Transfer of Aesthetic Values in Paintings to Photographs

Xiaoyan Zhang
School of Electrical and Electronic Engineering
2014
Transfer of Aesthetic Values in Paintings to Photographs

Xiaoyan Zhang

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University in partial fulfillment of the requirement for the degree of Doctor of Philosophy

2014
To my parents and
Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my supervisor Professor Kap Luk Chan and my co-supervisor Professor Martin Constable, who have supported me throughout my PhD study with their patience and knowledge, and encouraged me to explore my own way of research. During these years of studying under their supervision, I have learned a lot. I have been greatly impressed by their professional attitude towards research and teaching.

I would also like to acknowledge the seed fund and Ph.D. grant from the Institute for Media Innovation (IMI), Nanyang Technological University, Singapore. I am honored to be studying in IMI. IMI has continuously organized research seminars and invited renowned researchers around the world to give seminars and workshops, which have benefit all of us. I appreciate IMI for providing excellent environment and computing facilities for my research.

I would like to thank Junyan Wang for his valuable suggestions and discussion on various research problems in the process of my research. I am also indebted and thankful to Wei Huang, Li Wang, Bing Wang, and Tiandi Duan, among many other kind people who have helped me and taught me a lot.

Finally, I would like to thank my parents and my husband Wanxing Zhan. Without their support, this thesis would not be possible.
Contents

Acknowledgments i

Contents ii

Summary vi

List of Figures viii

List of Tables xxii

List of Abbreviations xxiii

List of Notations xxiv

1 Introduction 1

1.1 Background 1

1.2 Motivation 3

1.3 Objective 5

1.4 Major Contributions of the Thesis 7

1.5 Organization of the Thesis 9
2 Image Attributes and Aesthetic Aspects

2.1 Image Attributes .............................................. 11
   2.1.1 Color attributes ..................................... 11
   2.1.2 Geometric attributes ................................ 15

2.2 Aesthetic Aspects ............................................. 16
   2.2.1 Contrast .............................................. 17
   2.2.2 Atmospheric perspective ............................ 27
   2.2.3 Vignetting ........................................... 28
   2.2.4 Composition ......................................... 29

2.3 Concluding Remarks ........................................... 30

3 Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation 31

3.1 Introduction ............................................... 32

3.2 Related Work .............................................. 36

3.3 Paintings vs Photographs ................................. 38
   3.3.1 Depth planes ....................................... 38
   3.3.2 Contrast in paintings and photographs .......... 40

3.4 Contrast Manipulation ...................................... 42
   3.4.1 Reference selection ................................ 43
   3.4.2 Depth-aware contrast mapping ................... 45

3.5 Experiments and Discussion .............................. 55

3.6 Concluding Remarks ....................................... 74

4 Regional Contrast Manipulation for Portrait Photograph Enhance-
ment 77

4.1 Introduction ........................................ 78
4.2 Related Work .......................................... 81
4.3 Paintings vs Photographs .............................. 83
4.4 Example-based Portrait Photograph Enhancement .......... 86
  4.4.1 Reference selection ............................ 88
  4.4.2 Contrast mapping ................................ 92
4.5 Experiments and Discussion .......................... 96
4.6 Concluding Remarks ................................ 113

5 Vignetting Effect Manipulation 114
  5.1 Introduction ...................................... 114
  5.2 Related Work ..................................... 116
  5.3 Vignetting Effect in Paintings ...................... 117
  5.4 Transfer of Painter-style Vignetting to Photographs .......... 121
    5.4.1 Lightness weighting pattern and blending function .... 123
    5.4.2 Lightness weighting pattern correction .............. 127
    5.4.3 Content-aware interpolation ........................ 131
    5.4.4 Local contrast restoration ....................... 135
  5.5 Experiments ...................................... 136
  5.6 Concluding Remarks ................................ 142

6 Composition Improvement for Portrait Photographs 145
  6.1 Introduction ...................................... 146
  6.2 Related Work ..................................... 147
6.3 Proposed Composition Improvement Method ...............151
  6.3.1 Graph matching ....................................154
  6.3.2 Space cropping ......................................156
6.4 Experiments .............................................157
6.5 Concluding Remarks ....................................163

7 Conclusions and Recommendations ..........................167
  7.1 Conclusions ............................................167
  7.2 Recommendations for Future Research .................169

Author's Publications ........................................172

Bibliography ..................................................174
Summary

The current state-of-the-art research in computational vision and graphics offers many tools to change the look and feel of images. However, few academic studies are looking at how such operations can be deployed to improve the visual appeal of a photograph from an artistic viewpoint. This thesis focuses on studying the underlying attributes of visual appeal with the aid of pre-existent paintings, and developing algorithms and methods that facilitate the transfer of the organization of image attributes contributing towards aesthetics in paintings to photographs.

One of the aesthetic considerations in paintings is the way the tone and color values have been organized with respect to various notions of contrast. In this thesis, the following contrasts are reviewed and discussed: global contrast for the creation of global mood, local contrast to highlight the local details, contrast across depth planes to enhance depth perception, regional contrast and center-corner contrast for the focusing of attention. The thesis investigates these contrasts in landscape and portrait paintings. Composition, which is subject to geometric attributes such as target positioning and spacing, is also an important aesthetic aspect of painting. The thesis presents algorithms and methods that improve the depth perception through transfer of contrast across depth planes based on the principle of atmospheric perspective effect, and improve the focus of attention through figure-ground regional contrast transfer and painter-style vignetting transfer. The thesis also covers the cropping of a portrait photograph to improve its visual appeal. Example-based approach is used in the thesis to handle the diversity of content in paintings and photographs.

One of the contributions in this thesis addresses the enhancement of the atmospheric perspective effect in landscape photographs by the manipulation of depth-aware lightness and saturation contrast values. Based on a study which shows that
saturation contrast and lightness contrast inter- and intra- depth planes in paintings are purposefully organized and obey the principle of atmospheric perspective effect, an example-based method is developed to transfer such contrast organization in an example painting to a landscape photograph in an attempt to improve its visual appeal and illusion of depth. This contrast mapping is formulated as an optimization problem that simultaneously considers the desired inter-contrast, intra-contrast, and specific gradient constraints. Experimental results demonstrate the improved depth illusion and visual appeal in photographs.

For portrait painting, the regional contrast is the main attribute that makes the figure, especially the face, stands out in the painting. As another contribution of this thesis, a method for the manipulation of the regional contrast in portrait photographs is developed which makes use of pre-modern portrait paintings as aesthetic examples. The contrast organization in the example painting is transferred to the photograph by mapping the inter- and intra- regional contrast values. A novel piecewise nonlinear transformation curve is proposed to achieve the contrast mapping. Experimental results demonstrate that by using this proposed method, the visual appeal of portrait photographs are effectively improved and the face and the figure become more salient.

Vignetting is an effect that is manifested by being clear in the center and fading off at the edges by reducing the image's brightness or saturation at the periphery. The thesis analyzes the vignetting effect in paintings and photographs. The observation of the difference shows that the vignetting effect in paintings is more purposely presented. An algorithm is then explored to apply the lightness weighting derived from the vignetting effect in an example painting to a photograph.

The thesis also studies the composition of portrait paintings and develops an algorithm to improve the composition of a portrait photograph based on an example portrait painting. Pose, face direction and space around the target figure are the main elements considered in portrait composition. Space cropping technique is used to optimize the composition of the photograph based on the locations of body parts in the selected example painting.
List of Figures

1.1 Selected scenes from the movie “Star Wars”. This film was made in 1977, before the age of digital film making, but was digitally re-mastered in 1997. Images in the top row are from the original movie “Star Wars”, images in the second row are the corresponding scenes digitally re-mastered in 1997. 4


1.3 (a) Global contrast, (b) Local contrast, (c) Depth-aware contrast, (d) Center-corner contrast, (e) Regional contrast. 6

2.1 RGB color model. The figures are from [1]. (a) RGB color model, (b) RGB 24-bit color cube. 11

2.2 HSL and HSV cylinder models. The figures are from [2]. (a) HSL cylinder, (b) HSV cylinder. 12

2.3 Difference of L channel in HSL and CIELAB spaces. (a) Hue ramp and L channel of HSL space, (b) Hue ramp and L channel of CIELAB space. 14
2.4 Comparing image editing in LCH and HSL color spaces. (a) Original image, (b) L in LCH space, (c) L in HSL space, (d) Gamma correction of L in LCH space, (e) Gamma correction of L in HSL space. .................................................. 15

2.5 (a) Paintings, top: “A Coign of Vantage”, by Sir Lawrence Alma-Tadema, 1895, bottom: “The Poppy Field” by Vincent Van Gogh, 1888. (b) Normalized histograms of lightness, (c) Normalized histograms of saturation. .............................................................. 20

2.6 (a) Original photograph, (b) Histogram matching result using the painting in the top row in Figure 2.5 as reference, (c) Histogram matching result using the painting in the second row in Figure 2.5 as reference. (d) and (e) The comparison of the lightness and saturation distributions of the original photograph in (a), the result in (b) and the corresponding reference. (f) and (g) The comparison of the lightness and saturation distributions of the original photograph in (a), the result in (c) and the corresponding reference. .................... 21

2.7 (a) Landscape painting with outlined depth planes, (b) Left: Francisco Goya (1746 - 1828), “Portrait of King Ferdinand VIII” 1803, Right: the portrait divided into depth planes. .......................... 22

2.8 One example photograph with atmospheric perspective effect from [3]. ........................................................................................................ 27

2.9 The vignetting effect. (a) Original photograph, (b) Vignette in Photoshop®. .................................................................................... 29
3.1 An example result of atmospheric perspective effect enhancement using the proposed method. (a) Input photograph, (b) Reference painting, (c) Result. .......................... 32

3.2 Box plots\(^1\) of lightness intra-contrast values (a) and inter-contrast values (b) in paintings and photographs. .......................... 41

3.3 Box plots of saturation intra-contrast values (a) and inter-contrast values (b) in paintings and photographs. .......................... 42

3.4 The pipeline of the proposed method. .......................... 43

3.5 Reference recommendation. (a) Input w/ mask, (b) Recommended references. .......................... 44

3.6 The inter-contrast mapping curves. (a) \(m'_s > m_s\), (b) \(m'_s < m_s\). .......................... 48

3.7 Comparison of mapping curves. (a) Input, (b) Reference, (c) Shifting, (d) Linearly scaling, (e) Gamma-circular, (f) Linear-circular. 
\(q=0.5\) in (c)(d)(e)(f). .......................... 49

3.8 Comparison of mapping curves. (a) Original, (b) Gamma correction, (c) Gamma-circular, (d) Linear-circular. \(m'_s > m_s\). .......................... 50

3.9 (a) Photograph, (b) Edge by Canny detector, (c) Salient edges. .......................... 52

3.10 The influence of boundary area in intra-contrast mapping. (a) Input w/ mask, (b) Reference w/ mask, (c) W/ boundary, (d) W/o boundary. .......................... 52

3.11 (a) Reference w/ mask, (b) Input w/ mask, (c) Inter-contrast mapping result, (d) Result using \(E_d\) and \(E_g\) in Eq. (3.1), (e) Result with \(E_s\). .......................... 55
3.12 The influence of the parameter $q$. (a) Input w/ mask, (b) Reference w/ mask, (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) $q_L=1.0$, $q_C=1.0$.

3.13 The mean values of the four depth planes when using different parameters $q_L$ and $q_C$ in Figure 3.12.

3.14 The contrast manipulation results based on three selected references in Figure 3.5. The input with mask is in Figure 3.12(a). Images in the first row are the references with masks. 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.

3.15 The average matching error of inter-contrast and intra-contrast. The bars represent +/- one standard deviation. The inter-contrast is normalized to the range [-1,1], and the intra-contrast is normalized to the range [0,1].

3.16 Comparison with other contrast adjustment methods. (a) and (k) Input w/ mask, (b) and (l) Reference w/ mask, (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 optimization in [4], (i) and (s) Proposed ($q_C=0$), (j) and (t) Proposed ($q_C=q_{rc}$).
3.17 Comparison with depth-aware methods. (a) Input w/ mask, (b) Reference w/ mask, (c) HM in depth plane, (d) Method in [5], (e) Proposed. 

3.18 Comparison with the physical model used in [6]. The input photograph and reference are those used in Figure 3.1. (a) Estimated depth, (b) Result by [6], (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).

3.19 Comparison with style enhancement in [7]. (a) Input w/ mask, (b) Result in [7], (c) Proposed method. Image in (b) is the reference.

3.20 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 w/ mask, (b) Reference w/ mask, (c) Result.

3.21 Aesthetics assessment. (a) $M_b$, (b) $M_j$.

3.22 Ratio of aesthetic score.

3.23 Results using different partitions. (a) Input photograph, (b) Mask 1, (c) Mask 2, (d) Reference w/ mask, (e) Result using mask 1, (f) Result using mask 2.

3.24 An example demonstrating the limitation of the proposed method. (a) Input photograph, (b) Reference painting, (c) Result, (d) Close-up Region.

3.25 The comparison of pixel value distribution among the original photograph, reference and enhancement result in Figure 3.24. (a) Lightness histogram, (b) Saturation histogram.
3.26 Experimental results on night scene photographs. (a) Night scene paintings painted with moon light, (b) Night scene photographs. (c) Results using proposed method. 74

3.27 More results. The first, third and fifth rows: original photographs. The second, fourth and sixth rows: results using proposed method. 75

4.1 One of the results produced by the proposed method. (a) Original photograph, (b) Reference painting, (d) Enhancement result. 78

4.2 (a) “Portrait of Mariana Waldstein”, by Francisco de Goya (1746-1828). (b) “Mrs. Peter William Baker” by Thomas Gainsborough (1727-1788). (c) and (d) are photographs. (d) is cropped from an image in MIT-Adobe FiveK Dataset [8]. 80

4.3 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. 85

4.4 Framework of the proposed method. 87

4.5 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 initialization window for GrabCut. The blue and purple lines are the FG and BG brushes for refining the FG/BG segmentation. 89

4.6 The graph model. 90

4.7 Contrast mapping. (a) Transformation curve, (b) Convergence of maximum, minimum and mean values. 94
4.8 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. 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 $\alpha_l = 0.6, \alpha_s = 0.6, \beta_l = \beta_s = 0.5$. 97

4.9 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 for the two photographs and the corresponding results. 98

4.10 Influence of $\alpha, \beta$. (e)-(i): $\alpha_s = 0.4, \beta_l = \beta_s = 0.5$. (j)-(n): $\alpha_l = 0.5, \alpha_s = 0.4, \beta_s = 0.5$. 99

4.11 Comparison with other methods. (a) Reference paintings, (b) Input photographs, (c) Results by CAE in [9], (d) Results by AVE method, (e) Results by the proposed method. The photograph in the fifth row is from [9]. 101

4.12 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. 102

4.13 Comparison with the method in Chapter 3. (a) Original, (b) Reference, (c) Result by the proposed method, (d) Result by the method in Chapter 3. 105
4.14 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. ..................... 106

4.15 More experimental results. The first and fourth rows: reference paintings. The second and fifth rows: original photographs. The third and sixth rows: results. ..................... 107

4.16 Application using RGBD images. The depth from the video is the depth layer obtained by optical flow in [10]. ..................... 108

4.17 FG extraction from RGBD image. (a) RGB image, (b) Depth, (c) k-means clustering, (d) Mask of extracted FG. ..................... 110

4.18 FG extraction from RGBD image. (a) RGB image, (b) Depth, (c) k-means clustering. ..................... 110

4.19 (a) RGB image, same with Figure 4.18(a), (b) Extracted FG mask by removing the ground plane detected by the method in [11], (c) Extracted FG mask by removing the ground plane detected by the method in [12], (d) Reference painting, (e) Result using mask in (b), (f) Result using mask in (c). ..................... 111

4.20 More results for the application in RGBD images. (a) RGB images, (b) Depths, (c) Extracted FG masks, (d) Reference paintings, (e) Enhancement results. ..................... 112

5.1 The vignetting effect in paintings. (a) “Sacramento River Valley”, Albert Bierstadt, 1872. (b) “Lady Anne Hamilton later Duchess of Donegall”, Thomas Gainsborough, 1777-1780. ..................... 115
5.2 (a) The average lightness image of 920 paintings, (b) The average lightness image of 920 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) and (d) are the contrast enhanced version of (a) and (b). (e) and (f) are the plots of the average values of respectively 800 columns and 600 rows for the two average lightness images in (a) and (b).

5.3 (a) The average saturation image of 920 paintings, (b) The average saturation image of 920 photographs, (c) and (d) are the contrast enhanced version of (a) and (b).

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

5.5 (a) The average lightness of 117 full body portrait paintings, (b) The average lightness of 146 half body portrait paintings. (c) and (d) are the contrast enhanced version of (a) and (b). Portrait paintings with simple black background are excluded from the average lightness calculation.

5.6 The paintings in (b) and (d) are the top ranked paintings which are similar in the geometric structure with the photographs in (a) and (c) respectively.

5.7 The extracted lightness bias fields for the three paintings in Table 5.1.
5.8 The 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, (c) Lightness weighting, (d) Darken mode, (e) Multiply mode, (f) Overlay mode. 

5.9 Lightness weighting pattern correction. (a) Original photograph, (b) Example painting, (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.

5.10 Lightness weighting pattern correction. (a) Original photograph, (b) Example painting, (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.

5.11 Content-aware interpolation of the lightness weighting pattern.

5.12 (a) Original photograph, (b) Example painting, (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 by content-aware interpolation, (h) Result using weighting in (g).
5.13 (a) Original photograph, (b) Example painting, (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 by content-aware interpolation, (h) Result using weighting in (g). .......................... 134

5.14 Local contrast restoration. The original photograph is in Figure 5.9(a). (a) Example painting, (d) Lightness weighting pattern of (a), (b) Result without local contrast restoration, (e) Result with local contrast restoration, (c) and (f) Close-up regions outlined by red box in (b) and (e). .......................... 136

5.15 The original photograph is in Figure 5.12(a) and the example painting is in Figure 5.12(b). (a) Result by vignetting effect manipulation, (b) Result by atmospheric perspective effect enhancement, (c) Vignetting + atmospheric perspective effect enhancement, (d) Atmospheric perspective effect + vignetting enhancement. ........ 137

5.16 Comparison. (a) Original photograph, (b) Example painting, (c) Lightness weighting pattern extracted from the example painting, (d) Result by the proposed method, (e) Result by camera model 1, (f) Result by camera model 2, (g) Vignette in Photoshop®, (h) “Lo-fi” filter in Instagram®. ........................................ 139

5.17 Comparison. (a) Original photograph, (b) Example painting, (c) Lightness weighting pattern extracted from the example painting, (d) Result by the proposed method, (e) Result by camera model 1, (f) Result by camera model 2, (g) Vignette in Photoshop®, (h) “Lo-fi” filter in Instagram®. ........................................ 140
5.18 More experimental results on landscape photographs. (a) Original photographs, (b) Lightness weighting patterns, (c) Results.

5.19 Application on 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 in (e).

5.20 Two more examples of the application on portrait photographs. (a) Original photographs, (b) Lightness weighting patterns, (c) Results using weighting patterns in (b).

6.1 (a)-(c) Three paintings with different poses. The dark lines in the painting are the one-of-third lines and the white lines are the vertical and horizontal central lines. (d) The locations of faces and centers of the figures in a set of 120 full body paintings.

6.2 The framework of the proposed method.

6.3 (a) A photograph with extracted pose, (b) The graph for the photograph (a), (c) A painting with extracted pose, (d) The graph for the painting in (c). The nodes in blue are the body nodes, which express the body parts. The nodes in cyan are the boundary nodes. The red lines are the skeleton of the body. The yellow 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.
6.4 (a) Photographs with extracted poses, (b) Selected example paintings with poses, (c) The cropping results based on the selected example paintings. .............................. 158

6.5 (a) A photograph with extracted pose, (b) Selected full body painting with pose, (c) Selected small face half body painting with extracted pose, (d) Selected big face half body painting with extracted pose. (e)-(g) The cropping results using the three example paintings. 159

6.6 (a)-(c) Example paintings with extracted pose, (d) Input photograph with extracted pose, (e) Space cropping result (full body) by the proposed method based on example painting in (a), (f) Space cropping result (small face half body) by the proposed method based on example painting in (b), (g) Space cropping result (big face half body) by the proposed method based on example painting in (c), (h) The result by method ACDP in [13], (i) The result by method ABPE in [14]. .............................. 161

6.7 (a)-(b) Example paintings with extracted pose, (c) Input photograph with extracted pose, (d) Space cropping result (small face half body) by the proposed method based on example painting in (a), (e) Space cropping result (big face half body) by the proposed method based on example painting in (b), (f) The result by method ACDP in [13], (g) The result by method ABPE in [14]. .............................. 162

6.8 (a) The space around the figure, (b) Locations of face and center of left directed full body portrait paintings, (c) Locations of face and center of right directed full body portrait paintings. .............................. 163
6.9 Comparison with results using statistical models. (a) Input photographs with pose, (b) Selected example paintings with pose, (c) Cropping results by the proposed method, (d) Results using statistical models. .......................... 164

6.10 Using professional photographs as references. First row: Input photographs with poses, Second row: Selected example professional photographs with poses, Third row: Cropping results using the proposed method. ........................................ 165

6.11 More experimental results. First and fourth rows: Input photographs with pose, Second and fifth rows: Selected example portrait paintings with pose, Third and sixth rows: Cropping results using the proposed method. .............................. 166
List of Tables

3.1 User study result ........................................ 66
3.2 Intersection-over-union ratio of human drawing study for each photograph .................................. 72
4.1 User study result for the comparison of the proposed method with the CAE method in [9] and the AVE method. The percentage shows the number of users responding favorably to the respective method. 104
5.1 The example painting and average lightness of all the paintings in each of the three patterns P1, P2 and P11 ........................................ 121
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACQUINE</td>
<td>Aesthetic quality inference engine</td>
</tr>
<tr>
<td>BG</td>
<td>Background</td>
</tr>
<tr>
<td>CAE</td>
<td>Content-aware enhancement</td>
</tr>
<tr>
<td>DOG</td>
<td>Difference of Gaussian</td>
</tr>
<tr>
<td>FG</td>
<td>Foreground</td>
</tr>
<tr>
<td>FO</td>
<td>Non-face/skin part of the foreground objects</td>
</tr>
<tr>
<td>FS</td>
<td>Face and skin areas</td>
</tr>
<tr>
<td>GMM</td>
<td>Gaussian mixture model</td>
</tr>
<tr>
<td>HM</td>
<td>Histogram matching</td>
</tr>
<tr>
<td>MG</td>
<td>Middle-ground</td>
</tr>
<tr>
<td>MRF</td>
<td>Markov random field</td>
</tr>
<tr>
<td>NPR</td>
<td>Non-photorealistic rendering</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle swarm optimization</td>
</tr>
<tr>
<td>RGBD</td>
<td>Red, green, blue, depth</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>SIFT</td>
<td>Scale-invariant feature transform</td>
</tr>
</tbody>
</table>
List of Notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{\Delta}$</td>
<td>The luminance difference</td>
</tr>
<tr>
<td>$\bar{L}$</td>
<td>The average luminance</td>
</tr>
<tr>
<td>$I_{\text{max}}$</td>
<td>The maximum value of image $I$</td>
</tr>
<tr>
<td>$I_{\text{min}}$</td>
<td>The minimum value of image $I$</td>
</tr>
<tr>
<td>$f_x$</td>
<td>The $x-$ derivative of distribution $f$</td>
</tr>
<tr>
<td>$N_4(i)$</td>
<td>The four neighbors of pixel $i$</td>
</tr>
<tr>
<td>$</td>
<td>a</td>
</tr>
<tr>
<td>$\nabla D$</td>
<td>The gradient field of image $D$</td>
</tr>
<tr>
<td>$|\vec{v}|$</td>
<td>The norm of vector $\vec{v}$</td>
</tr>
<tr>
<td>$\alpha \in A$</td>
<td>$\alpha$ is an element of set $A$</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Background

With the prevalence of digital cameras and mobile devices, there is a dramatic growth in the amount of digital photographs captured and stored. In addition, the emergence of internet has allowed users to share and publish photographs within their social communities in real time. However, since most of camera users are amateurs, the quality of these photographs is often not so visually appealing to be shared directly. How can the aesthetic appearance of the photographs be improved? This has attracted the attention of computer vision and computer graphics research communities. Photo editing is a common post-processing step in digital photography even for professional photographers, aiming to improve the visual appearance of photos by touching up their values using computer techniques. Current computational vision and graphics research offers many tools which could contribute to improving the appearance of photos. Generally, the tools can be classified as manual/interactive editing, single photo automatic editing, learning-based editing and example-based editing systems.

The popular image-editing software, Adobe® Photoshop®, is used by amateurs
and professionals alike to edit their photographs manually. Using Photoshop®, photographs can be re-lighted, re-focused, and have their colors adjusted in a variety of ways. However, how to combine these operations to make the photos more aesthetically appealing needs professional skills and experience. In addition, local editing on selective regions of the image is a laborious process. It requires the user to first select the image region of interest, and then apply the edit to that region. It is a tricky task to adjust the value of one region while observing the relationship to its neighboring regions. Interactive semi-automatic photo-editing tools can save time for users to some degree. For example, adjustments on the color, tone or contrast can be propagated to a region by interactively drawing brush strokes [15–19]. However these interactive methods focus on the application of propagation technique responding to the user inputs, not on the principles that make one adjustment more aesthetically valuable than another.

Single photo automatic editing methods, e.g., contrast editing [4, 20, 21], color and tone correction [9], and composition optimization [13, 22] have been developed based on pre-defined principles or rules. It is difficult for users to control the extent of adjustment for different photos with due consideration of the semantic content.

Learning-based photo-editing methods are developed to render the photos to specific styles based on learned models or parameters [7, 8]. The models or parameters are trained using input-output image pairs where the output images are from professional photographers or artists. Currently, these learning-based methods mostly rely on the lower-level image statistics (e.g. color, intensity gradient) and they do not exploit object-level semantics except for the face as in [8]. Even the face is specifically considered in [8], but only global tone mapping function is used.

An alternative approach is to use example-based methods which transfer the aesthetic values from an example image to a target image. They can achieve aes-
thetic improvement by considering the object-level semantic content of the example image [23]. The selected example image should contain the required aesthetic values. Paintings, being a mode of creative expression, are carefully constructed by artists based on the visual perception. They provide us good examples to edit the photos for aesthetic appearance. Existing approaches in learning from traditional paintings, such as in painterly rendering [24, 25], color style transfer [26–28], image abstraction [29], aim to develop algorithms to emulate the surface characteristics of a painting such as its brush strokes, line drawings, and color. These approaches ignore the important structural qualities that inform the aesthetic values of an image. It is still a long way to understand how artists create aesthetics in paintings [30].

1.2 Motivation

In many contemporary movies and games, scenes are rendered to be more aesthetic and expressive by re-organizing the image values. Figure 1.1 shows one example of aesthetic editing in the movie “Star Wars”. This film was first made in 1977, before the age of digital film making, but was digitally re-mastered in 1997. By comparing the selected scenes before and after digital re-mastering, we can clearly see the effect of their aesthetic improvement. In the first scene (Figure 1.1(a)), the contrast of the ground and sky has been increased, the corners have been darkened, and the color has been shifted. There is a clear contrast change from the foreground to far away background mountain and there is a clear contrast between the corners to the center. In the second scene (Figure 1.1(b)), the contrast of the buildings has been increased, and the building area is brightened. In the third scene (Figure 1.1(c)), the sky has been given much more room, and the ground plane of the valley has been accentuated, showing a depth change from near to far. The change
in the movie was made by professional artists who were good not only with the software but also with pictorial skills. Such software and pictorial skills are beyond the reach of amateurs. All the changes on scenes of the movie are to make the values more organized.

The artists know how to organize values in an image to deliver a sophisticated and considered visual appearance. Painters can arrange light, form, color and composition in a considered and beautiful way. As in the examples shown in Figure 1.2, values across the 2D canvas plane and the scene depth plane are organized to produce the focus of attention and illusion of depth. The placement of objects is also organized to guide the eye of the viewer. Even the best photograph, which is derived mechanically from a scene, is not easily subject to such consideration. Is it possible to transfer the organization of such aesthetic values from paintings to digital photographs to serve the aesthetic improvement of the latter? This is an interesting question that motivates the work reported in this thesis.


1.3 Objective

The objective of the research reported in this thesis is to study the underlying attributes of the visual appeal in pre-existent paintings and develop algorithms and methods to facilitate the transfer of the organization of such image attributes from paintings to photographs for the purpose of aesthetic improvement. This thesis studies attributes in paintings through the following two main aesthetic aspects:

1. **Contrast.** One of the aesthetic considerations in paintings is the way the tone and color values have been organized with respect to various notions of contrast. The contrast that painters considered can be divided into global contrast for creating global mood, local contrast to highlight local details, contrast across depth planes for depth perception, regional contrast and center-corner contrast to create the focus of attention (see Figure 1.3). The thesis studies how these contrasts are organized in paintings and the difference of their organizations compared to photographs. Then, the thesis presents algorithms and methods to transfer these contrast organizations from paintings to photographs for aesthetic improvement. Specially, the thesis focuses on improving the depth perception through the manipulation of contrast across depth planes in line with the principle of atmospheric
perspective effect, and the focus of attention through figure-ground regional contrast manipulation and painter-style vignetting manipulation.

2. Composition. The composition of a painting is subject to geometric attributes such as target positioning and spacing. Composition takes an important role in image aesthetics. The thesis studies the placements of the face and the figure in portrait paintings. An algorithm is developed to transfer the composition of an example portrait painting to a portrait photograph.

One of the difficulties in applying the contrast and composition organization of paintings is that, although general observations of their organization can be made, it is a challenge to identify exactly the rules by which this organization operates. Therefore, instead of a rule-driven approach, the thesis advocates an example-based approach to handle the diversity of content in paintings and photographs.
1.4 Major Contributions of the Thesis

The major contributions of this thesis are summarized as follows:

- The exaggeration of the atmospheric perspective effect by artists in 2D images increases the illusion of depth thereby making the image more interesting. The first contribution in this thesis addresses the transfer of the atmospheric perspective effect in landscape paintings to landscape photographs by the manipulation of depth-aware lightness and saturation contrast values. The work is based on a statistical study on paintings. The statistical study clearly shows that the saturation contrast and lightness contrast inter- and intra-depth planes in paintings are exaggerated with a view to improving the visual appeal of the painting and the illusion of depth within it, while respecting the existing contrast relationships of a natural scene guided by the atmospheric perspective effect. The contrast manipulation is formulated as an optimization problem that simultaneously considers the desired inter-contrast, intra-contrast, and specified gradient constraints. Experimental results and a user study demonstrate that, by using the proposed algorithm, both the visual appeal and the illusion of depth of the photographs can be effectively improved.

- In a portrait, the regional contrast is the main attribute that makes the figure, especially the face, stands out in the painting. The second contribution in this thesis is the enhancement of regional contrast in snap-shot style portrait photographs by using pre-modern portrait paintings as aesthetic examples. The difference of the regional contrast organization in paintings and photographs is first studied and the outcome of this study provides us justification for the enhancement of regional contrast in a photograph based on an example painting. The contrast organization in an 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 area of the foreground figure, the non-face/skin part of the foreground, and the background region. A novel piecewise nonlinear transformation curve is proposed to achieve the desired contrast mapping. Experimental results demonstrate that, by using the proposed algorithm, the visual appeal of portrait photographs can be effectively improved and the face and the figure become more salient.

• Other than the depth-aware contrast, regional contrast and local contrast, the thesis also discusses center-corner contrast which corresponds to an effect that is generally known as the vignetting effect. Vignetting is an effect that is often seen as being clear in the center and fading off at the edges with reduced image brightness or saturation at the periphery. The thesis analyzes the vignetting effect in paintings and photographs. The observation from the analysis shows that the vignetting effect in paintings is more purposely presented with aesthetic view. As the third contribution of the thesis, an algorithm is explored to transfer the lightness weighting that generates the vignetting effect from an example painting to the photograph to improve the focus of attention.

• The fourth contribution of the thesis is to study the composition in portrait paintings and develop an algorithm to improve the composition of a portrait photograph based on an example portrait painting. Pose, face direction and space around the figure are the main elements considered in portrait composition. A graph model is proposed to describe these attributes for selecting an example painting with similar geometrical organization. Space cropping technique is used to improve the composition of the photograph by aligning to the composition in the example painting.
1.5 Organization of the Thesis

This thesis is organized as follows:

In Chapter 2, the color and geometric attributes, and the aesthetic aspects are studied. The differences of the various color spaces are also analyzed.

In Chapter 3, the depth-aware contrast organization in paintings and photographs are discussed and an algorithm is developed for the transfer of the atmospheric perspective effect in landscape paintings to photographs by the manipulation of depth-aware lightness and saturation contrast values.

Chapter 4 discusses the regional contrast in portrait. A simple and effective regional contrast mapping method is proposed to transfer the regional contrast from an example painting to the photograph.

Chapter 5 presents the vignetting effect manipulation in photographs driven by center-corner contrast relationship in an example painting.

In Chapter 6, an algorithm that improves the composition in portrait photographs based on an example painting is presented.

Chapter 7 summarizes the concluding remarks on this thesis and recommends several future research directions.
Chapter 2

Image Attributes and Aesthetic Aspects

Image attributes in a color image are generally understood as the Red (R), Green (G) and Blue (B) values of pixels. These are also known as the photometric attributes of an image. The distributions of these values are often the focus of study. In this chapter, the study of image attributes is beyond the RGB values of pixels with a focus on the perceptual concepts of colors. The color attributes and their representations in various color spaces are reviewed which provide the basis of their manipulation and the understanding of their effects on image aesthetics. This chapter also reviews geometric attributes. The notion of contrast, which expresses the organization of color attributes, is observed to be one of the many possible factors that are behind the visual appeal of an image. Itten’s color contrast is well-known in the art world and this chapter includes a discussion on how Itten’s contrast influences the research reported in this thesis. This chapter also presents the global manipulation of such contrast and discusses effects on the aesthetic appearance of an image. Composition, which is subject to the organization of geometric attributes, is also observed to be influencing the image aesthetics. The
effect of the composition is also discussed in this chapter.

2.1 Image Attributes

2.1.1 Color attributes

Lightness, saturation and hue are the three distinct perceptual attributes of color which can be calculated from the RGB values. Lightness is the subjective brightness perception of a color. Hue presents the color perceived by the observer and saturation refers to the purity of the hue [1]. Hue and saturation constitute the chromatic attribute of color and lightness is the achromatic attribute of color. For numerically measuring the colors, several color models, generally known as color spaces, are available.

Figure 2.1: RGB color model. The figures are from [1]. (a) RGB color model, (b) RGB 24-bit color cube.

**RGB color space.** RGB color space is an additive color space described using the three primary colors of red, green and blue. Any chromaticity in the RGB color model is produced by adding those three primary colors. This color space is often used in imaging and image display devices. A standard RGB space, known as
sRGB, is often used by many image formats. The color model is shown in Figure 2.1. Although the RGB color model describes colors numerically, it is not suitable for describing colors for the purpose of human interpretation.

**HSL and HSV color spaces.** HSL (hue, saturation, and lightness) and HSV (hue, saturation, and value) are two most common cylindrical-coordinate color models which rearrange the geometry of RGB in an attempt to be more perceptually relevant for human interpretation than the cartesian representation in RGB [2]. The models of these two color spaces are shown in Figure 2.2.

![HSL and HSV cylinder models. The figures are from [2]. (a) HSL cylinder, (b) HSV cylinder.](image)

The hue (H) values are calculated in the same way in both HSL and HSV spaces according to [2], which is expressed as

\[
H = 60 \times H' \tag{2.1}
\]

where

\[
H' = \begin{cases} 
0 & \text{if } \max(R, G, B) = \min(R, G, B) \\
\frac{G - B}{\max(R, G, B) - \min(R, G, B)} & \text{if } R = \max(R, G, B) \\
2 + \frac{B - R}{\max(R, G, B) - \min(R, G, B)} & \text{if } G = \max(R, G, B) \\
4 + \frac{R - G}{\max(R, G, B) - \min(R, G, B)} & \text{if } B = \max(R, G, B)
\end{cases} \tag{2.2}
\]
The lightness (L) and saturation (S) in HSL space are

\[ L_{HSL} = \frac{\max(R, G, B) + \min(R, G, B)}{2} \] (2.3)

\[ S_{HSL} = \begin{cases} 
0 & \text{if } \max(R, G, B) = \min(R, G, B) \\
\frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B) + \min(R, G, B) - 1} & \text{otherwise}
\end{cases} \] (2.4)

The value (V) and saturation (S) in HSV space are

\[ V_{HSV} = \max(R, G, B) \] (2.5)

\[ S_{HSV} = \begin{cases} 
\frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} & \text{if } \max(R, G, B) \neq 0 \\
0 & \text{otherwise}
\end{cases} \] (2.6)

Hue and saturation are related to the way in which humans perceive color. HSL and HSV spaces are very intuitive to use for image editing. Although H, S, L or H, S, V values are numerically independent as illustrated in Figure 2.3(a) in which it shows a clear color spectrum ramp but with a uniform lightness value, in real life, they are perceptually engaged. That is to say, if the hue of an object changes from blue to yellow it appears to become brighter. However, the lightness in HSL and HSV spaces measures these colors using a uniform value (as shown in Figure 2.3(a)). The HSL and HSV color spaces are not perceptually uniform. In other words, the perceived difference of two colors is not coherent with the Euclidean distance in the HSL or HSV color model.

**CIELAB and LCH color spaces.** Based on the artist Albert Munsell’s study of color perception, CIELAB model presents lightness value in the L channel and color in the a and b channels. The L channel of the hue ramp matches our perception of lightness (as shown in Figure 2.3(b)). CIELAB color space is considered as perceptually uniform. However, one of the features of CIELAB
space is that saturation and hue are not located in either of the a or b channel 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 colorfulness of a color.

Figure 2.4 shows the perceptual influence of the lightness change in LCH space and HSL space. The lightness in LCH space preserves the perceptual contrast much better than in HSL space (see the sun light reflection on the water in Figure 2.4(b)(c)). When using the same gamma correction (gamma is 0.7) in LCH space and HSL space to adjust the lightness contrast, the lightness in LCH space is corrected while preserving the original perceptual contrast as shown in Figure 2.4(d). However, the same adjustment in HSL space changes a lot in the dark region with reduced contrast in the sky region. The lightness change in the HSL space influences the perceptual colorfulness of the image. This example shows the advantage of the contrast manipulation in LCH space. Therefore, in the study of contrast in this thesis, the LCH color space is used to measure the contrast values. The chroma in LCH space is indicative of saturation.
Figure 2.4: Comparing image editing in LCH and HSL color spaces. (a) Original image, (b) L in LCH space, (c) L in HSL space, (d) Gamma correction of L in LCH space, (e) Gamma correction of L in HSL space.

2.1.2 Geometric attributes

Geometry within an image is the vehicle used to organize the space of that image. A painting is not simply a plane surface. There are 2D geometry and 3D geometry in a painting. The 2D geometry consists of center, corner and edge values. As a rule of thumb, elements of high pictorial value are placed somewhere near the center of the painting. The 3D geometry is usually understood by painters in terms of the traditional depth planes: foreground (FG), middle-ground (MG), background (BG) and sky separation. As a rule of thumb, elements of high pictorial value are placed somewhere in the FG of the painting. The space organization of 3D geometry produces the illusion of depth while the organization of 2D geometry produces the focus of attention by ways of placing elements.
2.2 Aesthetic Aspects

Along with computational aesthetics, some researchers have focused on the quantitative assessment of the aesthetics of art works [31]. As a pioneering work, Birkhoff proposed to measure aesthetics by using information theory [32]. Birkhoff bound the aesthetic perception with two criteria: order and complexity. For Birkhoff, the order is assumed to be elements such as: symmetry, rhythm, repetition, contrast, and the complexity is the amount of effort the human brain has to put into processing of an object. Based on Birkhoff’s work, researchers bring in originality, redundancy to replace order in the aesthetic measure [33, 34]. Rigau et al. defined a set of ratios based on information theory and Kolmogorov complexity to quantify the aesthetic process in paintings [35]. For the global aesthetics measure, they considered the creative process as the initial palette, the used palette and the final color distribution. The work also analyzed the image composition by measuring the spatial distribution of colors from a given palette through the decomposition of the image into regions. The same authors also applied these informational aesthetics measures to quantify the description of van Gogh’s works [36]. These informational aesthetics measures show that the aesthetics of artworks is related to the selection of color and the spatial distribution of color (composition). In recent publications, some researchers have proposed to formulate the measurement of the visual aesthetics of photographs [37, 38] and paintings [39] as a machine learning problem. They classify images into different aesthetic categories by extracting features to represent both the global characteristics and local characteristics from them. Among the features, the contrast of color (hue and saturation), contrast of lightness, and composition are always important features that influence the visual aesthetics measurement. This thesis looks into the following aesthetic aspects: contrast and composition in paintings, and study how they might be transferred to photographs.
2.2.1 Contrast

Contrast is defined as the difference in visual properties that makes an object distinguishable from other objects and the background [40]. Contrast as a principle in art is created by using elements (color, line, texture, shape, form, value, space) that conflict with one another [41]. In human visual perception of the real world, contrast is determined by the difference in the brightness and color of an object with other objects [40]. Painters operate within a world of relative values as the human visual system is more sensitive to contrast than absolute value. In making a painting, painters do not value the fidelity of paint to its optical source as much as its quality of relative difference to its neighbor upon the canvas [42]. It is an awareness of contrast that drives this process. They use contrast to make areas perceptually more lighter or darker to create the relative position, depth, and shading, so that, contrast is created to add visual interest and guide the eye to certain areas.

2.2.1.1 Itten’s contrast

General observations for how artists mentally organize contrast values are outlined by the work of Johannes Itten (1888-1967) [43]. Itten recognizes seven color contrasts as listed below:

1. **Contrast of Saturation.** It is created by selectively dulling the color.

2. **Contrast of Light and Dark** (lightness contrast). It expresses the brightness difference of light and dark values.

3. **Contrast of Extension** (Contrast of Proportion). This contrast is formed by assigning proportional sizes for colors. It focuses on the relative areas of two or more colors.

4. **Contrast of Complements.** Complementary contrast is created by using complementary colors that are pairs of colors opposite on the color wheel.
5. **Simultaneous Contrast.** It is the contrast between areas whose perceptual values interact with each other. Simultaneous Contrast is most intense when placing complementary colors side by side.

6. **Contrast of Hue.** It is illustrated by the undiluted colors in their most intense luminosity. Yellow/red/blue represents the strongest Contrast of Hue.

7. **Contrast of Warm and Cool (aka color temperature).** When some colors give a warm feeling (red, yellow, orange), some give a cool feeling (blue, green, purple). The Warm and Cool contrast is formed by the juxtaposition of warm and cool hues which creates the powerful dimension.

In the study on contrast reported in this thesis, it is assumed that any method that is resulted from this study could only effect pixel color values, not area size. So the Contrast of Extension is not included in the study because it is primarily a contrast of area size, not color. As Simultaneous Contrast is a compound expression of how the different values conflict with each other, it is not considered in the current study for being too complex and rooted in perceptual judgment. It is also not a major feature of art or photography except in late modernism such as in the work of the Op Artists [44]. Contrast of Hue, Contrast of Complements, and Contrast of Warm and Cool are about the difference of hues. The operation on these three contrasts will shift the color of the images, e.g. the blue sky may be changed into green. For preserving the natural property of the original photograph, the thesis only operates on the lightness and saturation contrast values while the hue is untouched.

### 2.2.1.2 Contrast in landscapes and portraits

As shown in Figure 1.3, the contrast of images can be considered in terms of their global contrast for creating the global mood, contrast across depth planes for the enhancement of depth perception, regional contrast and center-corner contrast for
the focus of attention, and local contrast to highlight the local details.

1. **Global contrast.** The global contrast is the value distribution of the whole image without spacial information. Histogram is one expression of the global distribution of lightness and saturation of an image. Therefore, the global lightness contrast and saturation contrast of the image are defined as histogram distributions and histogram matching is used to map the global contrast as in [45]. Given the normalized histogram of the input image \( \{P(s_k)\}_{k=0}^{N-1}\), and the normalized histogram of the destination image \( \{Q(r_k)\}_{k=0}^{N-1}\), the histogram matching is defined as follows [1].

\[
F(s_k) = \sum_{i=0}^{k} P(s_i), \; k = 0, 1, 2, \ldots, N - 1
\]  
(2.7)

\[
T(r_k) = \sum_{j=0}^{k} Q(r_j), \; k = 0, 1, 2, \ldots, N - 1
\]  
(2.8)

\[
\tilde{s}_k = \min_r \| T(r) - F(s_k) \|
\]  
(2.9)

where, \( N \) is the number of discrete levels, \( s_k \) and \( r_k \) are the number \( k \) discrete level of the input image and destination image respectively. Figure 2.5 shows two paintings and their lightness and saturation histograms. The histograms show that the two paintings have big difference in their global contrast organization. The painting in the top row has a large bright and un-saturated background region with small darker and saturated objects in foreground, while the lightness and saturation are evenly distributed across the painting in the bottom row. The histogram matching results on a photograph using the two paintings as references are shown in Figure 2.6(b)(c). The comparison of the lightness and saturation distributions in Figure 2.6(d)-(g) shows that the lightness and saturation of the photograph are mapped to those of the references. In other words, the result has similar global lightness and saturation contrast organization with the corresponding reference.
Chapter 2. Image Attributes and Aesthetic Aspects

2. Depth-aware contrast. In both landscape and portrait paintings, there is 3D geometric organization that illustrates the perception of depth. Figure 2.7 shows one landscape painting and one portrait painting that are divided into depth planes. The depth-aware contrast is the contrast organization within depth planes and between depth planes.

3. Regional contrast. Regional contrast is the contrast between neighboring regions in the 2D image plane. Creative regional contrast helps to guide the eye to the focus of attention. As in the portrait painting in Figure 1.3(e), the figure is in high lightness contrast to the background region and the face is in high lightness contrast to the surrounding regions. These contrast organizations attract our attention on the figure, especially the face.

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 a large space around the center.

Figure 2.5: (a) Paintings, top: “A Coign of Vantage”, by Sir Lawrence Alma-Tadema, 1895, bottom: “The Poppy Field” by Vincent Van Gogh, 1888. (b) Normalized histograms of lightness, (c) Normalized histograms of saturation.
Figure 2.6: (a) Original photograph, (b) Histogram matching result using the painting in the top row in Figure 2.5 as reference, (c) Histogram matching result using the painting in the second row in Figure 2.5 as reference. (d) and (e) The comparison of the lightness and saturation distributions of the original photograph in (a), the result in (b) and the corresponding reference. (f) and (g) The comparison of the lightness and saturation distributions of the original photograph in (a), the result in (c) and the corresponding reference.
Figure 2.7: (a) Landscape painting with outlined depth planes, (b) Left: Francisco Goya (1746 - 1828), “Portrait of King Ferdinand VIII” 1803, Right: the portrait divided into depth planes.

and the background (non-figure area) is often drawn simply in order to have a high contrast with the figure. Therefore, this thesis focuses on studying the depth-aware contrast in landscapes for the illusion of depth and focuses on exploring the regional contrast organization in portraits for the focus of attention on the figure.

4. **Center-corner contrast.** Center-corner contrast is also for the focus of attention. Important elements are placed in the center by artists while fading off the details close to the edge to avoid drawing the attention of the viewer away from the center. Artists are good at selecting a view point that naturally presents the important elements within the center by darkening the corners [46]. The center-corner contrast exists in both landscape and portrait paintings.

5. **Local contrast.** Local contrast serves to highlight local details. Even in the dark environment in the example painting shown in Figure 1.3(b), the foreground object is still depicted clearly by being in an overt lightness contrast with the neighboring area. The local contrast operation can be combined with the depth-aware contrast, regional contrast and center-corner contrast operations.
2.2.1.3 Contrast models

Currently, there is not a standard measurement model of lightness or saturation contrast. All existing contrast statistics would probably be better regarded as application tools than as descriptions of basic visual function.

In physics, the light and dark difference is called the luminance contrast. Since no contrast model is available to measure the perceptual difference of lightness, we therefore assume the lightness contrast can be measured in terms of the luminance contrast. Some methods have been reported to measure the luminance contrast in images. According to Weber-Fechner law [47, 48], the relationship between the physical magnitude of luminance and the perceived intensity is

\[ dp = k \frac{dS}{S} \]  

(2.10)

where \( S \) is the physical stimulus and \( dp \) is the differential change in perception, \( dS \) 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 [40]. Based on the relationship in Eq. (2.10), many definitions of contrast can be seen to follow the general relationship below:

\[ C = \frac{L_\Delta}{\bar{L}} \]

(2.11)

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

Weber contrast is commonly used in cases where small objects are present on
Chapter 2. Image Attributes and Aesthetic Aspects

a large uniform background [40]. It is defined as

\[ C_W = \frac{I - I_b}{I_b} \quad (2.12) \]

where \( I \) is the luminance of the object and \( I_b \) is the background luminance.

Michelson contrast is commonly used for cases where both bright and dark features are equivalent with similar size. It is defined as

\[ C_M = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}} \quad (2.13) \]

where \( I_{\text{max}} \) and \( I_{\text{min}} \) are the highest and lowest luminance respectively.

Luminance ratio is defined as

\[ C_{LR} = \frac{I_{\text{max}}}{I_{\text{min}}} \quad \text{or} \quad \lg C_{LR} = \lg \left( \frac{I_{\text{max}}}{I_{\text{min}}} \right) \quad (2.14) \]

It has often been applied to the same stimulus as Weber contrast.

These contrast models above are only suitable for simple images or contrast between regions. In complex images, the perceived contrast varies greatly across an image. Some models have been proposed to measure the perceived contrast in complex images and they are summarized as follows.

Root mean square (RMS) contrast [49] is the standard deviation of intensity values. It is defined as

\[ C_{RMS} = \sqrt{\frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I_{ij} - \bar{I})^2} \quad (2.15) \]

where \( I_{ij} \) is the \( i \)-th, \( j \)-th element of the image of size \( M \times N \), \( \bar{I} \) is the average intensity of the image. RMS contrast does not capture spatial information of the image.
Peli [49] 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

\[ c_i(x, y) = \frac{a_i(x, y)}{l_i(x, y)} \]  

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

Tadmor and Tolhurst [50] 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.

\[ C_1(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_c(x, y)} \]  
\[ C_2(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_s(x, y)} \]  
\[ C_3(x, y) = \frac{R_c(x, y) - R_s(x, y)}{R_c(x, y) + R_s(x, y)} \]

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 center and surround components respectively. \( I \) is intensity.

Rizzi et al. [51] proposed a local contrast measurement in digital images in 2005. It subsamples the image to various levels. At each level, the local contrast is calculated. The definition of the local contrast is

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

where \( m(l) \) and \( n(l) \) are the numbers of rows and columns on level \( l \), \( N \) is the number
of levels, $N^8(i)$ is the eight neighbors 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 [52], the authors improved the contrast measure by combining the multilevel approach with Tadmor and Tolhurst’s evaluation of a color stimulus [50]. The local contrast at each level is defined as

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

(2.21)

$DOG$ is the $C_3$ function in the Tadmor and Tolhurst method [50], $m(l)$ 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 color contrast is linearly combined with the luminance contrast in CIELAB space for color images.

A global contrast factor was used by Matkovic et al. [53] to measure richness of detail as perceived by a human. Global contrast factor is defined as

$$C_{GCF} = \sum_{i=1}^{N} w_i * C_i$$

(2.22)

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

Mantiuk et al. [54] used a low-pass contrast measurement to avoid halo artifacts 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 neighboring 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 neighboring pixels [39, 55], which is expressed as $C_S = S_i - S_j$, $S_i$ 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 color contrast considering the
chromatic and achromatic aspects. Rizzi et al. [52] calculated the contrast for L, a, b channels in CIELAB space separately. Then the color contrast was measured as a linear combination of the contrast values of L, a, and b. Marius Pedersen used Lab variance as a color contrast measure, which is calculated as the geometrical mean of the variance in each channel in the CIELAB color space [56].

2.2.2 Atmospheric perspective

Atmospheric perspective, also known as Aerial perspective, is the effect that an atmosphere has on the appearance of distant scenery. As the distance between an object and a viewer increases, the contrast of the object decreases (see the example photograph in Figure 2.8). The color of the object also becomes less saturated and lightness increases with distance [3]. Atmospheric perspective is a physical phenomenon derived from the scattering and absorption of light as it travels from the source to the viewer. Atmospheric perspective provides us a cue for depth perception [57]. In painting, atmospheric perspective is one technique used to create depth illusion by depicting distant objects as fading away and being less detailed as the painting in Figure 2.7(a). Such changes across depth planes obey the law of atmospheric perspective and give rise to depth-aware contrast organizations.
Atmospheric perspective has been used in computer graphics for landscape design [58,59] by modeling the light scattering. Jung et al. used an estimated atmospheric perspective physical model to improve the depth perception of a 2D image [6]. Different from the physical atmospheric perspective effect which only changes the influence of atmosphere based on distance, the atmospheric perspective effect in paintings is specially organized with an aesthetic view which is often expressed through exaggeration.

2.2.3 Vignetting

Vignetting is an effect that is seen as clear in the center and fading off at the edges by reducing the image brightness or saturation at periphery in photography (see the example in Figure 2.9(b)). Based on the source, vignetting can be classified as: natural vignetting, pixel vignetting, optical vignetting and mechanical vignetting [60–62]. Natural vignetting is the radial falloff due to geometric optics which causes different regions of the image plane receiving different irradiance. Pixel vignetting is caused by angle-dependence of digital sensors. Optical vignetting is caused by light paths blocked inside the lens body by the lens diaphragm. Mechanical vignetting is caused by the radial falloff due to light paths becoming blocked by other camera elements. Although vignetting is often an undesired effect caused by camera settings or lens limitations, it sometimes creates an artistic effect, such as drawing attention to the center of the image. In photographs, the vignetting effect can be created by camera settings or during post-processing. The “Lens Correction” filter in Photoshop® can also achieve the effect such as seen in the example in Figure 2.9(b). The physical vignetting shows a clear circle in the image.

The center-corner contrast organization by artists is another kind of vignetting effect, which is called the painter-style vignetting. Different from the physical...
Figure 2.9: The vignetting effect. (a) Original photograph, (b) Vignette in Photoshop®.

Vignetting (photo-style vignetting), artists are good at selecting a view point for the vista to naturally add shadows on the corner elements [46]. As shown in the example in Figure 1.3(d), the trees on the periphery and the clouds on the corner are painted dark naturally to attract the eye to the center. This painter-style vignetting also helps to enhance the expression of depth by guiding the eye from near to far through the center area.

2.2.4 Composition

Composition is the arrangement of visual elements within an image. It is subject to the geometric attributes of images. Composition takes an important role in image aesthetics. In order to achieve an aesthetic view, there are some principles for the organization of visual elements in both artworks and photography, such as shape and proportion, balance, golden ratio, focus etc. [63]. However no absolute rules can ensure good composition in images. Graphic artists use compositional principles as means of attracting attention and expressing concepts [64]. Artists place much less emphasis on compositional principles than graphic artists do. Photographers are more relying on their intuitive sense and vary substantially in their
approach to composition [64].

While in landscape composition, the placement of foreground elements and the organization of depth planes are important, the placement of the figure and the face are even more important in portrait composition [65]. Comparing with landscape, the composition of portrait is more focused. This thesis discusses the composition in portrait paintings and proposes a pose-related method to improve the composition of a portrait photograph based on an example portrait painting.

2.3 Concluding Remarks

This chapter reviews the color and geometric attributes of images. The discussion on the color spaces for the description of color attributes shows that the LCH color space is perceptual uniform and it is used to measure contrast values in the research reported in the thesis.

Contrast and composition are important aesthetic aspects in images. They are specially organized in paintings for various aesthetic effects. The effects of atmospheric perspective and vignetting correspond to depth-aware contrast and center-corner contrast respectively. These effects in paintings are specially organized for the illusion of depth and focus of attention. This chapter mentions a histogram matching method to transfer the global contrast. In the following chapters, the thesis focuses on the manipulation of depth-aware contrast, regional contrast and center-corner contrast while the local contrast is an integral part in the three contrast manipulations. Since the depth space is more important in landscapes and the expression of the figure and the face is more important in portraits, therefore the thesis focuses on studying the depth-aware contrast in landscapes for the illusion of depth and on exploring the regional contrast organization in portraits for the focus of attention. The composition in portraits is also studied.
Chapter 3

Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

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 interesting. 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 organization of such contrast in landscape paintings. The rational behind this manipulation is based on a statistical study which has clearly shown that the saturation contrast and lightness contrast inter-
Figure 3.1: An example result of atmospheric perspective effect enhancement using the proposed method. (a) Input photograph, (b) Reference painting, (c) Result.

and intra-depth planes in paintings are more purposefully organized than those in photographs. This contrast organization in paintings respects the existing contrast relationships within a natural scene governed by the atmospheric perspective effect and also exaggerates them sufficiently with a view to improving the visual appeal of the painting and the illusion of depth within it. The depth-aware lightness contrast and saturation contrast revealed in landscape paintings guide the mapping of such contrasts in photographs. This contrast mapping is formulated as an optimization 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.

3.1 Introduction

As indicated in Chapter 2, 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. 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 [57]. 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, in the game World of Warcraft, an algorithm drives the live updating of the atmospheric perspective effect upon objects as they move into the distance.

The atmospheric perspective effect 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, compressed, washed-out and generally unpleasing visual result. Many global and local image adjustment methods, and automatic and template-driven image improvement software (e.g. Instagram®, Adobe® Photoshop®, etc.) 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, centers, edges and regions, it is also clear that the photograph has an implicit depth dimension. It is obvious that depth is of particular importance in landscape images to create the feeling of space and distance. Thus, this chapter proposes to improve the visual appeal and illusion of depth in landscape photographs by enhancing the atmospheric perspective effect.

The physical model of atmospheric perspective has been developed by researchers [66,67]. However, the atmospheric condition and attenuation strength have to be determined for a specific photograph. Jung et al. used a physical model of atmospheric perspective to improve the depth perception of a 2D image [6]. The application of this physical model only increases the atmosphere effect simply based on distance. The contrast of both close object and distant object will be reduced by increasing the atmosphere effect. Additionally, users have to decide how much to enhance the atmosphere effect to create the desired result. However, for a flat and low contrast photograph with or without atmospheric perspective effect, the
contrast of close objects generally need to be increased while adjusting the relative
difference between objects. Although such sophisticated software as Photoshop®
provides powerful tools for users to manually edit the images, the skills needed
to effect the improvement in depth view are beyond the casual user. Even for an
experienced user, it is far from a trivial task. The users have to mask regions,
adjust each region differently to its neighboring regions and make decisions based
on the relative relationships of these regions. Clearly, the number of possible per-
mutations involved is legion and the chance of effecting an adjustment to the image
that damages its visual integrity is high.

Fortunately, painters have provided us plenty of meaningfully enhanced in-
stances 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 dis-
tance. The depth in a painting is of a simplified form, consisting of the traditional
discrete planes: foreground (FG), middle-ground (MG), background (BG) and sky
[68]. The atmospheric perspective is always exaggerated by painters to strengthen
the effect of the separation of depth planes, with a view to improving the visual
appeal and illusion of depth [69]. It is through a consideration of the relative
contrast values of these depth planes that the atmospheric perspective effect is
realized [69].

Therefore, this chapter proposes to learn how painters organize the exagger-
ation of the atmospheric perspective effect in landscape paintings by analyzing
the lightness and saturation contrast values according to traditional depth planes.
Also, this chapter explores how collections of these contrast organizations can be
passed onto acquired photographs through their corresponding depth planes thus
radically enhancing their visual appeal and the illusion of depth.

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 contrast are organized in paintings, a depth-based contrast analysis of a set of landscape paintings made by the Hudson River school painters was conducted. The paintings by Hudson River school artists are quintessential examples of a mastery of atmospheric perspective [70]. The contrast organization of these paintings was compared to a similar analysis of a collection of “snapshot” 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 organized [44]. This contrast organization in paintings respects the existing contrast relationships within a natural scene governed by atmospheric perspective and also exaggerates them with a view to improving its visual appeal and illusion of depth. 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 organization across depth planes of paintings as reference to enhance digital photographs.

One of the difficulties in applying the depth-aware contrast organization of paintings is that though general observations of their organization can be made, it is a challenge to identify exactly the rules by which this organization operates. Therefore, instead of a rule-driven approach or a physical model driven 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
Chapter 3. Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

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 to illustrate our idea is shown in Figure 3.1.

3.2 Related Work

Depth-aware enhancement. The research in [71] mimicked the shading characteristics of paintings to enhance the depth perception of scenes. This shading technique is effective in separating the overlapped objects. However, for long range 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, careful selection of point of view etc, to create the illusion of depth. Bailey [72] proposed to simulate artistic control of apparent depth in image by changing the lightness or color on the boundary which was similar as adding shading, or by replacing the color or lightness of objects to change the perceptual order. This technique is also suitable only to change the perceptual depth of overlapped objects. Jung et al. proposed a method by which the depth perception of a 2D image might be improved by enhancing such depth cues as: linear perspective, aerial perspective, focus, and shadow effects [6]. They made color 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.

Contrast enhancement. Traditional contrast enhancement seeks to improve
the quality of an image and enhance its general visibility. Its function is to make the
details in very dark or light regions more visible, or change the dynamic range of a
natural image whilst keeping the color unchanged. The methods that are currently
used for contrast enhancement are based on automatic global or local contrast
operations such as frequency transform [73, 74], histogram modification [75, 76],
gradient optimization [4] and Retinex filter [77]. An example-based contrast en-
hancement method by gradient mapping was proposed in [78]. These contrast
enhancement methods only enhance the contrast of global or local areas, and they
do not aim to visually emphasize the contrast of regions in relation to each other.
Interactive local contrast adjustment methods [17, 18] apply change to the con-
trast of interest regions or objects through the drawing of brush strokes. These
interactive adjustments need time and skill to master because users need to make
decisions based on the relative relationships of selected regions.

**Tonal and color adjustment.** Much work has been done on the enhance-
ment of the visual effect of an image through an adjustment of its tonal and color
values. Color transfer changes the color of the input by referring to a reference
image [27, 79, 80]. Differently, Cohen-Or et al. [81] proposed to adjust the colors
of a given photograph according to the so-called harmonic rules. Furthermore,
Wang et al. [82] presented a method for adjusting the color composition of an
image to match a predefined color theme. Instead of color, Bae et al. proposed
to manipulate the tone to convey the look and feel of a reference image by using
histogram matching in large-scale and texture transfer in detailed scale [83]. Dale
et al. [84] proposed to explore three global restorations: white balance, exposure
correction, and contrast enhancement, based on a set of “semantically” similar
images selected from a database. Though color transfer was performed between
corresponding regions that were segmented according to their visual context, only
global restoration was explored. Hwang et al. [85] proposed to locally enhance the
image by searching for the best transformation for each pixel. The search was done by finding the best candidate image pairs and then finding best matched pixels from these candidate image pairs. The transformation function derived from the best few candidate image pairs on the matched pixel was applied to the input pixel. These techniques directly perform tonal or color adjustments globally or locally across the 2D image plane. None of them has considered the change based on depth. Some methods attempt to personalize image enhancement based on learned styles or models [7, 8, 86–88]. These methods need a lot of input-output pairs to train models or parameters in order to perform the style mapping. For pre-modern painting styles, it is not possible to get such training pairs.

**Processing in gradient domain.** Contrast processing by operating on gradients has been used in high dynamic range compression [89], contrast enhancement [4, 78], and style enhancement [7]. Specifically, work in [7, 89] 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 are used to weight the gradient mapping.

### 3.3 Paintings vs Photographs

Before introducing the contrast organization in paintings and photographs, the separation of depth planes is discussed for depth-aware contrast analysis.

#### 3.3.1 Depth planes

In both painting and photography, the average landscape has three distinct depth planes: FG, MG and BG [68, 90]. The three depth planes are particularly impor-
tant for any landscape composition. FG is the plane closest to the viewer. Things of high pictorial value are usually placed somewhere in the FG. 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, it is considered separately from BG as a separate plane. Therefore, in the chapter, the BG denotes the distant part of the scene without the sky. The MG lies between the FG and BG. Figure 2.7(a) is an example of the separation of FG, MG, BG, and sky. However, for some framing requirements, the MG or BG or sky can be absent in the landscape. For landscape with only a FG and sky, it is meaningless to enhance the atmospheric perspective effect. Therefore, the enhancement of the atmospheric perspective effect in this chapter is targeted at landscapes with two depth planes (FG and BG), three depth planes (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 to partition the depth planes of photographs? In spite of the maturation and commercialization of depth acquisition technologies such as 3D cameras in recent years, technologies for long-range depth acquisition, which is the necessary requirement in the landscape research, are still much more limited. Some depth estimation algorithms are proposed to explore the depth from single image, e.g. “Make3D” [91]. These methods can roughly explore the depth of the scene, but it is still difficult to accurately segment the depth planes based on the estimated depth, especially the MG and BG. Therefore, this chapter proposes to segment the depth planes with minimal user interaction. The interactive segmentation method proposed in [92] can be used to outline the depth planes or they can be outlined manually. The users only need to give several strokes for each depth plane and a mask is generated for each depth plane automatically based on the strokes.
3.3.2 Contrast in paintings and photographs

In this section, a statistical study is conducted to measure the inter-contrast and intra-contrast of a group of paintings and a group of photographs. The group of paintings that is the subject of this study are made by the Hudson River Painters. The Hudson River Painters are a group of stylistically coherent American landscape artists of the late romantic period. They all painted roughly the same subject matter (the area around the Hudson River in the US). 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. Paintings by the Hudson River school artists are quintessential examples of a mastery of atmospheric perspective [70]. The group of photographs were acquired through a simple Google Image search using the search term “Hudson River”. They were thematically similar to the set of paintings. These search results were culled to remove those that had obviously been taken with special lenses, those that were obviously “Photoshoped” photos, and those that had been taken through a color filter. Assuming the saturation has the same property as lightness, the saturation contrast is measured in the same way as lightness.

The intra-contrast is measured by the Gradient Range. The Gradient Range is 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 [89]. The Gradient Range expresses the maximum local contrast of the plane. Based on the contrast definition in Eq. (2.11) which corresponds to the Weber-Fechner law, the inter-contrast of two planes with mean values $M_1, M_2$ is defined as $\frac{2(M_2 - M_1)}{M_1 + M_2}$.

The lightness intra-contrast values shown in Figure 3.2(a) 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 the 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. Apart from the difference in the way that the intra-contrast has been managed between paintings and photographs, there is also a clear difference in the way that their inter-contrast has been managed. The paintings have a clearer decrease in inter-contrast of lightness from MG-FG to Sky-BG (Figure 3.2(b)). Almost all the MG-FG, BG-MG, Sky-BG inter-contrast values of paintings are positive. This indicates that the lightness is increasing from FG to sky in paintings. This is also consistent with the phenomenon of atmospheric perspective.

![Intra-contrast values and Inter-contrast values in paintings and photographs](image1)

**Figure 3.2:** Box plots\(^1\) of lightness intra-contrast values (a) and inter-contrast values (b) in paintings and photographs.

The saturation contrast values in Figure 3.3 shows that the paintings are generally more purposely organized in their saturation contrast than the photographs.

---

\(^1\)A box plot is a way to visually summarize and compare groups of data [93]. It describes the distribution of data samples using a descriptively parameterized box. A box plot is suitable for data distribution that does not follow the normal distribution and where the mean and variance of the distribution are not good representations of the distribution. The Boxplot function in MATLAB was used to create the box plots. The central line in the box is the median of the distribution, and the edges of the box are the 25th and 75th percentiles. The two vertical lines extending from the central box indicate the remaining data outside the central box, which extend maximally to 1.5 times the height of the central box. The red dots are the remaining data which are called as outliers.
Figure 3.3: Box plots of saturation intra-contrast values (a) and inter-contrast values (b) in paintings and photographs.

The FGs of the paintings have a higher saturation intra-contrast than MG and BG consistently. The BG-MG inter-contrast values of saturation in paintings are more likely to be negative than those of photographs (see Figure 3.3(b)). This indicates that the saturation is more 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 organization across the depth planes in their lightness and saturation values. This difference shows that the paintings are more purposely organized in their contrast management than the photographs. This contrast organization in paintings corresponds to an exaggerated form of the physical phenomenon of atmospheric perspective.

### 3.4 Contrast Manipulation

Statistics in the preceding section has clearly shown the intentionally introduced lightness and saturation inter-contrast and intra-contrast organization in paintings which are not prominent in photographs. This part attempts to use this contrast organization of paintings as reference to enhance the acquired digital photographs, whilst keeping the hue unchanged. The pipeline of the proposed method is shown
3.4.1 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 general visual and structural accordance with each other. If they are not, then the contrast values of the photograph are at risk of being pushed further than is visually acceptable. Therefore, the reference painting is selected on the basis of having similar global contextual structure (i.e. it has similar objects such as tree, river, mountain and sky, and similar object arrangements within the scene) and similar context in each depth plane with the input photograph.

The image retrieval technique using quantized SIFT features and spatial pyramid matching scheme that approximates geometric correspondence has demonstrated effectiveness in retrieving semantically-similar outdoor images [84,94,95]. In the proposed reference selection scheme in this chapter, this technique is used to calculate the global contextual similarity and contextual similarity in depth planes. Specifically, a vocabulary of 200 words and 4 level pyramid descriptor is used to
calculate the global contextual similarity. The contextual similarity of each depth plane is measured using a vocabulary of 50 words and 1 level pyramid descriptor. In addition, a reference having a similar lightness level with the input photograph in each depth plane will be preferred. This would return a reference/input pair that are more likely to have been “captured” in similar environmental situations. This lightness constraint helps avoid changing the input too far away from its original appearance. The lightness of each depth plane is described by a 10 bins histogram. The similarity of the lightness histograms and the similarity of pyramid descriptors are calculated by the histogram intersection [94]. The final matching score is the linear combination of the spatial pyramid similarity $K_t$ and lightness similarity $K_l$ by a parameter $\beta \in [0, 1]$, which is defined as $K = \beta K_t + (1 - \beta) K_l$. By experimentation, $\beta = 0.8$ is found to consistently produce good results.

\[
K = \beta K_t + (1 - \beta) K_l.
\]

Figure 3.5: Reference recommendation. (a) Input w/ mask, (b) Recommended references.

Finally the top $n$ ranked paintings are selected as the recommended references. 800 Hudson river paintings with four clear depth planes were collected to form the painting database. Figure 3.5 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 does not have to
have similar color to the input photograph.

3.4.2 Depth-aware contrast mapping

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

The depth-aware contrast mapping has two objectives. One is to manipulate the mean values of the four depth planes in the photograph to align with the inter-contrast values in the reference painting. The other is to adjust the local contrast of each depth plane to acquire the intra-contrast value from the reference painting. To satisfy these two objectives, the depth-aware contrast mapping is modeled as an optimization problem. The energy function, $E(\cdot)$, is defined as follows.

$$E(f) = \sum_{i \in f} [E_d(i) + E_g(i) + E_s(i)]$$  \hspace{1cm} (3.1)

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

$$E_d(i) = w_d(i)[f(i) - d(i)]^2$$  \hspace{1cm} (3.2)

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

The second term is 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, therefore only the gradients on the edges are mapped to those of the reference whilst the gradients elsewhere are preserved. By operating the gradients on only
the edges can also avoid the block artifacts produced by the down sampling of images and the noise caused by the canvas texture in paintings. Therefore, the second term is defined as:

\[ E_g(i) = w_g(i)[f_x(i) - g_x(i)]^2 + w_g(i)[f_y(i) - g_y(i)]^2 \]  \hspace{1cm} (3.3)

where \( f_x \) and \( f_y \) denote the \( x- \) and \( y- \) derivative of \( f \). \( g_x \) and \( g_y \) are the target gradient after intra-contrast mapping. \( w_g \) is the weight for the gradient cost.

\( E_s(i) \) is a gradient constraint term, which is defined as

\[ E_s(i) = w_s(i) \left[ \frac{|g_x(i)|^2}{|F_x(i)|^\theta + \varepsilon} + \frac{|g_y(i)|^2}{|F_y(i)|^\theta + \varepsilon} \right] \]  \hspace{1cm} (3.4)

The gradient constraint is to penalize the large changes to the original gradient. \((F_x, F_y)\) is the gradient of the original input \( F \). \( \theta \) controls the sensitivity to the gradient of \( F \). \( \varepsilon \) is a small regularization constant. The weight \( w_s \) is adaptively defined based on the enhancement scale of the gradient in the intra-contrast mapping.

### 3.4.2.1 Inter-contrast Mapping

Inter-contrast measures the relative difference of mean values of depth planes. Therefore, the 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 is

\[ m_s' = \rho m_r \]  \hspace{1cm} (3.5)
The scale \( \rho \) is 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 plays 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 can be adjusted by \( M_t = M_s + q(M_r - M_s) \).

\( M_s \) is the mean value of the input \( F \). \( q \) is a scale in \([0, 1]\). When \( q = 0 \), the mean value of the output image is preserved to the original one, and when \( q = 1 \), it is moved to that of the reference. \( q \) can be 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 can be used. However, these two methods extend the value out of range \([0, 1]\) and over-compress or over-exaggerate the visual contrast, as the curve shows in Figure 3.6. Also, they are easy to produce artifacts in over-exposed area (see Figure 3.7(c) and (d)). Therefore, a non-linear transform is proposed. Gamma curve is a non-linear transform that is often applied to correct lightness. However, by using the gamma curve, it is easy to over-enhance the dark zone areas which often have hidden block or ring noise. In addition, it also compresses the visual contrast as it enhances more within the dark zone and less within the bright zone (see the magenta curve in Figure 3.6). To avoid over-enhancing of the dark zone area and preserve the visual contrast more effectively, a linear-circular mapping curve (the blue curve in Figure 3.6) is proposed to map the input plane to the target mean value. When \( m'_s > m_s \), the mapping function is

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

(3.6)

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

\[
I'(i) = \begin{cases} 
  y_0 - \sqrt{r - [I(i) - x_0]^2} & \text{if } I(i) \leq m_s \\
  \frac{1-m'_s}{1-m_s} \cdot [I(i) - m_s] + m'_s & \text{if } I(i) > m_s 
\end{cases}
\] (3.7)

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

**Figure 3.6:** The inter-contrast mapping curves. (a) \(m'_s > m_s\), (b) \(m'_s < m_s\).
Figure 3.7: Comparison of mapping curves. (a) Input, (b) Reference, (c) Shifting, (d) Linearly scaling, (e) Gamma-circular, (f) Linear-circular. $q=0.5$ in (c)(d)(e)(f).

In addition to the linear-circular mapping curve, the gamma-circular mapping curve proposed in [77] is also experimented. This mapping curve (the green curve in Figure 3.6) was designed to avoid over-enhancement of the hidden artifacts 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 Figure 3.6(b)). It can properly adjust the lightness contrast in an over-exposed scene whilst avoiding artifacts, as the results show in Figure 3.7(f). In this case, the gamma-circular nonlinear transform produces a similar result with the proposed method (comparing Figure 3.7(e)(f)). When increasing the lightness ($m'_s > m_s$), the proposed method can preserve the perceptual local contrast much better than gamma correction and gamma-circular mapping, as the example shows in Figure 3.8.

Due to adjusting the mean values of the four depth planes separately, the transition around the boundaries between the depth planes is not smooth. To solve this problem, a soft transition map $T$ is created for each plane using the
Figure 3.8: Comparison of mapping curves. (a) Original, (b) Gamma correction, (c) Gamma-circular, (d) Linear-circular. $m'_s > m_s$.

The constraint propagation method proposed in [17]. The constraint weight $w$ is

$$w(i) = \begin{cases} 
0 & \text{if } i \in BA \\
1 & \text{otherwise}
\end{cases} \quad (3.8)$$

where $BA$ is the area around the boundary partitioning planes. $BA$ is created by dilating the boundary with a $21 \times 21$ 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 $F$ is

$$F' = \frac{I'_f * T_f + I'_m * T_m + I'_b * T_b + I'_s * T_s}{T_f + T_m + T_b + T_s} \quad (3.9)$$

In the energy function Eq. (3.1), the desired image having the inter-contrast values of the reference painting is $d = F'$. 
3.4.2.2 Intra-contrast Mapping

According to perceptual studies [96], long coherent edges are perceptually salient to the human visual system even when their gradients are weak. Therefore, the intra-contrast mapping is operated on long and coherent edges (i.e. salient edges) instead of those with only strong magnitude. The long-edge detector proposed in [97] is 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 \( e^o \). Given the gradient \((L_x, L_y)\) of the lightness \( L \), the local gradient saliency calculated by

\[
S(i) = \sqrt{s_x(i)^2 + s_y(i)^2}
\]

\[
s_x(i) = \cos^2(e^o(i)) \cdot e^l(i) \cdot L_x(i)
\]

\[
s_y(i) = \sin^2(e^o(i)) \cdot e^l(i) \cdot L_y(i)
\]

Then, the salient edges are detected as:

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

where \( W \) is the mask of detected salient edges, and \( h \) is the saliency threshold. \( h \) can be calculated from \( \Theta(h) = t \), where \( \Theta \) is the cumulative histogram of the non-zero values in \( S \). \( t \) is in \([0,1]\). The default value of \( t \) is 0.7, which produces results mostly coherent with the perceptual salient edges. The edges detected by Canny detector and salient edges of an image are shown in Figure 3.9. The weak edges in the sky are detected by the salient edge detector but not by the canny edge detector. The edge detection results in other parts of the image are mostly similar by using the two detectors.

Gradient mapping on the edges is performed separately in the depth planes. The contrast on the boundary across depth planes influences the intra-contrast
mapping, as the example in Figure 3.10 shows. The over-enhancement of the cloud in the sky as shown in Figure 3.10(c) is due to the gradients on the boundary between the depth planes. However, when the boundary area (obtained by dilating the boundary line with a $11 \times 11$ structuring element) is excluded from the contrast mapping, the sky contrast is not influenced by the gradients on the boundary and the sky looks more natural (see Figure 3.10(d)).

After detecting the salient edges, the gradient mapping on the edges of each plane is performed by edge-length constrained histogram matching. The constraint
on the edge-length is to assure that the gradient of an edge is mapped to that of a corresponding edge with similar length in the reference painting. The length constraint is conducted by clustering the edges based on the length to three clusters $C_p, p = 1, 2, 3$ which represents the long, medium and short edges. Gaussian Mixture Model is used for the clustering and histogram matching is used to map the gradient distribution of the target to that of the reference in each cluster. The process of histogram matching is expressed as:

$$T_{C_p}(r_k) = \sum_{j=0}^{k} \frac{n_{C_p}^{r_j}}{n_{r_p}}, k = 0, 1, 2, \ldots, N - 1 \quad (3.14)$$

$$F_{C_p}(s_k) = \sum_{i=0}^{k} \frac{n_{C_p}^{s_i}}{n_{s_p}}, k = 0, 1, 2, \ldots, N - 1 \quad (3.15)$$

$$s_k^{C_p} = \min_r \|T_{C_p}(r_k) - F_{C_p}(s_k)\| \quad (3.16)$$

where $n_{C_p}^{r_j}$ is the number of pixels with value $r_j$ in cluster $C_p$ of reference plane, $n_{C_p}^{s_i}$ is the number of pixels with value $s_i$ in cluster $C_p$ of photograph plane. $n_{r_p}$ is the total number of pixels of cluster $C_p$ in reference plane, $n_{s_p}$ is the total number of pixels of cluster $C_p$ in photograph plane. $N$ is the number of discrete levels.

Applying this matching method, gradients at edges of the four planes in $F'$, which are $S_f, S_m, S_b, S_s$, are mapped to be $\tilde{S}_f, \tilde{S}_m, \tilde{S}_b, \tilde{S}_s$. If there is no salient edge in a reference plane, the gradients of the corresponding plane in $F'$ is unchanged.

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

$$\eta_f(P_f) = \tilde{S}_f/S_f, \quad \eta_m(P_m) = \tilde{S}_m/S_m$$

$$\eta_b(P_b) = \tilde{S}_b/S_b, \quad \eta_s(P_s) = \tilde{S}_s/S_s \quad (3.17)$$
\[ \eta = \max(\eta_f, \eta_m, \eta_b, \eta_s) - W_s + 1 \]  

where, \( P_f, P_m, P_b, P_s \) are the indexes of the salient edges of the four depth planes. \( W_s \) represents the mask of salient edges of the input photograph. Finally, the gradients after intra-contrast mapping are given by

\[ (G'_{sx}, G'_{sy}) = \eta \cdot (G_{sx}, G_{sy}) \]  

where, \((G_{sx}, G_{sy})\) are the gradients of the inter-contrast mapped result \( F' \). The target gradients in the second term of Eq. (3.1) are \( g_x = G'_{sx} \) and \( g_y = G'_{sy} \).

**3.4.2.3 Optimization Solver**

The minimization of the energy cost function \( E \) in Eq. (3.1) can be achieved by solving the linear function \( Af = b \), where

\[
A(i, j) = \begin{cases} 
-w_g(i) - \frac{w_s(i)}{(F(i) - F(j))^2 + \epsilon} & j \in N_4(i) \\
-w_d(i) - \sum_{k \in N_4(i)} A(i, k) & i = j \\
0 & \text{otherwise}
\end{cases}
\]  

\[ b(i) = w_d(i)d(i) - w_g(i) \sum_{k \in N_4(i)} (F(k) - F(i)) \]  

\( N_4(i) \) are the four neighbors of pixel \( i \). In the implementation, constant values are used for weights \( w_d = 0.01, w_g = 1 \). In order to penalize large changes to the original gradient, the weight \( w_s \) is defined as

\[ w_s(i) = \lambda \left( 1 - e^{-\frac{(\eta(i) - 1)^2}{\sigma^2}} \right) + \epsilon \]  

Therefore, a larger weight is used for pixels having a large gradient change and a small weight is used for pixels which are not on salient edges. The parameters are
Figure 3.11: (a) Reference w/ mask, (b) Input w/ mask, (c) Inter-contrast mapping result, (d) Result using $E_d$ and $E_g$ in Eq. (3.1), (e) Result with $E_s$

$\theta = 0.5$, $\lambda = 0.06$, $\varepsilon = 0.0001$ and $\sigma = 3$. The example in Figure 3.11 shows the effect of each part in Eq. (3.1). Figure 3.11(c) has the target inter-contrast values of lightness and saturation only, while the result using $E_d$ and $E_g$ in Figure 3.11(d) shows the enhancement on the local contrast. The comparison of the result in Figure 3.11(d)(e) illustrates clearly the influence of $E_g$. The large gradient changes on the salient edges which may create sharp edge artifacts are avoided in the result using $E_g$.

3.5 Experiments and Discussion

This section will analyze the performance of the proposed depth-aware contrast manipulation algorithm. The algorithm was implemented in MATLAB on a PC with an Intel 2.93GHz processor and 3GB RAM. For a photograph-reference image pair of a pixel dimension of 600×800, it takes around 35 seconds to obtain the results. Solving the optimization function is the most time-consuming step of the proposed algorithm. It is possible that by optimizing the code and using
acceleration schemes the computation time can be shortened.

The depth planes of paintings are not derived from the real depth map. They are, effectively, flat “cards” that serve to underpin the artists’ structural conventions. The depth planes of paintings in the database are drawn manually.

**Influence of the parameter q.** The parameter $q$ is expressed as $q_L$ for lightness and $q_C$ for saturation. Figure 3.12 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 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 Figure 3.13). When $q$ increases towards 1, the mean values of the four depth planes of the target are changed towards to those of the reference. Due to the constraint on the lightness in the reference selection, the selected references have similar lightness values to the input photograph. However, the saturation of the reference may be much higher than the input. The reference in the example in Figure 3.12 has distinctly greater saturation than the original photograph as the curves show in Figure 3.13. According to studies of visual ap-
Figure 3.13: The mean values of the four depth planes when using different parameters $q_L$ and $q_C$ in Figure 3.12.

...
by the Algorithm 1. The recommended \( q_C = q_{rc} = \frac{M_{Cr} - M_{Cs}}{M_{Cr} - M_{Cs}} \). \( M_{Cs}, M_{Cr} \) are the 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. \( \alpha = \alpha \cdot \frac{C_{rc}}{C_{ft}} \)
3. The target mean \( M_{Ct} = \alpha M_{Cs} \)
4. Update \( C \) of the photograph by Eq. (3.6) and Eq. (3.7), 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. (3.23), \( C_{ft} = \sigma_{a'b'} + 0.94 \mu_{C'} \)
   
   **if** \( |C_{ft} - C_{fr}| > e \) **then**
   
   Return to step 2
   
   **else**
   
   Output \( M_{Ct} = \alpha M_{Cs} \)
   
   **end if**

**Performance.** The contrast manipulation results of the photograph in Figure 3.12(a) using three recommended paintings are shown in Figure 3.14. The organization of inter-contrast and intra-contrast of lightness and saturation of the three reference paintings are consistent with the observation in the contrast analysis in Section 3.3.2, 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 enhancement result are successfully aligned from those of the photograph to those of the reference as the graphs in the third to sixth rows show. The visual organization 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 are calculated. The matching error is 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
Figure 3.14: The contrast manipulation results based on three selected references in Figure 3.5. The input with mask is in Figure 3.12(a). Images in the first row are the references with masks. 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.
of the test results are shown in Figure 3.15. The average matching errors of inter-contrast and intra-contrast are less than 0.03. This shows that enhanced target images by the proposed method exhibit significant similarity with the reference paintings in the depth-aware contrast organization governed by atmospheric perspective effect. 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. (3.1) to avoid large changes on the gradient.

**Comparison.** The proposed example-based depth-aware contrast manipulation method is compared with three existing example-based contrast adjustment methods in Figure 3.16. 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 is performed on $L$ and $C$ channels in LCH space. Bae’s two scale tone mapping method [83] is applied on the $L$ channel in CIELAB space. $a$ and $b$ channels are scaled by the enhancement scale of $L$ to adjust the saturation. Another method is the example-based contrast enhancement by gradient
mapping (CEGM) [78]. Three automatic contrast enhancement methods: “Lo-fi” filter in Instagram®, “Auto Contrast” in Adobe® Photoshop®, and local gradient optimization in [4] are also compared with the proposed contrast manipulation method in Figure 3.16.

The global contrast adjustment methods (HM and Bae’s method) only match
the global value distribution and have no spatial consideration. CEGM, “Auto Contrast” in Adobe® Photoshop®, and local gradient optimization in [4] methods only enhance the contrast in local area without regional consideration to depth. “Lo-fi” filter in Instagram® enhances both the local contrast and 2D center-corner difference. Differently, the proposed method enhances the local contrast and adjusts the contrast between depth planes. In the results by the proposed method, both the visual appeal and illusion of depth obeying the law of atmospheric perspective effect are enhanced. The results in (j) and (t) are the Chroma increased version of (i) and (s) respectively to match the colorfulness of the reference painting. The visual color “change” in the results in (j) and (t) comparing with 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, the method by histogram matching in each depth plane in the spatial domain is also experimented. The comparison with the histogram matching in depth plane and comparison with the previous depth-aware contrast manipulation method proposed in [5] are given in Figure 3.17. The results using histogram matching (see Figure 3.17(c)) are not so pleasing as a consequence of having more artifacts. The local contrast adjustment achieved using the proposed method is clearly more effective than that achieved using the method in [5]. The proposed method does not produce artifacts even in low quality image with compression blocks.

In order to show the advantage of the proposed atmospheric perspective effect enhancement method comparing with the atmospheric physical model used in [6], the physical model on lightness and saturation is applied to the test photographs and compared with the proposed method. Depth is estimated based on the outlined depth planes. The physical model is only used on the ground planes and not on the sky. Figure 3.18(b) shows the result that is produced by using the physical model
Figure 3.17: Comparison with depth-aware methods. (a) Input w/ mask, (b) Reference w/ mask, (c) HM in depth plane, (d) Method in [5], (e) Proposed.
on the photograph in Figure 3.1(a). The result in Figure 3.18(b) shows that both the close object and distant object will be affected when the atmosphere 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 so 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 it 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 to strengthen their separation, thereby the illusion of depth is enhanced. In the results by the proposed method, the contrast of FG is generally modified to be larger than that in the original and has a bigger difference with other depth planes.

![Figure 3.18](image)

**Figure 3.18:** Comparison with the physical model used in [6]. The input photograph and reference are those used in Figure 3.1. (a) Estimated depth, (b) Result by [6], (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).

The style enhancement in [7] enhances the input image using learned mapping parameters. The style-enhanced result in [7] is used as the reference and contrast manipulation on the input image is performed using the proposed method. The result is comparable with the style-enhanced result in the depth view with the exception of some local difference, as shown in the example in Figure 3.19.

**Application on other photograph categories.** Apart from landscapes with four depth planes, the proposed atmospheric perspective effect enhancement...
Chapter 3. Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

Figure 3.19: Comparison with style enhancement in [7]. (a) Input w/ mask, (b) Result in [7], (c) Proposed method. Image in (b) is the reference.

Figure 3.20: 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 w/ mask, (b) Reference w/ mask, (c) Result.

User study. A user study was carried out to evaluate the effectiveness of the proposed method on the enhancement of atmospheric perspective effect. The user study was designed to test the illusion of depth and visual appeal in the results. 30 participants (10 females and 20 males) with ages ranging from 20 to 35 were invited. All of them are with a normal vision. 25 of the test results
Table 3.1: User study result.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Selection Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhanced Result</td>
<td>Original</td>
</tr>
<tr>
<td>Which has the clearer illusion of depth?</td>
<td>82.4%</td>
</tr>
<tr>
<td>Which is more pleasing?</td>
<td>69.9%</td>
</tr>
</tbody>
</table>

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 the 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 perspective in 2D images. No further information on the goals of the experiment or the origin of the images was provided to the participants. Then the 25 image pairs were shown one by one to the participants.

The result of the user study is shown in Table 3.1. Higher percentages of responses were in favor of the enhanced results. By using a one-sample, one-tailed t-test on these responses, the chances that results by the proposed method were perceived as having a clearer illusion of depth was significantly larger than 50% (p ≪ 0.0001), and the chances that results by the proposed method were perceived as being more pleasing was also significantly larger than 50% (p < 0.0001). This concludes that the proposed method has significantly enhanced the illusion of depth and visual appeal of the photographs, thereby enhanced the atmospheric perspective effect.

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

**Aesthetics measure.** Here we want to show how the proposed method influence 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 [35]. They measure the selection of palette, the order in the color 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 color distribution. Given the normalized histogram of luminance $Y_{709}$, $P(x)$ ($x$ in $[0, 255]$), the global aesthetics measure using Shannon’s perspective is

$$M_b = \frac{H_{\text{max}} - H_p}{H_{\text{max}}}$$  \hspace{1cm} (3.24)

$$H_p = -\sum_x P(x)\log P(x)$$  \hspace{1cm} (3.25)

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

$$M_j = \frac{\sum_{i=1}^{4} \pi_i H(p_i)}{H_p}$$  \hspace{1cm} (3.26)

where $\pi_i$ is the area of region $i$, $H(p_i)$ is the entropy of region $i$ calculated by Eq. (3.25).

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
Chapter 3. Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

Figure 3.21: Aesthetics assessment. (a) $M_b$, (b) $M_j$.

Plots of the values are shown in Figure 3.21. The larger $M_b$ values of the input photographs comparing with those of the references reflect the higher color homogeneity in the photographs. The larger $M_j$ values of the references comparing with 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 with those of the references. In other words, the aesthetics of the photographs is enhanced to be close to that of the reference paintings by using the proposed method.

An aesthetic quality inference engine (ACQUINE) has been developed to auto-
matically rate natural photographs [100]. Aesthetic scores for the 25 photographs, 25 corresponding reference paintings and the results are collected by uploading them to acquine.alipr.com. ACQUINE is to rate the aesthetics for 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 proposed. 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:

$$r_o = \frac{S_O - S_I}{S_R - S_I}$$

(3.27)

When $r_o > 0$, the measured aesthetics of the result is changing close to the reference, and when $r_o = 1$ the measured aesthetics of the result is the same with the reference. The distribution of $r_o$ for the 25 test results in Figure 3.22 again shows that the aesthetics of the photographs is enhanced to be close to that of the reference paintings.

![Figure 3.22: Ratio of aesthetic score.](image)

**User study for depth plane partition.** Another user study to see to what
degree does people have consensus in partitioning a given image was conducted. 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 by an art expert 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 [101–103]. It is used here to measure the overlap of two people’s partitionings for the same image. The intersection-over-union ratio of region A and B is defined as

$$R = \frac{S(A \cap B)}{S(A \cup B)}$$  \hspace{1cm} (3.28)

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 put more weight on areas far away from the partition boundary. Because each photograph is partitioned by 20 participants, and each time two people’s partitioning are 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 3.2. 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 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 3.2, the partitions by the 20 participants have a very low intersection-over-union ratio. However, the 20 partitions generally can be classified to two types as the two masks shown in Figure 3.23(b)(c). Among the 20 participants, 6 of them outline FG, MG, and BG for the 13th photograph in a similar way as the mask shows in Figure 3.23(c). The results using the two different partitions are different (see Figure 3.23(e)(f)). This user study shows that the partitions from different people are not uniform. The partition result is mostly depended on personal understanding of the scene and not a standard rule. This is an unsolved problem in the research area of image segmentation.

![Image](image_url)

**Figure 3.23:** Results using different partitions. (a) Input photograph, (b) Mask 1, (c) Mask 2, (d) Reference w/ mask, (e) Result using mask 1, (f) Result using mask 2.

**Limitations.** While experiments have demonstrated the effectiveness of the proposed method, we observe a few failure cases. The sky is fragmented by tree
Table 3.2: Intersection-over-union ratio of human drawing study for each photograph.

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

branches as the example in Figure 3.24(a) shows, some of the small fragments are not detected as sky and some fine tree branches are also not segmented out in the sky plane. As a result, the color of some small sky fragments differs from the rest of the sky and the color of the miss-segmented fine tree branches differs from other parts of the tree (Figure 3.24(c)). Histograms of lightness and saturation of the enhancement result, original photograph and reference painting in Figure 3.24 are shown in Figure 3.25. The histograms show that, by using the proposed depth-aware contrast manipulation method, the lightness and saturation distribution ranges are moved closer to those of the reference. In other words, the lightness
Chapter 3. Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

Figure 3.24: An example demonstrating the limitation of the proposed method. (a) Input photograph, (b) Reference painting, (c) Result, (d) Close-up Region.

contrast and saturation contrast are enhanced to align with those of the reference painting.

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.

![Distribution of Lightness and Saturation](image)

**Figure 3.25:** The comparison of pixel value distribution among the original photograph, reference and enhancement result in Figure 3.24. (a) Lightness histogram, (b) Saturation histogram.

More experimental results by the proposed method are shown in Figure 3.27.

**Discussion on night scene photographs.** Atmospheric perspective is visible at night. The proposed method can also be used on night scene photographs. Two
experimental results on night photographs are shown in Figure 3.26. The two
night scene photographs are captured in a city with man-made lighting. However,
in pre-modern paintings, the night scene is drawn with only natural light (e.g.
moon light), and no man-made lighting. The two paintings in Figure 3.26(a) show
the moon light scene. The atmospheric perspective effect also exists in these night
scene paintings, where there is clear contrast organization between and within
depth planes to create the illusion of depth. After using the two night scene
paintings in Figure 3.26(a) as references to enhance the night scene photographs
in Figure 3.26(b), the illusion of depth is improved (see Figure 3.26(c)). Hence the
proposed atmospheric perspective enhancement method can successfully enhance
the atmospheric perspective effect in the night photographs.

3.6 Concluding Remarks

This chapter proposes an example-based depth-aware contrast manipulation method
to transfer the atmospheric perspective effect in landscape paintings to landscape
Chapter 3. Atmospheric Perspective Effect Transfer for Landscape Photographs through Depth-aware Contrast Manipulation

Figure 3.27: More results. The first, third and fifth rows: original photographs. The second, fourth and sixth rows: results using proposed method.

photographs. More specifically, the following contributions are reported in this chapter. First, a statistical study is conducted to explore the difference of contrast organization governed by atmospheric perspective between traditional paintings and “snapshot” photographs. The conclusion of the study is that the contrast governed by atmospheric perspective in paintings is more purposefully organized. This provides the justification for the proposed depth-aware contrast manipulation of photographs. Second, a novel linear-circular mapping curve is proposed to adjust the mean values to map the inter-contrast. The proposed mapping curve can preserve the perceptual contrast, and avoid creating over-enhanced artifacts in low
value zones. Third, an effective gradient mapping algorithm is developed to map the intra-contrast. Finally, an optimization problem is formulated that simultaneously considers the desired inter-contrast, intra-contrast, and specified gradient constraints. Experiments demonstrate that the enhanced target images exhibit significant similarity with the reference paintings in the depth-aware contrast organization. Additionally, a user study is conducted to assess the effectiveness of the proposed method. This study demonstrates that both the visual appeal and illusion of depth governed by atmospheric perspective effect in the photographs are enhanced using the proposed method.

In this chapter, four depth planes are 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 require several discrete depth planes and does not require the real depth which is a challenge to capture for long range landscape photographs.
Chapter 4

Regional Contrast Manipulation for Portrait Photograph Enhancement

This chapter proposes a method to manipulate the regional contrast in snap-shot 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 organization 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.
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

4.1 Introduction

Figure 4.1: One of the results produced by the proposed method. (a) Original photograph, (b) Reference painting, (c) Enhancement result.

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 picture or a natural scene. Similar to the strategy of using different modes or lenses to capture landscapes or portraits, different kinds of photos have a different subject focus and different formal needs, therefore they need to be edited in different ways. This chapter addresses the specific genre of portrait photographs which are photographs focusing on the depiction of a person. Whether you are traveling, enjoying a family event, or recording your glory moment, everyone ends up shooting portraits. Portrait photos constitute an important part of our personal photo collections.

In a portrait, the figure, especially the face, is predominant. However, the key to controlling the strength of the attraction of a portrait is not by just controlling the value of the center of interest, but rather the contrast among the face and skin areas of the foreground figure, the non-face/skin part of the foreground figure with its accessories, and the background [104]. The management of contrast among all these regions makes the figure or the face stand out. Therefore, the proposed por-
trait photograph enhancement method does not focus on the enhancement simply on the face, but on the relative relationships of the face and skin areas of the foreground figure, the non-face/skin part of the foreground figure with its accessories, and the background. A content-aware automatic photo enhancement method was proposed in [9] in order to enhance the face, sky and salient areas differently. However it only processed these regions separately and did not consider the relative relationships of these regions. In addition, the face and sky in all the photos were shifted to a pre-defined exposure or color. As a result, the contrast on the face was sometimes reduced and also the relative contrast between the face and background might be reduced. Although Adobe® Photoshop® provides a powerful tool for users to manually edit their images, the skills required to achieve this region-specific improvement are also beyond the casual user. Users have to mask regions and make decisions based on the relative relationships of these regions. Additionally, defining the relative relationships for these regions in order to make the figure, especially the face, more attractive is a skill that is not easy to acquire.

Fortunately, portrait paintings supply us with good examples which can be used to modify the relative relationships of these regions. Compared with photographs, which can be regarded as a projection of the physical nature, a traditional portrait painting is a very carefully constructed piece. Portrait artists are able to filter and manipulate the subject scene based on their visual perception [105]. Key to this is the consideration given to the organization of color contrast values [106]. In the example painting shown in Figure 4.2(a), the foreground figure is darker than the outdoor background (BG). In the painting in Figure 4.2(b), 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 region. Foreground (FG) in this chapter is comprised of the figure and its accessories. Despite the contrast between these regions (inter-contrast), there is also contrast organization within each region (intra-contrast) to emphasize the face, the pose of the body and also the richness of the BG. Differently, in snap-shot style photographs which are more casually framed (“point and shoot”), their lighting, colorfulness and the regional contrast are not well organized (e.g. the two photographs in Figure 4.2(c)(d)).

![Figure 4.2:](image)

(a) “Portrait of Mariana Waldstein”, by Francisco de Goya (1746-1828). (b) “Mrs. Peter William Baker” by Thomas Gainsborough (1727-1788). (c) and (d) are photographs. (d) is cropped from an image in MIT-Adobe FiveK Dataset [8].

In order to better understand the regional contrast organization of paintings, a statistical comparison on the contrast values of lightness and saturation has been conducted between photographs and paintings. Hue is not touched so as to preserve the natural appearance of the photographs. Differences between the paintings and the snap-shot photographs are observed with evidence that the intra-contrast and inter-contrast of the lightness and saturation in the paintings are more purposefully organized. This contrast organization in paintings purposely makes the figure, especially the face, stand out in the image, 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 organization in paintings.

Therefore, in this chapter, a method is developed by which the high-level lightness and saturation contrast organization of portrait paintings can be transferred to portrait photographs to enhance their visual appearance and the expression of figures. 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 with the depth-aware contrast manipulation discussed in Chapter 3, 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 enhancement results obtained from the proposed method is shown in Figure 4.1.

4.2 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 recognized emotions [107]. The artistic styles in the NPR system were based on painting materials, color palette and brush model. Some recent research also attempts to render a portrait photo to an artistic style [108–111]. However, the styles are limited to simulating abstraction, line drawing styles or organic models. Face relighting is another popular research topic about portrait. In [112], 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 one reference face image [113]. The work in [114] applied the face lighting template generated by artist on faces of photographs. All the mentioned research work above on the subject of portraits only focuses on the manipulation of the face. Differently, the method proposed in this chapter not only considers the face, but also the relationships of the face with the BG and other parts of the figure.

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 [21]. 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 artifacts 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 photo enhancement method proposed in [9] aims to enhance the face, sky and salient areas differently, but it does 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. [8] proposed to learn the photographic global tonal adjustment function from photographers to personalize the 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 color and tone mapping to stylize the image enhancement [7]. These meth-
ods mostly rely on lower-level image statistics (e.g. color, intensity gradient) and do not exploit object-level semantics other than the face in [8]. Even if the face is specifically considered in [8] for learning the tone mapping, only global tone mapping function is learned.

Joshi et al. proposed to improve the quality of faces in a personal photo collection by leveraging photos of the same person which have been selected for their quality [23]. This work only focused on performing both global and face-specific corrections. Hwang et al. [85] proposed to locally correct the color and tone of the 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 finding the best matched pixel from these candidate image pairs for each input pixel. The search was based on the local scene descriptors and context. Then, the mapping derived from the image pair on the matched pixel was applied to the input pixel. This example-based method needs the enhancement parameters for the example image pair known as \textit{a priori}. For pre-modern painting styles, it is not possible to know the enhancement parameters, therefore the method is not applicable.

4.3 Paintings vs Photographs

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

Portrait paintings collected for this study are chosen from artists in the 17-18 centuries such as: Thomas Gainsborough, Fransisco Goya, Edouard Manet, and
Jean-Auguste-Dominique Ingres. These artists usually maintained a high dynamic range in their paintings which was broadly similar to that of photographs. The database covers a variety of environments and relationships of FG and BG. In total, 300 portrait paintings are collected to form the database. 180 of them are half body portrait paintings and 120 of them are full body portrait paintings. 250 portrait photographs from the MIT-Adobe FiveK Dataset [8] and 50 from our personal collection are collected to form the photograph dataset.

Similar to the inter-contrast definition in Chapter 3, the inter-contrast between two regions with mean values \( M_1 \), \( M_2 \) is defined as \( \frac{2|M_2-M_1|}{M_1+M_2} \). Here only the absolute inter-contrast value is 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 is not used. This is because the portrait is very sensitive to the gradient change on the edge, especially around the face. Therefore, the intra-contrast of one region is defined based on the value in the spatial domain, which is \( \frac{I_{\text{max}}-I_{\text{min}}}{I_{\text{mean}}} \), where \( I_{\text{max}} \) and \( I_{\text{min}} \) are the maximum and minimum values of one region respectively, and \( I_{\text{mean}} \) is the mean value of the region.

Box plots of the mean values and contrast values are shown in Figure 4.3. The mean values show that the FS region is generally the brightest part of the painting and much more colorful 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 lightness of photographs. 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 4.3(b) 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.
### Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

#### Figure 4.3: 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.
In the intra-contrast comparison, the paintings and photographs are classified into two classes based on the complexity of the BG. One class is with a complex BG (intra-contrast is shown in Figure 4.3(c)), the other class is with a simple BG (intra-contrast is shown in Figure 4.3(d)). In the paintings 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 the photographs with a simple BG are larger than those with a complex BG. The reason is that the figure that has been set against a simple BG is more focused. These differences in intra-contrast organization between paintings and photographs show that artists exaggerate the contrast of regions in paintings. The intra-contrast of a complex BG is exaggerated to be larger, and the intra-contrast of a simple BG is reduced to be smaller.

In summary, the regional contrast of paintings is more purposely organized than that of photographs. Higher inter-contrast is used by painters to attract the focus of attention to the face or figure in portrait paintings. The intra-contrast of a region is also specially organized by painters to serve the visual appeal of the painting.

4.4 Example-based Portrait Photograph Enhancement

Statistics in the preceding section clearly shows the purposely created lightness and saturation inter-contrast and intra-contrast organizations of paintings. This section
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

reports the attempt to use this contrast organization as reference to enhance digital photographs. The framework of the proposed example-based portrait photograph enhancement method is shown in Figure 4.4.

![Figure 4.4: Framework of the proposed method.](image)

This example-based method starts with the segmentation of the FG/BG and face/skin area detection. Although a lot of research have been done on human detection, it is still a challenge to extract the accurate human shape in 2D images. Moreover, in half body portrait where only the face and part of the upper body are visible, most human detection methods do not work. This chapter proposes to use the GrabCut interactive segmentation [115] to segment the FG/BG automatically with an initialization window. One thing consistent in portraits is that, the front face is generally always facing the viewer. Therefore, the face is used as the most important information for the purpose of detecting the position of the figure. The Haar face detector is used to detect the face [116,117]. Then, based on the size and location of figures, an enlarged window that is likely to cover the FG area is defined to initialize the GrabCut. When the detected single face is around the
center of the image horizontally, the enlarged window centers at the face window. When the face window is in the right or left part of the image, more space is 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 [115], a user interface is provided for users to draw some strokes to refine the results. The face mask and skin areas are detected based on the skin color in the figure using the method in [118] with some morphological post-processings. Some results are shown in Figure 4.5.

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

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

4.4.1 Reference selection

The main objective of reference selection is to select reference paintings that can create more pleasing results as compared to the original input, while not pushing the original image too far away from its original natural property. Because the proposed portrait enhancement method is 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 modeled as a graph as shown in Figure 4.6. The features within regions $\Psi = \{\Psi_S, \Psi_O, \Psi_B\}$ and the relationships among regions $\Phi = \{\Phi_{SO}, \Phi_{SB}, \Phi_{OB}\}$ are used to select reference paintings, where $\Psi_S$, $\Psi_O$, and $\Psi_B$ are the features within FS, FO and BG respec-
Figure 4.5: 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 initialization window for GrabCut. The blue and purple lines are the FG and BG brushes for refining the FG/BG segmentation.

tively, and $\Phi_{SO}$, $\Phi_{SB}$, and $\Phi_{OB}$ 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 produce more pleasing results to the original input photograph should be ranked higher in the selection process in terms of graph similarity according to some distance metric. Therefore,
distance metric learning is proposed to determine the similarity of paintings and input photographs based on the graph model.

![Figure 4.6: The graph model.](image)

The features within each region are 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 $G_B$ is also calculated as a feature to express the complexity of the BG. Highly detailed and variation-rich image BG will have a high global contrast factor and simple BG will have a low global contrast factor [119]. In this manner the contrast factor can represent the difference between simple and complex BGs. The global contrast factor is calculated as the weighted average of local contrast values at various resolution levels (more details can be found in [119]). The relationships among regions $\Phi$ are 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

$$D(i, j) = \sum_{k=1}^{N} \alpha_k (v_i^k - v_j^k)^2$$  \hspace{1cm} (4.1)$$

where $\alpha$ is the parameter to linearly combine the distance of features. $v_i^k$ is the $k^{th}$ feature in the feature vector formed by concatenating the ordered graph attributes

...
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

of photograph $I_i$, and $v^k_j$ is the $k$-th feature in the feature vector of painting $R_j$.
The objective here is to determine the parameter $\alpha$ such that paintings which can
be more likely to create more pleasing results are closer to the input photograph
than others. As in [85, 87], the parameter $\alpha$ is learned by using a target distance
function $D_t(i, j)$. The parameter $\alpha$ is determined by minimizing the objective
function

$$\arg\min_\alpha \sum_{i,j} \|D(i, j) - D_t(i, j)\|^2$$ (4.2)

The target distance $D_t$ is 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 un-
acceptably enhanced photographs. Minimizing this objective function returns an
appropriate distance function that reflects how far the photograph and painting
should be in terms of their enhancement results. This objective function is con-
 vex 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) are used as references
to generate enhanced photographs. The target distance function is $D_t(i, j) = \frac{1}{S(i, j)}$, 
where $S(i, j)$ is the assigned score for enhanced photograph $I_i$ using painting $R_j$
as reference by an expert who has lot of experience on photo adjustment. It is a
challenge to score an image directly for its aesthetic value due to subjective pref-
ferences. However, it might be more convenient to classify enhancement results
to be as good, acceptable or unacceptable. Therefore, instead of directly giving
a score, an expert is asked to classify the enhancement for 32 half body portrait
photographs and 30 full body portrait photographs to three categories: good, ac-
ceptable and unacceptable with pre-set score values. The highest score is given
to good enhancement, lowest score is given to unacceptable enhancement, and the
score for acceptable enhancement is in the middle. For each half body portrait photograph, 180 enhanced photographs are obtained while 120 enhanced photographs are obtained for each full body portrait photograph. In all the enhanced photographs, an expert only classifies the ones that can be confidently assigned to one of the three categories for them, otherwise they are discarded. Finally, 2776 samples are collected to learn the distance metric for half body portraits and 1381 samples for full body portraits.

4.4.2 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 we could adjust the mean, maximum and minimum values of regions in photographs to map them to those in reference paintings for the contrast mapping. Given the target mean value $m_t$, maximum $I^t_{\text{max}}$, minimum $I^t_{\text{min}}$, the task of the contrast mapping is

$$I(m_s, I_{\text{max}}, I_{\text{min}}) \xrightarrow{f} I'(m_t, I^t_{\text{max}}, I^t_{\text{min}}) \quad (4.3)$$

Piecewise linear contrast stretching has been used in [120] 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 artifacts [18]. Moreover, this method may produce out of range mapping. To more effectively perform contrast mapping, an interpolating transformation curve is piecewisely defined and it is monotonically increasing and continuously differentiable. This interpolating transformation curve is known as histogram warping in [18,121]. For a given set of control points $(a_k, b_k, d_k)$, where $b_k = f(a_k)$,
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

$\delta_k = f'(a_k)$ which is the contrast adjustment at the key value $a_k$, the transformation curve is generated using a piecewise rational quadratic interpolating spline [122]:

$$f(x) = b_{k-1} + \frac{(r_k t^2 + d_{k-1}(1-t)t)(b_k - b_{k-1})}{r_k + (d_k + d_{k-1} - 2r_k)(1-t)t}$$ (4.4)

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 [18], the control points are specified by users while the control points are generated automatically based on the analysis of the histogram in [121]. Differently, in this chapter, three control points are automatically defined based on the original and target mean, minimum and maximum values. Specifically, the three control points are $(a_1, b_1, d_1) = (I_{\min}, I'_{\min}, d_1), (a_2, b_2, d_2) = (m_s, m'_s, d_2)$ and $(a_3, b_3, d_3) = (I_{max}, I'_{max}, d_3)$, 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 the work presented in [121], the contrast adjustments at the control points are 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

$$d_k = \left(\frac{b_k - b_{k-1}}{a_k - a_{k-1}}\right)^{g^l_k} \left(\frac{b_{k+1} - b_k}{a_{k+1} - a_k}\right)^{g^h_k}$$ (4.5)

where $g^l_k = \frac{F(a_k) - F(a_{k-1})}{F(a_{k+1}) - F(a_k)}$ and $g^h_k = \frac{F(a_k) - F(a_{k-1})}{F(a_{k+1}) - F(a_k)}$ are the slope weights. $F(x)$ is the cumulative distribution function. An example of the transformation curve is shown in Figure 4.7(a).

Good contrast within an image is not simply a case of the higher being the better. Artists are skilled at controlling the perceptually high contrast values of their paintings. However, giving high computational contrast to the photograph cannot ensure a good perceptual look. Therefore, the target inter-contrast and intra-contrast may not need to match exactly to those in the reference. This
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

Figure 4.7: Contrast mapping. (a) Transformation curve, (b) Convergence of maximum, minimum and mean values.

means that 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 original image and the user’s preference. Thus, two parameters $\alpha$, $\beta$ are used to control the inter-contrast and intra-contrast mapping subject to the original image and 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 is 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 is shifted toward to $m_r$ and the magnitude of shift is controlled by a weight $\alpha$ in $[0-1]$.

$$m'_s = m_s + \alpha(m_r - m_s)$$  \hspace{1cm} (4.6)

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

$$c'_s = c_s + \beta(c_r - c_s)$$  \hspace{1cm} (4.7)

where $\beta$ is a weight in $[0-1]$. Based on the definition of intra-contrast, the degree
of enhancement on the maximum-minimum will be \( \lambda = \frac{\mu}{\sigma m} \). Then the target maximum and minimum values are set as

\[
I^t_{\text{max}} = \lambda (I_{\text{max}} - m_r) + m_s^t
\]

(4.8)

\[
I^t_{\text{min}} = \lambda (I_{\text{min}} - m_r) + m_s^t
\]

(4.9)

where \( I_{\text{min}} \) and \( I_{\text{max}} \) 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 Figure 4.7(b), three iterations are sufficient to map the mean, maximum and minimum to the target values.

Finally, to smooth the boundary between regions and keep the change spatially coherent, edge-preserving smoothing [123] is 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 uses the following energy function:

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

(4.10)

where

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

(4.11)

\( D \) is the desired smoothed difference image, \( (D_x, D_y) \) is the gradient of \( D \), \( (g_x, g_y) \) is the gradient of \( F \). \( \theta \) controls the sensitivity to the gradient of \( F \). \( \varepsilon \) is a small regularizing constant. The smoothing is to make the gradients of the output \( D \) as small as possible, unless the input \( F \) has significant gradient. The weight \( w_s \) is defined as

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

(4.12)

\( \tau_1 \) is larger than \( \tau_2 \) 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 minimization of $E$ can be achieved using standard or weighted least-squares techniques like the conjugate-gradient method.

### 4.5 Experiments and Discussion

In this section, the performance of the proposed portrait photograph enhancement method will be analyzed. Among the 300 photographs, 62 of them are used for the distance metric training in the reference selection, the others are used as test photographs.

**Influence of reference.** The enhancement results of two photographs using five selected top ranked reference paintings are shown in Figure 4.8. 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 Figure 4.8), even though the face and figure also stand out in the images, the change is too far away from its original natural property (e.g. the pink T-shirt becomes black). Therefore, the reference selection method is effective for the portrait enhancement. While the two example photographs in Figure 4.8 are of outdoor scenes with complex BGs, Figure 4.9 shows two indoor photographs with relatively simple BGs. The first one is with bright BG and dark FG, and the other is with dark BG and bright FG. We can see that the selected reference paintings have similar natural properties with 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.
Figure 4.8: 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. 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 $\alpha_l = 0.6$, $\alpha_s = 0.6$, $\beta_l = \beta_s = 0.5$.

Effect of parameters. In the contrast mapping, the same $\alpha$, $\beta$ are used for the three regions FO, FS, and BG. Parameters for mapping lightness are expressed as $\alpha_l$, $\beta_l$, and for saturation $\alpha_s$, $\beta_s$. When $\alpha_l = \alpha_s$, $\beta_l = \beta_s$, they are indicated as $\alpha$, $\beta$. $\alpha$ and $\beta$ control the degree of match to the contrast of the reference. Figure 4.10 shows the effect on the appearance of the enhancement results for changing $\alpha_l$, $\beta_l$ from 0 to 1. When $\alpha_l$ increases from 0 to 1, the inter-contrast of lightness is moved towards to that of the reference (see Figure 4.10(c)). The figure stands out and becomes brighter as $\alpha_l$ increases in Figure 4.10. When $\beta_l$ changes from 0 to
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

Figure 4.9: 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 for the two photographs and the corresponding results.

1, the intra-contrast values of lightness in the three regions are moved to those of the reference correspondingly (see Figure 4.10(d)). More local details in the BG become visible when $\beta_l$ increases. The influence of $\alpha_s$, $\beta_s$ is in similar way.

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 the parameters around the default values to change the appearance according to the personal preference. Paintings are generally drawn with high contrast between regions and with high saturation values. If $\alpha = 1$, the enhanced photograph may become visually unacceptable. Generally $\alpha_l$ is in the range $[0.3, 0.7]$, and $\alpha_s$ is in $[0.2, 0.6]$. $\beta_l$ and $\beta_s$ 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 is compared to the content-aware enhancement (CAE)
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

Figure 4.10: Influence of $\alpha, \beta$. (e)-(i): $\alpha_s = 0.4$, $\beta_l = \beta_s = 0.5$. (j)-(n): $\alpha_l = 0.5$, $\alpha_s = 0.4$, $\beta_s = 0.5$. 

(a) Input photograph
(b) Reference
(c) Inter-contrast
(d) Intra-contrast
(e) $\alpha_l = 0$
(f) $\alpha_l = 0.2$
(g) $\alpha_l = 0.5$
(h) $\alpha_l = 0.8$
(i) $\alpha_l = 1.0$
(j) $\beta_l = 0$
(k) $\beta_l = 0.2$
(l) $\beta_l = 0.5$
(m) $\beta_l = 0.8$
(n) $\beta_l = 1.0$
method in [9] which has been reported to perform better than the global learning-based enhancement method in [8] and recent 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 by the proposed method are also compared with the results enhanced by using the average mean and contrast values. The average mean and contrast values are obtained from the study on contrast organizations in paintings shown in Figure 4.3. This is referred as the AVE method in this chapter. Figure 4.11 shows 5 results enhanced using the three methods. It shows that faces and figures in the results by the proposed method become more salient as compared to those in the results obtained by the CAE method. Moreover, the light falling on the face in the results obtained by the proposed method are more naturally presented. However, in the results on the second and third rows using the CAE method, the light falling on the face becomes unnatural because the lightness has been processed differently on the left side of the face, as compared to the right. The enhancement results using the AVE method in the first to third rows are comparable with those obtained by the proposed example-based method. However for the photographs in the forth and fifth rows, the proposed example-based method performs better than using the average values. The example-based method can handle the diversity of the scenes better.

**User studies.** In order to significantly evaluate the effectiveness and advantage of the proposed method, two user studies were conducted. The purpose of the first user study 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 user study 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
Figure 4.11: Comparison with other methods. (a) Reference paintings, (b) Input photographs, (c) Results by CAE in [9], (d) Results by AVE method, (e) Results by the proposed method. The photograph in the fifth row is from [9].

show that the results obtained using the proposed method were more pleasing and the figure and face become more salient. The second was to show that the results 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 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 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. 5 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. 5 paintings from the ranks in the range of 106-120 for full body portrait and 5 paintings in the ranks 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 can select “Group 1”, “Group 2” or “Group 3”.

![Figure 4.12: 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.](image)

The 40 photographs were divided into 7 sets. The first set had 4 photographs
while the other sets had 6. The comparison in each set contains 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 select to complete one, two or three sets of comparisons. Finally there were 38 participants, 21 were males and 17 were females. The age ranged from 20 to 45. Totally, 2070 samples were collected for the comparison of enhanced photographs vs originals. The average selection rate of enhanced photographs vs originals is shown in Figure 4.12(a). The bar graph shows that a significant majority of the responses (more than 90%) from the participants are in favor of the enhanced results by the proposed method in both questions. This demonstrates that, by using the proposed method, the visual appeal of the portrait photographs is effectively improved and the face and the figure 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 favor of the results obtained using the top ranked reference paintings. This concludes the effectiveness of the proposed reference selection method and the importance of reference selection.

The second user study was conducted to compare the results of enhancements by the proposed method against those by the CAE method in [9] and the AVE method. In [9], 55 results for images with face were reported. 36 of them which could be considered as portrait were used for this user study. The results by the proposed method and one of the other methods were shown as a pair side by side. The left/right ordering of the images in each pair was randomly generated. For
Table 4.1: User study result for the comparison of the proposed method with the CAE method in [9] and the AVE method. The percentage shows the number of users responding favorably to the respective method.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Proposed</th>
<th>ACE</th>
<th>The same</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>62.0%</td>
<td>30.6%</td>
<td>7.4%</td>
</tr>
<tr>
<td>Q2</td>
<td>66.8%</td>
<td>29.6%</td>
<td>3.6%</td>
</tr>
</tbody>
</table>

Proposed vs AVE

<table>
<thead>
<tr>
<th>Questions</th>
<th>Proposed</th>
<th>AVE</th>
<th>The same</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>68.4%</td>
<td>19.6%</td>
<td>12.0%</td>
</tr>
<tr>
<td>Q2</td>
<td>70.0%</td>
<td>22.1%</td>
<td>7.9%</td>
</tr>
</tbody>
</table>

each image pair, the participants were asked 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 select to complete one or two sets. There were a total of 29 participants aging 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 4.1. It shows that the proposed method are preferred by a significantly higher percentage of the responses than other methods. This is not surprising, because the method in [9] 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 Chapter 3, the gradient is used to map the intra-contrast. However, the portrait is very sensitive to the change of gradient especially that on the face. It is easy to see artifacts in the area of high gradient magnitudes in a portrait. In the example in Figure 4.13(d), there is clear artifacts 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 Chapter 3 in Figure 4.13 shows that the proposed method in this chapter is more suitable to manipulate the 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 skin area detection based on color may still be disturbed by the color of clothes. In the example shown in Figure 4.14(a), the color 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 Figure 4.14(b)). After regional con-

**Figure 4.13:** Comparison with the method in Chapter 3. (a) Original, (b) Reference, (c) Result by the proposed method, (d) Result by the method in Chapter 3.
Figure 4.14: 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.

Contrast manipulation using the proposed method, these regions differ slightly from the rest of the clothes region (see Figure 4.14(c)).

Another limitation is that artifacts near the transition boundary between the brightened FG region and the dark BG region may be visible. This is because the brightness of the FG region is extended to the neighboring BG region by the boundary smoothing if there is no clear boundary between FG and BG in the original photograph. The technique in [123] was designed to smooth an image while preserving edges with large gradients. In the photograph in Figure 4.14(d), there is no clear edge between the right lower arm and ground. Therefore the brightness on the lower arm is extended to the neighboring BG region by using the edge preserving smoothing (see Figure 4.14(e)). The smoothing around the boundary region is stronger than that of the non-boundary region. Hence, artifacts may be visible. Some noise may also become visible in face and skin areas after getting brightened using the regional contrast manipulation. More experimental results by the proposed method are shown in Figure 4.15.
Figure 4.15: More experimental results. The first and fourth rows: reference paintings. The second and fifth rows: original photographs. The third and sixth rows: results.
Chapter 4. Regional Contrast Manipulation for Portrait Photograph Enhancement

Application using RGBD images. Along with the development of techniques and commercial off-the-shelf cameras, an RGB color image with depth (D) information is more easily captured. While the Xbox 360 Kinect depth sensor [124] by Microsoft can only capture depth for indoor scenes, the light field camera (e.g. Raytrix camera [125]) 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 producing parallax. With depth, the figure can be easily extracted from the image. The framework of the proposed application using RGBD images is in Figure 4.16.

![Figure 4.16: Application using RGBD images. The depth from the video is the depth layer obtained by optical flow in [10].](image)

The main element that influences the FG extraction from depth is the ground plane. For the continuity of the ground and the connection with the feet, the FG can not be extracted directly using depth range threshold. However, the ground plane has a clearly different orientation to the figure. The orientation of a plane can be expressed by the normal vector or the gradient change on the plane. There are two options to divide the figure from the ground. One is to segment the planes 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 [126] ©2011 IEEE). Surface normals have been used for range image segmentation [127]. 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

\[
\vec{N} = \vec{P}_1 \times \vec{P}_2 \\
\vec{P}_1 = (x + w, y, z(x + w, y)) - (x, y, z(x, y)) \\
\vec{P}_2 = (x, y + w, z(x, y + w)) - (x, y, z(x, y))
\] (4.13)

where \((x, y)\) is the top-left corner pixel of the \((w+1)\) by \((w+1)\) support region. To handle the quantization 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 in clustering. The initial cluster centers and the number of clusters are defined based on the prior knowledge of the portrait image. In a portrait image, the bottom part is the ground, the top part is the upper BG, and the figure is near the center. Therefore, three cluster centers are initialized by the feature of the points in the bottom left corner, the feature of the top left corner points and the feature in the face area. For half body portrait, 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 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 be 2. An example of the clustering result is shown in Figure 4.17(c). Post-processing is performed using
Figure 4.17: FG extraction from RGBD image. (a) RGB image, (b) Depth, (c) k-means clustering, (d) Mask of extracted FG.

morphological operations to remove the miss-segmented small regions. The final mask of the figure is shown in Figure 4.17(d). The k-means clustering method can generate results in realtime. However, for a scene with complex BG, the method is not robust. In the example in Figure 4.18, the k-means clustering can not 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. Then, region grouping can be used to extract the FG by using a seed point as the center of the detected face area. The plane fitting method based
on normal vectors followed by graph-cut segmentation [11] can efficiently detect the ground plane in RGBD image. The extracted FG of the image in Figure 4.18(a) using region grouping after removing the ground plane detected by the method in [11] is shown in Figure 4.19(b). However, the method used in [11] is slower and non-realtime.

Based on the observation that the gradient of the depths on the figure and ground is much different in the y direction, as in [12], 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 filter response corresponds to the ground plane. The extracted FG of the
image in Figure 4.18(a) using region grouping after removing the detected ground plane is shown in Figure 4.19(c). This method can perform in realtime. However, the detected leg and foot areas are a little noisy compared to those produced using the method in [11]. The enhancement results using the extracted FGs by the two methods are shown in Figure 4.19(e)(f). They show that the proposed method is robust. Considering the speed, the $F$ filter proposed in [12] is selected to detect the ground plane for the FG extraction. Results for half body and two-person portrait photographs are shown in Figure 4.20. In the second row of Figure 4.20(e), the lower part of the photograph is adjusted to be darker to have a larger contrast against the bright background regions and FG. This contrast organization in the result is closer to that of the reference painting.

![Figure 4.20: More results for the application in RGBD images. (a) RGB images, (b) Depths, (c) Extracted FG masks, (d) Reference paintings, (e) Enhancement results.](image-url)
4.6 Concluding Remarks

This chapter proposes a method to improve the visual appeal and focus of attention of portrait photographs by manipulating the regional contrast. This improvement is guided by a selected portrait painting that is being used as an example. Experimental results and user study confirm that the figure and the face become more salient and the photograph is more pleasing by applying the proposed method. This chapter also discusses the application of the proposed method on RGBD images.
Chapter 5

Vignetting Effect Manipulation

This chapter discusses the vignetting effect in paintings and photographs which is created by organizing the center-corner contrast. First, the vignetting effect in paintings is compared to that in photographs by analyzing their lightness distributions. The observation of the differences between paintings and photographs shows that the vignetting effect in paintings is more purposely produced with regards to its aesthetic composition. The vignetting patterns in paintings are also analyzed. They show that the vignetting effect in paintings is naturally presented based on the geometry of the scene. Then an algorithm is developed to transfer the lightness weighting extracted from an example painting to a photograph to create the painter-style vignetting effect. Experiments show that the proposed algorithm can successfully transfer the vignetting effect from an example painting to a photograph and the vignetting effect is more naturally presented with regard to aesthetic composition comparing with popular software tools and camera models.

5.1 Introduction

As introduced in Chapter 2, vignetting is an effect that is clear in the center and fades off at the edges by reducing the image’s brightness or saturation at the
periphery in photography. The purposely introduced vignetting effect can create artistic effect, such as drawing attention of the viewer to the center of the image. In photographs, the effect can be created by camera settings or post-processing. This photo-style vignetting creates a clear circle in the image.

Artists also use vignetting in paintings by organizing center-corner contrast. The vignetting effect in painting, which is called painter-style vignetting, helps to draw attention of the viewer to the center (in portrait) or guide the eye of the viewer to the far away scene (in landscape), see the two example paintings in Figure 5.1. The painter-style vignetting is naturally presented by gradually fading off the elements near the periphery and does not leave a definite line on the border of the paintings [128]. This painter-style vignetting can be produced by fading off two or three areas unevenly or symmetrically in the periphery of the painting. In the example painting in Figure 5.1(a), the corners are willfully darkened to serve compositional needs. This vignetting is a common pictorial device well known to professional photographers and painters.

Based on the observation on the differences in photo-style vignetting and painter-style vignetting, this chapter attempts to study the painter-style vignetting and
explore an algorithm to apply this painter-style vignetting on photographs for the purpose of an artistic effect. To our knowledge, till now little attention has been given on the vignetting effect in paintings.

5.2 Related Work

Most of the research on vignetting is for the correction of vignetting in photographs caused by camera settings or lens limitations. Correcting the vignetting effect has a wide applications such as image segmentation, image mosaicking and camera calibration. The main methods used are radiometric calibration by estimating the radiometric response function and vignetting function from a sequence of collected images [129, 130], and vignetting function estimation from a single image by looking for slow consistent intensity variations in the radial direction [131, 132].

For the creation of a vignetting effect in photographs, Samii et al. mentioned to stylize false vignetting effect in images [133]. This work was done by learning experts’ artistic decision processes to enhance a new photograph. This learning method needs a set of training data that maintain the history of adjustments applied to every photograph by experts. It is a challenging work to collect such training data, especially for pre-modern painting styles. In photorealism rendering, Wu et al. [134] proposed to simulate the vignetting effect based on a lens model to estimate the effect of vignetting in Bokeh effect.

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®, “Lo-fi” filter in Instagram®. However, these photo-style vignetting effects enforce a darker or desaturated corner or edge without the consideration on the geometric structure of the image, and they create a clear unnatural circle around the image.
Different from the work mentioned above, this chapter proposes to learn the vignetting effect in paintings by analyzing the distribution of the lightness and then apply the learned lightness weighting pattern to photographs.

5.3 Vignetting Effect in Paintings

This section describes how vignetting effect is organized in paintings by analyzing the spacial distribution of values. Landscape paintings by Hudson River painters are used as the dataset for this analysis.

First, the pixel-by-pixel average values of the lightness of 920 paintings are calculated to see how the lightness is organized in the 2D image plane. The average lightness of 920 paintings are shown in Figure 5.2(a). In addition, the pixel-by-pixel average lightness values of 920 landscape photographs are calculated to compare with those of the paintings. 300 of the landscape photographs are collected from the MIT-Adobe FiveK Dataset [8] and others are collected from our personal collection. All the photographs are non-professional and without post-processing. By comparing the average lightness of paintings and photographs, we can clearly see that corners of the former have been significantly more weighted (being darker) than the latter. The changes of the average lightness along a column and a row have similar tendency in the paintings and photographs (see the curve in Figure 5.2(e)(f)). However, the difference of the lightness values in the center and periphery is more distinct in paintings than in photographs. 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 composition.

The pixel-by-pixel average saturation values of the 920 paintings and photographs are also calculated to see the change of the saturation in the 2D image plane (see Figure 5.3). There is a general tendency of decrease from the center
Figure 5.2: (a) The average lightness image of 920 paintings, (b) The average lightness image of 920 photographs. (a) and (b) are calculated by resizing all the paintings or photographs to the same size ($600 \times 800$) and then calculating the pixel-wise average of the lightness image. (c) and (d) are the contrast enhanced version of (a) and (b). (e) and (f) are the plots of the average values of respectively 800 columns and 600 rows for the two average lightness images in (a) and (b).

to the periphery in the average saturation of paintings. However, there is larger variation in the local distribution of saturation comparing with lightness. The de-
creasing tendency from center to periphery in saturation is not so consistent as in lightness. This chapter thus focuses on analyzing the organization of lightness in the vignetting effect.

![Figure 5.3: (a) The average saturation image of 920 paintings, (b) The average saturation image of 920 photographs, (c) and (d) are the contrast enhanced version of (a) and (b).](image)

In the painter-style vignetting, the lightness weighting on the periphery is not circularly even. More weight can be given in one, two or three corners. Based on the lightness weighting on corners in paintings, the vignetting is divided into 11 patterns as the simple sketches in Figure 5.4 show. The number under each pattern is the number of paintings that are in this pattern. It shows that the main patterns in the painter-style vignetting are P1, P2 and P11. P1 is called as left weighed pattern, and P2 is right weighed pattern while P11 is the normal pattern. An example painting in each of the three patterns and the average lightness of all
the paintings in each of the three patterns are given in Table 5.1. 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 relatively brighter corners provide us a sense of open space. In the normal pattern P11, the corners on the sky are relatively darker than the center, but they are not of the same darkness with the ground. This means that the darkness on the sky and ground should be processed differently. From the example painting and the corresponding average lightness weighting in each pattern in Table 5.1, we can see that the painters use the lightness weighting pattern in their paintings according to the geometry of the scene which might be enforced by selecting points of view.

![Lightness Weighting Patterns](image)

**Figure 5.4:** Lightness weighting patterns. In the sketch, D is darker and B is brighter. The number under each pattern is the number of paintings that are in this pattern in the 920 paintings. In pattern P11, there is no clear bias on the top or bottom two corners and the lightness in the top corners are in similar brightness or darkness.

In portrait paintings, the figure especially the face is the most important element which is placed near the center. The high contrast between the figure and background is usually used to draw the attention on the figure. The background is usually drawn simply dark or with faint details. The vignetting effect in the background of the portrait paintings is not so clear as in the landscape paintings (see the average lightness of the portrait paintings in Figure 5.5).

In summary, the vignetting effect in paintings is naturally presented by darkening objects on the periphery based on the geometry of the scene and drawing
Table 5.1: The example painting and average lightness of all the paintings in each of the three patterns P1, P2 and P11.

<table>
<thead>
<tr>
<th>Patterns</th>
<th>P1 (Left Weighted)</th>
<th>P2 (Right Weighted)</th>
<th>P11 (Normal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>209</td>
<td>192</td>
<td>449</td>
</tr>
<tr>
<td>Example painting</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Average lightness</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

attention on the center while preserving contrast on the dark objects. Additionally, the darkness on the sky and ground are introduced differently.

5.4 Transfer of Painter-style Vignetting to Photographs

The analysis of the vignetting effect in paintings clearly shows that the key technique used by artists to naturally present the vignetting effect is to introduce the lightness weighting based on the content of the scene. Although, the general lightness weighting patterns have been obtained from paintings, they cannot be simply used as the template to manipulate the lightness in photographs due to the variation of 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 with the photograph.
Given a photograph without vignetting, the objective of this chapter is to create the vignetting effect in the photograph. The objective can be formulated as

\[ L_{\text{out}} = f(L, W) \]  

(5.1)

where \( L \) is the lightness of the input photograph without vignetting, \( W \) is the lightness weighting map. \( L_{\text{out}} \) 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 of creating vignetting effect are to generate the lightness weighting map \( W \) and define the blending function \( f \). Based on the observation 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, or normal etc.)
2. Enhancing the contrast between center and corner.
3. Preserving the local contrast on the periphery and darken objects.
4. Manipulating the sky and ground differently.

5.4.1 Lightness weighting pattern and blending function

The selection of lightness weighting pattern is to select an example painting from the painting database and then extract the lightness weighting pattern from the example painting. Based on the assumption that images similar in content and geometry can have similar lightness weighting on corners, the example painting having such similarity with the input photograph is selected. Similar to example painting selection in Chapter 3, the example painting is selected based on its local correspondence with the input photograph. The recommended top ranked paintings which have similar geometric structure with the two example input photographs are shown in Figure 5.6. The user can select one of the top ranked paintings as the example painting.

After selecting the example painting, its lightness weighting pattern is extracted by estimating the lightness bias field. Given an image $Z$ with lightness bias, it can be modeled by 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

$$Z_i = I_i B_i$$  \hspace{1cm} (5.2)

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 optimization problem based on the constraints on $B$ and $I$. The estimation of $B$ can be achieved by the following optimization function.

$$B' = \arg \max_B P(B \mid Z) \propto \arg \max_B P(Z \mid B)P(B)$$  \hspace{1cm} (5.3)

The probability models of $P(Z \mid B)$ and $P(B)$ are constructed based on the sparse-
ness priors of the gradient probability distribution in [135]. They are used to estimate the lightness bias field. For the three paintings in Table 5.1, the extracted lightness bias fields are shown in Figure 5.7. The estimated lightness bias field presents a coarse 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 to blend it with the original lightness while satisfying the criteria 2 to 4? The three blending techniques, known as Darken, Multiply and Overlay in Photoshop®, are adopted here. Two layers are created following [136]. The upper layer contains the weighting pattern extracted from a painting and the lower year contains the input photograph. The
Figure 5.7: The extracted lightness bias fields for the three paintings in Table 5.1.

Darken function given in [136] is as follows:

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

(5.4)

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 is to simply multiply the lower layer value by the normalized upper layer value [136]. Given the range \( N \) of \( M \), the Multiply mode is

\[
L' = L \times \frac{M}{N}
\]

(5.5)

Eq. (5.5) shows that \( L'(i, j) \leq L(i, j) \) for every pixel \((i, j)\). Therefore, the Multiply mode makes the whole image darker than the original one.

Differently, the Overlay mode is to darken the region that is dark in \( M \) and

...
brighten the region that is bright in $M$ [136]. The influence of $M$ also depends on the value of $L$. The formulation is

$$L'(i, j) = \begin{cases} \frac{M(i, j)(N - L(i, j))}{2} + L(i, j) - (N - L(i, j)) & \text{if } L(i, j) > \frac{N}{2} \\ \frac{M(i, j)L(i, j)}{2} & \text{otherwise} \end{cases}$$ \hspace{1cm} (5.6)$$

The effects of the Darken, Multiply and Overlay modes can be clearly seen in the Figure 5.8. The Overlay mode satisfies the criteria for the painter-style vignetting. It brightens areas that are bright in the lightness weighting pattern and darkens areas that are dark in the lightness weighting pattern. It also darkens the sky and ground differently based on their original values. In addition, the local contrast is preserved to some degree.
5.4.2 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 lightness may become too bright or too dark, as the two examples in Figures 5.9 and 5.10 show. The result in Figure 5.9(d) is over-brightened on the sky and the result in Figure 5.10(d) is generally darker than both the original and the example painting. In order to avoid over-brightened or over-darkened regions, the lightness weighting 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 to have similar values with those of the example painting. Given the correction function $f_c(\cdot)$, the corrected weighting pattern is

$$M' = f_c(M) \quad (5.7)$$

First, the potential bright regions and dark regions after blending need to be determined to limit the values. 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 areas corresponding to the 25% brightest pixels in $M$ are considered as the potential bright areas $B_0$ in the blended output and the areas corresponding to the 25% darkest pixels in $M$ are considered as the potential dark areas $D_0$ in the blended output. The corresponding areas covered by $B_0$ and $D_0$ in the original lightness may have bright or dark pixel values. Therefore, the bright regions $B_s$ are determined as the corresponding areas of the 50% brightest pixels covered by $B_0$ in the
Figure 5.9: Lightness weighting pattern correction. (a) Original photograph, (b) Example painting, (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.
Figure 5.10: Lightness weighting pattern correction. (a) Original photograph, (b) Example painting, (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.
original lightness and the dark regions $D_s$ are the corresponding areas of the 80% darkest pixels that have a lightness value $L < 0.5$ and overlapped by $D_0$. The regions highlighted in red in Figures 5.9(e) and 5.10(e) are the bright regions determined after the lightness weighting. The regions highlighted in blue in Figures 5.9(e) and 5.10(e) 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 regions $B_r$ (red regions in Figures 5.9(f) and 5.10(f)) 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 Figures 5.9(f) and 5.10(f)).

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 is to satisfy:

$$A_{B_r} = f_o(A_{B_s}, M_B) \quad (5.8)$$
$$A_{D_r} = f_o(A_{D_s}, M_D) \quad (5.9)$$

where, $f_o(\cdot)$ is the Overlay blending mode in Eq. (5.6). $M_B$ 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 $M'_{B_s}$ and $M'_{D_s}$ 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 $a=[0, M'_{D_s}, M'_{B_s}, 1]$ and $b=[0, M_D, M_B, 1]$. Then, the nonlinear curve is interpolated based on the function in Eq. (4.4). 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 curves for the two examples are shown in Figures 5.9(g) and 5.10(g). By using the corrected lightness
weighting pattern, the over-brightened and over-darkened regions are successfully avoided in the results (see Figures 5.9(h) and 5.10(h)).

5.4.3 Content-aware interpolation

Although the selected example painting has similar geometry with the input photograph, there are still some variations in the spacial organization of the structure. As the two examples in Figures 5.12 and 5.13 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 the consideration of the variation in the content, the weighting pattern is interpolated adaptively to fit the structure of the photograph.

In landscape paintings, painters darken the foreground and sky differently, and also the foreground, middle-ground and background planes are darkened in different degrees. Here, “darken” is to make the periphery area of one plane darker, not to change the lightness of the whole depth plane as done in the atmospheric perspective effect manipulation. Similar to the Chapter 3, the landscape images are divided to depth planes, generally foreground (FG), middle-ground (MG), background (BG) and sky. The weighting pattern at each plane is 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 in the photograph. The registration is performed by shifting the location by a vector \((s_x, s_y)\) to match the center of the plane in the photograph and scaling its size by \(\alpha\). \( M_r \) 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
Figure 5.11: Content-aware interpolation of the lightness weighting pattern.

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 by the linear interpolation using the data from the combined weighting pattern. The hole-filled weighting pattern $M_s$ is not smooth on the boundary between regions. A smoothed weighting pattern $\tilde{M}_s$ is generated by using the edge-preserving smooth [123] based on the following optimization function:

$$\tilde{M}_s = \arg\min_{M_s} \sum_{p \in M_s} \{[\tilde{M}_s(p) - M_s(p)]^2 + w_s(p)h(\nabla \tilde{M}_s, \nabla L)\}$$  \hspace{1cm} (5.10)

$$h(\nabla \tilde{M}_s, \nabla L) = \frac{[\tilde{M}_s^x(p)]^2}{|g_x(p)|^p + \varepsilon} + \frac{[\tilde{M}_s^y(p)]^2}{|g_y(p)|^p + \varepsilon}$$  \hspace{1cm} (5.11)

The smoothing is to keep the gradient of the output $\tilde{M}_s$ as small as possible, unless the original lightness $L$ has 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 $L$. $\theta$ controls the sensitivity to the gradient
Figure 5.12: (a) Original photograph, (b) Example painting, (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 by content-aware interpolation, (h) Result using weighting in (g).

of $L$. $\varepsilon$ is a small constant. The weight $w_s$ is

$$w_s(p) = \begin{cases} \tau_1 & \text{if } p \in \text{boundary area} \\ \tau_2 & \text{otherwise} \end{cases} \quad (5.12)$$
Figure 5.13: (a) Original photograph, (b) Example painting, (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 by content-aware interpolation, (h) Result using weighting in (g).

$\tau_1$ is larger than $\tau_2$ 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 Figure 5.11. Using
the interpolated lightness weighting pattern, the left part of the sky in Figure
5.12(h) and the right part of the sky in Figure 5.13(h) are not as dark as the FGs.

## 5.4.4 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 mode can preserve the local details to some degree, the local contrast are still reduced in dark regions. In order to preserve the local details, an optimization function is used to restore the reduced local contrast in $\tilde{L}$. The lightness $L_o$ with restored local contrast is calculated as

$$L_o = \arg \min_{L_o} \sum_{i \in L_o} [E_d(i) + E_g(i)]$$  \hspace{1cm} (5.13)

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

$$E_d(i) = w_d(i)[L_o(i) - \tilde{L}(i)]^2$$  \hspace{1cm} (5.14)

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

$$E_g(i) = w_g(i)[f_x(i) - g_x(i)]^2 + w_g(i)[f_y(i) - g_y(i)]^2$$  \hspace{1cm} (5.15)

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 let the output to have the gradients of the original input. Since gradient is one expression of local contrast, the output with the restored gradients 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 Figure 5.14 shows the effect of the local
5.5 Experiments

In this section, the performance of the proposed vignetting effect manipulation method will be analyzed. First, the difference between the vignetting effect manipulation and atmospheric perspective effect enhancement is discussed. The atmospheric perspective effect enhancement is to enhance the difference of depth planes for creating the depth perception in 2D images. The values within a depth plane are adjusted in the same way. Differently, the vignetting effect is to adjust the contrast between center and corner to attract the attention of viewers to the center. In each depth plane, the region near the center is weighted differently from the region close to the periphery. The difference of the vignetting effect and at-
mospheric perspective effect can be clearly seen by comparing the two rendering results in Figure 5.15(a)(b). In paintings, the vignetting effect and atmospheric perspective effect are used together for aesthetic composition. 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 manipulated 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 Figure 5.15(c). If the atmospheric perspective effect is first enhanced followed by the vignetting effect manipulation, there is no exaggeration or reduction in the contrast (see Figure 5.15(d)).

Figure 5.15: The original photograph is in Figure 5.12(a) and the example painting is in Figure 5.12(b). (a) Result by vignetting effect manipulation, (b) Result by atmospheric perspective effect enhancement, (c) Vignetting + atmospheric perspective effect enhancement, (d) Atmospheric perspective effect + vignetting enhancement.

In order to show the advantage of the proposed vignetting effect manipulation
method, it is compared with two popularly used camera models for 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 (called camera model 1) is the natural vignetting model which is based on natural illumination radially falling-off due to geometric optics [137]. The formula of this vignetting model is

\[ V_m = V_o \cdot \cos^4 \left( \tan^{-1} \left( \frac{r}{f} \right) \right) \]  \hspace{1cm} (5.16)

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 center. The second camera model (called camera model 2) is the Kang-Weiss vignetting model proposed in [138]. In the two models, the focal length is set as \( f = 600 \). Figures 5.16 and 5.17 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 center. All the pixels in the same distance are darkened in the same way, 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 center. The pixels close to the periphery are enforced to be dark without consideration about the content. There is also a clear circle in the result. Differently, the vignetting effect in the result by the proposed method is more naturally presented without a clear circle and the sky and ground is darkened in different degrees following the aesthetic composition in the example painting.

Some more experimental results on landscapes are shown in Figure 5.18. The proposed vignetting effect manipulation method can also be applied for portrait photographs. Because the estimated bias field by [135] only describes the coarse distribution of the lightness, the bright face surrounded by dark regions cannot be extracted as bright in the bias field, see Figure 5.19(c). The face in the interpolated
Figure 5.16: Comparison. (a) Original photograph, (b) Example painting, (c) Lightness weighting pattern extracted from the example painting, (d) Result by the proposed method, (e) Result by camera model 1, (f) Result by camera model 2, (g) Vignette in Photoshop®, (h) “Lo-fi” filter in Instagram®.
Chapter 5. Vignetting Effect Manipulation

Figure 5.17: Comparison. (a) Original photograph, (b) Example painting, (c) Lightness weighting pattern extracted from the example painting, (d) Result by the proposed method, (e) Result by camera model 1, (f) Result by camera model 2, (g) Vignette in Photoshop®, (h) “Lo-fi” filter in Instagram®.
weighting pattern based on the outlined foreground and background is also dark, see Figure 5.19(d). According to the portrait painting analysis in Chapter 4, the face and skin areas are almost the brightest parts in the figure. Therefore, based on this prior knowledge on portrait paintings, the lightness weighting on the face and skin areas is corrected to be bright (set to the 95 percentile value of the lightness weighting on the figure). Figure 5.19 (e) 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 areas correction in Figure 5.19(f) shows similar lightness

Figure 5.18: More experimental results on landscape photographs. (a) Original photographs, (b) Lightness weighting patterns, (c) Results.
weighting with the example painting. In the original portrait photograph Figure 5.19(a), the bottom right hand is darker than the hand in the left part. In the lightness weighting pattern from the reference painting, the area around the bottom right hand is also darker than the left part. By using the lightness weighting pattern from the reference to the photograph, the left hand part is enhanced more than the right hand part. This lightness spatial distribution is similar with that of the reference painting. Moreover, in the result in Figure 5.19(f), the figure is brightened and the corners in the background become darker. This manipulation makes the figure stands out in the photograph and draws the attention of viewers to the figure which is the intended effect. Two more examples of the application on portrait photographs are shown in Figure 5.20.

5.6 Concluding Remarks

This chapter proposes an example-based vignetting effect manipulation method in photographs. The proposed method is based on the analysis of the vignetting effect in paintings which is called painter-style vignetting. The painter-style vignetting in paintings is more purposely produced for aesthetic composition and it is naturally presented based on the geometry of the scene. An algorithm is developed to apply the painter-style vignetting on photographs while satisfying four manipulation criteria observed from paintings. First, the lightness weighting pattern is extracted from the selected example painting which has similar geometry with the input photograph. Then, the Overlay mode is applied to blend the lightness weighting pattern with the original lightness to enhance the center-corner contrast and process the sky and ground differently based on their original values. In order to avoid over-brightened or over-darkened regions in the result, the extracted lightness weighting pattern is corrected using a nonlinear curve to limit the dark regions
Figure 5.19: Application on 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 in (e).

and bright regions in the blended result to have similar values with those of the example painting. A content-aware interpolation method is proposed to warp the lightness weighting pattern of the example painting to fit the structure of the photograph. The content-aware interpolation makes the algorithm to be robust to the variation of the geometric structure between the example painting and the input photograph. Finally, the local contrast is restored using an optimization function
Figure 5.20: Two more examples of the application on portrait photographs. (a) Original photographs, (b) Lightness weighting patterns, (c) Results using weighting patterns in (b).

to preserve the local contrast even in darkened objects. Experiments show that, by using the proposed method, the vignetting effect from the example painting is successfully transferred to the photograph. The comparison with popular software and camera models has shown the advantage of the proposed method.
Chapter 6

Composition Improvement for Portrait Photographs

This chapter studies the composition in portrait paintings and develops an algorithm to improve the composition of portrait photographs based on an example portrait painting. The observation from the portrait painting study shows that the placements of the face and the figure in the portrait painting are pose-related and not exactly rule-based. Based on this observation, this chapter develops an algorithm to improve the composition of a portrait photograph by learning the placements of the face and the figure from an example portrait painting. The example portrait painting having similar figure pose with the input photograph is selected. This example painting selection is modeled as a graph matching problem. Finally, space cropping is performed using an optimization 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.
6.1 Introduction

Composition, being one of the important aesthetic aspects that influence the visual quality, requires attention to be improved in photographs. In portrait paintings, where the placements of the figure and the face are the most important elements in the composition [65], the size of the figure and the position of the figure are specially organized by artists to draw all the attention on the subject. In contrast, in snap-shot style photographs by amateurs, the size of the figure and the spatial composition are often not well organized. When we are selecting personal photos for sharing in social networks or for printing for home display, we often have to discard some photos with a small face or with the subject in corners or edges even we like the good smile or beautiful pose in them. In this situation, the composition can be improved first before sharing or printing. Motivated by this, this chapter proposes to improve the composition of portrait photographs by optimizing the size and position of the figure in the image.

Currently, rule-based methods have been developed to improve the composition of portrait photographs. Rule-based methods are to optimize the composition based on some rule constraints (e.g. rule of thirds, one of fifth or center). The main feature they considered is the location of the face, because the face is the main focus of portrait. These rules can produce photographs more likely to be pleasing than placing the face in an arbitrary location. However, no absolute rules can ensure good composition in images. Artists and professional photographers frame the portrait depending more on their intuition and perception based on the content they want to describe, rather on the rules. In portrait paintings, artists use the emotion on the face and the pose to express the personality of the figure. When planning a portrait, the artist first carefully finds a natural pose for the subject that looks active [139, 140]. Then, the artist studies the subject for a facial expression that satisfies his concept of the subject’s essence [139]. Finally, the artist
will compose the portrait by selecting the best view [140]. In this final composition process after selecting the body pose and face pose, it is based on some rules and perception of the artist. In other words, the rules are dependent on body pose and face pose. Hence, we can say that the composition in portrait is pose-related, not exactly rule-based. Inspired by this, this chapter proposes to improve the composition of a portrait photograph according to an example portrait painting in which the figure has similar pose with that in the photograph. Space cropping is used to improve the composition while preserving the important subject.

Firstly, a graph model is developed to describe the features of a pose. The example painting is selected by searching a matching graph. Then, space cropping is conducted using an optimization function considering the location of each body part.

6.2 Related Work

**Composition in aesthetics measure.** Image composition takes an important role in image aesthetics. Almost all methods that measure the aesthetics of images use composition as part of their features. Rule of thirds is one popular rule used in the aesthetics measure [38, 141, 142]. It is the rule that the center of important region should be placed on one of the power points (the four cross points of the horizontal and vertical one-of-third lines). Instead of rule of thirds, the golden section or golden ratio also is used as an important composition rule in the aesthetics measure [143, 144]. Obrador et al. [145, 146] proposed to use simplicity of the scene and the balance among visual elements as the composition rules to estimate the aesthetics in images. The rule of thirds, golden mean, and golden triangles were used as the low-level visual balance rules. Features were extracted based on templates guided by these composition rules. All the methods above
measure the aesthetics of all image categories in the same way. In [147], different features were used for the aesthetics measure of different photograph categories (e.g. landscapes and portraits). In portrait photographs, the ratio of the face area to the image size was considered as one composition feature. The work in [148] focused on the aesthetics measure in photographs with face. The pose of the face and golden section composition rule were used as part of the features for the aesthetics measure. Similarly, the work in [149, 150] also evaluated the visual aesthetics in photographic portraiture. While the body ratio and face position were used as composition features in the aesthetics measure of portraits in [149], the face position, the ratio of face area to image size and the aspect ratio of image were considered as composition features in [150]. The portrait with a face placed on the top rule of thirds power points or centering on top power line or image center was considered as pleasing in [150]. The rule of thirds and size of face area were also used in [151]. From the existing aesthetics measuring methods reviewed above, generally we can conclude that the size of the face area and the location of the face are important composition features for the aesthetics measure of portrait images. However, they only measure the location of face based on some rules without the consideration of the body pose.

**Composition optimization.** Gooch et al. [152] proposed to create images based on the artistic composition rules on selection of format, viewpoint, and layout for 3D objects. The rules of thirds and fifths were applied as layout rules. They created good image composition by placing objects based on the viewing parameters computed by optimizing the composition rules. Dixon et al. [153] and Byers et al. [154] created a robot system that captures a well-composed photo based on detected faces. Four photography composition rules (the rule of thirds, the empty space, no-middle and edge rules) were used to capture a “good” picture. For a single person picture, the ideal framing was calculated by placing the face
slightly to the left or right of the vertical center line, ensuring the figure takes up two-thirds of the image height, and placing the center of the face slightly below one third of the image. For groups of people, the rule of thirds and empty space rule were used. Specifically, the centerline of the box covers all of the faces was placed slightly below one third of the image. While the robot photographer is to take the “good” picture based on the rules, most of the published research papers focus on improving the composition of captured photographs. There are two objectives in composition optimization. One is to improve the composition for aesthetics. Another is to change the aspect ratio or downsizing the image while preserving the important image features for adaptive display. This technique is called retargeting. Because the result by cropping may not contain all the important subjects in the limited aspect ratio or size, content-aware retargeting techniques such as seam-carving [155] and image warping [156] are proposed to adjust the relative position of important subjects in the image. These content-aware retargeting methods may easily create artifacts. Although content-aware methods are effective in the composition correction, cropping is still one of the most favored methods comparing to others for not introducing any artifacts [157].

Cropping is a popular method in composition optimization for aesthetics. Zhang et al. [13] presented an automatic photo-cropping algorithm by formulating auto cropping as an optimization problem. The energy function was defined based on the constraints on the face position, the center 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. For big face portrait, the face was preferred to be placed at the the cross points of the vertical golden division lines with the horizontal one-of-third lines. For small face portrait, the face was preferred to be placed at the cross points of lines of thirds. Santella et al. [158] proposed an interactive method for image cropping by using eye tracking to
identify the important content. An optimization framework that facilitates new composition rules (including the subject, avoiding cut through background objects, maximizing content area and placing content based on rule of thirds) was used to calculate the best cropping. A method combining crop and retarget was used in [22] to optimize the composition of photos. The rule of thirds, diagonal dominance, visual balance and salient-region size were used to formulate the optimization function. The methods in [22, 158] did not focus on the portrait photographs. Li et al. used the composition method in [22] to enhance the aesthetics of photos with face [14]. The above reviewed cropping methods are all rule-based. Differently, She et al. [159] and Zhang et al. [160] proposed to use learning-based method to perform cropping. In [159], the dictionary was learned for each image category based on the extracted graph-based visual saliency map, and sparse coding was used to find the best cropped region. In [160], a probabilistic model was first learned for small connected subgraphs (graphlets) of segmented regions. Then, the photo cropping was performed by searching a candidate based on the learned probabilistic model. These learning-based methods are suitable for images with multiple subregions for building a graph and they do not focus on the portrait photographs. According to the knowledge of the author on the state of the art, no one has used example-based method to achieve the cropping for composition improvement.

**Pose application.** Pose and position of the person were used in [161] to recommend professional photos with similar pose, position and composition. Similarly, we also use pose and relative position as features to select example painting. Then the selected example painting is used to guide the space cropping for improving the composition of the input photograph.
6.3 Proposed Composition Improvement Method

The pose and the placement of the face should always be considered together [162, 163]. Three example paintings with different face placements are given in Figure 6.1. In the painting Figure 6.1(a), the figure is centered at the vertical central line, but the face is shifted to the left. In the painting Figure 6.1(b), the face and figure are biased to the right of the image. While the faces in the two paintings Figure 6.1(a)(b) are all on the top of the third line, the face in Figure 6.1(c) is around the third line. The faces and centers of the figures in 120 full body paintings are drawn in Figure 6.1(d). It shows that the location of the face is not limited to the rule of thirds and the placement has a large variance.

![Figure 6.1: (a)-(c) Three paintings with different poses. The dark lines in the painting are the one-of-third lines and the white lines are the vertical and horizontal central lines. (d) The locations of faces and centers of the figures in a set of 120 full body paintings.](image)

The various placements of faces and figures in paintings indicate that the aesthetics of the portrait is not exactly rule-based. Faces are naturally placed based on the pose of figures in the paintings. This observation in paintings provides the justification for the improvement of the composition of portrait photographs based on the pose of the figure. An example painting with figure in the pose similar to that of the figure in the photograph is selected as the reference to guide the space cropping of the photograph. Here, the pose contains both the body pose and face...
pose. The face pose is comprised by the face tilt direction and the face rotation angle. The face tilt direction can be estimated together with the body pose by the pose estimation method. Therefore, in the following content, pose represents the body pose together with the face tilt direction, and the face rotation angle is extracted separately as the face direction. The framework of the proposed method is in Figure 6.2.

![Figure 6.2: The framework of the proposed method.](image)

Given an input portrait photograph, first the pose and face direction are extracted. The pose estimation method proposed in [164] is used to extract the pose of the figure and the face direction is calculated by the method proposed in [165]. Apart from the pose and face direction, spaces around the figure are also considered as features in the example painting selection. These space arrangement features are to make sure that the photograph can be cropped to have similar space arrangement as the selected example painting. Because of the pose and spaces around the figure can be described as the relative relationships of body parts and body parts with the boundaries of the image, therefore, a graph $G = (V, E)$ can be used to model the features, 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. One is body node which represents the body part, and the other is the boundary node as shown in Figure 6.3(b)(d). The features of the body node are the position and direction of the corresponding body part. The location of the center on the boundary is the feature for the boundary node. The face direction is set as one feature of the face node. The cost of the edge linking the body nodes is defined by the relative location and orientation of the corresponding body parts. The cost of the edge linking the body node with the boundary node is defined by their distance. For simplification, only the head, lower arm and lower leg nodes are linked with the boundary nodes. Two example graphs are shown in Figure 6.3.

Figure 6.3: (a) A photograph with extracted pose, (b) The graph for the photograph (a), (c) A painting with extracted pose, (d) The graph for the painting in (c). The nodes in blue are the body nodes, which express the body parts. The nodes in cyan are the boundary nodes. The red lines are the skeleton of the body. The yellow 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.

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

The objective of example painting selection is to find a painting that has a graph similar to the graph of the input photograph. Therefore, the example painting selection is posed as graph matching based on a similarity measure. Given the input graph $G_I$ and the graph of the $k$-th painting $G_k$, the graph matching is formulated as

$$\max_k S(G_I, G_k)$$

(6.1)

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

$$S(G_I, G_k) = \sum_i S^v_i(G_I, G_k) + \sum_j S^e_j(G_I, G_k)$$

(6.2)

where, $S^v_i$ is the similarity of the nodes $i$ and $S^e_j$ is the similarity of the edges $j$.

In the pose estimation by [164], the posterior marginal distribution of the position of each body part is estimated. Based on the posterior marginal distribution, the descriptors of each body node which are the position and orientation of the corresponding body part are calculated. For the body node $v^b_i$, the position descriptor is the posterior marginal distribution $P^v_i = P(l_i = (x, y, \theta))$, where $l_i$ is the body part $i$, $(x, y)$ is the scale-normalized location, and $\theta$ is the orientation. The orientation descriptor of $v^b_i$ is $P^\theta_i = \sum_{(x, y)} P(l_i = (x, y, \theta))$.

The descriptors of the edge $e^b$ linking two body nodes are the relative locations and relative orientations of the two corresponding body parts. The relative location descriptor of the edge $e^b_j$ which links the body parts $j_1$ and $j_2$ is $P^x_j = P(l^{x}_{j_1} - l^{x}_{j_2} = \delta)$. The relative orientation descriptor of the edge $e^b_j$ is $P^\theta_j = P(r(l^\theta_{j_1} - l^\theta_{j_2}) = \rho)$, where $r(\cdot)$ is a circular difference operator. Details of the calculation of the descriptors can be found in [164].

The face direction is described by 13 quantized degrees from $90^\circ$ to $-90^\circ$. The
positive degrees are for faces facing right, and negative degrees are for faces facing left [165]. The locations of the four boundary nodes $v^d$ are (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 is their distance in $x$ direction. The descriptor of the edge that links the body node with the horizontal boundary node is their distance in $y$ direction.

As in [164], the combined Bhattacharyya similarity $s(a, b) = \sum_j \sqrt{a(j) \cdot b(j)}$ is used to calculate the similarity of the position and orientation descriptors of body nodes $v^b$ and the similarity of the relative location and relative orientation descriptors of edges $e^b$. The similarity of two face directions $\theta_1$ and $\theta_2$ is calculated as

$$S_F(\theta_1, \theta_2) = e^{-\frac{(\theta_1 - \theta_2)^2}{\sigma^2}} \quad (6.3)$$

In the implementation, $\sigma = 45$.

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 make sure that there is enough space in the photograph for cropping to achieve the composition of the example painting. 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 $d_I$ and $d_k$ is

$$S^e_{q} = \begin{cases} 
1 & \text{if } d_I - d_k \geq 0 \\
\frac{1}{e^{-\frac{1}{d_I - d_k}}} & \text{if } d_I - d_k < 0 
\end{cases} \quad (6.4)$$

where $\sigma_d$ is the standard variance, and in the implementation $\sigma_d = 0.3$. The final similarity of two graphs is the linear combination of the similarities of all
descriptors for the nodes and edges, which is expressed as

\[ S(G_I, G_k) = \sum_i w_1 \cdot S^{b_i} + w_2 \cdot S_F + \sum_p w_3 \cdot S^{e_p^b} + \sum_q w_4 \cdot S^{e_q^d} \]  

(6.5)

where, \( S^{b_i} \) is the similarity of the number \( i \) body node, and \( S^{e_p^b} \) is the similarity of the \( p \)-th edge. A weight \( \omega_1 = 0.2 \) is used for the similarity of body node descriptors, a weight \( \omega_2 = 0.2 \) is used for the similarity of face descriptor, a weight \( \omega_3 = 0.4 \) is used for the similarity of body edge descriptors, and a weight \( \omega_4 = 0.2 \) is used for the similarity of edge \( e^d \). The painting that has the highest similarity score with the input photograph is selected as the example to guide the cropping of the input photograph.

### 6.3.2 Space cropping

This section presents the algorithm to improve the composition of the photograph based on the selected example painting. Because the size and aspect ratio of the figure are different in the photograph and example painting, we cannot directly crop the photograph by giving the same space around the figure as in the example painting. Hence, the space cropping is formulated as an optimization problem considering the location of each body part. Besides the marginal distribution of the position of each body part, the segmented location \((x, y)\) for each body part is also provided by the pose estimation method. The space cropping is to assign a similar location for each body part of the photograph with that of the example painting. The energy function for the space cropping is defined as

\[ E = \alpha_1 E_{pos} + (1 - \alpha_1) E_{face} \]  

(6.6)
The first term $E_{pos}$ is the pose constraint. Given the target location $v_i$ for the body part $i$ and the reference location $v_{ri}^*$ which is from the example painting, $E_{pos}$ is defined as

$$E_{pos} = \frac{1}{N} \sum_{i}^{N} \|v_i - v_{ri}^*\|$$  \hspace{1cm} (6.7)

where $N$ is the number of body parts. Due to the importance of the face, a face constraint $E_{face}$ is added in the energy function to prevent the face from shifting too much from the location in the painting. Given the target location $v_{face}$ of the face and the reference location $v_{rface}^*$ which is from the example painting, $E_{face}$ is defined as

$$E_{face} = \|v_{face} - v_{rface}^*\|$$  \hspace{1cm} (6.8)

The cropped rectangle is represented by the left top corner coordinate $(l, t)$, width $(w)$ and height $(h)$. Given a specified aspect ratio $a$, the height can be calculated as $h = a \times w$. Therefore, the optimization of the energy function can be reformulated as finding a vector of $(l, t, w)$. The particle swarm optimization (PSO) [166] method is used to seek the optimal solution by globally searching the minimum candidate of Eq. (6.6). In the implementation, the weight $\alpha_1 = 0.4$.

### 6.4 Experiments

The algorithm was implemented in MATLAB on a PC with an Intel 2.67GHz processor and 4GB RAM. For a photograph of a pixel dimension of $1000 \times 750$, it takes around 0.95 seconds to obtain the optimization result using PSO.

Figure 6.4 shows the cropping results of two photographs with figures in different poses. The pose of the figure in the selected example painting is similar to that in the input photograph and the location of the figure in the cropped result is similar to that in the example painting. This shows the effectiveness of the
proposed method. Additionally, we can see that the location of the figures in the two results are very different and they serves the requirement of their poses.

![Figure 6.4: (a) Photographs with extracted poses, (b) Selected example paintings with poses, (c) The cropping results based on the selected example paintings.](image)

For a full body portrait photograph, the proposed method can crop it to a full body portrait, or a half boy portrait with small face, arms and hands, or a half body portrait with big face. An example is shown in Figure 6.5. In the half body portrait painting with big face, generally only the upper body with upper arms are visible. The extracted poses for the lower arms using half body template are not meaningful. Therefore, in the selection of the example big face half body portrait painting, the lower arms are not considered.

To evaluate the performance of the proposed pose-related example-based com-
Figure 6.5: (a) A photograph with extracted pose, (b) Selected full body painting with pose, (c) Selected small face half body painting with extracted pose, (d) Selected big face half body painting with extracted pose. (e)-(g) The cropping results using the three example paintings.

position improvement method, it is compared with two rule-based automatic space cropping methods which are designed for photographs with face. One method is the auto cropping based on templates in [13] (ACDP method). The other method is the aesthetic-based photo editing method proposed in [14] (ABPE method). The comparison of the proposed method with the two rule-based methods are shown in Figures 6.6 and 6.7. The results in Figures 6.6(h) and 6.7(f) by the ACDP method are produced by using the small face template. For an input full body photograph, the proposed method can crop it to full body, or small face half body, or big face half body portraits as the Figure 6.6 shows. For a small face half body portrait photograph, the proposed method can also crop it to small face half body and big face half body portraits (see the Figure 6.7). However the two methods ACDP
and ABPE can only produce one kind of result. In the ACDP method, the region of interest (ROI) is detected by an attention model, which is one kind of saliency detection, to constrain the cropping. The detected ROI may lose some important body parts. In Figure 6.6(h), the feet are cropped out which leaves the figure unbalanced in the image. The ABPE method [14] only considers the face region as important information, hence in the result only the face and shoulder are visible while the other parts are all cropped out (see the results in Figures 6.6(i) and 6.7(g)). The faces are placed at the bottom one third point while leaving half of the space on the top for background. Although this placement of the face is better than its original placement, it is not reasonable considering the visual appearance of the whole image. The results by the proposed method are more encouraging comparing with the rule-based methods.

Besides the proposed composition improvement method based on the selected example painting, the method was also tested on cropping the portrait photograph based on the space arrangement from statistics. From portrait paintings, it is observed that more space is almost given on the left side of the figure in portrait paintings with face facing left while more space is given on the right side of the figure in portrait paintings with face facing right. Inspired by this, a study was conducted to explore the space arrangement in portrait paintings by dividing the portraits into 6 classes:

1. Facing left full body portrait
2. Facing right full body portrait
3. Facing left small face half body portrait
4. Facing right small face half body portrait
5. Facing left big face half body portrait
6. Facing right big face half body portrait

The space around the figure (see the Figure 6.8(a)), the face location, the center
and the size of face are measured in each of the 6 classes. Then gaussian mixture models (GMM) are used to fit the distributions of the space, center, face location and face size. The GMM models of the locations of the face and the center in the full body portrait paintings are shown in Figure 6.8(b)(c). Based on these statistical models, an energy function is formulated to crop photographs. The energy function is

\[ E = E_{TB} + E_{LR} + E_F + E_C + E_{FS} \]  

(6.9)

where, \( E_{TB} \) is the 2D GMM model of the top and bottom spaces, \( E_{LR} \) is the 2D GMM model of the left and right spaces, \( E_F \) is the GMM model of the face location, \( E_C \) is the GMM model of the center location, and \( E_{FS} \) is the one dimensional GMM model of the face size. PSO method is used to maximize the energy
Figure 6.7: (a)-(b) Example paintings with extracted pose, (c) Input photograph with extracted pose, (d) Space cropping result (small face half body) by the proposed method based on example painting in (a), (e) Space cropping result (big face half body) by the proposed method based on example painting in (b), (f) The result by method ACDP in [13], (g) The result by method ABPE in [14].

function for space cropping. This method creates baseline results (see the two examples in Figure 6.9(d)). However, by using the statistical models, the face and center are more likely to be around the vertical central line. Differently, the proposed example-based method can produce results with faces and centers in various locations. In the first example in Figure 6.9, the cropping result using statistical models is similar to the result by the proposed method. However, in the second example in Figure 6.9, the face is more biased from the central line in the result by the proposed method comparing with the result using statistical models. The face bias in this example is more pleasing for the pose requirement. In addition, the method using statistical models needs to segment the figure first for measuring the space around it. The proposed example-based method can produce results automatically without segmentation. Some more experimental results by the proposed
method are shown in Figure 6.11.

![Diagram](image)

**Figure 6.8:** (a) The space around the figure, (b) Locations of face and center of left directed full body portrait paintings, (c) Locations of face and center of right directed full body portrait paintings.

The proposed method can be easily extended to using professional photographs as the examples to guide the composition improvement for input photographs. 156 full body and 143 half body professional portrait photographs from the Corel Stock Photo database are used to form a professional portrait photograph dataset. The cropping results using the selected example professional photographs are shown in Figure 6.10.

### 6.5 Concluding Remarks

This chapter proposes a method to improve the composition of portrait photographs based on an example portrait painting. The example painting is selected based on the pose and face direction of the figure and the space around the figure. A graph model is constructed for the example painting selection. Space cropping technique is used to improve the composition of the input photograph based on the selected example painting. The space cropping is formulated as an optimization problem. The experimental results show that the proposed method performs
better than rule-based methods, and can create full body, small face half body and big half body portraits for an input full body photograph. However, the proposed method relies on the accuracy of the pose estimation. There is still a big room for improving the pose estimation in complex scenes, but it is beyond the scope of this thesis.
Figure 6.10: Using professional photographs as references. First row: Input photographs with poses, Second row: Selected example professional photographs with poses, Third row: Cropping results using the proposed method.
Figure 6.11: More experimental results. First and fourth rows: Input photographs with pose. Second and fifth rows: Selected example portrait paintings with pose. Third and sixth rows: Cropping results using the proposed method.
Chapter 7

Conclusions and Recommendations

7.1 Conclusions

This thesis studies the manipulation of contrast and composition in photographs in order to achieve the transfer of the organization of such aesthetic image attributes from paintings for aesthetic improvement on the photographs.

Depth-aware contrast manipulation is an approach taken and this leads to several algorithms. The depth-aware contrast is the contrast organization within depth planes and between depth planes. The artists exaggerate the atmospheric perspective effect in 2D images by organizing the depth-aware contrast to increase the illusion of depth, thereby making the image more interesting. To learn quantitatively how the depth-aware contrast values are organized in paintings, a statistical contrast analysis is performed on a set of landscape paintings. The contrast organization of these paintings is compared to a similar analysis of a collection of snap-shot landscape photographs. The statistical study shows that the paintings have significantly more purposeful organization in the depth-aware contrast
than photographs. The organization corresponds to an exaggerated form of the physical phenomenon of atmospheric perspective. Based on the finding, an algorithm is developed to enhance the atmospheric perspective effect in landscape photographs by manipulating the depth-aware contrast based on the contrast organization in a selected reference painting. The depth-aware contrast mapping is formulated as an optimization problem that simultaneously considers the desired inter-contrast, intra-contrast, and specified gradient constraints. Finally, a user study is conducted to evaluate the effectiveness of the proposed method. This user study demonstrates that both the visual appeal and illusion of depth guided by atmospheric perspective effect of the photographs are enhanced using the proposed method.

In portrait, the regional contrast is the main attribute that makes the figure, especially the face be the focus of attention while the depth planes are simplified to FG and BG. This thesis proposes a method to manipulate the regional contrast in snap-shot style portrait photographs by using pre-modern portrait paintings as aesthetic examples. The example portrait painting is selected using a learned distance metric. The contrast organization 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 user study demonstrates that the visual appeal of the portrait photographs are effectively improved and the face and the figure become more salient using the proposed regional contrast manipulation method.

Vignetting effect is also used by artists to create the focus of attention in the paintings by organizing the center-corner contrast and it can be applied jointly with depth-aware contrast manipulation mentioned in the previous two paragraphs. The thesis analyzes the vignetting effect in paintings and photographs. The observa-
tion of the difference between paintings and photographs shows that the vignetting
effect in paintings is more naturally presented based on the geometry of the scene.
Then the thesis explores an algorithm to apply the lightness weighting extracted
from an example painting to a photograph for creating the painter-style vignetting
effect while satisfying four manipulation criteria observed from the paintings. Ex-
periments show that the proposed method can successfully transfer the vignetting
effect from the example painting to the photograph and the effect is more naturally
presented comparing with popular software and camera models.

Composition takes an important role in image aesthetics. The thesis studies
the composition in portrait paintings and has observed that the placements of the
face and the figure in portrait paintings are pose-related. Motivated by this, the
thesis proposes to improve the composition of a portrait photograph by learning
the placements of the face and the figure from an example portrait painting. The
example painting is selected using a graph model by graph matching. Finally
space cropping is conducted according to an optimization 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.

7.2 Recommendations for Future Research

In the depth-aware contrast and regional contrast manipulations, the local contrast
is also adjusted by manipulating the intra-contrast. In the center-corner contrast
manipulation, the local contrast restoration is also adjusting the local contrast.
However, these local contrast adjustments are only to manipulate the relative
difference of neighboring pixels based on their original relationship, not to change
the relationship of local regions. In paintings, the contrast of local regions is
also organized by painters to describe the local details. A method to transfer the
local contrast organization in paintings to photographs can be explored. First,
the image can be segmented using the existing segmentation methods, such as
graph-cut, mean-shift, and GMM to small regions. Then, Markov random field
(MRF) model can be used to model the segmented regions as nodes to control the
relationship of neighboring regions. Due to lacking of the prior knowledge of the
contrast relationship between neighboring regions and relationship between nodes
and observations, currently there is no solution to this local contrast mapping.
However, the current research work has clearly showed the importance to learn the
local relationship of regions.

The contrast mapping in both landscape and portrait photographs do not touch
the hue contrast. Although hue has a smaller influence on the depth perception
comparing lightness and saturation, hue plays a most important role in controlling
the mood of the scene (especially in the sky area). So the hue contrast mapping
for controlling the emotion and mood is another potential research direction. In
addition, a hue, lightness and saturation dependent contrast mapping method
needs to be developed instead of operating the three attributes separately.

In the review of contrast definition models, it has been found that there were
no formal definitions for these contrasts (global contrast, local contrast, lightness
contrast, saturation contrast and also hue contrast). All the current contrast mod-
els would probably be better regarded as application tools than as descriptions of
basic visual function. It would be an interesting research project to aim at formal
specifications of all the mentioned contrasts.

For the composition, the thesis only focuses on the analysis of composition in
portrait photographs and paintings. The composition in landscape images is also
an important aspect that influences the aesthetics. While the placements of the
figure and the face are important elements in portrait composition, the placement
of foreground elements and the organization of depth planes are much important in landscape images. Therefore, improving the composition in landscape photographs and study the criteria on the generation of a more aesthetic related depth map can be one potential research direction.

One of the most significant challenges in computational aesthetics is the evaluation of the claimed validity of proposed methods. Some methods for automatic image quality assessment are proposed in recent years, e.g. ACQINE [100], Content-based photo quality assessment [147], etc. They use the organization of global and regional features to assess the quality of images. These methods can measure the quality of images in a general view. However, they are not suitable for an issue related evaluation. Currently, user study is a common way to access the validity of the proposed method for a special task. Similar with other computational aesthetic research, assessment of the validity of our proposed methods is also one challenge problem. In the future, work needs to be done to develop a quantitative measurement for assessing the results for each task.
Author’s Publications


6. Xiaoyan Zhang, Martin Constable, and Ying He, “On the Transfer of Painting Style to Photographic Images through Attention to Colour Contrast”, in Pro-

Bibliography


quality assessment and enhancement based on visual aesthetics,” in Proc. of

[145] P. Obrador, L. Schmidt-Hackenberg, and N. Oliver, “The role of image com-
position in image aesthetics,” in Proc. of International Conference on Image

[146] P. Obrador et al., “Towards category-based aesthetic models of pho-
tographs,” in Proc. of the 18th International Conference on Advances in

in Proc. of International Conference on Computer Vision, 2011, pp. 2206–
2213.


[149] L. Huang et al., “Personalized portraits ranking,” in Proc. of the 19th ACM

[150] S. S. Khan and D. Vogel, “Evaluating visual aesthetics in photographic por-
traiture,” in Proc. of the Eighth Annual Symposium on Computational Aes-

[151] S. Battiato et al., “Aesthetic scoring of digital portraits for consumer ap-

[152] B. Gooc et al., “Artistic composition for image creation,” in Proc. of the


