

Xiaoyan Zhang · Martin Constable
Kap Luk Chan · Jinze Yu
Wang Junyan

Computational Approaches in the Transfer of Aesthetic Values from Paintings to Photographs

Beyond Red, Green and Blue

 Springer

Computational Approaches in the Transfer of Aesthetic Values from Paintings to Photographs

Xiaoyan Zhang · Martin Constable
Kap Luk Chan · Jinze Yu
Wang Junyan

Computational Approaches in the Transfer of Aesthetic Values from Paintings to Photographs

Beyond Red, Green and Blue

 Springer

Xiaoyan Zhang
College of Computer Science and Software
Engineering
Shenzhen University
Shenzhen
China

Jinze Yu
Institute of Industrial Science
The University of Tokyo
Tokyo
Japan

Martin Constable
Royal Melbourne Institute of Technology
Ho Chi Minh
Vietnam

Wang Junyan
University of Southern California
Los Angeles, CA
USA

Kap Luk Chan
Tolendata Singapore R&D Centre Pvt. Ltd.
Singapore
Singapore

ISBN 978-981-10-3559-3 ISBN 978-981-10-3561-6 (eBook)
DOI 10.1007/978-981-10-3561-6

Library of Congress Control Number: 2017944299

© Springer Nature Singapore Pte Ltd. 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Colour is uncontainable. It effortlessly reveals the limits of language and evades our best attempts to impose a rational order on it... To work with colour is to become acutely aware of the insufficiency of language and theory - which is both disturbing and pleasurable.

David Batchelor, 'Chromophobia', 1997

Preface

Paintings are assumed to be a special class of image for the quality of their aesthetic appeal. In a painting, little is left to chance, with colour, form, composition and suchlike all being given careful consideration by the painter, who is a highly trained individual. In this way, a painting offers an aesthetic experience that is unlikely to be found in a photograph.

New generations of image capture devices, such as Google Goggles and the light field camera, promise a future in which the formal attributes of a captured scene may be made available for editing to a degree that has hitherto been the exclusive territory of painting. In this sense, paintings and photographs are converging, and it therefore seems an opportune time to address them in comparison to each other. This book details cutting-edge work demonstrating how some of the aesthetic attributes of a painting may be transferred to a photograph using the latest computational approaches.

The research described in this book was the result of a collaboration between engineers and an artist and was a substantial learning experience for all. For the artist, the engineering contributed an empirical dimension to the classic literature on the topic, most of which hails from a time before an image could be subject to such exact evaluation. For the engineers, the complex and multidimensional way that artists fabricate images was a source of constant surprise.

For its emphasis on technical process, this book will be of primary appeal to engineers. However, we can imagine it also being of relevance to artists seeking an empirical understanding of the aesthetic image.

A Quick Tour of the Contents

Part I (Part I on p. 1) discusses the image attributes and aesthetic aspects of paintings. Introduced in Chap. 1 on p. 3 are the two painting genres that are the focus of this book: the portrait and the landscape. In Chap. 2 on p. 9 the topic of colour is introduced. This introduces the six structural contrasts that are the

foundational core of our research: global contrast, local contrast, centre-corner contrast, regional contrast, neighbouring regional contrast and depth-aware contrast. In Chap. 3 on p. 33 is presented a summary of ways in which contrast may be computationally modelled. In Chap. 4 on p. 39 the topic of composition is addressed, with a special focus on vignetting.

Part II (Part II on p. 53) details a number of methods by which the aesthetic values of paintings may be transferred to photographs. These are as follows:

- **Transfer of the global contrast** (Chap. 5 on p. 55). In this early work, we analyse the global hue, saturation and lightness contrasts of paintings by two well-known artists: Vincent Van Gogh (Dutch 1853–1890) and Sir Lawrence Alma-Tadema (English 1836–1912). The style of a painting is considered as a high-level organisation of these contrasts. This organisation is transferred to a photographic image by histogram matching of saturation and lightness and also by dominant hue alignment and spread stretch along the RYB colour wheel.
- **Transfer of the atmospheric perspective effect** (Chap. 6 on p. 71). The atmospheric perspective effect is a natural phenomenon, relating to the effects of atmospheric distortion upon the appearance of distant objects. We show that artists exaggerate this phenomenon in order to improve the visual appeal of the painting and the illusion of depth. The contrast manipulation is formulated as an optimisation problem that simultaneously considers the desired inter-contrast, intra-contrast and specified gradient constraints.
- **Transfer of regional contrast from portrait paintings** (Chap. 7 on p. 107). The function of a portrait is to draw attention of the viewer to the sitter's face. We show that painters serve this function through attention to the inter- and intra-regional contrast values of the painting. In transferring this contrast structure from paintings to photographs, we employ a novel piecewise nonlinear transformation curve.
- **Transfer of the pose-related composition in portraits** (Chap. 8 on p. 137). Painters exhibit substantial discretion in the manner in which they frame a figure within a portrait. Pose, face direction and space around the figure are the main elements considered in portrait composition. A graph model is proposed that describes these attributes, the purpose of which is to select a painting with similar geometrical organisation. Space cropping technique is used to improve the composition of the photograph by aligning its composition to that of the selected painting.
- **Transfer of the vignetting effect** (Chap. 9 on p. 155). Emulating the natural fade-off of human vision, vignetting manifests as a darkening of the edges and corners of a painting, and a lightening of its centre. We compare the vignetting effect in paintings with that of photographs. The observation from the analysis shows that the vignetting effect in paintings is more purposely organised. Based on the observation, an algorithm is explored to transfer the lightness weighting from an example painting to a similar photograph.

- **Defining hue contrast** (Chap. 10 on p. 179). This chapter addresses the thorny subject of global hue contrast. There are well-understood methods to define global lightness and saturation contrast, which all reference such intensity-based values as average, least and most. Hue cannot be understood in such terms and is therefore not definable using the same approach. Using as reference Johannes Itten’s colour contrasts, we present an approach by which the global hue contrast of an image may be quantified. It is particularly suitable to visualise the global hue contrast of large sets of images. We apply this definition on sets of aesthetically related images, thereby expressing their high-level hue organisation.
- **Interactive local hue contrast** (Chap. 11 on p. 191). Hue has an inherent semantic dimension, the hue of an object frequently being linked to its meaning. For this reason, it is very perilous to edit the hue of an image. This chapter proposes a way in which the hue of an image may be manipulated without its semantic meaning being damaged. The colour of an object frequently presents not as a single hue, but rather a range of hues. Hence a green tree may consist of yellow-greens and blue-greens. Taking the cue from the traditional painters practice of ‘breaking’ their colours, we present a means by which this range may be compressed or stretched. We develop a tool based on superpixels that can interactively select a range of hues, to make them available for this manipulation.

Shenzhen, China
Ho Chi Minh, Vietnam
Singapore, Singapore
Tokyo, Japan
Los Angeles, USA
2017

Xiaoyan Zhang
Martin Constable
Kap Luk Chan
Jinze Yu
Wang Junyan

Acknowledgements

We would like to acknowledge all the fantastic editors and their support staff, who worked those long hours and toiled so very hard on this book.

Thanks also to the Institute for Media Innovation, Nanyang Technological University (NTU), Singapore, for providing the PhD grant to complete part of the work. A special thanks to Prof. Michael S. Brown, of the National University of Singapore (NUS) for all his support and guidance, and to Prof. Martin Reiser for his support and valuable input.

Contents

Part I Image Attributes and Aesthetic Aspects

1	The Portrait and Landscape Genres	3
1.1	The Portrait Genre	3
1.2	The Landscape Genre	5
1.3	Concluding Remarks	8
	References	8
2	The Colour Attributes of Paintings	9
2.1	Painters, Paint and Colour Systems	9
2.2	Colour Spaces	10
2.3	Colour Contrast	16
2.3.1	Itten's Contrasts	17
2.3.2	Structural Contrasts	18
2.3.3	Hue and Harmony	25
2.4	Concluding Remarks	30
	References	31
3	Computational Models of Colour Contrast	33
3.1	Contrast Models for Simple Images	33
3.2	Contrast Models for Complex Images	34
3.3	Contrast Models for Hue Contrast	36
	References	37
4	The Geometric Attributes of Paintings	39
4.1	Composition in Paintings	42
4.1.1	The Rule of Thirds	42
4.1.2	Region of Interest and the Horizon	44
4.2	Vignetting	46
4.3	Concluding Remarks	51
	References	51

Part II Transfer of Aesthetic Values

- 5 Global Colour Style Transfer 55**
 - 5.1 Global Hue, Saturation and Lightness Analysis 56
 - 5.2 Hue Mapping 58
 - 5.2.1 RGB to RYB Conversion 58
 - 5.2.2 Dominant Hue Extraction 60
 - 5.2.3 Dominant Hue Alignment and Mapping 61
 - 5.3 Saturation and Lightness Mapping 62
 - 5.4 Experiments and Discussion 62
 - 5.5 Concluding Remarks. 70
 - References. 70
- 6 Atmospheric Perspective Effect Transfer for Landscape Photographs 71**
 - 6.1 Introduction 71
 - 6.2 Related Work 74
 - 6.3 Paintings Versus Photographs. 76
 - 6.3.1 Depth Planes 76
 - 6.3.2 Contrast in Paintings and Photographs 76
 - 6.4 Contrast Manipulation. 78
 - 6.4.1 Reference Selection 79
 - 6.4.2 Depth-Aware Contrast Mapping 81
 - 6.5 Experiments and Discussion 89
 - 6.6 Concluding Remarks. 102
 - References. 104
- 7 Regional Contrast Manipulation for Portrait Photograph Enhancement 107**
 - 7.1 Introduction 107
 - 7.2 Related Work 110
 - 7.3 Paintings Versus Photographs. 111
 - 7.4 Example-Based Portrait Photograph Enhancement 113
 - 7.4.1 Reference Selection 114
 - 7.4.2 Contrast Mapping 117
 - 7.5 Experiments and Discussion 120
 - 7.6 Concluding Remarks. 133
 - References. 134
- 8 Composition Improvement for Portrait Photographs 137**
 - 8.1 Introduction 137
 - 8.2 Related Work 138
 - 8.3 Proposed Composition Improvement Method 139
 - 8.3.1 Graph Matching. 141
 - 8.3.2 Space Cropping 143

- 8.4 Experiments 144
- 8.5 Concluding Remarks. 152
- References. 152
- 9 Vignetting Effect Transfer 155**
 - 9.1 Introduction 155
 - 9.2 Related Work 156
 - 9.3 Vignetting Effect in Paintings 157
 - 9.4 Transfer of Painter-Style Vignetting to Photographs 161
 - 9.4.1 Lightness Weighting Pattern and Blending Function 161
 - 9.4.2 Lightness Weighting Pattern Correction 163
 - 9.4.3 Content-Aware Interpolation 166
 - 9.4.4 Local Contrast Restoration 170
 - 9.5 Experiments and Discussion 171
 - 9.6 Concluding Remarks. 176
 - References. 178
- 10 Defining Hue Contrast 179**
 - 10.1 Introduction 179
 - 10.2 Related Work 180
 - 10.3 Defining Hue Contrast 180
 - 10.4 Our Approach. 181
 - 10.5 Computational Details. 183
 - 10.6 Results and Interpretations 184
 - 10.7 Evaluation Through a User Study. 188
 - 10.8 Concluding Remarks. 189
 - References. 189
- 11 Interactive Local Hue Contrast Manipulation. 191**
 - 11.1 Introduction 191
 - 11.2 Related Work 194
 - 11.3 Hue Manipulation Framework 195
 - 11.3.1 System Overview 195
 - 11.3.2 Segmentation. 197
 - 11.3.3 Hue Histogram Selection. 197
 - 11.3.4 Hue Contrast Manipulation Operators 199
 - 11.4 Experimental Results 200
 - 11.4.1 Hue Compression and Hue Stretch 200
 - 11.4.2 Combining the Three Operations. 202
 - 11.5 Discussion and User Feedback 203
 - 11.6 Concluding Remarks. 204
 - References. 204

Part I
Image Attributes and Aesthetic Aspects

Chapter 1

The Portrait and Landscape Genres

Abstract While contemporary paintings come in every conceivable form, pre-modern paintings were understood in terms of only five categories or ‘genres’, with some genres being more important than others. These genres, in descending order of importance, were: history painting (which were large multigure compositions addressing significant topics), portrait, genre, landscape and still life. This book examines two of those genres: the landscape and the portrait. Paintings in general are supremely organised things, the function of this organisation being to create a memorable visual impact. For landscapes, this impact is achieved largely through accentuating the illusion of depth that they convey. For portraits, the face and the figure are of prime importance. This chapter introduces the histories of these genres and examines them for their functional differences.

Keywords Portraits · Landscapes · Painting

1.1 The Portrait Genre

The social function of a portrait is to display the likeness of an individual, together with their personality and social standing [2]. Through the history of portrait painting, this function has remained the same. For this reason, their visual appearance has changed very little through the ages. The painted portrait genre underwent a decline in popularity that started with the introduction of the photographic portrait, a medium that to a large degree replaced painting. The prime formal requirement of a portrait is to draw attention to the visible skin areas of the body, especially the face.

Portraits are very alike in the way that their subject matter is formally treated, with little to distinguish one portrait from another in the particularities of their construction. However, distinct subtypes may be identified. Based on how much of the human body is visible, portraits can be classified as being ‘full-length’ or ‘half-length’. The half-length portrait contains the upper body, while the whole body of the figure is visible in the full-length portrait. The half-length can also include the ‘head only’ category, wherein only the head and shoulders are visible.

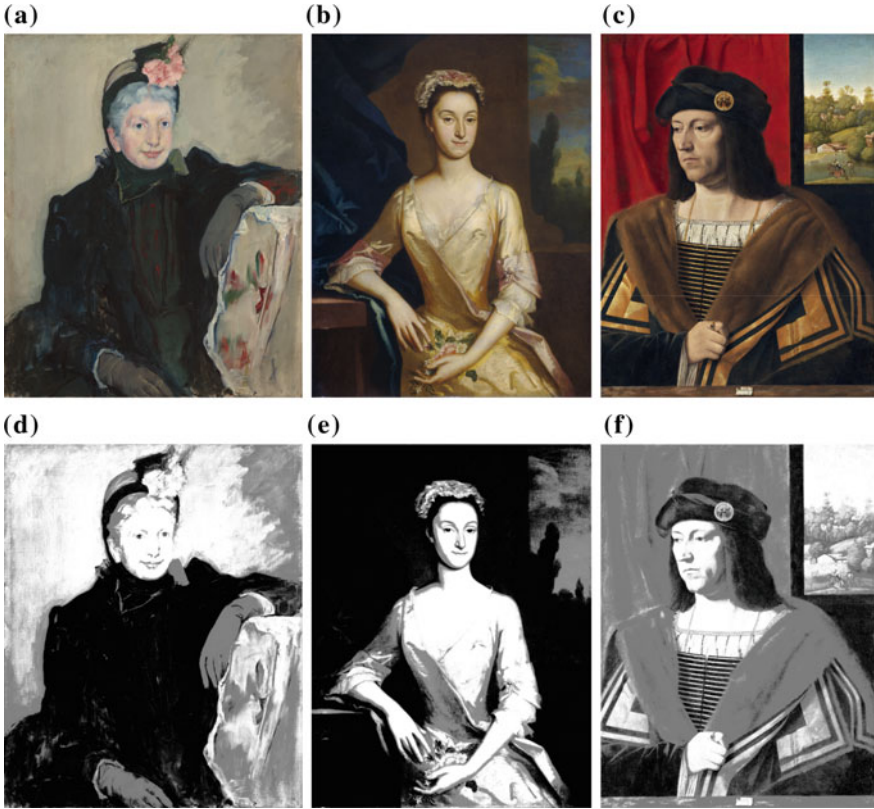


Fig. 1.1 The three tone plan types. **a** Dark against light ('Portrait of an Elderly Lady', Mary Cassatt (1844–1926), 1887). **b** Light against dark ('Portrait of a Lady', Joseph Highmore (1692–1780), 1730). **c** FG and BG of equal value ('Portrait of a Gentleman', Bartolomeo Veneto (1502–1531), 1520). **d**, **e** and **f** show their respective lightness maps. Courtesy National Gallery of Art, Washington

Spatially, portraits are very simple, being divisible into two depth planes: foreground (FG) and background (BG). Based on the contrast relationship of these regions, they can be classified into three tone plans: dark FG against light BG, light FG against dark BG, and FG and BG of equal lightness [3]. These are shown in Fig. 1.1. In each of these cases, the face is highlighted as light against dark.

Full-length portraits can be classified according to the complexity of their backgrounds. A complex background is usually, though not exclusively, associated with outdoor scenes, as shown in Fig. 1.2.



Fig. 1.2 Two full-length portraits, each employing a different style of background. **a** Featuring a simple background: ‘Portrait of the Countess of Chincon’, Francisco Goya, 1798. BY-SA Image courtesy of Rufus46 on Wikimedia Commons. **b** Featuring a complex background: ‘Portrait of Anne, Countess of Chesterfield’, Thomas Gainsborough (1727–1788), c1778. Courtesy the Getty Research Institute

1.2 The Landscape Genre

Landscape painting depicts natural scenery and generally encompasses a wide angle of view, which is broadly equitable to the field of human vision at its widest (around 90 degrees). It might include landforms, living elements upon that landform and weather-based phenomena such as sky, clouds and the sun. Typically, the spatial organisation of a landscape is more complex than that of a portrait, involving a middle ground (MG) and sky region, in addition to the FG and BG. Supporting this, the art historian Malcolm Andrews describes the landscape painter as an accomplished ‘spatialist’, with a host of pictorial tricks at his disposal [1].

In the history of landscape painting, the function of landscapes has changed significantly. Early Classical landscape paintings served as little more than backdrops to human dramas, most of which were simple morality tales drawn from the Bible or from Classical sources. They were relatively plain and frequently featured Classical ruins or buildings in the background, and figures in the foreground. With the advent

(a)



(b)



Fig. 1.3 **a** A typical landscape of the Classical style: ‘Diana and Her Nymphs in a Landscape’, Laurent de La Hyre (1606–1656), 1644. Courtesy the Getty Research Institute. **b** A typical landscape of the Romantic style: ‘Mount Corcoran’, Albert Bierstadt (1830–1902), c.1876. Courtesy National Gallery of Art, Washington. Note the difference in treatment, with the contrasts and atmospheric ambience of the Romantic painting being far more dramatic

of Romanticism (which commenced in the late 1700s), more emphasis was placed on the extremes of mood embodied by the landscape, and the narrative importance of foreground action was lessened. In effect, the landscape itself became the dramatic

subject. Serving this, the depth of the landscapes was exaggerated, and the extremes of atmosphere were heightened, mainly through the use of vivid skies and dramatic weather conditions. In a Romantic era painting, the sky is an important component of a landscape, having a substantial effect upon the emotional ambience of the scene [4]. It is frequently a ‘special case’ region, with the artist taking exceptional liberty with its colour organisation. In Fig. 1.3 is shown for comparison an example of a Classical and a Romantic landscape. The Romantic landscape is now understood as being archetypal of the genre.

The popularity of painted landscape suffered less than the painted portrait from the invention of photography. However, the form that it took underwent substantial change. From being a practice that took place in the painter’s studio, painted over many weeks on a very large canvas, it became one which took place outdoors, painted over the course of a single day on a small canvas. This form of painting is referred to as ‘en plein air’ (Fr. ‘in open air’). The foremost practitioners of this practice were the Impressionists, who favoured it for the opportunity it gave them to experiment with novel configurations of colour.

The landscapes which this book addresses are sourced mostly from the Romantic era, particularly those done by the American artists known as the Hudson River Painters. These are a group of stylistically coherent American landscapists from the late Romantic period (mid-nineteenth century). They were chosen for the heightened spatial and atmospheric effects that is a feature of their work, and also for the closeness of their tonal pallet to photographs. The average Impressionist painting (see Fig. 1.4) features very shallow tones and an unrealistic treatment of hue. This is too removed from the photographic norm to be consistently useful in our work on style transfer.



Fig. 1.4 ‘Wheatstacks, Snow Effect, Morning’, Claude Monet (1840–1926), 1891. Courtesy the Getty Research Institute. A typical painting of the Impressionist school, with shallow tones, and unrealistic treatment of hue

1.3 Concluding Remarks

Portrait painting is a relatively simple genre. Most portraits are spatially very similar, being composed of simple FG/BG pairing. The main function of a portrait is to draw attention to the sitter's face. Landscape painting is spatially more complex and less predictable. The main function of a landscape is to elicit the impression of deep space.

References

1. Andrews, M.: *Landscape and Western Art*. Oxford University Press, Oxford (1999)
2. Friedlander, M.J.: *Landscape, Portrait, Still-Life: Their Origin and Development*. Schocken Books, New York (1963)
3. Loomis, A.: *Creative Illustration*. Titan Books, London (1947)
4. Thornes, J.E.: *John Constable's Skies: A Fusion of Art and Science*. The University of Birmingham Press, United Kingdom (1999)

Chapter 2

The Colour Attributes of Paintings

Abstract In order to understand a phenomenon, one must be able to navigate it: to name its parts and to know its up from its down. Colour is a great challenge in this respect, being dimensionally very complex. This chapter examines the ways in which colour is understood in the visual art domain. Of particular focus is contrast and the many forms that it takes. Firstly, we examine RYB, HSL and other colour spaces and describe the role that they played in our work. We also detail the work of Johannes Itten (1888-1967) whose writings on colour contrast informed a lot of our research. A painter's colour-thinking is high-level, addressing the structure-based contrast properties of a painting. This structure takes many forms: per region, per object, as a whole across the entire painting, between regions and objects, etc. These we describe and exemplify. We also address the role that they played in our research. The so-called colour harmony describes as high-level principles the difference between 'good' and 'bad' colour contrasts. We detail historical and contemporary ideas on the subject, much of which has its roots in the work of early twentieth-century mystics. Using this work as a starting point, and the work of the Impressionist painters as reference, we offer our own simple rules of hue harmony.

Keywords Colour harmony · Contrast · Colour theory

2.1 Painters, Paint and Colour Systems

In the engineering domain, the image attributes of a digital colour image are generally understood through the Red, Green and Blue (RGB) values of its pixels. These are known as the photometric attributes of an image. In the art domain, photometrication is less straightforward, requiring reference to multiple systems.

Painters are particularly familiar with the physical property of paint. Paint is composed of a binder (essentially a glue), an extender (to give bulk) and the pigment. It is the pigment that supplies the colour to the paint. The name of the pigment may be derived from the name of the element or compound from which it is derived

(e.g. cadmium red), from a natural object which embodies that colour (e.g. lemon yellow) or practically anything else (such as Payne's Grey, named after the artist William Payne). Each paint has different physical properties (drying times, glossiness, etc.), different behaviour (e.g. alizarin crimson is dark when applied thickly, but lighter when applied thinly) and different cost (emerald green used to be made from emeralds).

From artist to artist, there is a huge amount of variance in what particular paints are employed. This variance is across time (some pigments being unavailable before the modern era) [12], and also differs from artist to artist [6] (e.g. some artists favour a more saturated pallet than others).

The painter organises these paints according to their position in the red, yellow and blue (RYB) hue wheel (otherwise known as the 'painter's' hue wheel). The particular spatial ordering of the colours around the RYB wheel serves as a reference point from which particular colour mixes and juxtapositions may be derived. The hue, saturation and lightness (HSL) colour space offers a different organization framework that is more perceptually-based.

Hence, an artist considers paint in a manner that is informed by chemistry, physics and physiology. Dimensionally complex indeed! Colour can also be understood in abstract terms: as a set of descriptive coordinates.

2.2 Colour Spaces

It was Sir Isaac Newton (1643–1727) who first visually organised colour as values within a circle (i.e. a 'colour wheel') [23]. A colour wheel expresses unique hues as discrete 'pie slices' within a wheel. The colour wheel is spatially limited, being unable to simultaneously accommodate a graduation from fully saturated to black as well as from fully saturated to white. Addressing this limitation, in 1810 the German artist Phillip Otto Runge (1777–1810) described in a letter to Johann Wolfgang Goethe (1749–1832) a system of defining colour as values within a sphere [25]. Runge's colour sphere (shown in Fig. 2.1) expressed variations in hue across the circumference of the sphere and variations in lightness and saturation across its latitudinal axis, with black at its north pole, and white at the south. This was the first colour space, being a volume within which colour values may be defined.

Contemporary colour spaces serve the same general purpose as Runge's, yet are organised in different ways and call upon different colour dimensions. Each being a volume, they are all employ at least three values (i.e. three-dimensional cartesian coordinates).

RYB colour space. Red, yellow and blue (RYB) is the traditional painter's colour space. In this space, colour is defined as the selective absorption (i.e. subtraction) of different parts of the colour spectrum. For this reason, it is known as a subtractive space. Though all possible colours may be in theory defined by the three subtractive primaries, in practice white and black are usually added as separate pigments. The RYB colour wheel is different to the RGB colour wheel for the colour complimentary pairs that define it. This difference is shown in Fig. 2.2 where within the RYB wheel,

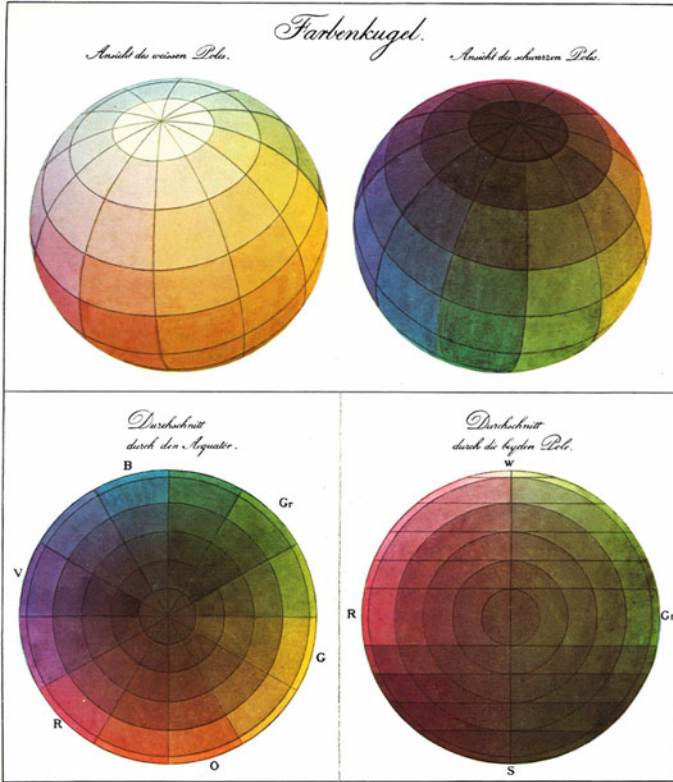


Fig. 2.1 Runge's colour sphere

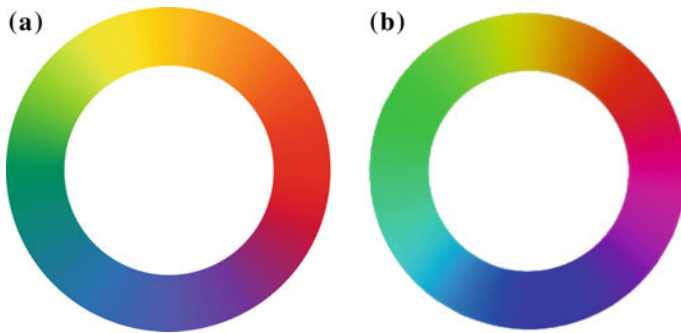


Fig. 2.2 a The RYB colour wheel. b The RGB colour wheel. Note the difference of complimentary colour pairs

the complimentary of red is green, while within the RGB wheel, the complimentary of red is cyan. The red/green, yellow/purple and blue/orange complimentary pairs are perceptually defined, being 'retto contrario' (It. 'exactly opposite'). Admixes of the RGB complimentary pairs produce neutral greys. In our work, RYB colour space is referenced when we wish to define a colour according to its perpetual complimentary.

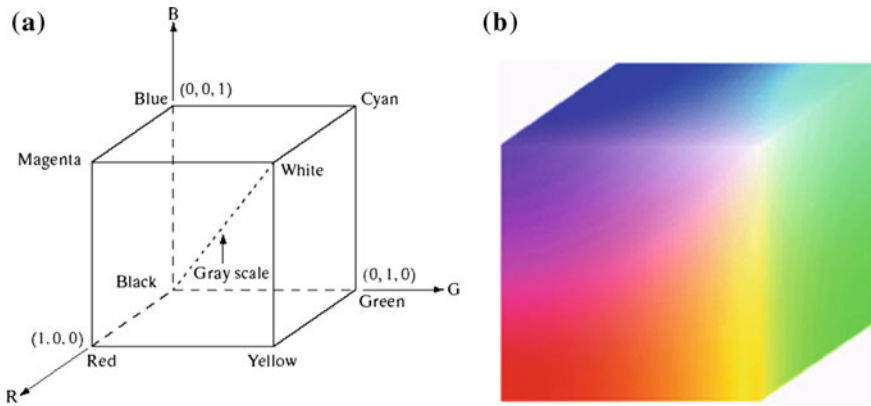


Fig. 2.3 a The RGB colour model (the figures are from [13]). b The RGB 24-bit colour cube

RGB colour space. Red, green and blue (RGB) is an additive colour space, wherein different colours are defined by different admixtures of light. This colour space is used in digital imaging and image display devices. A variant of RGB space, known as sRGB, is used in many digital image formats. The colour model is shown in Fig. 2.3. Human vision employs a colour model similar to RGB [14] and is defined by the relative sensitivity of eye cells to these hues. Despite this, RGB is not suitable for describing colours for the purpose of human interpretation. For that, HSL or HSV colour spaces are far more suitable. In our work, images are read-in and written-out in RGB colour space.

HSL colour spaces. Hue, saturation and value (HSV) or Hue, saturation and lightness (HSL) are the three distinct perceptual attributes of colour, with hue and saturation being together its chromatic component. They are far more human-readable than RGB or any other colour model [18] and are the standard point of reference for artists and designers when they conceptualise their colour thinking. They are particularly suitable for digital artists, as the HSL values may be digitally ‘dialled up’ or ‘dialled down’, something that is not possible in traditional painting.

- **Lightness** is the perceived intensity of a colour. It occupies a uniquely important position in the colour triumvirate, being self-sufficient from chromaticity. In other words, it is easy to imagine an image without a chromatic component, but practically impossible to imagine the inverse. Lightness is measured in degrees of intensity and is bound at each end by zero and maximum.
- As lightness is an expression of a colour’s intensity, **hue** is the specific wavelength of that light. Hue is expressed as terms: words that define a particular range of the visible electromagnetic spectrum [22]. The hue range extends from purple (short wavelength) to red (long). The span of hue difference is navigated according to perceptually unique hues, which are red, green, blue and yellow. Hue has a strong semantic dimension. For example, a tree should not be blue and a sky should not be green.



Fig. 2.4 **a** Jean-François Millet, ‘Portrait of Louise-Antoinette Feuardent’ (1841). Courtesy the Getty Research Institute. **b** A normalised map of its saturation. **c** A normalised map of its lightness. *Note the similarity of these maps*

- **Saturation** defines the purity of the hue [13]. Similar to lightness, saturation is measured as degrees of intensity, bound at each end by zero and maximum. Zero saturation nullifies chromaticity, rendering it achromatic.

The way that these values interact within a painting is noteworthy. In most paintings made before the modern era (which can be approximately defined as starting in the 1860s), the mapped saturation values look very similar to those of the lightness [7]. In Fig. 2.4 a painting by the French artist Jean-François Millet (1814–1875) is shown together with its lightness and saturation maps. Their similarity is clearly apparent.

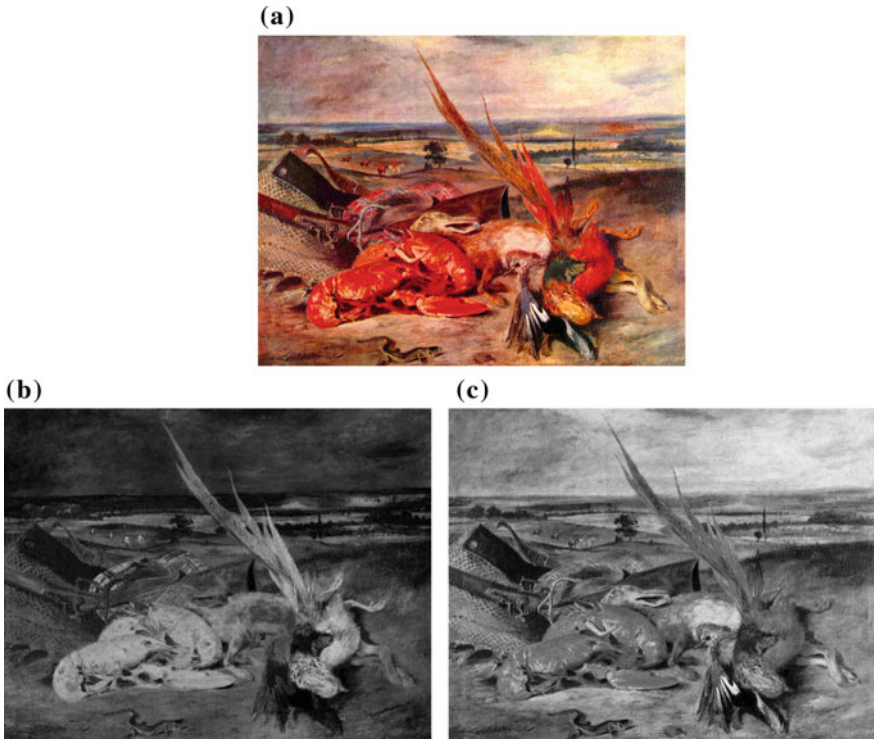


Fig. 2.5 **a** ‘Still Life with Lobster’, Eugène Delacroix, 1826. Courtesy of The Yorck Project, Wikimedia Commons. **b** A normalised map of its saturation. **c** A normalised map of its lightness. *Note* the dramatic difference between these two maps

In 1839, Michel-Eugène Chevreul (1786–1889) wrote ‘The Laws of Contrast of Colour’ [5]. Though it was intended for use by the design and printing industries, it had a significant impact upon the way that painters thought about colour. Taking the lead in this new thinking, Eugène Delacroix (1798–1863) was the first to knowingly and aggressively unlink saturation from lightness. Fig. 2.5 shows a painting by Delacroix together with its lightness and saturation maps. There is strong difference between their perceptual impact, the saturation map being very eventful, while the lightness map is evenly distributed. This difference is further discussed in Sect. 2.3.2 on p. 18.

Mapped as colour spaces, HSL and HSV form cylinders. The models of these two colour spaces are shown in Fig. 2.6.

In our work, hue is sourced by converting to HSL/HSV colour model. We also used HSL/HSV as our ‘thinking space’, for the discussion of the aesthetic values of an image.

Munsell, CIELab and LCH colour spaces. It excited eighteenth-century colour theorists, such as Johann Wolfgang Goethe, that according to Runge, colour could be expressed as a sphere. This was in agreement with their understanding of the universe

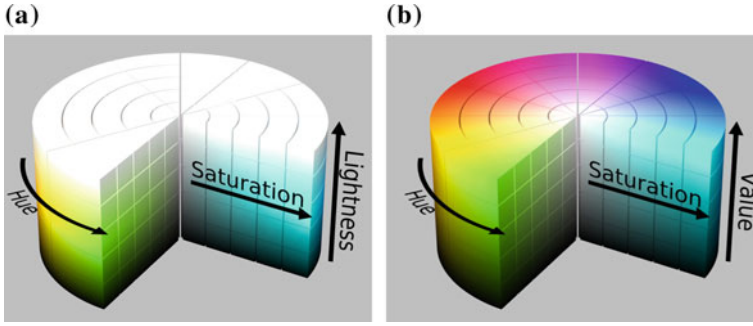
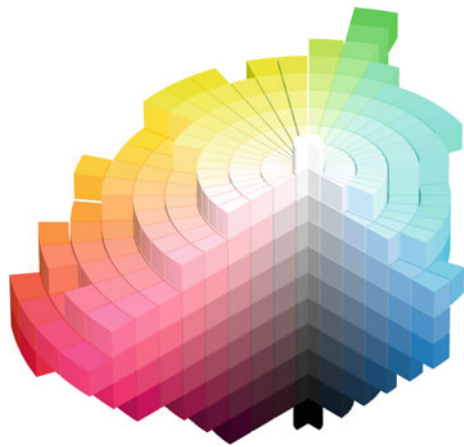


Fig. 2.6 HSL and HSV cylinder models (image from Wikimedia Commons). **a** The HSL cylinder. **b** The HSV cylinder

Fig. 2.7 Munsell’s colour space visualised. Courtesy Wikimedia Commons



as being a layered and ordered thing, with God placed firmly in the centre (the so-called ‘harmony of the spheres’). It took another artist, Albert Henry Munsell (1858–1918), to prove how wrong this assumption was. In 1905, he developed the Munsell colour space [21]. It was painstakingly mapped as degrees of ‘barely noticeable difference’ using human perception as the point of reference. Munsell showed that far from being a regular form, when mapped as a volume, colour was extremely uneven (see Fig. 2.7). For example, there are more perceptually different degrees of ‘greenness’ than there are of ‘yellowness’.

In mapping the hues around the vertical axis of the space, Munsell arranged them such that any n° spread of hue values in one part of the wheel contains the same number of perceptually different hue values as those from another part of the wheel with the same degree of spread. Shown in Fig. 2.8 is the Munsell hue wheel (a) compared to an RGB hue wheel (b). Note how the green span within the Munsell hue wheel is larger than that within the RGB wheel. This corresponds to the human experience of green, which exists in more perceptually different shades than any other

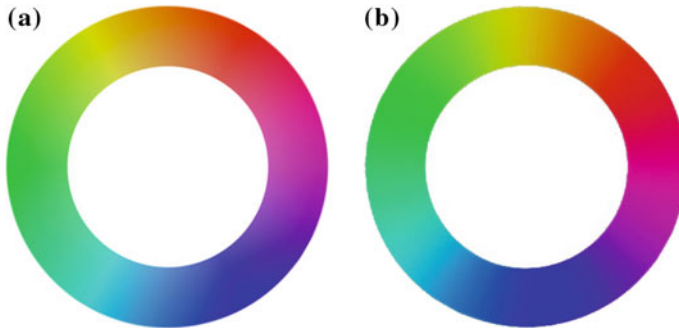


Fig. 2.8 **a** The Munsell colour. **b** The RGB colour wheel. Note the larger spread of the green span in the Munsell wheel

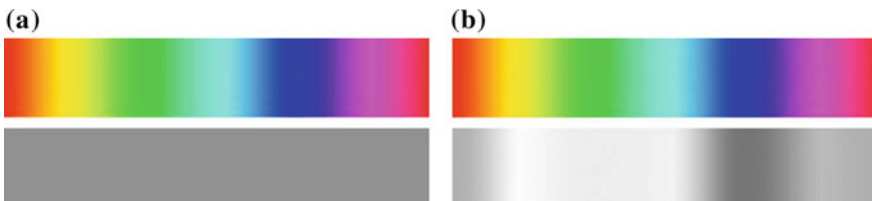


Fig. 2.9 The difference of the L channel in HSL and CIELAB spaces. **a** The hue ramp and L channel of HSL space. **b** The hue ramp and L channel of CIELAB space

hue. Munsell colour space was used in our work in the defining of global hue contrast (Chap. 10 of 179).

The CIELAB model is a direct descendant of Muncell colour space. It presents lightness in the L channel and colour in the a and b channels. The L channel of the hue ramp matches our perception of lightness (as shown in Fig. 2.9b). One of the features of CIELAB space is that saturation and hue are not located in either the a or b channels but are expressed as relationships between these two channels. In order to be expressed as hue, saturation and lightness, the CIELAB space is converted to LCH (lightness, chroma, hue) space by transforming the Cartesian a-b coordinates of CIELAB to the polar C-H coordinates of LCH. Chroma and saturation are related for expressing the colourfulness of a colour. In our study of depth-aware contrast, regional contrast and center-corner contrast, the LCH colour space is employed. The chroma in LCH space is indicative of saturation.

2.3 Colour Contrast

Simply put, contrast is a property of difference. In order for an object to be perceptually apparent, it is required to contrast its environment. This will likely be a contrast of lightness, but may also be a contrast of hue and/or saturation.

In visual art, the aesthetic organization of an image is managed through attention to its contrasts. These contrasts may involve any perceivable attribute of the image: theme, texture, size, etc. This is a fundamental principle of all art objects. For example, in an action movie, the drama is driven by the contrast of good and evil. It is the contrast of colour which this section addresses, and which is of prime importance in our work.

2.3.1 *Itten's Contrasts*

In 1961, Johannes Itten (1888–1967) wrote ‘The Elements of Colour’ [16], in which he described seven colour contrasts. This book can be found in most art school libraries and has had a huge influence upon the way that colour is understood in the art domain.

1. **Contrast of Saturation.** This refers to the contrasting of regions of different saturation values.
2. **Contrast of Light and Dark.** This refers to the contrasting of regions of different lightness values.
3. **Contrast of Extension.** This refers to the way in which two or more regions, each of different colour, are effected by size. For example, a saturated region of the same size as an unsaturated region will perceptually overwhelm that region. Their relationship may be brought into balance by increasing the size of the unsaturated region.
4. **Contrast of Complements.** Complementary colours are pairs of colours that are located on opposite sides of the colour wheel. Examples of complementary RYB pairs are red/green, yellow/purple and blue/orange.
5. **Simultaneous Contrast.** First described by Chevreul, Simultaneous Contrast refers to a perceptual effect whereby one region of colour may change the apparent colour properties of another region. Hence, a mid grey against a dark grey will seem lighter than the same mid grey against a light grey.
6. **Contrast of Hue.** This refers to the amount of perceptually different hue values in an artwork. Simply put: an artwork with few perceptually different hue values has low Contrast of Hue, while one with many perceptually different hue values has high Contrast of Hue.
7. **Contrast of Warm and Cool** (aka colour temperature). This contrast assumes that some colours are perceived as warmer/cooler than others. Perversely, Itten describes two contradictory forms of temperature contrast: relative (such as the warm yellow of a sunflower versus the cold yellow of a lemon) and absolute. The latter divides the RYB hue wheel in half along the green/purple axis. Colours in the yellow to red-violet half are warm, colours in the other half are cool.

The first focus of our research served to examine how the contrast values of an artwork may be quantified. The second focus was to examine how these contrast values may be transferred to photographs (so-called style transfer).

Itten's contrasts are used as a point of reference in much of our work. However, not all of his contrasts were applicable to our purposes. We assume that a style transfer will only effect the colour of pixels, not their placement. For this reason, the Contrast of Extension was not included in our study because it is primarily a contrast of size, not colour. We also did not consider Simultaneous Contrast to be, in the strict sense, a colour contrast. Rather it is a perceptual phenomenon (an optical illusion). Additionally, it is also not a major feature of art or photography, except in late modernism such as exemplified by the work of the Op Artists [8].

Though we use Itten as a point of reference, he is not without his failings. There are many errors and inconsistencies in his work. For example, he describes yellow as being 'lighter' than blue. This is clearly erroneous, as any hue has the potential to be lighter or darker than any other hue. However, he did an effective job at gathering known wisdoms of the time and remains a point of reference for engineers, seeking canon literature on the topic of pictorial aesthetics [9, 10, 24].

2.3.2 Structural Contrasts

Itten's Elements of Colour is a collection of general observations describing the basic dimensions of colour contrast. They do not on their own address how the contrast of a painting is structured. This section addresses six structure-based contrasts: global contrast, center-corner contrast, local contrast, regional contrast, neighbouring regional contrast and depth-aware contrast. The notable thing about these contrasts is that they all serve distinctly different aesthetic functions within a painting.

1. **Global contrast.** Global contrast is a property of the entire image. A histogram is a visual expression of the global distribution of lightness or saturation within an image. Shown in Fig. 2.10 are two paintings, together with their lightness histograms. The chromaticity has been nullified in both paintings to make their lightness values easier to review. The first extends from black to white, with a predominance of darker values. The second extends from dark grey to light grey, with a predominance of lighter values.

The tonal range and distribution that Van Gogh employed is known by artists as *high key* and that employed by the unknown artist is known as *low key* [3]. Global contrast defines the primary impact of this range and distribution. This impact comes before any high-level consideration of content or structure, and may therefore be said to be a reflection of the *painting as an object*.

It can be simply observed that artists exaggerate upon nature. This is as true for painters in the handling of contrast as it is for writers and film-makers in the extreme fictions they craft. However, this exaggeration is not simply a linear increase (i.e. 'the more the better'). The fugitive nature of skies, and their importance in establishing the emotional mood of a landscape, legitimatises any creative liberties that the artist wishes to take with them. This makes them suitable case studies when examining how artists exaggerate contrast. In Fig. 2.11 is shown the RMS contrast value of 30 skies (RMS contrast is described in Sect. 3.2 on p. 34). These are sourced from the work of the Hudson River school of painters. Clearly, many of the paintings are more



Fig. 2.10 a ‘Portrait of a Young Lady’ (by an unknown artist, c1835) (left). b Van Gogh’s ‘Girl in White’ (1890). Courtesy National Gallery of Art, Washington. Inset their lightness histograms. Note the difference in their lightness distribution, which is a difference of global contrast

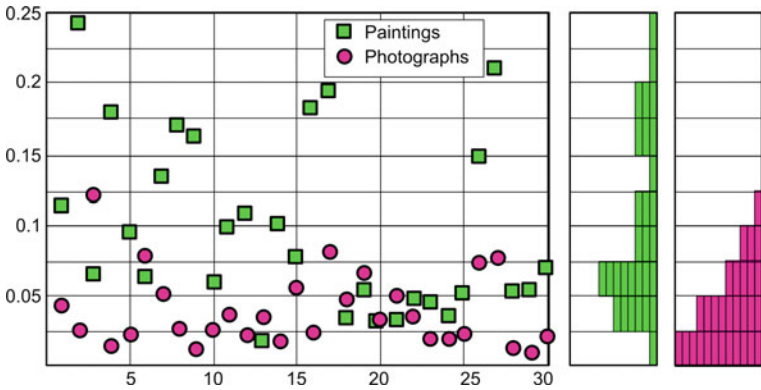


Fig. 2.11 The RMS contrast value of 30 skies, sourced from the Hudson River school of painting, compared to that of skies sourced from 30 photographs. Distribution histograms are on the right. The difference in histograms evidences the artist exaggerating upon nature

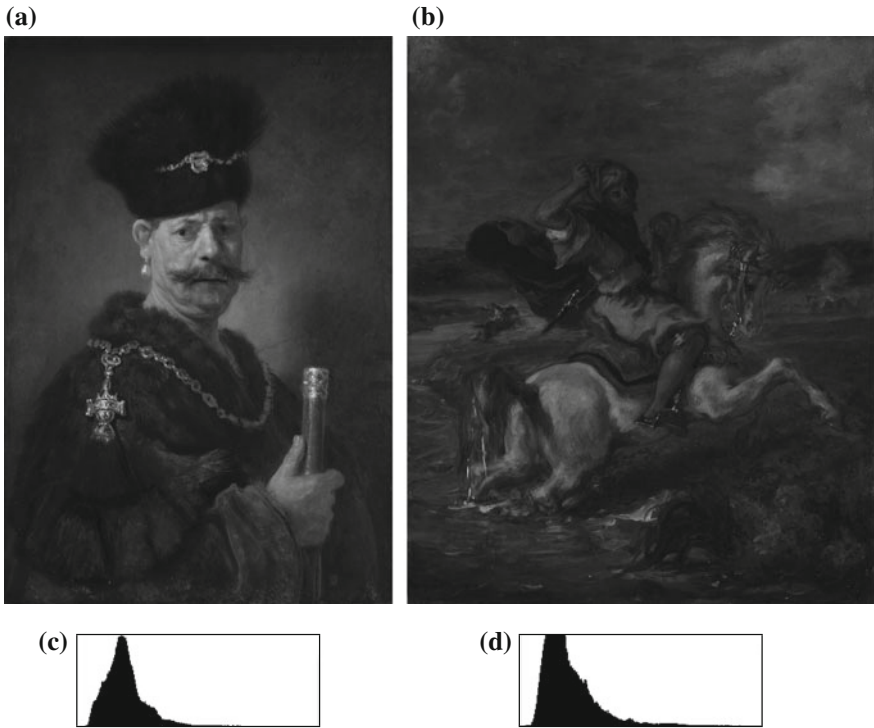


Fig. 2.12 **a** ‘A Polish Nobleman’, Rembrandt Van Rijn (1606–1669), 1637. Courtesy National Gallery of Art, Washington. **b** ‘Moroccan Horseman Crossing a Ford’, Eugeène Delacroix, (1798–1863), 1850. Courtesy the Getty Research Institute. **c** and **d** are their histograms. Note the general similarity of their histograms, yet the difference of their perceptual impact. This is the result of the difference in the spatial distribution of their values, with those of ‘Moroccan Horseman Crossing a Ford’ being even-form, and ‘A Polish Nobleman’ being event-form

contrasty than the photographs. However, many of the paintings are also of very low contrast. Hence, the exaggeration employed by painters is one where similarity is as likely to be exaggerated as difference. In the words of the noted art theorist, Rudolph Arnheim: ‘...a configuration of colours will strive either toward contrast or toward assimilation’ [27].

The shortcoming of employing a histogram as a means of evaluating global contrast is that it lacks a spatial dimension. In Fig. 2.12 two paintings are presented: one by Rembrandt, the other by Delacroix. Despite the similarity of their histograms, these paintings are visually dissimilar in their perceptual impact. The lightness of the Delacroix painting varies evenly across the surface (even-form), while that of the Rembrandt painting is clumpy and focused around the *event* of the face (event-form). This exemplifies the spatial component of global contrast.

This difference between even-form and event-form global contrast is particularly important when comparing the global saturation of an image with its global lightness.

It can be seen in Fig. 2.5 that the painter has enforced a strong difference between the spatial distribution of these two values. This even/event-form difference between the saturation and lightness distribution is a distinguishing feature of impressionist and post-impressionist painting and has had a huge impact upon our contemporary visual aesthetic.

In Chap. 3 on p. 33, computational models of colour contrast are discussed. However, there is currently no computational model we are aware of that satisfactorily differentiates between even-form and event-form global contrast.

2. **Local contrast.** Global contrast varies across contiguous regions of a painting, giving us local variations of global contrast, which we term local contrast. In a typical landscape painting, the lightness values of the foreground will range from black to white, while in the background they will range from dark grey to light grey. In the painting shown in Fig. 2.13 the lightness values of the foreground region extend from a near black to a near white, while those in the background region range from dark grey to light grey.

In a landscape painting, such differences in local contrast are used to enhance the impression of space. In a portrait painting, they are used to draw attention of the viewer to the face.

3. **Centre-corner contrast.** Centre-corner contrast is the contrast between the edges and corners of a painting and its approximate centre [3]. The centre and the corners of a painting are attributes of its picture plane, which is shown in Fig. 4.1. Centre-corner contrast can be found in both landscapes and portrait paintings and is an important component of the vignette effect, which is discussed in Chap. 4, on p. 39. The key function of the vignette and centre-corner contrast is to enhance the depth of the painting and to draw attention to the ‘region of interest’ (see Sect. 4.1.2 on p. 44).

4. **Regional contrast.** Regional contrast is the contrast between two colours and is therefore an entirely relative value (i.e. a may be lighter than b but darker than c). For an artist, relative judgements have substantial value in the process of painting during which an artist will evaluate the colours of specific regions not just for their fidelity to those of the subject, but also for their difference to other regions [2]. In other words, the artist seeks to be faithful to the relative organisation of the colours in a scene, not their absolute value. Local contrast relies upon pairwise comparison, which is the most efficient manner in which a human can make evaluative judgements [26]. Staggered sets of pairwise comparisons are employed by painters when they are establishing the colour values of their paintings. In the pictogram shown in Fig. 2.14, the piece of paper under the table is *darker* than that which is on top of the table, yet *lighter* than the floor under the table. This pairwise ordering emulates a simple sorting algorithm (*a comparison sort*), using which an artist sorts the colours within the painting into an agreeable configuration.

5. **Neighbouring regional contrast.** As regional contrast is relative, it cannot be employed to evaluate the aesthetic value of a painting without first defining the terms by which the two compared regions are selected. It is therefore hard to address in a general manner. However, there is a subset of regional contrast that exclusively concerns neighbouring regions (see Fig. 2.15). This contrast serves to attract our eye



Fig. 2.13 a ‘Pastoral Landscape’, Asher Brown Durand, (1796–1886), 1861. Courtesy National Gallery of Art, Washington. b Regions within the FG and BG shown in comparison with histograms inset. The tones of the FG range from *black* to near-*white*, whereas those of the BG extend from *light grey* to *dark grey*. This is a difference of local contrast, and in landscape painting is used to enhance depth

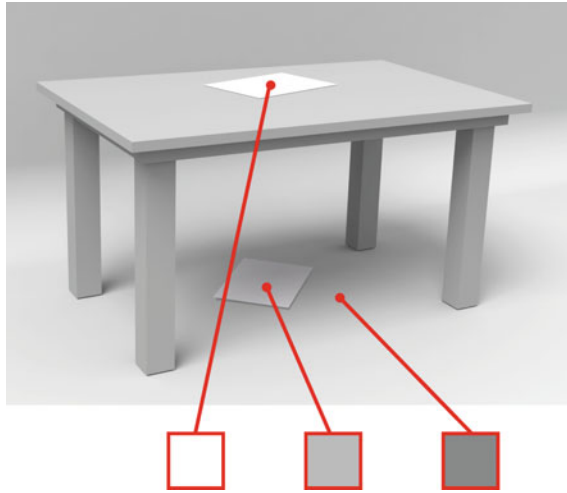


Fig. 2.14 In this pictogram, the piece of paper under the table is *darker* than that on *top* of the table, yet *lighter* than the floor. This is an example of the pairwise sorting algorithm that a painter makes in the process of painting

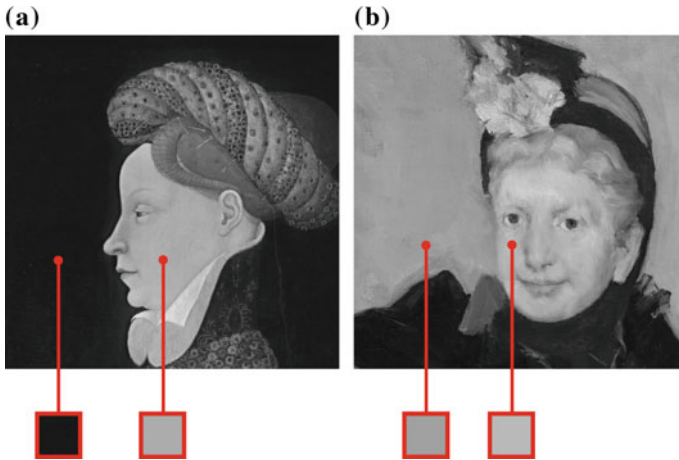


Fig. 2.15 **a** Detail of ‘Profile Portrait of a Lady’, Franco-Flemish, c1410. This is an example of high neighbouring regional contrast of the face against the background **b** Detail of ‘Portrait of an Elderly Lady’, Mary Cassatt (1844–1926), c1887. This is an example of low neighbouring regional contrast of the face against the background. Courtesy National Gallery of Art, Washington

to the centre of attention and is especially important in portrait painting, where it is employed to emphasise the head against the background.

6. Depth-aware contrast. In both landscape and portrait paintings, there is geometric organisation that supports the illusion of depth. Real-world depth exists as infinite degrees of difference between the viewer and the furthest object in the scene.

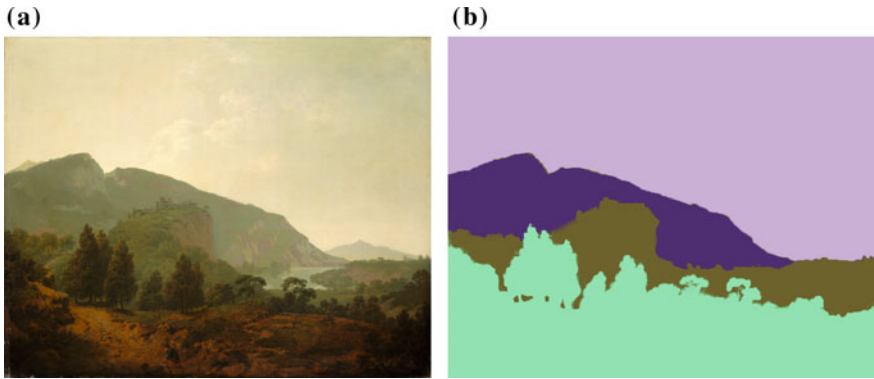


Fig. 2.16 **a** Joseph Wright (1734–1797), ‘Italian Landscape’ 1790. Courtesy National Gallery of Art, Washington. **b** Its depth planes shown as coloured regions

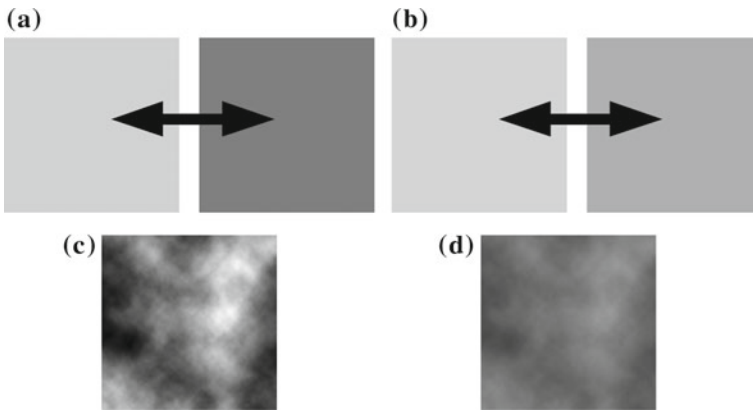


Fig. 2.17 **a** High inter-contrast. **b** Low inter-contrast. **c** High intra-contrast. **d** Low intra-contrast

However, artists generally render depth as a few simplified planes. A portrait tends to employ no more than two depth planes: a FG and BG (i.e. figure/not-figure). A landscape may employ more depth planes, but usually no more than four (FG, MG, BG and sky, see Fig. 2.16). The depth-aware contrast of a painting is the contrast organisation that exists *within* the depth planes, and *between* the depth planes. These we term, respectively, ‘intra-contrast’ and ‘inter-contrast’ (see Fig. 2.17).

In landscape painting, the maintenance of the illusion of depth is of high importance, whereas in portrait painting, the expression of the figure and the face is prime. The figure in a portrait painting inhabits the middle vertical, while the background often serves only to provide a high contrast to the figure. Therefore, this book focuses on studying the depth-aware contrast organisation of landscapes and the regional and neighbouring local contrast organisation of portraits.

2.3.3 Hue and Harmony

The preceding section described the various dimensions of contrast in a painting, which on their own do not describe an aesthetic evaluation of an image. Such is offered by colour harmony, which can be understood as describing *good* contrast. Since painting was first written about, there has been a declared assumption that some configurations of colour are better than others. However, defining exactly what constitutes this quality is a task that is still a work in progress, and one that is perhaps beyond reach.

Frequently cited on this topic is the notable work done by Matsuda. For a number of years he performed field studies on the fashion habits of female Japanese university students. He proposed a set of hue and tone templates (Figs. 2.18 and 2.19), the hue templates employing the RGB hue wheel. Any hue or tone/saturation configuration that conforms to the shaded regions of these templates is assumed to be harmonious.

Though these are frequently cited in reference to paintings and photographs [19, 24], their value outside of the domain of flat design is questionable. The ‘i type’ hue template would correspond to a painting of a single hue, something not found outside of late modernism. Similarly, ‘I type’, ‘T type’ and ‘X type’ configurations in painting are very rare. These templates are also very regular, while the hue histograms of paintings are typically very noisy and irregular.

Matsuda’s tone templates are similarly inappropriate for the aesthetic evaluation of painting. Almost all pre-modern painting features lightness (tone) that extends from black to white (an exception is shown in Fig. 2.10). This also is true of photography, and is the working assumption that is made by Photoshop’s ‘Auto Correct’ adjustment. Hence, only the ‘Value contrast’ and ‘Maximum contrast’ template would be applicable to paintings. We therefore observe that any consistent correspondence

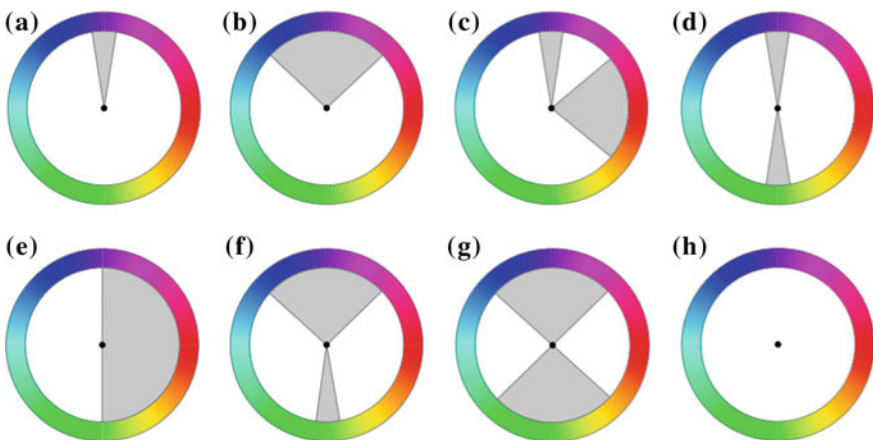


Fig. 2.18 Matsuda’s hue templates. **a** i type. **b** V type. **c** L type. **d** I type. **e** T type. **f** Y type. **g** X type. **h** N type

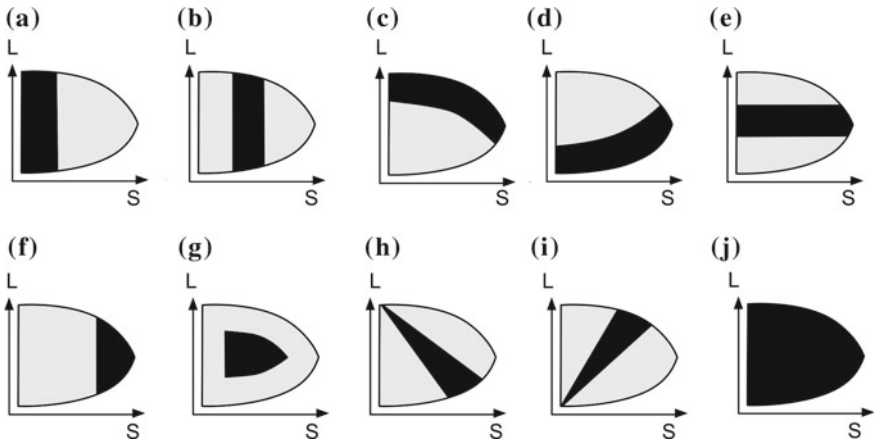


Fig. 2.19 Matsuda's tone templates. **a** Value contrast. **b** Medium chroma and value contrast. **c** Upper arc contrast. **d** Lower arc contrast. **e** Chroma contrast. **f** High chroma and equivalent value harmony. **g** Neutral tint and equivalent value harmony. **h** Triangle contrast. **i** Inverted triangle contrast. **j** Maximum contrast

between Matsuda's templates and the colour organisation of paintings is probably wishful thinking.

Adobe's advanced colour harmony picker 'Color CC' works in RYB colour space. Drawing from traditional literature on the topic, it references the following classic colour schemes:

- **Analogous:** colours chosen for their similarity of hue.
- **Monochromatic:** colours chosen for their identical hue (variance is supplied through differences of lightness).
- **Triadic:** three colours located at points evenly dispersed around the hue wheel.
- **Complementary:** colours chosen for their opposition on the RYB hue wheel.
- **Compound:** sometimes also called the split-complementary, this is a variant of the complementary wherein one of the complementary pairs is an analogous spread.

There is some correspondence between Adobe's system and Matsuda's. For example, Matsuda's *i* type is a good match for the Monochromatic, and the *V* type for the Analogous. However, there are also significant dissimilarities. There is no match for the Triadic in Matsuda's schemes, and Matsuda and Adobe reference different hue wheels. In this respect, we must remark that we find Matsuda's use of the RGB wheel curious, as this colour wheel has no value in canon art literature, where the RYB wheel is the norm for evaluation of complementary pairs. The general impression that is given is that both Matsuda and Adobe are recalling the same poorly remembered event in a manner that is fuzzy and inconsistent. This we find to be symptomatic of artistic colour theory in general. Supporting this, Burchett [4] and Westland [28] both found that despite a plethora of enquiry into the topic, which extends over many centuries, there is no general agreement as to exactly what colour harmony *is*.

Even Goethe's massive tome on so-called colour theory [27] is, by most informed accounts, nothing of the kind, being more the work of a poet than a fully-fledged theory [14].

These criticisms notwithstanding, Matsuda's and Adobe's systems reflect a general agreement that colour in art and design operates in a manner that is, in some mysterious way, subject to high-level ordering principles.

The subject of colour harmony is most often addressed in relation to the hue component of colour. Hue is very difficult to exactly quantify. The lightness and saturation of an image are understood through such values as least, most and average. However, the human experience of hue is defined as differences in *terms* (red, yellow and green). The word term was first used in this context by the perceptual psychologist Nault [22]. It was he who demonstrated that children, when reading maps, generally associate change in lightness and saturation with change in amount, such as depth and degree. However, change in hue they associate with a change in quality, such as the difference between water, land and coral. Consequent to this, no hue is inherently greater than another. Nor, in a hue event as complex as a painting, can there be such thing as an average hue, any more than there can be an average nationality in a room full of people. In other words, hue is dimensionally complex: not as easy to conceptualise as saturation or lightness.

Of pivotal importance in most colour harmony literature is the notion of complementary pairs [2, 16, 27]. In physics, a complementary pair is one that when mixed together produces a neutral (i.e. a grey). This mixing may be optical, such as Maxwell's discs [20], or through admixtures of paint. For artists, a far more important property of a complementary pair is that they are as different to each other as can possibly be. This property has little empirical substance, maximum perceptual difference of hue being difficult to measure, but it nonetheless has a long history in the arts. In his book 'The Art of Painting', Leonardo Da Vinci (1452–1519) describes the colour pairs blue/orange and red/green as being opposite (*retto contrario*, literally 'exactly opposite') [11]: a clear reference to artist's complementary colours. This predates any conceptualisation of the colour wheel.

Harmony in the visual arts is a notion that was originally borrowed from the world of music, which had long recognised that certain configurations of musical notes seem more agreeable to the ear than others. The first notable proponent of colour harmony was Johann Wolfgang von Goethe (1749–1832), a poet and philosopher who, in his 'Theory of Colour', saw himself as an antidote to Newton and his empiricism. Goethe placed the physiological experience of colour above its physical nature. The first time he mentions colour harmony is in reference to the complementary pairs:

...the eye especially demands completeness, and seeks to eke out the colourific circle in itself. The purple or violet colour suggested by yellow contains red and blue; orange, which responds to blue, is composed of yellow and red; green, uniting blue and yellow, demands red; and so through all gradations of the most complicated combinations. /When in this completeness the elements of which it is composed are still appreciable by the eye, the result is justly called harmony.

Goethe is simply making the case that the aesthetic appeal of the complementary pair is one of ‘completeness’: that all three primaries are *implied* by a complementary pair (the secondary component of the pair being itself a mix of two primaries). This idea clearly excited him, and his book is full of assertions that any one hue cannot be experienced in isolation: that it naturally gravitates towards this completion which he terms harmony.

In his writing, Goethe constantly positions colour and colour harmony as something that reflects a larger universal order. Throughout the history of its usage, the word harmony is never far from such a spiritual reading. Pythagoras (c. 580–500 BC) described harmony as a mathematical order that found reflection in the order in which the planets crossed the heavens (the so-called harmony of the spheres). It is perhaps for this reason that Goethe’s work resonated greatly with early European modernists, many of whom were proponents of what we would now term ‘new age’ religions. The noted colourist, Wassily Kandinsky (1866–1944) was a Theosophist, and Itten himself was a proponent of Mazdaznanism, which was a fire cult. Goethe was a huge influence of Theosophy (H. Blavatsky) and Anthroposophy (R. Steiner), both of which are cultish in their outlook. They all wrote about colour in a manner that bordered on the religious. For example, in ‘Concerning the Spiritual in Art’ Kandinsky remarked that ‘orange is red brought nearer to humanity by yellow’ [17]. Goethe described the colour ‘yellow-red’ as being a favourite of ‘savage nations’ and ‘impetuous, robust and uneducated men’ and Itten himself describes the colours of a stained glass window as offering a ‘...direct invitation to higher spirituality’.

In summary, colour harmony has a lot of dubious baggage that is more suited to the religious domain than the creative or empirical. If colour harmony is a consistent and prescriptive system, then predictions could be made of it. For example: given the same problem (e.g. an incomplete painting), different artists would produce the same solution. We have seen no evidence that this is possible. Despite this, some general truths on the aesthetic application of hue in painting can be stated. These general rules of avoidance are detailed in Table 2.1. Importantly, these are more efficiently stated not as prescriptions (as in Matsuda and Adobe), but as avoidances. The difference between a prescription and an avoidance is interesting. Though ‘follow the path’ may seem to casual observation to be effectively the same as ‘keep off the grass’, in practice avoidance offers greater freedom of choice (e.g. ‘keep off the grass, but otherwise do as thou wilt’). It also allows for apparently contradictory directives to be observed.

Each of these three rules of avoidance we illustrate using examples from the Impressionist or Post-Impressionist era. These artists were famous for their rule-challenging experiments with hue and therefore present a limit-case test bed of these general truths. These rules are further substantiated by our work on hue contrast, as described in Chap. 10 on p. 179.

Table 2.1 Hue contrast in paintings: three general rules of avoidance

Rule	Description
Avoidance of hue imbalance	Very few paintings employ hue values from only one half of the RYB hue wheel. Related to this is the rarity of monochrome paintings (i.e. paintings that employ only one predominate hue). Of 100 Post-Impressionist paintings reviewed, only six featured hue histograms that did not spread further than 180 degrees around the RYB hue wheel. An example is shown in Fig. 2.20, together with its hue values expressed within a RYB hue wheel. It can be seen that all its hues are variants of yellow
Avoidance of hue balance	Very few paintings employ equal amount of hue values from one side of the RYB wheel as from the other. Of 100 Post-Impressionist paintings reviewed, only five obviously featured such balance. An example is shown in Fig. 2.21, together with its hue values expressed within a RYB hue wheel. This principle corresponds effectively to the original ‘rule of thirds’ as first stated by Sir Joshua Reynolds’s (1723–1792): ‘...in colours, whether of the warm or cold kind, there should be one of each which should be apparently principal and predominate over the rest’. (see Sect. 4.1.1 on p. 41)
Avoidance of hue ubiquity	Very few paintings employ every single hue. An exception is shown in Fig. 2.22, together with its hue values expressed within a RYB hue wheel. This can be evidenced by performing a Google image search on the term ‘garish colours’ (garish being a ‘taste’ word with pejorative connotations [15]). The results heavily feature ‘rainbow’ images which employ every single hue in the spectrum. Of 100 Post-Impressionist paintings reviewed, 15 displayed full hue ubiquity. Of these, seven were of the pointillist school, which was a group of painters that was defined by this very practice

(a)

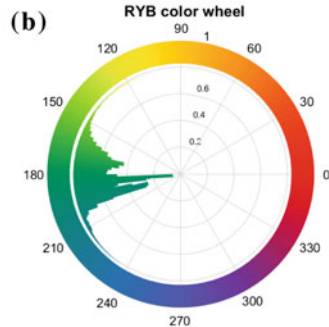


Fig. 2.20 a ‘Roses’, Van Gogh (1853–1890), 1890. Courtesy National Gallery of Art, Washington. b Its hue values expressed within a RYB hue wheel, showing hue imbalance: all hue values laying within one half of the hue wheel

(a)

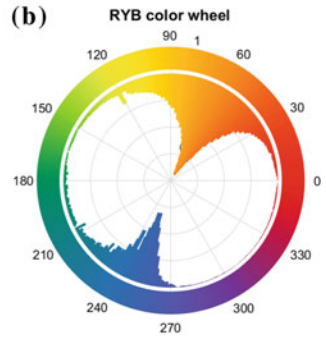


Fig. 2.21 **a** ‘Fisherman’s House at Varengeville’, Edward Monet (1840–1926), 1882. Courtesy The Yorck Project, Wikimedia Commons. **b** Its hue values expressed within a RYB hue wheel showing hue balance: hues evenly distributed into two opposing masses of similar span

(a)

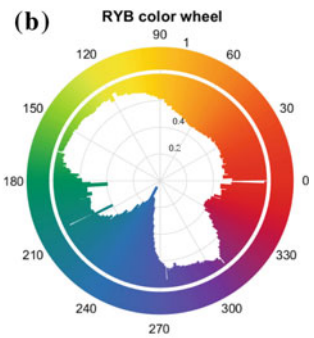


Fig. 2.22 **a** ‘The Bay’, Paul Signac (1863–1935), 1906. Courtesy Szilas, Wikimedia Commons. **b** Its hue values expressed within a RYB hue wheel, showing hue ubiquity: hues in the spectrum having been employed

2.4 Concluding Remarks

This chapter described the many ways that colour in the art domain may be expressed. Clearly, it is a multidimensional phenomenon, with no single approach being applicable in all circumstances. We discuss colour spaces as different ways in which to model colour as a volume, each with their own particular application. We address Itten’s contrasts, which are a collection of general observations that describe the basic

dimensions of colour contrast. We present six structure-based contrasts: global contrast, centre-corner contrast, local contrast, regional contrast, neighbouring regional contrast and depth-aware contrast. Each of these serves a particular aesthetic function within a painting. We conclude with an examination of the history and practice of so-called colour harmony. Though we doubt that such a thing can be exactly codified, we nonetheless propose three general rules of artistic hue contrast. These are framed as avoidances rather than prescriptions.

This work serves as foundation for our research on transferring the style from paintings to photographs.

References

1. Arnheim, R.: *Art and Visual Perception* Univ. of California Pr. (2004)
2. Arnheim, R.: *Visual Thinking*. University of California Press, Berkeley (1969)
3. Barnard, G.: *The Theory and Practice of Landscape Painting in Water Colours*. W. S. Orr and Company (1855)
4. Burchett, K.: *Color Harmony Color Research & Application* (2002)
5. Chevreul, M.E.: *De la loi du contraste simultané des couleurs et de l'assortiment des objets colors*. Translated into English by Charles Martel as the principles of harmony and contrast of colours Longman, Brown, Green, and Longmans (1855)
6. Constable, M.: <http://opticalenquiry.com/paint/?p=224>
7. Constable, M.: *Contrast Turps Banana*, London, U.K. **14**, 50–57 (2013)
8. Constable, M., Zhang, X.: Depth-based analyses of landscape paintings and photographs according to Itten's contrasts. In: *Proceedings of the 4th Pacific-Rim Symposium on Image and Video Technology*, pp. 481–486 (2010)
9. Corridoni, J.M., Del Bimbo, A., Pala, P.: Image retrieval by color semantics. *Multimed. Syst.* **7**(3), 175–183 (1999)
10. Daniel, C.O., Olga, S., Tommer, L.: Colour harmonization. In: *ACM SIGGRAPH* pp. 624–630 (2006)
11. Da Vinci, L.: *The art of painting*. Trans. Richter, J. P. Dover Publications (1957)
12. Feller, R.L. (ed.): *Artists Pigments. A Handbook of Their History and Characteristics* Washington, D.C.: National Gallery of Art (1986)
13. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 3rd edn. Pearson Education (2007)
14. Helmholtz, H.V.: LXXXI. On the theory of Compound Colours *Philosophical Magazine Series* **4**, 4, Iss. 28 (1852)
15. Hermeren, G.: *Aesthetic Value*. Boulder, Colo.: Westview Press. 1995
16. Itten, J.: *The Art of Colour: The Subjective Experience and Objective Rationale of Colour*. Reinhold, New York (1961)
17. Kandinsky, W.: *Concerning the Spiritual in Art*. 1911
18. Kuehni, R.: *Color Space and its Divisions: Color Order from Antiquity to the Present* Wiley-Interscience, 2003
19. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. *IEEE J. Sel. Top. Signal Process.* (2009)
20. Maxwell, J.C.: *Experiments on Colour as Perceived by the Eye, with Remarks on Colour-Blindness* (1855)
21. Munsell, A.H.: *A Color Notation*. G. H. Ellis Co., Boston (1905)
22. Nault, W.H.: *Children's Map Reading Abilities*. Geographic Society of Chicago, Newsletter, III (1967)
23. Newton, I.: *The Principia: Mathematical Principles of Natural Philosophy*. University of California Press (1999)

24. Nishiyama, M., Okabe, T., Sato, I., Sato, Y.: Aesthetic quality classification of photographs based on color harmony. In: Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition. (2011)
25. Runge, P.O.: Die Farben-Kugel oder Construction des Verhlnisses aller Mischungen der Farben zueinander, und ihrer vollstndigen Affinit, mit angehgtem Versuch einer Ableitung der Harmonie in den Zusammenstellungen der Farben (Eng. Color sphere or construction of the relationship of all mixtures of colours, and their complete affinity, with an attached essay on the derivation of harmony in color compositions). Hamburg: Perthes (1810)
26. Thurstone, L.L.: A Law of Comparative Judgment *Psychological Review* **34**, 273–286 (1927)
27. von Goethe, J.W., Eastlake, C L. Goethe's Theory of Colours: Translated From the German (1840)
28. Westland, S., Laycock, K., Cheung, V., Henry, P., Mahyar, F: *Colour Harmony Colour: Design & Creativity* (2007)

Chapter 3

Computational Models of Colour Contrast

Abstract In the previous chapter, colour contrast was examined as phenomena within the art domain. In this chapter, we review computational models that might be employed to define these phenomena. In the engineering domain, there is not a standard measurement model of lightness, saturation or hue contrast. All existing models would probably be better regarded as application tools rather than as descriptions of basic visual function. Colour contrast is broadly understood as being one of two things: a relationship between two things (i.e. simple contrast) or the variance of a property within a single thing (i.e. complex contrast).

Keywords Colour contrast · Contrast models · Hue contrast

3.1 Contrast Models for Simple Images

The following contrast models are suitable for simple images or for contrast between regions.

In physics, the difference between light and dark is called the luminance contrast. Since no contrast model is available to measure the perceptual difference of lightness, we therefore assume that lightness contrast can be measured in terms of the luminance contrast. According to Weber–Fechner law [2], the relationship between the physical magnitude of luminance and the perceived intensity is

$$dp = k \frac{dS}{S} \tag{3.1}$$

where S is the physical stimulus, dp is the differential change in perception, and dS is the differential change in the stimulus. k is a constant factor. This relationship tells us that a small difference is negligible if the average luminance is high, while the same small difference can give us clear contrast if the average luminance is low.

Based on the relationship in Eq. (3.1), many definitions of contrast can be seen to follow the general relationship below:

$$C = \frac{L_{\Delta}}{\bar{L}} \quad (3.2)$$

where L_{Δ} is the luminance difference and \bar{L} is the average luminance. Some contrast models defined according to this relationship are summarised as follows.

Weber contrast. This is commonly used in cases where a small object is present on a large uniform background [2]. It is defined as

$$C_W = \frac{I - I_b}{I_b} \quad (3.3)$$

where I is the luminance of the object and I_b is the background luminance.

Luminance ratio. This is frequently applied to the same stimulus as Weber contrast. It is defined as

$$C_{LR} = \frac{I_{\max}}{I_{\min}} \text{ or } \lg C_{LR} = \lg \left(\frac{I_{\max}}{I_{\min}} \right) \quad (3.4)$$

Michelson contrast. This is commonly used for cases where both bright and dark features are equivalent, each being of similar size to the other. It is defined as

$$C_M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (3.5)$$

where I_{\max} and I_{\min} are the highest and lowest luminance, respectively.

3.2 Contrast Models for Complex Images

In complex images, the perceived contrast varies greatly across the image plane. The following models propose to measure this contrast.

Root mean square. Root-mean-square (RMS) contrast [8] is the standard deviation of intensity values. It is defined as

$$C_{RMS} = \sqrt{\frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_{ij} - \bar{I})^2} \quad (3.6)$$

where I_{ij} is the i -th, j -th element of the image of size $M \times N$ and \bar{I} is the average intensity of the image. RMS contrast does not capture the spatial information of the image.

Peli [8] proposed a local band-limited contrast definition for contrast measurement in complex images. This contrast definition considers the spatial frequency content of the image. The contrast at each band of spatial frequencies is

$$c_i(x, y) = \frac{a_i(x, y)}{l_i(x, y)} \quad (3.7)$$

where $a_i(x, y)$ is the i -th band-pass filtered image and $l_i(x, y)$ is the low-pass filtered image containing all energy below the band i .

Tadmor and Tolhurst [11] proposed to calculate the local contrast by simulating the visual response to stimulus contrast in natural scenes. Based on the Difference of Gaussian (DOG) receptive-field model, three contrast models are proposed as follows.

$$C_1(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_c(x, y)} \quad (3.8)$$

$$C_2(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_s(x, y)} \quad (3.9)$$

$$C_3(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_c(x, y) + R_s(x, y)} \quad (3.10)$$

where $R_c(x, y) = \sum_i \sum_j G_C(i - x, j - y)I(i, j)$, $R_s(x, y) = \sum_i \sum_j G_S(i - x, j - y)I(i, j)$, $G_C(x, y)$, $G_S(x, y)$ are 2D Gaussian models to express the receptive-field centre and surround components, respectively. I is intensity.

Rizzi et al. [9] proposed a local contrast measurement of digital images. It subsamples the image to various levels. At each level, the local contrast is calculated. The definition of the local contrast is

$$C_R = \frac{1}{N} \sum_l \left[\frac{1}{m^{(l)}n^{(l)}} \sum_i \left(\sum_{j \in N8(i)} \frac{L_i^{(l)} - L_j^{(l)}}{8} \right) \right] \quad (3.11)$$

where $m^{(l)}$ and $n^{(l)}$ are the numbers of rows and columns on level l , N is the number of levels, $N8(i)$ is the eight neighbours of pixel i and L is L channel in CIELAB space. This measure is simple and has a lower computational complexity than DOG-based methods. In 2008 [10], the authors improved the contrast measure by combing the multilevel approach with Tadmor and Tolhurst's evaluation of a colour stimulus [11]. The local contrast at each level is defined as

$$C_{RSC}^c = \frac{1}{N} \sum_l \left[\frac{1}{m^{(l)}n^{(l)}} \sum_i DOG^{(l)}(i) \right] \quad (3.12)$$

DOG is the C_3 function in the Tadmor and Tolhurst method [11], $m^{(l)}$ and $n^{(l)}$ are the numbers of rows and columns on level l , c is the channel which can be L, a or b

channel in CIELAB space. Isoluminant colour contrast is linearly combined with the luminance contrast in CIELAB space for colour images.

A global contrast factor was used by Matkovic et al. [6] to measure richness of detail as perceived by a human. Global contrast factor is defined as

$$C_{GCF} = \sum_{i=1}^N w_i * C_i \quad (3.13)$$

where C_i is the average local contrast for level i , local contrast is the average difference of one pixel with its 4-neighbour pixels, N is the number of levels and w_i is the weight for level i .

Mantiuk et al. [5] used a low-pass contrast measurement to avoid halo artefacts at sharp edges. The low-pass contrast is defined as $G_{ij} = \log_{10}(L_i/L_j)$, where L_i and L_j are luminance values of neighbouring pixels i and j .

Compared to luminance contrast, contrast models for saturation are hard to find. Generally, the saturation contrast is defined as the saturation difference of two regions or two neighbouring pixels [3, 12], which is expressed as $C_S = S_i - S_j$, where S_i and S_j are the saturation values of pixels i and j or the average saturation of regions i and j . Some models are designed to measure the colour contrast considering the chromatic and achromatic aspects. Rizzi et al. [10] calculated the contrast for L, a, b channels in CIELAB space separately. Then, the colour contrast was measured as a linear combination of the contrast values of L, a and b. Marius Pedersen used Lab variance as a colour contrast measure, which is calculated as the geometrical mean of the variance in each channel in the CIELAB colour space [7].

3.3 Contrast Models for Hue Contrast

There is little discussion or agreement as to what constitutes hue contrast. When formulating his colour space, Munsell painstakingly mapped hue as degrees of 'barely perceptual difference' using human perception as the point of reference (see Sect. 2.2 on p. 10).

Machajdik [4] defined the range of hues within a region or an image as the degrees of spread hue of its hue histogram, as mapped within an RYB hue wheel. A large hue spread angle indicates high hue contrast and a small one indicates low. The contrast of hue (hue difference) between two regions is defined as the $d = \min(|h_i - h_j|, 360 - |h_i - h_j|)$ where h_i is the mean hue of the region i [4].

The problem with defining hue contrast is one of dimensionality. The lightness and saturation global contrast of an image generally rely upon its average, maximum and minimum values being knowable. Conversely, in defining the global hue of an image, there is no such thing as an average hue, nor a maximum or minimum one. Hence, other means are required for the task. In our work, the hue contrast is defined with reference to the values hue variety and hue antagonism [1]. This is described in Chap. 10 on p. 179.

References

1. Constable, M., Wang, J., Chan, K.L., Zhang, X.: A definition of global hue contrast in artworks. *J. Int. Colour Assoc.* **14**, 28–36 (2015)
2. Hecht, S.: The visual discrimination of intensity and the Weber–Fechner law. *J. Gen. Physiol.* **7**, 235–267 (1924)
3. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. *IEEE J. Sel. Top. Signal Process.* **3**(2), 236–252 (2009)
4. Machajdik, J., Hanbury, A.: Affective image classification using features inspired by psychology and art theory. In: *Proceedings of MM* (2010)
5. Mantiuk, R., Myszkowski, K., Seidel, H.P.: A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.* **3**, 286–308 (2006)
6. Matkovic, K., Neumann, L., Neumann, A., Psik, T., Purgathofer, W.: Global contrast factor - a new approach to image contrast. In: *Proceedings of Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging*, pp. 159–168. (2005)
7. Pedersen, M., Rizzi, A., Hardeberg, J.Y., Simone, G.: Evaluation of contrast measures in relation to observers perceived contrast. In: *Proceedings of Fourth European Conference on colour in Graphics, Imaging and Vision*, pp. 253–258. (2008)
8. Peli, E.: Contrast in complex images. *J. Opt. Soc. Am.* **7**(10), 2032–2040 (1990)
9. Rizzi, A., Algeri, T., Medeghini, G., Marini, D.: A proposal for contrast measure in digital images. In: *Proceedings of Second European Conference on colour in Graphics, Imaging and Vision*, pp. 187–192. (2004)
10. Rizzi, A., Simone, G., Cordone, R.: A modified algorithm for perceived contrast measure in digital images. In: *Proceedings of Fourth European Conference on Colour in Graphics, Imaging and Vision*, pp. 249–252. (2008)
11. Tadmor, Y., Tolhurst, D.: Calculating the contrasts that retinal ganglion cells and LGN neurons encounter in natural scenes. *Vis. Res.* **40**, 3145–3157 (2000)
12. Yelizaveta, M., Chua, T., Irina, A.: Analysis and retrieval of paintings using artistic colour concepts. In: *Proceedings of IEEE International Conference Multimedia and Expo*, pp. 1246–1249 (2005)

Chapter 4

The Geometric Attributes of Paintings

Abstract Within a painting, two forms of geometry can be identified: 2D and 3D. 2D geometry is defined by such things as the corner, the centre, the edge and the middle. 3D geometry is defined as simplified planes, such as foreground, middle ground and background and is particularly important in enhancing the depth of a landscape painting. These geometries function as structural scaffolding around which the painter builds their contrasts. The composition of a painting is covered by its 2D geometry. We examine the rule of thirds which, through an examination of its history, is shown to be more complex and interesting than its current form. We also look at the placement of the region of interest and of the horizon and find them to be governed by simple principles. A photographer's vignette is simple: a darkening of the edges and corners, and perhaps also a lightening of the centre. In contrast, a painter's vignette is complex and involves interplay between the 2D and 3D geometries of the painting.

Keywords Depth planes · The vignette · Composition · Rule of thirds · Horizon · Region of interest

We can observe that a painting is composed of both 2D and 3D geometry. The 2D geometry occurs on the flat plane of the picture and consists of such things as centre, corner and edge values [3]. In the art domain, this is referred to as the picture plane (see Fig. 4.1) and describes an imaginary plane between the artist and the scene, the limits of which, define the limits of the painted scene. The 3D geometry of a painting is an illusional attribute, crafted by the skill of the painter. It is defined by the distance of the viewer from the objects in the scene.

Though depth can be expressed in infinite degrees of distance, artists commonly express it as a limited number of flat planes, arranged along the z-axis. These are termed depth planes and might include the FG, MG, BG and sky (see Fig. 4.2).

Defining the limits and location of the depth planes is not always straightforward. The boundary of these planes is mostly defined by the limits of an object or group of objects (e.g. a hill or a group of trees). However, some regions will likely blend into each other, with their bounds becoming consequently indistinct. Additionally, sometimes a clearly defined object, such as a river running into the distance, will span more than one depth plane. All this can make the exact limits, and even the number

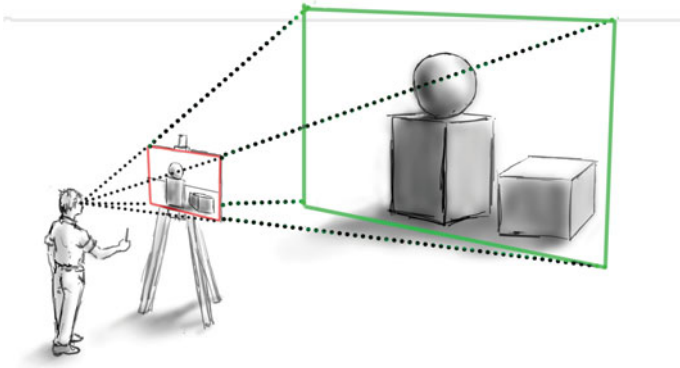


Fig. 4.1 The picture plane (in red) which, when projected, describes the limits of the scene as perceived by the artist (in green) (color figure online)

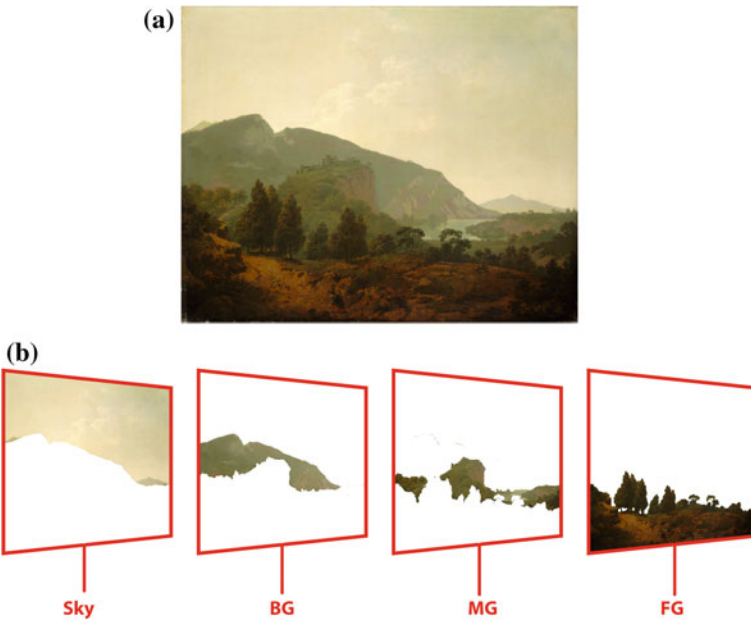


Fig. 4.2 a 'Italian Landscape', Joseph Wright (1734–1797), 1790. b The outlined depth planes of the same painting. Courtesy National Gallery of Art, Washington

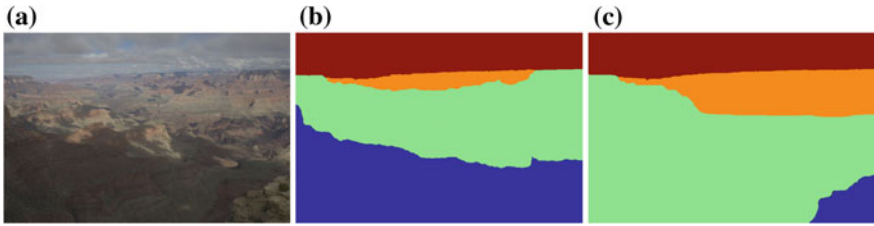


Fig. 4.3 A landscape **a**, the depth planes of which have been partitioned in different ways by different people **(b)** and **(c)**

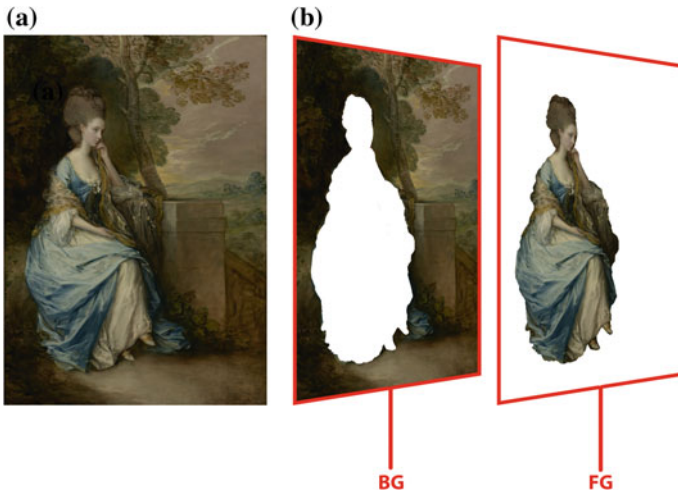


Fig. 4.4 **a** ‘Portrait of Anne, Countess of Chesterfield’, Thomas Gainsborough, c1778. Courtesy the Getty Research Institute. **b** The outlined depth planes of the same painting

of these planes, difficult to define. This is shown in Fig. 4.3, where a photograph (a) has been partitioned differently by two different people (b) and (c).

The 3D geometry plays a more significant role in landscape painting than portrait painting, with the average landscape employing more depth planes than the average portrait (see Fig. 4.4). This is because depth is the implicit subject of most landscape paintings.

The 2D geometry of a painting is defined by its centre, corners, edges and suchlike. It plays an important role in portrait painting, in which the figure invariably lays close to, or upon, the middle vertical. In a portrait painting, depth contrast plays a secondary role to local and regional contrast, both of which are primarily 2D contrasts.

4.1 Composition in Paintings

In art, composition is the configuration of the 2D geometry of a painting in order to serve an aesthetic purpose. In portrait paintings, composition serves to highlight the focus of attention, which is the face. In landscape paintings, it serves to enhance the painting's apparent depth.

4.1.1 *The Rule of Thirds*

The so-called rule of thirds has a grand reputation as defining compositional balance in a painting or a photograph. The function of this section is to draw attention to the history of this rule, and to a degree debunk it. In common understanding, it refers to the need to locate important objects on one of the four cross points of the horizontal and vertical one-third divisions. However, if the history of this rule is examined, something more complex and interesting comes to light.

Painters generally avoid symmetry for the flattening effect that it has upon the form of the painting. In his book 'Annotations on The art of painting of Du Fresnoy' [4], the British artist Sir Joshua Reynolds (1723–1792) said:

Permit not two conspicuous lights to shine with rival radiance in the same design. The same right judgment which proscribes two equal lights, forbids any two objects to be introduced of equal magnitude or force, so as to appear to be competitors for the attention of the spectator. This is common; but I do not think it quite so common, to extend the rule so far as it ought to be extended: even in colours, whether of the warm or cold kind, there should be one of each which should be apparently principal and predominate over the rest. It must be observed, even in drapery, that two folds of the same drapery be not of equal magnitude. (p 96)

Clearly he is expressing the opinion that symmetry is a 'bad thing'. Though he presents this as a rule, it should be clear that Reynolds was not prescribing a way of doing things, rather stating a 'case use' which should be generally avoided. In other words, it was a rule of avoidance, not prescription. It was also extremely general: referring to colour, form, light etc. In 1797, and in response to Reynold's observation, the English artist John Thomas Smith (1766–1833) restated the same principle [13]. It was Smith who also gave it the name by which it is now known.

Analogous to this 'Rule of thirds', (if I may be allowed so to call it) I have presumed to think that, in connecting or in breaking the various lines of a picture, it would likewise be a good rule to do it, in general, by a similar scheme of proportion; for example, in a design of landscape, to determine the sky at about two-thirds; or else at about one-third, so that the material objects might occupy the other two... I have found the ratio of about two thirds to one third, or of one to two, a much better and more harmonising proportion, than the precise formal half, the too-far-extending fifths and, in short, than any other proportion whatever. (p. 16)

In contrast to Reynolds, Smith offers a more prescriptive approach. Nonetheless, he is at pains to stress its approximate nature ('*about* one third'). Since then, the rule of thirds has (devolved?) into something far more specific, referring only to the placement of the subject within a scene, and has also become entirely prescriptive.

Fig. 4.5 The golden ratio visually expressed

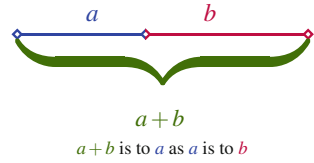
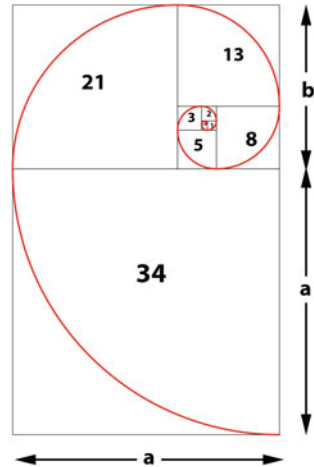


Fig. 4.6 A rectangle constructed from the golden ratio, inside of which is a recursive spiral



We would argue that the rigidly applied 1:2 ratio that the contemporary rule of thirds prescribes is as questionable as a 1:1 ratio, being effectively another form of the same symmetry that Reynolds found so objectionable.

Research by Amirshahi et al. supports this observation. They found the rule of thirds to be not be as important in the evaluation of the visual quality of artworks as commonly assumed [1].

What Reynolds and Smith were probably responding to was the influence upon artists of the golden ratio. This is a geometrically derived proportion first codified by the Ancient Greeks which may be easily confused with a thirds division by the casual observer. It refers to the division of a line into two parts: a & b , where the ratio of $(a + b)$ to a is equal to the ratio a to b (or $r = \frac{a+b}{a} = \frac{a}{b}$). Visually this can be expressed as shown in Fig. 4.5.

From this ratio may be derived 2D and 3D geometric forms that are assumed to be aesthetically agreeable. A key property of this value is that it is recursive, i.e. it will never return any of the symmetries that Reynolds found so aesthetically troublesome. Shown in Fig. 4.6 is a rectangle derived from the golden ratio. If a square is constructed inside the rectangle, the remaining region would be of the same proportion as the original rectangle. Performed recursively, each square can be then bisected, with the result forming an infinitely recursive spiral (shown in red).

Some artists knowingly employ this so-called golden geometry. The impressionist painter Georges Pierre Seurat (1859–1891) ‘attacked every canvas by the golden

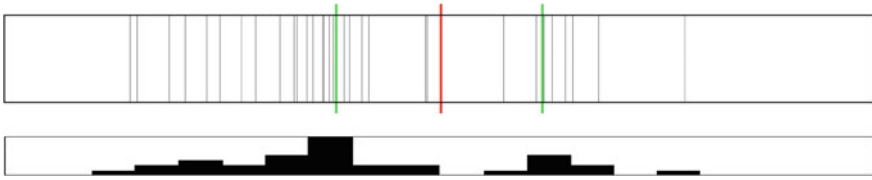


Fig. 4.7 The result of asking 31 art students and art teachers to place a mark dividing a 100 cm length of dowelling. The *red line* marks the location of the middle, and the two *green lines* mark the location of the *left* and *right* golden ratio (color figure online)

section’ [6], and Nicolas Poussin (1594–1665) was famous for employing golden geometry very prescriptively, as did Piero della Francesca (unknown–1492) [10].

Not all artists knowingly employ the golden geometry, but the following experiment conducted in an art school shows its influence, whether conscious or otherwise. Thirty-one art students and art teachers were each given a length of dowelling 100 cm in length (dowelling is thin round-section timber). They were asked to each place a mark on the timber, thus dividing it into two sections. This mark was to be at a point they found ‘balanced but not symmetrical’. This distinction is a significant one and corresponds to the principle of informal (asymmetrical) and formal (symmetrical) balance [8]. As Reynolds and others have noted, painters strongly disfavour symmetry. The orientation of the stick was noted (i.e. whether they marked to the left or the right of the centre). Figure 4.7 shows the all the responses superimposed upon each other. It is clear that there is a significant clustering towards the left half of the stick. Less obvious, but still apparent, is a clustering towards the left golden ratio divide.

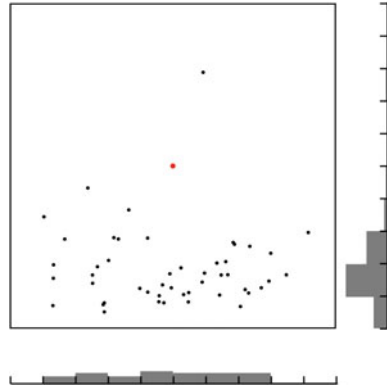
4.1.2 *Region of Interest and the Horizon*

As a general rule, humans perceive what they are looking at as either a scene or an object [9]. A portrait may be regarded as a depiction of the human object, while a landscape is clearly a depiction of a scene. Humans perceive a scene from the outside-in: beginning with its global generalities, and ending with its local specifics. These are in the form of an object or cluster of objects and are usually of strong narrative importance, such as a human figure in a bare landscape or a ship in a seascape. This is known as the ‘region of interest’ (ROI), sometimes also referred to as its ‘focal point’, ‘centre of attention’ or ‘point of interest’. This ROI is visually emphasised by its various contrast properties.

The need for a landscape painting to feature a ROI is so strong that paintings without such foci were almost unheard of until the 1820s. It was Jean-Baptiste-Camille Corot (1796–1875) and later the Impressionists who primatised the landscape over its narrative content, and in the process did away with the ROI.

Figure 4.8 shows the location of the ROI of 48 landscape paintings. There is a clear tendency for it to be located in the bottom third of the painting (46:2) and a

Fig. 4.8 The ROI of 48 landscape paintings. The *red dot* marks the location of the centre of the painting (color figure online)



slight tendency for it to be located to the left of the painting (25:23). However, the only clear rule being observed is that it is not located at the centre of the painting.

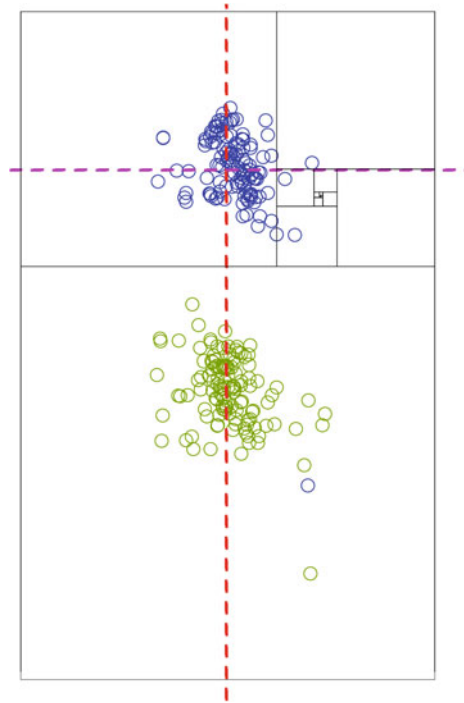
The horizon line of a painting serves to define the viewer’s implicit point of view. A high horizon line suggests that the viewer is looking down from a height, while a low horizon line implies that they are looking up. Shown in Fig. 4.9 are the horizon lines of 53 landscape paintings, sourced from the Hudson River school of landscape painters. The data is more coherent than that of the ROI, with significant clustering evident around the middle (in red), and the lower golden ratio division (in green). However, again the clearest rule being observed is one of avoidance, with very few horizons being located in the upper half of the painting.

In portrait paintings, the middle vertical is heavily employed, with key points of the figure being located on or near this line. Shown in Fig. 4.10 are the positions of the centre of the face (in blue) and the centre of the body (in green) of 120 full-length portrait paintings. These were sourced from such pre-modern painters as Édouard Manet (1832–1888), Jean-Auguste-Dominique Ingres (1780–1867), Thomas Gainsborough (1727–1788) and Francisco Goya (1746–1828). It can be seen that the

Fig. 4.9 The horizon line of 53 landscape paintings. Note the clustering of horizons below the middle horizontal (in red). The *green lines* represent the *upper* and *lower* golden ratio division (color figure online)



Fig. 4.10 The placement of the head (*blue*) and the centre of body (*green*) in 120 full-length portrait paintings. Note the clustering of both sets around the middle vertical, and the centering of the head cluster at the first twist of the golden ratio spiral (color figure online)



centre of the body and face both cluster around the centre vertical, shown as a dotted red line. Overlaid are nested rectangles derived from the golden ratio, from which it can be seen that the face cluster is located at the first twist of the spiral (shown as a dotted purple line). There is a slight right-hand bias of the head cluster. This is addressed in Sect. 4.2 and shown in Fig. 4.14.

In summary, the ROI of a landscape is defined according to how effectively it avoids the centre, while that of a portrait is defined by how well it observes the middle vertical and key points along that vertical. The horizon is defined by how well it avoids the top half of the painting, and to a lesser degree its middle horizontal.

4.2 Vignetting

This section employs figures and data from Sect. 9.3 on p. 157.

Vignetting is a visual effect that manifests as a clarity in or around the centre of an image, and a fading-off towards its edges. It can be the result of many phenomena [7, 11]. Exclusively a property of digital cameras, ‘pixel vignetting’ is caused by the angle-dependence of digital sensors. ‘Mechanical vignetting’ is caused by physical obstructions in front of the lens, such as a poorly placed lens hood or a slow moving

Fig. 4.11 An example of a vignette. ‘Fanciful Landscape’, Thomas Doughty (1793–1856), 1834. Courtesy National Gallery of Art, Washington



shutter. ‘Optical vignetting’ is found in compound lenses and is caused by lens elements being occluded by other lens elements. All these forms of vignetting are usually regarded as undesired effects caused by poor camera operation.

A vignette is also a natural outcome of the way humans perceive the world: instinctively turning their heads to place objects of importance within the approximate centre of their field of vision. Additionally, the periphery of a human’s field of vision is indistinct and ill-perceived (so-called peripheral vision). Certainly, it is not bound by the clear rectangle, as in a painting or a photograph. The result is a scene which is clear towards the centre and indistinct towards the periphery: effectively a vignette.

Emulating this natural property, photographers employ the vignette as an aesthetic effect. The aesthetic photographic vignette effect can be created by adjusting camera settings or by post-processing. The ‘Lens Correction’ filter in Photoshop can also achieve the effect. The post-processed and camera-based vignette are both simple in form, being effectively dark semi-transparent ovoids placed *on top of the image*.

Painters also enhance the visual appeal of their paintings by the use of a vignette (see example Fig. 4.11). Though the exact centre of a painting is avoided by the artist, it can be observed that elements of high pictorial value are generally placed somewhere near the centre. This is the so-called region of interest (see Sect. 4.1.2 on page 44). Landscape painters are very good at selecting viewpoints that favour this centre-weighted arrangement [6], which the vignette serves to accentuate.

Though in principle the 2D picture plane and the 3D depth planes of a painting spatially conflict with each other, in practice there is frequently co-dependence between them [2]. Hence, the centre-corner contrast of a painted vignette is usually supported by depth-aware contrast (see Sect. 2.3.2 on p. 18), with select regions of the depth planes being manipulated to enhance the centre-corner effect. Working together, centre-corner and depth-aware contrast combine to produce the impression of a spiral that runs down the Z axis of the painting. Projected from the front, this manifests as a vignette with one open top corner.

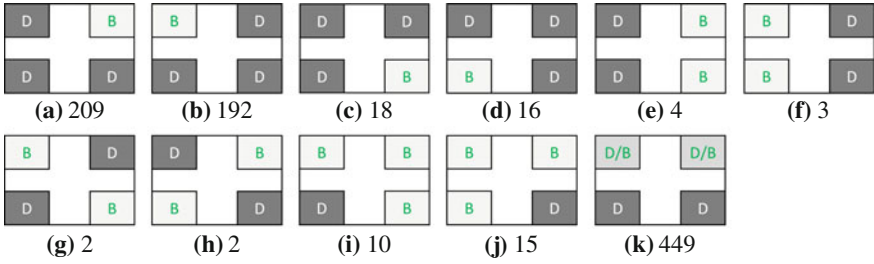


Fig. 4.12 Pictograms of the landscape painting corner lightness weighting patterns. In each of the 11 patterns, *D* and *B* denote *darker* and *brighter* respectively. The number under each pattern denotes the number of paintings the corner lightness weighting patterns of which correspond to that pattern. The dataset consisted of 920 paintings. In pattern **k**, there is no clear bias on the *top* or *bottom* two corners and the lightness in the *top corners* is of similar brightness or darkness

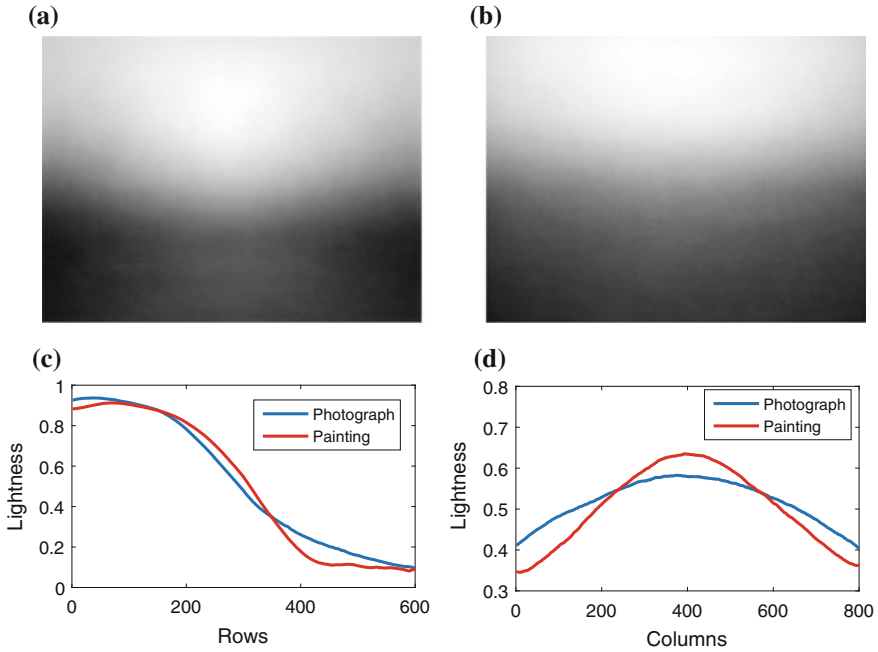


Fig. 4.13 **a** The averaged and normalized lightness of 920 landscape paintings. **b** The averaged and normalized lightness of 920 photographs. These are calculated by resizing all the paintings or photographs to the same size (600×800) and then calculating the pixel-wise average of the lightness image. Plots **c** and **d** present the average values, respectively, of 800 columns and 600 rows for the two average lightness images in **(a)** and **(b)**. The paintings exhibit a slight but noticeable bias of lightness on the *right*, while the photographs show no such bias

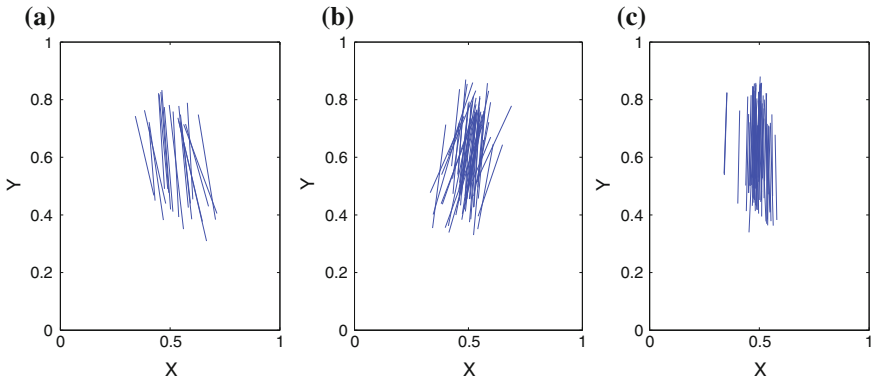


Fig. 4.14 The orientation of 120 portrait paintings. **a** 20 were oriented towards the *left*. **b** 51 were oriented towards the *right*. **c** 49 were oriented *vertically*

Interestingly, it is more common for a vignette to feature an open top right corner than top left. This open cornering is detailed in Fig. 4.12. Further supporting this right-weighted bias, Fig. 4.13 shows the average lightness of 920 landscape paintings. Slight but clear bias is apparent. For comparison, the same number of photographs of similar content has been examined, where no such weighting can be seen. Generally, it can be observed from these figures, that paintings follow the same lightness distribution of photographs, but their light regions are lighter and their dark regions are darker.

This lateral bias is also evident in portraits, but takes a different form. In Fig. 4.14, 120 full-figure portraits were examined for the orientation of the figure. This was understood as the axis that joined the centre of the body to the centre of the head. These centres are shown superimposed on each other in Fig. 4.10. Rightwards and vertical alignment occur as frequently as each other, yet leftwards alignment is generally avoided.

Arnheim [3] and others [5, 12] have remarked on this lateral bias, which appears to be related to our reading and writing direction [3].

In Fig. 4.15b, c is shown a Photoshop-style vignette which features an even darkening towards the corners. Figure 4.15d, e shows a vignette made by a trained digital artist that emulates that of a traditional painting. The darkening towards the corners lays across the 2D plane, but is also derived in part by 3D geometry within the scene (e.g. the trees on the left hand side). Additionally, the centre has been lightened, the contrast decreases towards the far depth planes and increases towards the foreground.

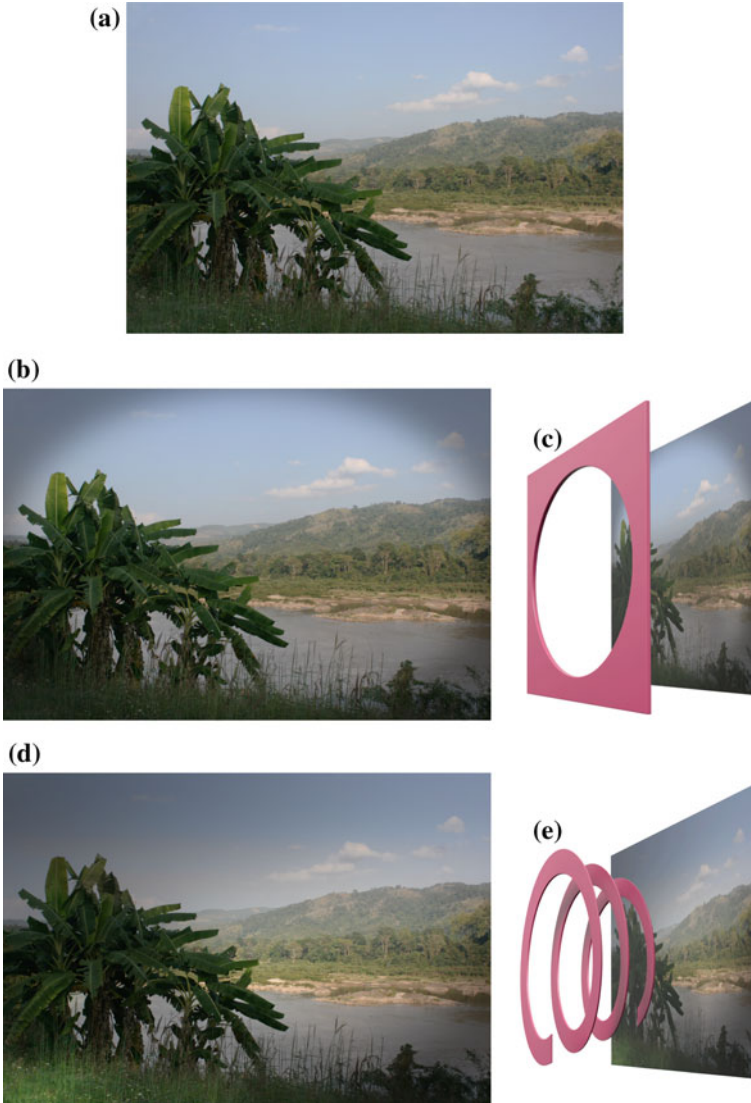


Fig. 4.15 The difference between a standard vignette effect and a painter's vignette. **a** Original photograph. **b** A standard vignette effect in Photoshop®. **c** The standard vignette effect shown as a *flat circle* in front of the image. **d** An emulation of a painter's vignette made in Photoshop®. **e** The painter's vignette shown as a *spiral* running down the centre of the image

4.3 Concluding Remarks

This chapter examined the 2D and 3D geometry of paintings. Composition is the means by which this geometry is organised in order to enhance the aesthetic value of the painting. We examine the history and practice of the rule of thirds and conclude that though as a steadfast rule it may be of limited truth, it is likely a result of the influence of such recursive geometries as the golden ratio. Key aspects of composition, the ROI and the horizon, we observe as being governed by opposing principles: avoidance and observance. Vignetting we show is being derived from the perceptual biasing of the centre of human vision, together with a fading-out of perception towards the periphery. We show that the painter's vignette, compared to that of a photographer, is superior in its complexity and function.

References

1. Amirshahi, S.A., Hayn-Leichsenring, G.U., Denzler, J., Redies, C.: Evaluating the rule of thirds in photographs and paintings. *Art Percept.* **2**, 163–182 (2014)
2. Arnheim, R.: *Art and Visual Perception* Univ. of California Pr. (2004)
3. Chahboun, S., Flumini, A., González, C. P., McManus, I. C., & Santiago, J.: *Reading and Writing Direction Effects on the Aesthetic Appreciation of Photographs* (2016)
4. Dufresnoy, C. A., Reynolds, J.: *The Art of Painting of Charles Alphonse Du Fresnoy (1783)*
5. Gaffron, M.: Right and left in pictures. *Art Q.* **13**, 312–331 (1950)
6. Ghyka, M.C.: *The geometry of art and life.* Courier Corporation (1946)
7. Goldman, D.B.: Vignette and exposure calibration and compensation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(12), 2276–2288 (2010)
8. Landa, R.: *Graphic Design Solutions* (2010)
9. Oliva, A.: Gist of the scene. *Neurobiol. Atten.* **696**(64), 251–258 (2005)
10. Pedoe, D.: *Geometry and the Liberal Arts.* St. Martin's Press, New York (1976)
11. Ray, S.: *Applied Photographic Optics.* Focal Press, Massachusetts (2002)
12. Shimamura, A.: *Experiencing Art: In the Brain of the Beholder* (2013)
13. Smith. J. T.: *Remarks on Rural Scenery; with Twenty Etchings of Cottages, From Nature; and Some Observations and Precepts Relative to the Pictorial* (1797)

Part II
Transfer of Aesthetic Values

Chapter 5

Global Colour Style Transfer

Abstract In this chapter, we examine the global chromatic and achromatic aspects of a painting. Their organisation is particular to each artist, with work by that artist showing consistency in its contrast organisation. This organisation forms a significant part of their ‘style’. Contrast can be broken down into its chromatic aspects (hue and saturation) and its achromatic aspect (lightness). We propose a method by which this style may be transferred from a painting to a photograph. The principle behind our method is to transfer the inherent structures of the contrast properties. This differentiates our method from a simple colour transfer, using which the hues of an image may be shifted, and in doing so damages the semantic properties of the photograph.

Keywords Global colour contrast · RYB colour space · Dominant hue

It is clear that an artist’s style is consistent through most of their career. This style, influenced by the habits and customs of their time, is also unique to that artist. Much of that style may be defined as the unique structure of the contrast properties which they employ. We seek to transfer the important aspects of that structure onto a photograph.

Our operation on colour is different from colour transfer [1, 5, 7, 8] which serves to shift the absolute colour values of a target image to match those of a reference image. The principle of style transfer is to transfer the contrast *structure* whilst retaining the dominant colour of the original photograph. The motivation of our work is similar to that which seeks to impose colour harmonisation on an image [2], but the harmony is not based on prescribed rules, rather the colour scheme of a reference painting.

First, we analyse the global hue, saturation and lightness contrast organisations of paintings from two well-known yet very different artists: Vincent Van Gogh (1853–1890) and Sir Lawrence Alma-Tadema (1836–1912). The popular Dutch post-impressionist, Van Gogh, is well known for the vigour and clarity of his forms and his expressive use of colour. In contrast, the work of the Dutch/British classicist Alma-Tadema depicts Roman classical life with consummate technical skill and employs a colour pallet that is colourful without being garish.

The saturation and lightness characteristics are transferred to a photographic image by histogram matching. The hue characteristics are transferred via dominant

hue alignment and spread stretch along the (Itten) RYB colour wheel. In the hue mapping, we map the hue of a painting to the RYB colour wheel in order to analyse its colour style as understood in terms that artist's would be familiar with. The conversion from RYB \Rightarrow RGB is based on the work of Gosset and Chen [4] with our extension for RGB \Rightarrow RYB mapping. Dominant hues are then extracted in an RYB colour wheel. The saturation and lightness distributions are analysed in HSV colour space. See Sect. 2.2 on p. 10 for details of these colour spaces.

5.1 Global Hue, Saturation and Lightness Analysis

The global colour organisation of an image describes the colour distribution characteristic of the entire image. Figure 5.1 shows the lightness, saturation and hue values of two paintings: one by each of our subject artists. In the Alma-Tadema, the figures are of high lightness contrast, set against a predominantly light background. The saturation is similar in character, being high saturation contrast against low. Differently, the Van Gogh features quite small saturation and lightness contrasts.

We selected 15 paintings from Lawrence Alma-Tadema and 15 paintings from Vincent Van Gogh. The average lightness and saturation distributions of the 15 paintings were calculated separately as shown in Fig. 5.2. These averages support our observations.

Compared with the distribution of the lightness and saturation, hue is more complex and difficult to define as a set of general differences between artists. This issue is addressed in Sect. 2.3.3 on p. 25. The painting 'Ask Me No More' contains a wide spread of blue, and a small spread of red-purple and yellow-green. In contrast, the painting 'The Night Cafe' by Van Gogh has clear dominant colours: red, yellow, green and green-blue. The colour schemes, which orders the relationship between dominant colours, selected by the painter influence the so-called harmony of the painting. Colour schemes are one way to express hue contrast. A colour scheme is defined by its spread width and the relative position of its dominant hues. It is rare for an artist to employ a single colour scheme in their entire body of work. Therefore, in our painting style transfer, we use a single example painting from one artist to change the contrast of the photograph instead of using one coherent rule derived from an artist.

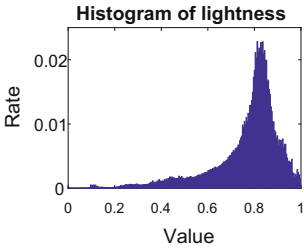
The proposed method for style transfer involved the extraction of the hue, saturation and lightness contrast characteristics of a painting and a photograph, followed by the mapping of these attributes to those from the painting. The hue was measured within the RYB colour wheel, and dominant hues were extracted to define its colour scheme. The global saturation and lightness contrast were reflected by the saturation and lightness histogram distributions, which were measured in HSV space.



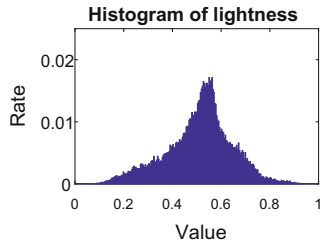
(a) Alma-Tadema's 'Ask Me No More' (1906)



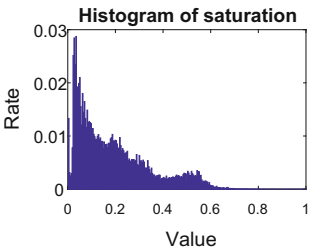
(b) Van Gogh's 'The Night Cafe' (1888)



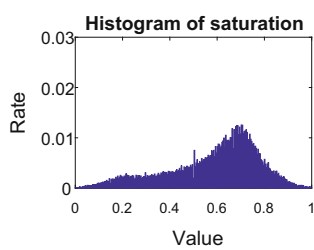
(c) Lightness histogram of 'Ask Me No More'



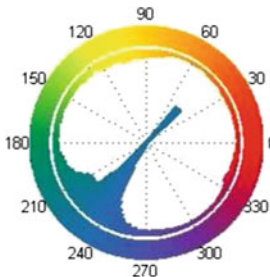
(d) Lightness histogram of 'The Night Cafe'



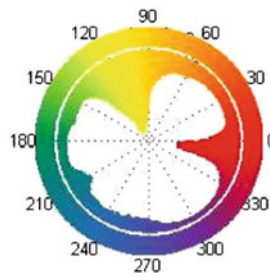
(e) Saturation histogram of 'Ask Me No More'



(f) Saturation histogram of 'The Night Cafe'



(g) Hue histogram of 'Ask Me No More'



(h) Hue histogram of 'The Night Cafe'

Fig. 5.1 Global colour organisation comparison of Alma-Tadema's 'Ask Me No More' (1906) and Van Gogh's 'The Night Cafe' (1888). Courtesy Wikimedia Commons

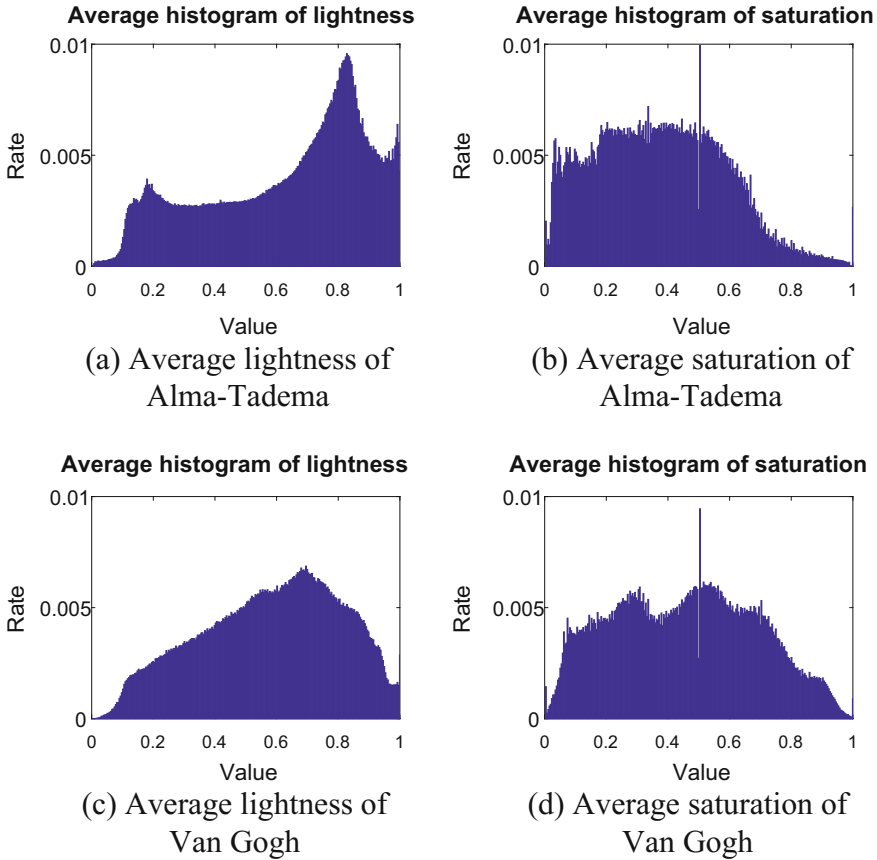


Fig. 5.2 The average lightness and saturation distributions of 15 paintings from Alma-Tadema and 15 paintings from Van Gogh

5.2 Hue Mapping

For our purposes, the hue of an image should be expressed in the RYB colour space (i.e. the so-called artists' colour space). This is described in Chap. 2 on p. 10. It should also express clearly dominant hues. In our hue mapping, these are then aligned from the source painting to the target photograph.

5.2.1 RGB to RYB Conversion

In order to be coherent with Itten's colour wheel, we corrected the values of the eight corner colours of the RYB colour cube (in Fig. 5.3a) proposed by Gosset and Chen [4]. Based on the conversion model using linear interpolation, we obtained the RYB colour wheel (see Fig. 5.3b, c).

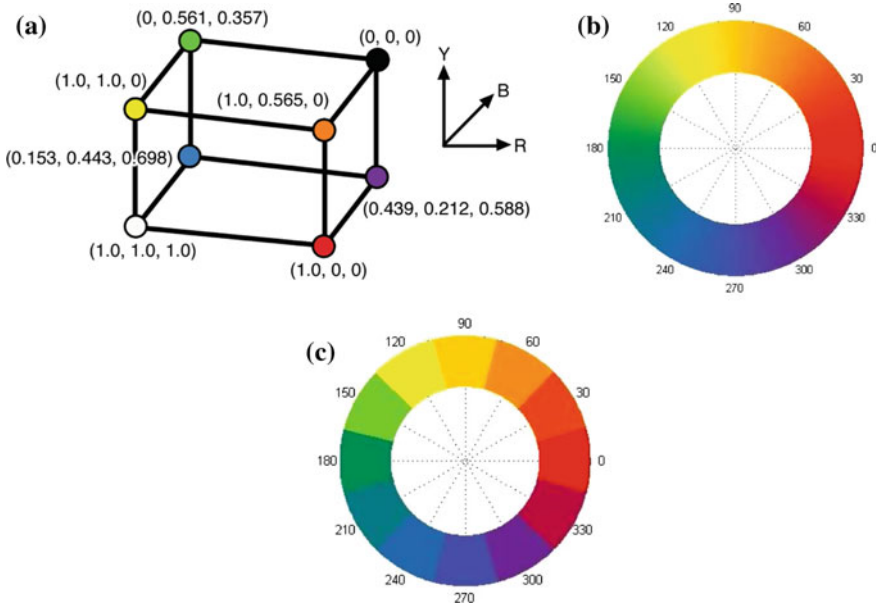


Fig. 5.3 a RYB interpolation cube. b The continuous RYB colour wheel. c The 12-hue colour wheel. The values are coincident with Itten’s 12 hue colour circle(color figure online)

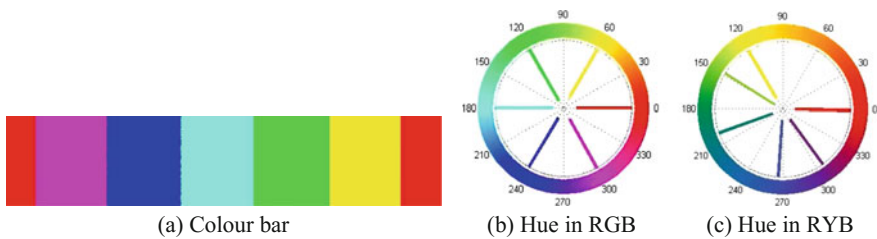


Fig. 5.4 Conversion of a colour bar in RGB to RYB space

To process an RGB digital image in RYB colour space, it needs first to be converted. The conversion from RGB to RYB was achieved with reference to Newton’s Method [6]. In the conversion process, some RGB colours were out of range. These were moved to the nearest colours in the RYB wheel. For example, when converted to RYB, pure RGB blue $(0, 0, 1.0)$ is $(-0.47, -1.70, 1.04)$. The normalisation process began by subtracting the smallest value, and hence we had $v_1 = (-0.47, -1.70, 1.04) - (-1.70, -1.70, -1.70) = (1.23, 0, 2.74)$. If the largest value of v_1 was bigger than 1.0, then $v_2 = v_1/\max(v_1) = (0.44, 0, 1.00)$. As the normalisation was linear, so the hue value remained unaffected. The RYB cube model is similar in structure to the RGB cube model, so we adopted the conversion algorithm from RGB to HSL to calculate the hue in RYB.

Figure 5.4 illustrates the conversion using one fully saturated colour bar in RGB space. The mapped hue shows that the relative position of colours in RYB is significantly different from that in RGB. Pure green in RGB is biased towards yellow in RYB, and pure blue in RGB is biased towards purple.

5.2.2 Dominant Hue Extraction

Itten describes three hue-specific contrasts: complementary contrast, warm and cool contrast and hue contrast (described in Chap. 2, on p. 17). These may all be expressed positionally within a circular hue histogram. Complementary contrast and warm and cool contrast are linear combinations of two-level factor means. For normal hue contrast, three dominant hues are sufficient to express most of the relationship in the colour wheel. So three dominant hues were extracted in the measure of global hue contrast in RYB colour wheel, and the three dominant hues constituted the colour scheme used in the painting.

In black and white areas there is evidently no colour, likewise in grey areas. Very low saturation values account for all these null colour exceptions. For this reason, pixels with saturation less than a threshold value (e.g. 0.1) were ignored in the dominant hue extraction.

There is a lot of variation of value within one area of a painting caused by such things as differences in canvas absorbance, canvas primer, brush technique and colour layering. Relative to a painting, a photograph presents as far more contiguous and uniform. In order to maintain consistence in dominant hue clustering, we first smoothed the painting using a mean-shift-based smoothing algorithm [3]. Then k-means clustering with $k = 3$ was used to cluster the pixel values from the smoothed image in RYB. The clustering texture of each pixel was defined as $(H, S = 1)$ in a polar coordinate space (H, S) . Finally, it was converted to the Cartesian coordinates (x, y) for clustering. The hue extraction result of one painting is shown in Fig. 5.5.

When an image contains only complementary contrast with two dominant colours, one of the three clusters partitioned by k-means may have extremely few pixels. In this case, this cluster was merged with the nearest cluster.

After k-means clustering, the size of a cluster, $s(i)$, $i = 1, 2, 3$, is given by the number of pixels in the cluster i . A weight, $w(i)$, is assigned to each cluster i ,

$$w(i) = \frac{s(i)}{\max(s)}, i = 1, 2, 3 \quad (5.1)$$

where $s = \{s(i)\}$.

To show the relative size of each cluster in the RYB wheel, we defined a sector covering the range of hue spanned by the cluster. The radius (distance from the circumference) of the sector was determined by its weight calculated from the Eq. 5.1. Then, a nonlinear transform of the weight was performed to visualise a small cluster in the RYB colour wheel:

$$w(i) = \begin{cases} \sqrt{0.5} \times \sqrt{w(i)} & w(i) \leq 0.5 \\ w(i) & otherwise \end{cases} \tag{5.2}$$

The nonlinear transform amplified the weight less than 0.5. The arc of the sector shows the range of hue, i.e. the spread of hue, of each cluster. The visualisation of the dominant hue distributions in Fig. 5.5a–c is shown in Fig. 5.5e. This characterises the colour scheme used in the painting.

5.2.3 Dominant Hue Alignment and Mapping

After extracting the dominant hues and finding the hue spread width and the mean value for each dominant hue region in the photographic image, they were compared to those extracted from the painting. The pairing of dominant hues between the painting and photograph was based on their relative position in the colour wheel after alignment.

The basic principle of the alignment is to minimise the total rotation angle of all the dominant hues in the colour wheel. m_{r1}, m_{r2}, m_{r3} denote the mean positions of dominant hues of the reference painting. And m_{s1}, m_{s2}, m_{s3} denote the mean positions of dominant hues of the target photograph. The optimisation alignment rotated one of the dominant hues of the reference painting $m_{rp}(p = 1/2/3)$ to the position (mean value) of one dominant hue of the photograph $m_{sq}(q = 1/2/3)$ to minimise the total sum of rotation angles, i.e.,

$$\arg \min_{\alpha, p, q} \sum_{M_{ij}} \alpha_{M_{ij}}, i, j = 1, 2, 3 \tag{5.3}$$

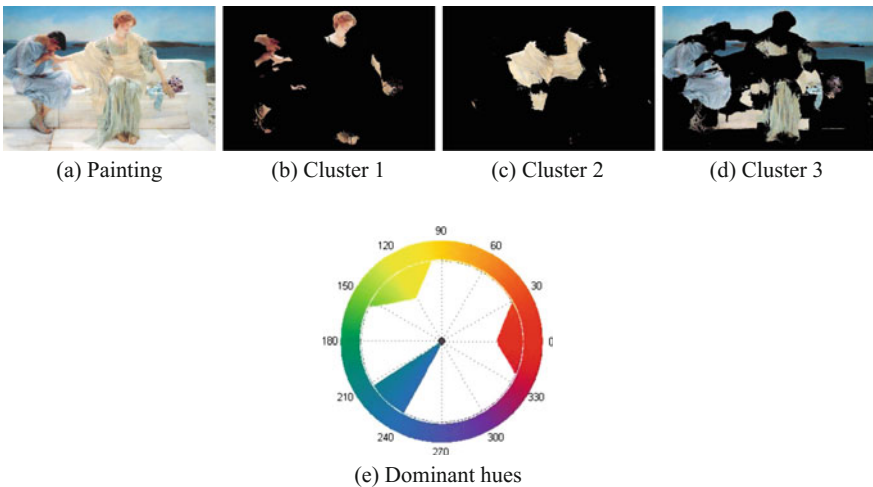


Fig. 5.5 Dominant hue extraction of the painting ‘Ask Me No More’ in a

M_{ij} is the matched pair of the i -th hue of the reference with the j -th hue of the source which have the nearest distance after the dominant hues of the reference rotate a angle α . $\alpha_{M_{ij}}$ is the rotation angle of matched pair M_{ij} , which is $\alpha_{M_{ij}} = (m_{ri} + \alpha - m_{sj}) \bmod 360$.

The hue distribution of each dominant hue of the source was scaled by the hue spread of the corresponding dominant hue in the reference painting and rotated to the mean position of the aligned corresponding dominant hue. The mapping function is

$$H(i) = \left\{ [sH(i) - smean] \times \frac{wdr}{wds} + rmean + \alpha \right\} \quad (5.4)$$

where $sH(i)$ is the i -th dominant hue distribution of the source image with mean value $smean$. $rmean$ is the mean value of the aligned corresponding dominant hue of the reference. wdr and wds are the spread widths of the reference dominant hue and source dominant hue, respectively. The final mapped hue value will be $H(i) \bmod 360$.

When there were 2 clusters in the reference and 3 clusters in the source, two source dominant hues were paired with the same dominant hue of the reference. In this case, the two source dominant hues were rotated to the coverage range of the paired reference dominant hue, instead of rotating to the mean value.

5.3 Saturation and Lightness Mapping

The histogram distributions of lightness and saturation reflected the global saturation and lightness contrast of the images. So the overall saturation and lightness distributions were extracted. The image was converted from RGB to HSV space, and the histograms in S and V were used to represent the overall distribution of saturation and lightness, respectively. Histogram matching was used to map the saturation and lightness values in the photographic image to those from a painting. By doing so, the distributions of saturation and lightness were changed to match those of the painting. This accomplishes the transfer of saturation and lightness attributes of the style.

5.4 Experiments and Discussion

Paintings by Van Gogh and Alma-Tadema were used as reference in our experiment to test the proposed algorithm. The saturation and lightness distributions of these paintings were extracted in HSV colour space, and the dominant hues were extracted in the RYB colour space. Four target photographs were used (Fig. 5.6) to show the effectiveness of the proposed algorithm. They were chosen for the naturalness of their content: featuring skies, water and foliage. Such content was particularly vulnerable to shifts in hue, which qualified them as suitable test subjects. For a



Fig. 5.6 The four target photographs used in our experiments

given photographic image, similar attributes were computed. Histogram matching was performed on the lightness and saturation attributes first. Following this, the hue mapping was performed using the method in Sect. 5.2 on the saturation and lightness modified image.

Figures 5.7, 5.8, 5.9 and 5.10 show four paintings (reference paintings 1–4) and their hue, saturation and lightness distributions. The first two paintings in Figs. 5.7 and 5.8 are by Alma-Tadema. In them, we see large regions of low saturation and high lightness, with a consequent lack of perceptually apparent hue. In the centre regions, there are rich colours. Reference painting 1 contains red, yellow and green. Reference painting 2 spreads from blue-purple to red to orange. The clothes and skin of these two paintings have high saturation contrast and high lightness contrast. The consistent saturation and lightness style are clearly discernible within the relative histogram distributions of these paintings as shown in Figs. 5.7 and 5.8c, d.

The last two paintings in Figs. 5.9 and 5.10 are by Van Gogh. Vivid red, yellow and green can be clearly seen in the painting ‘The Poppy Field’ (Fig. 5.9), and vivid blue and yellow are apparent in the painting ‘Starry Night Over the Rhone’ (Fig. 5.10). Compared with Alma-Tadema’s paintings, the saturation and lightness distribution are even and therefore of low contrast.

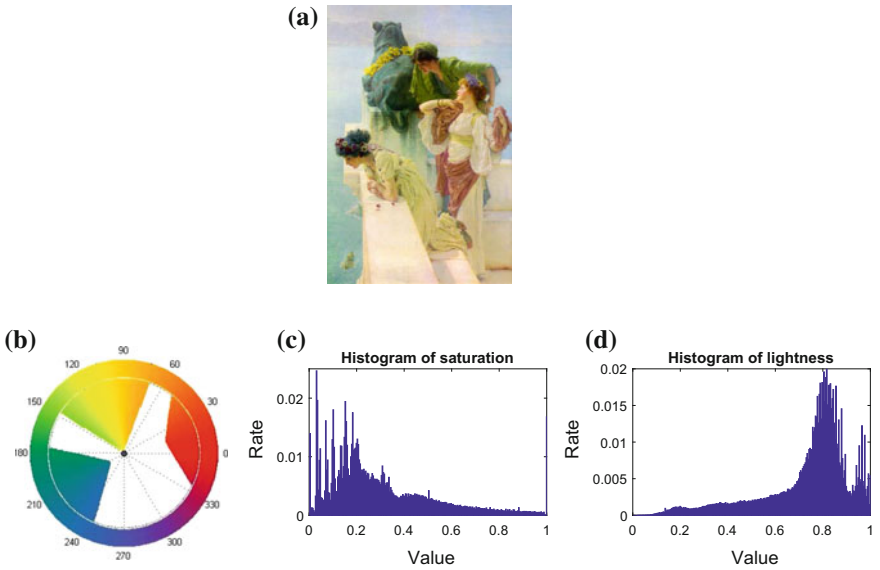


Fig. 5.7 Reference painting 1. Hue (b), saturation (c) and lightness (d) distributions of reference painting 1 (a): Alma-Tadema's 'A Coign of Vantage' (1895), Courtesy The Athenaeum

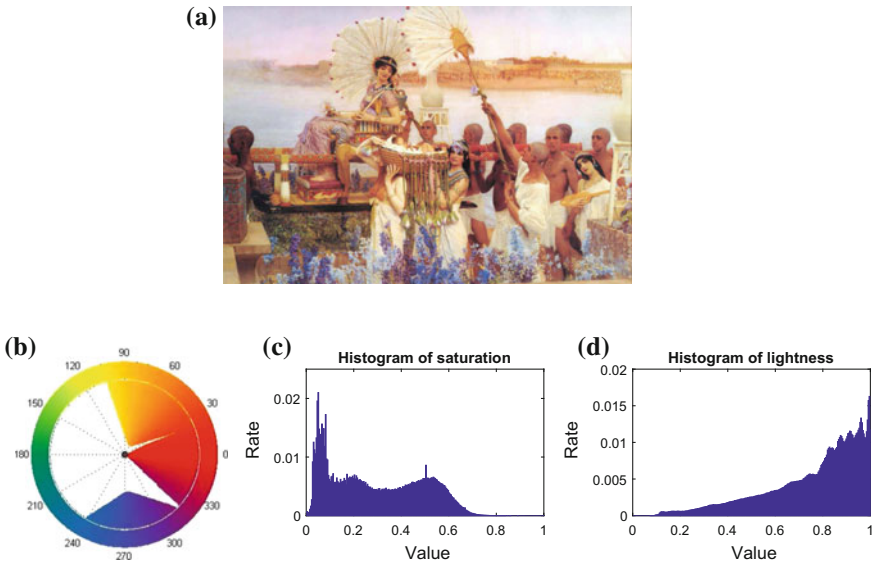


Fig. 5.8 Reference painting 2. Hue (b), saturation (c) and lightness (d) distributions of reference painting 2 (a): Alma-Tadema's 'The Finding of Moses' (1904), Courtesy Wikimedia Commons

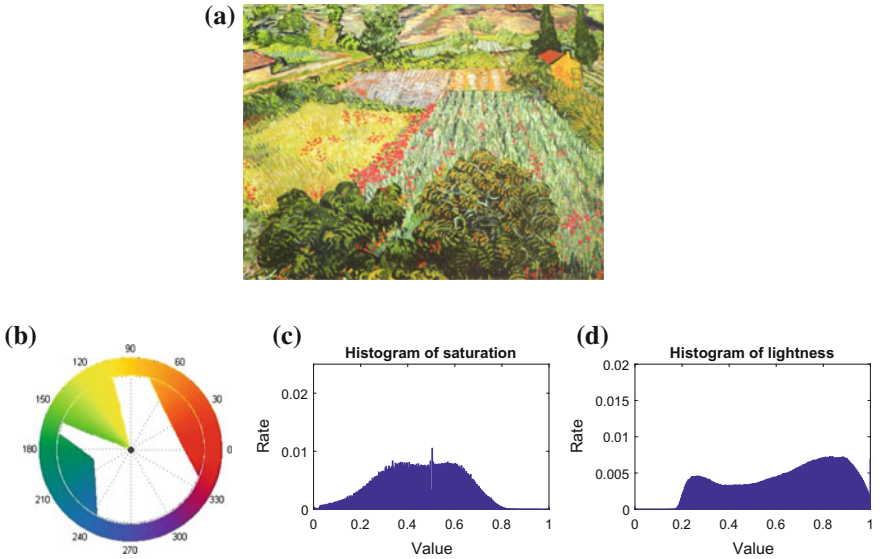


Fig. 5.9 Reference painting 3. Hue (b), saturation (c) and lightness (d) distributions of reference painting 3 (a): Van Gogh's 'The Poppy Field' (1888), Courtesy Wikimedia Commons

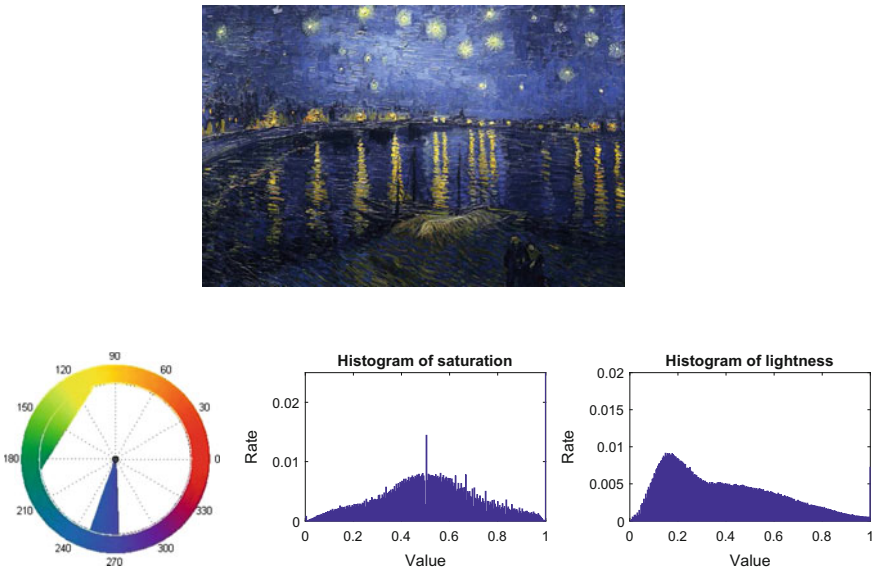


Fig. 5.10 Reference painting 4. Hue (b), saturation (c) and lightness (d) distributions of reference painting 4 (a): Van Gogh's 'Starry Night Over the Rhone' (1888), Courtesy of Tianyu Lu, Flickr



Fig. 5.11 Painting style transfer results using two reference paintings which have distinct different saturation and lightness contrast. References used are: painting 1 (shown in Fig. 5.7) and painting 3 (shown in Fig. 5.9)

The results shown in Fig. 5.11 use the reference painting 1 (shown in Fig. 5.7) and painting 3 (shown in Fig. 5.9). The two reference paintings are of similar colour schemes; however, the spread width and relative positions of their dominant hues are different, and also their saturation and lightness contrast distributions are different.

Comparing the results in Fig. 5.11d with e, and Fig. 5.11g with h, which are from the same target photograph using different references, we can see obvious difference in their colour appearance. The saturation and lightness distributions before and after style transfer of the source image 1 in Fig. 5.11 are compared with the distribution of the reference, as shown in Figs. 5.13 and 5.14c, d. It can be seen that the saturation and lightness distributions of the style-transferred images are similar to the corresponding reference paintings, even the hue mapping in RYB space has some influence on the saturation and lightness distributions. In order to show the influence of hue mapping, the lightness and saturation matching results before hue mapping of the

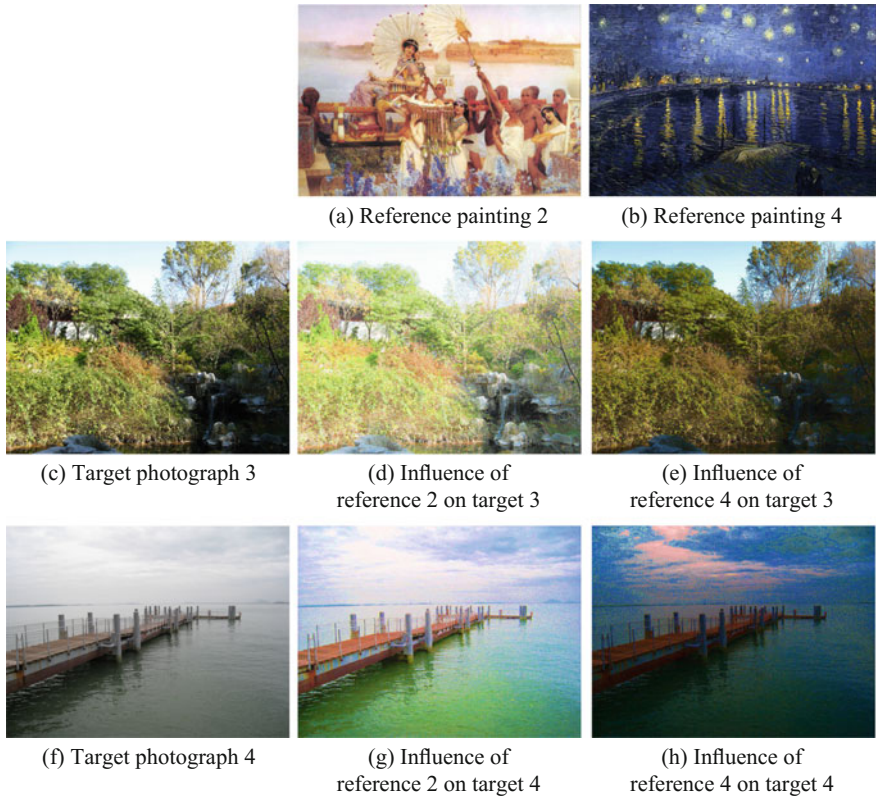


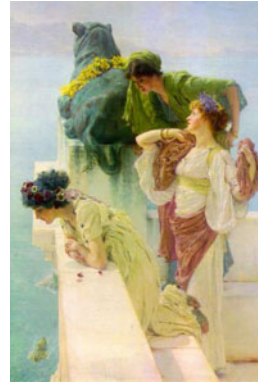
Fig. 5.12 Painting style transfer results using two reference paintings which have distinct different saturation and lightness contrast. References used are: painting 2 (shown in Fig. 5.8) and painting 4 (shown in Fig. 5.10)

target photograph 1 using references 1 and 3 are shown in Figs. 5.13 and 5.14e. Compared with the results after hue mapping in Fig. 5.11d, e, clear colour difference can be seen in the green grass in Fig. 5.11d and in the sky in Fig. 5.11e. However, the adjusted colour is not unnatural as it has received only a small rotation to its neighbouring colour in order to adjust the global contrast. But when hue mapping was performed in the hue wheel from RGB space, the results appear strange as the hue arrangement in RGB is less perception coherent (see the results in Figs. 5.13 and 5.14f).

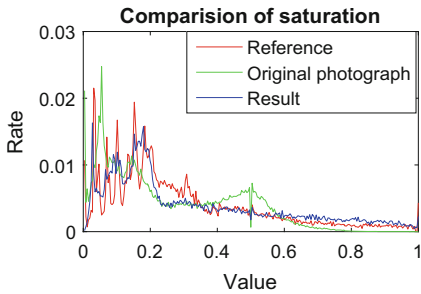
The experiment in Fig. 5.12 used the reference painting 2 in Fig. 5.8 and painting 4 in Fig. 5.10. The two reference paintings have distinct differences in hue distribution. Painting 2 has a wide spread of hue from purple, red, orange, yellow. In contrast, painting 4 has a narrow blue and yellow-green. From the result in Fig. 5.12d, we can see that the resultant image using reference 1 has a wide colour spread around the



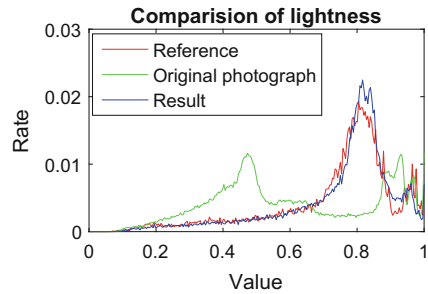
(a) Target photograph 1



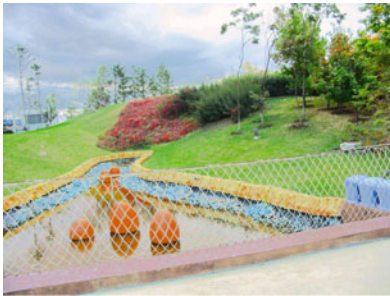
(b) Reference painting 1



(c) Comparison of saturation



(d) Comparison of lightness



(e) Mapping of lightness and saturation



(f) Mapping of hue in RGB space

Fig. 5.13 Comparisons of the mapping of saturation, lightness and hue: Target photograph 1 to reference painting 1

dominant colour, through orange, yellow, green and blue. In appearance, the result from reference 4 has a narrow colour spread, green and blue (Fig. 5.12e).

The experiment results confirm the advantages of our proposed approach over existing approaches, such as harmonisation and recolour. The operation of hue mapping in our research is similar to harmonisation; however, it is based on an example painting instead of a template, and the hue is measured in artist’s colour space to

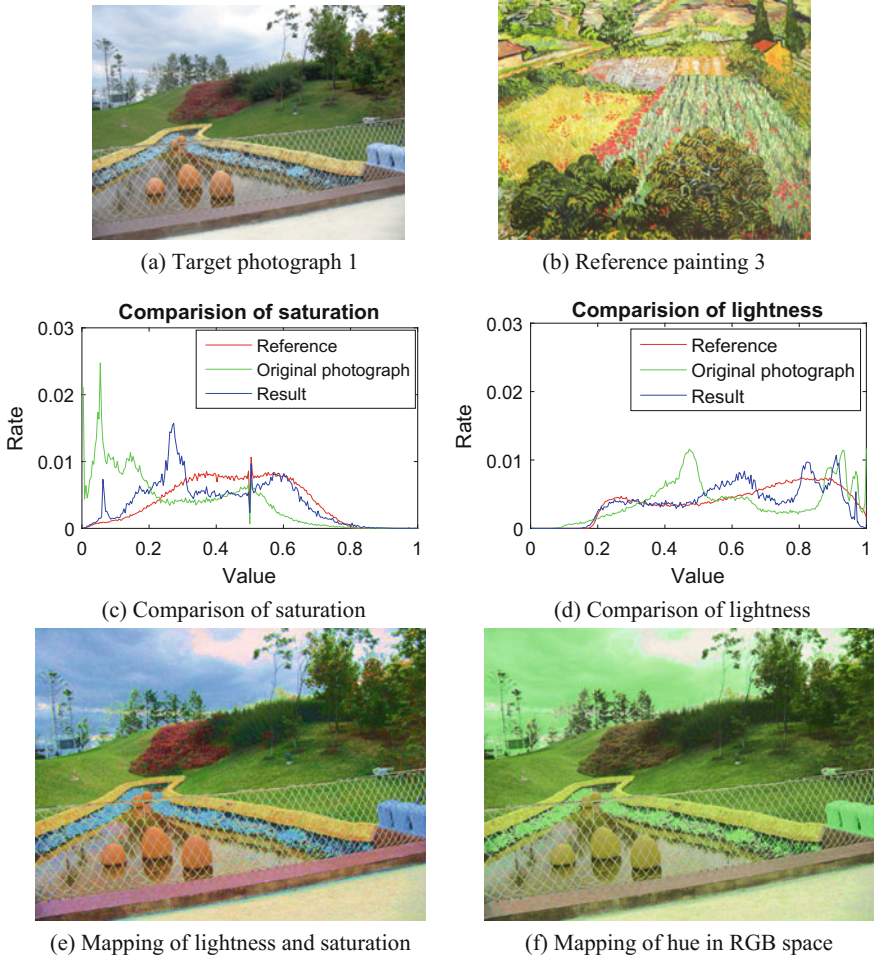


Fig. 5.14 Comparisons of the mapping of saturation, lightness and hue: Target photograph 1 to reference painting 3

express the colour contrast as perceived and practiced by an artist. Compared with results derived from a simple recolor, our results do not transfer the colours of the reference painting to the target photograph. Instead, they match the global contrast, which expresses the relative position distribution in the wheel, as opposed to its absolute value.

However, the proposed algorithm has some limitations. First, there is not a uniform colour space to express the hue, saturation and lightness values by which an artist operates. We mapped the hues in RYB space and the saturation and lightness in HSV space converted from RGB. The three attributes are therefore handled separately. In some cases, this produced artefacts in the results. In future work, we will consider

a hue depended lightness and saturation mapping in a uniform artist colour space. Second, the method can only transfer the global style of the painting through attention to the global contrast, without the consideration of the content and centre of attention. In future research, local contrast will be considered as a factor in the style of painting.

5.5 Concluding Remarks

This chapter presented a method of transferring the global colour contrast style of a painting to a photographic image. The contrast style of a painting was considered for its hue, saturation and lightness. The transfer of saturation and lightness attributes was by histogram matching using the histograms obtained from the HSV space converted from RGB. The last step of the transfer was hue mapping. We studied the painter's RYB colour wheel and extended the bilateral conversion between RGB and RYB. The dominant hues were extracted from discernible hues using k-means clustering to constitute the colour schemes driven by colour contrast. Dominant hue pairing was determined by optimisation alignment, and the hue spread width defined the scaling of hue range. Experiments on a number of photographic images with paintings from two well-known artists demonstrated the feasibility and performance of our proposed style transfer method.

References

1. Chang, Y., Saito, S., Nakajima, M.: Example-based colour transformation of image and video using basic colour categories. *IEEE Trans. Image Process.* **16**(2), 329–336 (2007)
2. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. In: *ACM SIGGRAPH*, pp. 624–630 (2006)
3. Comaniciu, D., Meer, P.: Mean-shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 603–619 (2002)
4. Gossett, N., Chen, B.Q.: Paint inspired colour mixing and compositing for visualization. In: *Proceedings of the IEEE Symposium on Information Visualization*, pp. 113–118(2004)
5. Greenfield, G.R., House, D.H.: A Palette-driven approach to image colour transfer. In: *Computational Aesthetics in Graphics, Visualization and Imaging*, pp. 91–98 (2005)
6. Michael, T.H.: *Scientific Computing an Introductory Survey*, 2nd edn. McGraw Hill, New York (1997)
7. Neumann, L., Neumann, A.: Colour style transfer techniques using hue, lightness and saturation histogram matching. In: *Computational Aesthetics in Graphics, Visualization and Imaging*, pp. 111–122. (2005)
8. Tai, Y.W., Jia, J.Y., Tang, K.C.: Local colour transfer via probabilistic segmentation by expectation-maximization. In: *Proceedings 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 747–754 (2005)

Chapter 6

Atmospheric Perspective Effect Transfer for Landscape Photographs

Abstract The atmospheric perspective effect is a physical phenomenon relating to the effect that atmosphere has on distant objects, causing them to be lighter and less distinct. The exaggeration of this effect by artists in 2D images increases the illusion of depth, thereby making the image more appealing. This chapter addresses the enhancement of the atmospheric perspective effect in landscape photographs, by the manipulation of depth-aware lightness and saturation contrast values. The form of this manipulation follows the organisation of such contrast in landscape paintings. The rationale behind this manipulation is based on a statistical study which has shown clearly that the saturation contrast and lightness contrast between and within the depth planes in paintings are more purposefully organised than those in photographs. This contrast organisation in paintings respects the existing contrast relationships within a natural scene governed by the atmospheric perspective effect, yet also exaggerates upon them. In our approach, the depth-aware lightness and saturation contrast revealed in landscape paintings guides the mapping of contrasts in photographs. This contrast mapping is formulated as an optimisation problem that simultaneously considers the desired inter-contrast, intra-contrast, and some gradient constraints. Experimental results demonstrate that by using this proposed method, both the visual appeal and the illusion of depth in the photographs are effectively improved.

Keywords Atmospheric perspective · Vignetting · Landscapes · Painting · Style

6.1 Introduction

Atmospheric perspective is an effect that the atmosphere has on the appearance of those aspects of a natural scene that are far from a viewer. It is a physical phenomenon derived from the scattering and absorption of light as it travels from the source to the viewer. As the distance between an object and the viewer increases, both the

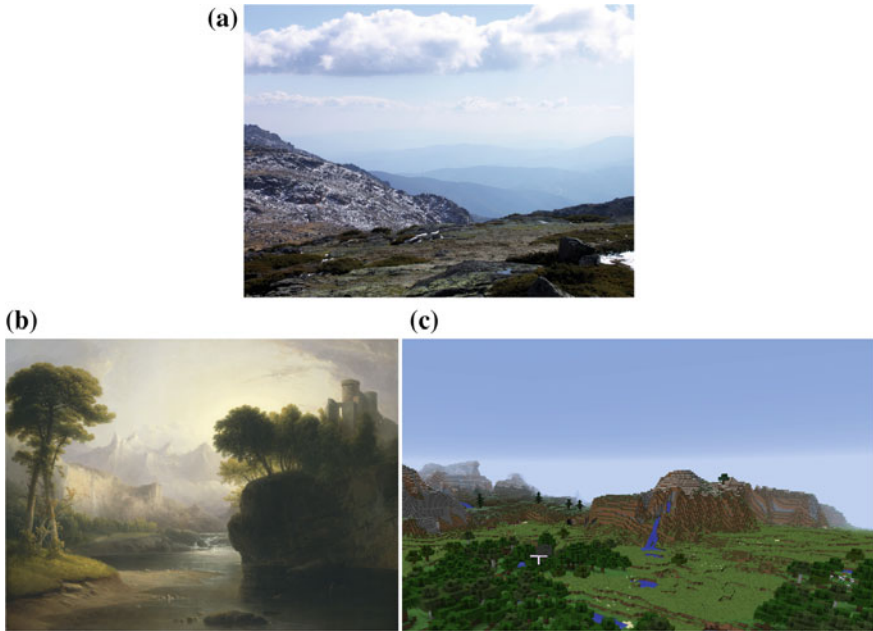


Fig. 6.1 **a** The atmospheric effect as evident in a natural scene (Courtesy Wikimedia Commons). **b** The atmospheric effect as evident in a painting (An example of a vignette. ‘Fanciful Landscape’, Thomas Doughty, 1834 (Courtesy National Gallery of Art, Washington). **c** Atmospheric effect in the game ‘Minecraft’

contrast between objects and background (inter-contrast) and the contrast within objects (intra-contrast) decrease.

In images, one function of atmospheric perspective is to provide the viewer with a cue from which they can infer the illusion of depth within a 2D image [25]. Additionally, the atmospheric perspective effect can make images more visually interesting and exciting by eliciting the impression of dramatic weather. These attributes make the depiction and exaggeration of atmospheric perspective a popular approach in the design of art, films and games. For example, it is used in digital games, where it is calculated ‘on the fly’ as the player walks through an environment (Fig. 6.1c).

Atmospheric perspective exists even on clear days. However, when capturing photographs, the lighting conditions, the skill of the camera operator and the settings of the camera can all have an adverse effect upon the appearance of the photograph, which leads to a flat, tonally compressed, washed-out and generally unappealing visual result. Many global and local image adjustment methods, and automatic and template-driven image improvement software (e.g. Instagram, Adobe Photoshop) have been developed to enhance the appearance of the captured photograph. One common shortcoming of these methods is that they perform an overall adjustment on the surface of the image. Though it is clear that a photograph is flat and rectangular in shape, consisting of such two dimensional attributes as corners, centres, edges and

regions, it is also clear that a photograph has an implicit depth dimension. Depth is of particular importance in landscape images. Thus, this chapter proposes to improve the visual appeal and illusion of depth in landscape photographs by enhancing the atmospheric perspective effect.

Atmospheric perspective has been physically modelled by researchers [23, 24]. However, these rely upon the atmospheric condition and attenuation strength being determined for a specific photograph. Jung et al. used a physical model of atmospheric perspective to improve the depth perception of a 2D image [17]. This physical model increases the atmosphere effect based on the absolute distance of objects in the scene from the camera. Hence, the contrast of both close objects and distant objects will be reduced by increasing the atmosphere effect. Additionally, users have to decide how much to enhance the atmosphere effect in order to create the desired result. However, for a flat and low-contrast photograph with or without atmospheric perspective, the contrast of close objects generally needs to be increased, while adjusting the relative difference between objects. Although such sophisticated software as Photoshop provides powerful tools for users to manually edit images, the skills needed to perform such an improvement across the depth plane are beyond the casual user. Even for an experienced user, it is far from a trivial task. Users have to mask regions, adjust each region differently and make decisions based on the relative relationships of these regions. Clearly, the number of possible permutations involved is legion and the chance of producing an adjustment to the image that damages its visual integrity is high.

Fortunately, painters have provided us plenty of meaningfully enhanced examples of atmospheric perspective from which much can be learned. In painting, atmospheric perspective is used to create the illusion of depth by simulating the natural changes effected by the atmosphere on the value of objects seen at a distance. The depth in a painting is of a simplified form, consisting of the traditional discrete depth planes: foreground (FG), middle-ground (MG), background (BG) and sky [30]. The atmospheric perspective is exaggerated by painters in order to strengthen the apparent separation of depth planes, with a view to increasing the illusion of depth [11]. It is through a consideration of the relative contrast values of these depth planes that the atmospheric perspective effect is realised [11].

Therefore, this chapter proposes to learn how painters organise the exaggeration of the atmospheric perspective effect in landscape paintings by analysing the lightness and saturation contrast values according to traditional depth planes. This chapter also explores how collections of these contrast organisations can be passed onto acquired photographs through their corresponding depth planes, thus radically enhancing their visual appeal and the illusion of depth.

Painters operate within a world of relative values for the human visual system is more sensitive to contrast than absolute value. In making a painting, a painter does not value the fidelity of paint to its optical source as much as its quality of relative difference to its neighbour upon the canvas [1]. It is an awareness of contrast that drives this process. This is different to the average 'snapshot' landscape photograph, as captured by an amateur, which are often of poor visual quality with little to no high-level aesthetic organisation.

To a painter, the atmospheric perspective effect can be implemented using two forms of depth-aware contrast considerations. One is intra-contrast, which refers to the contrast values within an individual depth plane. The other is inter-contrast, which refers to the contrast values between depth planes. To learn quantitatively how these two types of depth-aware contrasts are organised in paintings, a depth-based contrast analysis of a set of landscape paintings made by the Hudson River school painters was conducted. This analysis was compared to a similar analysis of a collection of ‘snapshot’ style landscape photographs which had been taken in the Hudson River area. Differences between the paintings and the snapshot landscape photographs were observed, with evidence that the intra-contrast and inter-contrast of lightness and saturation of the paintings were more purposefully organised. This contrast organisation in paintings respects the existing contrast relationships within a natural scene and also exaggerates upon them, with a view to improving the visual appeal and illusion of depth within the painting. This observation provides the justification for the proposed enhancement of photographs through separate contrast adjustments to their depth planes. This chapter proposes to use this lightness and saturation contrast organisation across depth planes of paintings as reference to enhance digital photographs.

One of the difficulties in applying depth-aware contrast organisation of paintings is that though general observations of their organisation can be made, it is a challenge to identify exactly the rules by which this organisation operates. Therefore, instead of a rule-driven approach or one driven by a physical model approach, an example-based approach is proposed to enhance the atmospheric perspective effect of photographs. Based on the assumption that images with similar content should have similar contrast arrangement, the example painting is selected from a painting database based on the similarity of its contextual structure to the input photograph. Referring to the selected reference painting as an example, the depth-aware enhancement is then accomplished by manipulating the inter-contrast and intra-contrast of lightness and saturation. One example result is shown in Fig. 6.2.

6.2 Related Work

Depth-aware enhancement. The research detailed in [20] mimicked the shading characteristics of paintings in order to enhance the perceived depth of scenes. This technique is effective when performed on a simple scene consisting of overlapping objects. However, to create the illusion of depth, landscape painters might also employ such approaches as intelligent application of atmospheric perspective, selective application of linear perspective, the selective overlap of objects and regions, and careful selection of point of view. Bailey [3] proposed to simulate the artistic control of apparent depth in an image by changing the lightness or colour on the boundary, or by replacing the colour or lightness of objects to change their perceptual order. This technique is also only suitable for the purpose of changing the perceptual depth of simple overlapped objects. Jung et al. proposed a method by which the

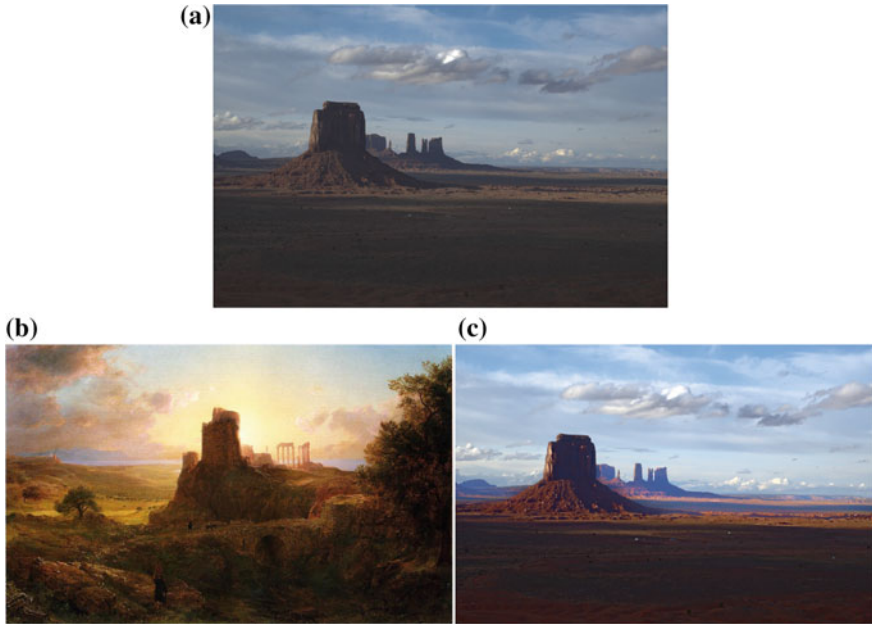


Fig. 6.2 An example result of atmospheric perspective effect enhancement using the proposed method. **a** Input photograph. **b** Reference painting (CC Image courtesy of fredericedwinchurch.org). **c** Result

depth perception of a 2D image might be improved by enhancing such depth cues as: linear perspective, aerial perspective, focus and shadow effects [17]. They made colour transformations based on the atmospheric perspective effect upon the hue, saturation and intensity of the scene. The shadow effects were applied only to the region of interest (ROI) with a blurring of the background performed to enhance the visual focus upon the ROI. This depth perception enhancement method is valid for scenes with a clear focal object. However, in landscapes without such focus, this method is not applicable.

Processing in gradient domain. Contrast processing by operating on gradients has been used in high dynamic range compression [13], contrast enhancement [15, 21] and style enhancement [31]. Specifically, work described in [13, 31] manipulated the gradients of edges to adjust the local contrast. The proposed intra-contrast mapping in this chapter is also conducted by operating the gradients on edges. Differently, the proposed method operates gradients upon the edges deemed to be salient instead of those with a strong magnitude and the length of those edges is used to weight the gradient mapping.

6.3 Paintings Versus Photographs

In order to explore the contrast organisation in paintings, we conducted a comparison of the lightness and saturation contrast organisation of the four depth planes of paintings and photographs. Before introducing the contrast organisation in paintings and photographs, the separation of depth planes is discussed.

6.3.1 *Depth Planes*

Depth exists in infinite degrees between the viewer and the most distant object in a scene (see Chap. 4 on p. 39). However, in practice, artists (painters and professional photographers) consider depth as existing in discrete planes, typically FG, MG and BG [29, 30]. These planes are, effectively, flat ‘cards’ that serve to underpin the artists’ structural conventions. They are particularly important for landscape images, where they are used to enhance the depth of the scene. The FG is the plane, which is closest to the viewer, and usually contains the objects of highest pictorial interest. The BG is the plane that is the most distant from the viewer. The BG also encompasses the sky. In recognition of the importance of the sky, we consider it separately from the BG as a discrete plane. Some landscape paintings have only a FG and sky, and for these, it is meaningless to enhance the atmospheric perspective effect. Therefore, the effect is exclusively targeted at landscapes that feature two (FG and BG), three (FG, MG, sky or FG, BG, sky), or four depth planes (FG, MG, BG and sky). However, as a general case, the proposed method considers landscapes with four depth planes. The application on other landscape categories will be discussed in the experiment section.

How may photographs be partitioned into depth planes? Long-range depth acquisition is an ongoing topic in landscape research. However, in spite of the maturation and commercialisation of such technologies, existing solutions are still unsatisfactory. Some depth estimation algorithms are proposed to explore the depth from single image, e.g. ‘Make3D’ [27]. These methods can roughly explore the depth of a scene, but it is still difficult to accurately segment the depth planes based on this estimated depth, especially that of the MG and BG. Therefore, this chapter proposes to segment the depth planes with minimal user interaction. The interactive segmentation method proposed in [6] can be used to outline the depth planes, or they can be outlined manually. The users only need to give a few strokes for each depth plane, and a mask is generated automatically for each depth plane based on these strokes.

6.3.2 *Contrast in Paintings and Photographs*

In this section, we describe a statistical study which was conducted to measure the inter-contrast and intra-contrast of a group of paintings and a group of photographs.

The group of paintings that was the subject of this study derived from the Hudson River painters. These were a group of stylistically coherent American landscape artists from the late romantic period (the mid-nineteenth century), and all painted approximately the same subject matter (the area around the Hudson River in the USA). Paintings done in that time used a wide range of saturation and lightness values, being approximately equal to the range of lightness and saturation values of photographs. This makes them suitable candidates for the purpose of style transfer. Paintings done in the twentieth century (e.g. Impressionist and Post-Impressionist) would be too shallow in their tonalities for this purpose. Paintings by Hudson River school artists are quintessential examples of masterly atmospheric perspective [22]. The group of photographs were acquired through a simple Google Image search using the search term ‘Hudson River’, which ensured that they were thematically similar to the set of paintings. These search results were culled to remove those that had been taken with special lenses, those that had been obviously ‘Photoshopped’ and those that had been taken through a colour filter. Assuming the saturation had the same property as lightness, the saturation contrast was measured in the same way as was lightness.

The intra-contrast was measured by the gradient range. The gradient range was defined as the width of the gradient histogram of the logarithms of lightness or saturation values. The gradient field in the logarithm domain of lightness directly corresponds to local contrast [13]. The gradient range expresses the maximum local contrast of the plane. Based on the contrast definition in Eq. (3.2) which corresponds to the Weber–Fechner law, the inter-contrast of two planes with mean values M_1, M_2 was defined as $\frac{2(M_2 - M_1)}{M_1 + M_2}$.

The lightness intra-contrast values shown in Fig. 6.3a of the depth planes clearly illustrate a general tendency in both the paintings and the photographs towards a lessening of lightness intra-contrast from the FG to the sky. This tendency is consistent with the natural phenomenon of atmospheric perspective. However, the FG, MG and BG differences in the paintings are far more clearly defined than those in the photographs. In other words, the painters exaggerated the natural tendency of change in the intra-contrast values from the FG to the sky. Additionally, the contrast in the FG is generally higher in the paintings than in the photographs.

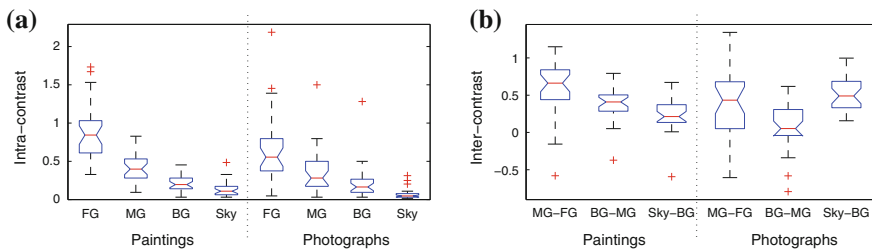


Fig. 6.3 Difference in lightness contrast values between paintings and photographs. **a** Box plots of intra-contrast values. **b** Box plots of inter-contrast values

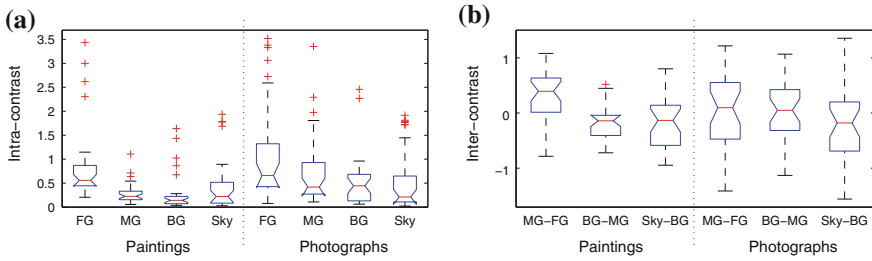


Fig. 6.4 Difference in saturation contrast values between paintings and photographs. **a** Box plots of intra-contrast values. **b** Box plots of inter-contrast values

The difference in their lightness inter-contrasts is more complex. The MG-FG inter-contrast of the paintings is larger than that of the photographs, while the Sky-BG inter-contrast is smaller. Additionally, almost all the MG-FG, BG-MG, Sky-BG inter-contrast values of the paintings are positive. The net result is a clearer decrease in the painting inter-contrast across the Z-axis from FG to Sky (Fig. 6.3b). This is also consistent with the phenomenon of atmospheric perspective.

The saturation contrast values in Fig. 6.4 shows that the paintings are generally more purposely organised in their saturation contrast than the photographs. The FGs of the paintings have a consistently higher saturation intra-contrast than the MGs and BGs. The BG-MG inter-contrast values of saturation in paintings are more likely to be negative than those of photographs (see Fig. 6.4b). This indicates that the saturation is clearly decreasing from MG to BG in paintings. However, a clear decreasing tendency from MG-FG to Sky-BG in the inter-contrast of saturation of both paintings and photographs cannot be observed.

In summary, the study shows that the paintings and photographs exhibit significantly different inter-contrast and intra-contrast organisation across the depth planes in their lightness and saturation values. This difference shows that the paintings are more purposely organised in their contrast management than the photographs. This contrast organisation in paintings corresponds to an exaggerated form of the physical phenomenon of atmospheric perspective.

6.4 Contrast Manipulation

Statistics in the preceding section clearly show intentionally introduced lightness and saturation inter-contrast and intra-contrast organisation in paintings which are not prominent in photographs. This section attempts to use this contrast organisation in paintings as reference to enhance the acquired digital photographs, while keeping the hue unchanged. The pipeline of the proposed method is shown in Fig. 6.5.

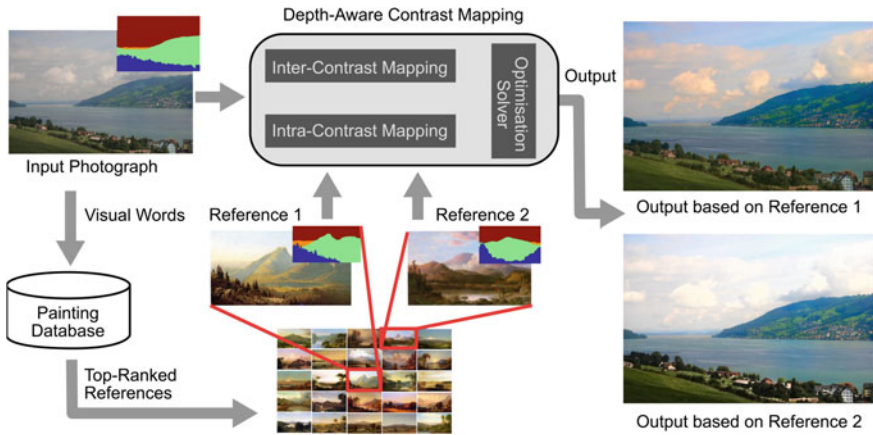


Fig. 6.5 The pipeline of the proposed method

6.4.1 Reference Selection

With the assumption that images with similar content should have similar contrast arrangement, the reference painting and the input photograph should be in visual and structural accordance with each other. If they are not, then the contrast values of the photograph are at risk of being changed more than is visually acceptable. Therefore, the reference painting was selected on the basis of having similar global contextual structure (i.e. it had similar objects such as trees, river, mountain and sky, and similar object arrangements within the scene) and similar content in each depth plane to that of the input photograph.

The image retrieval technique used a quantised SIFT features and spatial pyramid matching scheme that approximated geometric correspondence and has demonstrated effectiveness in retrieving semantically similar outdoor images [9, 16, 18]. In the proposed reference selection scheme, this technique was used to calculate the global contextual similarity and contextual similarity in depth planes. Specifically, a vocabulary of 200 words and four-level pyramid descriptor was used to calculate the global contextual similarity. The contextual similarity of each depth plane was measured using a vocabulary of 50 words and one-level pyramid descriptor. In addition, a reference having a similar lightness level to that of the input photograph in each depth plane was preferred. This would return a reference/input pair that were more likely to have been ‘captured’ in similar environmental situations. This lightness constraint helped avoid changing the input too far away from its original appearance. The lightness of each depth plane was described by a 10 bin histogram. The similarity of the lightness histograms and the similarity of pyramid descriptors were calculated by the histogram intersection [18]. The final matching score was the linear combination of the spatial pyramid similarity K_t and lightness similarity K_l by a parameter $\beta \in [0, 1]$, which was defined as $K = \beta K_t + (1 - \beta) K_l$. By experimentation, $\beta = 0.8$ was found to consistently produce good results.

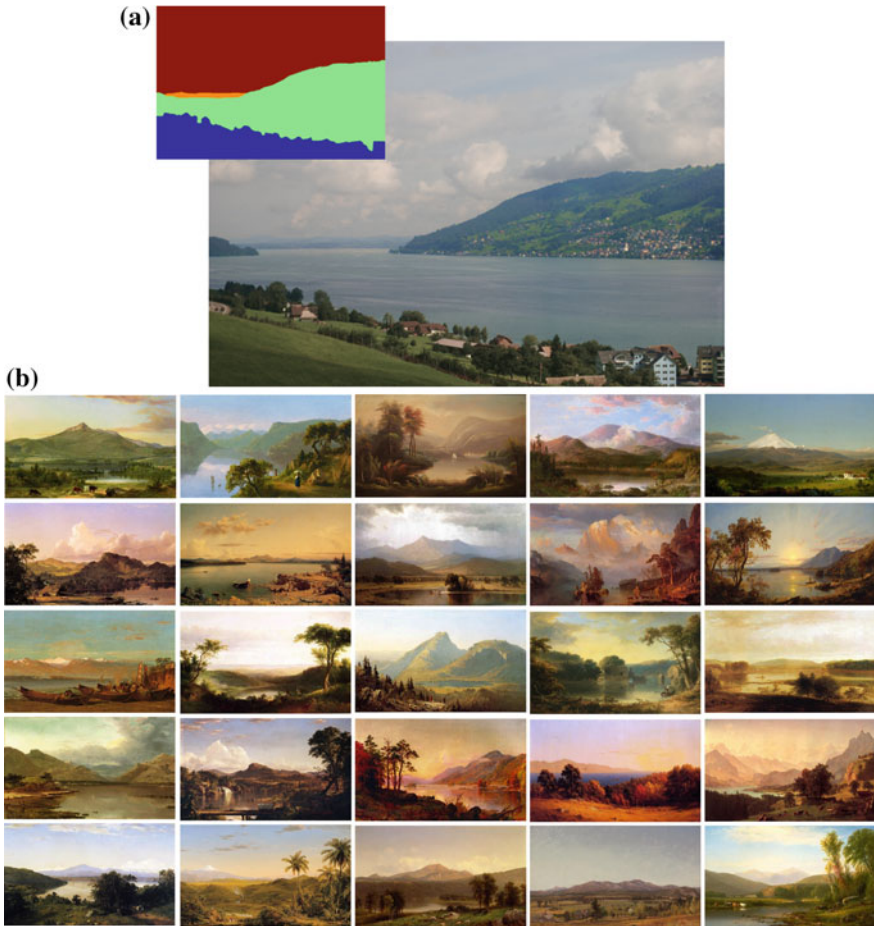


Fig. 6.6 Reference recommendation. **a** Input with mask. **b** Recommended references

Finally, the top n ranked paintings were selected as the recommended references. 800 Hudson River paintings with four clear depth planes were collected to form the painting database. Figure 6.6 shows the 25 recommended references for one input photograph. We can see the recommended top-ranked references have a similar semantic structure with the input photograph. The user can select one of the recommended paintings as a reference to manipulate the contrast. Note that, hue similarity is not required here and hence the selected reference is not of similar colour to the input photograph.

6.4.2 Depth-Aware Contrast Mapping

After selecting the reference painting, the lightness and saturation inter-contrast and intra-contrast of the photograph were manipulated based on the corresponding values in the reference painting. The operations on lightness and saturation were performed separately.

The depth-aware contrast mapping had two objectives. One was to manipulate the mean values of the four depth planes in the photograph so as they align with the inter-contrast values of the reference painting. The other was to adjust the local contrast of each depth plane so as it matched the intra-contrast value from the reference painting. To satisfy these two objectives, the depth-aware contrast mapping was modelled as an optimisation problem. The energy function, $E(\cdot)$, was defined as follows:

$$E(f) = \sum_{i \in f} [E_d(i) + E_g(i) + E_s(i)] \quad (6.1)$$

where f was the output image having the target inter-contrast and intra-contrast. i was a pixel in f . The first term was the inter-contrast mapping constraint, which was defined as:

$$E_d(i) = w_d(i)[f(i) - d(i)]^2 \quad (6.2)$$

where d was the desired image having the inter-contrast values of the reference painting. w_d was the weight for the inter-contrast mapping constraint.

The second term was the cost function of intra-contrast mapping. Based on the definition of intra-contrast, the intra-contrast mapping operates on the gradients. Since gradients along edges can sufficiently affect the overall local image contrast, only the gradients on the edges were mapped to those of the reference while the gradients elsewhere were preserved. Operating the gradients only on the edges can also avoid the block artefacts produced by the down sampling of images and the noise caused by the canvas texture in paintings. Therefore, the second term was defined as:

$$E_g(i) = w_g(i)[f_x(i) - g_x(i)]^2 + w_g(i)[f_y(i) - g_y(i)]^2 \quad (6.3)$$

where f_x and f_y denoted the x and y derivative of f . g_x and g_y were the target gradient after intra-contrast mapping. w_g was the weight for the gradient cost.

$E_s(i)$ was a gradient constraint term, which was defined as

$$E_s(i) = w_s(i) \left[\frac{|g_x(i)|^2}{|F_x(i)|^\theta + \varepsilon} + \frac{|g_y(i)|^2}{|F_y(i)|^\theta + \varepsilon} \right] \quad (6.4)$$

The gradient constraint served to penalise large changes to the original gradient. (F_x, F_y) was the gradient of the original input F . θ controlled the sensitivity to the gradient of F . ε was a small regularisation constant. The weight w_s was adaptively defined based on the enhancement scale of the gradient in the intra-contrast mapping.

6.4.2.1 Inter-Contrast Mapping

Inter-contrast measures the relative difference of mean values of depth planes. Therefore, inter-contrast mapping can be conducted by moving the mean value of the input depth plane as a proportion of the mean value in the corresponding reference depth plane. Given one depth plane I of the input F (L or C channel in LCH space) with mean value m_s and the corresponding plane I_r of the reference with mean value m_r , the target mean value after inter-contrast mapping was

$$m'_s = \rho m_r \quad (6.5)$$

The scale ρ was set based on the desired change to the absolute lightness or saturation values of the output image. In addition to contrast, the absolute value also played a role in the appearance of the image. Given the mean value of the whole reference image M_r and the target mean value of the whole output image M_t , then $\rho = \frac{M_t}{M_r}$. The mean value of the output was adjusted by $M_t = M_s + q(M_r - M_s)$. M_s was the mean value of the input F . q was a scale in $[0-1]$. When $q = 0$, the mean value of the output image was preserved to be in accordance with its original value, and when $q = 1$, it was moved to that of the reference. q was used by the user to control the desired enhancement of the mean value of the whole image.

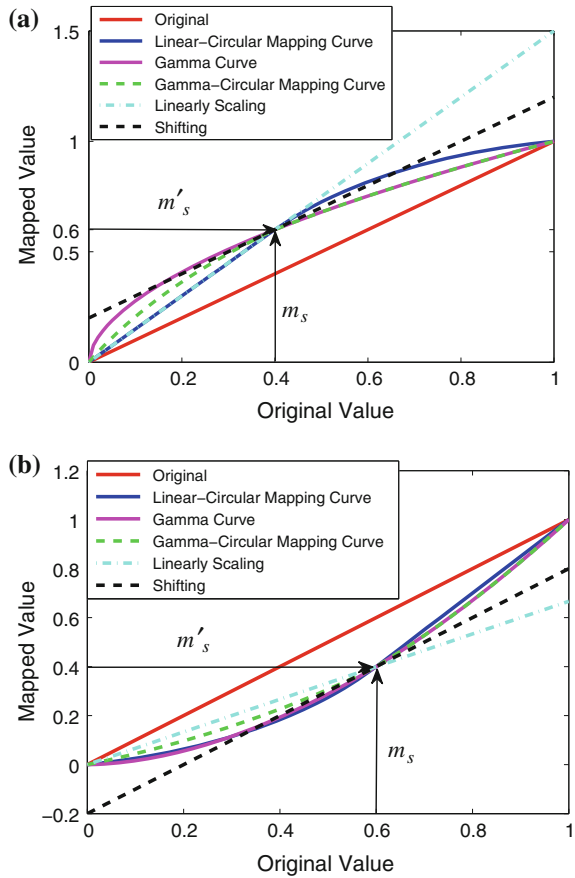
To directly map the mean values, mean value shifting or linearly scaling could have been used. However, these methods would have extended the values out of range $[0, 1]$ and over-compressed or over-exaggerated the visual contrast, as the curve shown in Fig. 6.7. Also, they were likely to produce artefacts in over-exposed areas (see Fig. 6.8c, d). Therefore, a nonlinear transform was proposed. The gamma curve is a nonlinear transform that is often applied to correct lightness. However, by its use, it is easy to over-enhance dark areas which often have hidden block or ring noise. In addition, it also compresses the visual contrast as it enhances more within dark zones than within bright zones (see the magenta curve in Fig. 6.7). To avoid over-enhancing of dark zones and to preserve the visual contrast more effectively, a linear-circular mapping curve (the blue curve in Fig. 6.7) was used to map the input plane to the target mean value. When $m'_s > m_s$, the mapping function was

$$I'(i) = \begin{cases} \frac{m'_s}{m_s} \cdot I(i) & \text{if } I(i) \leq m_s \\ y_0 + \sqrt{r - [I(i) - x_0]^2} & \text{if } I(i) > m_s \end{cases} \quad (6.6)$$

where x_0 , y_0 , r were calculated from the circle $(x - x_0)^2 + (y - y_0)^2 = r$ which passed point $(1, 1)$ and was tangent with the linear line in point (m_s, m'_s) . When $m'_s < m_s$, the mapping function was

$$I'(i) = \begin{cases} y_0 - \sqrt{r - [I(i) - x_0]^2} & \text{if } I(i) \leq m_s \\ \frac{1 - m'_s}{1 - m_s} \cdot [I(i) - m_s] + m'_s & \text{if } I(i) > m_s \end{cases} \quad (6.7)$$

Fig. 6.7 The inter-contrast mapping curves. **a** $m'_s > m_s$.
b $m'_s < m_s$



Here, x_0, y_0, r were calculated from the circle $(x - x_0)^2 + (y - y_0)^2 = r$ which passes point $(0,0)$ and was tangent with the linear line in point (m_s, m'_s) . After one iteration of the mapping, the mean value was close to m'_s , but usually not exactly equal to m'_s . Within 3–5 iterations, the mean value usually successfully matched m'_s .

In addition to the linear-circular mapping curve, the gamma-circular mapping curve proposed in [7] was also employed. This mapping curve (the green curve in Fig. 6.7) was designed to avoid over-enhancement of the hidden artefacts in low value zones and preserve the contrast information in high value zones more effectively.

When $m'_s < m_s$, the proposed mapping curve is similar to the gamma curve (see Fig. 6.7b). It can properly adjust the lightness contrast in an over-exposed scene while avoiding artefacts, as the results shown in Fig. 6.8f. In this case, the gamma-circular nonlinear transform produced a similar result to that of the proposed method (comparing Fig. 6.8e, f). When increasing the lightness ($m'_s > m_s$), the proposed method preserved the perceptual local contrast much better than gamma correction and gamma-circular mapping, as the example shown in Fig. 6.9.



Fig. 6.8 Comparison of mapping curves. **a** Input. **b** Reference (CC Image courtesy The Athenaeum). **c** Shifting. **d** Linearly scaling. **e** Gamma-circular. **f** Linear-circular. $q = 0.5$ in c–f

Due to the mean values of the four depth planes having been adjusted separately, the transition around the boundaries between the depth planes was not smooth. To solve this problem, a soft transition map T was created for each plane using the constraint propagation method proposed in [19]. The constraint weight w was

$$w(i) = \begin{cases} 0 & \text{if } i \in BA \\ 1 & \text{otherwise} \end{cases} \quad (6.8)$$

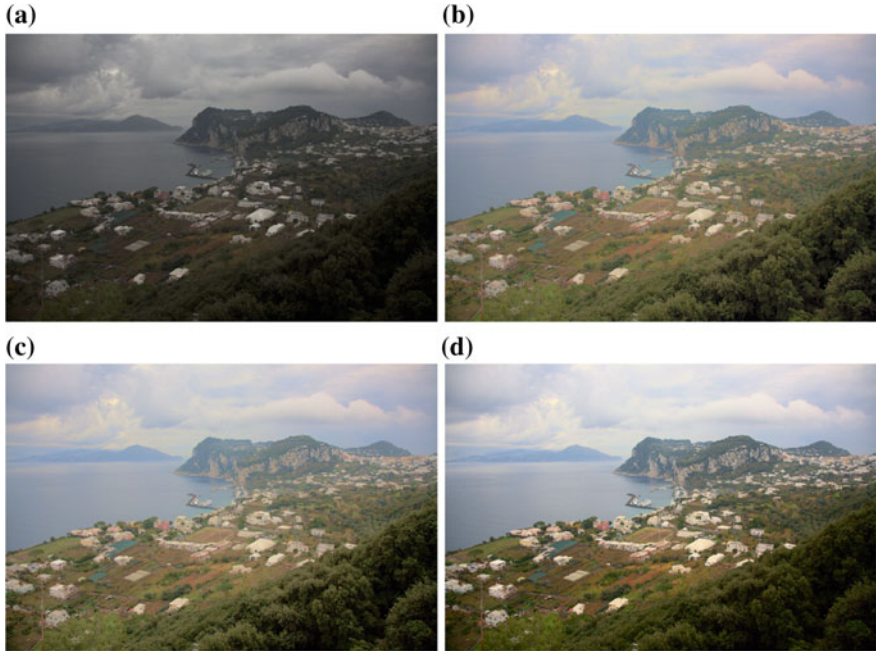


Fig. 6.9 Comparison of mapping curves. **a** Original. **b** Gamma correction. **c** Gamma-circular. **d** Linear-circular. $m'_s > m_s$

where BA was the area around the boundary partitioning planes. BA was created by dilating the boundary with a 21×21 structuring element. Given the inter-contrast mapped results of the four depth planes, I'_f, I'_m, I'_b, I'_s , and the transition maps, T_f, T_m, T_b, T_s , the final inter-contrast mapping result of the input F was

$$F' = \frac{I'_f * T_f + I'_m * T_m + I'_b * T_b + I'_s * T_s}{T_f + T_m + T_b + T_s} \tag{6.9}$$

In the energy function Eq. (6.1), the desired image having the inter-contrast values of the reference painting was $d = F'$.

6.4.2.2 Intra-Contrast Mapping

According to perceptual studies [4], long coherent edges are perceptually salient to the human visual system even when their gradients are weak. Therefore, intra-contrast mapping was operated on long and coherent edges (i.e. salient edges) instead of those with only strong magnitude. The long-edge detector proposed in [5] was used to detect the length of the dominant edge running through each pixel e^l and the local orientation of the dominant edge at each pixel e^o .

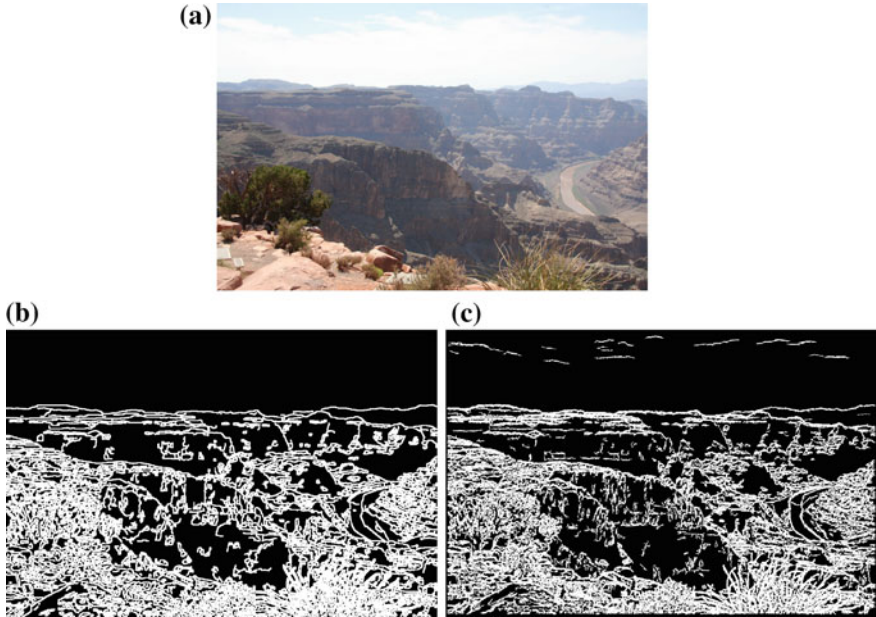


Fig. 6.10 a Source photograph. b Edge detected by Canny detector and c detected by Salient edges

Given the gradient (L_x, L_y) of the lightness L , the local gradient saliency was calculated by

$$S(i) = \sqrt{s_x(i)^2 + s_y(i)^2} \quad (6.10)$$

$$s_x(i) = \cos^2(e^o(i)) \cdot e^l(i) \cdot L_x(i) \quad (6.11)$$

$$s_y(i) = \sin^2(e^o(i)) \cdot e^l(i) \cdot L_y(i) \quad (6.12)$$

Then, the salient edges were detected as:

$$W(i) = \begin{cases} 1 & \text{if } S(i) > h \\ 0 & \text{otherwise} \end{cases} \quad (6.13)$$

where W was the mask of detected salient edges, and h was the saliency threshold. h can be calculated from $\Theta(h) = t$, where Θ was the cumulative histogram of the nonzero values in S . t was in $[0,1]$. The default value of t was 0.7, which produced results mostly coherent with perceptual salient edges. Edges detected by the Canny detector and the salient edges of an image are shown in Fig. 6.10. The weak edges in the sky were detected by the salient edge detector but not by the Canny edge detector. The edge detection results in other parts of the image were mostly similar using both methods.

Gradient mapping on the edges was performed separately in the depth planes. The contrast on the boundary across depth planes influenced the intra-contrast mapping. After detecting the salient edges, the gradient mapping on the edges of each plane was performed by edge-length constrained histogram matching. The constraint on the edge-length served to ensure that the gradient of an edge is mapped to that of a corresponding edge of similar length in the reference painting. The length constraint was conducted by clustering the edges based on the length to three clusters C_p , $p = 1, 2, 3$, which represented the long, medium and short edges. Gaussian mixture model was used for the clustering, and histogram matching was used to map the gradient distribution of the target to that of the reference in each cluster. The process of histogram matching was expressed as

$$T_{C_p}(r_k) = \sum_{j=0}^k \frac{n_j^{C_p}}{n_r^{C_p}}, k = 0, 1, 2, \dots, N - 1 \quad (6.14)$$

$$F_{C_p}(s_k) = \sum_{i=0}^k \frac{n_i^{C_p}}{n_s^{C_p}}, k = 0, 1, 2, \dots, N - 1 \quad (6.15)$$

$$\tilde{s}_k^{C_p} = \min_r \|T_{C_p}(r) - F_{C_p}(s_k)\| \quad (6.16)$$

where $n_j^{C_p}$ was the number of pixels with value r_j in cluster C_p of reference plane, $n_i^{C_p}$ was the number of pixels with value s_i in cluster C_p of photograph plane, $n_r^{C_p}$ was the total number of pixels of cluster C_p in reference plane, $n_s^{C_p}$ was the total number of pixels of cluster C_p in photograph plane, and N was the number of discrete levels. Applying this matching method, gradients at edges of the four planes in F' , which were S_f, S_m, S_b, S_s , were mapped to be $\tilde{S}_f, \tilde{S}_m, \tilde{S}_b, \tilde{S}_s$. If there was no salient edge in a reference plane, the gradients of the corresponding plane in F' was unchanged.

Though gradients of each plane were mapped separately to those of the corresponding plane in the reference, a scale matrix η was calculated. $\eta_f, \eta_m, \eta_b, \eta_s$ were initialized as matrices of zeros which had the same size as the image dimensions of the target photograph.

$$\begin{aligned} \eta_f(P_f) &= \tilde{S}_f/S_f, & \eta_m(P_m) &= \tilde{S}_m/S_m \\ \eta_b(P_b) &= \tilde{S}_b/S_b, & \eta_s(P_s) &= \tilde{S}_s/S_s \end{aligned} \quad (6.17)$$

$$\eta = \max(\eta_f, \eta_m, \eta_b, \eta_s) - W_s + \mathbf{1} \quad (6.18)$$

where P_f, P_m, P_b, P_s were the indexes of the salient edges of the four depth planes. W_s represented the mask of salient edges of the input photograph.

Finally, the gradients after intra-contrast mapping were given by

$$(G'_{sx}, G'_{sy}) = \eta \cdot (G_{sx}, G_{sy}) \quad (6.19)$$

where (G_{sx}, G_{sy}) were the gradients of the inter-contrast mapped result F' . The target gradients in the second term of Eq. (6.1) were $g_x = G'_{sx}$ and $g_y = G'_{sy}$.

6.4.2.3 Optimisation Solver

The minimisation of the energy cost function E in Eq. (6.1) was achieved by solving the linear function $\mathbf{A}f = b$, where

$$\mathbf{A}(i, j) = \begin{cases} -w_g(i) - \frac{w_s(i)}{(|F(i) - F(j)|^\theta + \varepsilon)} & j \in N_4(i) \\ w_d(i) - \sum_{k \in N_4(i)} \mathbf{A}(i, k) & i = j \\ 0 & \text{otherwise} \end{cases} \quad (6.20)$$

$$b(i) = w_d(i)d(i) - w_g(i) \sum_{k \in N_4(i)} (F(k) - F(i)) \quad (6.21)$$

$N_4(i)$ were the four neighbours of pixel i . In the implementation, constant values were used for weights $w_d = 0.01$, $w_g = 1$. In order to penalise large changes to the original gradient, the weight w_s was defined as

$$w_s(i) = \lambda \left(1 - e^{-\frac{(\eta(i)-1)^2}{\sigma}} \right) + \varepsilon \quad (6.22)$$

Therefore, a larger weight was used for pixels having a large gradient change and a small weight was used for pixels which were not on salient edges. The parameters

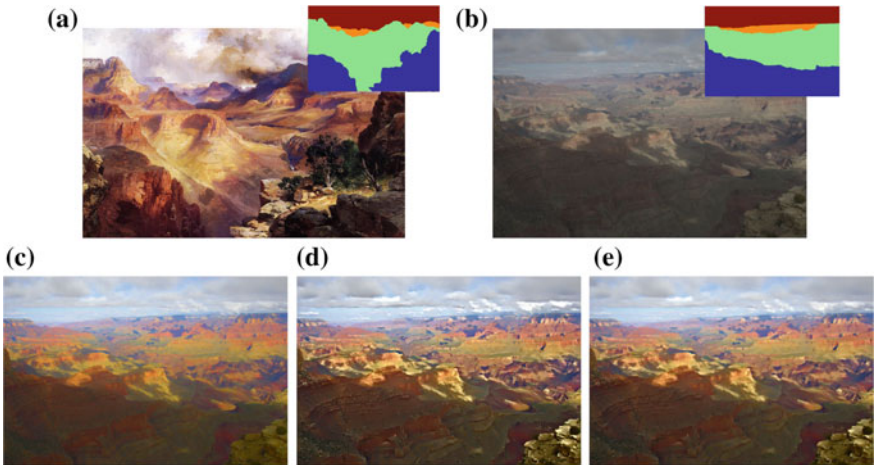


Fig. 6.11 **a** Reference with mask (CC Image courtesy The Athenaeum), **b** Input with mask. **c** Inter-contrast mapping result. **d** Result using E_d and E_g in Eq. (6.1). **e** Result with E_s

were $\theta = 0.5$, $\lambda = 0.06$, $\varepsilon = 0.0001$ and $\sigma = 3$. The example in Fig. 6.11 shows the effect of each part in Eq. (6.1). Figure 6.11c has the target inter-contrast values of lightness and saturation only, while the result using E_d and E_g in Fig. 6.11d shows the enhancement on the local contrast. The comparison of the result in Fig. 6.11d, e illustrates clearly the influence of E_g . The large gradient changes on the salient edges which may create sharp edge artefacts were avoided in the result using E_g .

6.5 Experiments and Discussion

This section analyses the performance of the proposed depth-aware contrast manipulation algorithm. The algorithm was implemented in MATLAB on a PC with an Intel 2.93 GHz processor and 3 GB RAM. For a photograph–reference image pair of a pixel dimension of 600×800 , it took around 35 s to obtain the results. Solving the optimisation function was the most time-consuming step of the proposed algorithm. It is possible that by optimising the code and using acceleration schemes, the computation time can be shortened. The depth planes of paintings in the database were drawn manually.

Influence of the parameter q . The parameter q is expressed as q_L for lightness and q_C for saturation. Figure 6.12 shows the influence of the parameter q on the appearance of the result. When $q = 0$, the mean value of the image is preserved to match that of the original. There is only a relative adjustment on the mean values of the four depth planes, to match the inter-contrast of the reference (see the mean values of the four depth planes in Fig. 6.13). When q increases towards 1, the mean values of the four depth planes of the target are changed towards those of the reference. As a result of the constraint on the lightness in the reference selection, the selected references have similar lightness values to that of the input photograph. However, the saturation of the reference may be much higher than the input. The reference in the example in Fig. 6.12 has distinctly greater saturation than the original photograph as the curves shown in Fig. 6.13. According to studies of visual appeal, saturated picture are more pleasing [32]. However, ‘How much saturation is sufficient?’ is subject dependent. The user can control q_C to adjust the value of saturation within the output image. In the current implementation, the default parameter is $q_L = 1$, and the parameter q_C is recommended by matching the colourfulness of the photograph to that of the reference. The definition of colourfulness combines both the colour variance and the chroma magnitude [14]:

$$C_f = \sigma_{ab} + 0.94\mu_C \quad (6.23)$$

where $\sigma_{ab} = \sqrt{\sigma_a^2 + \sigma_b^2}$ is the trigonometric length of the standard deviation in CIELAB space. μ_C is the mean of C in LCH space. Given the colourfulness values of reference C_{fr} and input C_{fs} , the target mean saturation value M_{Ct} is recommended

by Algorithm 1. $q_C = q_{rc} = \frac{M_{Ct} - M_{Cs}}{M_{Cr} - M_{Cs}}$, M_{Cs} , M_{Cr} are the recommended mean values of the C in the input photograph and reference, respectively.

Algorithm 1 Recommended M_{Ct}

1. Initially $\alpha = 1$, the target colorfulness value $C_{ft} = C_{fs}$
 2. $\alpha = \alpha \cdot \frac{C_{fr}}{C_{ft}}$
 3. The target mean $M_{Ct} = \alpha M_{Cs}$
 4. Update C of the photograph by Eqs. (6.6) and (6.7), get C'
 5. Update a, b in CIELAB space, $a' = a \cdot \frac{C'}{C}$, $b' = b \cdot \frac{C'}{C}$
 6. Update C_{ft} by Eq. (6.23), $C_{ft} = \sigma_{a'b'} + 0.94\mu_{C'}$
- if** $|C_{ft} - C_{fr}| > \epsilon$ **then**
 Return to step 2
else
 Output $M_{Ct} = \alpha M_{Cs}$
end if
-

Performance. The contrast manipulation results of the photograph in Fig. 6.12a using three recommended paintings are shown in Fig. 6.14. The organisation of the inter-contrast and intra-contrast of lightness and saturation of the three reference paintings is consistent with the observation in the contrast analysis in Sect. 3.3.3, except that the intra-contrast values of saturation in FG, MG and BG of the second painting are very similar (no decreasing tendency). Different references produce different results. Applying the proposed contrast manipulation method, the inter-contrast and intra-contrast of the enhanced photographs are successfully aligned to those of the reference, as the graphs in the third to sixth rows show. The visual organisation of the enhanced photograph is similar to that of the corresponding reference painting with respect to depth planes.

In order to measure the effectiveness of the proposed depth-aware contrast mapping algorithm, the matching error of inter-contrast and intra-contrast of 70 test results was calculated. The matching error was defined as the absolute difference between the targeted contrast value from the reference painting and the actual output contrast value in the enhanced photograph. The average matching errors of the test results are shown in Fig. 6.15. The average matching errors of inter-contrast and intra-contrast are less than 0.03. This shows that images enhanced by the proposed method exhibit significant similarity with their corresponding reference paintings. The FG and sky are more likely to have a larger intra-contrast matching error. This is due to the gradient constraint E_s in the energy function Eq. (6.1), in order to avoid large changes on the gradient.

Comparison. The proposed example-based depth-aware contrast manipulation method was compared with three existing example-based contrast adjustment methods shown in Fig. 6.16. Histogram matching (HM) is the traditional method used in image processing to adjust the value of an image based on a reference. Here, histogram matching was performed on L and C channels in LCH space. Bae's two-scale tone mapping method [2] was applied on the L channel in CIELAB space. a and b channels were scaled by the enhancement scale of L to adjust the saturation.

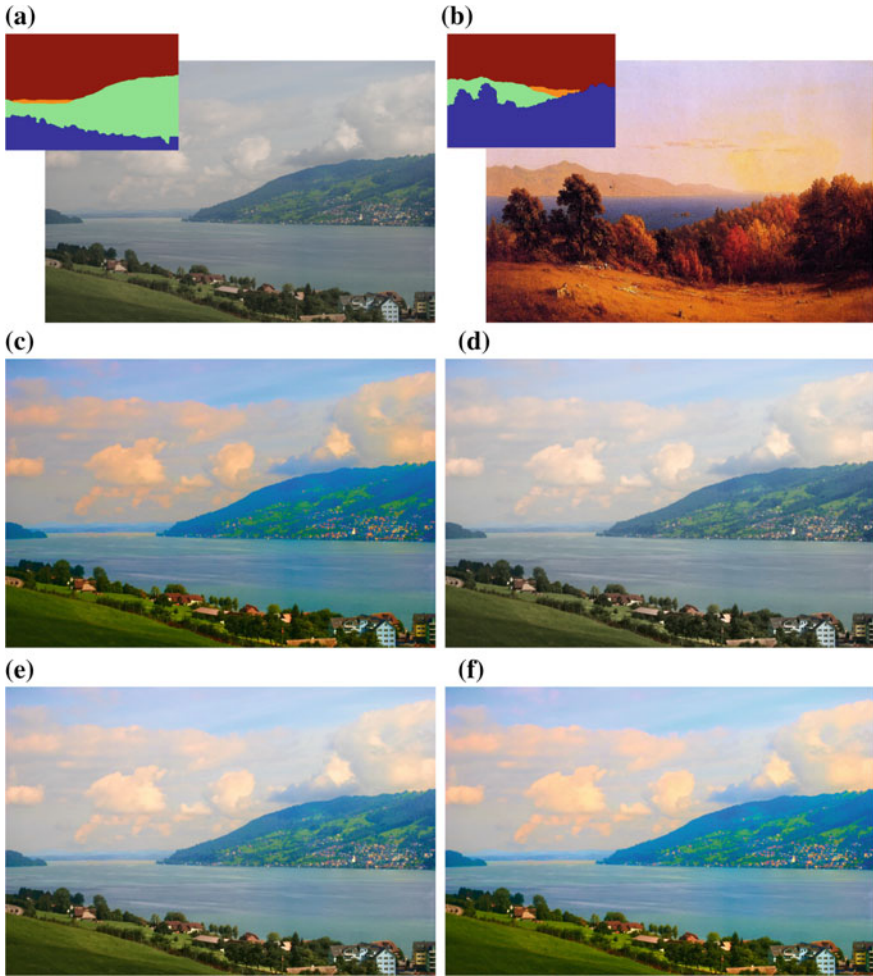
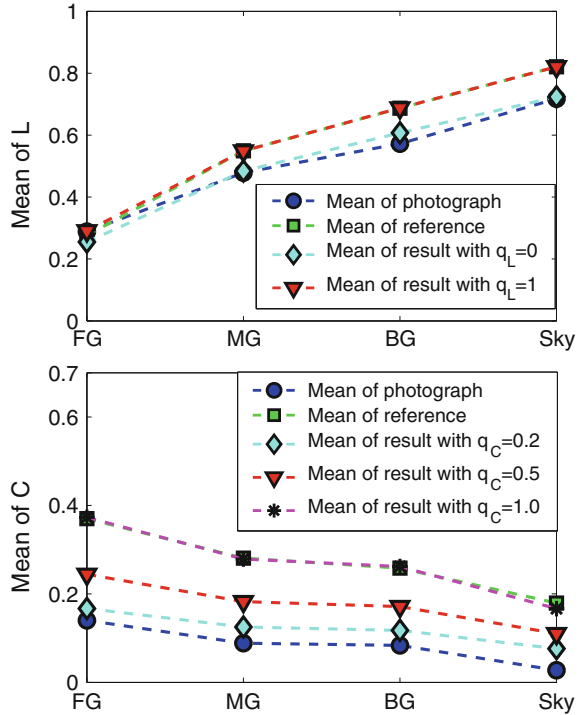


Fig. 6.12 The influence of the parameter q . **a** Input with mask. **b** Reference with mask (CC Image courtesy The Athenaeum). **c** $q_L = 0, q_C = 1.0$. **d** $q_L = 1.0, q_C = 0.2$. **e** $q_L = 1.0, q_C = 0.5$. **f** $q_L = 1.0, q_C = 1.0$

Another method was the example-based contrast enhancement by gradient mapping (CEGM) [15]. Three automatic contrast enhancement methods, the ‘Lo-fi’ filter of Instagram[®], ‘Auto Contrast’ in Adobe[®] Photoshop[®] and local gradient optimisation in [21], were also compared with the proposed contrast manipulation method in Fig. 6.16.

HM and Bae’s global contrast adjustment methods only match the global value distribution and have no spatial consideration. CEGM, ‘Auto Contrast’ in Adobe[®] Photoshop[®] and local gradient optimisation in [21] only enhance the contrast in local areas without regional consideration of depth. The ‘Lo-fi’ filter in Instagram[®]

Fig. 6.13 The mean values of the four depth planes when using different parameters q_L and q_C in Fig. 6.12



enhances both the local contrast and 2D centre/corner difference. Differently, the proposed method enhances the local contrast and adjusts the contrast between depth planes. In the results achieved using the proposed method, the visual appeal, the illusion of depth and the observance of the law of atmospheric perspective effect are all enhanced. The results in (j) and (t) are the Chroma increased version of (i) and (s), respectively, to match the colourfulness of the reference painting. The visual apparent colour change in the results in (j) and (t), as compared to the original input photographs (a) and (k), is due to the adjustment on the Chroma. In this process, the numeric value of the hue is in fact not changed.

For depth-aware contrast mapping, histogram matching of each depth plane in the spatial domain was also performed. A comparison with the histogram matching and with the depth-aware contrast manipulation method proposed in [33] is given in Fig. 6.17. The results using histogram matching (see Fig. 6.17c) are not so pleasing as a consequence of having more artefacts. Local contrast adjustment achieved using the proposed method is clearly more effective than that achieved using the method described in [33]. The proposed method does not produce artefacts even in a low-quality image with compression blocks.

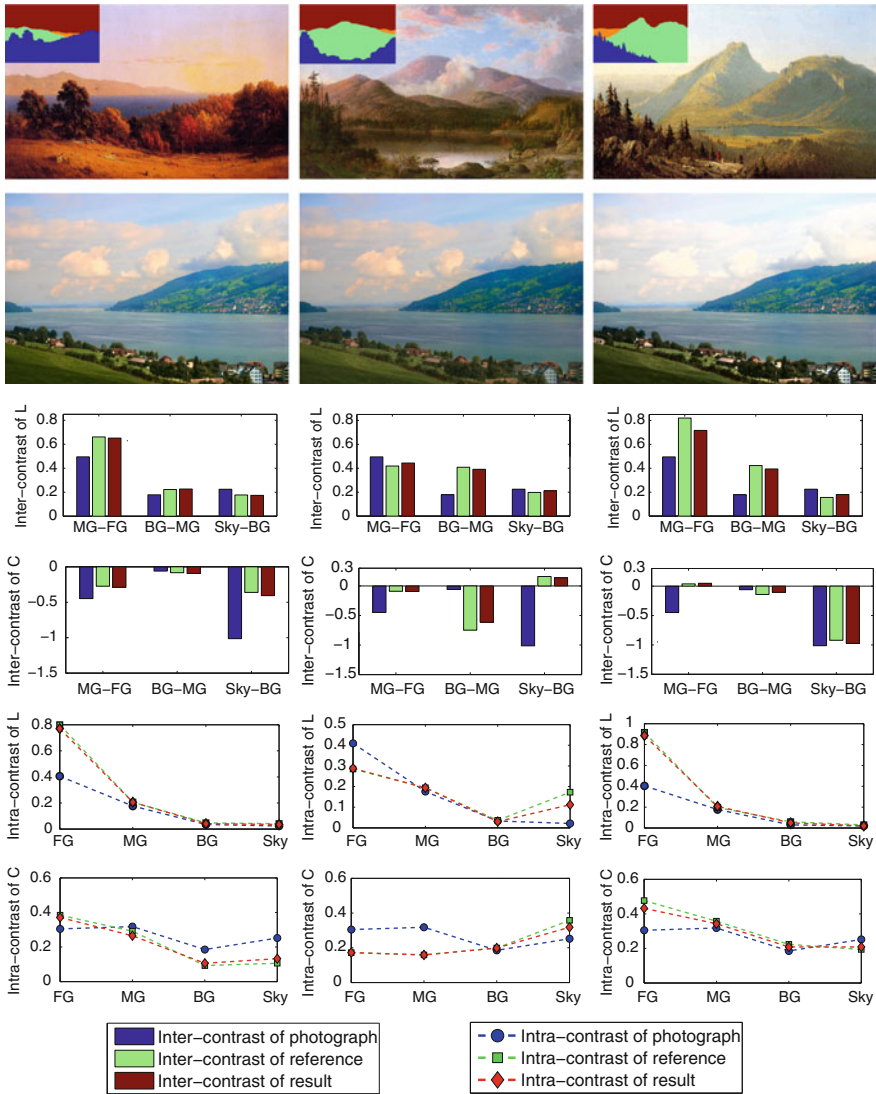


Fig. 6.14 The contrast manipulation results based on three selected references in Fig. 6.6. The input with mask is in Fig. 6.12a. Images in the first row are the references with masks (Second painting: Courtesy Wikimedia Commons. Third painting: CC Image courtesy of Sanford Robinson Gifford org). Images in the second row are the corresponding results. The bars in the third and fourth rows are to compare the inter-contrast of L and C among input, references and results for the three cases. The curves in the fifth and sixth rows are to compare the intra-contrast of L and C among input, references and results. Parameters: $q_L = 1$, $q_C = 0.60, 0.70, 0.68$ corresponding to the three references from *left to right*

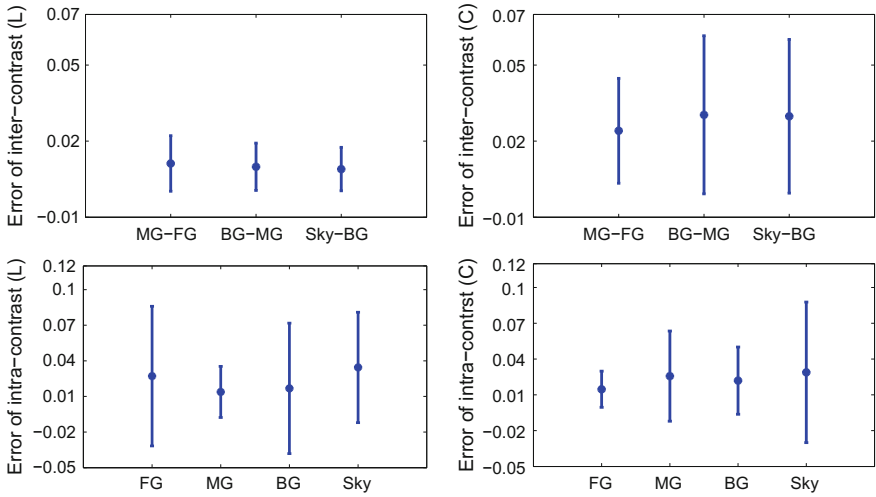


Fig. 6.15 The average matching error of inter-contrast and intra-contrast. The bars represent \pm one standard deviation. The inter-contrast is normalised to the range $[-1, 1]$, and the intra-contrast is normalised to the range $[0, 1]$

In order to show the advantage of the proposed atmospheric perspective effect enhancement method, it was compared with the atmospheric physical model used in [17]. Depth was estimated based on the outlined depth planes. The physical model was used only on the ground planes and not on the sky. Figure 6.18b shows the result that is produced by using the physical model on the photograph shown in Fig. 6.2a. The result in Fig. 6.18b shows that both close objects and distant objects will be affected when the atmospheric effect is increased. However, even if a large attenuation strength and a heavy atmospheric condition is used, the separation of the FG, MG and BG is still not as clear as in the result produced by the proposed method. The reason is that the physical model only reduces the contrast of the scene based on distance and does not adjust the relative values of depth planes to enhance the illusion of depth. Differently, the proposed method adjusts the relative difference of depth planes according to the reference painting in order to strengthen their separation, thereby enhancing the illusion of depth. In the results achieved by the proposed method, the contrast of the FG is generally modified to be larger than it was originally and is substantially different to the other depth planes.

Application to other photograph categories. Aside from landscapes with four depth planes, the proposed atmospheric perspective effect enhancement method can also be used for landscapes with two depth planes (FG and BG) or three depth planes (FG, MG, sky or FG, BG, sky). Two examples are illustrated in Fig. 6.19.

User study. A user study was carried out to evaluate the effectiveness of the proposed method. The user study was designed to test the illusion of depth and visual appeal in the results. Thirty participants (10 females and 20 males) with ages ranging from 20 to 35 were invited. All of them had normal vision. Twenty-five of



Fig. 6.16 Comparison with other contrast adjustment methods. **a** and **k** Input w/mask, **b** and **l** Reference w/mask (CC Image courtesy The Athenaeum), **c** and **m** Global HM, **d** and **n** Bae's method, **e** and **o** CEGM, **f** and **p** "Lo-fi" filter in Instagram[®], **g** and **q** "Auto Contrast" in Adobe[®] Photoshop[®], **h** and **r** local gradient optimisation in [21], **i** and **s** proposed ($q_C = 0$), **j** and **t** Proposed ($q_C = q_{rc}$)

the test results were randomly selected for this user study, including a variety of scenes. Each test result and its corresponding original photograph were shown as a pair side by side. The left/right ordering of the images in each pair was randomly generated. For each image pair, the participants were asked to choose their answers responding to two questions: 'Which has the clearer illusion of depth?' and 'Which is more pleasing?'. The participants can choose 'left' or 'right' as responses to each question according to their first impression, and they were not allowed to go back to previously viewed images. The study began with an introductory page that showed two example images to introduce the illusion of depth produced by atmospheric

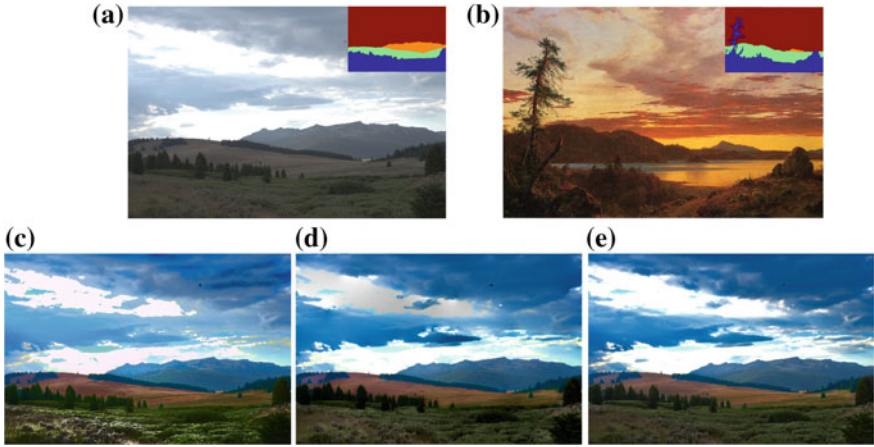


Fig. 6.17 Comparison with depth-aware methods. **a** Input with mask. **b** Reference with mask (CC Image courtesy The Athenaeum). **c** HM in depth plane. **d** Method in [33]. **e** Proposed



Fig. 6.18 Comparison with the physical model used in [17]. The input photograph and reference are those used in Fig. 6.2. **a** Estimated depth. **b** Result by [17]. **c** Result by the proposed method with the parameter $q_C = 0$. The mean saturation is preserved to the original one to make the result comparable with that in **b**

perspective in 2D images. No further information on the goals of the experiment or the origin of the images was provided to the participants.

The result of the user study is shown in Table 6.1. Higher percentages of responses were in favour of the enhanced results. By using a one-sample, one-tailed t-test on these responses, the chance that results achieved by the proposed method were perceived as having a clearer illusion of depth was significantly larger than 50% ($p \ll 0.0001$), and the chances that results by the proposed method were perceived as being more pleasing were also significantly larger than 50% ($p < 0.0001$). This shows that the proposed method has significantly enhanced the illusion of depth and visual appeal of the photographs.

All the 25 test results were uniformly selected by the participants as having a clearer illusion of depth (selection rate larger than 50%). However, 5 results were selected as not so pleasing as the original photographs (selection rate smaller than 50%) even though they had a clearer illusion of depth. According to the feedback

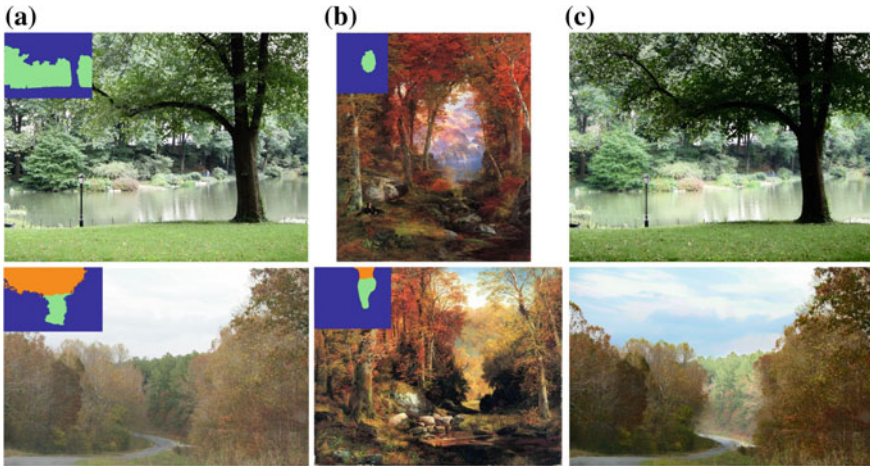


Fig. 6.19 Application on photographs with two or three depth planes. The *top row* is the example with two depth planes. The *second row* is the example with three depth planes. **a** Input with mask. **b** Reference with mask (CC Image courtesy The Athenaeum). **c** Result

Table 6.1 User study result

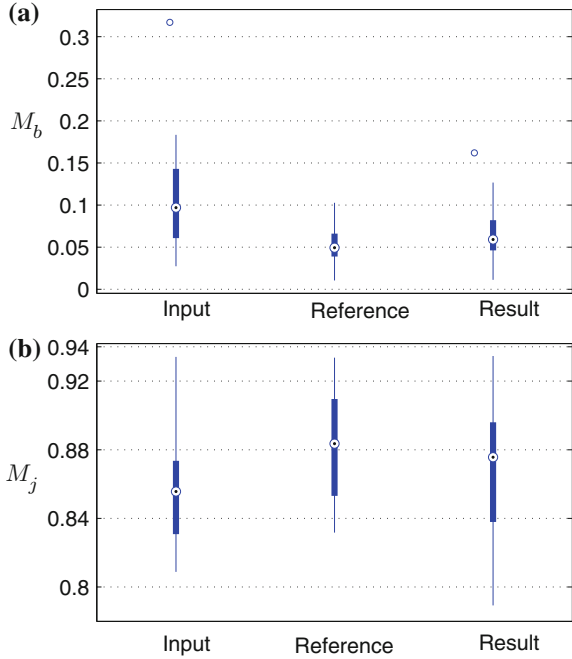
Questions	Selection rate	
	Enhanced result	Original
Which has the clearer illusion of depth?	82.4%	17.6%
Which is more pleasing?	69.9%	30.1%

of the participants, this was due to their saturation values being over-enhanced. The user study also shows that the visual appeal of images is a more subjective preference compared to the perception of depth.

Aesthetics measure. Here, we want to show how the proposed method influenced the aesthetics of the photographs. For the measurement of aesthetics, a set of ratios have been used to quantify the aesthetic experience based on information theory and Kolmogorov complexity [26]. They measure the selection of palette, the order in the colour distribution and the transition from the palette to object. Here, Shannon’s perspective is used to measure the influence of the proposed contrast manipulation method on the aesthetics. Kolmogorov’s perspective and Zurek’s perspective are not measured because the proposed contrast manipulation method has little influence on the order of colour distribution. Given the normalised histogram of luminance Y_{709} , $P(x)$ (x in $[0, 255]$), the global aesthetics measure using Shannon’s perspective is

$$M_b = \frac{H_{\max} - H_p}{H_{\max}} \tag{6.24}$$

Fig. 6.20 Aesthetics assessment. **a** M_b . **b** M_j



$$H_p = - \sum_x P(x) \log P(x) \tag{6.25}$$

The maximum entropy for luminance is $H_{\max} = 8$. The palette similarity of the four depth planes is measured as

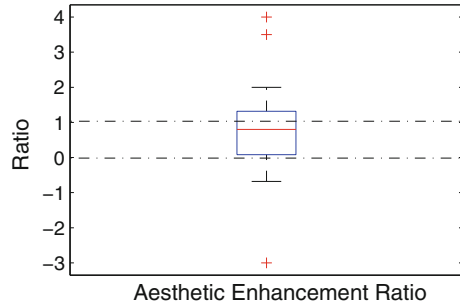
$$M_j = \frac{\sum_{i=1}^4 \pi_i H(p_i)}{H_p} \tag{6.26}$$

where π_i is the area of region i and $H(p_i)$ is the entropy of region i calculated by Eq. (6.25).

The M_b and M_j values of the 25 test photographs, 25 corresponding reference paintings and the results are calculated to compare their differences. The box plots of the values are shown in Fig. 6.20. The larger M_b values of the input photographs, as compared to those of the references, reflect the higher colour homogeneity in the photographs. The larger M_j values of the references, as compared to those of the photographs, reflect the higher similarity between the palettes of the four depth planes in the paintings. The distributions of M_b and M_j values of the results are very similar to those of the references. In other words, the aesthetic value of the photographs has been enhanced to be closer to that of the reference paintings.

An aesthetic quality inference engine (ACQUINE) has been developed to automatically rate natural photographs [10]. Aesthetic scores for the 25 photographs, 25

Fig. 6.21 Ratio of aesthetic score



corresponding reference paintings and the results are collected by uploading them to acquine.alipr.com. ACQUINE serves to rate the aesthetic value of natural professional photographs, not for paintings. Therefore, the given aesthetic scores for images may not really reflect their aesthetics (some paintings get lower aesthetic scores than the input photographs). Hence, to remedy this, a relative measure reflecting the change of aesthetic scores is employed. Given the aesthetic scores of input (S_I), reference (S_R) and result (S_O), a ratio is used to measure the relative change of aesthetic value, which is defined as:

$$r_o = \frac{S_O - S_I}{S_R - S_I} \quad (6.27)$$

When $r_o > 0$, the measured aesthetic value of the result is converging to that of the reference, and when $r_o = 1$, the measured aesthetics of the result is the same as that of the reference. The distribution of r_o for the 25 test results in Fig. 6.21 again shows that our approach changed the aesthetic value of the photographs to move it closer to that of the reference paintings.

User study for depth plane partition. Another user study was conducted in order to see to what degree people have consensus in partitioning a given image. There were 20 participants in the user study. Each participant was required to perform depth-based partition on 24 landscape photographs selected from the database. The participants were required to partition each photograph into FG, MG, BG and sky by drawing three boundary lines using a purposely developed user interface. Before starting partitioning, several completed partitions made by an experienced artist were given as examples to show the meaning of FG, MG and BG. The selection of photographs was based on the assumptions that the images may be reasonably suitable for partitioned into FG, MG, BG and sky and the images did not contain complicated details such that drawing the boundary of each partition would be relatively easy.

Intersection-over-union ratio (known as Jaccard similarity coefficient) is one measurement for the overlap of the segmented region with the ground truth in semantic segmentation and object detection [1, 8, 12]. It is used here to measure the overlap of two people's partitioning for the same image. The intersection-over-union ratio of region A and B is defined as

Table 6.2 Intersection-over-union ratio of human drawing study for each photograph

Photographs	Average/standard deviation		
	FG	MG	BG
1	0.77/0.17	0.60/0.29	0.73/0.29
2	0.68/0.24	0.34/0.37	0.56/0.32
3	0.69/0.26	0.54/0.34	0.41/0.28
4	0.88/0.12	0.57/0.40	0.56/0.31
5	0.79/0.29	0.49/0.32	0.53/0.43
6	0.73/0.15	0.20/0.28	0.30/0.32
7	0.81/0.19	0.59/0.38	0.41/0.26
8	0.83/0.27	0.75/0.34	0.69/0.32
9	0.90/0.18	0.96/0.06	0.97/0.02
10	0.55/0.32	0.30/0.36	0.38/0.34
11	0.65/0.30	0.59/0.35	0.65/0.32
12	0.80/0.11	0.30/0.29	0.24/0.29
13	0.55/0.39	0.37/0.32	0.43/0.30
14	0.77/0.24	0.60/0.38	0.54/0.38
15	0.64/0.29	0.68/0.36	0.73/0.31
16	0.84/0.14	0.38/0.33	0.56/0.18
17	0.90/0.11	0.57/0.32	0.75/0.33
18	0.76/0.23	0.59/0.35	0.55/0.34
19	0.64/0.27	0.52/0.34	0.16/0.24
20	0.74/0.20	0.65/0.34	0.54/0.31
21	0.74/0.21	0.45/0.32	0.86/0.16
22	0.62/0.26	0.36/0.31	0.33/0.35
23	0.99/0.01	0.80/0.19	0.62/0.33
24	0.77/0.28	0.47/0.39	0.38/0.28
Average	0.75/0.22	0.53/0.32	0.54/0.29

$$R = S(A \cap B)/S(A \cup B) \quad (6.28)$$

where $S(A)$ refers to the area of region A. The boundary of partitioning is not very reliable, so we would put less weight on areas around partition boundary and more on areas far away from the partition boundary. Because each photograph is partitioned by 20 participants, and each time two people's partitioning is selected to calculate the intersection-over-union ratio, there would be $(20 \times 19)/2 = 190$ results for each image. Hence, the average of these 190 results is taken as a measurement of consistency of partitioning for each image. The standard deviation is also recorded. The average and standard deviation of the intersection-over-union ratios of FG, MG and BG are shown in Table 6.2. Sky is very easy to identify, and the average ratio of sky is almost 1, so it is not shown here. The table shows that: 1. the average

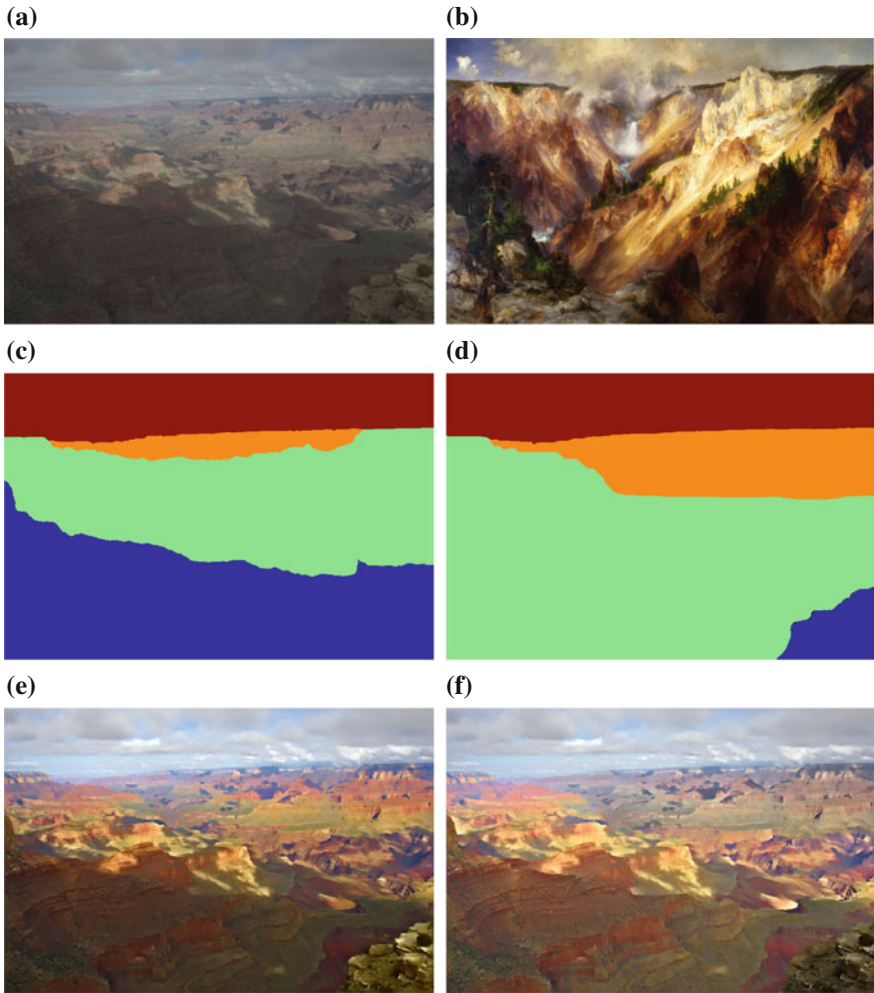


Fig. 6.22 Results using different partitions. **a** Input photograph. **b** Reference (CC Image courtesy Wikimedia Commons). **c** Mask 1. **d** Mask 2. **e** Result using mask 1. **f** Result using mask 2

intersection-over-union ratios of FG are generally larger than those of MG and BG; 2. the standard deviations for MG and BG statistics are quite large. For the 13th photograph in the Table 6.2, the partitions by the 20 participants have a very low intersection-over-union ratio. However, the 20 partitions generally can be classified into two types as the two masks shown in Fig. 6.22b, c. Among the 20 participants, 6 of them outline FG, MG and BG for the 13th photograph in a similar way as the mask shown in Fig. 6.22c. The results using the two different partitions are different (see Fig. 6.22e, f). This user study shows that the partitions from different people are not uniform. The partition result is mostly dependent on personal understanding of

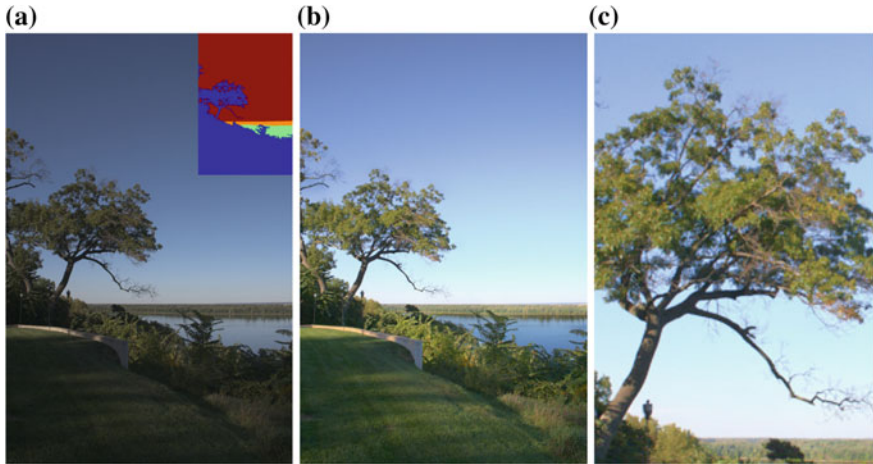


Fig. 6.23 An example demonstrating the limitation of the proposed method. **b** Input photograph. **b** Result. **c** Close-up Region of **b**

the scene and not a standard rule. This is an unsolved problem in the research area of image segmentation.

Limitations. While experiments have demonstrated the effectiveness of the proposed method, we observe a few failure cases. The sky is fragmented by tree branches as the example in Fig. 6.23a shows some of the small fragments are not detected as sky and some fine tree branches are also not segmented out of the sky plane. As a result, the colour of some small sky fragments differs from the rest of the sky and the colour of the mis-segmented fine tree branches differs from other parts of the tree (Fig. 6.23c).

In addition, for a captured photograph with only a simple sky that has almost no texture (e.g. over-exposed sky), the proposed method cannot adjust the local contrast of the sky to match the local contrast of a contrast-rich sky in a painting.

More experimental results by the proposed method are shown in Fig. 6.24.

6.6 Concluding Remarks

This chapter proposed an example-based depth-aware contrast manipulation method to transfer the atmospheric perspective effect from landscape paintings to landscape photographs. More specifically, the following contributions were reported in this chapter. First, a statistical study was conducted to explore the difference of contrast organisation governed by atmospheric perspective between traditional paintings and ‘snapshot’ style photographs. The conclusion of the study was that the contrast governed by atmospheric perspective in paintings was more purposefully organised. This provided the justification for the proposed depth-aware contrast manipulation of



Fig. 6.24 More results. The first and third rows: the original photographs. The second and fourth rows: the results using proposed method

photographs. Second, a novel linear-circular mapping curve was proposed to adjust the mean values to map the inter-contrast. This curve preserved the perceptual contrast, and avoided creating over-enhanced artefacts in low value zones. Third, an effective gradient mapping algorithm was developed to map the intra-contrast. Finally, an optimisation problem was formulated that simultaneously considers the desired inter-contrast, intra-contrast and specified gradient constraints. Experiments demonstrated that the enhanced target images exhibited significant similarity to the reference paintings in their depth-aware contrast organisation. Additionally, a user study was conducted to assess the effectiveness of the proposed method. This study demonstrated that both the visual appeal and illusion of depth governed by atmospheric perspective effect in the photographs were enhanced using the proposed method.

In this chapter, four depth planes were used to introduce the proposed depth-aware contrast manipulation. However, the method can easily be extended to other numbers of depth planes. The proposed method only requires several discrete depth planes and does not require the real depth, which is a challenge to capture for long-range landscape photographs.

References

1. Arbelaez, P., Hariharan, B., Gu, C., Gupta, S., Bourdev, L., Malik, J.: Semantic segmentation using regions and parts. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3378–3385 (2012)
2. Bae, S., Paris, S., Durand, F.: Two-scale tone management for photographic look. *ACM Trans. Graph.* **25**(3), 637–645 (2006)
3. Bailey, R.: Perception-guided image manipulation. Ph.D. Dissertation, Department of Computer Science and Engineering, Washington University in St. Louis (2007)
4. Beaudot, W., Mullen, K.: How long range is contour integration in human color vision? *Vis. Neurosci.* **15**, 51–64 (2003)
5. Bhat, P., Zitnick, L., Cohen, M., Curless, B.: Gradientshop: a gradient-domain optimization framework for image and video filtering. *ACM Trans. Graph.* **29**(2), 10:1–10:15 (2010)
6. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. Proceedings of the IEEE International Conference on Computer Vision **1**, 105–112 (2001)
7. Chen, S., Beghdadi, A.: Natural enhancement of color image. *EURASIP J. Image Video Process.* **2010**, 2:1–2:30 (2010)
8. Csurka, G., Larlus, D., Perronnin, F.: What is a good evaluation measure for semantic segmentation? In: Proceedings of the British Machine Vision Conference, pp. 1–11 (2013)
9. Dale, K., Johnson, M.K., Sunkavalli, K., Matusik, W., Pfister, H.: Image restoration using online photo collections. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2217–2224 (2009)
10. Datta, R., Wang, J.: Acquire: aesthetic quality inference engine - real-time automatic rating of photo aesthetics. In: Proceedings of the ACM Multimedia Information Retrieval, pp. 421–424 (2010)
11. Dunning, W.V.: *Changing Images of Pictorial Space: A History of Spatial Illusion in Painting*. Syracuse University Press, Syracuse (1991)
12. Everingham, M., Gool, L.V., Williams, C.K.I., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. *Int. J. Comput. Vis.* **88**, 303–338 (2010)
13. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. *ACM Trans. Graph.* **21**(3), 249–256 (2002)
14. Hasler, D., Susstrunk, S.: Measuring colourfulness in natural images. In: *IS&T/SPIE Electronic Imaging 2003: Human Vision and Electronic Imaging VIII*, vol. 5007, pp. 87–95 (2003)
15. Huang, H., Xiao, X.Z.: Example-based contrast enhancement by gradient mapping. *Visual Comput.* **26**, 731–738 (2010)
16. Johnson, M.K., Dale, K., Avidan, S., Pfister, H., Freeman, W.T., Matusik, W.: Cg2real: improving the realism of computer generated images using a large collection of photographs. *IEEE Trans. Visual Comput. Graph.* **17**, 1273–1285 (2011)
17. Jung, J.I., Lee, J.H., Shin, I.Y., Moon, J.H., Ho, Y.S.: Improved depth perception of single-view images. *ECTI Trans. Electr. Eng. Electron. Commun.* **8**(2), 164–172 (2010)
18. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2169–2178 (2006)
19. Lischinski, D., Farbman, Z., Uyttendaele, M., Szeliski, R.: Interactive local adjustment of tonal values. *ACM Trans. Graph.* **25**, 646–653 (2006)
20. Luft, T., Colditz, C., Deussen, O.: Image enhancement by unsharp masking the depth buffer. *ACM Trans. Graph.* **25**, 1206–1213 (2006)
21. Majumder, A., Irani, S.: Perception-based contrast enhancement of images. *ACM Trans. Appl. Percept.* **4**(3), 1–22 (2007)
22. Mattingly, D.B.: *The Digital Matte Painting Handbook*. Wiley Publishing Inc, New York (2011)
23. Narasimhan, S.G.: *Models and Algorithms for Vision through the Atmosphere*. Columbia University, Ph.D. thesis (2003)

24. Narasimhan, S.G., Nayar, S.K.: Vision and the atmosphere. *Int. J. Comput. Vision* **48**, 233–254 (2002)
25. O’Shea, R.P., Blackburn, S.G., Ono, H.: Contrast as a depth cue. *Vision. Res.* **34**(12), 1595–1604 (1994)
26. Rigau, J., Feixas, M., Sbert, M.: Informational aesthetics measures. *IEEE Comput. Graph. Appl.* **28**(2), 24–34 (2008)
27. Saxena, A., Sun, M., Ng, A.Y.: Make3d: learning 3d scene structure from a single still image. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(5), 824–840 (2009)
28. Sen, D., Pal, S.K.: Automatic exact histogram specification for contrast enhancement and visual system based quantitative evaluation. *IEEE Trans. Image Process.* **20**(5), 1211–1220 (2011)
29. Sheppard, R.: *Landscape Photography: From Snapshots to Great Shots*. Peachpit Press, San Francisco (2012)
30. Sievers, A.H.: *Master Drawings from Smith College Museum of Art*. Hudson Hills Press, London (2000)
31. Wang, B., Yu, Y., Xu, Y.Q.: Example-based image color and tone style enhancement. *ACM Trans. Graph.* **30**(4), 64:1–64:12 (2011)
32. Yendrikhovskij, S.N., Blommaert, F.J.J., de Ridder, H.: Perceptually optimal color reproduction. In: *Proceedings of the 6th Color Imaging Conference: Color Science, Systems, and Applications*, vol. 3299, pp. 274–281 (1998)
33. Zhang, X., Constable, M., Chan, K.L.: Aesthetic enhancement of landscape photographs as informed by paintings across depth layers. In: *Proceedings of the IEEE International Conference on Image Processing*, pp. 1137–1140 (2011)

Chapter 7

Regional Contrast Manipulation for Portrait Photograph Enhancement

Abstract This chapter proposes a method to manipulate the regional contrast in snapshot style portrait photographs by using pre-modern portrait paintings as aesthetic examples: to improve the visual appeal and focus of attention of the photographs. The example portrait painting is selected based on a comparison of the existing contrast properties of the painting and those of the photograph. The contrast organisation in the selected example painting is transferred to the photograph by mapping the inter- and intra-regional contrast values of the regions, such as the face and skin areas of the foreground figure, the non-face/skin part of the foreground and the background region. A novel piecewise nonlinear transformation curve is used to achieve this contrast mapping. Finally, the transition boundary between regions is smoothed to achieve the final results. Experimental results demonstrate that, by using this proposed method, the visual appeal of portrait photographs is effectively improved and the face and the figure become more salient.

Keywords Portraits · Painting · Photographs

7.1 Introduction

Most existing approaches in image editing improve the appearance of images globally or locally in the same way for all image categories, regardless whether it is a portrait or a natural scene. Similar to the photographer's strategy of using different modes or lenses to capture landscapes or portraits, different kinds of photographs have different subject focus and different formal needs, and therefore need to be edited in different ways. This chapter addresses the specific genre of portrait photographs which are photographs that depict a person. Whether you are travelling, enjoying a family event or recording your glory moment, everyone ends up shooting portraits. Portrait photographs constitute an important part of our personal photograph collection.

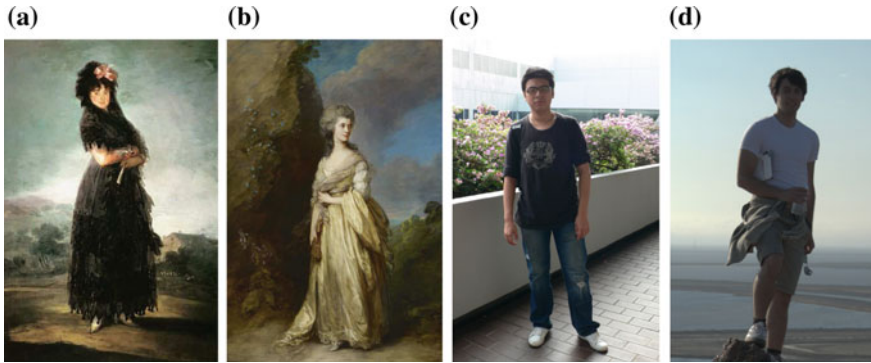


Fig. 7.1 **a** ‘Portrait of Mariana Waldstein’, by Francisco de Goya (1746–1828). **b** ‘Mrs. Peter William Baker’ by Thomas Gainsborough (1727–1788). Courtesy Wikimedia Commons. **c** and **d** are photographs. Image **d** is cropped from an image in MIT-Adobe FiveK Dataset [2]

In a portrait, the figure, especially the face, is of prime importance. However, the key to controlling the strength of the attraction of a portrait is not by just controlling the absolute value of the region of interest, but rather the contrast *between* and *within* the face and skin areas of the foreground figure, the non-face/skin part of the foreground figure with its accessories and the background [21]. The management of this network of contrast serves to make the figure and face stand out. Therefore, the proposed portrait photograph enhancement method does not solely focus on the enhancement on the face, but on managing these relative relationships. A content-aware automatic photograph enhancement method was proposed in [15] in order to enhance the face, sky and salient areas differently. However, it processed these regions separately and did not consider their interrelationships. In addition, the face and sky in all the photographs were shifted to a pre-defined value. As a result, the contrast on the face is sometimes reduced as well as the relative contrast between the face and its background. Although Adobe Photoshop provides a powerful tool for users to manually edit their images, the skills required to achieve this region-specific improvement are beyond the casual user, for which users would have to mask regions and make high-level creative decisions.

Fortunately, portrait paintings supply us with good examples which can be used as references in the modification of the relative relationships of these regions. Compared to a photograph, which can be regarded as a projection of optical appearance, a traditional portrait painting is very carefully constructed. Portrait artists are able to filter and manipulate the subject scene based on their visual perception [7]. Key to this is the consideration given to the organisation of colour contrast values [1]. In the example painting shown in Fig. 7.1a, the foreground figure is darker than the outdoor background (BG). In the painting shown in Fig. 7.1b, the outdoor BG is darker than the figure. Despite the difference in strategies that have been employed, in all the three paintings, the artists have maintained a visual emphasis on the figure. The face and skin areas (FS) are almost the brightest parts of the figure and are in a high contrast with the non-face/skin part of the foreground (FO) and the BG regions. Foreground (FG) in this chapter is comprised of the figure and its accessories. In addition to



Fig. 7.2 One of the results produced by the proposed method. **a** Original photograph. **b** Reference painting (CC Image courtesy of <https://www.wikiart.org>). **c** Enhanced result

the contrast organisation *between* these regions (inter-contrast), there is also contrast organisation *within* each region (intra-contrast) to emphasise the face, the pose of the body and also the richness of the BG. Differently, in snapshot style photographs which are more casually framed ('point and shoot'), their lighting, colourfulness and the regional contrast are usually not well organised (e.g. the two photographs in Fig. 7.1c, d).

In order to better understand the regional contrast organisation of paintings, a statistical comparison of the contrast values of lightness and saturation is conducted between photographs and paintings. Hue is not considered, as any manipulation of this value could negatively impact the natural appearance of the photographs. Differences are observed which evidence that the intra-contrast and inter-contrast of the lightness and saturation in the paintings are more purposefully organised. This contrast organisation purposely makes the figure, especially the face, 'stand out', while preserving the contrast within each region in order to enhance the visual appeal of the whole image. This observation provides justification for the enhancement of portrait photographs through contrast adjustment to their regions according to the corresponding contrast organisation in paintings.

In summary, a method is developed by which the high-level lightness and saturation contrast organisation of portrait paintings can be transferred to portrait photographs in order to enhance their visual appearance. This is done with respect to the contrast values of the three regions FS, FO and BG. These contrast values are considered both inter- and intra-regionally. Similar to the depth-aware contrast manipulation discussed in Chap. 6, an example-based method is proposed to manipulate the regional contrast of photographs. A selected reference portrait painting is used as an example to guide the enhancement of the portrait photograph. One of the results obtained from the proposed method is shown in Fig. 7.2.

7.2 Related Work

Portrait rendering. With the recent development of painterly rendering, portrait rendering has become an interesting topic in the computer vision and computer graphics research communities. Colton proposed to use a non-photorealistic rendering (NPR) system to automatically produce portraits based on recognised emotions [6]. The artistic styles in the NPR system were based on painting materials, colour palette and brush models. Some recent research attempts to render a portrait photograph in an artistic style [3, 19, 31, 34]. However, these styles are limited to simulating abstraction, line drawing or organic models. Face relighting is another popular portrait-related research topic. In [12], artistic face lighting templates were learned from a dataset of professional and amateur portrait photographs. The lighting template was based on the light distribution and shading of the face, e.g. left-weighted or right-weighted. Chen et al. developed an algorithm to relight the face based on a reference face image [4]. The work in [5] applied the face lighting template generated by artists onto the faces of photographs. All this research work only focuses on the manipulation of the face. Differently, the method proposed in our work not only considers the face, but also the relationships of the face with the BG and other parts of the figure. This is an important distinction, as painters rarely evaluate objects or features absolutely (i.e. in isolation), preferring instead a relative approach.

Content-aware image enhancement. While most common single image manipulation methods only automatically enhance the image globally or locally without considering the content, several content-aware methods have been proposed to process regions differently. Rivera et al. proposed to adjust the contrast differently in the dark, middle and bright regions [24]. The contrast in each region was mapped separately, and the mapping function was generated based on the analysis of the content. Finally, the mapping results were merged using a weighting function. This context-aware method aimed to reduce the artefacts and other unnatural effects in the resultant images. The considered regions were split based on the intensity value and had no high-level semantic meaning. In addition, only lightness was adjusted. Differently, the content-aware automatic photograph enhancement method proposed in [15] aimed to enhance the face, sky and salient areas differently, but it did not consider the relative relationships of these regions.

To circumvent the difficulty in defining the enhancement level for different images in single image manipulation methods, learning-based methods have been proposed to enhance images so that they match specific styles. Bychkovsky et al. [2] proposed to learn the photographic global tonal adjustment function from photographers in order to personalise tone adjustment. The global tone adjustment function was trained on input–output images pairs, where the output images were those enhanced by professional photographers. Similarly, Wang et al. applied learned colour and tone mapping in order to stylise the image enhancement [29]. These methods mostly rely on lower-level image statistics (e.g. colour, intensity gradient) and do not exploit object-level semantics other than the face in [2], and even then only its global tone mapping function is learned.

Joshi et al. proposed to improve the aesthetic appeal of faces in a personal photograph collection by leveraging photographs of the same person which have been selected for their quality [13]. This work only focused on performing both global and face-specific corrections. Hwang et al. [11] proposed to locally correct the colour and tone of an image by searching for the best transformation for each pixel using image pairs before and after enhancement. The search was done by finding the best few candidate image pairs and then identifying the best matching pixels from these pairs. The search was based on local scene descriptors and context. The mapping derived from the image pair on the matched pixel was then applied to the input pixel. This example-based method requires that the enhancement parameters for the example image pair are known a priori. For pre-modern painting styles, it is not possible to know the enhancement parameters, therefore the method is not applicable.

7.3 Paintings Versus Photographs

A statistical comparison of the inter-contrast values and intra-contrast values of lightness and saturation was performed between a set of photographs and a set of paintings. In addition, the mean values of lightness and saturation of regions were also compared to show the region-specific difference in the distribution of these values.

Portrait paintings collected for this study were chosen from artists in the 17th–18th centuries such as Thomas Gainsborough (1727–1788), Francisco Goya (1746–1828), Edouard Manet (1832–1883) and Jean-Auguste-Dominique Ingres (1780–1867). These artists maintained a high dynamic range in their paintings which was broadly similar to that of photographs. Most paintings done in the twentieth century (e.g. the Impressionists and Post-Impressionists) would not have this quality. The database covers a variety of environments and relationships of FG and BG. In total, 300 portrait paintings were collected to form the database. A total of 180 of them were half-body portrait paintings, and 120 of them were full-body portrait paintings. A total of 250 portrait photographs from the MIT-Adobe FiveK Dataset [2] and 50 from our personal collection were collected to form the photograph dataset.

Similar to the inter-contrast definition in Chap. 6, the inter-contrast between two regions with mean values M_1 , M_2 was defined as $\frac{2|M_2 - M_1|}{M_1 + M_2}$. Here only the absolute inter-contrast value was considered because the order of the mean values can be clearly seen from the plot of mean values. In the intra-contrast definition, which measures the local contrast of the region, the gradient range was not used. This was because the portrait is very sensitive to the gradient change on edge regions, especially those around the face. Therefore, the intra-contrast of one region was defined based on the value in the spatial domain, which was $\frac{I_{max} - I_{min}}{I_{mean}}$, where I_{max} and I_{min} were the maximum and minimum values of one region, respectively, and I_{mean} was the mean value of the region.

Box plots of the mean and contrast values are shown in Fig. 7.3. The mean values show that the FS region is generally the brightest part of the painting and much more

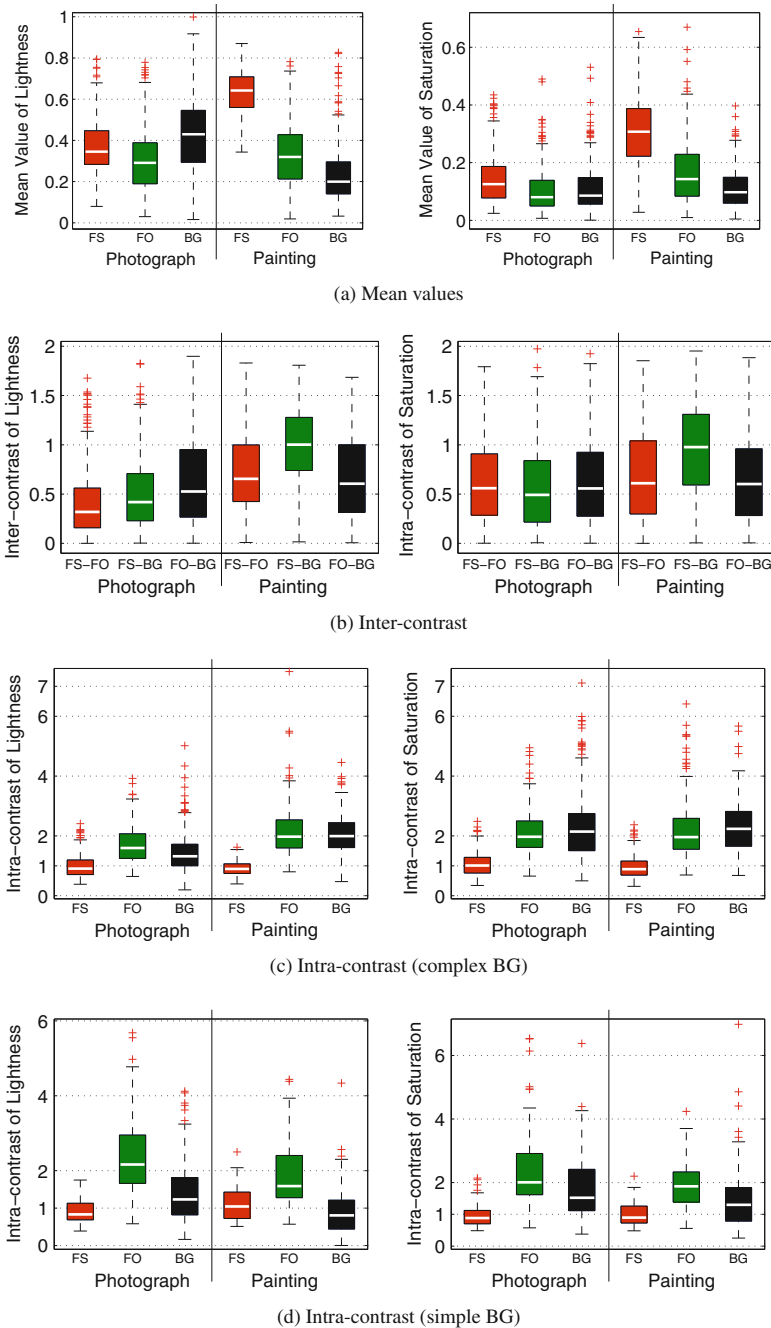


Fig. 7.3 Box plots of mean values and inter-contrast values of 300 paintings and 300 photographs in **a** and **b**. **c** Box plot of intra-contrast values of paintings and photographs with a complex BG. **d** Box plot of intra-contrast values of paintings and photographs with a simple BG

colourful than the FO and BG. This is coherent with the objective of portraiture, which is to focus on the expression of the figure, especially the face. However, this tendency cannot be observed in the lightness of photographs, where the BG is generally the lightest region. Although the saturation of photographs has this tendency, the difference is not as clear as that of paintings. The inter-contrast describes the relative difference of mean values. Figure 7.3b clearly shows that the inter-contrast values of lightness and saturation between FS and FO, FS and BG of paintings are much higher than those of the photographs.

In the comparison of intra-contrast, the paintings and photographs were classified into two classes based on the complexity of the BG. One class features a complex BG (intra-contrast is shown in Fig. 7.3c), the other features a simple BG (intra-contrast is shown in Fig. 7.3d). In those with a complex BG, the intra-contrast values of lightness are more likely to be higher in FO and BG than those of the photographs. However, for the intra-contrast of saturation, there is no clear difference between paintings and photographs. For the simple BG class, in both paintings and photographs, the intra-contrast of the FO is larger than that of the BG. However, the intra-contrast values of the FO and BG of paintings are smaller than those of photographs. One interesting point is that the lightness and saturation intra-contrast values of the FO in photographs with a simple BG are larger than those with a complex BG. These differences in intra-contrast organisation between paintings and photographs show that artists exaggerate the contrast of regions in paintings. The intra-contrast of a complex BG is exaggerated, and the intra-contrast of a simple BG is reduced.

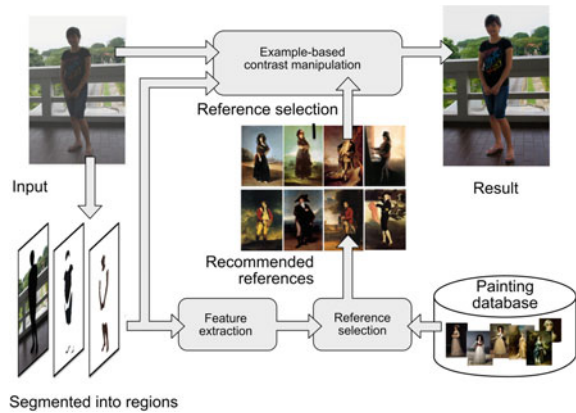
In summary, the regional contrast of paintings is more purposely organised than that of photographs. Both an enhancement and reduction of contrast are used by painters to attract the focus of attention to the face or figure in portrait paintings.

7.4 Example-Based Portrait Photograph Enhancement

Statistics in the preceding section clearly show the purposely created lightness and saturation inter-contrast and intra-contrast organisations of paintings. This section reports the attempt to use this contrast organisation as reference for the enhancement of digital photographs. The framework of the proposed example-based portrait photograph enhancement method is shown in Fig. 7.4.

This example-based method starts with the segmentation of the FG/BG and FS area detection. Although a lot of research has been done on human detection, it is still a challenge to accurately extract the human form in 2D images. Moreover, in half-body portraits, where only the face and part of the upper body are visible, most human detection methods do not work. In our work, GrabCut interactive segmentation [25] was used to segment the FG/BG automatically with an initialisation window. One thing consistent in portraits is that the face is always facing the viewer. Therefore, the face was used as the most important information for the purpose of detecting the position of the figure. The Haar face detector was used to detect the face [16, 28]. Then, based on the size and location of figures, an enlarged window that was likely to

Fig. 7.4 Framework of the proposed method



cover the FG area was defined in order to initialise the GrabCut. When the detected single face was around the centre of the image horizontally, the enlarged window centered at the face window. When the face window was in the right or left part of the image, more space was given to the left or right side of the enlarged window, respectively. For an image with a complex BG, the segmented FG may not be so accurate. As in [25], a user interface was provided for users to draw some strokes to refine the results. The face mask and skin areas were detected based on the skin colour in the figure using the method in [8] with some morphological post-processing. Some results are shown in Fig. 7.5.

Given the segmented BG region, detected face and skin areas, and the FO region, the portrait painting database was then searched for the best matching reference. The similarity of paintings and photographs was judged based on their original contrast values.

Using a user-selected painting among the top-ranked reference paintings as an example, the regional contrast in the portrait photograph was then enhanced. The operation on lightness and saturation was performed separately in the same way. Finally, the transition boundary between regions was smoothed.

7.4.1 Reference Selection

The main objective of the reference selection is to select reference paintings that can be used to visually improve the original input, while not pushing the original image too far away from its original natural property. Because the proposed portrait enhancement method serves to enhance the relative relationships of regions, therefore the reference paintings should be selected with reference to these relative relationships. The relationship among regions can be modelled as a graph as shown in Fig. 7.6. The features within regions $\Psi = \{\Psi_S, \Psi_O, \Psi_B\}$ and the relationships among

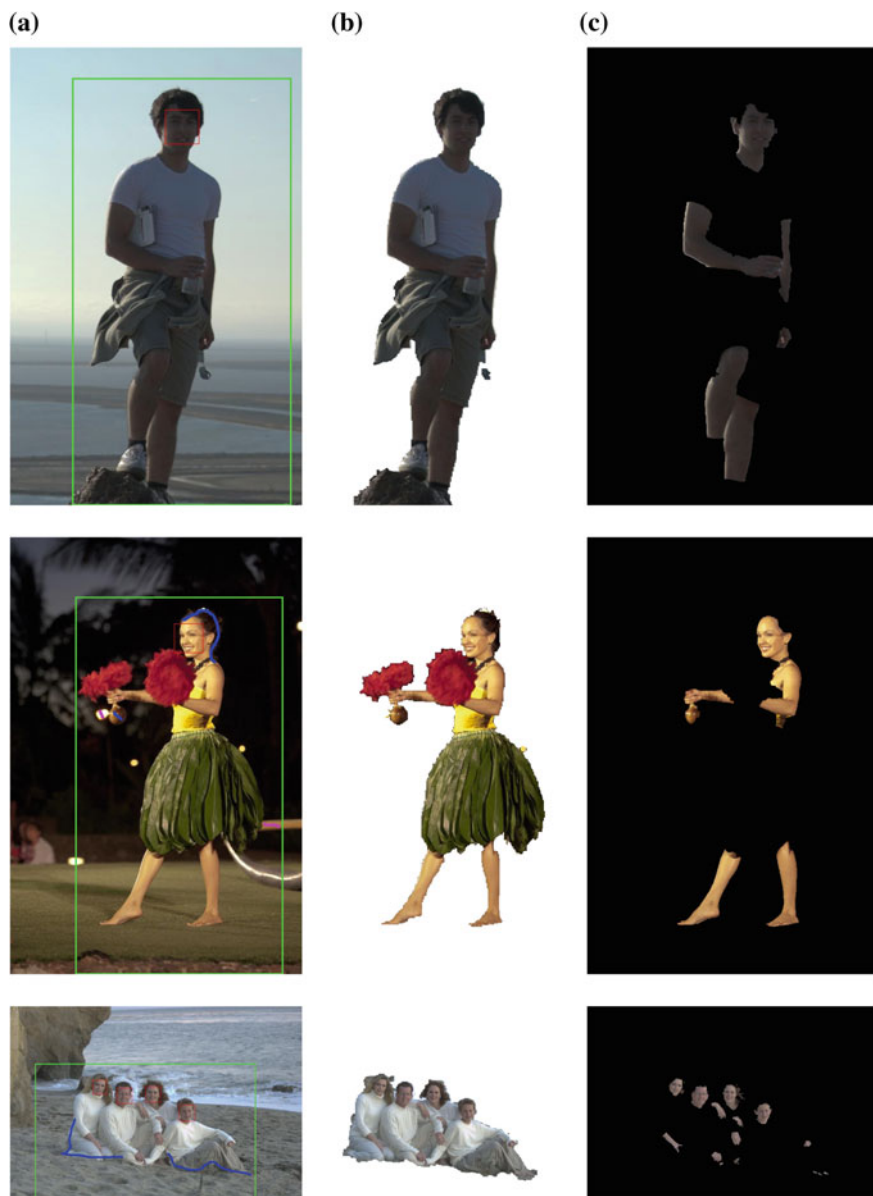
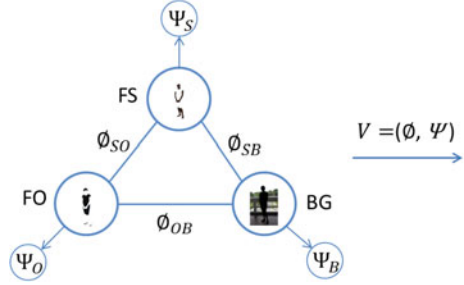


Fig. 7.5 FG/BG segmentation and face/skin area detection. **a** Input photograph. **b** Extracted FG. **c** Face and skin. The *red* window in the input photograph is the face window, the *green* window is the initialisation window for GrabCut. The *blue* and *purple* lines are the FG and BG brushes for refining the FG/BG segmentation (colour figure online)

Fig. 7.6 The graph model



regions $\Phi = \{\Phi_{SO}, \Phi_{SB}, \Phi_{OB}\}$ are used to select reference paintings, where Ψ_S, Ψ_O and Ψ_B are the features within FS, FO and BG, respectively, and Φ_{SO}, Φ_{SB} and Φ_{OB} are the relationship features between FS and FO, FS and BG, FO and BG, respectively. After calculating the graph $V = (\Phi, \Psi)$ for each image, the task is to determine the similarity of the graphs between input photographs and paintings. Particularly, paintings that can be used to visually improve the original input photograph should be ranked higher in the selection in terms of graph similarity according to a distance metric. Therefore, distance metric learning was employed to determine the similarity of paintings and input photographs based on the graph model.

The features within each region were mean values of lightness and saturation, the 10-bins histograms of lightness and saturation, the maximum (95th percentile value), minimum (5th percentile value) values of lightness and saturation distributions, and the intra-contrast values of lightness and saturation. The BG global contrast factor G_B was also calculated as a feature to express the complexity of the BG. A highly detailed and variation-rich BG will have a high global contrast factor, and simple BG will have a low global contrast factor [18]. In this manner, the contrast factor can represent the difference between simple and complex BGs. This was calculated as the weighted average of local contrast values at various resolution levels (more details can be found in [18]). The relationships among regions Φ were defined by the inter-contrast values of lightness and saturation between FS and FO, FS and BG, FO and BG.

Given the features, the distance metric between the input photograph I_i and painting R_j is

$$D(i, j) = \sum_{k=1}^N \alpha_k (v_i^k - v_j^k)^2 \quad (7.1)$$

where α is the parameter to linearly combine the distance of features. v_i^k is the k -th feature vector formed by concatenating the ordered graph attributes of the photograph I_i , and v_j^k is the k -th feature in the feature vector of painting R_j . The objective here is to determine the parameter α such that paintings which can be more likely to create more pleasing results are closer to the input photograph than others. As in [11, 14], the parameter α was learned by using a target distance function $D_t(i, j)$. The parameter α was determined by minimising the objective function

$$\arg \min_{\alpha} \sum_{i,j} \|D(i, j) - D_t(i, j)\|^2 \quad (7.2)$$

The target distance D_t was defined based on the assigned score by a human expert for an enhanced photograph from a photograph–painting pair. Smaller distance signifies more pleasingly enhanced photographs, and larger distance signifies unacceptably enhanced photographs. Minimising this objective function returns an appropriate distance function that reflects how close the photograph and painting should be in terms of the enhancement result. This objective function is convex, and the unique optimum can be easily found by running a gradient descent procedure.

To learn the distance metric, a collection of 300 portrait paintings (180 half-body portrait paintings and 120 full-body portrait paintings) were used as references. The target distance function was $D_t(i, j) = \frac{1}{S(i, j)}$, where $S(i, j)$ was the assigned score for the enhanced photograph I_i using painting R_j as reference by an expert who had lot of experience on photograph adjustment. It was a challenge to directly assign an aesthetic value score to an image. It was more convenient to classify enhancement results as being ‘good’, ‘acceptable’ or ‘unacceptable’. An expert was therefore asked to classify the enhancement of 32 half-body portrait photographs and 30 full-body portrait photographs into these three categories and assigned to each was a pre-set score value. For each half-body portrait photograph, 180 enhanced photographs were obtained, while 120 enhanced photographs were obtained for each full-body portrait photograph. In all the enhanced photographs, an expert only classified those that could be confidently assigned to one of the three categories, otherwise they were discarded. In total, 2776 samples were collected in order to learn the distance metric for half-body portraits and 1381 samples for full-body portraits.

7.4.2 Contrast Mapping

Based on the definition of intra-contrast, adjusting the maximum, minimum and mean values of lightness or saturation distribution can change the intra-contrast of a region. Meanwhile, adjusting the mean value of each region changes the inter-contrast of regions. More formally this means that the mean, maximum and minimum values of regions in photographs can be changed so as to cohere with corresponding mapped regions in reference paintings. Given the target mean value m_s^t , maximum I_{max}^t , minimum I_{min}^t , the task of the contrast mapping is

$$I(m_s, I_{max}, I_{min}) \xrightarrow{f} I'(m_s^t, I_{max}^t, I_{min}^t) \quad (7.3)$$

Piecewise linear contrast stretching has been used in [33] to adjust the mean, maximum and minimum values of regions. This piecewise linear transformation is not continuously differentiable across the histogram span, which may cause contouring artefacts [19]. Moreover, this method may produce out of range mapping.

To more effectively perform contrast mapping, an interpolating transformation curve was defined piecewise which was monotonically increasing and continuously differentiable. This interpolating transformation curve was known as histogram warping in [10, 19]. For a given set of control points (a_k, b_k, d_k) , where $b_k = f(a_k)$, $d_k = f'(a_k)$ which was the contrast adjustment at the key value a_k , the transformation curve was generated using a piecewise rational quadratic interpolating spline [26]:

$$f(x) = b_{k-1} + \frac{(r_k t^2 + d_{k-1}(1-t)t)(b_k - b_{k-1})}{r_k + (d_k + d_{k-1} - 2r_k)(1-t)t} \tag{7.4}$$

where $r_k = \frac{b_k - b_{k-1}}{a_k - a_{k-1}}$, and $t = \frac{x - a_{k-1}}{a_k - a_{k-1}}$ for $x \in [a_{k-1}, a_k]$.

In [19], the control points were specified by users, while in [10] they were automatically generated based on the analysis of the histogram. Differently, in our work, three control points were automatically defined based on the original and target mean, minimum and maximum values. Specifically, the three control points are $(a_1, b_1, d_1) = (I_{min}, I_{min}^t, d_1)$, $(a_2, b_2, d_2) = (m_s, m_s^t, d_2)$ and $(a_3, b_3, d_3) = (I_{max}, I_{max}^t, d_3)$, and two endpoints are $(0, 0, d_0)$ and $(1, 1, d_4)$. The definition of the contrast adjustment at each point controls the shape of the transformation curve. To fit the piecewise linear curve as presented in [10], the contrast adjustments at the control points were defined as the geometric mean of the slopes weighted by the probability mass of the slopes' intervals. Given $d_0 = \frac{b_1}{a_1}$, $d_4 = \frac{1-b_3}{1-a_3}$, the contrast adjustments $d_k, k = 1, 2, 3$ are

$$d_k = \left(\frac{b_k - b_{k-1}}{a_k - a_{k-1}} \right)^{g_k^l} \left(\frac{b_{k+1} - b_k}{a_{k+1} - a_k} \right)^{g_k^h} \tag{7.5}$$

where $g_k^l = \frac{F(a_k) - F(a_{k-1})}{F(a_{k+1}) - F(a_{k-1})}$ and $g_k^h = \frac{F(a_{k+1}) - F(a_k)}{F(a_{k+1}) - F(a_{k-1})}$ are the slope weights. $F(x)$ is the cumulative distribution function. An example of the transformation curve is shown in Fig. 7.7a.

Effective contrast within an image is not simply a case of ‘the more the better’. Artists are skilled at controlling the perceptually high contrast values of their paint-

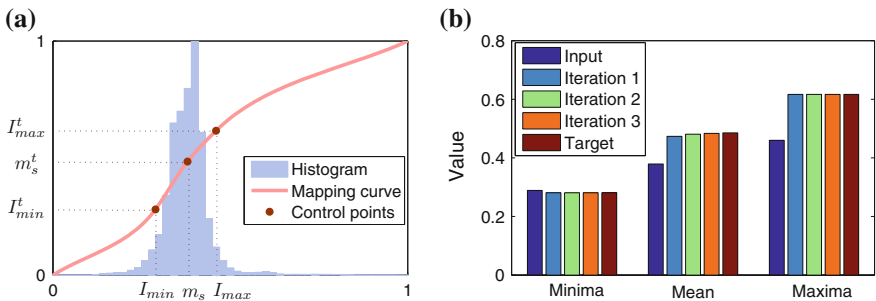


Fig. 7.7 Contrast mapping. **a** Transformation curve. **b** Convergence of maximum, minimum and mean values

ings. However, imposing high computational contrast onto a photograph cannot ensure a perceptually agreeable result. Therefore, the target inter- and intra-contrast may not need to match exactly that of the reference. In other words, the reference painting provides the template by which the photograph will be changed, but how close the match is to this template depends on the limits of the original image and the user's preference. Thus, two parameters α , β were used to control the inter- and intra-contrast mapping subject to the original image and to the user's preference. Since the inter-contrast is the relative difference of mean values, and the mean lightness and saturation values also influence the appearance of the image, the inter-contrast mapping was then performed by moving the mean value of each region to somewhere close to that of the reference. Given the mean value m_s of one region I of the photograph and the mean value m_r of the corresponding region I_r in the reference, the mean value was shifted towards to m_r , and the magnitude of shift was controlled by a weight α in [0-1].

$$m_s^t = m_s + \alpha(m_r - m_s) \quad (7.6)$$

Given the intra-contrast of I and the intra-contrast of I_r as c_s , c_r respectively, the target intra-contrast is

$$c_s^t = c_s + \beta(c_r - c_s) \quad (7.7)$$

where β is a weight in [0-1]. Based on the definition of intra-contrast, the degree of enhancement on the maximum–minimum will be $\lambda = \frac{c_s^t m_s^t}{c_r m_s}$. Then, the target maximum and minimum values are set as

$$I_{max}^t = \lambda(I_{max}^r - m_r) + m_s^t \quad (7.8)$$

$$I_{min}^t = \lambda(I_{min}^r - m_r) + m_s^t \quad (7.9)$$

where I_{min}^r and I_{max}^r are the minimum and maximum values of I_r . The target maximum and minimum values are limited to the range [0.01, 0.99]. As shown in Fig. 7.7b, three iterations are sufficient to map the mean, maximum and minimum to the target values.

Finally, to smooth the boundary between regions and to keep the change spatially coherent, edge-preserving smoothing [9] was performed on the difference image between the input F (L or C channel in LCH space) and contrast mapping result \tilde{F} , $d = \tilde{F} - F$. It used the following energy function:

$$E = \sum_{p \in D} \{[D(p) - d(p)]^2 + w_s(p)h(\nabla D, \nabla F)\} \quad (7.10)$$

where

$$h(\nabla D, \nabla F) = \frac{|D_x(p)|^2}{|g_x(p)|^\theta + \varepsilon} + \frac{|D_y(p)|^2}{|g_y(p)|^\theta + \varepsilon} \quad (7.11)$$

D is the desired smoothed difference image, (D_x, D_y) is the gradient of D , (g_x, g_y) is the gradient of F . θ controls the sensitivity to the gradient of F . ε is a small regularising constant. Smoothing is employed to make the gradients of the output D as small as possible, unless the input F had a significant gradient. The weight w_s is defined as

$$w_s(p) = \begin{cases} \tau_1 & \text{if } p \in \text{boundary area} \\ \tau_2 & \text{otherwise} \end{cases} \quad (7.12)$$

τ_1 is larger than τ_2 for smoothing the boundary area between regions. In the implementation, the parameters are $\theta = 1.2$, $\varepsilon = 0.0001$, $\tau_1 = 0.3$ and $\tau_2 = 0.05$. The minimisation of E can be achieved using standard or weighted least-squares techniques like the conjugate-gradient method.

7.5 Experiments and Discussion

In this section, the performance of the proposed portrait photograph enhancement method is analysed. Among the 300 photographs, 62 of them were used for the distance metric training in the reference selection, and the others were used as test photographs.

Influence of reference. The enhancement results of two photographs using five selected top-ranked reference paintings are shown in Fig. 7.8. The results using the reference paintings selected by the reference selection method are encouraging. By applying the proposed enhancement method, more details of the figure, especially around the face area, are visible and the figure stands out in the images. However, if a painting not selected by the reference selection method is used as the reference (as the paintings in the seventh column of Fig. 7.8), even though the face and figure also stand out, the change is too far away from its original natural property (e.g. the pink T-shirt becomes black). This indicates that the reference selection method is effective. While the two example photographs in Fig. 7.8 are of outdoor scenes with complex BGs, Fig. 7.9 shows two indoor photographs with relatively simple BGs. The first one has a bright BG and dark FG, and the other has a dark BG and bright FG. We can see that the selected reference paintings have similar natural properties to those of the input photograph. In other words, the photograph with a dark FG and bright BG is more likely to be improved using a reference painting with similar dark FG and bright BG.

Effect of parameters. In the contrast mapping, the same α and β are used for the three regions FO, FS and BG. Parameters for mapping lightness are expressed as α_l , β_l , and for saturation α_s , β_s . When $\alpha_l = \alpha_s$, $\beta_l = \beta_s$, they are indicated as α , β . α and β control the degree of match to the contrast of the reference. Figure 7.10 shows the effect on the appearance of the enhancement results for changing α_l , β_l from 0 to 1. When α_l increases from 0 to 1, the inter-contrast of lightness is moved towards to that of the reference (see Fig. 7.10c). The figure stands out and becomes brighter as α_l increases in Fig. 7.10. When β_l changes from 0 to 1, the intra-contrast values

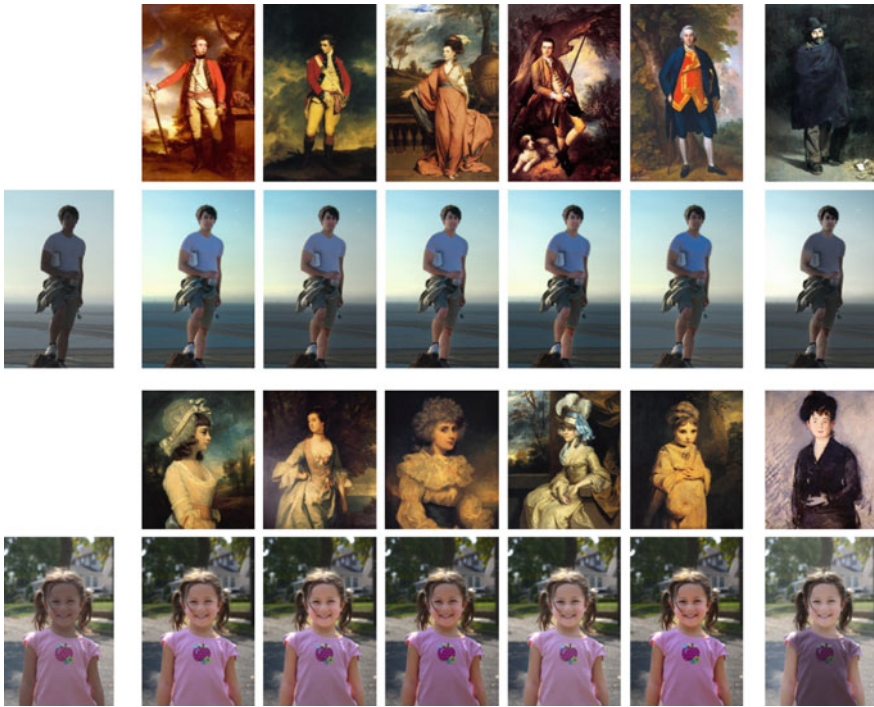


Fig. 7.8 Experiment on two outdoor photographs with complex BGs. The first column: original photographs. The second to the sixth columns: 5 of the selected reference paintings for the two photographs and the corresponding results (CC Image courtesy of Wikimedia Commons, and <https://www.wikiart.org>). The seventh column: paintings not recommended as reference and the corresponding results. The parameters for the first example are $\alpha_l = 0.5, \alpha_s = 0.4, \beta_l = \beta_s = 0.5$. The parameters for the second example are $\alpha_l = 0.6, \alpha_s = 0.6, \beta_l = \beta_s = 0.5$

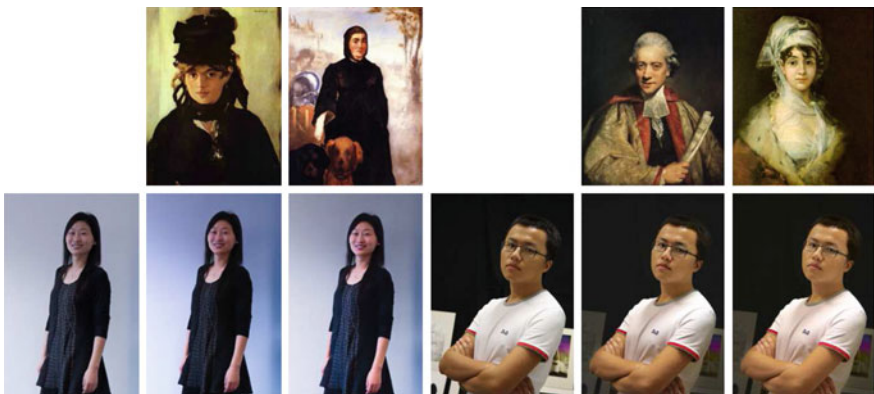


Fig. 7.9 Experiment on two indoor photographs with simple BGs. The first and fourth columns: original photographs. The second to third and fifth to sixth columns: two of the selected reference paintings (Courtesy Wikimedia Commons) for the two photographs and the corresponding results

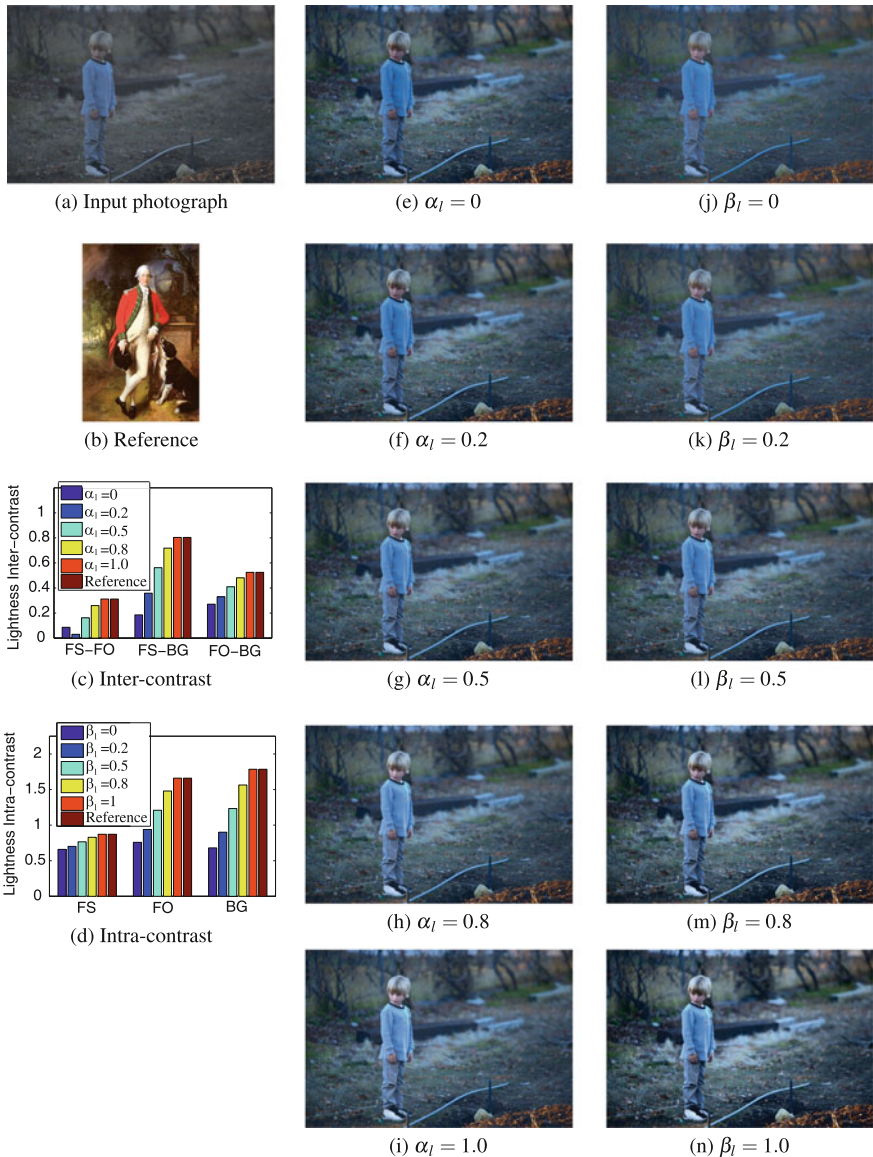


Fig. 7.10 Influence of α, β . **e-i** $\alpha_s = 0.4, \beta_l = \beta_s = 0.5$. **j-n** $\alpha_l = 0.5, \alpha_s = 0.4, \beta_s = 0.5$. The reference painting (b), cc courtesy Wikimedia Commons

of lightness in the three regions are moved to those of the reference correspondingly (see Fig. 7.10d). More local details in the BG become visible when β_l increases. The influence of α_s and β_s follows a similar form.

In the reference selection, the enhanced photographs used for distance metric learning are obtained by using the default parameters $\alpha_l = 0.5$, $\alpha_s = 0.4$, $\beta_l = \beta_s = 0.5$. Therefore, with the selected reference paintings ranked by the proposed reference selection method, acceptable results can be obtained using these default parameters. The user can further adjust these to change the appearance according to personal preference. Paintings are generally drawn with high contrast between regions and with generally high saturation values. If $\alpha = 1$, the enhanced photograph may become visually unacceptable. Generally, α_l is in the range $[0.3, 0.7]$, and α_s is in $[0.2, 0.6]$. β_l and β_s are suggested to be set in a similar way, and generally they are in the range $[0.3, 1]$.

Comparison. To assess the performance of the proposed portrait photograph enhancement method, it was compared to the content-aware enhancement (CAE) method in [15]. The CAE method had been reported to perform better than the global learning-based enhancement method in [2]. It was also compared to popular software tools ('I'm Feeling Lucky' of Google's Picasa, 'Auto Correct' of Microsoft's Office Picture Manager and 'Auto Smart Fix' of Photoshop). In addition, results achieved using the proposed method were also compared with results enhanced by using the average mean and contrast values. The average mean and contrast values were obtained from the study of contrast organisation in paintings shown in Fig. 7.3. This is referred to as the AVE method. Figure 7.11 shows 5 results enhanced using the three methods. It shows that faces and figures in the results achieved using the proposed method become more salient as compared to those obtained by the CAE method. Moreover, the light falling on the face in the results obtained by the proposed method is more naturally presented. However, in the results on the second and third rows using the CAE method, the light falling on the face appears unnatural because the lightness was processed differently on the left side of the face, to the right. The enhancement results using the AVE method in the first to third rows were comparable with those obtained by the proposed example-based method. However, for the photographs in the fourth and fifth rows, the proposed example-based method is shown to have performed better than the average values method. This indicates that the example-based method can handle a diversity of scenes more effectively.

User studies. In order to evaluate the effectiveness and advantage of the proposed method, two user studies were conducted. The purpose of the first was to evaluate the effectiveness of the proposed reference selection method and the effect of the proposed portrait photograph enhancement method. The purpose of the second was to compare the effect of the proposed portrait photograph enhancement method with other methods.

For the first user study, 40 of the test photographs were randomly selected from a variety of scenes. This user study had two objectives. The first was to show that results obtained using the proposed method were more pleasing and that the figure and face were more salient. The second was to show that results obtained using the top-ranked reference paintings were more desirable than those obtained using other paintings. For the first objective, 5 reference paintings were randomly selected from the top 15 paintings for each of the 40 selected test photographs. The results obtained using the 5 reference paintings were all compared with the original photograph. Each

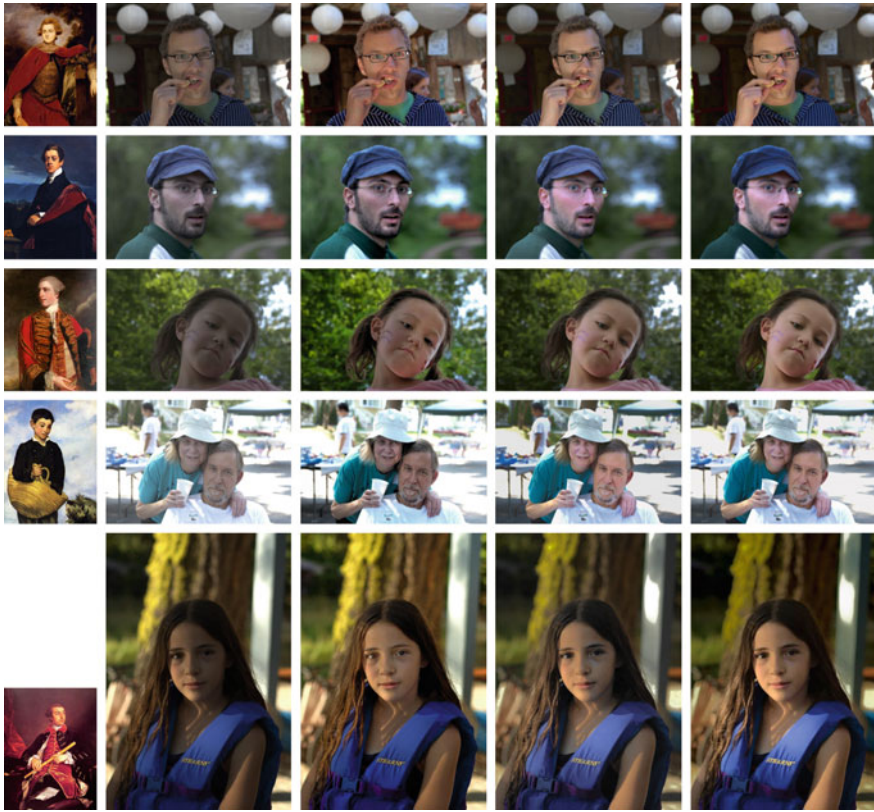


Fig. 7.11 Comparison with other methods. **a** Reference paintings (CC Image courtesy The Athenaeum, Wikimedia Commons). **b** Input photographs. **c** Results by CAE in [15]. **d** Results by AVE method. **e** Results by the proposed method. The photograph in the fifth row is from [15]

result and its corresponding original photograph were shown as a pair side by side. The left/right ordering of the images in each pair was randomly generated. For each image pair, the participants were asked to answer two questions. The first question Q1 was ‘In which image, are the figure and face more salient?’ and the second Q2 was ‘Which image is more pleasing?’. The participants could choose ‘left’ or ‘right’ as the response to each question, according to their first impression. No further information on the goals of the experiment or the origin of the images was provided to the participants.

For the second objective, the results using the reference paintings selected from the middle and bottom ranges of the ranking were compared with those obtained using the top-ranked reference paintings. Five paintings from the ranks in the range of 51–65 for full-body portrait and 5 paintings from the ranks in the range of 65–81 for half-body portrait were randomly selected as the middle-ranked reference paintings. Five paintings from the ranks in the range of 106–120 for full-body portrait and 5

paintings in the range of 166–180 for half-body portrait were randomly selected as the bottom-ranked reference paintings. The three groups of results obtained using the top-, middle- and bottom-ranked reference paintings were arranged in three rows. The order of the group arrangement was randomly generated. Participants were asked to select the group that they prefer more. The participant could select ‘Group 1’, ‘Group 2’ or ‘Group 3’.

The 40 photographs were divided into 7 sets. The first set consisted of 4 photographs while the other sets consisted of 6. The comparison in each set contained the comparison of the result obtained using the top-ranked reference painting with the original side-by-side and the comparison of the results of the three groups obtained using the top-, middle- and bottom-ranked reference paintings. In the first set, there were 24 comparisons, and in the other 6 sets, there were 36 comparisons. Google Form was used to create the webpage for the user study. The participants connected to the webpage using their own computers and monitors. Each participant could elect to complete one, two or three sets of comparisons. There were 38 participants, 21 were males and 17 were females, their ages ranging from 20 to 45. In total, 2070 samples were collected for the comparison of enhanced photographs versus originals. The average selection rate of enhanced photographs versus originals is shown in Fig. 7.12a. The bar graph shows that for both questions, a significant majority of the responses (more than 90%) indicate a preference for the results achieved using the proposed method. This demonstrates that, by using the proposed method, the visual appeal of the portrait photographs has been effectively improved and the face and the figure have become more salient. In the comparison among the three groups of enhanced photographs obtained by using top-, middle- and bottom-ranked reference paintings, respectively, more than 80% of the responses are in favour of the results obtained using the top-ranked reference paintings. This demonstrates the effectiveness of the proposed reference selection method and the importance of appropriate reference selection.

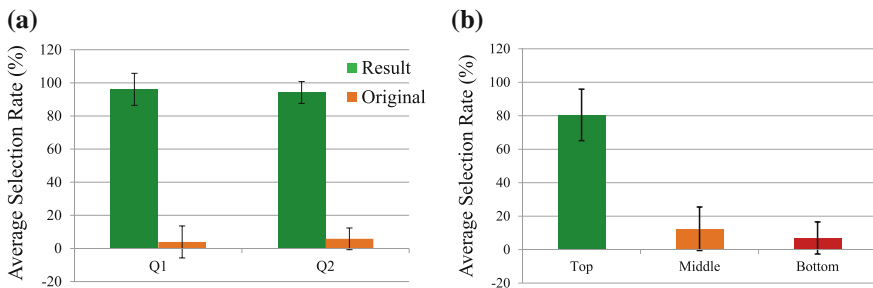


Fig. 7.12 User study result for the evaluation of the proposed method. **a** Selection rate for the comparison between enhanced photographs and originals. **b** The comparison of the selection rates of the three groups of photographs enhanced by using top-, middle- and bottom-ranked reference paintings respectively

Table 7.1 User study result for the comparison of the proposed method with the CAE method in [15] and the AVE method. The percentage shows the number of users responding favourably to the respective method

Proposed versus ACE [15]	Questions	Proposed	ACE	The same
	Q1	62.0%	30.6%	7.4%
	Q2	66.8%	29.6%	3.6%
Proposed versus AVE	Questions	Proposed	AVE	The same
	Q1	68.4%	19.6%	12.0%
	Q2	70.0%	22.1%	7.9%

The second user study was conducted to compare the results achieved using the proposed method against those by the CAE method in [15] and the AVE method. In [15], 55 images featuring a face were reported, 36 of which could be considered as a portrait and which could be used for this user study. The results by the proposed method and one of the 2 other methods were shown as a pair side by side. The left/right ordering of the images in each pair was randomly generated. For each image pair, the participants were required to answer the same two questions as in the first user study. Differently, the participants could choose ‘Left’, ‘Right’ or ‘They look the same’ as responses to each question according to their first impression. The 36 images were divided into 2 sets. Each participant could elect to complete one or two sets. There were a total of 29 participants ranging from 20 to 35, 18 of them were male and 11 of them were female. For each comparison, there were on average 18 responses. The result of the user study is shown in Table 7.1. It shows that results achieved using the proposed method are preferred by a significantly higher percentage of the respondees than other methods. This is not surprising, as the method in [15] cannot handle the relationship between regions and the AVE method does not consider the diversity of scenes.

Discussion. In the depth-aware contrast manipulation of landscape photographs covered in Chap. 6, the gradient is used to map the intra-contrast. However, the portrait is very sensitive to the change of gradient especially when performed on the face, as artefacts are easily produced in areas of high gradient magnitudes. In the example shown in Fig. 7.13d, there are clear artefacts in the boundary of the face area and the clothes area, and the highlight area on the face is not smooth. The comparison between the proposed method in this chapter and the method in Chap. 6 in Fig. 7.13 shows that the proposed method in this chapter is more suitable for the manipulation of contrast in portrait photographs.

While experiments have demonstrated the effectiveness of the proposed method, a few limitations are observed. The quality of the results relies on the success of FG/BG segmentation and face/skin area detection. Although interactive segmentation can reduce the segmentation errors, and the skin area detection in the segmented FG is also much more robust than the detection in the whole image, in some situations the

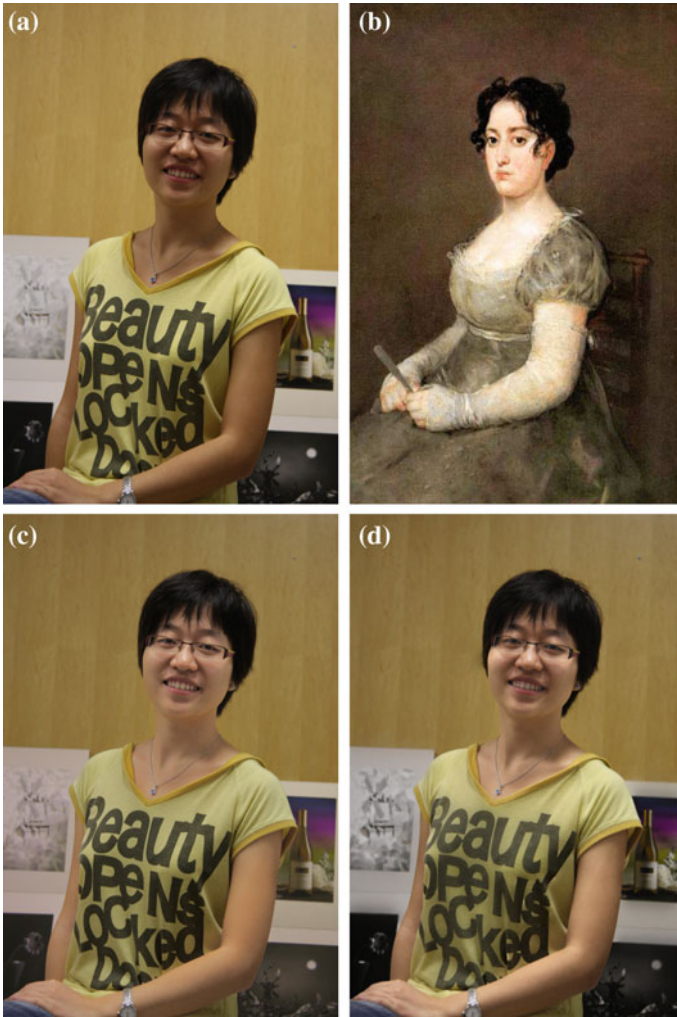


Fig. 7.13 Comparison with the method in Chap. 6. **a** Original. **b** Reference (CC Image courtesy Wikimedia Commons). **c** Result achieved using the proposed method. **d** Result achieved using the method in Chap. 6

skin area detection based on colour may still be disturbed by the colour of clothes. In the example shown in Fig. 7.14a, the colour and lightness of the clothes are very close to those of the skin areas. As a result, some regions of the clothes are detected as skin (see Fig. 7.14b). After regional contrast manipulation using the proposed method, these regions differ slightly from the rest of the clothes region (see Fig. 7.14c).

Another limitation is that artefacts near the transition boundary between the brightened FG region and the dark BG region may be visible. This is because the brightness

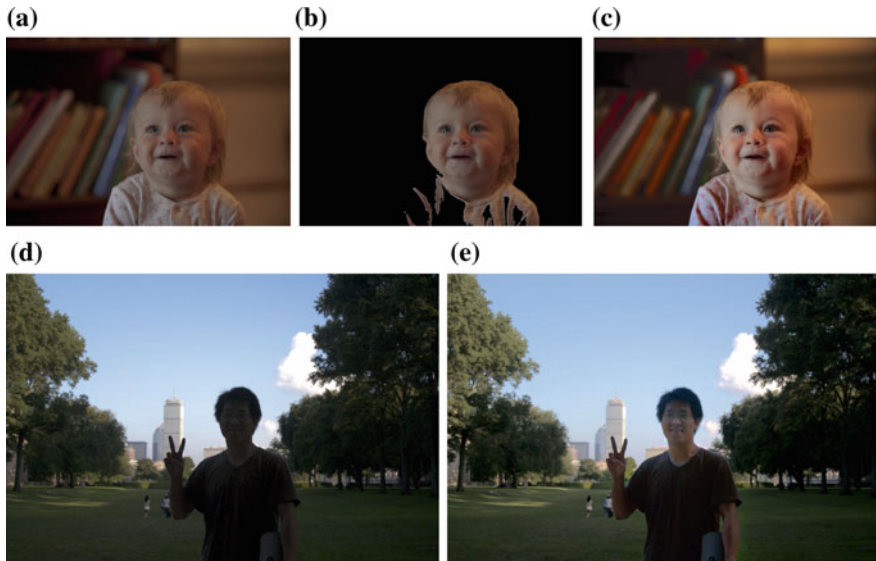


Fig. 7.14 Two examples demonstrating the limitation of the proposed method. **a** and **d** Original photographs. **b** Detected face and skin areas in **a**, **c** and **e** Results

of the FG region is extended to the neighbouring BG region by the boundary smoothing if there is no clear boundary between FG and BG in the original photograph. The technique in [9] was designed to smooth an image while preserving edges with large gradients. In the photograph in Fig. 7.14d, there is no clear edge between the right lower arm and ground. Therefore, the brightness on the lower arm is extended to the neighbouring BG region as a result of the edge-preserving smoothing (see Fig. 7.14e). The smoothing around the boundary region is stronger than that of the non-boundary region. Hence, artefacts may be visible. Some noise may also become visible in face and skin areas after being brightened using the regional contrast manipulation. More experimental results by the proposed method are shown in Fig. 7.15.

Application using RGBD images. New development of techniques, in commercial hardware, has made the capture of RGB/depth (RGBD) information more easily achievable. While the Xbox 360 Kinect depth sensor [20] by Microsoft can only capture depth for indoor scenes, the light field camera (e.g. Raytrix camera [23]) can capture the depth in both indoor and outdoor situations. The depth of scenes can also be calculated from videos by purposely planned camera motion that produces parallax. With depth, the figure can be easily extracted from the image. The framework of the proposed application using RGBD images is in Fig. 7.16.

The main element that influences the success of FG extraction from depth is the ground plane. As a consequence of the continuity of the ground plane and its connection with the feet, the FG cannot be extracted directly using a depth range threshold. However, the ground plane has a clearly different orientation to that of the figure. The orientation of a plane can be expressed by the normal vector or the

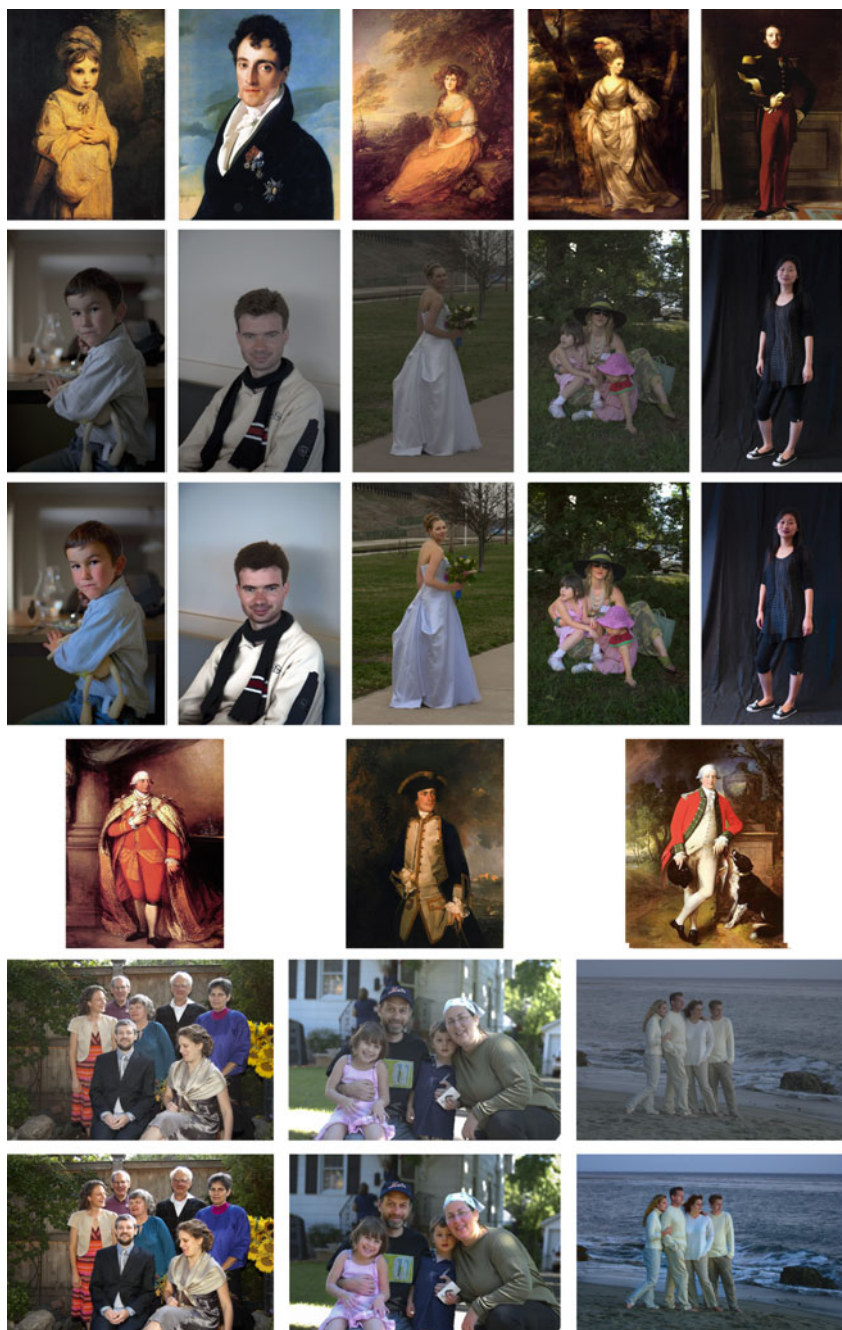


Fig. 7.15 More experimental results. The first and fourth rows: reference paintings (CC Image courtesy of <https://www.wikiart.org>, Wikimedia Commons, The Athenaeum). The second and fifth rows: original photographs. The third and sixth rows: results

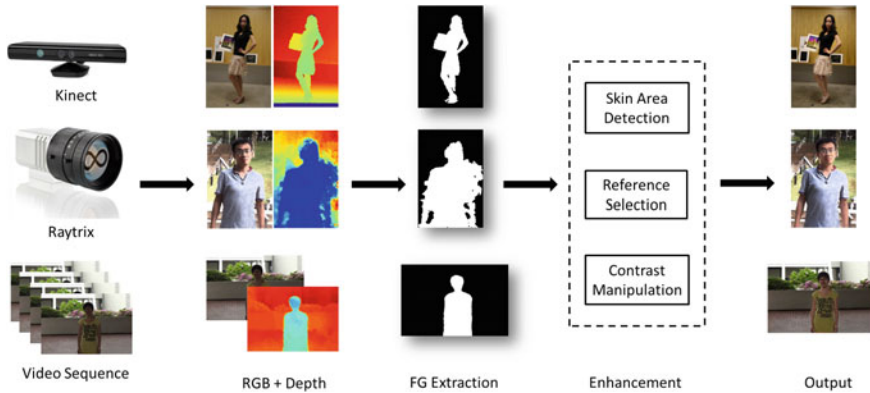


Fig. 7.16 Application using RGBD images. The depth from the video is the depth layer obtained by optical flow in [17]

gradient change on the plane. There are two options to divide the figure from the ground. One is to segment them based on the difference of normal vectors and depths. The other is to fit the ground plane based on the prior knowledge about the normal direction of the ground plane and then remove the ground plane.

Option 1: Segmentation (published in [32] ©2011 IEEE). Surface normals have been used for range image segmentation [22]. Similarly, the normal vector and depth are used as features to segment FG and BG. The normal vector of a flat plane is calculated by the cross product of two intersected vectors lying on the plane. Given the depth z , the normal vector of a plane, defined by a support region of $(w + 1)$ by $(w + 1)$, is calculated as

$$\begin{aligned}\vec{N} &= \vec{P}_1 \times \vec{P}_2 \\ \vec{P}_1 &= (x + w, y, z(x + w, y)) - (x, y, z(x, y)) \\ \vec{P}_2 &= (x, y + w, z(x, y + w)) - (x, y, z(x, y))\end{aligned}\quad (7.13)$$

where (x, y) is the top-left corner pixel of the $(w + 1)$ by $(w + 1)$ support region. To handle the quantisation error in depth data, w is set as $w = 6$. Then the normal vector of each pixel in the plane will be expressed as $\frac{\vec{N}}{\|\vec{N}\|} = (a, b, c)$.

After calculating the normal vector for each pixel, the feature (a, b, c, z) is used for region clustering. In this work, k-means method is chosen for its simplicity and speed. The initial cluster centres and the number of clusters are defined based on the prior knowledge of the portrait image. In a portrait image, certain assumptions can be made: the bottom part is the ground, the top part is the upper BG, and the figure is near the centre. Therefore, three cluster centres are initialised using the feature of the points in the bottom-left corner, the feature of the top-left corner points and the feature of the face area. For half-body portraits, the ground may not be in the scene. In this case, the bottom-left corner will be part of the BG or the figure. Therefore, if the

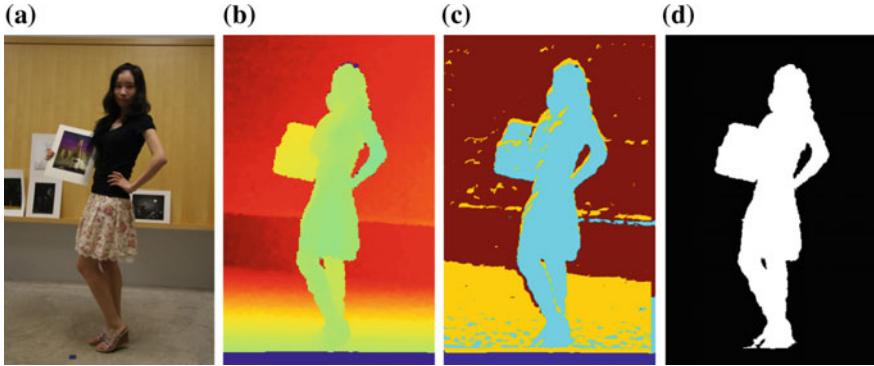


Fig. 7.17 FG extraction from RGBD image. **a** RGB image. **b** Depth. **c** k-means clustering. **d** Mask of extracted FG

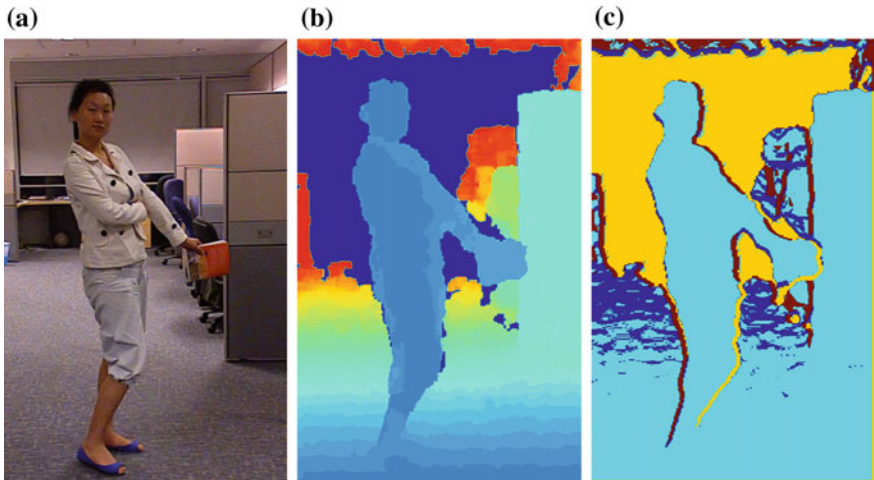


Fig. 7.18 FG extraction from RGBD image. **a** RGB image. **b** Depth. **c** K-means clustering

depth of the bottom-left corner is very close to that of the top-left corner or the face (difference is smaller than 10% depth range), it means that no ground is contained in the scene and the cluster number is reduced to 2. An example of the clustering result is shown in Fig. 7.17c. Post-processing is performed using morphological operations to remove the mis-segmented small regions. The final mask of the figure is shown in Fig. 7.17d. The k-means clustering method can generate results in realtime. However, for a scene with a complex BG, the method is not robust. In the example in Fig. 7.18, the k-means clustering cannot segment the FG from the ground plane.

Option 2: Ground fitting. Since the main element that influences the FG extraction is the ground plane, therefore the ground plane can be detected first and then removed. Region grouping can then be used to extract the FG by using a seed point

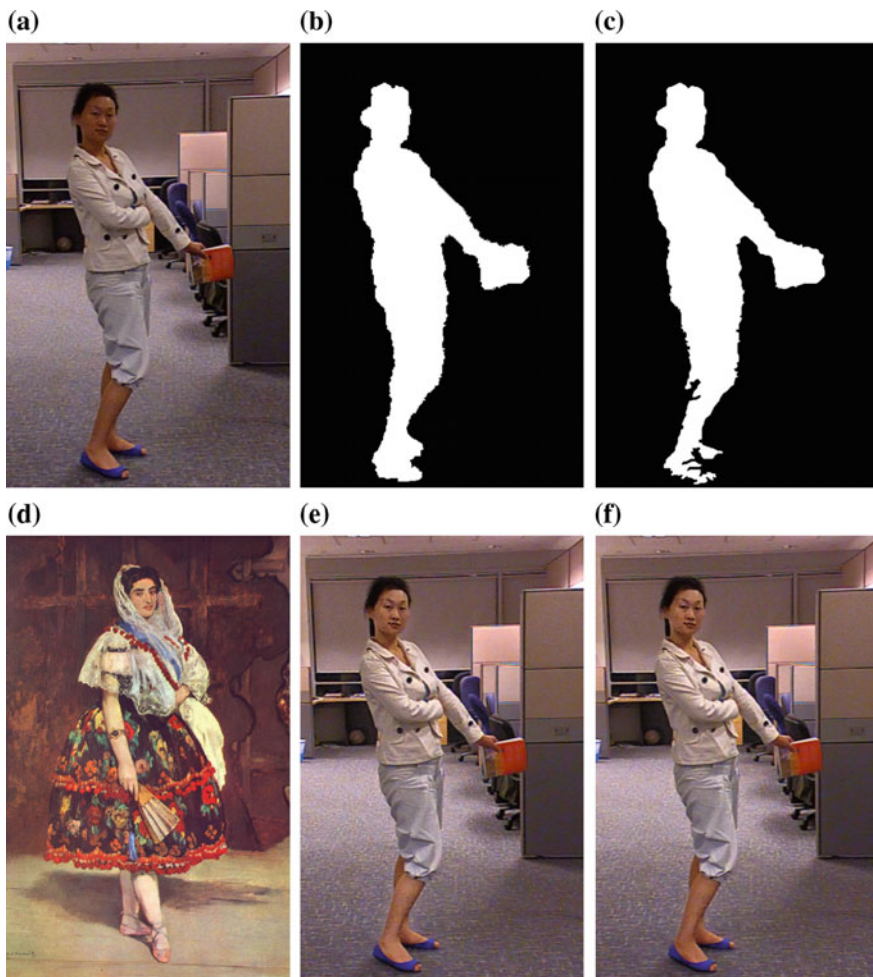


Fig. 7.19 **a** RGB image, same with Fig. 7.18a. **b** Extracted FG mask by removing the ground plane detected by the method in [27]. **c** Extracted FG mask by removing the ground plane detected by the method in [30]. **d** Reference painting (CC Image courtesy Wikimedia Commons). **e** Result using mask in **b**. **f** Result using mask in **c**

as the centre of the detected face area. The plane fitting method based on normal vectors followed by graph-cut segmentation [27] can efficiently detect the ground plane in an RGBD image. The extracted FG of the image in Fig. 7.18a using region grouping after removing the ground plane detected by the method in [27] is shown in Fig. 7.19b. However, the method used in [27] is slower and non-realtime.

Based on the observation that the gradient of the depths on the figure and ground is different in the y direction, as in [30], a filter $F = [1, 1, 1, -1, -1, -1]$ can be applied to delineate the ground plane from the foot of the figure. The thresholded

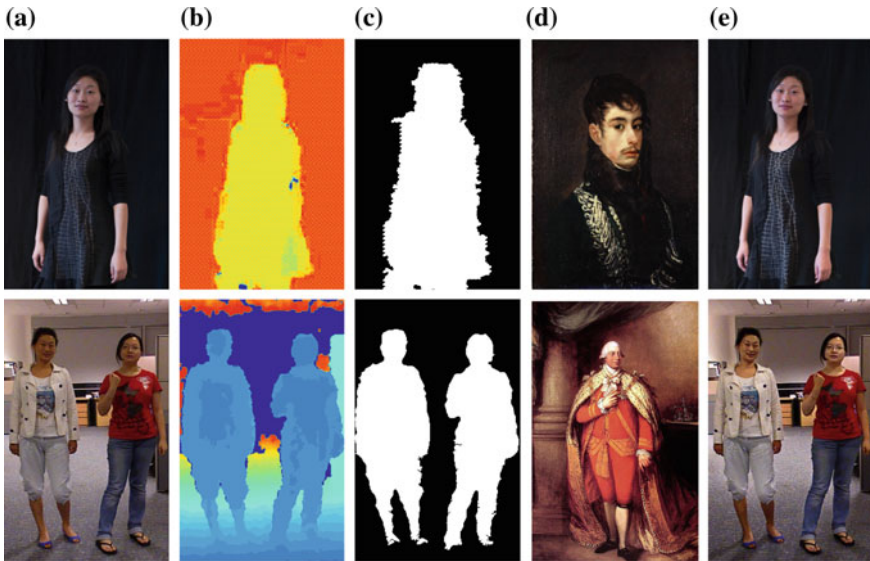


Fig. 7.20 More results for the application in RGBD images. **a** RGB images. **b** Depths. **c** Extracted FG masks. **d** Reference paintings (CC Image courtesy Wikimedia Commons). **e** Enhancement results

filter response corresponds to the ground plane. The extracted FG of the image in Fig. 7.18a using region grouping after removing the detected ground plane is shown in Fig. 7.19c. This method can perform in realtime. However, the detected leg and foot areas are a noisy as compared to those produced using the method in [27]. The enhancement results using the extracted FGs by the two methods are shown in Fig. 7.19e, f. They show that the proposed method is robust. Considering the speed, the F filter proposed in [30] is selected to detect the ground plane for the FG extraction. Results for half-body and two-person portrait photographs are shown in Fig. 7.20. In the second row of Fig. 7.20e, the lower part of the photograph is adjusted so that it is darker: to have a larger contrast against the bright background regions and FG. This contrast organisation in the result is closer to that of the reference painting.

7.6 Concluding Remarks

This chapter proposed a method to improve the visual appeal and focus of attention of portrait photographs by manipulating their regional contrast. This improvement was guided by a selected portrait painting that was used as an example. Experimental results and a user study confirmed that the figure and the face become more salient, and the photograph was more pleasing as a result of using the proposed method. This chapter also discussed the application of the proposed method on RGBD images.

References

1. Arnheim, R.: *Art and Visual Perception: A Psychology of the Creative Eye*, 2nd edn. University of California Press, Berkeley and Los Angeles, CA (1974)
2. Bychkovsky, V., Paris, S., Chan, E., Durand, F.: Learning photographic global tonal adjustment with a database of input/output image pairs. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 97–104 (2011)
3. Cao, Y., Wang, Z.X.: Automatic generation of portrait. In: *Proceedings of International Conference on Audio, Language and Image Processing*, pp. 735–738 (2008)
4. Chen, J., Su, G., He, J., Ben, S.: Face image relighting using locally constrained global optimization. In: *Proceedings of the 11th European Conference on Computer Vision*, pp. 44–57 (2010)
5. Chen, X., Jin, X., Zhao, Q., Wu, H.: Artistic illumination transfer for portraits. *Proc. Eurographics Symp. Render.* **31**, 1425–1434 (2012)
6. Colton, S., Valstar, M.F., Pantic, M.: Emotionally aware automated portrait painting. In: *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts*, pp. 304–311 (2008)
7. DiPaola, S.: Exploring a parameterized portrait painting space. *Int. J. Art Technol.* **2**(1–2), 82–93 (2009)
8. Faizan: face detection using skin color (2010). <http://www.mathworks.com/matlabcentral/fileexchange/26849-face-detection-using-skin-color>
9. Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.* **27**(3), 67:1–67:10 (2008)
10. Grundland, M., Dodgson, N.A.: Automatic contrast enhancement by histogram warping. *Comput. Imaging Vis.* **32**, 293–300 (2004)
11. Hwang, S.J., Kapoor, A., Kang, S.B.: Context-based automatic local image enhancement. In: *Proceedings of the European Conference on Computer Vision*, vol. 1, pp. 569–582 (2012)
12. Jin, X., Zhao, M., Chen, X., Zhao, Q., Zhu, S.C.: Learning artistic lighting template from portrait photographs. In: *Proceedings of the 11th European Conference on Computer Vision*, pp. 101–114 (2010)
13. Joshi, N., Matusik, W., Adelson, E.H., Kriegman, D.J.: Personal photo enhancement using example images. *ACM Trans. Graph.* **29**(2), 12:1–12:15 (2010)
14. Kang, S., Kapoor, A., Lischinski, D.: Personalization of image enhancement. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (San Francisco, CA, USA)*, pp. 1799–1806 (2010)
15. Kaufman, L., Lischinski, D., Werman, M.: Content-aware automatic photo enhancement. *Comput. Graph. Forum* **31**(8), 2528–2540 (2012)
16. Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. *Proceedings of IEEE International Conference on Image Processing* **1**, 900–903 (2002)
17. Liu, C.: Beyond pixels: exploring new representations and applications for motion analysis. Doctoral Thesis, Massachusetts Institute of Technology (2009)
18. Matkovic, K., Neumann, L., Neumann, A., Psik, T., Purgathofer, W.: Global contrast factor - a new approach to image contrast. In: *Proceedings of Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging*, pp. 159–168 (2005)
19. Meng, M., Zhao, M.T., Zhu, S.C.: Artistic paper-cut of human portraits. In: *Proceedings of ACM Multimedia 2010 International Conference*, pp. 931–934 (2010)
20. Microsoft: Kinect for Windows. <http://www.microsoft.com/en-us/kinectforwindows/>
21. Perkins, M.: *Professional Portrait Lighting: Techniques and Images from Master Photographers*. Amherst Media, Inc. (2006)
22. Pulli, K., Pietikainen, M.: Range image segmentation based on decomposition of surface normals. In: *Proceedings of 8th Scandinavian Conference on Image Analysis*, vol. 2, pp. 893–899 (1993)
23. Raytrix: Raytrix. <http://www.raytrix.de/>

24. Rivera, A.R., Ryu, B., Chae, O.: Content-aware dark image enhancement through channel division. *IEEE Trans. Image Process.* **21**(9), 3967–3980 (2012)
25. Rother, C., Kolmogorov, V., Blake, A.: "grabcut" - interactive foreground extraction using iterated graph cuts. *Proc. of ACM SIGGRAPH* **23**, 309–314 (2004)
26. Sarfraz, M., Al-Mulhem, M., Ashraf, F.: Preserving monotonic shape of the data using piecewise rational cubic functions. *Comput. Graph.* **21**(1), 5–14 (1997)
27. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from rgbd images. In: *Proceedings of the 12th European Conference on Computer Vision*, pp. 746–760 (2012)
28. Viola, P., Jones, M.J.: Rapid object detection using a boosted cascade of simple features. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I-511–I-518 (2001)
29. Wang, B., Yu, Y., Xu, Y.Q.: Example-based image color and tone style enhancement. *ACM Trans. Graph.* **30**(4), 64:1–64:12 (2011)
30. Xia, L., Chen, C.C., Aggarwal, J.K.: Human detection using depth information by kinect. In: *Proceedings of Workshop on Human Activity Understanding from 3D Data in conjunction with CVPR*, pp. 15–22 (2011)
31. Yang, M., Lin, S., Luo, P., Lin, L., Chao, H.: Semantics-driven portrait cartoon stylization. In: *Proceedings of IEEE 17th International Conference on Image Processing*, pp. 1805–1808 (2010)
32. Zhang, X., Chan, K.L., Constable, M.: Depth-based reference portrait painting selection for example-based rendering. In: *Proceedings of IEEE International Conference on Cybernetics and Intelligent Systems*, pp. 105–110 (2011)
33. Zhang, X., Constable, M., Chan, K.L.: Example-based contrast enhancement for portrait photograph. In: *Proceedings of International Conference on Pattern Recognition*, pp. 943–946 (2012)
34. Zhang, Y., Gao, L., Zhang, S.Y.: Feature-based automatic portrait generation system. In: *Proceedings of World Congress on Computer Science and Information Engineering*, pp. 6–10 (2009)

Chapter 8

Composition Improvement for Portrait Photographs

Abstract Composition has a great impact upon the visual quality of a photograph. This chapter studies the composition in portrait paintings and proposes an algorithm to improve the composition of portrait photographs based on an example portrait painting. From a study of portrait painting, it can be shown that the placement of the face and the figure in portrait paintings is pose-related. Based on this observation, our algorithm improves the composition of a portrait photograph by referencing the placement of the face and the figure from an example portrait painting. The example portrait painting is selected based on the similarity of its figure pose to that of the input photograph. This similarity measure is modelled as a graph matching problem. Finally, space cropping is performed using an optimisation function to assign a similar location for each body part of the figure in the photograph with that in the example portrait painting. The experimental results demonstrate the effectiveness of the proposed method. A user study shows that the proposed pose-based composition improvement is preferred more than the rule-based methods.

Keywords Composition improvement · Pose · Portrait

8.1 Introduction

In portrait paintings, where the placement of the figure and the face are the most important elements of the composition [9], the size and the position of the figure and the face are specially organised by artists to draw the attention of the viewer to those regions.

In contrast, in snap-shot style photographs by amateurs, these same properties are frequently not well-organised. At a basic level, selection based on composition is done as a matter of course. When selecting personal photographs for sharing in social networks or for printing for the purpose of home display, we often discard photographs in which the face is too small or where the subject occupies or intersects the corners or edges of the image. This we might do despite the expression and lighting of the face being otherwise pleasing. In such a situation, the composition can be improved before use.

Motivated by this, we propose to improve the composition of portrait photographs by optimising the size and placement of the figure in the image. From a study of portrait painting, it is observed that the placement of the face and the figure in a portrait painting is pose-related. Based on this observation, an algorithm is developed to improve the composition of a portrait photograph by using a portrait painting as an exemplar. This painting is selected based on the similarity of its pose to that of the input photograph, the similarity measure being modelled as a graph matching problem. Finally, space cropping is performed using an optimisation function to assign a similar location for each body part of the figure in the photograph to that of the corresponding body part in the example portrait painting. The experimental results demonstrate the effectiveness of the proposed method. A user study shows that the proposed pose-based composition improvement is preferred over existing rule-based methods.

8.2 Related Work

Composition in aesthetics measure. Image composition has an important role in image aesthetics. Almost all methods that quantify the aesthetics of images consider composition. This may involve the so-called rule of thirds [2, 5, 6, 13, 18] or the golden ratio [3, 14, 16]. These are described in Sect. 4.1.1 p. 41. In addition to this, the aesthetic measure of portraiture is also assessed with respect to the pose of the face [14], the ratio of the face area to the image size [13, 19], and the ratio of the body length to the face length [10]. All this work is in general agreement with the idea that the size and placement of the face are important for the aesthetic evaluation of portrait images. However, these methods only measure the location of the face based on rules, without consideration of the body pose.

Composition optimisation. There are two aspects of composition optimisation. One is to place the figure agreeably within the borders of the image, the other is to change the format of the image to suite the content (i.e. the proportion of the vertical of the image to its horizontal). Cropping can perform both these tasks. However, cropping an image may remove important parts of it. To address this, content-aware retargeting techniques such as seam carving [1] and image warping [24] are proposed to adjust the relative position of important subjects in the image. However, these methods may easily create artefacts, and cropping remains the favoured photographer's approach [20].

Zhang et al. [25] presented an automatic photograph-cropping algorithm by formulating auto-cropping as an optimisation problem. The energy function was defined based on the constraints on the face location, the centre of the region of interest (ROI) and the cropped areas of the face and ROI. The placement of the face was based on templates classified according to facial information. Santella et al. [21] proposed an interactive method for image cropping by using eye tracking to identify important content. An optimisation framework that facilitated composition rules (including the subject, avoiding cutting through background objects, maximising content areas and placing content based on the rule of thirds) was used to calculate the best crop. The

rule of thirds, diagonal dominance, visual balance and salient-region size were used to optimise the composition of photographs [17]. Li et al. used the composition method in [17] to enhance the aesthetics of photographs with a face [15]. The above-reviewed cropping methods are all rule-based. Differently, She et al. [22] and Zhang et al. [26] proposed to use learning-based methods to perform cropping. These learning-based methods are suitable for images that feature many objects, using which a relationship graph may be built. They do not focus on portrait photographs. According to the knowledge of the authors, no published work has used an example-based method to achieve cropping for the purpose of the composition improvement of portraits.

Human pose retrieval. Ferrari et al. originally proposed a human pose retrieval method based on pose estimation to retrieve people using their pose from images or videos [8]. The full version was further published in [7]. Human pose estimators were used for the pose estimation. It was capable of estimating upper-body pose and full-body pose. In order to improve the pose estimation performance, body part detectors were used for the body pose estimation and further applied for pose retrieval [11]. Although it was effective, only upper-body pose was estimated. In our work, the full-body pose needs to be estimated. Therefore, we develop our pose similarity measure method based on the work in [8].

8.3 Proposed Composition Improvement Method

The observation of composition in portrait paintings provides the justification for the improvement of the composition of portrait photographs based on the pose of

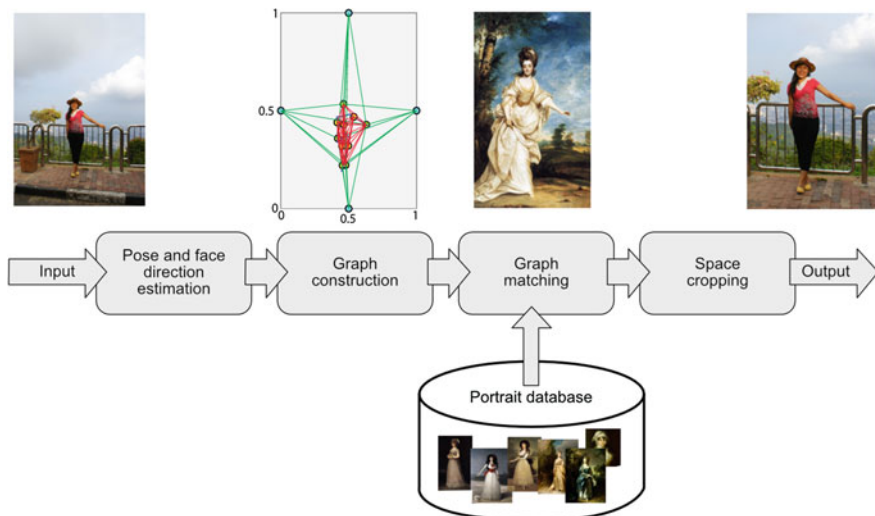
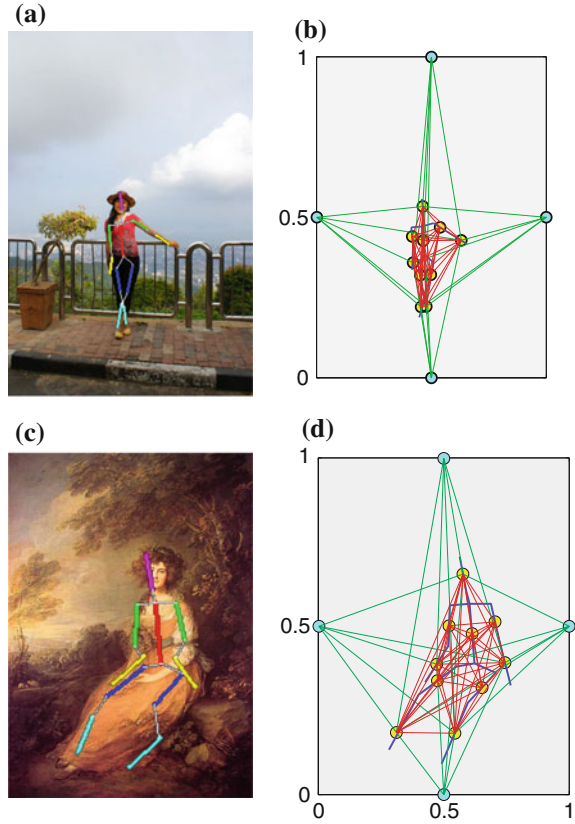


Fig. 8.1 The framework of the proposed method

Fig. 8.2 **a** A photograph with extracted pose. **b** The graph for the photograph **a**. **c** A painting with extracted pose (CC Image courtesy The Athenaeum). **d** The graph for the painting in **c**. The *yellow points* are the body nodes, which express the body parts. The *green points* are the boundary nodes, the *blue lines* are the skeleton of the body, the *red lines* are the edges linking the body nodes and the *green lines* are the edges linking the head, lower arm and lower leg nodes with the boundary nodes (color figure online)



the figure. The framework of the proposed method is shown in Fig. 8.1. An example painting is selected, the pose of which is similar to that of the photograph. Here, pose represents the body pose together with the face tilt direction and is extracted using the state-of-the-art pose estimation method proposed in [8]. The face rotation angle is extracted separately as the face direction by the method proposed in [27]. In addition to the pose and face evaluation, spaces around the figure were also considered. In artist's vernacular, these are known as negative spaces [23] and are considered in order to ensure that the photograph can be cropped in such a manner as that it matches the example painting.

Because the position of the pose and negative spaces can be described in relative terms, a graph $G = (V, E)$ can be used to model these relationships, where V is the node of the graph and E is the edge that links the connected nodes. There are two types of nodes in the graph: the body node which represents specific body parts and the boundary node distributed on the boundaries of the image. The descriptors of the body node are the position and direction of the corresponding body part. Specifically, the face node is also described by the face direction. The 4 boundary

nodes are described by the location of the boundary centre. The weight of the edge linking the body nodes is defined by the relative location and orientation of the corresponding body parts. The weight of the edge linking the body node with the boundary node is defined by their distance from each other. For simplification, only the head, lower arm and lower leg nodes are linked with the boundary nodes. Two example graphs are shown in Fig. 8.2.

After the graph is constructed, the example painting is selected by searching the painting database based on a similarity measure for graph matching. Then, space cropping is performed by using an optimising function.

8.3.1 Graph Matching

After constructing the graph to describe the pose and spaces around the figure, the example painting is selected by finding a painting that has a graph similar to that of the input photograph. Therefore, the example painting selection can be achieved by graph matching based on a similarity measure. Given the input graph G_I and the graph of the k -th painting G_k in a database, the graph matching is formulated as

$$\max_k S(G_I, G_k) \tag{8.1}$$

While the correspondences of nodes and edges of two graphs are fixed, the similarity of two graphs can be written as

$$S(G_I, G_k) = \sum_i S_i^v(G_I, G_k) + \sum_j S_j^e(G_I, G_k) \tag{8.2}$$

where, $i = 1, \dots, N$, N is the number of nodes, and $j = 1, \dots, M$, M is the number of edges. S^v is the similarity of the nodes and S^e is the similarity of the edges. The definition of similarities of edges and nodes is based on the definition of the descriptors.

In the pose estimation by [8], the posterior marginal distribution of the position of each body part was estimated. Based on the posterior marginal distribution, the position and orientation of each body node were calculated. For the body node v_i^b , the position descriptor was the posterior marginal distribution $P_{pos}^{v_i^b} = P(l_i = (x, y, \theta))$, where l_i was the body part i , (x, y) was the scale-normalised location, and θ was the orientation. The orientation descriptor of v_i^b was measured by $P_{ori}^{v_i^b} = \sum_{(x,y)} P(l_i = (x, y, \theta))$.

The descriptors of the edge e^b linking two body nodes were the relative *location* and relative *orientation* of the two corresponding body parts. The relative location descriptor of the edge e_j^b which links the body parts j_1 and j_2 was $P_{pos}^{e_j^b} = P(l_{j_1}^{xy} - l_{j_2}^{xy} = \delta)$. The relative orientation descriptor of the edge e_j^b was

$P_{ori}^{e^b} = P(r(I_{j_1}^\theta - I_{j_2}^\theta) = \rho)$, where $r(\cdot)$ was a circular difference operator. Details of the calculation of the descriptors could be found in [8].

The face direction was described by 13 quantised degrees from 90° to -90° . Positive degrees are for faces facing right, and negative degrees are for faces facing left [27]. The locations of the four boundary nodes v^d were (0.5, 0), (1, 0.5), (0.5, 1) and (0, 0.5). The descriptor of the edge e^d that links the body node with the vertical boundary was their distance in the x direction. The descriptor of the edge that links the body node with the horizontal boundary node was their distance in the y direction.

As in [8], the combined Bhattacharyya similarity $s(a, b) = \sum_j \sqrt{a(j) \cdot b(j)}$ was used to calculate the similarities of the position and orientation descriptors of the body nodes v^b and the similarities of the relative location and relative orientation descriptors of edges e^b . The similarity of the two face directions θ_1 and θ_2 was measured as

$$S_F(\theta_1, \theta_2) = e^{-\left(\frac{\theta_1 - \theta_2}{\sigma}\right)^2} \quad (8.3)$$

where σ controlled the sensitivity of the similarity to the difference of face directions.

The locations of the boundary nodes are the same for all the graphs, therefore it is not necessary to calculate the similarity of corresponding boundary nodes. The constraint on the distance of the boundary node with the body node is to ensure that there is enough space in the photograph for cropping in order to match the composition of the example painting. If the distance in the photograph is greater than the corresponding distance in the painting, the similarity would be the maximum value; otherwise, the similarity decreases in proportion to the increase of their difference. Based on this definition, the similarity of the q -th e^d edge from the input graph with the corresponding e^d edge from one painting graph with descriptors d_l and d_k was

$$S_q^{e^d} = \begin{cases} 1 & \text{if } d_l - d_k \geq 0 \\ e^{-\left(\frac{d_l - d_k}{\sigma_d}\right)^2} & \text{if } d_l - d_k < 0 \end{cases} \quad (8.4)$$

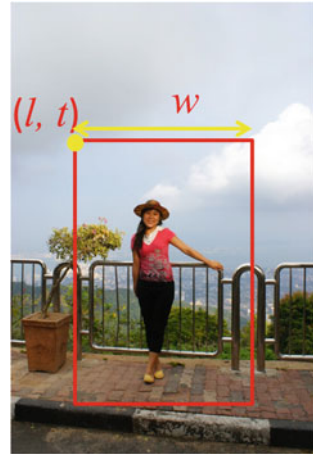
where σ_d was the standard variance.

After calculating the similarity of body nodes S^{v^b} , the similarity of face directions S_F , the similarity of edges linking body nodes S^{e^b} and the similarity of edges linking body nodes with boundary nodes S^{e^d} , the final similarity of the two graphs was the linear combination of these similarities, which was expressed as

$$S(G_I, G_k) = w_1 \sum_n S_n^{v^b} + w_2 \cdot S_F + w_3 \sum_p S_p^{e^b} + w_4 \sum_q S_q^{e^d} \quad (8.5)$$

where $n = 1, \dots, N^b$, N^b was the number of body nodes. $p = 1, \dots, M^b$, M^b was the number of edges linking body nodes. $q = 1, \dots, M^d$, M^d was the number of edges linking body nodes with boundary nodes. The painting that had the highest similarity score to that of the input photograph was selected as the exemplar: to guide the cropping of the input photograph.

Fig. 8.3 Space cropping to find a cropping window (l, t, w)



8.3.2 Space Cropping

This section presents an algorithm to improve the composition of the photograph based on the selected example painting. Space cropping was used to assign a similar location for each body part of the photograph and to match the location of equivalent body parts in an example painting, by defining a suitable cropping window (see Fig. 8.3). Because the size and aspect ratio of the photograph was different to that of the example painting, it could not be directly cropped by matching its negative space to that of the example painting. Hence, the space cropping was formulated as an optimisation problem that considered the location of each body part. The energy function for the space cropping was defined as

$$E = \alpha E_{pos} + (1 - \alpha) E_{face} \tag{8.6}$$

The first term E_{pos} was the pose constraint. Given the target location v_i for the body part i and the reference location v_i^r which was from the example painting, E_{pos} was defined as

$$E_{pos} = \frac{1}{N^b} \sum_i^{N^b} \|v_i - v_i^r\| \tag{8.7}$$

Due to the importance of the face in a portrait, a face constraint E_{face} was added to the energy function to prevent it from shifting too much from its location as defined by the painting. Given the target location v_{face} of the face and the reference location v_{face}^r which was from the example painting, E_{face} was defined as

$$E_{face} = \|v_{face} - v_{face}^r\| \tag{8.8}$$

The pose constraint and the face constraint was linearly combined by a constant scale α .

The cropped rectangle was represented by the left top corner coordinates (l, t) , width (w) and height (h) . Given a specified aspect ratio a , the height could be calculated as $h = a \times w$. Therefore, the optimisation of the energy function could be reformulated as finding a vector of (l, t, w) . The particle swarm optimisation (PSO) [12] method was used to seek the optimal solution by globally searching the minimum candidate of Eq. (8.6).

8.4 Experiments

For the graph matching, the parameters used in our implementation were determined empirically as $\sigma = 45$, $\sigma_d = 0.3$, $\omega_1 = 0.2$, $\omega_2 = 0.2$, $\omega_3 = 0.4$ and $\omega_4 = 0.2$. The weight $\alpha = 0.4$ was used in the energy function (8.6) for space cropping.

The 300 portrait paintings collected in Chap. 7 were used to form the painting database. These consist of classic examples from some of art history's most accomplished portrait painters. The database covers a variety of pose and environments; 180 of them are half-body portrait paintings and 120 of them are full-body portrait paintings. Based on the size of face in the image, the half-body portrait paintings are classified as being 'small face half-body portrait' and 'large face half-body portrait' (face length/image height is greater than 0.2). These images are in height greater than 500 pixels. Based on the study of aspect ratios of portrait paintings, the aspect ratios for full-body portraits and small face half-body portraits was defined as $a = 1.5$, and for large face half-body portraits $a = \frac{4}{3}$. The test portrait photographs were collected from the MIT-Adobe FiveK Dataset [4] and our personal collections.

Performance. Figure 8.4 shows the cropping results of some photographs with figures in different poses. The pose of the figure in the selected example painting in Fig. 8.4 is similar to that of the corresponding input photograph. From composition improvement results in Fig. 8.4, we can see that the location of the figure in the result is similar to that of the corresponding example painting. This shows the effectiveness of the proposed method. Additionally, we can see that the locations of the figures in the two composition improvement results are very different and serves the requirement of their poses.

For a full-body portrait photograph, the proposed method can down-crop to a half-body portrait with small face, or a half-body portrait with large face (see the example shown in Fig. 8.5). In the half-body portrait painting with large face, generally only the upper body with upper arms are visible. The extracted poses for lower arms are not meaningful. Therefore, in the selection of the example large face half-body portrait painting, the lower arms are not considered. For a small face half-body portrait photograph, the proposed method can also down-crop to a small face half-body or large face half-body portrait.

Comparison with rule-based methods. To evaluate the effectiveness of the proposed composition improvement method based on pose, the results were compared with those achieved using two rule-based automatic composition improvement

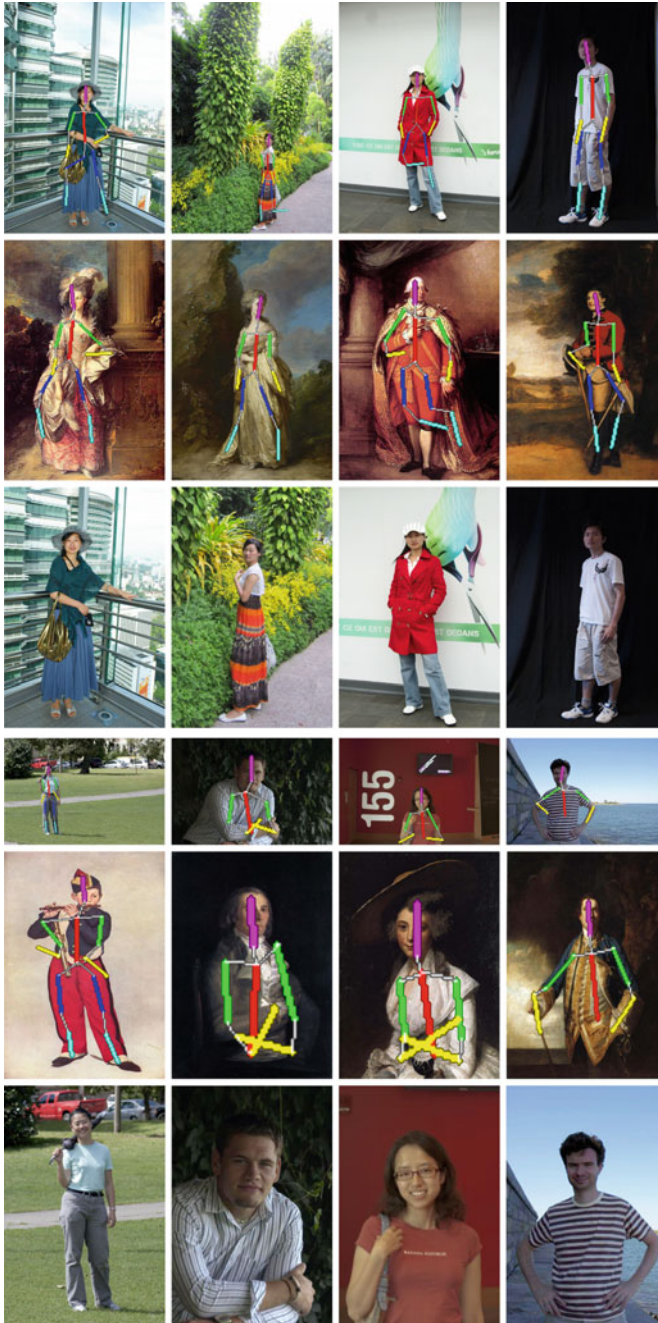


Fig. 8.4 Experimental results. First and fourth rows: input photographs with pose. Second and fifth rows: selected example portrait paintings with pose (CC Image courtesy The Athenaeum, Wikimedia Commons, <https://www.wikiart.org>, <http://www.artelista.com/>). Third and sixth rows: cropping results using the proposed method



Fig. 8.5 a A photograph with extracted pose. b Selected full-body painting with pose (CC Image courtesy of <https://www.wikiart.org>). c Selected small face half-body painting with extracted pose (CC Image courtesy of Wikimedia Commons). d Selected large face half-body painting with extracted pose (CC Image courtesy of <https://www.wikiart.org>). e–g The cropping results using the three example paintings

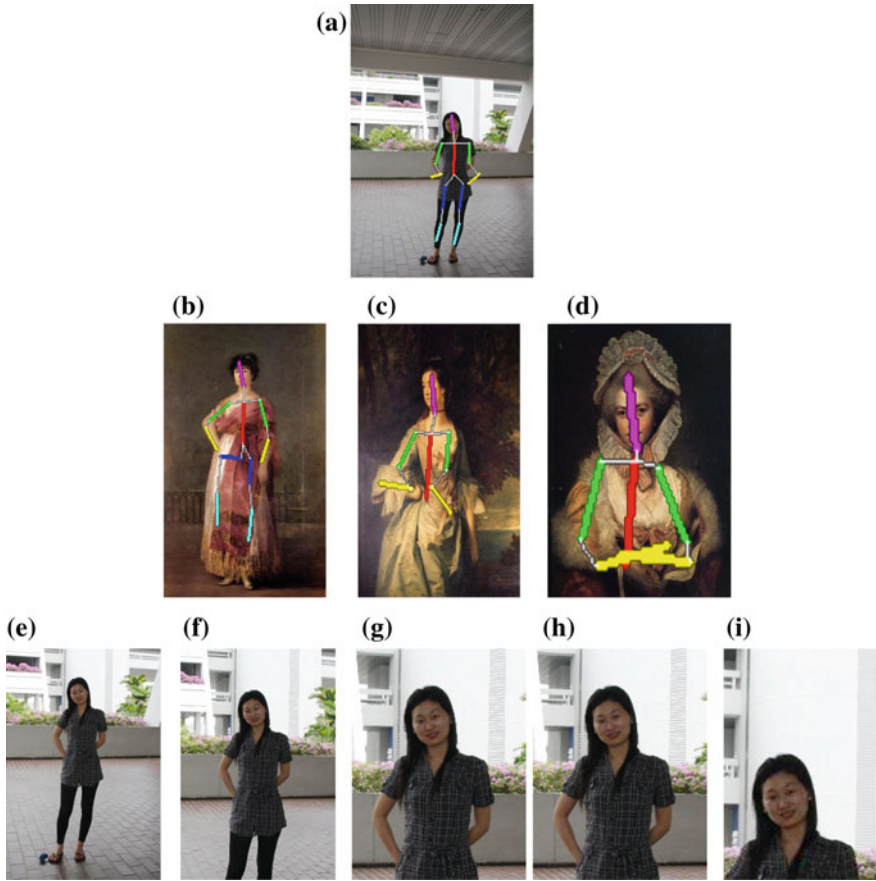


Fig. 8.6 a Input photograph with extracted pose. b–d Example paintings with extracted pose (CC Image courtesy of Wikimedia Commons). e Space cropping result (full-body) achieved using the proposed method based on example painting b. f Space cropping result (small face half-body) achieved using the proposed method based on example painting b. g Space cropping result (large face half-body) achieved using the proposed method based on example painting in d. h The result achieved using the ACDP method in [25]. i The result achieved using the ABPE method in [15]

methods. These two methods were designed for photographs which feature a face. One method was the auto-cropping based on templates in [25] (ACDP method). The definition of the template was rule-based. The other method was the aesthetic-based photo-editing approach proposed in [15] (ABPE method). The composition improvement results for a photograph are shown in Fig. 8.6. The result in Fig. 8.6h is produced by the ACDP method using the small face template. The selection of the template is based on the size of the face in the original photograph. In the ACDP method, the region of interest (ROI) is used to constrain the cropping in order to prevent the photograph from being cropped too aggressively. The ROI is detected by

an attention model, which is one kind of saliency detection. The detected ROI may lose some important body parts. In Fig. 8.6h, the feet are cropped-out which leaves the figure unbalanced in the image. The ABPE method [15] only considers the face region as important information; hence, in the result, only the face and shoulders are visible while other parts are all cropped-out (see the result in Fig. 8.6i). In the result shown in Fig. 8.6i, the face is placed at the bottom one-third point. Although this placement of the face is better than its original placement, it is far from being visually agreeable. Results achieved using the proposed method are more encouraging compared with those achieved using the rule-based methods.

Comparison with a method that employs statistical models. Results achieved using the proposed method were compared with those achieved using cropping based on statistics. From our study of portrait paintings, it was observed that more space was generally given to the left side of the figure in paintings in which the head was facing left, while more space was given to the right side of the figure in which the head was facing right. Inspired by this, a study was conducted to explore the space arrangement in portrait paintings by dividing full-body portrait paintings, small face half-body portrait paintings and large face half-body portrait paintings into left-directed and right-directed classes based on the facing direction of the head.

The spaces around the figure (see the Fig. 8.7a), the face location, the size of the face and the centre of the body were measured in each of the six classes. Then, Gaussian mixture models (GMM) were used to fit the distributions of these features. The GMM models of locations of the faces (measured by the centres of the face windows) and the centres of the bodies in the full-body portrait paintings are shown in Fig. 8.7b, c. Based on these statistical models, an energy function was formulated

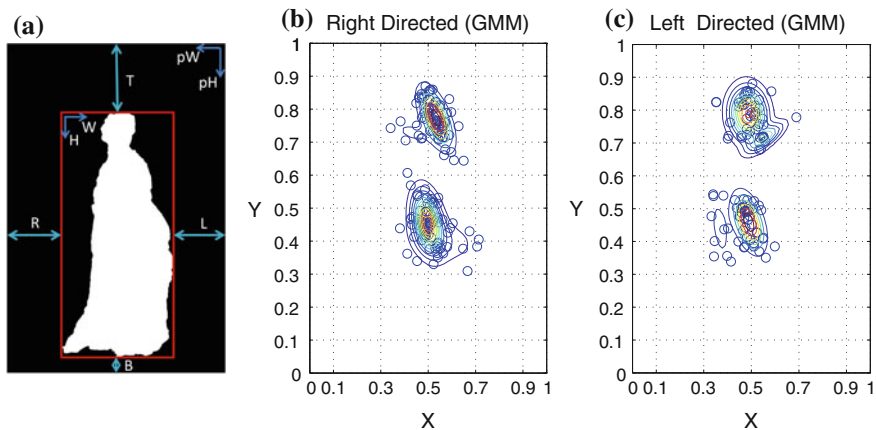


Fig. 8.7 **a** The spaces around the figure are: L (*left space* of the figure), R (*right space* of the figure), T (*top space*), B (*bottom space*). H and W are the height and width of the figure, respectively, and pH and pW are the height and width of the image, respectively. **b** GMM models of centres of faces and centres of bodies of right-directed full-body portrait paintings. **c** GMM models of centres of faces and centres of bodies of left-directed full-body portrait paintings

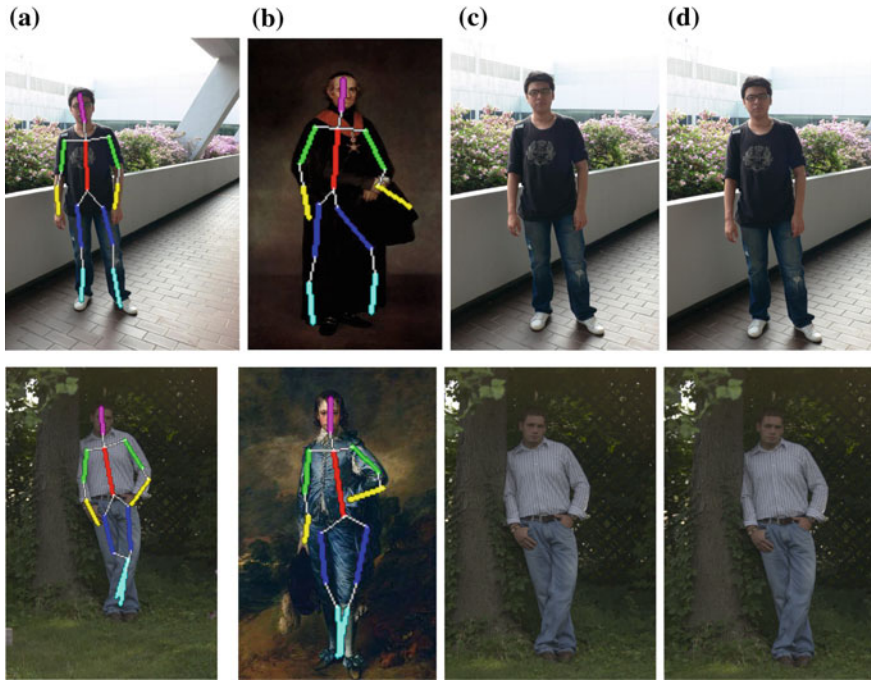


Fig. 8.8 Comparison with results using statistical models. **a** Input photographs with poses. **b** Selected example paintings with poses (CC Image courtesy of Wikimedia Commons). **c** Cropping results by the proposed method. **d** Results using statistical models

to crop the photographs. This energy function was

$$E = E_{TB} + E_{LR} + E_F + E_C + E_{FS} \tag{8.9}$$

where, E_{TB} was the 2D GMM model of the top and bottom spaces, E_{LR} was the 2D GMM model of the left and right spaces, E_F was the GMM model of the face locations, E_C was the GMM model of the centres of bodies and E_{FS} was the one-dimensional GMM model of the face size. The PSO method was used to maximise the energy function for finding an optimising cropping window. This method is based on statistical models and creates baseline results (see the two examples in Fig. 8.8d). However, by using statistical models, the face and the body are more likely to be placed around the vertical central line. Differently, the proposed example-based method can produce results with faces and bodies placed in various locations. In the top example in Fig. 8.8, the cropping result using a statistical model is similar to the result achieved using the proposed method. However, in the bottom example in Fig. 8.8, the face is more biased from the central line in the result achieved using the proposed method as compared with that achieved using a statistical model. The face bias in this example is more visually pleasing. In addition, statistical models require

Table 8.1 User study result for the comparison of the proposed method with the ACDP [25] and ABPE [15] methods and with the method using statistical models

Ours versus ACDP	Ours	ACDP	Same
	68.3%	22.2%	9.5%
Ours versus ABPE	Ours	ABPE	Same
	81.4%	16.4%	2.2%
Ours versus Statistical Method	Ours	Statistical Method	Same
	59.7%	30.0%	10.3%

that the figure is first segmented in order that the spaces around it be measured. The proposed example-based method can produce results automatically without segmentation.

User Study. In order to significantly evaluate the effectiveness and advantage of the proposed approach, a user study was conducted to compare composition improvements achieved using our proposed method with those achieved using the ACDP [25], ABPE [15] and the statistical models; 20 of the test photographs were randomly selected for this user study. We used Google Forms to create the webpage for the user study. The participants connected to the webpage using their own computers and monitors.

The results obtained by our proposed method and one of other methods were shown as a pair side-by-side. The left/right ordering of the images in each pair was randomly generated. At the starting page of the user study, a simple introduction to composition in portraiture was given. Then, for each image pair, the participants were asked to answer a question ‘Which is better composed and more suitable for printing for home display or for making a photograph album?’. The participants could choose ‘Left’, ‘Right’ or ‘They look the same’ as a response according to their first impression. There were a total of 18 participants aged from 20 to 35. Eleven of them were male and 7 of them were female. For the comparison of our proposed method with each of the three other methods, 360 responses were collected. The result of the user study is summarised in Table 8.1. It shows that our results are preferred by a significantly higher percentage of respondees to those achieved using other methods. The positive responses to our results are 4 times greater than those achieved using ABPE method. The fewer positive responses to the ABPE method are due in part to the fact that it only considers the face region as being important information and forces the face to the one-third point without considering the other body parts. In some cases, the face is placed on the bottom one-third point (as the example in Fig. 8.6i). This results in the figure only taking up a small proportion of the total image area, the remainder consisting of background. This comparison between our proposed method and the ABPE method demonstrates that it is important to consider the face and body together in composition improvement. It is not surprising that our results are preferred more in comparison with results achieved using the ACDP method, because ACDP is easily disturbed by background objects. The method using



Fig. 8.9 **a** A painting without accurate post-extraction by the use of the original pose estimation method [8]. CC Image courtesy The Athenaeum. The pose extraction in **b** is conducted on a manually segmented foreground area

statistical models gets a higher percentage of positive responses than the ACDP and ABPE methods in comparison with the proposed method and could produce effective baseline results; however, it is not as good as the proposed example-based method in the processing of images with diverse scenes.

Limitations. The experiments and user study demonstrated the effectiveness of the proposed approach. However, the proposed method relies on the accuracy of the pose estimation. Using the method in [8], this accuracy is influenced by the background colour. This is because the pose estimation is conducted in the foreground area which is segmented by using the GrabCut method based on colour. Some images do not return accurate pose extraction results (see Fig. 8.9a). In the image in Fig. 8.9a, the colour of the background is very similar to that of the clothes of the person, and GrabCut's estimation of the foreground area spreads into a large area of the background. However, if a manual or interactive segmentation method is used to segment the foreground area accurately, the accuracy of pose estimation is increased greatly. The extracted pose in Fig. 8.9b is based on a manually segmented foreground region. It is much more accurate than that shown in Fig. 8.9a.

8.5 Concluding Remarks

This chapter proposed a method to improve the composition of a portrait photograph based on an example portrait painting. The example painting was selected based on the pose, the face direction of the figure and the spaces around the figure. A graph model was constructed for the example painting selection. Space cropping was used to improve the composition of the input photograph based on the selected example painting. The space cropping was formulated as an optimisation problem. The experimental results and a user study showed that the proposed method performs better than rule-based methods. However, the proposed method relies on the accuracy of the pose estimation. There is still room for improving the pose estimation in complex scenes, but it is beyond the scope of this chapter.

References

1. Avidan, S., Shamir, A.: Seam carving for content-aware image resizing. *ACM Trans. Graph.* **26**(3), 10:1–10:9 (2007)
2. Battiato, S., Moltisanti, M., Ravi, F., Bruna, A.R., Naccari, F.: Aesthetic scoring of digital portraits for consumer applications. In: *Proceedings of SPIE, Digital Photography IX*, vol. 8660, pp. 7:1–7:10 (2013)
3. Bhattacharya, S., Sukthankar, R., Shah, M.: A framework for photo-quality assessment and enhancement based on visual aesthetics. In: *Proceedings of International Conference on Multimedia*, pp. 271–280 (2010)
4. Bychkovsky, V., Paris, S., Chan, E., Durand, F.: Learning photographic global tonal adjustment with a database of input/output image pairs. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 97–104 (2011)
5. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: *Proceedings of European Conference on Computer Vision*, pp. 288–301 (2006)
6. Dhar, S., Ordonez, V., Berg, T.L.: High level describable attributes for predicting aesthetics and interestingness. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1657–1664 (2011)
7. Eichner, M., Marin-Jimenez, M., Zisserman, A., Ferrari, V.: 2D Articulated human pose estimation and retrieval in (almost) unconstrained still images. *Int. J. Comput. Vis.* **99**, 190–214 (2012)
8. Ferrari, V., Marin-Jimenez, M., Zisserman, A.: Pose search: Retrieving people using their pose. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2009)
9. Hess, A.: *Borders Composition Digital Field Guide*. Wiley Publishing, New York (2010)
10. Huang, L., Xia, T., Wan, J., Zhang, Y., Lin, S.: Personalized portraits ranking. In: *Proceedings of the 19th ACM International Conference on Multimedia*, pp. 1277–1280 (2011)
11. Jammalamadaka, N., Zisserman, A., Jawahar, C.V.: Human pose search using deep poselets. In: *Proceedings of International Conference on Automatic Face and Gesture Recognition*. (2015)
12. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948 (1995)
13. Khan, S.S., Vogel, D.: Evaluating visual aesthetics in photographic portraiture. In: *Proceedings of the Eighth Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, pp. 55–62 (2012)

14. Li, C., Gallagher, A.C., Loui, A.C., Chen, T.: Aesthetic quality assessment of consumer photos with faces. In: Proceedings of International Conference on Image Processing, pp. 3221–3224 (2010)
15. Li, C., Loui, A.C., Chen, T.: Towards aesthetics: a photo quality assessment and photo selection system. In: Proceedings of the ACM international conference on Multimedia, pp. 827–830 (2010)
16. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. *IEEE J. Sel. Top. Signal Process.* **3**(2), 236–252 (2009)
17. Liu, L., Chen, R., Wolf, L., Cohen-Or, D.: Optimizing photo composition. *Comput. Graph. Forum* **29**, 469–478 (2010)
18. Luo, Y., Tang, X.: Photo and video quality evaluation: focusing on the subject. In: Proceedings of European Conference on Computer Vision, pp. 386–399 (2008)
19. Luo, W., Wang, X., Tang, X.: Content-based photo quality assessment. In: Proceedings of International Conference on Computer Vision, pp. 2206–2213 (2011)
20. Rubinstein, M., Gutierrez, D., Sorkine, O., Shamir, A.: A comparative study of image retargeting. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* **29**(6), 160:1–160:9 (2010)
21. Santella, A., Agrawala, M., DeCarlo, D., Salesin, D., Cohen, M.: Gaze-based interaction for semi automatic photo cropping. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 771–780 (2006)
22. She, J., Wang, D., Song, M.: Automatic image cropping using sparse coding. In: Proceedings of the First Asian Conference on Pattern Recognition, pp. 490–494 (2011)
23. Thomas, J.: *The Art of Portrait Drawing: Learn the Essential Techniques of the Masters*. North Light Books (2006)
24. Wang, Y.S., Tai, C.L., Sorkine, O., Lee, T.Y.: Optimized scale-and-stretch for image resizing. *ACM Trans. Graph.* **27**(5), 118:1–118:8 (2008)
25. Zhang, M., Zhang, L., Sun, Y., Feng, L., Ma, W.: Auto cropping for digital photographs. In: Proceedings of the IEEE Conference on Multimedia and Expo, pp. 438–441 (2005)
26. Zhang, L., Song, M., Zhao, Q., Liu, X., Bu, J., Chen, C.: Probabilistic graphlet transfer for photo cropping. *IEEE Trans. Image Process.* **22**, 802–815 (2012)
27. Zhu, X., Ramanan, D.: Face detection, pose estimation and landmark estimation in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2879–2886 (2012)

Chapter 9

Vignetting Effect Transfer

Abstract This chapter discusses how the vignetting effect of paintings may be transferred to photographs, with attention to centre-corner contrast. First, the lightness distribution of both is analysed. The results show that the painter's vignette is more complex than that achieved using common digital post-processing methods. Specifically, it is shown to involve both the 2D geometry and 3D geometry of the scene. An algorithm is then developed to extract the lightness weighting from an example painting and transfer it to a photograph. Experiments show that the proposed algorithm can successfully perform this function. The resulting vignetting effect is more naturally presented with regard to aesthetic composition as compared with vignetting achieved with popular software tools and camera models.

Keywords Vignetting · Center-corner contrast · 2D geometry · 3D geometry

9.1 Introduction

As introduced in Sect. 4.2 on p. 46, the effect of a vignette is to darken the outside of the image and lighten the centre. The purposely introduced vignette serves to draw the attention of the viewer to the middle of the image, and to emulate the fade-off of natural human vision towards the periphery. In photographs, the effect can be created by changing the camera settings or by post-processing. This photograph-style vignetting darkens the image towards its corners, with a clear transition between the dark periphery and the light middle. Importantly, the effect is entirely 2D across the flatness of the picture plane.

Artists also employ vignetting in paintings by organising centre-corner contrast. The vignetting effect in painting is also known as the *vista* (lat. view). In portraits, it draws attention of the viewer to the subject (Fig. 9.1a), while in landscapes, it guides their eye to the far distance (Fig. 9.1b). Different to the camera vignette, the painter's vignette is integrated into the 3D form of the painting, and does not leave a notably visible transition between the dark and light regions [4]. This painter's vignette can be produced by fading off two or three discrete regions unevenly or symmetrically in the periphery of the painting.



Fig. 9.1 Vignetting effect in paintings. **a** 'Portrait of Anne, Countess of Chesterfield', Thomas Gainsborough, c1778. Courtesy the Getty Research Institute. **b** 'Fanciful Landscape', Thomas Doughty, 1834. Courtesy National Gallery of Art, Washington

Based on our observations of the differences between photograph-style and painter-style vignetting, this chapter attempts to study painter-style vignetting and proposes an algorithm to apply it to photographs for the purpose of their aesthetic improvement.

9.2 Related Work

Most research on this subject seeks to correct unwanted vignetting caused by camera settings or lens limitations. To determine the degree of vignetting, the main methods used are radiometric calibration by the estimation of the radiometric response function and acquiring the vignetting function from a sequence of collected images [6, 9], and vignetting function estimation from a single image by looking for intensity variations in the radial direction [3, 15, 16].

For the creation of a vignetting effect in photographs, Samii et al. proposed to stylise a false vignetting effect in images [11]. This work was done by learning experts' decision processes when enhancing a photograph. This learning method requires a set of training data that maintains the history of expert adjustments applied to every photograph. It would be a challenging task to collect such data for pre-modern painting styles. In photorealistic rendering, Wu et al. [13] proposed to simulate the vignetting effect based on a lens model in order to estimate the effect of vignetting in the Bokeh effect. For style transfer, Liu et al. proposed to transfer a vignetting effect [10]. The vignetting function in an example image was modelled by a sixth-order even polynomial, which was applied to the user's photograph by simply multiplying

each pixel of the photograph by the vignetting function. The vignetting function in [10] was a camera vignetting function that affected image pixels based on the distance from the pixel to the optical centre.

With the development of digital imaging techniques, vignetting has become a popular effect in some software tools, such as the ‘Lens Correction’ filter in Photoshop® and the ‘Lo-fi’ filter in Instagram®. However, these photograph-style vignetting effects enforce a darker or desaturated corner or edge without consideration of the interior geometric structure of the photograph, and they also create a clear unnatural circle around the centre of the photograph.

Different from the preceding work, this chapter proposes to learn the vignetting effect in paintings by analysing the spatial distribution of the lightness and then adaptively applying the learned lightness weighting pattern to photographs.

9.3 Vignetting Effect in Paintings

This section describes how the vignetting effect is organised in paintings by analysing the spatial distribution of values. Landscape paintings by the Hudson River painters (a group of Romantic era artists from the mid-nineteenth century) are used as the dataset for this analysis.

First, the pixel-by-pixel average lightness values of 920 paintings are calculated to see how the lightness is organised across the 2D image plane. In addition, the pixel-by-pixel average lightness values of 920 landscape photographs are calculated and compared with those of the paintings. Three hundred of the landscape photographs are collected from the MIT-Adobe FiveK Dataset [2], and others are collected from our personal collection. The results are shown in Fig. 9.2a, b. For the purpose of clarity, these figures have been normalised using Photoshop’s auto-contrast function. All the photographs are non-professional and without post-processing. By comparing the average lightness of paintings and photographs, we can clearly see that the corners of the former have been significantly more weighted (being darker) than the latter. The changes of the average lightness along columns and rows evidence similar tendency (see the curve in Fig. 9.2c, d). However, the difference of the lightness values in the centre and periphery is more distinct in paintings than in photographs. The middle of the paintings is brighter, while their periphery is darker. In other words, the vignetting effect exists in both paintings and photographs, but the vignetting effect in paintings is more purposely produced (exaggerated) for aesthetic reasons.

The pixel-by-pixel average saturation values of the 920 paintings and photographs are also calculated to assess the change of saturation on the 2D image plane (see Fig. 9.3). There is a general tendency for the saturation of the paintings to decrease from the centre to the periphery. However, there is larger variation in the local distribution of saturation as compared to that of lightness, with the decrease in saturation from centre to periphery being not so consistent as that of lightness. This chapter therefore focuses on analysing the organisation of lightness for the purpose of producing a vignetting effect.

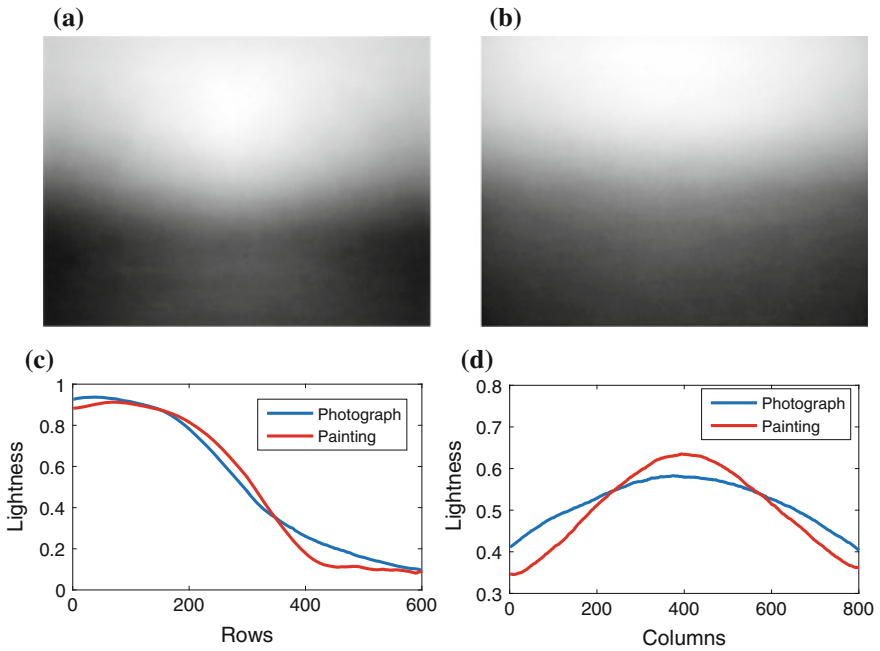


Fig. 9.2 **a** Averaged and normalised lightness of 920 landscape paintings. **b** Averaged and normalised lightness of 920 landscape photographs (**a** and **b** are calculated by resizing all the paintings or photographs to the same size (600×800) and then calculating the pixel-wise average of the lightness image). **c**, **d** are the plots of the average values of, respectively, 800 columns and 600 rows for the two average lightness images in **a** and **b**

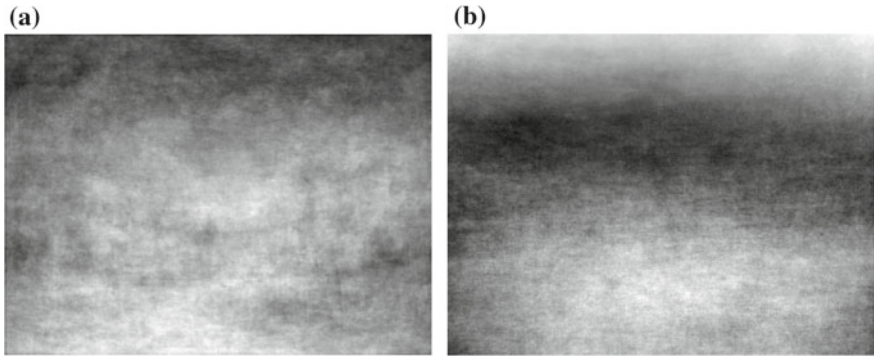


Fig. 9.3 **a** Averaged and normalised saturation image of 920 paintings. **b** Averaged and normalised saturation of 920 photographs

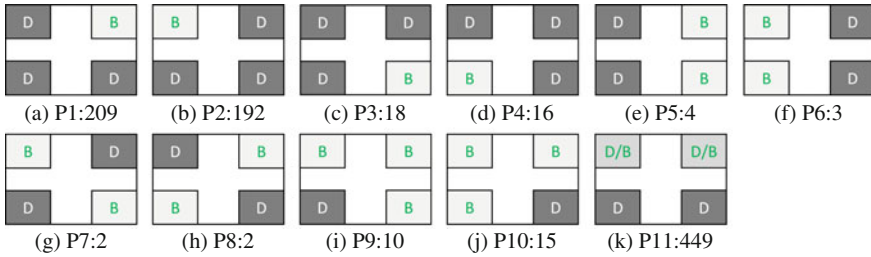


Fig. 9.4 Lightness weighting patterns of 920 paintings. In the sketch, D is darker and B is brighter. The number under each pattern is the number of paintings that employ this pattern. In pattern P11, there is no clear bias on the top or bottom two corners and the lightness in the top corners is in similar brightness or darkness

Table 9.1 Example painting and average lightness of all the paintings in each of the three patterns P1, P2 and P11. CC Image courtesy Wikimedia Commons, The Athenaeum

Patterns	P1 (Left Weighted)	P2 (Right Weighted)	P11 (Normal)
No.	209	192	449
Example painting			
Average lightness			

In painter-style vignetting, the lightness weighting on the periphery is not even, with more weight frequently being given to one or other of the corners. Based on the lightness weighting of corners in paintings, the vignetting is divided into 11 patterns as the pictograms in Fig. 9.4 show. The number under each pattern is the number of paintings that exhibit this pattern. It shows that the predominant patterns are P1, P2 and P11. An example painting and average lightness of these paintings are given in Table 9.1. The paintings in pattern P1 have a relatively brighter top right corner, and the paintings in pattern P2 have a relatively brighter top left corner. The function of one corner being brighter than the others is to open up the depth space of the scene by implying a spiral running down the z-axis of the painting. In the even-weighted pattern P11, the corners of the sky are relatively darker than the centre, but this middle/corner difference is relatively greater in the ground. This implies that the

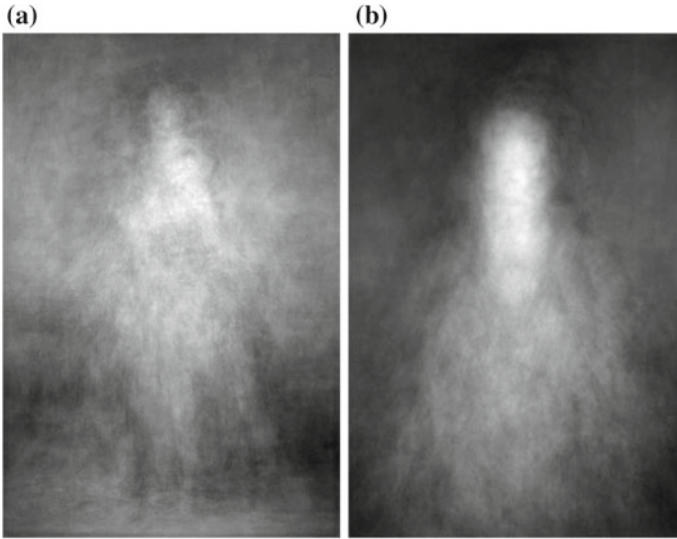


Fig. 9.5 **a** Averaged and normalised lightness of 117 full-body portrait paintings. **b** Average lightness of 146 half body portrait paintings. Portrait paintings with simple *black* background are excluded from the average lightness calculation

darkness of the sky and ground should be processed differently. Additionally, the darkness on different depth planes should also be processed differently (observed from Fig. 9.2). From the example painting and the corresponding average lightness weighting in each pattern in Table 9.1, we can see that the painters use whatever lightness weighting pattern in their paintings as corresponds to the geometry of the scene. However, there appears to be slight bias towards the top right being lighter than the top left. This corresponds to observations made by Shimamura [12].

In portrait paintings, the figure and face is the most important element and is usually placed along the centre vertical. The high contrast between the figure and background is used to draw attention to the figure. The background is usually dark, and monochrome or with indistinct ('fuzzy') details. The vignetting effect in the background of portrait paintings is not as clear as that of landscape paintings (see the average lightness of the portrait paintings in Fig. 9.5).

In summary, the vignetting effect in paintings derives from a darkening of objects on the periphery of the bounds based on the 3D geometry of the scene, and serves to draw attention to the middle while preserving the contrast of dark objects. Additionally, the darkness on the sky and depth planes is introduced differently.

9.4 Transfer of Painter-Style Vignetting to Photographs

The analysis of the vignetting effect in paintings clearly shows that a key technique used by artists is to introduce lightness weighting based on the geometry and content of the scene. Although general lightness weighting patterns have been obtained from paintings, they cannot be simply used as a template to manipulate lightness in photographs due to the variation of their semantic content. Therefore, this chapter proposes to manipulate the vignetting effect in a photograph based on an example painting which is similar in content and geometry to that of the target photograph.

Given a photograph without vignetting, the objective of the work described in this chapter is to create a vignetting effect in a photograph. This objective can be formulated as

$$L_{out} = f(L, W) \tag{9.1}$$

where L is the lightness of the input photograph without vignetting, W is the lightness weighting map. L_{out} is the output lightness with vignetting effect. $f(\cdot)$ is the blending function of the lightness weighting map and the original lightness. Therefore, the main tasks are to generate a lightness weighting map W and to define the blending function f . Based on our observations from paintings, there are four criteria for generating the painter-style vignetting in photographs.

1. Selecting the lightness weighting pattern (e.g. left or right weighted)
2. Enhancing the centre-corner contrast (see centre-corner in Sect. 2.3.2).
3. Preserving the local contrast on the periphery and darken objects (see local contrast in Sect. 2.3.2).
4. Manipulating the sky and depth planes differently.

9.4.1 Lightness Weighting Pattern and Blending Function

Firstly, an example painting was selected from the painting database and its lightness weighting pattern was extracted. The example painting was selected based on the assumption that images similar in content and geometry will require similar lightness weighting on their corners. As with the example painting selection described in Chap. 6 on p. 71, the example painting was selected based on its local correspondence with the input photograph. In Fig. 9.6, one example photograph is shown, together with its recommended top-ranked paintings. The user is given the choice to select one of the top-ranked paintings as the choice example.

After selecting the example painting, its lightness weighting pattern was extracted by estimating the lightness bias field. Given an image Z with lightness bias, it can be modelled as the product of the bias-free image I and the bias field B . The value of the pixel i in Z can be expressed as

$$Z_i = I_i B_i \tag{9.2}$$

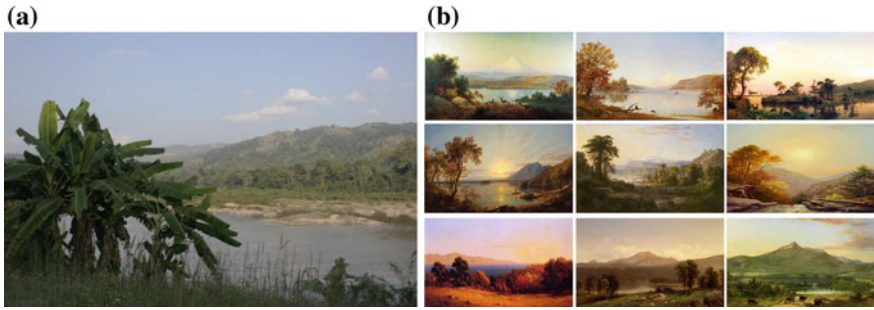


Fig. 9.6 Painting groups **b** are those top-ranked as being similar in geometric structures to the photograph in **a**

The task of bias field estimation is to estimate B together with I based on the value Z . This ill-posed problem can be formulated as an optimisation problem based on the constraints on B and I . The estimation of B can be achieved by the following optimisation function.

$$B' = \arg \max_B P(B | Z) \propto \arg \max_B P(Z | B)P(B) \quad (9.3)$$

The probability models of $P(Z | B)$ and $P(B)$ are constructed based on the sparseness priors of the gradient probability distribution in [14]. They are used to estimate the lightness bias field. For the three paintings in Table 9.1, the extracted lightness bias fields are shown in Fig. 9.7. The estimated lightness bias field presents a coarse

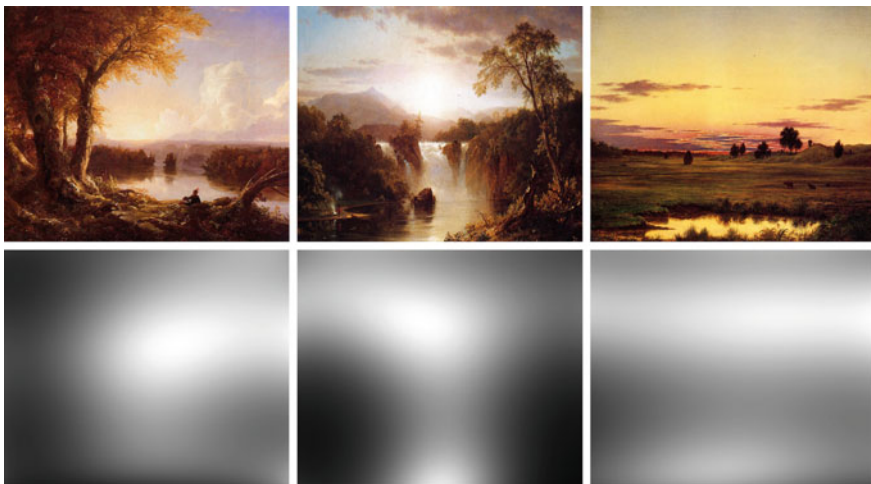


Fig. 9.7 Extracted lightness bias fields for the three paintings in Table 9.1

lightness distribution in the 2D image plane, indicating the dark regions and bright regions. The estimated lightness bias field from the example painting is used as the initial lightness weighting pattern M for the input photograph.

Given the extracted lightness weighting pattern, how is it to be blended with the original lightness while satisfying criteria 2 to 4? The three blending algorithms, known as Darken, Multiply and Overlay, are adopted here. Two layers are created following [1]. The upper layer contains the weighting pattern extracted from a painting, and the lower layer contains the input photograph. The Darken function given in [1] is as follows:

$$L'(i, j) = \begin{cases} M(i, j) & \text{if } M(i, j) < L(i, j) \\ L(i, j) & \text{otherwise} \end{cases} \tag{9.4}$$

where L is the lightness value in the lower layer, M is the lightness weighting pattern value in the upper layer. L' is the output lightness after the blending. Here, M and L are in the same value range $[0, N]$.

The Multiply mode serves simply to multiply the values of the two layers [1]. Given the range N of M , the Multiply mode is

$$L' = L \times \frac{M}{N} \tag{9.5}$$

Equation (9.5) shows that $L'(i, j) \leq L(i, j)$ for every pixel (i, j) . Therefore, in nonlinear colour space, the Multiply mode makes the whole image darker than the original.

Differently, the Overlay mode serves to both darken and lighten the output [1]. The influence of M also depends on the value of L . The formulation is

$$L'(i, j) = \begin{cases} \frac{M(i, j)(N-L(i, j))}{\frac{N}{2}} + L(i, j) - (N - L(i, j)) & \text{if } L(i, j) > \frac{N}{2} \\ \frac{M(i, j)L(i, j)}{\frac{N}{2}} & \text{otherwise} \end{cases} \tag{9.6}$$

The effects of the Darken, Multiply and Overlay modes can be clearly seen in Fig. 9.8. The Overlay mode satisfies the criteria for painter-style vignetting. It brightens areas that are bright in the lightness weighting pattern and darkens areas that are dark. In addition, local contrast is preserved to some degree.

9.4.2 Lightness Weighting Pattern Correction

The extracted lightness weighting pattern M from the example painting only indicates the relative brightness and darkness of regions. If it is directly used to weight the lightness of the photograph, the result may be too bright or too dark. The sky in Fig. 9.9d is generally darker than both the original and the example painting. In order to avoid over-brightened or over-darkened regions, the lightness weighting

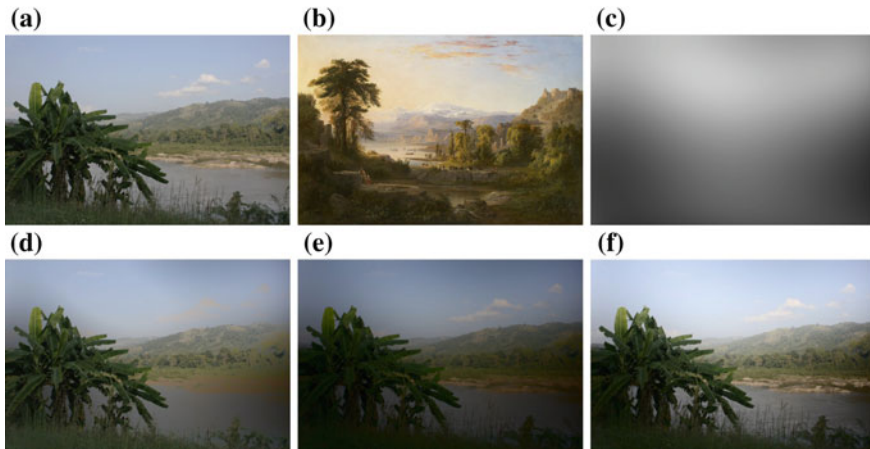


Fig. 9.8 Effects of the Darken, Multiply and Overlay modes. The blending is only conducted on the lightness of the original photograph. **a** Original photograph. **b** Example painting (CC Image courtesy Wikimedia Commons). **c** Lightness weighting. **d** Darken mode. **e** Multiply mode. **f** Overlay mode

pattern extracted by the bias field estimation is corrected using a nonlinear curve. The objective of the correction is to limit the dark regions and bright regions of the result, in order that they have similar values to those of the example painting. Given the correction function $f_c(\cdot)$, the corrected weighting pattern is

$$M' = f_c(M) \quad (9.7)$$

First, the intended bright regions and dark regions after blending need to be determined in order that their values be limited. Based on the definition of the Overlay blending mode, the bright regions in the blended output correspond to the regions that are bright in both the original input photograph and the lightness weighting pattern. The dark regions in the blended output correspond to the regions that are dark in both the original input photograph and the lightness weighting pattern. The regions corresponding to the brightest 25th percentile in M are considered as the intended bright regions B_0 in the blended output, and the regions corresponding to the darkest 25th percentile in M are considered as the intended dark regions D_0 in the blended output. The corresponding regions covered by B_0 and D_0 of the original lightness may consist of bright or dark pixel values. Therefore, the bright regions B_s are determined as the corresponding regions of the lightest 50th percentile covered by B_0 in the original lightness and the dark regions D_s are the corresponding areas of the darkest 80th percentile that have a lightness value $L < 0.5$ and overlapped by D_0 . The regions highlighted in red in Fig. 9.9e are the bright regions determined after the lightness weighting. The regions highlighted in blue in Fig. 9.9e are the dark regions determined after the lightness weighting. The same percentage of the brightest pixels in the example painting covered by B_0 are determined as the bright

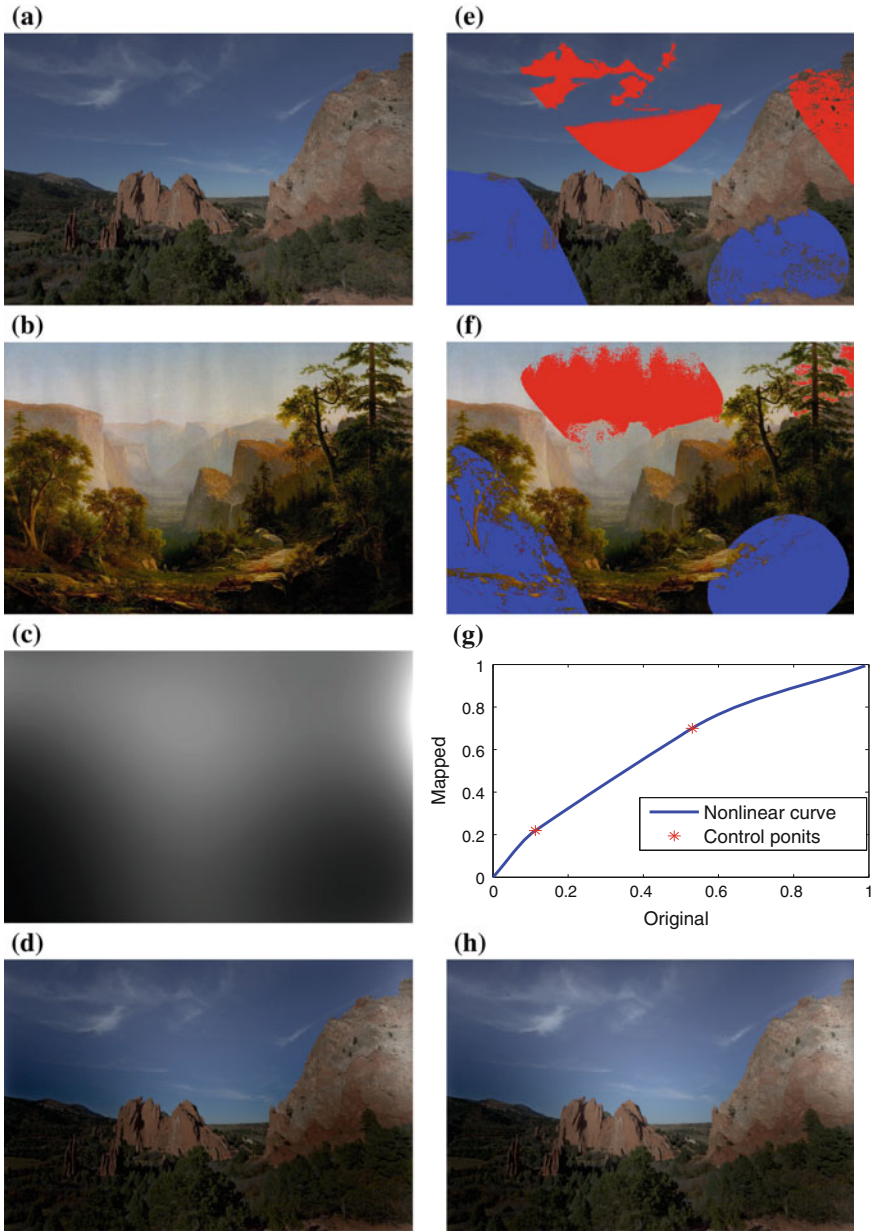


Fig. 9.9 Lightness weighting pattern correction. **a** Original photograph. **b** Example painting (CC Image Courtesy of <https://www.wikiart.org>). **c** Extracted lightness weighting pattern of **b**. **d** Result using **c** without correction. **e** Original photograph with highlighted points. **f** Example painting with highlighted points. **g** Correction curve. **h** Result using corrected lightness weighting pattern

regions B_r (red regions in Fig. 9.9f), and the same percentage of the darkest pixels in the example painting covered by D_0 are determined as the dark regions D_r (blue regions in Fig. 9.9f).

Given the average value A_{B_s} of the regions B_s in the original lightness and the average value A_{D_s} of the regions D_s , the average value A_{B_r} of the regions B_r and the average value A_{D_r} of the regions D_r , the lightness weighting correction serves to satisfy:

$$A_{B_r} = f_o(A_{B_s}, M_B) \quad (9.8)$$

$$A_{D_r} = f_o(A_{D_s}, M_D) \quad (9.9)$$

where $f_o(\cdot)$ is the Overlay blending mode in Eq. (9.6). M_B and M_D are the desired weightings for the lightness values A_{B_s} and A_{D_s} , respectively. M_B and M_D can be calculated using the inverse function of $f_o(\cdot)$. Given the current weightings M_B^o and M_D^o for A_{B_s} and A_{D_s} which are the average values of the corresponding regions of B_s and D_s in M , the control points are $a = [0, M_B^o, M_B^o, 1]$ and $b = [0, M_D^o, M_D^o, 1]$. Then, the nonlinear curve is interpolated based on the function in Eq. (7.4). Given the corrected lightness weighting pattern M' and the blending mode $f_o(\cdot)$, we get the weighted lightness $L' = f_o(L, M')$. The nonlinear curve for an example is shown in Fig. 9.9g. By using the corrected lightness weighting pattern, the over-darkened regions are successfully avoided in the result in Fig. 9.9h.

9.4.3 Content-Aware Interpolation

Although the selected example painting has a similar geometry to that of the input photograph, there are still some variations in the spatial organisation of the structures. As the two examples in Figs. 9.11 and 9.12 show, the foreground darkness in the example painting is weighted on the sky area of the photograph when the lightness weighting pattern is used globally as a mask. Therefore, instead of using the weighting pattern globally without consideration of the variation of content, the weighting pattern is interpolated adaptively to fit the structure of the photograph.

In landscape paintings, the FG and sky are of different darkness, and also the FG, MG and BG planes are darkened to different degrees. Here, the ‘darken’ blend mode serves to make the peripheral area of a plane darker, not to change the lightness of the whole depth plane as in our atmospheric perspective effect manipulation (see Chap. 6 on p. 71). However, similar to this work, the landscape images are divided into depth planes: FG, MG, BG and sky. The weighting pattern of each plane is then generated based on the lightness weighting of the corresponding plane in the example painting.

Given the weighting of one plane in the example painting M_r , it is first registered with the corresponding plane of the photograph. The registration is performed by shifting the location by a vector (s_x, s_y) to match the centre of the plane in the photograph and scaling its size by α . M_r is scaled to the same width and height of the corresponding region in the photograph. After getting the registered weighting M_r^R ,

a surface S is constructed for the 3D points $(x, y, M_r^R(x, y))$. Then, the weighting of the corresponding plane in the photograph is linearly interpolated from the surface S . Finally, the weighting pattern of the photograph is generated by combing the weighting patterns for all the depth planes. Holes in the combined weighting pattern are filled using the data from the combined weighting pattern through linear interpolation. The hole-filled weighting pattern \tilde{M}_s is not smooth on the boundary between regions. A smoothed weighting pattern \tilde{M}_s is generated by using an edge-preserving smoothing [5] based on the following optimisation function:

$$\tilde{M}_s = \arg \min_{\tilde{M}_s} \sum_{p \in \tilde{M}_s} \{[\tilde{M}_s(p) - M_s(p)]^2 + w_s(p)h(\nabla \tilde{M}_s, \nabla L)\} \quad (9.10)$$

$$h(\nabla \tilde{M}_s, \nabla L) = \frac{|\tilde{M}_s^x(p)|^2}{|g_x(p)|^\theta + \varepsilon} + \frac{|\tilde{M}_s^y(p)|^2}{|g_y(p)|^\theta + \varepsilon} \quad (9.11)$$

The smoothing serves to keep the gradient of the output \tilde{M}_s as small as possible, unless the original lightness L has a significant gradient. $(\tilde{M}_s^x, \tilde{M}_s^y)$ is the gradient of \tilde{M}_s , and (g_x, g_y) is the gradient of L . θ controls the sensitivity to the gradient of L . ε is a small constant. The weight w_s is

$$w_s(p) = \begin{cases} \tau_1 & \text{if } p \in \text{boundary area} \\ \tau_2 & \text{otherwise} \end{cases} \quad (9.12)$$

τ_1 is larger than τ_2 for smoothing the boundary area. In the implementation, the parameters are set as $\theta = 0.5$, $\varepsilon = 0.0001$. $\tau_1 = 0.1$ and $\tau_2 = 0.001$. The effect of each step in the content-aware interpolation can be clearly seen in Fig. 9.10. Using the interpolated lightness weighting pattern, the left part of the sky in Fig. 9.11h and the right part of the sky in Fig. 9.12h are not as dark as the FGs.

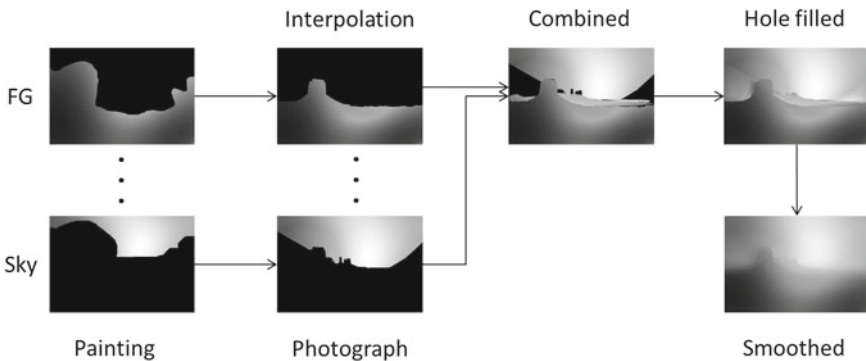


Fig. 9.10 Content-aware interpolation of the lightness weighting pattern

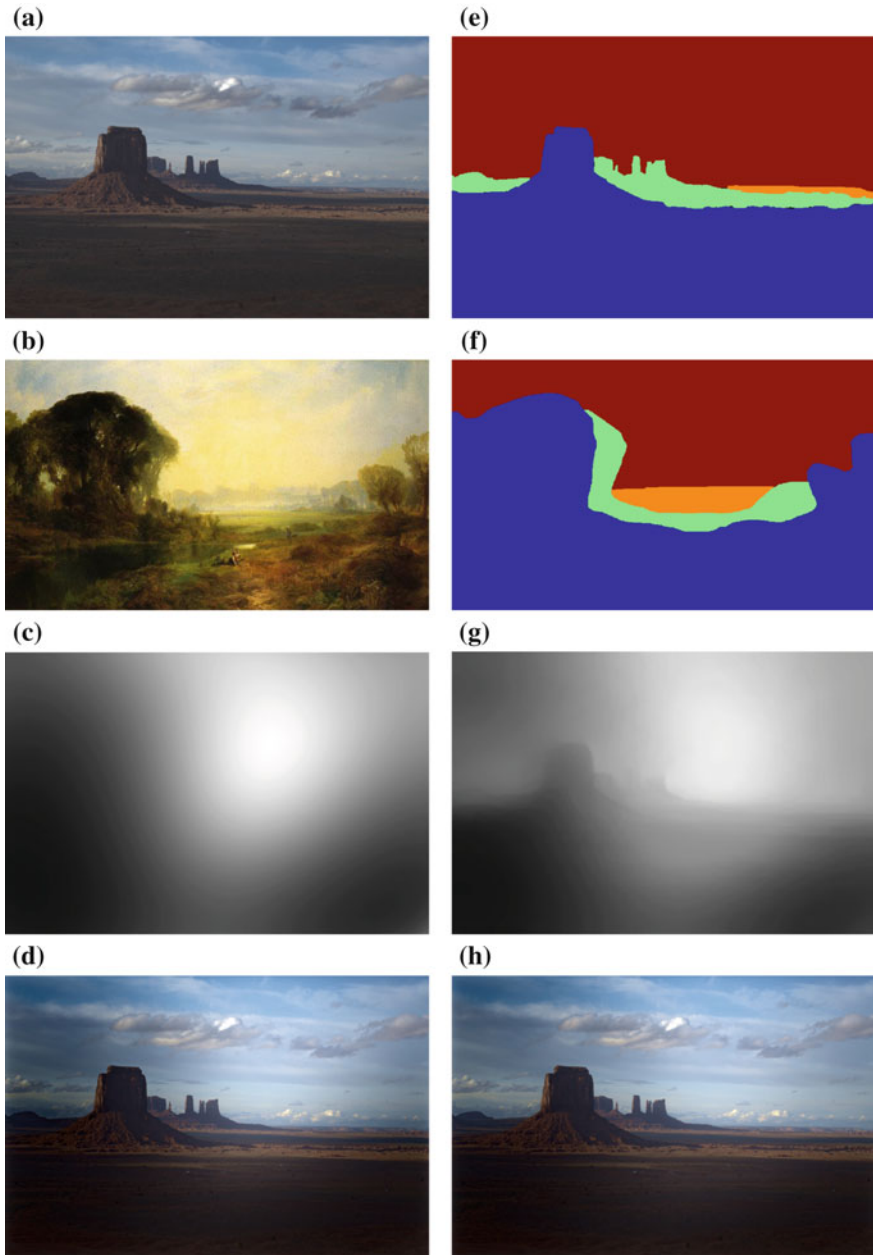


Fig. 9.11 **a** Original photograph. **b** Example painting (CC Image courtesy The Athenaeum). **c** Lightness weighting pattern of the example painting. **d** Result using weighting in **c**. **e** Mask of the original photograph in **a**. **f** Mask of the example painting in **b**. **g** Lightness weighting pattern after content-aware interpolation. **h** Result using weighting in **g**

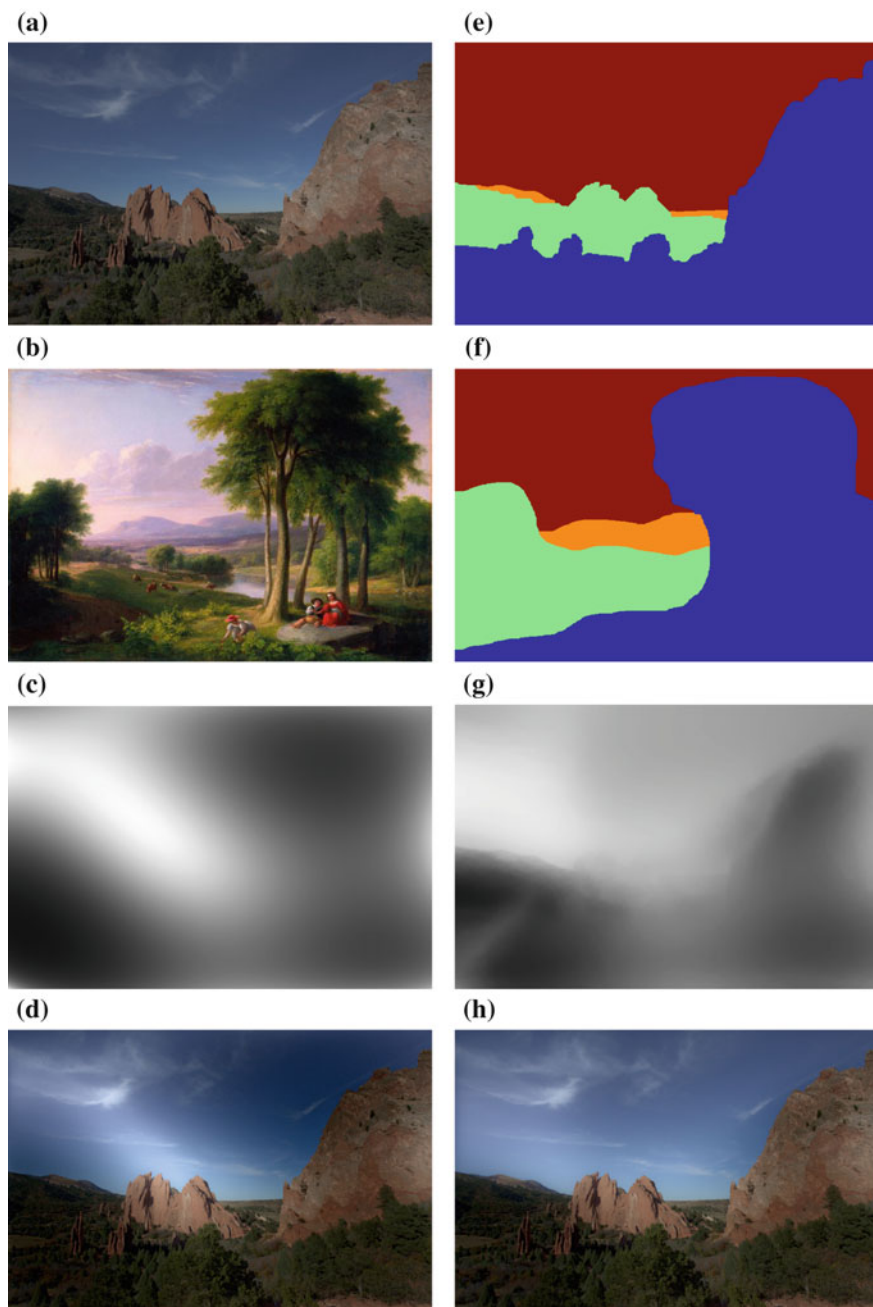


Fig. 9.12 **a** Original photograph. **b** Example painting (CC Image courtesy Wikimedia Commons). **c** Lightness weighting pattern of the example painting. **d** Result using weighting in **c**. **e** Mask of the original photograph in **a**. **f** Mask of the example painting in **b**. **g** Lightness weighting pattern after content-aware interpolation. **h** Result using weighting in **g**

9.4.4 Local Contrast Restoration

Using the content-aware interpolated lightness weighting pattern \tilde{M}_s and the blending mode $f_o(\cdot)$, the weighted lightness is calculated as $\tilde{L} = f_o(L, \tilde{M}_s)$. Although the Overlay blend mode can preserve local details to some degree, local contrast is still reduced in dark regions. In order to preserve the local details, an optimisation function is used to restore the reduced local contrast in \tilde{L} . The lightness L_o with restored local contrast is calculated as

$$L_o = \arg \min_{L_o} \sum_{i \in L_o} [E_d(i) + E_g(i)] \quad (9.13)$$

The first term E_d is to constrain the output to have the similar value with \tilde{L} , which is defined as

$$E_d(i) = w_d(i)[L_o(i) - \tilde{L}(i)]^2 \quad (9.14)$$

The second term E_g is the constraint on the gradient, which is defined as

$$E_g(i) = w_g(i)[f_x(i) - g_x(i)]^2 + w_g(i)[f_y(i) - g_y(i)]^2 \quad (9.15)$$

where f_x and f_y denote the x - and y -derivatives of L_o . g_x and g_y denote the x - and y -derivatives of the original lightness L . This constraint is to ensure that the output has the same gradient of the original input. Since gradient is one expression of local contrast, the output with the restored gradient has the local contrast of the original. w_d and w_g are the weights for the two constraints. In the implementation $w_d = 0.01$, $w_g = 1$. The example in Fig. 9.13 shows the effect of the local contrast restoration.

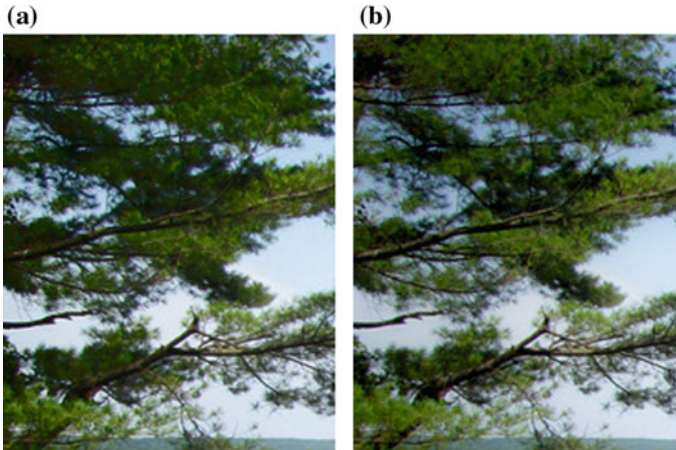


Fig. 9.13 Local contrast restoration. **a** Close-up of region of result without local contrast restoration. **b** Close-up of region of result with local contrast restoration

9.5 Experiments and Discussion

In this section, the performance of the proposed vignetting effect manipulation method is analysed. Firstly, the difference between the vignetting effect manipulation and atmospheric perspective effect enhancement (in Chap. 6 on p. 71) is discussed. The atmospheric perspective effect enhancement serves to enhance the difference of depth planes and thereby enhance the impression of depth in 2D images. The values within a depth plane are adjusted in the same way. Differently, the vignetting effect serves to adjust the contrast between centre and corner in order to attract the attention of viewers to the centre. In doing so, it is aware of the limits of the picture plane. In each depth plane, the region near the centre is weighted differently from the region close to the periphery. The difference of the vignetting effect and atmospheric perspective effect can be clearly seen by comparing the two rendering results in Fig. 9.14a, b. In paintings, the vignetting effect and atmospheric perspective effect are used together for aesthetic composition. Depending upon the particularities of the scene and on the artist's intention, one effect might be more evident than the other. For example, featureless desert terrains are better served by the atmospheric perspective effect, while wooded European landscapes are better served by the vignetting effect.

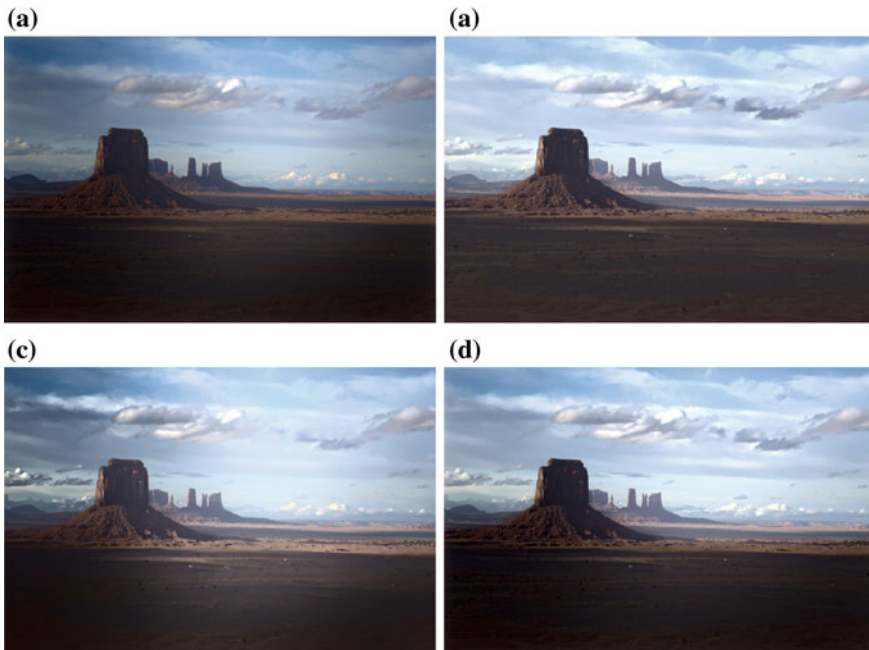


Fig. 9.14 The original photograph is shown in Fig. 9.11a and the example painting is shown in Fig. 9.11b. **a** Result achieved using vignetting effect manipulation. **b** Result achieved using atmospheric perspective effect enhancement. **c** Vignetting + atmospheric perspective effect enhancement. **d** Atmospheric perspective effect + vignetting enhancement

Given one example painting, the vignetting and atmospheric perspective effects can also be enhanced together in the photograph. However, the order of the two enhancements affects the final appearance. If the vignetting effect is first applied followed by the enhancement of atmospheric perspective effect, the vignetting effect is amplified in the depth-aware contrast manipulation for the atmospheric perspective effect enhancement, see the result in Fig. 9.14c. If the atmospheric perspective effect is first enhanced followed by the vignetting effect manipulation, there is no exaggeration or reduction in the contrast (see Fig. 9.14d).

In order to show the advantage of the proposed vignetting effect manipulation method, it is compared with two popularly used camera models commonly used to creating vignetting, and the vignetting effect of two popular tools (the vignette in Photoshop[®] and the ‘Lo-fi’ filter in Instagram[®]). The first camera model (camera model 1) was a natural vignetting model which is based on natural illumination radially falling-off due to geometric optics [7]. The formula of this vignetting model is

$$V_m = V_o \cdot \cos^4 \left(\tan^{-1} \left(\frac{r}{f} \right) \right) \quad (9.16)$$

where V_m is the measured illumination by the sensor and V_o is the incoming illumination. f is the focal length of the camera, and r is the distance of the pixel from the centre. The second camera model (camera model 2) was the Kang–Weiss vignetting model proposed in [8]. In the two models, the focal length was set as $f = 600$. Figures 9.15 and 9.16 show two examples of the comparison. The two camera models and the vignette in Photoshop[®] only darken the pixels based on the distance from the centre. All pixels the same distance from the centre are darkened to the same degree, no matter whether they are sky pixels or ground pixels. There is a clear circle in the result. The ‘Lo-fi’ filter in Instagram[®] darkens the periphery pixels and also brightens the centre. The pixels close to the periphery are darkened without consideration of the content. There is also a clear circle in the result. Differently, the vignetting effect achieved using the proposed method is more naturally presented without a clear circle and the sky is darkened to a different degree to the ground, following the aesthetic composition in the example painting.

Some more experimental results on landscapes are shown in Fig. 9.17. The proposed vignetting effect manipulation method can also be applied to portrait photographs. Because the estimated bias field by [14] only describes the coarse distribution of the lightness, a bright face surrounded by dark regions cannot be extracted as bright in the bias field, see Fig. 9.18c. The face in the interpolated weighting pattern based on the outlined foreground and background is also dark, see Fig. 9.18d. According to the portrait painting analysis in Chap. 7 on p. 107, the face and skin areas are usually the brightest parts in the figure. Therefore, based on this prior knowledge of portrait paintings, the lightness weighting on the face and skin areas is brightened (set to the 95 percentile value of the lightness weighting on the figure). Figure 9.18e is the lightness weighting pattern with face and skin areas correction. We can see that the result using the lightness weighting pattern with face and skin correction in Fig. 9.18f shows similar lightness weighting as that of the



Fig. 9.15 Comparison of different vignetting effect manipulation methods. **a** Original photograph. **b** Example painting (CC Image courtesy of the The Athenaeum). **c** Lightness weighting pattern extracted from the example painting. **d** Result achieved using the proposed method. **e** Result achieved using camera model 1. **f** Result achieved using camera model 2. **g** Vignette achieved using Photoshop®. **h** Result achieved using ‘Lo-fi’ filter in Instagram®



Fig. 9.16 Comparison of different vignetting effect manipulation methods. **a** Original photograph. **b** Example painting (CC Image courtesy The Athenaeum). **c** Lightness weighting pattern extracted from the example painting. **d** Result achieved using the proposed method. **e** Result achieved using camera model 1. **f** Result achieved using camera model 2. **g** Result achieved using Photoshop®. **h** Result achieved using the 'Lo-fi' filter of Instagram®

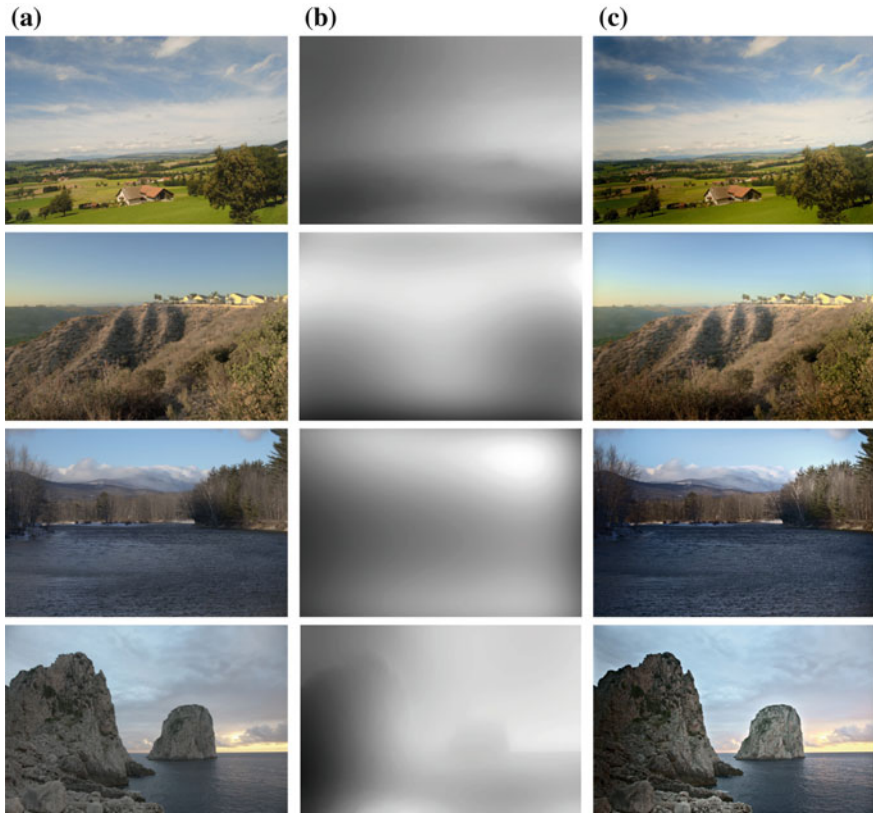


Fig. 9.17 More experimental results on landscape photographs. **a** Original photographs. **b** Lightness weighting patterns. **c** Results

example painting. In the original portrait photograph Fig. 9.18a, the lower-right half is darker than the lower-left half. In the lightness weighting pattern derived from the reference painting, the area around the bottom-right half is also darker than the bottom-left half. By applying the lightness weighting pattern from the reference to the photograph, the left half is enhanced more than the right. This lightness spatial distribution is similar to that of the reference painting. Moreover, in the result shown in Fig. 9.18f, the figure is brighter and the corners in the background are darker. This manipulation makes the figure stand out in the photograph and draws the attention of viewers to the figure, which is the intended effect. Two further portrait examples are shown in Fig. 9.19.



Fig. 9.18 Application of the approach to portrait photographs. **a** Original photograph. **b** Example painting. **c** Lightness weighting pattern extracted from the example painting. **d** Interpolated lightness weighting pattern without face and skin areas correction. **e** Interpolated lightness weighting pattern with face and skin areas correction. **f** Result using weighting pattern **e**

9.6 Concluding Remarks

This chapter proposed an example-based vignetting effect manipulation method in photographs. The proposed method was based on the analysis of the vignetting effect in paintings, which was more purposely produced for aesthetic composition and was naturally presented based on the geometry of the scene. An algorithm was developed to apply the painter-style vignetting to photographs, while satisfying four manipulation criteria observed from paintings. First, the lightness weighting pattern was extracted from an example painting chosen for the similarity of its geometry to that of the input photograph. Then, the Overlay blend mode was applied to blend the



Fig. 9.19 Two more examples of the application to portrait photographs. **a** Original photographs. **b** Lightness weighting patterns. **c** Results using weighting patterns in **b**

lightness weighting pattern with the original lightness to enhance its centre-corner contrast. The sky and ground were processed differently based on their original values. In order to avoid over-brightened or over-darkened regions in the result, the extracted lightness weighting pattern was corrected. This was done using a nonlinear curve to limit the dark and bright regions in the blended result to match the lightness of the example painting. A content-aware interpolation method was proposed to warp the lightness weighting pattern of the example painting to fit the structure of the photograph. The content-aware interpolation made the algorithm immune to the differences of the geometric structure between the example painting and the input photograph. Finally, local contrast was restored using an optimisation function to preserve the local contrast in darkened regions. Experiments showed that, by using the

proposed method, the vignetting effect from the example painting was successfully transferred to the photograph. Comparison with popular software and camera models showed the advantage of the proposed method.

References

1. Binary-Artist.com: Photoshop Tutorials. http://www.binary-artist.com/photoshop_tutorials.php
2. Bychkovsky, V., Paris, S., Chan, E., Durand, F.: Learning photographic global tonal adjustment with a database of input/output image pairs. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 97–104 (2011)
3. Cho, H., Lee, H., Lee, S.: Radial bright channel prior for single image vignetting correction. In: Proceedings of European Conference Computer Vision, pp. 189–202 (2014)
4. Edin, R., Jepsen, D.: Color Harmonies: Paint Watercolors Filled with Light. North Light Books, an imprint of F+W Media, Inc. (2010)
5. Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.* **27**(3), 67:1–67:10 (2008)
6. Goldman, D.B.: Vignette and exposure calibration and compensation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(12), 2276–2288 (2010)
7. Jia, J., Tang, C.K.: Tensor voting for image correction by global and local intensity alignment. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(1), 36–50 (2005)
8. Kang, S.B., Weiss, R.: Can we calibrate a camera using an image of a flat textureless lambertian surface? *Proc. of European Conf. Comput. Vis.* **2**, 640–653 (2000)
9. Kim, S.J., Pollefeys, M.: Robust radiometric calibration and vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(4), 562–576 (2008)
10. Liu, Y., Cohen, M., Uyttendaele, M., Rusinkiewicz, S.: Autostyle: automatic style transfer from image collections to users’ images. *Comput. Graph. Forum* **33**(4), 21–31 (2014)
11. Samii, A., Althoff, T.: Iterative learning: Leveraging the computer as an on-demand expert artist. In: CS281A Statistical Learning Theory (Michael Jordan and Martin Wainwright) and CS294-69 Image Manipulation and Computational Photography (Maneesh Agrawala), University of California, Berkeley pp. 1–11 (2011)
12. Shimamura, A.P.: *Experiencing Art: In the Brain of the Beholder*. Oxford University Press, Oxford (2015)
13. Wu, J., Zheng, C., Hu, X., Xu, F.: Rendering realistic spectral bokeh effect due to lens stops and aberrations. *Vis. Comput.* **29**(1), 41–52 (2012)
14. Zheng, Y., Grossman, M., Awate, S., Gee, J.: Automatic correction of intensity nonuniformity from sparseness of gradient distribution in medical images. In: Proceedings of the 12th International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 852–859 (2009)
15. Zheng, Y., Lin, S., Kambhmettu, C., Yu, J., Kang, S.B.: Single-image vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(12), 2243–2256 (2009)
16. Zheng, Y., Lin, S., Kang, S.B., Xiao, R., Gee, J.C., Kambhmettu, C.: Single-image vignetting correction from gradient distribution symmetries. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(6), 1480–1494 (2013)

Chapter 10

Defining Hue Contrast

Abstract Quantifying the style of visual art has long been a challenging yet largely overlooked problem. Of key importance in the style of a painting is the manner in which the artist has employed hue. We propose a means to quantify global hue contrast as two values: hue variety and hue antagonism. These were derived from Johannes Itten's work on colour contrast. Using our approach, the hue contrast of a single image can be visually expressed as a single point in a plot. For this reason, the approach is particularly suitable for the comparative analysis of groups of artworks, such as those by two or more artists or an artist at different times in their career. We describe two case studies of such an analysis, and the results agree with observations. We also include a user study that supports the validity of our proposed approach.

Keywords Hue contrast · Global contrast · Artistic style · Cinematics

10.1 Introduction

The task of defining global hue contrast is problematic and can be summarised simply using the following analogy. Consider a room full of people. The height of these people can be understood according to such values as the shortest, tallest, median and the standard deviation. This is analogous to the lightness or the saturation global contrast of an image, which can be defined using a range of techniques, most of which call upon maximum, minimum and average values. Defining the global hue contrast is analogous to defining the range of nationalities represented in that room. We cannot average nationality, and there is no such thing as a 'least' or 'greatest' nationality.

The need to define global hue contrast is clear. Of the three perceptual components of colour (hue, saturation and lightness), hue is generally understood to be the most impactful. Supporting this is the common artist's saying that 'lightness does all the work, colour gets all the glory'. However, hue plays a larger part in the stylistic characterisation of some paintings than others. Paintings made before the early modern age (in art heralded by the work of Eugène Delacroix (1798–1863)) relied heavily on lightness contrast for their impact. Delacroix's work leads directly to the

Impressionists and Post-Impressionists, whose work was characterised by the use of shallow tones and wide hue variance to test our approach, two sets of paintings by the same artist are examined for their relative use of hue: Pablo Picasso's (1881–1973) Blue Period paintings and Rose Period paintings [2].

Certainly, the world that we live in today is more colourful than that of our grandparent's time, with a ubiquity of bright lighting and cheap synthetic colours. We therefore also examine two contemporary artworks: Pixar's animation 'Finding Nemo' [4] and DreamWorks's animation 'Sharks Tale' [13] for their relative use of hue.

10.2 Related Work

Much work has been done on the use of hue in art and design. This has tended to focus on the subject of so-called colour harmony, which is discussed in Sect. 2.3.3 on p. 25. Moretti and Lyons [11] and Neumann and Nemcsics et al. [12] have developed tools to generate sets of colour swatches selected for their harmonic accordance. Through an analysis of fashion and art, Matsuda [10] developed a set of eight harmonic colour schemes which represent how harmonious configurations of colour typically present within a colour wheel. This work addresses colour harmony in art. In the context of colour and art, we understand the word harmony to refer to 'good contrast'. Such work is concerned with the classification of artworks into harmony types (e.g. 'this artwork employs an analogous harmony, whereas that artwork employs a split complementary harmony'). Such analysis is difficult to generalise, being unsuitable for the purpose of classifying large groups of images. Differently, we sought a general definition, one that could be expressed simply, consistently and numerically. Such an approach would be more suitable to express the hue properties of image sets.

10.3 Defining Hue Contrast

In his seminal book 'The Art of Color' [7], Johannes Itten describes seven colour contrasts of which three are both global and specifically address hue:

1. **The contrast of hue.** This refers to the amount of perceptually different hue values in an artwork.
2. **The contrast of warm and cool.** Perversely, this has two readings: warm and cool as a relative value (as in the warm yellow of a sunflower versus the cold yellow of a lemon) and warm and cool as an absolute value. The latter assumes that the RYB hue wheel be bifurcated along an axis, with yellow to red-violet being warm and all other values being cool.
3. **The contrast of complements.** This describes the contrast of values located on opposite sides of the RYB hue wheel.

Itten describes two other contrasts, which may indirectly involve hue: simultaneous contrast and the contrast of extension. These are relative contrasts and therefore not of relevance to this enquiry, as can the contrast of warm and cool in its relative form.

These three contrasts can be condensed. The contrast of complements may be combined with absolute temperature contrast, as there is no complementary contrast that is not also an absolute warm and cool contrast. We may therefore generalise Itten's hue contrasts to address only two properties: the number of perceptually different hue values and the property of 'oppositeness': (values on opposite sides of the red, yellow and blue (RYB) hue wheel). These we term, respectively, **hue variety** and **hue antagonism**.

To measure hue variety, we employ the Muncell hue wheel. Muncell mapped his colour space according to 'barely perceptual difference'. The result is a hue wheel that is perceptually uniform, meaning that there are as many perceptually different hue values in any pie-slice of the wheel as there are in a pie-slice of the same degree of spread in any other part of the wheel. The spread of the hue histogram of an image expressed within this wheel would therefore accurately express the amount of perceptually different hue values in that image. Muncell colour space is discussed in Sect. 2.2 on p. 10.

To measure hue antagonism, we employ the RYB hue wheel (see Sect. 2.2 on p. 10). The RYB hue wheel is defined by complementary pairs which have been known by artists since at least Leonardo Da Vinci's time. An RYB complementary pair is at a maximum state of perceptual difference, a property Leonardo termed 'retto contrario' literally 'exactly opposite'. The hue histogram of an image expressed within this wheel would therefore accurately express the antagonistic properties of that image. RYB colour space is discussed in Sect. 2.2 on p. 10.

10.4 Our Approach

Firstly, grey, black and white (i.e. de-saturated) regions have no perceptual hue. To eliminate the influence of such values, pixels with a saturation less than an absolute threshold rate are ignored, leaving only those with a perceptually apparent hue. Following this, we utilise the different properties of the Muncell and RYB hue wheels to define hue variety and hue antagonism. The process is outlined below, and full technical details can be found in the supplementary material.

1. To calculate the hue variety of an image, its hue is mapped as a circular histogram within a Muncell hue wheel. In total, the wheel consists of 360° of hue difference. The hue variety of an image is calculated as being the total number of those degrees of hue difference that lie within the spread of its histogram. The perceptually uniform nature of the Muncell wheel means that this figure is a reliable indicator of the amount of perceptually different hue values within the image. This process

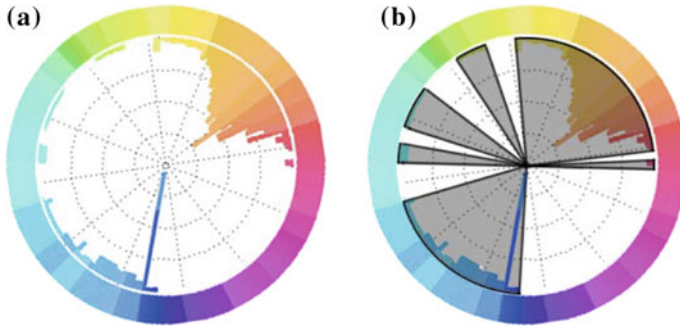


Fig. 10.1 Hue values expressed as circular histogram within a Munsell hue wheel (*left*). Hue variety is defined as the sum of the spread angles (the *shaded grey* regions in the figure on the *right*) (color figure online)

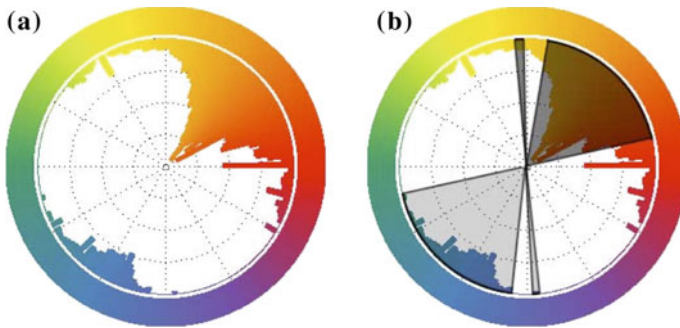


Fig. 10.2 Hue values expressed as circular histogram around an RYB hue wheel (*left*). (*Right*) the hue values in *dark grey* are in an antagonistic relationship to the values in *light grey*. The hue antagonism is calculated as the angle of spread of this antagonism in ratio to 180 (color figure online)

is visually summarised in Fig. 10.1, where the hue variety may be understood as the sum of the grey areas as expressed against a possible maximum of 360.

2. To calculate the hue antagonism of an image, the hue is mapped as a circular histogram within the RYB hue wheel. Any pair of hue values within this histogram may be considered an antagonistic pair (in complete antagonism). To measure the hue antagonism, the wheel is bisected at any axis, and the hue of the half-histogram on one side of the wheel is examined to determine whether there is a corresponding hue in the half-histogram on the opposing half. As the RYB hue wheel presents hue in its artistic antagonistic pairing, this value is therefore an indicator of the amount of hue antagonism within the artwork. This process is visually summarised in Fig. 10.2, where the dark grey regions represent hue that is in an antagonistic relationship to hue within the half of the histogram on the opposite half.
3. The hue variety and hue antagonism is visualised in a plot with the former expressed on the X-axis and the latter on the Y-axis.

As each artwork returns a single point on the global hue contrast plot, the hue characteristics of many artworks may be evaluated easily in relation to each other. This is particularly useful in the comparison of large groups of stylistically consistent images such as work by different artists, or by artists at different times in their careers.

10.5 Computational Details

The circular hue histograms: The hue antagonism is measured on the RYB hue wheel. The hue variety is measured on the Munsell hue wheel. Since all digital images lie in RGB colour space, we have first to convert the RGB colour wheel to RYB and Munsell colour wheels.

The calculation of the lookup table for RGB–RYB hue conversion

- Step 1. The computation of the RGB to RYB conversion:** The R, Y and B values are converted from RGB space using the tri-linear interpolation in [6, 14].
- Step 2. The preparation of the RYB hue wheel:** The RYB hue wheel in range (0, 359) is divided into $360 \times S$ units, $S = 5$ in this work.
- Step 3. The computation of the RYB to HSL conversion:** According to the RYB hue wheel, and setting the saturation and lightness as 1, the HSL from RYB conversion can be achieved according to [5].
- Step 4. The computation of the histogram on H of RYB:** The hue value of RGB is obtained by converting RGB to HSL according to the RGB–HSV conversion in [5].

N.B.: We have to make sure each RGB value has a corresponding RYB hue value. For those RGB values that have no RYB hue value in the RYB hue wheel, their corresponding RYB hue values are set to those of their nearest neighbour in the RYB hue wheel. This procedure gives a lookup table for the RYB–RGB hue conversion. By using this lookup table, the RYB hue value of an RGB colour can be found.

The calculation of the Munsell circular hue histogram for RGB image: The calculation of Munsell circular hue histogram is much easier, since there exists source code for the calculation [1].

The calculation of hue variety: The hue variety is calculated on the Munsell circular hue histogram. The hue spread is calculated as the sum of all the regions of spread within the circular histogram.

We define the hue spread as:

$$y = \frac{N_t}{S}$$

where N_t is calculated as: $N_t = \sum_{i=1}^{360 \times S} I_H(i)$ where $I_H(i) = 1$, if $H(i) > th$, otherwise $I_H(i) = 0$. th is the threshold to ignore noise.

The calculation of hue antagonism: The hue antagonism is calculated on the RYB circular hue histogram. For every point within any 180° span of the wheel, we ask whether there is a corresponding point on the opposite side of the wheel. The

existence of a pair of opposite hue values at point θ for $\theta < 180 \times S$ can be measured by $I_H(\theta) \cdot I_H(\theta + 180 \times S)$. When the value is 1, there is opposition at point θ , and when the value is 0, no opposition. The ratio of the number of oppositions to 180 gives us the hue antagonism.

The definition for hue antagonism is:

$$\rho = \frac{\sum_{\theta=1}^{360 \times S} I_H(\theta) \cdot I_H(\theta + 180 \times S)}{180 \times S}$$

where the numerator is the number of oppositions in the circular histogram, $I_H(\theta)$ equals to 1 if the value of the histogram at the position θ is bigger than th . Otherwise, it is 0.

10.6 Results and Interpretations

We applied our global hue contrast definition on two sample datasets, each composed of two subsets. Each subset was composed of stylistically coherent images (e.g. a set of paintings all made by the same artist, at the same stage in their life) and was selected for their significant similarity to their partner set (e.g. the partner subset consisting of paintings all made by the same artist, at a different stage in their life).

The two set pairs were derived from the work of the painter Pablo Picasso [2], and the animations ‘Finding Nemo’ (Pixar) [4] and ‘Shark Tale’ (Dreamworks) [13]. They are detailed in Table 10.1.

In the case of the painting subsets, we used as source material the highest quality digital reproductions available online. The frames used in the animation subsets were taken directly from the DVD of these movies and were rendered as image sequences of maximum quality and size. The results are here presented together with our interpretations.

The painting subsets (see Fig. 10.6) evidence the artist employing clearly different global hue contrast strategies, yet in both sets these strategies were unusual. Specifically, the following key points were observed:

Table 10.1 The datasets

Sets	Subsets	Size of subsets
Paintings by P. Picasso (1881–1973)	Blue Period paintings (1901–1904)	74 paintings
	Rose Period paintings (1904–1906)	74 paintings
Undersea animations	Pixar’s ‘Finding Nemo’ (2003)	144,513 frames
	DreamWorks’s ‘Shark’s Tale’ (2004)	129,020 frames

- The Rose Period paintings are generally more monochrome than the Blue Period paintings. This is evidenced by the massing of points at the bottom of the plot. A monochrome painting is an example of extreme hue imbalance (see ‘Avoidance of hue imbalance’, in Table 2.1 in Sect. 2.3.3 on p. 25). An example of a monochrome Rose Period painting is shown in Fig. 10.3.
- Unlike the Blue Period paintings, the Rose Period never exhibit full hue variety, otherwise known as hue ubiquity or garishness (see ‘Avoidance of hue ubiquity’ in Table 2.1 in Sect. 2.3.3 on p. 25). This is evidenced by the dearth of points in the top right-hand corner of the plot. An example of a Blue Period painting with full hue variety is shown in Fig. 10.4.
- The Rose Period paintings have significantly more hue antagonism than the Blue Period paintings (as evidenced by the slight massing of the points towards the left of the plot). An example of a Rose Period painting with high hue antagonism is shown in Fig. 10.5.

In both the Blue Period and Rose Period subsets, there are gaps midway in the plots (regions *A* and *B*). These indicate that the artist had actively favoured either configurations of low or high hue antagonism, broadly corresponding to a perceptual ‘likeness’ and ‘oppositeness’ of hue values. Supporting this, the art theorist Rudolph Arnheim said that for a painter: ‘...a configuration of colours will strive either towards contrast or towards assimilation’ [1]. These gaps are at different locations in the two sets, yet are clear in both and are most clearly shown in the antagonism histogram bar on the left of the plot.

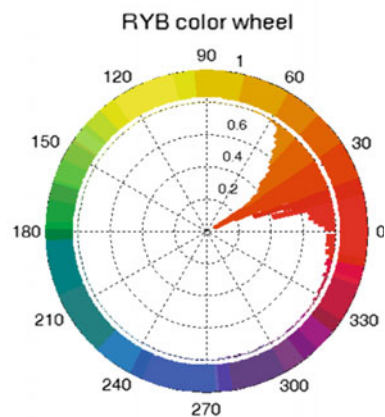


Fig. 10.3 **a** ‘Nude with a Pitcher’, Pablo Picasso, 1906 (*Rose Period*). CC Image courtesy of Lindsey Davis, Flickr. **b** The hue histogram of the painting. Note the monochrome hue (color figure online)

(a)

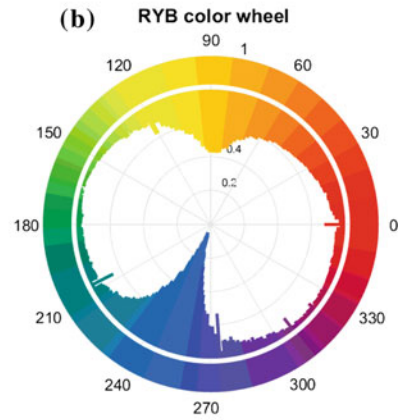


Fig. 10.4 a ‘Women in the Loge’, Pablo Picasso, 1901 (*Blue Period*). CC Image courtesy of Matt Dertinger, Flickr. b The hue histogram of the painting. Note the high hue variety (color figure online)

(a)

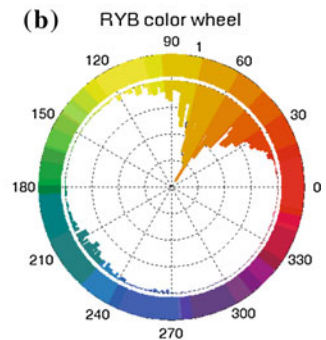


Fig. 10.5 a ‘Family of Saltimbanques’, Pablo Picasso, 1905 (*Rose Period*). CC Image courtesy of Cliff Dertinger, Flickr. b The hue histogram of the painting. Note the apparent hue antagonism (color figure online)

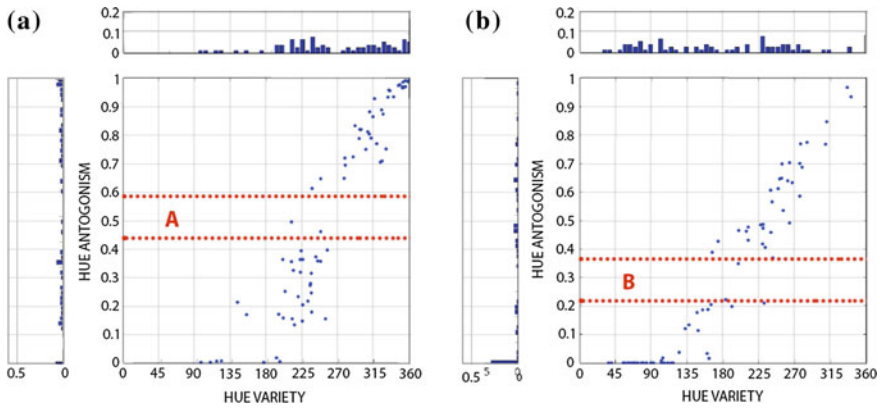


Fig. 10.6 **a** Global hue contrast results of Picasso’s Blue Period. CC Image courtesy of Cliff, Flickr. **b** Global hue contrast results of Picasso’s Rose Period. A total of 74 paintings are examined in each set. Regions ‘A’ and ‘B’ mark apparent gaps, between clusters of high hue antagonism and low

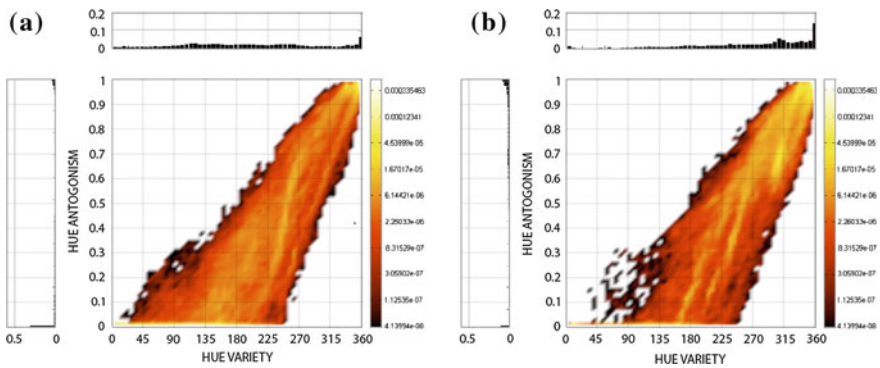


Fig. 10.7 **a** Global hue contrast results of ‘Finding Nemo’ by Pixar studios. **b** ‘Shark Tale’ by DreamWorks studios

The animation set (Fig. 10.7) was composed of a very large number of frames. This resulted in significant colocation of data points. Regions of dense colocation are shown in bright yellow.

Of the global hue contrast of the undersea animation set, the following key points were observed:

1. In both subsets, there is heavy colocation at the top corner of the plot, indicating that the animators had employed full hue variety. However, it is clear that ‘Shark Tale’ employs this strategy more than ‘Finding Nemo’. This colocation extends into the lower–upper regions, indicating a favouring of configurations featuring high hue variety.
2. ‘Finding Nemo’ employs more monochrome (i.e. lower hue antagonism) than ‘Shark Tale’.

3. The bottom-left corners of the ‘Shark Tale’ plot is more thinly populated than that of ‘Finding Nemo’. Frames within this region do not have many distinct hue values, yet also feature as much hue antagonism as their limited hue range will allow. The reason for this difference becomes clear when we consider that the colour of the two main protagonists in ‘Finding Nemo’ is orange who are frequently pictured against a plain blue sea (blue and orange are an antagonistic pair).
4. In both plots, there are yellow ‘rivers’: contiguous regions of heavy collocation. These are oriented slightly obliquely to the dataset. This indicates a biasing of these frames, most of which were derived from continuous sequences within the animations, in favour of low hue opposition.
5. Aside from these differences, both animations present nearly all-possible combinations of variety and antagonism, but it can be generally observed that the distribution of combinations is far more even in ‘Finding Nemo’.

10.7 Evaluation Through a User Study

To test our approach, a user study was performed. From each of the subsets, two images were selected for the proximity of their plotted global hue contrast values. These two images we term the ‘match pair’. A third image was selected, the global hue contrast result of which was not so far away from the match pair as to be evidently different, nor so close as to be easily confused with the pair. The result was three images (an image group), two of which were close in their global hue contrast (as defined by our approach). Thirty of such image groups were presented to 61 undergraduate art students.

For each image group, one of the match pairs was presented to the students as a ‘master’ reference image. The students were asked to select one from the remaining two images for its similarity to the master image. This similarity was to be on the basis of its hue character. They were told to ignore differences in lightness, saturation, content and composition. To prepare them, they were shown four examples of match pairs. They were told nothing of the basis on which our principle operates (variety and antagonism).

To visually evaluate the hue values of an image to the exclusion to all other values is something that even a trained artist would find difficult. The test was therefore a tough one for our procedure. Despite this, the results show that there is significant correspondence between the students’ pairing choices and the pairs derived from our approach.

In Fig. 10.8, a box plot depicting the distribution of successful pairings out of the total 30 candidate image sets is shown. Should the pairing have been entirely random, a median of 15 would have resulted, with a wide spread of values either side (i.e. a large interquartile range). Instead, we see a median of 20 with a small interquartile range. Tellingly, only one student scored less the mid range of 15.

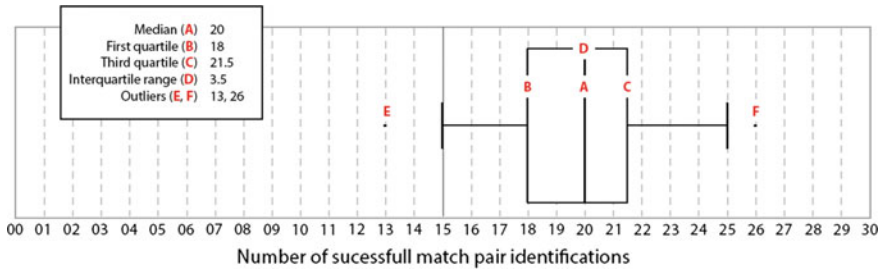


Fig. 10.8 A box plot showing the range of successful pair match identifications per student

10.8 Concluding Remarks

In this chapter, we present the global contrast of hue as being difficult to define, existing global contrast definitions being unequal to its dimensional complexity. Using the work of Johannes Itten as a point of reference, we offer two values: hue variety and hue antagonism. These together define hue global contrast. Enabling the definition of hue variety, we employ Munsell colour space for its ability to define hue in a perceptually uniform manner. Enabling the definition of hue antagonism, we employ RYB colour space for its ability to define antagonistic hue pairs which conform to the artistic norm. A user study supports the validity of our approach.

References

1. Centore, P.: Colour Theory for Painters. <http://www.99main.com/~centore/> (2010)
2. Chevalier, D.: Picasso, the Blue and Rose Periods. Crown Publishers, New York (1991)
3. Feisner, E.A.: Colour: How to Use Colour in Art and Design. Laurence King, London (2000)
4. Finding Nemo: Dir. Stanton A and Unkrich L. Pixar (2003)
5. Ford, A., Roberts, A.: Colour Space Conversions. Westminster University, London (1998)
6. Gossett, N., Chen, B.Q.: Paint inspired colour mixing and compositing for visualization. In: Proceedings of the IEEE Symposium on Information Visualization, pp. 113–118 (2004)
7. Itten, J.: The Art of Color: The Subjective Experience and Objective Rationale of Color. Wiley, New York (1974)
8. Kuehni, R.: Color Space and its Divisions: Color Order from Antiquity to the Present. Wiley-Interscience, Hoboken (2003)
9. Mahnke, F.: Color. Environment and Human Response. John Wiley, New York (1996)
10. Matsuda: Color Design. Asakura Shoten (in Japanese). (1995)
11. Moretti, G., Lyons, P.: Tools for the selection of colour palettes. In: The SIGCHI-NZ Symposium, pp. 1318. New York, USA: ACM (2002)
12. Neumann, L., Nemcsics, A., Attila, C.: Computational Color Harmony based on Coloroid System. <http://en.scientificcommons.org> (2008)
13. Shark Tale: Dir. Bergeron B, Jenson V and Letterman R. DreamWorks (2004)
14. Zhang, X., Constable, M., He, Y.: On the transfer of painting style to photographic images through attention to colour contrast. In: Proceedings of the 4th Pacific-Rim Symposium on Image and Video Technology, pp. 414–421 (2010)

Chapter 11

Interactive Local Hue Contrast Manipulation

Abstract The lightness and saturation of an image may be subject to adjustment for the purpose of aesthetic improvement. Normally, this is in the form of a contrast adjustment. Such an adjustment would work upon the maximum, minimum and mean values of the lightness and/or saturation. The third value in the colour triumvirate is hue, but existing image improvement approaches rarely manipulate this value for the reason of its importance to the semantic reading of an image. Looking to painters for inspiration, we propose a method by which a region of contiguous hue in an image may be subject to one of two contrast adjustments: hue spread and hue compression. These are broadly analogous to adjustments of the relative form of Itten's contrast of temperature. We develop a tool, which employs superpixels, using which a user may segment regions of similar hue, at varying scales. This allows the user to quickly select semantically related regions. These local regions also provide the basic region of operation for three proposed hue operations: (1) hue spread, (2) hue compression and (3) hue shift. We show that when combined with superpixel segmentation, these operations are capable of increasing, decreasing and changing regions of local hue contrast in a manner that is aesthetically agreeable. We demonstrate our framework on a variety of input images and discuss its evaluation from the feedback of several expert users.

Keywords Hue contrast manipulation · Image improvement · Superpixels

11.1 Introduction

Saturation and lightness values can be defined as pixel-wise variations of 'amount', and manipulation of these values acts upon this amount through their average, least and greatest values. Hue can only be measured as an amount if regarded as a range of increasing wavelength along the length of the electromagnetic spectrum. However, a more conventional and perceptually correct way of presenting hue is one where the two ends of this sequence are bent round to join each other, so forming the colour wheel [8]. Ranges within this sequence are organised according to terms, which are perceptually unique and name-based: red, green, blue and yellow. This arrangement

reflects the non-hierarchical nature of the hue range, with no single hue term being greater than another, and such notions as average, least and greatest being invalid. Consequently, hue cannot be operated upon in the same manner as lightness and saturation.

While lightness contrast and saturation contrast are fairly well understood as computational problems [5], hue is unique for the close relationship it has to the meaning of an object [10]. For example, foliage needs to be green and ripe tomatoes need to be red. Edits to the hue of an image risk damaging this semantic relationship. Another problem is that there is little discussion or agreement as to what constitutes hue contrast.

To better understand hue contrast, we turned our attention to painters. A painting is differentiated from a photograph by the degree to which its author exaggerates upon nature, as demonstrated in Chap. 6. on p. 71. In this chapter, we reference two artists, both notable for the liberties that they took with hue.

Edward Hopper (1882–1967) was an American artist notable for the flatness of his colour: featuring regions of very low hue modulation. In Fig. 11.1 is shown how he painted the grass as a contiguous region of single hue.

The paintings of Georges Seurat (1859–1891) were painted in the Pointillist style, and feature dense variations of hue. In Fig. 11.2 is shown how he painted the grass as a contiguous region of many hues.

The targeted hue regions are semantically defined, in this case the green of the grass. This corresponds to the role of hue in human perception, frequently signifying difference of kind [7] (e.g. blue = sky, grass = green). Neither treatment impacts upon the perceived dominant hue, which both present as a yellow/green.

Seurat was employing a technique that classical painters had known for centuries: the ‘breaking’ of colours [3]. This refers to the practice of breaking-up contiguous regions of colour with variations of hue, while preserving the dominant hue. A contiguous region of unvariegated hue would draw the attention of the viewer to the canvas surface of the painting, which is antithetical to the function of a painting,

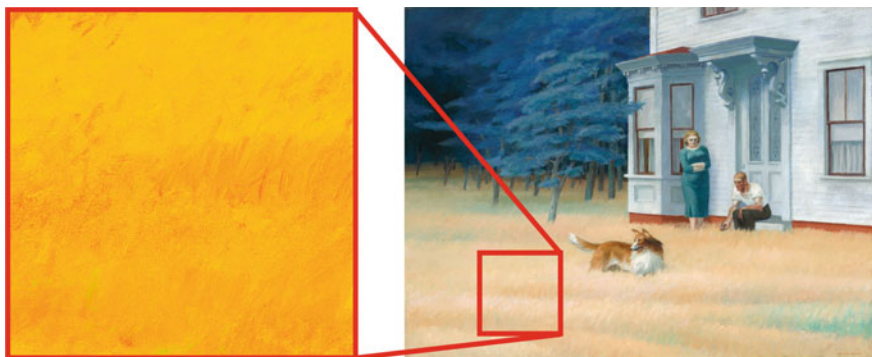


Fig. 11.1 ‘Cape Cod evening’, Edward Hopper, 1939. CC Image courtesy of Irina, Flickr. Note how the grass has been painted with a narrow range of hues (the *inset* has been saturated to make the hue more apparent)

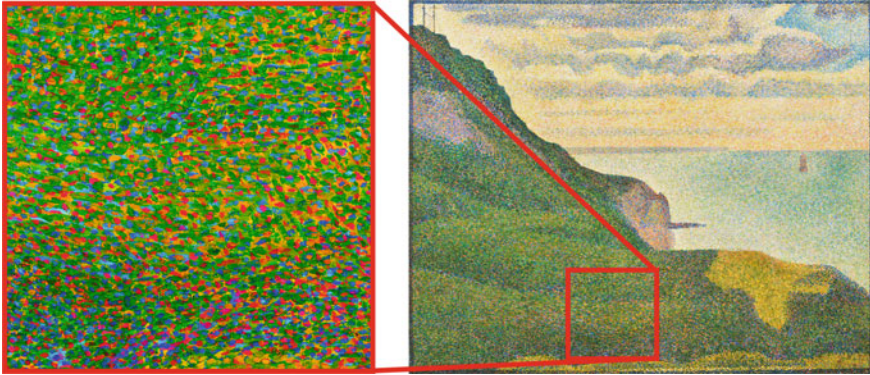


Fig. 11.2 ‘Seascape at Port-en-Bessin, Normandy’ Georges Seurat, 1888. Courtesy National Gallery of Art, Washington. Note how the grass has been painted with a wide range of hues, ranging from *blue* to *yellow*, and even spots of *red* (the *inset* has been saturated to make the hue more apparent) (colour figure online)

which is to create the illusion of depth. Hopper was knowingly working against this tradition when he painted using a single monochrome hue.

Seurat’s approach is also supported by literature. Of the seven colour contrasts described by Itten [4], three relate directly to hue:

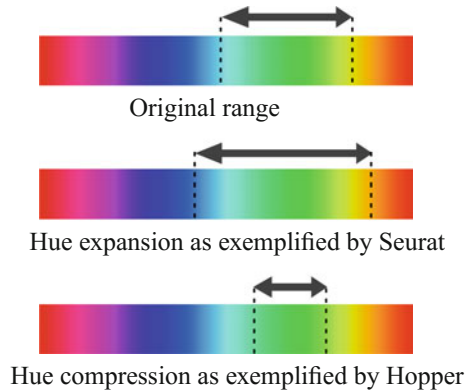
1. Contrast of complements,
2. Contrast of hue and
3. Contrast of warm and cool

Itten’s contrasts are described in full in Sect. 2.3.1 on p. 17.

The contrast of complements refers to a contrast between two colours on opposite sides of the RYB hue wheel. The contrast of warm and cool assumes that colours on the yellow to red-violet side of the wheel are warm, and colours on the violet to yellow-green half of the wheel are cool. In practice, there is little difference between this contrast and the contrasts of complements (complimentary pairs are also warm and cool pairs). The contrast of hue refers broadly to the number of perceptually different hues in an image: a rainbow representing maximum hue contrast.

Contrast of warm and cool also manifests in a relative manner. For example, the warm yellow of a sunflower may be contrasted against the cool yellow of a lemon. It is this contrast that Seurat employed, with the green of the grass being rendered as a range of hues that extend from a warm yellow-green to a cool blue-green. Seurat’s breaking of colours may be understood as a form of relative warm and cool contrast. Such ‘bracketing’ of local hues around a dominant hue may also be framed as an analogous contrast, though this term is more usually employed to signify a global relationship.

Fig. 11.3 A pictogram showing the difference between the range of hues employed by the artists Seurat and Hopper



Seurat’s treatment of hue we may term hue stretching, with that employed by Hopper being hue compression. These are simply represented by the pictograms of Fig. 11.3.

Contribution Inspired by this material, we propose an interactive image editing framework that allows the user to perform hue stretching and hue compression on a selection of hues. As stated, such hue treatment is specific to local regions. To this end, we incorporate a superpixel-based interactive segmentation method that allows the user to easily select image regions at different superpixel scales (see Fig. 11.7). The user may then compress or stretch these selected hues. To this, we add the ability to shift the selected hue range.

11.2 Related Work

Most existing approaches in the aesthetic manipulation of hue concern themselves with global hue harmony according to aesthetic principles. Colour harmony we address in Sect. 2.3.3 on p. 25. Cohen-Or et al. [2] employed a method that uses a template-driven approach to harmonise images according to models developed by Matsuda [6] and Tokumaru et al. [14]. Wang et al. [15] proposed a knowledge-based system which incorporates established colour design rules into a comprehensive interactive framework to aid users in the selection of colours for scene objects. Nishiyama et al. [9] used a similar approach for the automatic aesthetic quality classification of photographs. Most other approaches that do not involve colour harmony were example-based (e.g. [11–13, 16]), requiring an appropriate reference image in order to function.

These approaches rely on established harmonic distributions and do not give as much flexibility to hue manipulation as our proposed method. Moreover, any attempt to impose hue harmony rules upon an image runs the risk of changing the semantic content of that image.

11.3 Hue Manipulation Framework

In this section, we present the core concepts of our approach.

11.3.1 System Overview

Our framework combines two main operations: image segmentation and hue adjustment. Figure 11.4 shows a screenshot of our software framework. Our method provides two significant operations for hue: hue compression and hue stretch. To this, we add the ability to shift a range of hue (i.e. to rotate it around the hue wheel). On its own, hue shift is not a significant operation, being trivial to effect in most image manipulation software; however, we found that in conjunction with a hue shift or compression, useful results may be achieved. These will be shown in Sect. 11.4. The framework of the system is shown in Fig. 11.5.

We chose to perform our colour operations in the HSI colour space which presents hue as a separate channel. HSI was derived from RGB colour space in the following manner:



Fig. 11.4 The user interface developed for the hue contrast manipulation

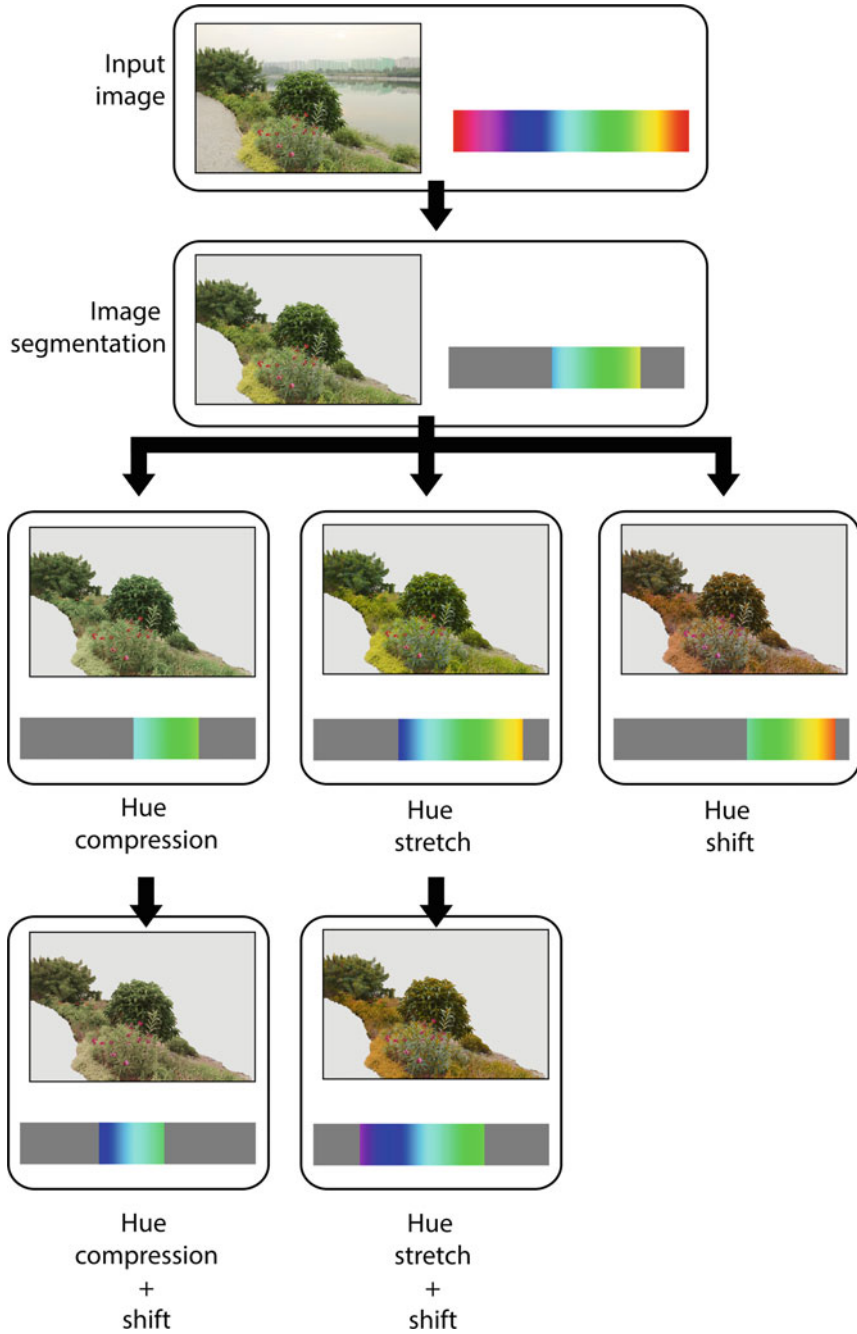


Fig. 11.5 The framework of the proposed system

$$\alpha = \frac{1}{2}(2R - G - B), \beta = \frac{\sqrt{3}}{2}(G - B), H = \text{atan2}(\beta, \alpha),$$

$$m = \min(R, G, B), I = \frac{1}{3}(R + G + B), S = \begin{cases} 0 & \text{if } I = 0, \\ 1 - \frac{m}{I} & \text{otherwise} \end{cases}, \quad (11.1)$$

where R, G, B are red, green and blue values in the RGB space and H, S, I are the hue, saturation and brightness values in the HSI space, respectively.

For brightness and saturation values equal to zero, the value of hue is not meaningful. We therefore only process the pixels with saturation (S) and brightness (I) larger than a certain threshold. Here, we define the threshold to be $t_s = 0.10$ and $t_i = 0.15$ (assuming that S and I are both within the range $[0,1]$), respectively.

We discuss each component of our framework in detail in the following sections.

11.3.2 Segmentation

As discussed in Sect. 11.1, hue manipulation is best performed as a local regional adjustment. This means that it will be necessary for the user to segment the image for hue processing. To this end, we use a segmentation strategy based on superpixels. In particular, we use the simple linear iterative clustering (SLIC) method [1] for superpixel segmentation that is spaced using colour information. The result is a group of superpixels of similar hue. This selection may be contiguous, as in the green of a bush, or non-contiguous, as in many flowers of similar hue. This selection is in agreement with human perception, since the items that we are considering are regions defined by likeness of hue and location, and not individual pixels. The user selects regions by simply drawing a scribble through the superpixels they want to select. Moreover, our system allows the user to control the size of the superpixels through a simple parameter selection. Larger superpixels increase the speed in which objects can be selected, while smaller superpixels provide access to finer grain regions. Examples of different superpixel segmentation are shown in Fig. 11.6.

The resulting superpixel segmentation will provide the basic unit of hue manipulation, being applied to each superpixel independently.

11.3.3 Hue Histogram Selection

Our tool allows users to manipulate a range of hue in a selected region. To isolate the hue range, the user can select a pixel in the image the hue value of which will be shown on the colour wheel. A default interval centred at that hue value will be selected for the next step of hue contrast manipulation (we set 15° distance from both sides of the indicated hue value by default), since our manipulation operations are processed for a range of hues. The user can also manually adjust the two sides

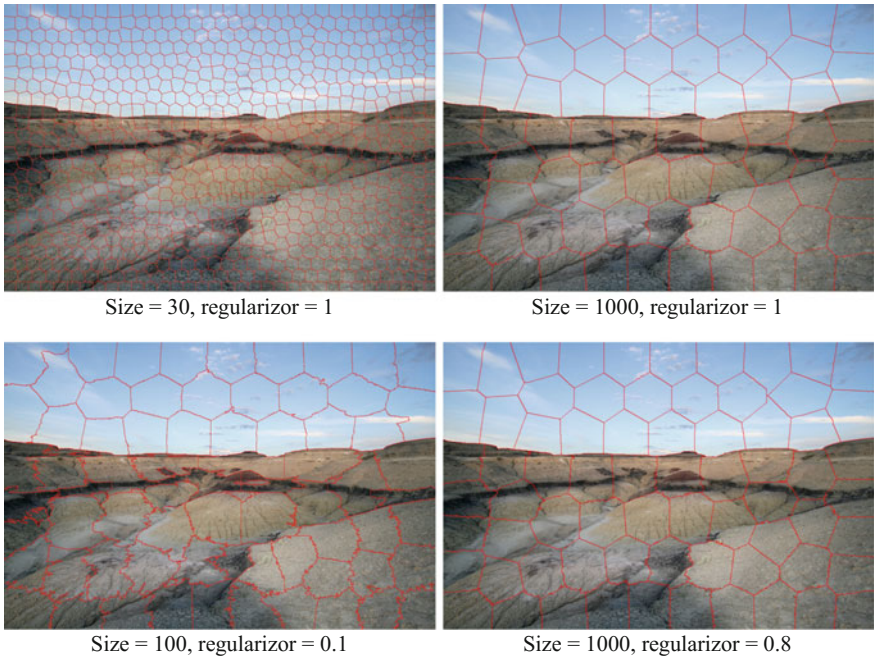


Fig. 11.6 Effect of segmentation parameters: size and regularizer

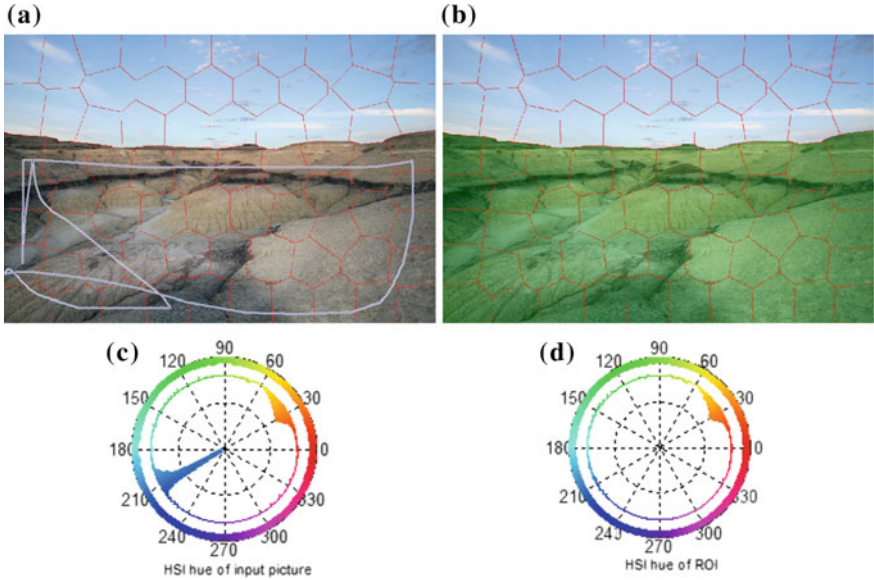


Fig. 11.7 **a** The user selects a range of superpixels by drawing a scribble around them. **b** Selected superpixels are highlighted in *green*. **c** The hue histogram of the whole image. **d** The hue histogram of the selected region (colour figure online)

of the interval being processed by adjusting the position of the starting hue (denoted as A) and ending hue (denoted as B) on the hue wheel.

Our system will automatically calculate the hue histogram of the selected region and denote it in the hue ramp using the markers A and B , which facilitates the next step of hue contrast manipulation. A local hue contrast manipulation result is shown in Fig. 11.7 (by the algorithms in Sect. 11.3.4). Following the automatic selection of A and B , users can adjust markers to obtain accurately the part of the histogram they wish to process. This can also be used to filter the part of the histogram they wish to keep untouched. Through the use of superpixels in this way, we can locally manage hue in a meaningful manner.

11.3.4 Hue Contrast Manipulation Operators

After selecting a region of interest (or the whole image in the case of a global contrast manipulation on an image of a globally predominant hue), users can decide to employ one of the three designed operators: hue compression, hue stretch or hue shift. We provide two categories of operations: a scale parameter and an adjustment interval (see Fig. 11.8).

As indicated in Sect. 11.3.3, markers A and B denote the two sides of the region that will be processed. The resulting interval is denoted by the markers A' and B' . Stated generally, the hue axis is considered as being infinitely long with a periodicity of 360° , and $A < B$, $A' < B'$ on the hue axis (e.g. Fig. 11.8). These 3 operators refer to the centre of interval which is $\frac{A+B}{2}$.

In the case of hue contrast manipulation by the setting of scale parameters, for hue compression and hue stretch, users can input a scale t that reflects the degree to which the selected section of the hue histogram is compressed or stretched. The resulting $[A', B']$ is calculated as $A' = \frac{A+B}{2} - t\frac{B-A}{2}$ and $B' = \frac{A+B}{2} + t\frac{B-A}{2}$. Moreover,

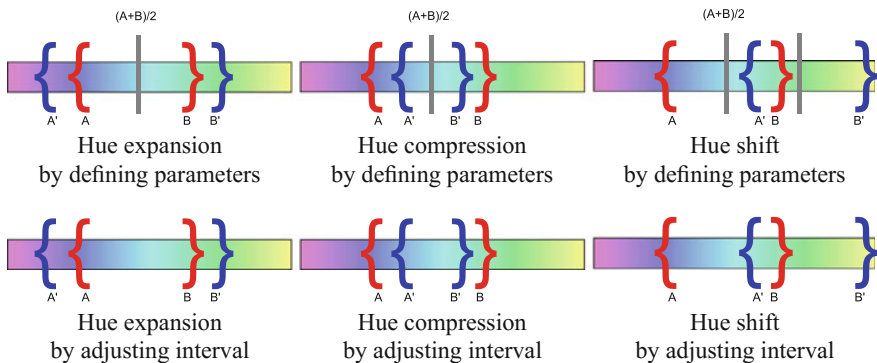


Fig. 11.8 Effect of hue manipulation by defining parameters and adjusting interval

to make the two operations meaningful, t needs to be nonzero for a hue compression and to be of a value such that $B' - A'$ will not be larger than 360° for a hue stretch. For hue shift, users can input a shift amount C reflecting the colour that they want to shift the selected region to. The resulting $[A', B']$ becomes $[A + C, B + C]$.

For a hue manipulation achieved by adjusting the interval $[A, B]$, users input the positions of A' and B' directly on the hue wheel. A 1-D Example for the 3 operators is shown in Fig. 11.8. Performing hue contrast manipulation by the setting of scale parameters gives users a quantitative control over the result, whereas performing hue contrast manipulation by adjusting the interval gives users the possibility to fine tune the result.

Given the hue values $H(x, y)$ of pixel (x, y) defined in the region of interest whose histogram is bound by A and B , hue contrast manipulation aims to transform the current histogram of the hue function $H(x, y)$ within $[A, B]$ to a new histogram of interval $[A', B']$. We update the hue values $H'(x, y)$ using the following rule:

$$H' = \frac{B' - A'}{B - A}(H - A) + A'. \quad (11.2)$$

The *modulo* operation is needed to ensure that the final hue value H' is located within a meaningful range $[0, 360]$.

11.4 Experimental Results

The following shows several results from our software. In Sect. 11.5, we provide feedback from expert users on their experience with hue operations and their opinions of our software.

11.4.1 Hue Compression and Hue Stretch

Hue compression has the aesthetic effect of lessening the difference of hue within local regions (Fig. 11.9d). This is very different from that which can be achieved with the tool in stretch mode (Fig. 11.9c). Using a hue compression, the naturally variegated appearance of a tropical jungle can be compressed so that it more closely resembles the monochromatic hue of a European forest.

In practice, it was found that our approach was particularly useful when applied to scenes containing foliage. The reasons for this is that human vision is particularly sensitive to greens, there being a greater range of greens perceivable than any other hue. Another example of foliage hue adjustment is shown in Fig. 11.10.

The ideal image for the hue stretching and compression is one where there is existing hue variety within the target region. This is in line with expectations, hue stretching and compression being conceptually impossible within a region of entirely

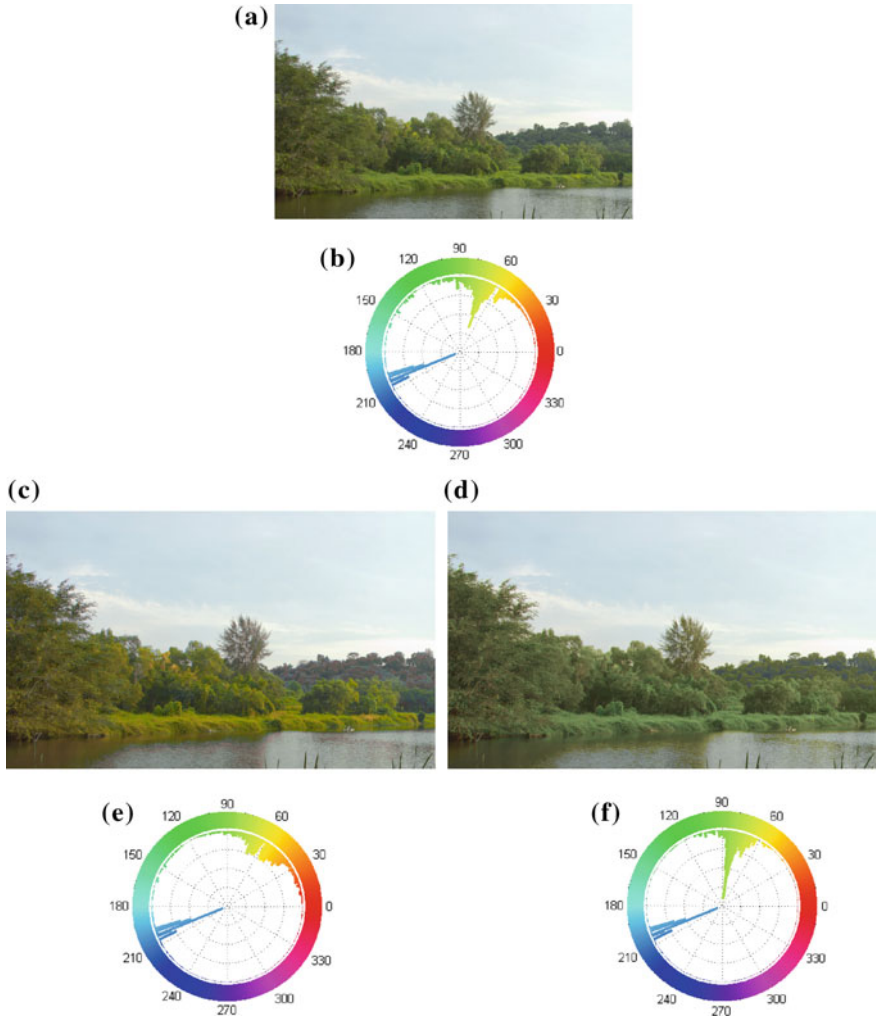


Fig. 11.9 **a** The source image. **b** Histogram of source image. **c** Detail of hue compression result. **d** Detail of hue stretch result. **e** Histogram of hue compression result. **f** Histogram of hue stretch result

homogenous hue. Photographs featuring foliage, rocks, soil and other natural phenomenon are excellent examples of scenes with strong local hue variation that can be amplified through hue stretching. Most man-made materials (plastic, fabric, painted surfaces, etc.), being generally homogeneous in their hue values, do not make good candidates for our approach.

Generally, good hue stretch results were possible if the target hue range did not span more than one hue term and if the degree of stretch did not also span more than one hue term on either side of the hue spectrum for one object or group of similar

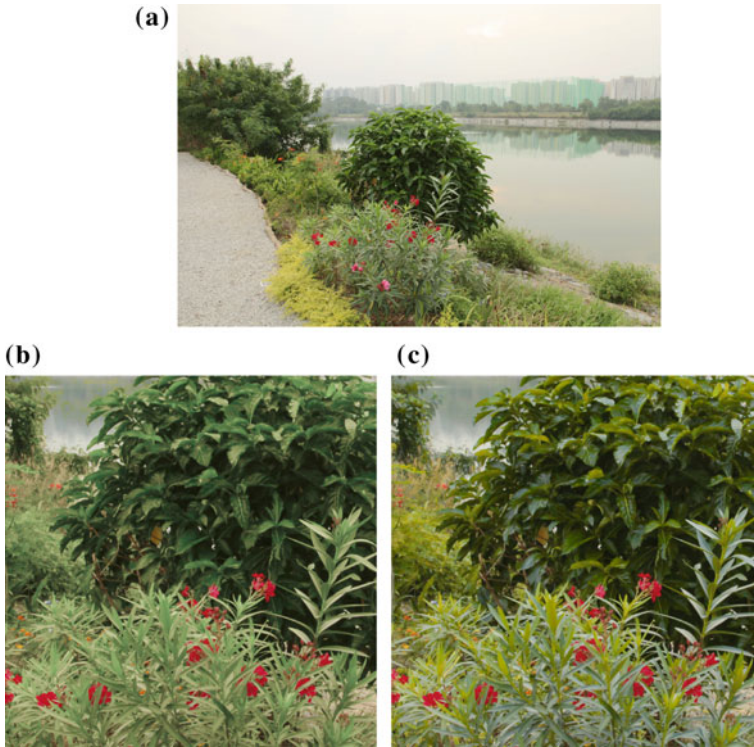


Fig. 11.10 **a** The source image. **b** Detail of hue compression result. **c** Detail of hue stretch result. **d** Detail of hue compression result. **e** Histogram of hue stretch result. **f** Histogram of hue compression result

objects. Greater spread sometimes produced unnatural-looking results. This is again in line with expectations regarding the importance of hue to the semantic reading of an image.

11.4.2 *Combining the Three Operations*

In Fig. 11.11 is shown a hue compression performed on a wide range of hues: the many different warm hues of tulips. Interestingly, these hues are not contiguous, but all involve semantically connected objects. The hue compression renders all the flowers as one dominant hue. This compression is followed by a hue shift, which changes the dominant hue.



Fig. 11.11 **a** The source image. **b** A hue compression. **c** A hue compression followed by a hue shift.

11.5 Discussion and User Feedback

There is a number of image editing software solutions that allow hue manipulation. For example, Adobe Photoshop, Affinity Photo and GIMP enable the user to change the hue of an image through hue shifting. These are generally very hard to control and can easily impart an unnatural look to an image. More specific software solutions exist. HueShifter¹ is a tool created specifically to change the hue of a region of an image in a manner that does not affect its perceptual lightness.

In general, such software offers the ability to manipulate hue in the following ways: shift (i.e. rotated a hue or range of hues around the hue wheel), replacement, saturation or desaturation. Hue stretching and compression is not offered. Compared to existing software, our framework is unique in its coupling of segmentation and the three hue manipulations: compression, stretching and shifting using a single unified interface.

By the selective replacement of many hues, something that might resemble a hue stretch or compress may be achieved. However, the results are very difficult to control and invariably very noisy. Furthermore, they are conceptually very different to our approach. In Fig. 11.12 is shown the closest that a Photoshop professional could achieve to the result shown in Fig. 11.11c. Better results required the use of multiple operations, combined with complex layer masks.

Feedback was gathered from a group of ten professionals in the film-making and animation industries. These professionals were used to using a wide range of professional software, much of it developed in-house and unavailable to the general public. Among this group was one who supervised an Oscar award winning visual effects team and another whose work was recognised with a prestigious Annie award. They therefore constitute a very capable group of users who operate at the very highest level of professional practice. The principles behind the tool were described to them, and a sample of results was shown. They were asked whether they thought the tool would be useful and whether they could reproduce the same effects using existing tools. All of them were convinced of the usefulness of our tool and that its function was not naturally served by any existing solutions. Results achieved using such methods usually require masking before they are usable. Additionally, hue expansion proved

¹<http://www.bergdesign.com/hueshifter/>.

Fig. 11.12 The closest result that we could achieve in a single operation to the result shown in Fig. 11.11c. A complex compound operation might have achieved better results at the cost of a noisy result



to be more achievable than hue compression, though the results achieved were not as subtle as those achieved using the proposed method.

Others thought that in the right circumstances, a hue replace would suffice if moderated with a mask. If a slight hue spread or compression was needed, then it could sometimes produce good results. Neither of the suggested approaches were as intuitive to manage as ours.

11.6 Concluding Remarks

We address local hue contrast as an ill-understood phenomenon, being outside of the scope of existing digital image manipulation software. Using the work of painters as a point of reference, we develop an image editing framework to adjust the aesthetic quality of an image by manipulating its local hue contrast. Modifications made to hue with existing photoediting tools can easily impart an unnatural look to the image. Considering that hue contrast is poorly understood and non-trivial to affect, we propose three hue contrast manipulation operations for colour manipulation. To realise our idea, we developed a user interaction tool that incorporates image segmentation to provide the user the possibility to deal with hue in a flexible and powerful way. The before and after values of the hue adjustment are presented in an intuitive and easy-to-understand manner. We demonstrate the effectiveness of our methods by comparing the results with those produced using Photoshop.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intel.* **34**, 2274–2282 (2012)
2. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. In: *ACM SIGGRAPH*, pp. 624–630 (2006)
3. Couture, T.: *Mthode Et Entretiens D’atelier* (1867)
4. Itten, J.: *The Art of Color: The Subjective Experience and Objective Rationale of Color*. Wiley, New York (1974)

5. Majumder, A.: Perception-based contrast enhancement of images. *ACM Trans. Appl. Percept* (2007)
6. Matsuda, Y.: *Color design*. Asakura Shoten (1995)
7. Nault, W.H.: *Childrens Map Reading Abilities*. Geographic Society of Chicago, Newsletter, III (1967)
8. Newton, I.: *Newton, Isaac. 1730. Opticks* (1979)
9. Nishiyama, M., Okabe, T., Sato, I.: Aesthetic quality classification of photographs based on color harmony. In: *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 33–40 (2011)
10. Ostergaard, A.: Some effects of color on naming and recognition of objects. *J. Exp. Psychol. Learn. Mem. Cogn.* **11**, 579–587 (1985)
11. Pitie, F., Kokaram, A., Dahyot, R.: N-dimensional probability density function transfer and its application to color transfer. In: *IEEE International Conference on Computer Vision* (2005)
12. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Graph. Appl.* **21**, 34–41 (2001)
13. Tai, Y.W., Jia, J., Tang, C.K.: Soft color segmentation and its applications. *IEEE Trans. Pattern Anal. Mach. Intel.* **29**, 1520–1537 (2007)
14. Tokumaru, M., Muranaka, N., Imanishi, S.: Color design support system considering color harmony. In: *Proceedings of IEEE International Conference on Fuzzy System* (2002)
15. Wang, L., Giesen, J., McDonnell, K., Zolliker, P., Mueller, K.: Color design for illustrative visualization. *IEEE Trans. Vis. Comput. Graph.* **14**, 1739–1754 (2008)
16. Zhang, X., Chan, K.L., Constable, M.: Atmospheric perspective effect enhancement of landscape photographs through depth-aware contrast manipulation. *IEEE Trans. Multimed.* **16**, 653–667 (2014)