer with a 2GHz processing unit the calculation time required for an input of 3205 observer points and 51 building polygons was roughly three days.

Based on the table that resulted from this application, Figure 11 presents the SVF values obtained and plotted in a map. As the results are expressed in percentage values, a continuous scale of colors was applied at every 10% increment. The colors vary from red to green with red indicating a SVF of 0%, while green indicates a SVF of 100%.

Figure 11. SVF values for all observer point in the Gualtar Campus

An analysis of the results shows that the SVF in this University Campus have high values, pointing out that the distances between buildings are in general well planned. Also their number of stocks is low, in a maximum of four floors. In an environmental analysis, this could indicate the high possibility of solar and natural day lighting access, as well as a high radiation heat exchange. The whole situation contributes to the user thermal and lighting comfort in the Campus. Lower values are although also noticeable. They correspond to the influence of the cover of the pedestrian path structure that links buildings (in the oldest part of the Campus).

For a future investigation, an evaluation of the Campus performance that integrates environmental parameters as solar access, day lighting, sound levels, air temperatures, surface temperatures and radiant temperatures could use the SVF as an environmental quality indicator. In other words, besides evaluating an actual situation this tool can help on future decisions and in predicting the impact of new buildings on sky view factors, before any actual site intervention on the Campus.

Here it is important to highlight that the SVF values plotted in a map help the visualization and create a proper database for integrating other environmental factors. All new information get by the use of this extension is generated in GIS software. Then, a continuing surface of SVF values based on the resulting table is an easy task to handle. This is what has been done and presented in Figure 12.

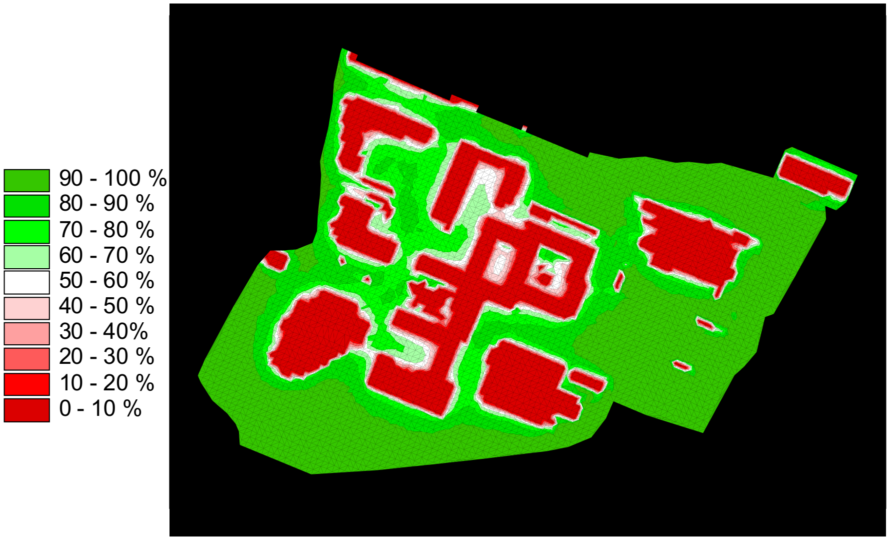


Figure 12. Continuous SVF surface of the Gualtar Campus

On the other hand, also important is the fact not mentioned before that the extension presents a limitation. Considering the inherent performance of the ArcView 3.2, the developed routine does not allow a good calculation and representation of trees or any urban element that should be represented by polygon with bottom section plans narrower than the top plans. These kinds of polygons (or buildings) are right now considered as elements with homogeneous section plans, which are simulated by their largest section. Consequently, for places with this kind of feature SVF values lower than the real ones are determined by running 3DSkyView2. On the contrary, when polygons have larger sections on their lower parts than on their upper parts, this is not verified. As a consequence, for this application on the University Campus the presence of trees was ignored.

5. CONCLUSIONS

The 3DSkyView2 integrated to other tools is a powerful software for a decision making process focusing on environmental aspects.

Both versions of 3DSkyView emphasize the potential of GIS as an important supporting tool in urban thermal analysis. Specifically for this new extension, the automatic determination of sky view factors for several observers simultaneously is its main point. Before, with the first version, it was only possible to do that by running the extension as many times as the number of observers considered. Now, the data are presented in a unique table that associates the results for each observer.

For the application here conducted on a University Campus, the tool demonstrated its ability and potentiality as a decision support tool. Adopting a dense net with an increment of 5 meters it was possible to extrapolate an analysis from an individual point of view to a general analysis of the whole Campus.

Future efforts are being directed to new outputs, such as solar diagrams or integrating ground level contour lines as part of the scenarios. The latter could reduce the time demanded to prepare the input data.

Furthermore, the software *ArcView GIS 3.2* itself has offered some limitations during the development of the extension and, moreover, there is already another version for this software. This new family, now *ArcView 8.x* and *9.x*, does not applies any more the *Avenue* programming language. That implies that a translation to the actual code of Visual-Basic for Applications should be studied in the near future in order to make the extensions also available to the more recent versions of *Arcview*.

ACKNOWLEDGEMENTS

The authors would like to demonstrate their gratitude to the Brazilian agencies CAPES (Post-Graduate Federal Agency), FAPESP (Foundation for the Promotion of Science of the State of São Paulo), CNPq (Brazilian National Council for Scientific and Technological Development), and FUNDUNESP (Foundation for the Development of the State University of São Paulo), and to the Portuguese agency GRICES (Directorate General in the Ministry of Science and Higher Education) for their cooperation in different phases of this research.

REFERENCES

Bärring, L., J.O. Mattsson, and S. Lindqvist, 1985, "Canyon geometry, street temperatures and urban heat island in Malmö, Sweden", *Journal of Climatology*, (5), p. 433-444.

Chapman, L., J.E Thornes, and A.V. Bradley, 2001, "Rapid determination of canyon geometry parameters for use in surface radiation budgets", *Theoretical and Applied Climatology*. 69, p. 81-89. Available in: http://www.cert.bham.ac.uk/research/urgent/canyongeometry.pdf.

Chapman, Lee, 2000, "Improved one dimensional energy balance modeling utilizing sky-view factors determined from digital imagery", in: *Proceedings of the 10th SIRWEC Conference,* Davos, Switzerland. Available in: http://www.sirwec.org/conferences/davos2000.html.

Johnson, G.T. and I.D.Watson, 1984, "The determination of view-factors in urban canyons", *Journal of Climate and Applied Meteorology*, (23), p.329-335.

Oke, T.R., 1981, "Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations", *Journal of Climatology*, 1(1-4), p.237-254.

Oke, T.R., 1988, "Street design and urban canopy layer climate", *Energy and Buildings*, 11(1-3), p.103-113.

Ratti, C.F. and P. Richens, 1999, "Urban texture analysis with image processing techniques",.in: *Proceedings of the CAADFutures 99, Atlanta*, USA, Kluwer Academic Publishers, Boston.

Souza, L.C.L., 1996, *The influence of urban geometry on air temperatures at the pedestrian level*. PhD Dissertation. University of São Paulo, São Carlos. 125 p (in Portuguese).

Souza, L.C.L, D. S. Rodrigues, and J.F.G. Mendes, 2003, "A 3D-GIS extension for sky view factors assessment in urban environment", in:: *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management – CUPUM 2003, Sendai, Japan*, p. 27-29.

Steyn, D.G., 1980, "The calculation of view factors from fisheye-lens photographs", *Atmosphere-Ocean*, 18(3), p.254-258.

Relationship between Convenience Store Robberies and Road Environment

Masahiro Murakami, Kotaro Higuchi, and Akihiro Shibayama
Kogakuin University, Japan

Keywords: Convenience Store Robbery, GIS, Graph Theory, Indexes of Graphs and Networks.

Abstract: This study focuses on road environment around robbed convenience stores and elucidates the road characteristics that are conducive to convenience store robberies by using GIS and indexes of graphs and networks. The method of this study is as follows: Five convenience store robberies, which occurred in Metropolitan Tokyo, were selected from newspaper reports. Then, road networks within a 1-kilometer radius of the robbed convenience stores were extracted from digital maps with a scale of one-twenty five hundredth (Geographical Survey Institute and Bureau of City Planning Tokyo Metropolitan Government). After adding the road networks and the attributes such as road width, we investigate the road characteristics using GIS and indexes of graphs and networks. Finally, we demonstrate several factors associated with convenience store robberies based on this compiled information.

1. INTRODUCTION

1.1 Background and purpose

In Japan, the incidence of crime has been increasing recently. In particular, the number of convenience store robberies has rapidly increased, and the rate of criminal arrests has concurrently decreased sharply, revealing that the police can not fully cope with the present situation (see Figure 1). Crimes are known to occur in locations with spatial factors conducive to

J.P. van Leeuwen and H.J.P. Timmermans (eds.), Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, 341-356.

criminal acts, as has been explained in earlier studies on CPTED (Crime Prevention through Environmental Design). Therefore, robbed convenience stores potentially should also have spatial factors conducive to robberies. It is necessary to demonstrate these factors in order to prevent future convenience store robberies.

However, there are few previous studies on this topic in Japan, and to the best of our knowledge, one study was presented by Kashiwabara et al (1997). They extracted several factors using KJ method that are associated with convenience store robberies, but these factors are qualitative and conceptual. In addition, they did not examine road environment in detail although they indicated that robbers regard road environment as an important aspect in order to successfully escape their pursuers.

This study focuses on the road environment around robbed convenience stores and elucidates the road characteristics that are conducive to convenience store robberies by using GIS and indexes of graphs and networks.

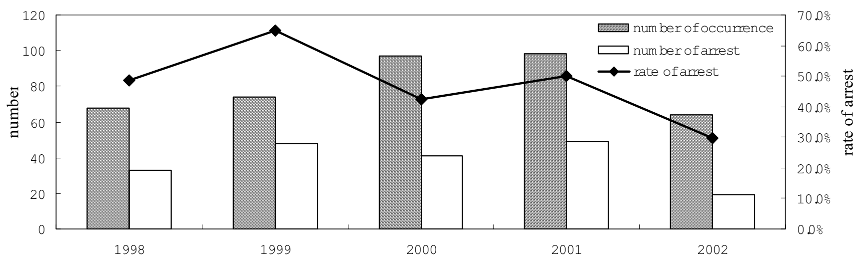


Figure 1. Trend of convenience store robberies in Metropolitan Tokyo (1998-2002)

1.2 Method

In Japan, detailed information on convenience store robberies is usually not made public. In newspapers, partial information (e.g. the date and time of the robbery, the name and address of a robbed convenience store, the characteristics of the robber, the amount of money stolen) is printed, but not all convenience store robberies are covered. Information collected only from newspapers was used in this study. The method of this study is as follows.

Five convenience store robberies, which occurred in Metropolitan Tokyo, were selected from newspaper reports (see in Table 1). Then, road networks within a 1-kilometer radius of the robbed convenience stores were extracted from digital maps with a scale of one-twenty five hundredth (Geographical Survey Institute and Bureau of City Planning Tokyo Metropolitan Government), as shown in Table 1. After adding the road networks and the attributes such as road width, we investigate the road characteristics using GIS and indexes of graphs and networks. Finally, we demonstrate several factors associated with convenience store robberies based on this compiled information.

GIS software ArcGIS 8.1 with Spatial Analyst (Esri Co., Inc., Japan) and SIS ver.6.0 (Informatix Co., Inc., Japan) was use to organize data and analyze road networks.

Table 1. Information on robbed convenience stores

Name of convenience store		Address of convenience store	Occurrence
Case 1	Poplar Wakabayashi	Wakabayashi, Setagaya-ku, Tokyo, Japan	Dec., 2002
Case 2	Ampm Komone	Komone, Itabashi-ku, Tokyo, Japan	Nov., 2002
Case 3	Ampm Nakano-Kamitakada	Nakano, Nakano-ku, Tokyo, Japan	Aug., 2002
Case 4	Seven-Eleven Minami-Ikebukuro	Minami-Ikebukuro, Toshima-ku, Tokyo, Japan	Aug., 2002
Case 5	Sunkus Takadanobaba	Takadanobaba, Shinjuku-ku, Tokyo, Japan	Aug., 1999

2. PREVIOUS STUDIES

As part of this study, the related literature was divided into two parts: crime prevention (2.1) and road network evaluation (2.2).

2.1 Related literature on crime prevention

Kashiwabara et al. selected 24 convenience store robberies, which had occurred from 1990 to 1995 in Tokyo, and examined associated factors using the KJ method. The extracted qualitative factors, such as regional characteristics and population density, were used to classify the urban environment around the convenience stores into several groups. They further arranged the characteristics of the groups (Kashiwabara, Itoh et al, 1997).

Crowe analyzed the floor plans and surroundings of several convenience stores, which are in U.S., with different locations or building shapes from the viewpoint of CPTED (Crowe, 1991).

Some previous studies focused on detecting hot spots of snatchings and residential burglaries by methods such as calculating the kernel density estimation with GIS (Harada, 2002) and investigating the patterns of criminal distribution by calculating the spatial autocorrelation with GIS (Shimada, Suzuki et al, 2002). In Japan, there are still few studies on crime prevention that utilize GIS. In this study, the kernel density estimation with GIS is used as one of the means of analyzing road characteristics that are conducive to convenience store robberies.

2.2 Relevant literature on road network evaluation

A considerable number of studies have been conducted on road network evaluations. The main studies are as follows.

Toi and Yoshitake compared β-value and γ-value with the proposed Node Distance Distribution and effectively applied the latter to some cities (Toi and Yoshitake, 1992).

Takagi et al. estimated road networks of underground shopping arcades by using some indexes of graphs and networks. In conclusion, analysis combining these indexes is effective for examining structural differences between road networks (Takagi, Taniguchi et al., 1991).

In this study, the formations of road networks around robbed convenience stores were evaluated based on the above-mentioned studies.

3. DATA ARRANGEMENT

Based on digital maps with a scale of one-twenty five hundredth (Geographical Survey Institute and Bureau of City Planning Tokyo Metropolitan Government), road networks within a 1-kilometer radius of robbed convenience stores were extracted using GIS (see Figure 2). After individually measuring the width of the roads, we added the information to the attributes of the roads and made a database for investigating road characteristics that are conducive to convenience store robberies.

Figure 2. Example of road networks

4. CHARACTERISTICS OF ROAD ENVIRONMENT

4.1 General statistics

The general statistics of road networks extracted according to methods in the previous sections are shown in Table 2. Case 1, Case 2, and Case 3 were designated as being in a residential district (hereafter, residential group), and Case 4 and Case 5 were designated as being in a commercial district (hereafter, commercial group). The residential group had more intersections and roads than the commercial group. As for the total length of road, the residential group tended to be greater. Road density (m/m^2) was higher in the residential group than in the commercial group. Characteristics with each group were rather similar.

Table 2. General Statistics on road networks

Index	Case 1	Case 2	Case 3	Case 4	Case 5
Number of intersections	1788	1697	2027	1595	1386
Number of roads	2401	2237	2699	2125	1766
Total of road length (m)	99,851	97,703	103,042	87,783	79,613
Road density * (m/m^2)	0.032	0.031	0.033	0.028	0.025

* The area of each district is 3,140,000 m^2 (mean the area of a circle 1 kilometer in radius).

4.2 Road network conditions

We focused on the road width and examined the characteristics of the road networks around the robbed convenience stores (see Table 3).

The ratio of total road length to road width showed that roads with widths of 4 to 6 meters tended to be very common in all districts. In addition, all districts had many roads with widths of less than 6 meters though the ratio of those roads differs in degree. In Case 3, roads with widths of less than 6 meters accounted for about 70% of all roads. There are many narrow roads in the road networks around the robbed convenience stores. According to the crime statistics by Metropolitan Police Department (2002), convenience store robberies occur with considerably frequency between 1 a.m. to 6 a.m. This is because robbers plan to approach and run away from convenience stores without being witnessed by passers-by. On narrow roads, there is generally little traffic from midnight to early morning, making districts with many narrow roads advantageous to robbers. It is thought that having many narrow roads around convenience stores is one of the factors conducive for convenience store robberies.

Table 3. Characteristics of road networks

Type	Item	Total	Road width					
			RW1	RW2	RW3	RW4	RW5	RW6
Case 1	Number of roads	2,401	188	1,117	638	247	153	58
	Total of road length (m)	7,364	7,364	45,162	27,337	11,506	6,433	2,049
	Ratio (%)	100.0	7.4	45.2	27.4	11.5	6.4	2.1
Case 2	Number of roads	2,237	176	809	724	405	48	75
	Total of road length (m)	97,703	6,649	32,474	35,084	17,749	1,843	3,904
	Ratio (%)	100.0	6.8	33.2	35.9	18.2	1.9	4.0
Case 3	Number of roads	2,699	522	1,417	360	246	108	46
	Total of road length (m)	103,042	19,393	53,930	13,590	10,096	4,113	1,920
	Ratio (%)	100.0	18.8	52.3	13.2	9.8	4.0	1.9
Case 4	Number of roads	2,125	279	809	395	277	161	204
	Total of road length (m)	87,783	9,229	31,306	16,611	12,905	7,758	9,973
	Ratio (%)	100.0	10.5	35.7	18.9	14.7	8.8	11.4
Case 5	Number of roads	1,766	243	872	300	134	97	120
	Total of road length (m)	79,613	10,741	36,831	14,225	6,475	4,690	6,652
	Ratio (%)	100.0	13.5	46.3	17.9	8.1	5.9	8.4

RW1: less than 4 meters
RW2: from 4 meters to 6 meters
RW3: from 6 meters to 8 meters
RW4: from 8 meters to 12 meters
RW5: from 12 meters to 20 meters
RW6: more than 20 meters

After drawing concentric circles around the robbed convenience stores at intervals of 100 meters, we calculated the ratio of total road length to road width in each interval (see in Figure 3). These calculations showed that the whole district, as well as the neighborhood of the robbed convenience stores, had many roads of widths less than 6 meters. In other words, narrow roads, which have little traffic from midnight to early morning, are broadly distributed around the robbed convenience stores. Thus, it is thought that the road networks in the robbed districts were advantageous for committing robberies at convenience stores.

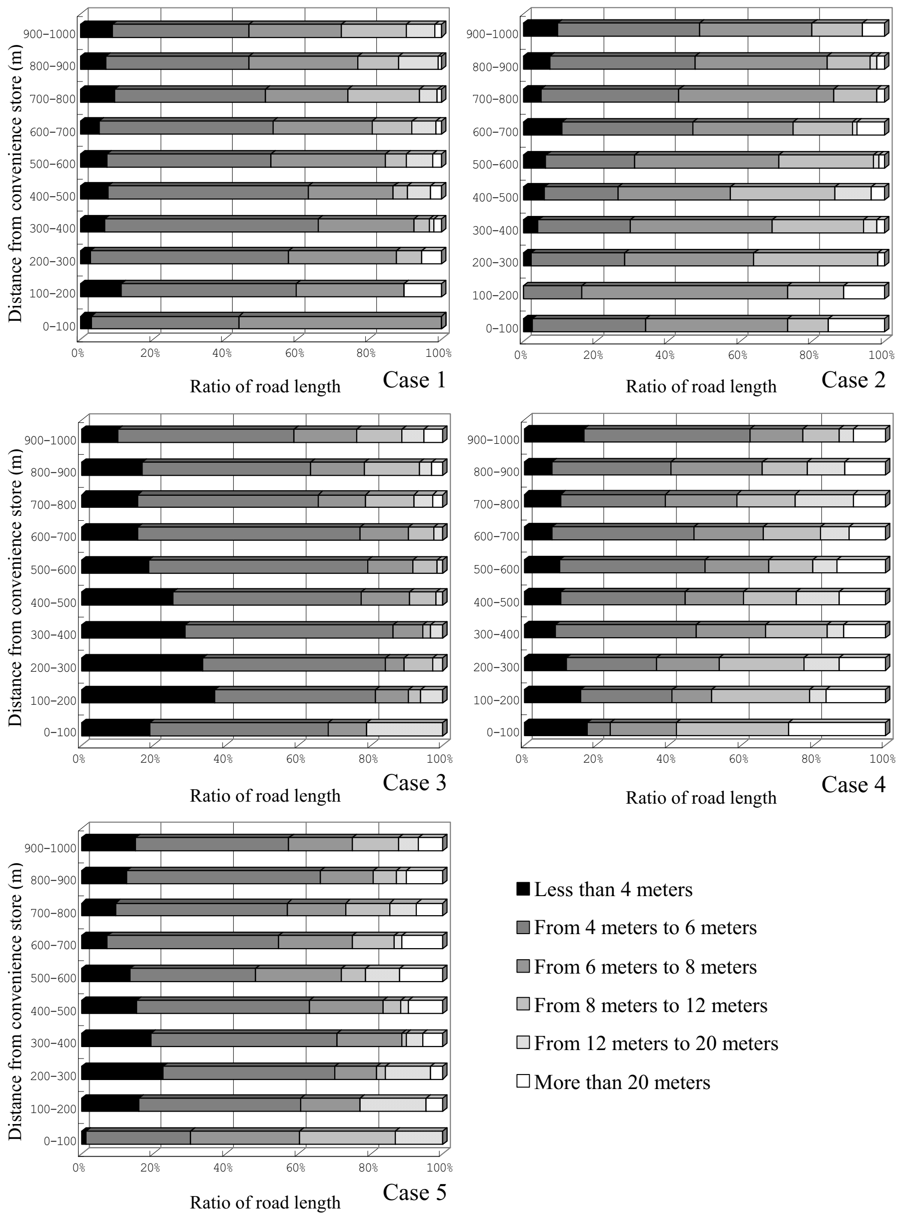
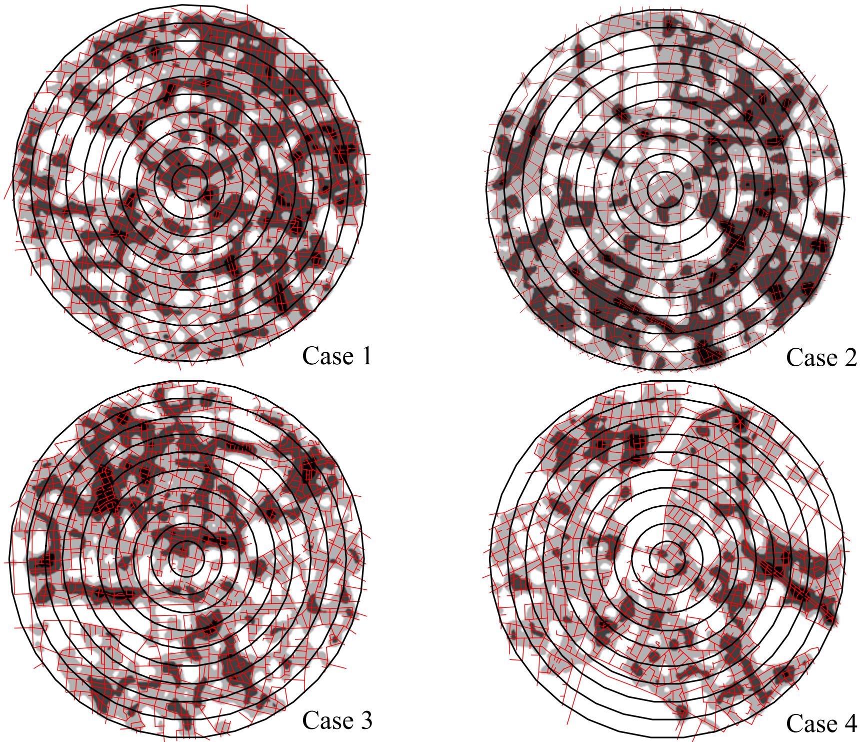
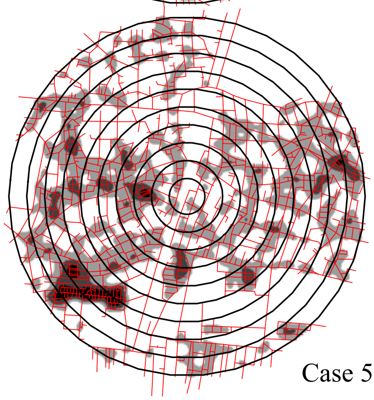


Figure 3. Ratio of total road length to road width

4.3 Kernel density estimation

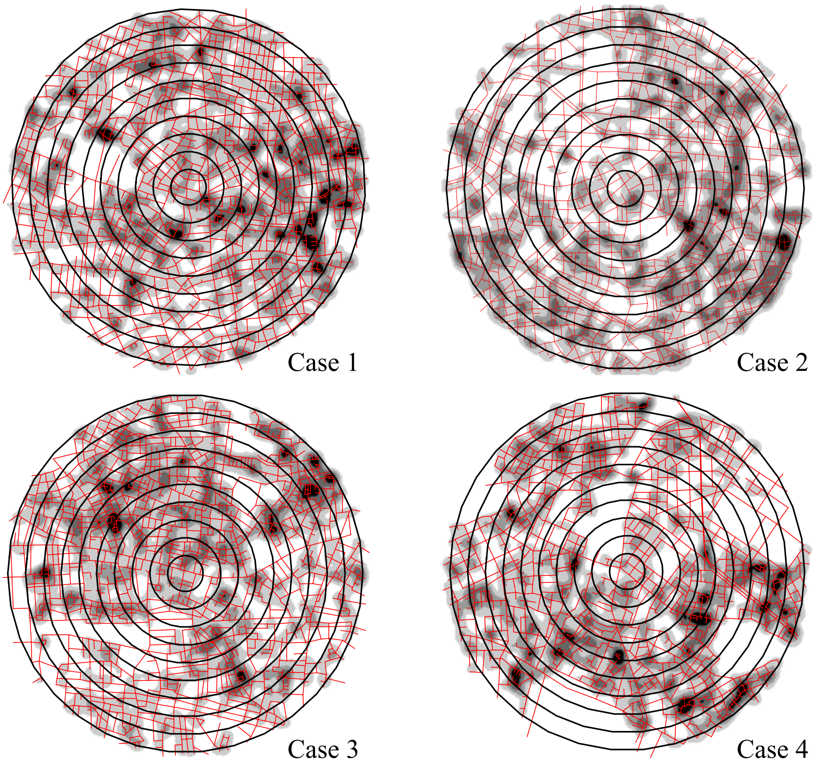
The kernel density of the roads of less than 6 meters in width and intersections was calculated using ArcGIS 8.1 with Spatial Analyst (see Figures 4 and 5). Based on this result, the distribution density of these characteristics can be shown visually.

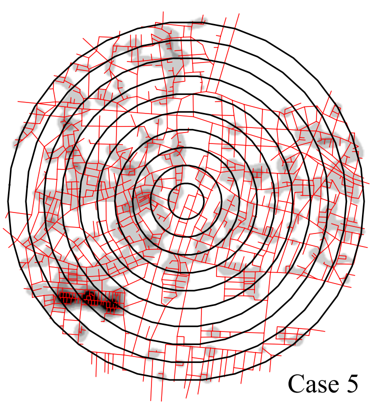




*As a shade deepens, the kernel density becomes high.
**The concentric circles are drawing around the robbed convenience stores at intervals of 100 meters.

Figure 4. Kernel density of roads with widths less than 6 meters





*As a shade deepens, the kernel density becomes high.

**The concentric circles are drawing around the robbed convenience stores at intervals of 100 meters

Figure 5. Kernel density of intersections

In Figure 4, areas around the robbed convenience stores have a widely and continuously ranging high density of roads with widths of less than 6 meters allow robbers to approach and run away from convenience stores without being witnessed by passers-by from midnight to early morning. Furthermore, when comparing these high-density areas of roads with widths of less than 6 meters and many intersections, we find that these areas have both narrow roads and many intersections distributed throughout the district. In such areas, robbers may easily escape their pursuers due to the density of narrow roads and poor visibility.

Thus in these areas of narrow streets and numerous intersections that range widely around the robbed convenience store, robbers may elude their pursuers and escape without being witnessed by passers-by from midnight to early morning. Based on this result, it is thought that the road environment around robbed convenience stores is potentially conducive to these crimes.

5. CHARACTERISTICS OF ROAD NETWORK FORM

Indexes of graphs and networks, as shown Table 4, are widely used to evaluate the formations of road networks (Takagi, Taniguchi et al, 1991, Toi and Yoshitake, 1992). In this chapter, the road networks around the robbed convenience stores were analyzed using those indexes (Table 5).

Table 4. Indexes of graphs and networks

Index		Equation
Evaluation index on graph theory	β-value	$\beta = e / v$
	γ-value	$\gamma = e / 3(v-2)$
Evaluation index on distance	Diameter	$T = \max . d(i, j)$
	η-value	$\eta = L / e$
	π-value	$\pi = L / T$
	A_i	$A_i = \sum_{j=1}^{v} d(i, j) / (v-1)$
	Average of A_i	$D = \sum_{i=1}^{v} \sum_{j=1}^{v} d(i, j) / v(v-1)$

Where, e is the number of roads, v is the number of intersections, L is total of road length, and $d(i, j)$ is the shortest distance between intersections.

5.1 Evaluation by indexes on graph theory

The γ-value represents the overall connectivity of the road network. Overall connectivity (about 40%) was quite similar for Cases 2, 3, and 4, although Case 1 and Case 5 showed a slightly different tendency.

The β-value represents the partial connectivity inside the road network. Due to the many blind alleys in the road network of Case 5, it shows a lower value compared with the other districts, which tended to be very similar. Since robbers may be easily cornered, they do not usually select districts with many dead-ends. Then, why did the convenience store robbery occur in the district? Perhaps the robber had superior knowledge of the district. Therefore, it must be recognized that convenience store robberies may also occur in districts with many dead ends.

Next, the β-value was examined from a different viewpoint. The number of roads connecting at one intersection is typically twice the β-value (Takagi, Taniguchi et al., 1991). Twice the β-value was calculated to be nearly 3.0 in all districts, showing that there are many T-intersections in the road networks. The poor visibility around T-intersections allows robbers to elude their pursuers there. It is thought that the number of T-intersections around convenience stores is one of the spatial factors associated with convenience store robberies.

Table 5. Calculation result of indexes of graphs and networks

Index	Case 1	Case 2	Case 3	Case 4	Case 5
β-value	1.34	1.32	1.33	1.33	1.27
γ-value	0.45	0.44	0.44	0.44	0.43
Diameter	2615	2605	2887	2722	2947
η-value	41.6	43.7	38.2	41.3	45.1
π-value	38.2	37.5	35.7	32.2	27.0
Average of A_i	1108	1130	1120	1132	1147
Standard deviation of A_i	165	170	185	165	186

5.2 Evaluation by indexes on distance

The following indexes were synthetically examined (Takagi, Taniguchi et al., 1991): the average road length in a road network (η-value), the intricacy of a road network (π-value), the overall traffic efficiency of a road network (average of A_i) and the uniformity of road length (standard deviation of A_i).

Case 1, Case 2, and Case 4 tended to be similar, with the following differences.

The road network of Case 1 is considerably intricate, but traffic efficiency is high (low average of A_i, 1108) and road length is uniform as a whole (low standard deviation of A_i, 165); that is, the shape and size of block is regular as a whole. Because of this, even a robber with a poor sense of the locality may manage to successfully escape.

As for Case 2, the road network is intricate, and the traffic efficiency is low (high average of A_i, 1130). In addition, η-value (43.7) and the standard deviation of A_i (170) also tend to be comparatively high. Compared to the network in Case 1, the size of the block is large on the average and the shape and size of block is comparatively irregular.

As for Case 4, the road length is uniform (low standard deviation of A_i, 165). Compared with Case 1, this road network is simple (π-value, 32.2), but traffic efficiency is very low (high average of A_i, 1132). Based on this, the road network of Case 4, but not that of Case 1, is advantageous to robbers who do not have a sense of the locality for escaping their pursuers.

Case 3 and Case 5 were found to be quite different.

As for Case 3, the road network is intricate (moderately high π-value, 35.7), but that has high traffic efficiency (low average of A_i, 1120). However, the short road length (low η-value, 38.2) and the standard deviation of A_i (185) show that the size of block is much smaller on the average and the shape and size of block is considerably irregular. Though this road network has high traffic efficiency, it is thought that robbers may not escape successfully if they lack a superior sense of knowledge of the district.

As for Case 5, π-value (27.0) is much lower. However, the average of A_i (1147) tends to be higher due to number of blind alleys mentioned earlier. In addition, η-value of 45.1 is much higher and the standard deviation of A_i (185) is considerably high. Thus, Case 5 tends to be quite different from Case 3, but Case 3 and Case 5 are similar in that it is difficult to readily escape without knowledge of the locality.

As shown above, we found that the road networks around the robbed convenience stores have many structural differences, but we failed to demonstrate common characteristics to all road networks using these methods. We have identified the following reasons: first, only a small number of convenience store robberies were included in this study; second, we did not have access to the means of escape (on foot, or by bicycle, motorcycle or car) for the robbers and could not analyze this important point; third, the level of knowledge of the locality was not available.

Next, the frequency distribution of A_i was examined (see Figure 6). From this frequency distribution, we analyzed how the areas of high intersection density are distributed within the road networks (Toi and Yoshitake, 1992). In short, a peak indicates an area of high density, and the value of A_i that the peak is appeared means the distribution of that area. For example, Case 1, Case 2, Case 3, and Case 4 have a peak A_i value of 931, indicating that high-density areas are located on the contour line of that value (see Figure 7a, 7b, 7c, 7d, 7e).

Figure 6 shows that not only do we find many peaks in all districts, but we also find several peaks with the same A_i value. The former shows that the areas with a high density of intersections are distributed widely around the robbed convenience stores in all districts, and the latter shows that there are some similarities in the distribution. Although the distribution in high-density areas where robbers may elude their pursuers is depicted visually in the former chapter, here we succeeded to precisely identify it. Furthermore, it was found that the distribution of those areas tended be similar.

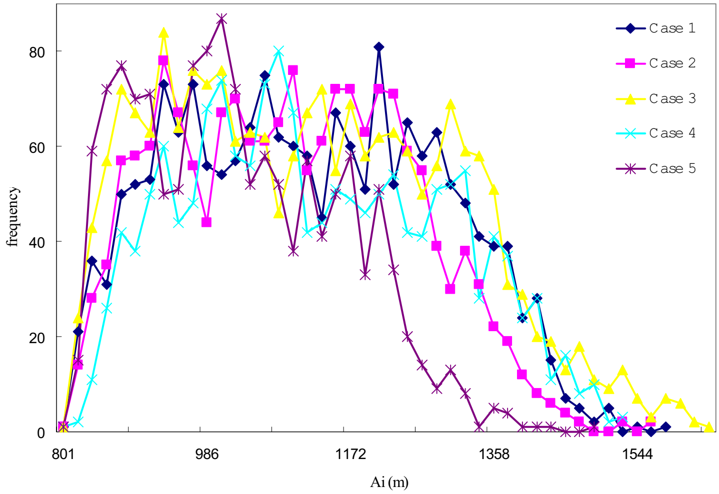


Figure 6. Frequency distribution of A_i

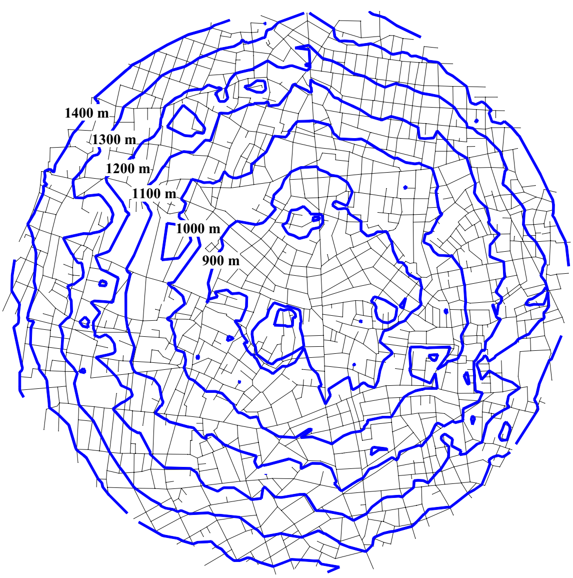


Figure 7a. Contour of A_i (Case 1)

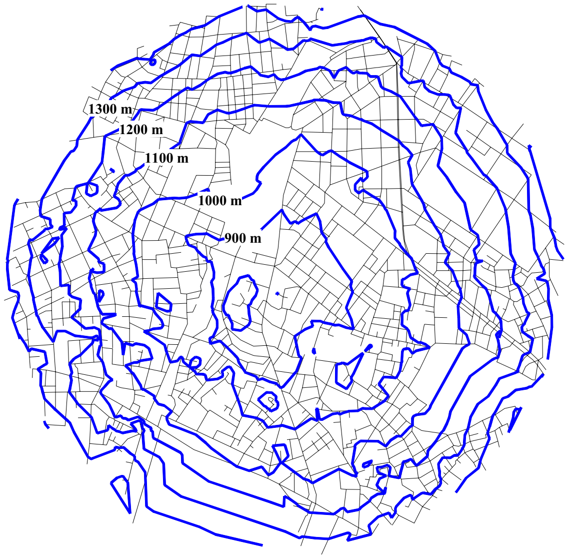


Figure 7b. Contour of A_i (Case 2)

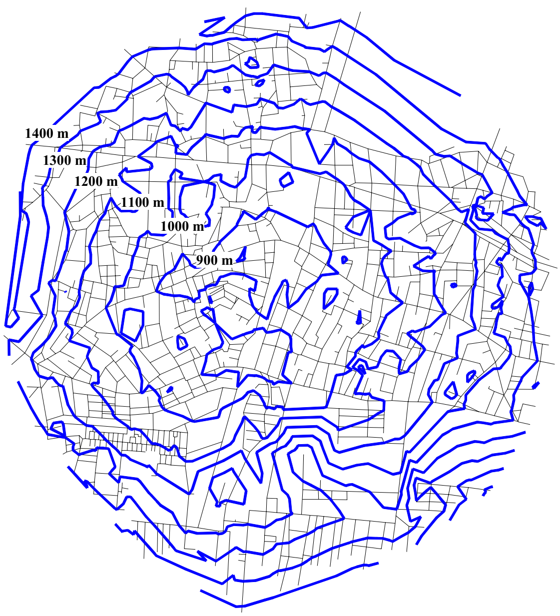


Figure 7c. Contour of A_i (Case 3)

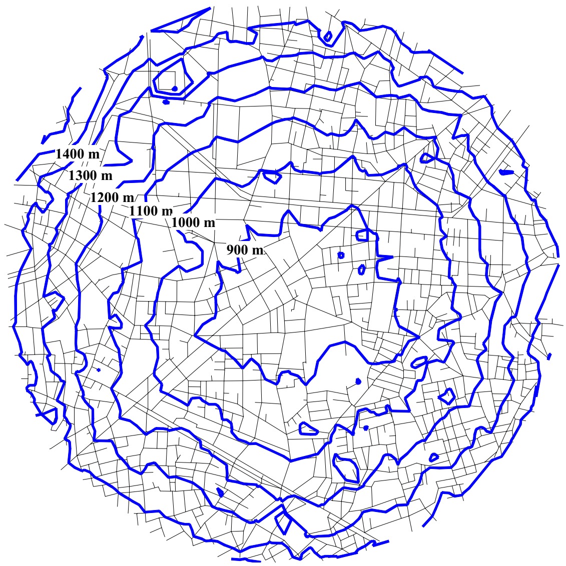


Figure 7d. Contour of A_i (Case 4)

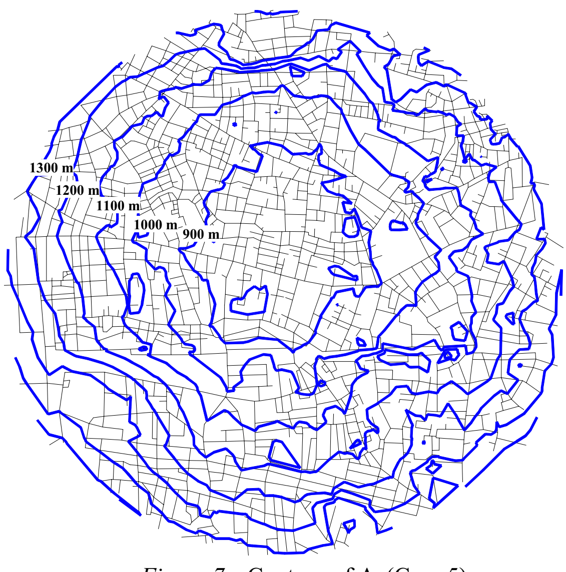


Figure 7e. Contour of A_i (Case 5)

6. CONCLUSION

Five convenience store robberies, which occurred in Metropolitan Tokyo, were selected from newspaper reports, and the road networks around the robbed convenience stores were investigated using GIS and indexes of graphs and networks. We succeeded to demonstrate factors that influence the incidence of convenience store robberies. Based on this result, we found that the road networks around the robbed convenience stores potentially had common characteristics. GIS and indexes of graphs and networks are effective techniques for explaining road characteristics that are conducive to convenience store robberies. In this study, we constructed a scheme based on the available information. In the future, it will be desirable to have a tool that can objectively identify likely targets of convenience store robberies.

ACKNOWLEDGEMENT

This study was supported by the Obayashi Foundation and the Grant-in-Aid for Project Research in Kogakuin University in Japan.

REFERENCES

Kashiwabara, T., A. Itoh, and T. Omi, 1997, "A Study of the Relationship between Crime Prevention for Convenience Store and Urban Factors in Tokyo Metropolitan Area", *Papers on City Planning*, No.32, p. 715-720.

Crowe, T., D., 1991, *Crime Prevention Through Environmental Design*, Butterworth-Heinemann Ltd., Oxford.

Harada, Y., 2002, "Crime Prevention and GIS", *2001 CPTED Workshop: JUSRI Report No.15*, p. 12-27.

Shimada, T., M. Suzuki, and Y. Harada, 2002, "Analysis of Spatial Crime Patterns Using Moran's I Statistics", *Theory and Applications of GIS*, Vol.10, No.1, p. 49-57.

Toi, S. and T. Yoshitake, 1992, "Network Form Evaluation in City Region Using Node Distance Distribution", *Papers on City Planning*, No.27, p. 271-276.

Takagi, M., H. Taniguchi, and J.S. Kim, 1991, "Basic Study of Indexes of Graphs and Networks and Analysis of Street Structures of Underground Towns by Applying Them", *Journal of Archi. Plann. Environ. Engng*, No.422, p. 37-44.

Metropolitan Police Department, 2002, *Crime Condition in Tokyo Metropolitan 2002*, Metropolitan Police Department, Japan.

AUTHOR INDEX