Uncovering vein patterns from color skin images for personal identification in forensic investigation

Tang, Chaoying

2012

Tang, C. (2012). Uncovering vein patterns from color skin images for personal identification in forensic investigation. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/51059>

<https://doi.org/10.32657/10356/51059>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**UNCOVERING VEIN PATTERNS FROM COLOR
SKIN IMAGES FOR PERSONAL IDENTIFICATION
IN FORENSIC INVESTIGATION**

TANG CHAOYING

SCHOOL OF COMPUTER ENGINEERING

2012

**UNCOVERING VEIN PATTERNS FROM COLOR
SKIN IMAGES FOR PERSONAL IDENTIFICATION
IN FORENSIC INVESTIGATION**

TANG CHAOYING

School of Computer Engineering

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

2012

Abstract

Recent technological advances have allowed for a proliferation of digital evidence images. Using these images as evidence in legal cases (e.g. child sexual abuse, child pornography and masked gunmen) can be very challenging, because the faces of criminals or victims are not visible. Although large skin marks and tattoos have been used, they are ineffective in some legal cases, because the skin exposed in evidence images have neither unique tattoos nor enough skin marks for identification. The blood vessel between the skin and the muscle covering most parts of the human body is a powerful biometric trait, because of its universality, permanence and distinctiveness. Traditionally, it was impossible to use vein patterns for forensic identification, because they were not visible in color images. All the current vein recognition systems developed by companies and research laboratories rely on near infrared (NIR) imaging devices to capture high quality vein patterns from hand and wrist, where skin is relatively thin, for commercial applications. Up until now, no one studies vein patterns for forensic identification because no method has been developed to visualize vein patterns hidden in color images. The primary aim of this research is to develop algorithms for visualizing vein patterns hidden in color images so that criminal and victim identification can be performed based on vein patterns. The secondary aim of this research is to develop algorithms for removing JPEG blocking artifacts in skin images that adversely affect forensic recognition.

We propose two approaches for uncovering vein patterns from color skin images. The first approach is based on RGB-NIR mapping. It extracts information from a pair of synchronized color and NIR images and uses a neural network (NN) to map RGB values to NIR intensities. Furthermore, we design an automatic intensity adjustment scheme for illumination compensation and an NN weight adjustment scheme for improving the robustness of the approach. Using an automatic matching algorithm, we match resultant images from the RGB-NIR mapping approach and find that its matching result is comparable to the result from matching NIR images, which are always considered as ground truth of vein patterns.

In the second approach, we use principles of optics and skin biophysics for uncovering vein patterns. It inverts the process of skin color formation in an image and derives the corresponding biophysical parameters, where veins can be observed. Based on this approach we develop four optical models for simulating skin color formation. They are all based on the radiative transfer equation which quantitatively describes transport of light in the human skin. The first and second optical models use the Kubelka-Munk (K-M) model to approximate the solution of the radiative transfer equation, whose exact analytical solution has not yet been obtained for complex and multiple scattering media such as human skin. In these two models, we assume that the optical properties of human skin is determined by three layers – the stratum corneum, the epidermis, and the dermis, and veins are located in the dermis. To overcome the limits of the first model, the second optical model uses a color optimization scheme and the automatic intensity adjustment scheme. The third and fourth optical models use Reichman’s solution to the radiative transfer equation. In the fourth optical model, we add the fourth layer, the hypodermis consisting of adipose and blood vessels to the skin structure. Because none of the models can provide exact and complete vein patterns, we propose a method to fuse the vein patterns obtained from different models. Experimental evaluations show that the fusion results are much better than any of the single models and also the RGB-NIR mapping approach. Its matching result is even better than matching NIR images.

Furthermore, we develop two specific approaches to remove blocking artifacts in JPEG-compressed skin images. The first one is a maximum-a-posteriori (MAP)-based approach which formulates skin image deblocking as an estimation problem, and embeds statistical information of skin images into a MAP model to perform the estimation. The second one is a knowledge-based approach which extracts prior knowledge of skin images from a training set, and uses it to infer original blocks in compressed evidence images. Two inference schemes, a block synthesis algorithm and an indexing mechanism are also proposed for this approach. Both approaches guarantee that the resultant and compressed images have the same quantized DCT coefficients. Experimental results demonstrate that the approaches perform better than other methods.

In this research, we break the limit of traditional vein recognition and show its potential for forensic analysis. According to our best knowledge, no one did similar research before.

Acknowledgement

I would like to thank all those people who contributed in many ways to my Ph.D. study and made it an unforgettable experience for me.

First of all, I give my most sincere gratitude to my supervisor, Professor Adams Wai-Kin Kong. This work would not have been possible without his support, guidance and encouragement. I am very thankful to him for giving me the opportunity to start my Ph.D. study at the Nanyang Technological University. He encouraged me to delve into a new research field that no one has tried before. During the past four years, he has put countless time and effort on my research, and has continuously given me insightful advices and constructive suggestions. He showed me the passion, diligence, integrity and perseverance that a true researcher should have. He gave me vision, enthusiasm and confidence. His unflinching dedication to research will always inspire me. I am extremely indebted to him for his understanding of my difficulties in the last period of Ph.D. study. Heartfelt thanks also go to his wife, Kannis, for her great help during my most difficult time in Singapore.

I gratefully acknowledge the technician of the Forensic and Security Laboratory, Christina, for her assistance and helpfulness. Her skill and professionalism in maintaining the laboratory greatly sped up the experiments in the research. I warmly thank Professor Chai Quek for his attention and care during the four years of my Ph.D. study. I would also like to thank Professor Kin Choong Yow, Professor David Cho, and Professor Maylor Leung for their advice and encouragement. I am also thankful to my teachers in Nanjing University of Aeronautics and Astronautics. My thanks are due to Professor Jiang Ju, Professor Gong Huajun, and Professor Shen Chunlin for their recommendation and encouragement.

I also thank all my friends and colleagues for their help and support. Many results described in this thesis would not have been obtained without a close collaboration with my colleague, Hengyi. He is more like my younger brother, and has given me countless assistance and happiness.

I owe a great deal of appreciation and gratitude to other members in our research team, Arfika, Xiaojie, Rubha, Quang, Xingpeng, Yanzhu, Frodo, Abhik, Han and Hamid. I am so lucky to have been able to work with you all. My thanks go in particular to Hamid for his valuable help and suggestions in the procedure of submitting thesis. I am especially indebted to all my Chinese friends for providing such a caring and fun-filled environment.

I want to thank present and past members of the Forensic and Security Laboratory: Raj, Dilip, Pravin, Atiqur, Weihua, Alex, Than and Deepak. Our talks and discussions on research and other random topics have been an enjoyable and unforgettable experience for me.

Words fail me when I try to express my appreciation to my family. I extend warm thanks to my parents and parents-in-law. Their unending love and unwavering support had made my Ph.D. journey possible. They are always beside me during the happy and hard moments. Special thanks go to my cousin and his wife for giving me the feeling of family in Singapore. I reserve my deepest appreciation to my love, Biao. He knows me better than I know myself. I truly thank him for always sticking by my side during the inevitable ups and downs in the past years. The journey becomes easier because we travel together.

Contents

Abstract	i
Acknowledgement	iii
List of Figures	viii
List of Tables	xiv
List of Publications and Patent	xv
Chapter 1 Introduction	1
1.1 A Challenging Personal Identification Problem	1
1.2 Objectives	3
1.3 Vein Recognition and IR Imaging	3
1.4 Challenges and Solutions	5
1.5 Organization of the Thesis	6
Chapter 2 Background	8
2.1 Human Skin Anatomy	8
2.1.1 The Epidermis	9
2.1.2 The Dermis	10
2.1.3 The Hypodermis	11
2.1.4 Other Parts of the Human Skin	11
2.2 Approaches for JPEG-Compressed Image Deblocking	12
2.2.1 The Preprocessing Methods	12
2.2.2 The Postprocessing Methods	14
2.3 Approaches for Vein Uncovering	20
2.3.1 Image Enhancement Methods	20
2.3.2 Independent Component Analysis-Based Method	23
2.4 The K-M Model and Its Application in Skin Optics	26

2.4.1 The K-M Model	26
2.4.2 The Skin Optics Based on the K-M Model.....	28
2.5 Vein Pattern Extraction and Automatic Matching Algorithms	29
2.5.1 A Vein Pattern Extraction Algorithm	29
2.5.2 A Three-Step Vein Matching Algorithm	33
2.6 The Database.....	36
Chapter 3 Skin Image Deblocking.....	39
3.1 An MAP-Based Approach	41
3.1.1 An MAP Formulation	41
3.1.2 A Potential Function Connecting Skin Statistics and the MAP Model.....	43
3.1.3 Optimizing the Potential Function	46
3.1.4 Experimental Evaluations	47
3.2 A Knowledge-Based Approach	51
3.2.1 Prior Knowledge in Skin Images	52
3.2.2 The Two Inference Algorithms Based on Prior Knowledge.....	56
3.2.3 Experimental Evaluations	66
3.3 Summary	78
Chapter 4 An RGB-NIR Mapping Approach to Uncover Vein Patterns	80
4.1 An RGB-NIR Mapping Network.....	81
4.2 Automatic Adjustment of Image Intensity	82
4.3 Automatic Adjustment of Neural Network Weight	83
4.4 Experimental Evaluations	88
4.4.1 Evaluation by Visual Observation	88
4.4.2 Evaluation of Deblocking and Vein Uncovering	93
4.4.3 Evaluation by Automatic Vein Matching	97
4.5 Summary	101
Chapter 5 K-M Model-Based Approach to Uncover Vein Patterns.....	102

5.1 The First Optical Model	103
5.1.1 Overview of the Approach	104
5.1.2 The Skin Color Formation	106
5.1.3 An Inversion of the Skin Color Formation	114
5.1.4 Experimental Evaluations	115
5.2 The Second Optical Model	126
5.2.1 Parameter Range Optimization through Color	127
5.2.2 Experimental Evaluations	129
5.3 Summary	140
Chapter 6 Reichman Model-Based Approach to Uncover Vein Pattern.....	141
6.1 The Third Optical Model	142
6.1.1 Model Description.....	142
6.1.2 Experimental Evaluations	144
6.2 The Fourth Optical Model.....	154
6.2.1 Model Description.....	154
6.2.2 Experimental Evaluations	157
6.3 Summary	167
Chapter 7 Fusion of Models to Uncover Vein Pattern.....	168
7.1 Description of the Fusion Method	168
7.2 Experimental Evaluations	173
7.3 Summary	176
Chapter 8 Conclusion.....	177
8.1 Summary of the Main Contributions	177
8.2 Future Work	179
References.....	182

List of Figures

Fig. 1.1 Federal prosecutions of child sex exploitation offenders [3].	2
Fig. 1.2 Percentage of suspects prosecuted in different crimes [3].	2
Fig. 1.3 The difference between genetically identical vein patterns.	5
Fig. 1.4 A diagram of the proposed scheme.	6
Fig. 2.1 Illustration of human skin structure. (www.medicinenet.com)	9
Fig. 2.2 The enhancement results by contrast stretching (example 1).	22
Fig. 2.3 The enhancement results by contrast stretching (example 2).	22
Fig. 2.4 The enhancement results by Laplacian transform.	22
Fig. 2.5 The enhancement results by Photoshop.	22
Fig. 2.6 Light reflection model for dielectric materials.	24
Fig. 2.7 Successful application of ICA-based method.	24
Fig. 2.8 Failed results from ICA-based method (example 1).	25
Fig. 2.9 Failed results from ICA-based method (example 2).	25
Fig. 2.10 An example image that ICA-based method cannot converge.	25
Fig. 2.11 Illustration of light transport based on the K-M model.	26
Fig. 2.12 Illustration of the process of light transport in a two-layered material.	28
Fig. 2.13 A preprocessing result.	30
Fig. 2.14 The Gabor filter outputs.	31
Fig. 2.15 The process of vein pattern extraction.	33
Fig. 2.16 The three-step matching result.	35
Fig. 2.17 Equipment setup for data collection.	37
Fig. 2.18 The image capture process.	37
Fig. 2.19 Interface of the image capture software package.	38
Fig. 3.1 The effects of blocking artifacts on pigmented skin marks and veins.	40

Fig. 3.2 The smoothing effect of traditional deblocking methods on skin images.	40
Fig. 3.3 Illustration of cliques.	43
Fig. 3.4 Illustration of cdfs from different cases.	45
Fig. 3.5 Illustration of the change of d_{mc_k} after compression.	46
Fig. 3.6 Evaluation of the MAP-based approach (example 1).	49
Fig. 3.7 Evaluation of the MAP-based approach (example 2).	50
Fig. 3.8 Some images in the training set.	54
Fig. 3.9 A raw image and a pair of skin images obtained from it.	54
Fig. 3.10 QDCT coefficient matrix, effective coefficients, and index vector.	55
Fig. 3.11 A compressed block and its corresponding original blocks.	55
Fig. 3.12 Spatial and frequency neighborhoods.	56
Fig. 3.13 A Markov network model.	58
Fig. 3.14 The performance of the Markov-model-based algorithm.	59
Fig. 3.15 Illustration of hybrid neighborhood in the testing process.	61
Fig. 3.16 Illustration of hybrid neighborhood in training set construction.	61
Fig. 3.17 The performance of the one-pass algorithm.	61
Fig. 3.18 The corresponding blocks from Fig. 3.17.	62
Fig. 3.19 Illustration of the block synthesis algorithm.	64
Fig. 3.20 The block synthesis results.	64
Fig. 3.21 Evaluation of the knowledge-based approach in the Y component.	67
Fig. 3.22 Evaluation of the knowledge-based approach in the U component.	68
Fig. 3.23 Evaluation of the knowledge-based approach in the V component.	68
Fig. 3.24 Evaluation of the knowledge-based approach in color images (example 1).	69
Fig. 3.25 Evaluation of the knowledge-based approach in color images (example 2).	69
Fig. 3.26 Evaluation of the knowledge-based approach in color images (example 3).	70
Fig. 3.27 Subjective evaluation results from the three dermatological professionals.	72
Fig. 3.28 Subjective evaluation results from the other participants.	73

Fig. 3.29 The average of similarity measures.	75
Fig. 3.30 The standard deviation of similarity measures.	76
Fig. 4.1 A pair of synchronized images of a male's left forearm obtained by the 2-CCD multi-spectral camera.....	81
Fig. 4.2 (a) A schematic diagram of the mapping network to uncover vein patterns and (b) a schematic diagram of the generic NN mapping model.....	82
Fig. 4.3 Illustration of extracted vein patterns (example 1).	83
Fig. 4.4 Illustration of extracted vein patterns (example 2).	84
Fig. 4.5 Illustration of vein segments.....	86
Fig. 4.6 The NN weight adjustment result (example 1).	86
Fig. 4.7 The NN weight adjustment result (example 2).	87
Fig. 4.8 The NN weight adjustment result (example 3).	87
Fig. 4.9 Visual evaluation of the RGB-NIR mapping approach (example 1).	89
Fig. 4.10 Visual evaluation of the RGB-NIR mapping approach (example 2).	89
Fig. 4.11 Visual evaluation of the RGB-NIR mapping approach (example 3).	90
Fig. 4.12 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 1).	90
Fig. 4.13 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 2).	91
Fig. 4.14 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 3).	91
Fig. 4.15 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 4).	92
Fig. 4.16 Vein uncovering results from a deblocked image (example 1).	94
Fig. 4.17 Vein uncovering results from a deblocked image (example 2).	95
Fig. 4.18 Vein uncovering results from a deblocked image (example 3).	96
Fig. 4.19 Some images in our database which have tattoos or too much hair.	97
Fig. 4.20 Cumulative match characteristic curves of vein patterns uncovered by the RGB-NIR mapping approach on the smaller testing set.	99

Fig. 4.21 Cumulative match characteristic curves of vein patterns uncovered by the RGB-NIR mapping approach on the larger testing set.....	99
Fig. 5.1 A schematic diagram of the optical approach to uncover vein patterns.	106
Fig. 5.2 Illustration of the process of light transport in human skin.	107
Fig. 5.3 The spectral absorption coefficient curve of the baseline skin.	108
Fig. 5.4 The spectral extinction coefficient curve of β -carotene.....	108
Fig. 5.5 The spectral scattering coefficient curve of the stratum corneum.	109
Fig. 5.6 The spectral extinction coefficient curves of eumelanin and pheomelanin.....	110
Fig. 5.7 The spectral extinction coefficient curves of hemoglobin and bilirubin.	112
Fig. 5.8 The spectral power distributions of illuminants D65, A, and F1-F6.	113
Fig. 5.9 The spectral response functions of Nikon D70s and Cannon 500D.	114
Fig. 5.10 Visual evaluation of the first optical model (example 1).....	118
Fig. 5.11 Visual evaluation of the first optical model (example 2).....	119
Fig. 5.12 Visual evaluation of the first optical model (example 3).....	120
Fig. 5.13 Visual evaluation of the first optical model (example 4).....	121
Fig. 5.14 Visual evaluation of the first optical model (example 5).....	122
Fig. 5.15 An example of manual matching in the first session.	124
Fig. 5.16 An example of manual matching in the second session.	125
Fig. 5.17 An example of the vein uncovering results using the first optical model.....	127
Fig. 5.18 Illustration of parameter range optimization.	128
Fig. 5.19 Data preparation for training a three-layered feed-forward neural network.....	128
Fig. 5.20 Visual evaluation of the second optical model (example 1).....	130
Fig. 5.21 Visual evaluation of the second optical model (example 2).....	131
Fig. 5.22 Visual evaluation of the second optical model (example 3).....	132
Fig. 5.23 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 1).	133
Fig. 5.24 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 2).	134

Fig. 5.25 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 3).	135
Fig. 5.26 Comparison between the extracted vein patterns from the first and second optical models (example 1).....	136
Fig. 5.27 Comparison between the extracted vein patterns from the first and second optical models (example 2).....	137
Fig. 5.28 Comparison between the extracted vein patterns from the first and second optical models (example 3).....	137
Fig. 5.29 Cumulative match characteristic curves of vein patterns uncovered by the first and second optical models.	139
Fig. 6.1 Visual evaluation of the third optical model (example 1).	145
Fig. 6.2 Visual evaluation of the third optical model (example 2).	146
Fig. 6.3 Visual evaluation of the third optical model (example 3).	147
Fig. 6.4 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 1).	148
Fig. 6.5 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 2).	149
Fig. 6.6 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 3).	150
Fig. 6.7 Comparison between the extracted vein patterns from the first, second, and third optical models (example 1).....	152
Fig. 6.8 Comparison between the extracted vein patterns from the first, second, and third optical models (example 2).....	152
Fig. 6.9 Comparison between the extracted vein patterns from the first, second, and third optical models (example 3).....	153
Fig. 6.10 Cumulative match characteristic curves of vein patterns uncovered by the first, second, and third optical models.....	153
Fig. 6.11 A schematic diagram of the fourth optical model to uncover vein patterns.	156
Fig. 6.12 Visual evaluation of the fourth optical model (example 1).	158
Fig. 6.13 Visual evaluation of the fourth optical model (example 2).	159

Fig. 6.14 Visual evaluation of the fourth optical model for images taken from different viewing angles and poses (example 1).	160
Fig. 6.15 Visual evaluation of the fourth optical model for images taken from different viewing angles and poses (example 2).	161
Fig. 6.16 Comparison between the extracted vein patterns from the four optical models (example 1).	163
Fig. 6.17 Comparison between the extracted vein patterns from the four optical models (example 2).	164
Fig. 6.18 Comparison between the extracted vein patterns from the four optical models (example 3).	165
Fig. 6.19 Cumulative match characteristic curves of vein patterns uncovered by the four optical models.	166
Fig. 7.1 The extracted vein patterns and the fusion result (example 1).	170
Fig. 7.2 The extracted vein patterns and the fusion result (example 2).	171
Fig. 7.3 The extracted vein patterns and the fusion result (example 3).	172
Fig. 7.4 Cumulative match characteristic curves of vein patterns obtained by the multi-optical-model fusion method.....	174

List of Tables

Table 1.1 Sex exploitation suspects declined for prosecution by U.S. attorneys, by reason.	2
Table 3.1 Automatic skin mark detection rate from different algorithms.....	77
Table 3.2 Means of blockiness index from different algorithms.	78
Table 4.1 The rank-one and rank-10% identification accuracies from the image mapping approach on the smaller testing set.	100
Table 4.2 The rank-one and rank-10% identification accuracies from the image mapping approach on the larger testing set.....	100
Table 5.1 Constant skin parameters.	116
Table 5.2 Manual matching results (%).	126
Table 5.3 The rank-one and rank-10% identification accuracies from the first and second optical models.	139
Table 6.1 The rank-one and rank-10% identification accuracies from the first, second, and third optical models.	154
Table 6.2 The rank-one and rank-10% identification accuracies from the four optical models. .	166
Table 6.3 The rank-one and rank-10% identification accuracies from different uncovering models.	175

List of Publications and Patent

Publications

1. Chaoying Tang, Adams Kong, Noah Craft, “Using a Knowledge-Based Approach to Remove Blocking Artifacts in Skin Images for Forensic Analysis”, *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1038-1049, 2011 (12 pages).
2. Chaoying Tang, Adams Kong, Noah Craft, “Uncovering Vein Patterns from Color Skin Images for Forensic Analysis”, *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 665-672, 2011 (accepted as oral presentation, acceptance rate of 3.5% for oral, 8 pages).
3. Chaoying Tang, Adams Kong, Noah Craft, “A Knowledge-Based Algorithm to Remove Blocking Artifacts in Skin Images for Forensic Analysis”, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1928-1931, 2011.
4. Chaoying Tang, Adams Kong, Noah Craft, “Removing JPEG Artifacts in Skin Images for Forensic Analysis”, *American Academy of Forensic Sciences, 63rd Annual Scientific Meeting*, Chicago, pp. 142-143, 2011 (Extended abstract, Oral presentation).
5. Chaoying Tang, Hengyi Zhang, Adams Kong, Noah Craft, “Visualizing Vein Patterns from Color Skin Images based on Image Mapping for Forensics Analysis”, *International Conference on Pattern Recognition*, 2012 (accepted) .
6. Hengyi Zhang, Chaoying Tang, Adams Kong, Noah Craft, “Matching Vein Patterns from Color Images for Forensic Investigation”, *IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2012 (oral presentation, 8 pages, best reviewed paper and Honeywell best student paper award).

Patent

1. Adams Wai Kin Kong, Noah Ames Craft, Chaoying Tang and Hengyi Zhang, “Method and Apparatus for Extracting Blood Vessel Patterns in Color Images for Personal Identification”, US provisional patent US 61/479,795.

Chapter 1

Introduction

1.1 A Challenging Personal Identification Problem

Recent technological advances have allowed for a proliferation of digital evidence (e.g. images of child pornography and violent protest). This media can be used as investigation hints, but commonly the faces of criminals in the images cannot be seen since they are covered or obstructed. Therefore the traditional identification technologies such as face recognition systems cannot handle this digital evidence. These cases might include images from child pornography and masked terrorists. United States Bureau of Justice Statistics reported that child sex offenses are the fastest growing offenses of the Federal criminal caseload [1]. The average rate of increase is 15%. During 2006, U.S. attorneys handled 3,661 suspects related to child sex offenses. Child pornography constituted 69%, followed by sexual abuse (16%) and sex transportation (14%). Fig. 1.1 shows the number of federal prosecutions of child sex exploitation offenders and illustrates two important points: that child pornography is the driving force for the increase of child sex offenses, and that the number of prosecuted suspects is much less than the number of suspects. More precisely, 2,039 suspects were prosecuted for federal child sex offenses in 2006, which is only 56% of total suspects handled by U.S. attorneys. Fig. 1.2 shows that among major categories of crimes, sexual exploitation also has the lowest percentage of suspects prosecuted among all offenses [1]. A declination for prosecution by U.S. attorneys does not mean that the suspect did not commit a federal offense. The major reason for these declinations is weak/inadmissible evidence shown in Table 1.1. Due to a lack of effective identification technology, a huge amount of child pornography has been posted on the Internet. In Canada alone, 30,000 cases of child pornography have been reported [2]. The U.S. Customs Service estimated that 100,000 websites offer child pornography [3]. Child pornography and other child sexual offenses are prohibited all

around the world. Anyone who possesses, makes, prints, publishes, distributes, sells or imports child pornography commits a criminal offense. Accurate biometric identification in cases of child sexual exploitation, especially child pornography, would greatly enhance the ability of prosecutors to achieve convictions. In addition to child pornography, personal identification based on evidence images or videos may be applicable to many other crimes such as masked criminals. It is not difficult to note that many criminals in recent violent protests around the world, including occupy Rome and 2011 English riots, wore masks to cover their faces for fear of possible identification and prosecution, but a good portion of their skin was observable.

Fig. 1.1 Federal prosecutions of child sex exploitation offenders [3].

Fig. 1.2 Percentage of suspects prosecuted in different crimes [3].

Table 1.1 Sex exploitation suspects declined for prosecution by U.S. attorneys, by reason. [3]

Basic for declination	Total	Sex abuse	Child pornography	Sex transportation
Number of declinations	1,419	330	898	191
Lack evidence of criminal intent	19.7%	14.2%	22.7%	15.2%
Prosecuted on other charges	15.0%	3.6%	18.7%	17.3%
No federal offense evident	4.9%	6.4%	4.1%	6.3%
Weak/inadmissible evidence	30.4%	53.0%	24.3%	20.4%
Agency request	8.3%	2.7%	10.5%	7.3%
Other*	21.6%	20.0%	19.7%	33.5%

* Other includes lack of investigative resources, jurisdiction or venue problems, witness problems, and unspecified reasons.

1.2 Objectives

There are several approaches to identify criminals based on partial skins in images collected from crime scenes, such as tattoos and skin marks. In a legal case, Dr. Adams Kong, PhD, an assistant professor at the Nanyang Technological University and Dr. Noah Craft, MD, PhD, an associate professor at the University of California at Los Angeles were recruited by the US Department of Justice as expert witnesses for verifying a Caucasian criminal in images of child pornography. In this case, they used skin marks to identify the criminal [4]. In addition to skin marks, tattoos are useful in some cases [5]. However, these two features are ineffective in some of these cases because the skin exposed in evidence images may have neither unique tattoos nor enough skin marks for personal identification [6]. Some people, especially gang members, can have the same tattoo pattern, while many do not have tattoos. Skin marks suffer from a similar problem. Therefore, the traditional identification systems such as face recognition systems and tattoo retrieval systems cannot handle these evidence images.

Vein pattern is a hard biometric trait. If vein patterns can be uncovered from skin images collected from crime scenes, forensic investigators can use this biometrics for criminal identification. The objective of this research is to uncover vein patterns from color skin images for forensic investigation. Traditionally, vein patterns were obtained by infrared (IR) imaging. In the next section, we will give an introduction to the related technologies.

1.3 Vein Recognition and IR Imaging

Vein recognition is a biometric technology that performs personal verification and identification based on the vast network of blood vessels under skin surface [7]. It is a unique and stable biometric trait with strong immunity to forgery. It was found that identical twins with the same DNA sequence have different vein patterns (see Fig. 1.3). The vast network of blood vessels is believed to be “hardwired” into the body at birth, and remains relatively unaffected by aging, except for predictable growth, as with fingerprints [8]. In addition, as the blood vessels are hidden

underneath the skin and are almost invisible, they are much harder to be duplicated compared with other biometric traits.

Vein biometrics has received considerable attention in the last decade. All the current systems developed in laboratories and companies use non-invasive IR or laser imaging technologies to obtain high quality vein patterns in hand (e.g. finger and palm) and wrist for commercial applications. Some hand vein systems have been applied to automated teller machine and physical access control [9]. However, there is no study about uncovering body vein pattern (BVP) hidden in skin images captured by consumer cameras for personal identification.

There are two types of IR imaging technologies: far infrared (FIR) whose wavelength spans between 6 and 14 μm , and near infrared (NIR) whose wavelength spans between 0.76 and 2 μm . FIR imaging systems detect the electromagnetic radiation emitted from the human body in the FIR spectrum [8]. It is based on the fact that veins emit more FIR radiation, because blood has higher temperature than surrounding tissue. NIR imaging technology is based on the fact that the hemoglobin in blood absorbs incident NIR radiation with wavelength in the range of 700-900 nm. If NIR light is shone on the skin and imaged by a NIR camera, then the veins about 3 mm under the skin surface can be captured. FIR is very sensitive to ambient temperature and humidity conditions, and also the temperature variation of the human body. Because of heat radiation, blood vessels and their surrounding tissue have similar temperature. Consequently, FIR imaging systems can only capture major vein patterns with relatively low contrast. NIR imaging systems produce images with better quality and are more tolerant to external environments and subject's condition. They are widely used in the current vein recognition systems.

Previously, vein pattern was solely used for commercial applications, where NIR imaging was employed to capture high quality NIR images for recognition. In this thesis, we extend applications of vein patterns to forensic analysis, where no IR images can be collected. We will present a set of algorithms and approaches to address related problems and finally uncover vein patterns hidden in color images for personal identification.

Fig. 1.3 The difference between genetically identical vein patterns.

(a) and (b) are respectively the left and right forearms of the same subject. These vein patterns are generated by the same DNA sequence.

1.4 Challenges and Solutions

This thesis is the first one in the world about uncovering body vein patterns from color skin images taken by consumer cameras for forensic investigation. Because veins lie underneath the skin surface, they are almost invisible in the color images to the naked eye. Therefore, it is a completely new and very challenging research direction. As we mentioned before, all the current vein recognition techniques are dependent on either hand or wrist veins captured by IR imaging systems. This research targets uncovering vein patterns from entire human body and extends the applications of vein patterns to forensic analysis, where no IR images can be collected.

Fig. 1.4 shows our proposed solution scheme. In this research, skin images collected from crime scenes are considered as inputs. Firstly, the quality of images will be checked. At this moment, we do not seek an automatic approach to separate high quality images from low quality ones. Then, if the images are highly compressed by JPEG algorithm, we will apply our algorithms to remove the blocking artifacts. Next, we will use our scheme and some existing technique to compensate lighting variation. These two steps are considered as preprocessing. Once the quality is acceptable, we use our algorithms to uncover hidden vein patterns. These vein patterns will be extracted and then matched with the vein patterns extracted from NIR images and/or color images collected from potential criminals (e.g. previous inmates and suspects), whose BVP records are stored in the criminal database. The color images in the criminal database are usually

with good quality, so no preprocessing is needed and vein uncovering is performed directly on the images.

Fig. 1.4 A diagram of the proposed scheme.

In this thesis, we mainly concentrate on developing algorithms for removing JPEG blocking artifacts especially for skin images, and uncovering vein patterns from color images. We propose two approaches for removing JPEG blocking artifacts from skin images: one is a MAP-based approach and the other is a knowledge-based approach. We develop two approaches for uncovering vein patterns from color skin images. The first one is based on RGB-NIR mapping. The second one is based on the principles of optics and skin biophysics. Using the second approach we develop four optical models and a multi-model fusion method. For algorithm design, development and evaluation, we construct a large skin image database constituted by corresponding RGB and NIR images. The images are collected on two separate occasions, at an interval of around two weeks.

1.5 Organization of the Thesis

The rest of this thesis is organized as follows. Chapter 2 gives a review of the related work, and reports the construction process and characteristics of our current skin image database. Chapter 3 presents two skin-specific JPEG deblocking approaches. Chapter 4 presents the vein uncovering

approach based on image mapping. Chapter 5 describes the vein uncovering approach based on skin optics, and the first and second optical models. Chapter 6 discusses the third and fourth optical models. Chapter 7 presents a multi-optical-model fusion method to obtain better uncovering performance. Experimental evaluations with analyses and discussions are provided in chapters 3 to 7. In the last chapter we conclude the thesis and offer the future work.

Chapter 2

Background

In this chapter, we will give an overview of the background and the previous work related to this thesis. We will also introduce our image capture system and current database. In section 2.1, we will give a short introduction to the skin anatomy. In sections 2.2 and 2.3, we will review the current technologies for removing JPEG blocking artifacts and uncovering vein patterns, respectively. Because we use the K-M model to develop one of our vein uncovering approaches, we will give a brief review of the model and its application in skin optics in section 2.4. In the later chapters, we will use a vein pattern extraction algorithm and an automatic vein matching algorithm to evaluate the uncovered vein patterns. Therefore, we will provide an introduction to the extraction and matching algorithms in section 2.5. For algorithm design, development and evaluation, we construct a large skin image database constituted by corresponding RGB and NIR images. In section 2.6, we will present our image capture system and current database.

2.1 Human Skin Anatomy

In this section, we give an overview of the physiological structure of the human skin, which is the fundamental to the whole research. More detailed information about the anatomy and functions of the human skin can be found in [10]-[14].

Skin is the largest organ of human beings and forms an interface between the body and the environment. It occupies around 12% ~ 16% of the total body mass, and covers an area of about $1.2 \sim 2 \text{ m}^2$ in adults [11]. Skin has the functions of thermoregulation, protection, metabolism, and tactile sensing. The thickness of skin is dependent on body sites and wear-and-tear, and its

variation contributes to the overall skin appearance. Skin can be divided into three major layers: the epidermis, the dermis and the hypodermis. Fig. 2.1 illustrates the human skin structure.

Fig. 2.1 Illustration of human skin structure. (www.medicinenet.com)

2.1.1 The Epidermis

The epidermis is the outermost part of the skin and can be subdivided into five layers (from superficial to the deepest): the stratum corneum, the stratum lucidum, the stratum granulosum, the stratum spinosum, and the stratum germinativum. Its average thickness is about 0.2 mm. There is no vein or capillary in the epidermis. The main substances in this layer that absorb and scatter light are melanin pigments [10][12].

- **Stratum Corneum:** This is the outermost layer, also called the horny layer. Its major components are hexagonal-shaped flattened dead cells full of keratin fibers. They are pushed towards the surface and gradually die and break down. A layer of sebum covers the stratum corneum. It is an oily-waxy substance created by the sebaceous glands which keeps the layer flexible and water resistant.
- **Stratum Lucidum:** This layer is also called the clear layer. It links the underlying stratum granulosum and the above stratum corneum.

- **Stratum Granulosum:** This layer is also called the granular layer. The keratinization process produces granules with lipids which prevent fluid loss from the body.
- **Stratum Spinosum:** This layer is also called the spiny layer. It is composed of several sublayers of new cells created by the underlying layer: the stratum germinativum.
- **Stratum Germinativum:** This layer contains the cells which are necessary in the regeneration of the above layers. The cells created in the mitotic processes of this layer move upwards through the above layers in the progressive maturation process of keratinization.

Melanin Pigments: They are the most important optical components in the epidermis. Melanin pigments are produced in the stratum germinativum, and can also be found in the stratum spinosum. The physiological function of melanin is to protect the underlying tissues from ultraviolet rays. Therefore the spectral absorption of melanin increases towards shorter wavelength. There are two types of melanin pigments: the black to dark-brown eumelanin and the yellow to reddish-brown pheomelanin. All individuals have different amount of melanin pigments.

2.1.2 The Dermis

The dermis represents about 90% of the skin mass. Its main functions are to provide nutrients, thermoregulation, and physical strength to the epidermis. The major components of the dermis are collagen, elastin fibers, nerves, and capillaries. Its average thickness is about 1 ~ 4 mm. The dermis can be subdivided into two layers: the papillary dermis and the reticular dermis. The main substances that absorb and scatter light in the dermis are hemoglobin, carotene, and bilirubin [10].

- **Papillary Dermis:** The major function of this layer is thermoregulation. Another important function of this layer is to provide nutrients to the epidermis. This layer contains a large number of nerves, capillaries, water, and cells. Collagen fibers form a fine network in this layer.

- **Reticular Dermis:** This layer contains fewer nerve and capillaries, and has a dense network.

The connective tissue in this layer gives elasticity and strength to skin. Hair follicles, glands, and sensory nerve endings are also contained in this layer.

Hemoglobin, Carotene and Bilirubin: The most important pigment that absorbs and scatters light in the dermis is hemoglobin. It can be categorized in two groups: oxy-hemoglobin (HbO_2) and deoxy-hemoglobin (Hb). The former is bound to oxygen and the latter is bound to nothing. Their absorption spectra are different. Carotene is an orange pigment which contributes to the overall skin color. However, its effect is much less than that of melanin and hemoglobin. Its presence in the body is mainly determined by food consumption. Bilirubin is a breakdown product of worn out blood cells. Hemoglobin in red blood cells breaks down into heme and globin. Heme is converted into bilirubin and carried to the liver where it is broken down and excreted. When the liver does not work properly, excessive bilirubin would make the skin turn a yellow-orange color.

2.1.3 The Hypodermis

The hypodermis is also called the subcutaneous tissue. Its major components are loose connective tissue and quantities of fat, which allow free movement of skin over the underlying tissue. It also contains nerves and blood vessels which are larger than those in the dermis [15]. Usually it is considered that the hypodermis plays no role in the optical properties of human skin. However, veins can be found in the transition of dermis and hypodermis, which has an influence on the optical properties of human skin. To improve the performance of the uncovering algorithm, the hypodermis is included in our fourth optical model.

2.1.4 Other Parts of the Human Skin

The following parts are superficial attributes of the human skin [10][12].

- **Hair Follicles:** Hair follicles reside in the dermis. Their bottom end is close to the boundary between the dermis and the hypodermis. Each follicle is associated with a bundle of muscle fibers to erect the hair for the purpose of thermoregulation.
- **Sweat Glands:** Sweat glands are simple tubular structures which open at the skin surface and produce sweat. They are located in almost all regions of the skin.
- **Sebaceous Glands:** Sebaceous glands are usually associated with hair follicles and secrete an oily liquid, called sebum, which keeps skin soft and smooth. They are located all over the body except hand palms and foot soles. There are more sebaceous glands in the areas such as face, scalp, chest, and back.

2.2 Approaches for JPEG-Compressed Image Deblocking

The purpose of image compression is to reduce the number of bits to represent a digital image for efficient storage and transmission. JPEG (Joint Photographic Experts Group) is the first international digital image compression standard for multilevel still images, both grayscale and color [16]. Because each block is transformed and quantized independently, the compressed images have blocking artifacts which are very annoying especially at high compression ratios. There has been a considerable amount of research on removing blocking artifacts. They can be classified into two strategies [17]. The first one uses preprocessing methods at the encoder side, while the second one uses postprocessing methods at the decoder side.

2.2.1 The Preprocessing Methods

Preprocessing methods have been commonly utilized in modern speech and audio coding. However, they have less application in the area of image or video coding. They can be classified into two groups: prefiltering and special encoding.

2.2.1.1 Prefiltering

Prefiltering techniques remove unnoticeable details in source images so that less information has to be coded. Therefore, the quality of compressed images can be enhanced under a given decoding algorithm and a bit rate constraint [17]. Tran et al. proposed a prefilter which flattens the input data of each DCT block to make them as homogeneous as possible [18]. Liang et al. proposed a block boundary pre/postfiltering to remove artifacts [19].

2.2.1.2 Special Encoding

Another preprocessing approach is based on modification of the encoder. Perceptual-based coding is a technique that uses a human visual system (HVS) model to optimize bit allocation so that the visible degradation is minimized [17]. Macq proposed a weighted optimum bit allocation method for orthogonal transforms [20]. Nill combined the DCT transform with a HVS model and obtained better performance in terms of compression ratio and image quality [21]. Bruckstein et al. defined a criterion based on sampling factor of an input image and optimized it for better performance [22]. Chang et al. proposed a direction-adaptive partitioned block transform to exploit the directional features in color images to improve coding performance [23]. Yamatani et al. proposed a poly-harmonic local cosine transform which modifies both the encoder and decoder of the JPEG baseline method [24].

A significant weakness of the preprocessing methods is that they modify the encoder. It is impractical for our forensic application because we have no control over the photographing process in acts of crime.

2.2.2 The Postprocessing Methods

Postprocessing techniques at the decoder side obtained a lot of attention and became an active research area in the late 1990s. Their advantage over preprocessing methods is that they can remove compression artifacts neither increasing the bit rate nor modifying the coding procedure. Generally speaking, the postprocessing algorithms can be classified into two groups: image enhancement and image restoration [17]. Image enhancement algorithms aim at improving the perceived quality subjectively. These algorithms take into account human visual sensitivities and the special structure of blocking artifacts. The image restoration algorithms try to recover the lost information in JPEG compression.

2.2.2.1 Postprocessing by Image Enhancement

The major concern about quality in most applications of visual communication is not how close a resultant image is to its original one, but how well it is perceived by a human observer [17]. Therefore, deblocking methods can be designed to match human perception. Postprocessing by image enhancement is this kind of approach which aims at smoothing the artifacts instead of restoring pixels back to their original values.

A. Postfiltering

Because JPEG blocking artifacts are high frequency components, a direct solution is to apply low-pass filters to the regions where artifacts occur [25]. Filtering can be performed either in spatial domain or in frequency domain. To avoid image blurring, a number of adaptive filtering techniques based on local statistics have been proposed. Generally speaking, block classification is always used as an initial step to distinguish different types of regions [26]. Then different filters (either linear or nonlinear) are adopted according to the label information.

Filtering in Spatial Domain: Reeves and Lim proposed to apply a symmetric 3×3 Gaussian space-invariant filter to the pixels along the block boundaries [27]. This is considered to be the first ever filtering method. However, it blurs the image. Ramamurthi and Gersho proposed a two-dimensional (2-D) filter to reduce the grid noise in the monotone area, and a one-dimensional (1-D) filter to reduce the staircase noise along the edges [28]. Hsu and Chen proposed an adaptive separable median filter which not only reduces the blocking artifacts, but also preserves the edges [29]. Lee et al.'s approach includes three components: an adaptive filter, a corner outlier detection, and a replacement scheme [30]. Averbuch et al. applied weighted sums on pixel quartets located at the intersection of four blocks [31]. The weights are obtained from a constrained two-dimensional function. Zhai et al. proposed a method which performs filtering in shifted windows [32]. Chou et al. proposed a fast algorithm that removes blockiness through nonlinear smoothing [95].

Filtering in Frequency Domain: Blocking artifacts are direct consequences of quantization errors of the DCT coefficients. Restoring these coefficients is a natural and more accurate approach to tackle the problem. The restored value must lie within its corresponding quantization interval. Chen et al. proposed an adaptive weighting mechanism that integrates the masking effect of the HVS [33]. The DCT coefficients are filtered in different ways for the blocks with different frequency properties. The quantization constraint is finally applied before the inverse DCT transform. Zeng applied zero-masking to the DCT coefficients in some shifted windows located on the boundaries of a target block [34]. The boundaries are modeled as a 2D step function contaminated by i.i.d. (independent and identically-distributed) noise with zero mean and a small variance. Singh et al. put forth a method that detects blocking artifacts based on HVS and block statistics [35]. They removed the artifacts in smooth regions by appropriately modifying a few DCT coefficients, and applied an edge-preserving smoothing filter to the non-smooth regions such as genuine edges. Chang and Chen proposed an algorithm that adopts a shift window within four adjacent DCT blocks to reduce computational complexity [36]. The artifacts are eliminated by modifying several DCT coefficients in the shifted window. Foi et al. proposed a filtering

approach based on the shape-adaptive DCT (SADCT) transform [94]. They performed hard-thresholding and Wiener filtering in SADCT domain, with an arbitrarily-shaped transform's support which is adaptively defined for every point in the image.

Hybrid Filtering: Many researchers designed other filtering methods that either combine both spatial and frequency information, or integrate some other algorithms (e.g., neural networks and fuzzy schemes). Luo and Ward proposed an adaptive approach which works in both DCT and spatial domains [37]. For smooth regions, the original pixel level in the same block is continuous. This method uses this property and the correlation between neighboring blocks to reduce the discontinuity across the boundaries in smooth regions. For texture and edge regions, it applies an edge-preserving smoothing filter. Tsai et al. used optimal decimation and interpolation filters [38]. They discovered that optimal filter searching is a separable least squares problem, which can be solved by variable projection method. Vo et al. proposed an adaptive fuzzy filter based on the variation of pixel values [39]. Fuzzy filter is a special case of bilateral filters which removes the artifacts and retains the sharpness of real edges at the same time [39]. Kwon et al. proposed a more accurate filtering algorithm which exploits the nonlinear learning property of neural networks [40].

We have reviewed three types of postfiltering methods. Because of their low computational complexity, they are the most popular deblocking methods. For these methods, adaptability is necessary to avoid image blurring. However, the adaptation to local statistics relies on a good classification scheme, which is usually sensitive to quantization and compression ratio. Poor estimation and adaptation can cause degradation of image quality.

B. Alternating Current (AC) Prediction

Another postprocessing approach is to recover quantized AC coefficients. Niss predicted low-frequency AC coefficients from Direct Current (DC) coefficients of a given block and its eight neighboring blocks, but remained high-frequency AC coefficients unchanged [41]. Lakhani found

that the values of AC coefficients predicated by Niss' algorithm are very coarse and improved it by restricting the interpolation region to 16×16 pixels, instead of 24×24 [42]. In 2008, he developed additional prediction equations for different configurations of the interpolation area [43]. AC prediction methods can considerably reduce the blocking artifacts in smooth areas, but the prediction is not reliable in the neighborhood of edges [17].

C. Enhancement Based on Perceptual Model

According to the perceptual model of HVS, each perceptual channel has a threshold depending on background of input signals. If the contrast of an input signal is lower than the threshold, it cannot be detected, which is known as the masking effect [44]. Macq et al. proposed a postprocessing scheme based on this property [44]. They separated a compressed image into several perceptual channels and eliminated the signals in the channels whose noise is above the visibility thresholds. Then they used the signals in the remaining channels to reconstruct the image. A major problem with this method is that the visibility thresholds are determined by comparing a compressed image with its original image. Because the original image is not available, these thresholds are actually obtained by estimation, which may not have reliable results.

2.2.2.2 Postprocessing by Image Restoration

For the image restoration approaches, artifact removal is an ill-posed image recovery process. Many classical image restoration methods have been adjusted for JPEG image deblocking. They can be classified into two categories: criterion-based methods and constraint-based methods.

A. Criterion-Based Methods

The basic concept of these methods is to find an optimal image from a JPEG compressed image based on some predefined criteria. For generic images, the best criterion should be perceptual

image quality. However, modeling perceptual image quality is still an on-going research topic. The mean square error and posterior probability are the commonly used criteria [17].

Criterion of Minimum Mean Square Error (MMSE): The minimization of mean square error can be achieved by using linear or nonlinear estimators. Nakajima et al. proposed an adaptive restoration algorithm [45], which estimates the quantization information and pixel values based on local statistics and a linear estimator. One important assumption of this method is that the compression error is zero-mean white noise uncorrelated with the original image. Hong et al. proposed a non-iterative method which classifies small local boundary regions based on their intensity distributions. Then appropriate linear predictors are chosen to estimate the corresponding boundary pixels in the regions [46]. The major problem with these methods is that mean square error is not very well matched to perceptual visual quality [47].

Criterion of Maximum A Posteriori (MAP) probability: The criterion in MAP-based restoration is to maximize the probability density of the original image conditioned on the compressed image [17]. In other words, this restoration is actually a process of looking for a candidate original image with maximum a priori probability. The a priori probability is usually modeled by a Gibbs distribution associated with a potential function, which plays a crucial role in the restoration. Stevenson and co-workers were the first research group investigating MAP-based JPEG deblocking methods [48]-[49]. They used Huber minimax function as the potential function, i.e., the original image was modeled as a Huber Markov random field (MRF). This potential function enforces neighboring pixels to have similar gray levels. Gradient projection method was adopted to perform the optimization. Luo et al. chose the same potential function, but used iterative conditional mode (ICM) to do the optimization [50]. Sun et al. modeled original images as a high order MRF based on the fields of experts (FoE) framework, which is more expressive [51]. Özcelik used a non-stationary Gauss-Markov model [52] to formulate the problem, and mean field annealing to seek an optimal solution. Li and Kuo proposed a multiscale MAP method, in which a compressed image was restored from coarse to fine scales [53]. Their

method improves global appearance and reduces artifacts at coarse scales, while preserves sharpness of edges at fine scales.

B. Constraint-Based Methods

The basic idea behind these methods is to impose constraints on compressed images and to restore them accordingly [17]. Constraints can be obtained from prior knowledge, such as the compression algorithm and image properties. The projection onto convex sets (POCS) method is a typical example. Its constraint set is a closed convex set whose elements are designed according to prior knowledge of original images. Given m properties from prior knowledge, m closed convex sets can be constructed. An iterative process can be used to find a possible solution which is in the intersection of all the sets. In POCS-based deblocking methods, two constraint sets are commonly used. One is the frequency constraint set determined by the quantization interval of each DCT coefficient, and the other is the spatial constraint set that captures the smoothness properties of original images. Up to now, various new constraints have been designed. Rosenholtz and Zakhor chose the band-limited property as a constraint set [54]. Yang et al. selected the discontinuities along block boundaries as a constraint set [55]. Weerasinghe et al. used the uniform property in smooth regions as a constraint set, which allows different degrees of smoothing in different image regions, while preserving image edges [56]. Park et al. defined dynamic ranges and structures of block boundaries as constraint sets [57].

POCS-based methods have three limitations: Firstly, they have no optimization criterion [58]. When all the constraints are satisfied, the projection iteration will stop. Secondly, users are required to define the projection operation for every constraint set. Improper operations may lead to divergence, and provide even worse results. Thirdly, their computational complexity is extremely high, because they are iterative algorithms and perform forward and inverse DCT calculations in each iteration.

We have reviewed the current techniques for JPEG image deblocking. They were designed for generic images, and therefore, prior knowledge from target images is hard to be extracted and used to recover the lost information. In addition, the compressed images and corresponding resultant images from these methods may have different quantized DCT (QDCT) coefficients. In other words, the difference between original (uncompressed) images and resultant images may be even larger than that between original images and compressed images in terms of QDCT coefficients. In chapter 3, we will present novel approaches to remove blocking artifacts especially for skin images.

2.3 Approaches for Vein Uncovering

The major part of this thesis is to develop specific algorithms for uncovering body vein patterns from color images. Currently there is no algorithm or software package especially designed for this purpose. We found two types of related approaches: image enhancement methods and independent component analysis (ICA)-based method. Physically, a skin image is formed by the light reflected from the skin. Image enhancement methods tackle the problem from the image processing point of view, while the ICA-based method addresses it from the optical point of view.

2.3.1 Image Enhancement Methods

Image enhancement aims at improving image quality for human visual perception, and for vision applications. After enhancement, pixel values are usually modified so that the resultant image is visually pleasing, or more suitable for analyses [59].

Traditional enhancement methods for grayscale images include contrast stretching, gray scale slicing, and histogram equalization. However, we cannot directly generalize these techniques to color images, because the color components do not exist in grayscale images. A possible solution is to apply the grayscale enhancement methods only to the R channel, because

veins are most visible in this channel. For skin with very low concentration of melanin (e.g. Caucasian skin) and fat, this approach is useful. Fig. 2.2 (a) shows a thigh of a Caucasian female, and Fig. 2.2 (b) shows its resultant R channel. However, for skin with high concentration of melanin and/or fat, which includes skin of Asians and Africans, this method does not work. Fig. 2.3 (a) shows a thigh with relatively high concentration of melanin and fat, and Fig. 2.3 (b) shows the resultant image. It is still very difficult to observe the vein patterns.

An alternative strategy is to utilize all the color information. However, processing each chromatic channel separately by the same greyscale technique has been proven to be inapplicable because of color shifts [60]. Fig. 2.4 (b) is a resultant image from an image enhancement algorithm based on Laplacian transform [61]. Its original image is given in Fig. 2.4 (a). Although the color contrast is increased a lot, it is still difficult to locate the veins. In the past two decades, some advanced techniques have been proposed for contrast enhancement and color image denoising, such as rank-order filters [62], locally adaptive filters [63], methods based on partial differential equations [64]-[65], bilateral filtering [66], retinex method based on the image formation model [67]-[70], etc. They try to enhance natural images under the presumption that the hue attribute is kept unaltered. This principle makes these methods unsuitable for vein pattern uncovering.

Some commercial image software packages nowadays have functions for color image enhancement, but none of them targets our application. Figs. 2.5 (b) and (c) are two resultant images from Photoshop, and Fig. 2.5 (a) is their original image. These results are still not satisfactory.

Fig. 2.2 The enhancement results by contrast stretching (example 1).
(a) is an image of a thigh with low concentration of melanin and fat; (b) is the enhanced R channel of (a).

Fig. 2.3 The enhancement results by contrast stretching (example 2).
(a) is an image of a thigh with high concentration of melanin and fat; (b) is the enhanced R channel of (a).

Fig. 2.4 The enhancement results by Laplacian transform.
(a) is an image of a thigh with high concentration of melanin and fat; (b) is the enhanced image of (a).

Fig. 2.5 The enhancement results by Photoshop.
(a) is a skin image; (b) is the enhanced image generated by the function “equalize”; (c) is the enhanced image generated by modifying the contrast.

2.3.2 Independent Component Analysis-Based Method

The majority of measured quantities in nature are mixtures of other quantities. The followings are some classical examples: (i) voice signals in a room where several people are talking at the same time, (ii) an electroencephalogram signal which is affected by many different brain regions, and (iii) a person's height, which is determined by many different genetic and environmental factors [71]. As a member of blind source separation (BSS) methods, ICA can recover the independent source signals from a set of measured quantities [72].

The light reflected from skin is also a mixture of several signals. According to the Dichromatic Reflection Model, the light reflected from a point on a dielectric material is a mixture of the light L_s reflected at the material surface, and the light L_b reflected from the material body [73]. Fig. 2.6 illustrates this process. Human skin is a multilayer dielectric material. The light that we receive from skin can be modeled as a linear combination of the light reflected from its surface, and the light refracted and reflected from its inner parts.

Tsumura used ICA to analyze and synthesize the color and texture of human skin [74]-[77]. An input image is first separated into reflections of light from surface and body. Then, the signal from body is decomposed into two components that correspond to the distributions of melanin in the epidermis and hemoglobin in the dermis [10]. These distribution maps can provide useful information of veins. It should be emphasized that Tsumura's method is for computer graphic applications.

An important assumption of this ICA-based method is the linearity among quantities of pigments. However, the complex structure of the skin very likely violates this assumption [75]. It is also assumed that spectral absorbance of the pigments does not change spatially. Nevertheless, hemoglobin has two states: oxy-hemoglobin (HbO_2) and deoxy-hemoglobin (Hb), which have different absorption coefficients. Consequently, this method may be useful for some skin images (Fig. 2.7), but useless for others (Fig. 2.8). Another important problem with this method is that it may not converge when applied to some hairy or highly blocky skin images. Fig. 2.9 shows a

compressed image and its results. Fig. 2.10 shows an image on which the ICA-based method cannot converge.

We have reviewed two classes of related techniques, and pinpointed their weaknesses for our target application. These techniques were not designed for uncovering vein patterns from color skin images. In chapters 4 ~ 7, we will propose two types of approaches to uncover vein patterns hidden in color images. One type is based on image mapping, and the other type is based on human skin biophysics and its optical properties.

Fig. 2.6 Light reflection model for dielectric materials.

Fig. 2.7 Successful application of ICA-based method.

(a) is a color image; (b) is the 1st component (melanin); (c) is the 2nd component (hemoglobin).

Fig. 2.8 Failed results from ICA-based method (example 1).
(a) is a color image; (b) is the 1st component (melanin); (c) is the 2nd component (hemoglobin).

Fig. 2.9 Failed results from ICA-based method (example 2).
(a) is a compressed color image; (b) is the 1st component (melanin); (c) is the 2nd component (hemoglobin).

Fig. 2.10 An example image that ICA-based method cannot converge.

2.4 The K-M Model and Its Application in Skin Optics

2.4.1 The K-M Model

Physical models that quantitatively describe light transport in tissues have been developed. Radiative transfer equation (RTE) is the most accurate one, which describes the process of light transport in a scattering medium (e.g. human tissue) [10]. However, at this moment, its exact analytical solution cannot be obtained for the complex and multiple scattering medium such as human skin, because no biophysical experiments were performed to estimate the required parameters. Fortunately, there are some practical models that approximate light transport in tissue. The K-M model is the most popular and simplest one. The K-M model assumes that the radiation passing through a scattering medium is divided into two diffuse fluxes: one travels in the forward direction, and the other travels in the backward direction, which are respectively denoted as I and J , in Fig. 2.11.

Fig. 2.11 Illustration of light transport based on the K-M model.

According to the K-M model, at a distance x from the surface, the changes in flux over an infinitesimal distance dx for I and J are:

$$dI = -(k + s)I dx + sJ dx, \quad (2.1)$$

$$dJ = (k + s)Jdx - sI dx, \quad (2.2)$$

where $k = 2\mu_a$ and $s = 2\mu_s$. μ_a (μ_s) is the absorption (scattering) coefficient of the layer at a particular wavelength, i.e., the fraction of radiation absorbed (scattered) per unit path length.

Using the boundary conditions ($J_d = 0$), the reflectance $R = J_0 / I_0$ and transmittance $T = I_d / I_0$ shown in Fig. 2.11 are obtained as:

$$R = \frac{(1 - \beta^2)(e^{Kd} - e^{-Kd})}{(1 + \beta)^2 e^{Kd} - (1 - \beta)^2 e^{-Kd}}, \quad (2.3)$$

$$T = \frac{4\beta}{(1 + \beta)^2 e^{Kd} - (1 - \beta)^2 e^{-Kd}}, \quad (2.4)$$

where

$$K = \sqrt{k(k + 2s)}, \quad (2.5)$$

$$\beta = \sqrt{\frac{k}{k + 2s}}. \quad (2.6)$$

The reflected light thus depends on three parameters of the material: the absorption coefficient, μ_a , the scattering coefficient, μ_s , and the material thickness, d . In general, the two optical coefficients are wavelength-dependent, i.e., $\mu_a(\lambda)$ and $\mu_s(\lambda)$.

For an n -layered material, let the reflectance and transmittance for the i^{th} layer be R_i and T_i , respectively. The total reflectance, $R_{12\dots n}$ and transmittance, $T_{12\dots n}$ can be computed by the recursive equations:

$$R_{12\dots n} = R_{12\dots n-1} + \frac{T_{12\dots n-1}^2 R_n}{1 - R_{12\dots n-1} R_n}, \quad (2.7)$$

$$T_{12\dots n} = \frac{T_{12\dots n-1} T_n}{1 - R_{12\dots n-1} R_n}. \quad (2.8)$$

Fig. 2.12 illustrates the process of light transport in a two-layered material.

Fig. 2.12 Illustration of the process of light transport in a two-layered material.

2.4.2 The Skin Optics Based on the K-M Model

The K-M model has been frequently used in modeling optical properties of different materials. In skin optics, Anderson was the first to use it to model optical properties of the dermis tissue [79]-[80]. His team also did biomedical experiments to verify their results. Wan et al. extended this model to compute the absorption and scattering coefficients of the epidermis tissue, taking into account both collimated and diffuse incident irradiance [81]. Diffey proposed a method based on the K-M model which considers both forward and backward scattering, and allows changes in the refractive index at the air / skin interfaces [82]. Gemert et al. further distinguished the cases of dominant absorption, dominant scattering, and scattering about equal to absorption [83]. They found that skin layers show strongly forward scattering. Cotton and Claridge used the K-M model to aid the categorization of pigmented skin lesions from color images [84]. They further derived diagnostically important factors from the internal structure and composition of pigmented skin lesions [85]. Their model takes into account the presence of melanin and hemoglobin. Doi and Tominaga modeled human skin as a two-layered turbid material [86]. They used the K-M equations with six unknown parameters to estimate the reflectance. These parameters were obtained by fitting the estimated reflectance to empirical values using the least squares method. Donner and Jason presented a model for light diffusion in multi-layered translucent materials [86]. They took into account multiple scattering between slabs by using a variant of the K-M

model in the frequency space. Störring used the K-M model to investigate skin color formation under changing and mixed illumination [88]. The K-M model was also used to reveal the structure of pigmented skin lesions for diagnosis [85] and predicate optical properties of the skin for laser surgery [79]-[80].

The current research on skin optics comes from different areas and has distinct purposes. Those developed by computer graphics researchers aim at skin appearance modeling, those developed by computer vision researchers aim at skin segmentation, while those developed by medical researchers aim at skin property estimation and dermatological diagnosis. None of these applications are related to forensics. *To our best knowledge, there are no prior reports of similar research or applications to uncover vein patterns from color images for forensic investigation.*

2.5 Vein Pattern Extraction and Automatic Matching Algorithms

In the later chapters, we will use a vein pattern extraction algorithm and a three-step automatic vein matching algorithm to evaluate the uncovered vein patterns. The algorithms were developed by Zhang et al. [115]. In this section, we give an introduction to the two algorithms.

2.5.1 A Vein Pattern Extraction Algorithm

The vein pattern extraction algorithm can be divided into four parts: preprocessing, local information estimation, enhancement, and representation. The aim of preprocessing is to alleviate problems from uneven illumination. To estimate the local directional information of veins, a Gabor filter bank is used to generate two information maps. Then, structural information of veins is captured from these information maps for vein pattern enhancement. Finally, the vein patterns are represented by a set of points for matching.

2.5.1.1 Preprocessing and Local Information Estimation

First, the contrasts of images are normalized by the contrast-limited adaptive histogram equalization (CLAHE) method [116]. It divides an image into small regions and stretches the histogram of each region into a desired distribution. In this algorithm, each image is divided into 8×8 blocks and uniform distribution is used as an objective histogram. Fig. 2.13 (b) shows the preprocessing result of Fig. 2.13 (a).

Fig. 2.13 A preprocessing result.
(a) is an original NIR image; (b) is the result of CLAHE.

Gabor filters have been proven to be a powerful tool to capture local information. The algorithm applies a filter bank made up of the real parts of 16 Gabor filters with different scales and orientations to NIR images and uncovered images to capture their local information. Only the real parts of the Gabor filters are used because veins are dark ridges in NIR and uncovered images from which vein patterns will be extracted. A real part of a Gabor filter in the spatial domain is defined as

$$G(x, y, \lambda_{mk}, \theta_k, \sigma_m, \gamma) = \frac{\gamma}{2\pi\sigma_m^2} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma_m^2}\right) \cos\left(2\pi \frac{x'}{\lambda_{mk}}\right), \quad (2.9)$$

where $x' = x\cos\theta_k + y\sin\theta_k$, and $y' = -x\sin\theta_k + y\cos\theta_k$ are the rotated coordinates with orientation $\theta_k = k\pi / 8$; λ_{mk} represents the wavelength of the sinusoidal component; σ_m is the standard deviation of the elliptical Gaussian window along x' direction; γ is the spatial aspect ratio; $m \in \{1, 2\}$ is the scale index and $k \in \{1, 2, \dots, 8\}$ is the orientation index. The DC component of the filter is removed so that it is robust against brightness variation and its power is normalized. Let G_{DC} be the processed filter, $Img(x, y)$ denote an NIR image or an uncovered image and $F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}$ denote a filtered $Img(x, y)$, which can be obtained by:

$$F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y) = -G_{DC}(x, y, \lambda_{mk}, \theta_k, \sigma_m, \gamma) * \text{Img}(x, y), \quad (2.10)$$

where $*$ represents an operation of a two-dimensional convolution. Thus, for an input image, 16 filtered images are generated by the filter bank.

Each filter is sensitive to veins with a particular direction, frequency and scale and therefore, the orientation and scale from the filter giving maximum response can be used to represent the orientation and scale of the vein, respectively. Although the DC of the filters is removed, the filter response still depends on image contrast, which is highly influenced by poses and illumination conditions. Thus, $F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y) / Q_m(x, y)$, where Q_m is a local image power around the point (x, y) , serves as a local information estimator [117]. Q_m is the square root of the summation of pixel value square around the point (x, y) . Note that Q_m depends on the size of the filters. The orientation of veins can be obtained from:

$$O(x, y) = \arg \max_{\theta_k, m, k} [F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y) / Q_m(x, y)], \quad (2.11)$$

which is called the orientation map. In addition,

$$E(x, y) = \max_{m, k} [F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y) / Q_m(x, y)], \quad (2.12)$$

which is called the response map, indicates the quality of local information. Fig. 2.14 (b) and (c) illustrate the Gabor filter outputs. We can observe vein patterns from the response and orientation maps.

Fig. 2.14 The Gabor filter outputs.

(a) is an original NIR image; (b) and (c) are respectively the response and orientation maps of (a).

2.5.1.2 Enhancement and Representation

Response maps, especially the ones generated from uncovered images, are very noisy because of low image quality. Structural information of veins in the orientation maps is used to enhance the response maps. In orientation maps, short lines and small pixel groups are likely to be hairs or other noise, while long lines and large pixel groups give us more confidence that they are veins. These characteristics are used to design the enhancement scheme. First the connected component labeling method [127] is used to group pixels into components based on their orientations. All pixels in the same component are connected and have the same orientation. Then a threshold of the filter response is used to find out potential pixels and potential vein components. After that the potential vein components are further connected based on orientation difference and spatial distance between each pair of two groups. More clearly, if their angular distance is smaller than or equal to $\pi / 8$ and their minimum distance is smaller than a threshold, they are connected. We iteratively search through all the component pairs and assign the same labels to connected vein components and generate a map of grouped labels Ψ . $g(i) = \{ (x, y) \mid \Psi(x, y) = i \}$ represents pixels with the same label i and $|g(i)|$ represents its size in term of number of pixels. A weighting function $W(x, y)$ based on the size of the groups is defined as

$$W(x, y) = \kappa \cdot \left\{ 1 + \exp \left(- \frac{|g(i)|}{\mu} \right) \right\}^{-1}, \quad (2.13)$$

where $(x, y) \in g(i)$. The parameter μ controls suppression and enhancement regions and the parameter κ is a magnification factor. This weighting function is applied to the response map $E(x, y)$ point-wisely.

The enhanced vein patterns are binarized using Otsu's method and then skeletonized. Finally, the vein patterns are uniformly sampled and represented by a set of points. The boundaries of body parts are also sampled and represented in the same way. Fig. 2.15 illustrates the vein pattern extraction process.

Fig. 2.15 The process of vein pattern extraction.

(a) is an uncovered image; (b) is the CLAHE result of (a); (c) and (d) are respectively the response and orientation maps obtained from (b); (e) is the enhanced image; (f) is the binarized image of (e); (g) is the skeletonized image of (f); (h) is the sampled points from the vein patterns in (g), and (i) shows the sampled points from both the vein patterns and the body boundaries.

2.5.2 A Three-Step Vein Matching Algorithm

A three-step vein matching algorithm was specially designed to handle noisy vein patterns. Given two vein patterns, two sets of points in $\mathbb{R}^2$ are used to represent them. One is denoted as a model set M with size of N_m by 2, and the other is denoted as a data set D with size of N_d by 2. The point sets from their body parts' boundaries are respectively denoted as M_{BD} with size of N_{mbd} by 2 and D_{BD} with size of N_{dbd} by 2. $T_R(\bullet, \tau_R)$ represents a rigid transformation controlled by a parametric vector τ_R and $T_N(\bullet, \tau_N)$ represents a non-rigid transformation controlled by a parametric vector τ_N . In general, point matching methods determine a parametric vector τ in a transformation T , which is applied to the model set, i.e., $T(M, \tau)$, so that the two point sets are aligned optimally and their difference can be measured. The matching scheme uses two explicit steps to clear noise and outliers. The first step is to remove points outside the overlapping area and the second step is to remove outliers and noise in the overlapping area. The final step is to deal with non-linear distortion and calculate the dissimilarity between two patterns.

Step 1 is to remove points outside the overlapping area of M and D . Boundaries of body parts are simple yet contain a good amount of shape information for rough alignment. A rigid transformation T_R is used to describe the deformation between the two sets of boundary points M_{BD} and D_{BD} . The rigid Coherent Point Drift (CPD) method is used to determine the alignment parameters τ_1 . Using the rigid transformation parameter τ_1 , the vein pattern point set M is transformed into $M_{T1} = T_R(M, \tau_1)$. Then the overlapping area B of M_{T1} and D is determined. Points that are located outside B are regarded as outliers and removed. The pruned vein pattern point sets are denoted as M_1 and D_1 . If the numbers of points in these two sets are less than a threshold, they are considered from two different vein patterns.

Step 2 is to remove outliers inside M_1 and D_1 . A rigid transformation T_R is used to roughly align M_1 and D_1 . Note that M_1 is regarded as a model set in this alignment. Using the rigid CPD method, the resultant parametric vector τ_{21} and correspondence c_{21} between M_1 and D_1 can be obtained. With the roughly aligned vein patterns $T_R(M_1, \tau_{21})$ and D_1 , the distances between corresponding points in $T_R(M_1, \tau_{21})$ and D_1 are calculated. If the distances are very large, the corresponding points in $T_R(M_1, \tau_{21})$ are considered outliers. All the points in $T_R(M_1, \tau_{21})$ are sorted in ascending order according to their distances and the top 80% of points in $T_R(M_1, \tau_{21})$ are retained. They are denoted as M_2 . The previous process is repeated to remove outliers in D_1 . The rigid CPD method is applied to align D_1 and M_2 . In this alignment, D_1 is considered as a model set. As with the previous case, only 80% of points in D_1 are retained and distances between corresponding points in $T_R(D_1, \tau_{22})$ and M_2 are used as a selection condition. The remaining points in D_1 are denoted as D_2 .

Step 3 is to match M_2 and D_2 with a non-rigid transformation. In this step, the non-rigid CPD method with a Gaussian kernel as a regularizer is used to match M_2 and D_2 . With the non-rigid transformation T_N and corresponding parameters τ_3 generated by the CPD method, $T_N(M_2, \tau_3)$ and D_2 can be aligned optimally. As with the previous steps, a correspondence vector c_3 can be obtained. Note that multiple points in M_2 can correspond to the same point in D_2 and data points with no correspondence are recognized as outliers.

Finally, the dissimilarity of two vein patterns is calculated based on the correspondence vector c_3 generated in Step 3. Let (x_{mi}, y_{mi}) in $T_N(M_2, \tau_3)$ and (x_{di}, y_{di}) in D_2 be a correspondence in c_3 , their Euclidean distance be e_i and r be a set containing the smallest 80% of e_i . A dissimilarity function defined as $s = |r|^{-1} \times \sum_{ei \in r} e_i$, where $|r|$ represents the number of elements in the set r , is used in the scheme.

Fig. 2.16 The three-step matching result.

(a) and (b) are vein patterns uncovered from color images from same person collected at different occasions; (c) is the points sampled from two vein patterns and their boundaries; (d) is the result of Step 1; (e) is the rigid registration result of Step 2; (f) and (g) are the results of outlier removal in the two sets; (h) is the result of final non-rigid matching.

2.6 The Database

To carry out our research, a skin image database is essential for algorithm design, development and evaluation. However, there is no skin image database publicly available. We construct a large skin image database constituted by corresponding RGB and NIR images. In this section, we give a brief introduction to our image capture system.

Our data acquisition system includes two cameras: a digital single-lens reflex camera (Nikon D70s or Canon 500D) and a multi-spectral 2-CCD camera (JAI's AD-080 CL). The digital single-lens reflex (DSLR) camera is used to capture color images in the visible spectrum. The multi-spectral 2-CCD camera is used to capture simultaneous images from both the visible and NIR light spectra through a single lens using two channels. The first channel captures visible light using a Bayer mosaic color imager, while the second channel captures NIR light using a monochrome imager. The color imager captures conventional visible bands (red, green and blue), and the monochrome imager detects what the human eyes cannot see. The 2-CCD camera not only enables us to measure both visible and NIR light, but also offers easy set-up and alignment. The equipment setup and the collection process are illustrated in Fig. 2.17 and Fig. 2.18, respectively.

We collect images from arms of each participant in two sessions, apart from around two weeks. For each arm, we collect six images, two from the DSLR camera and four from the 2-CCD camera in each session. To simulate photos collected from crime scenes, we also take some images from other body positions and some non-standard images from each participant with different viewing angles and poses.

An image capture software package, which connects the two cameras to a computer, was developed by Hangyi Zhang and other researchers. Fig. 2.19 shows its interface. With only one click, the software package collects three images, one from the DSLR camera and two from the multi-spectral 2-CCD camera, assigns well-defined file names, and saves them in proper

locations. It significantly reduces the workload of data collection. The current database contains images contributed by participants from diverse racial, age, occupation, and BMI groups.

Due to the development process of our database, in the following chapters, we will use databases of different sizes to evaluate our algorithms.

Fig. 2.17 Equipment setup for data collection.
(a) front view; (b) back view.

Fig. 2.18 The image capture process.

Fig. 2.19 Interface of the image capture software package.

Chapter 3

Skin Image Deblocking

With the advance of digital camera technology, we can easily collect a large number of high resolution photos. Storing and transmitting these photos are troublesome, because of their sizes. Most compact digital cameras can produce images with size of 12 mega pixels. Some DSLR cameras can even produce images with over 20 mega pixels. The standard JPEG algorithm installed in consumer cameras is commonly used to compress these images. A well-known disadvantage of this algorithm is the blocking artifact. Using biometric traits from the skin for criminal and victim identification highly depends on the quality of evidence images, because the size of these traits in the images is usually very small. As a result of compression, vein patterns can be broken, and skin marks can be blurred, or even totally removed, especially under high compression ratios. Fig. 3.1 illustrates the effects of blocking artifacts on pigmented skin marks and vein patterns. Directly applying our vein uncovering algorithms to a compressed skin image will produce blocky and disconnected vein patterns, which are almost impossible to discern or extract. Therefore, an important preprocessing step is to remove blocking artifacts in JPEG-compressed skin images and improve the image quality before any forensic analyses.

Chapter 2 reviews the current deblocking technologies. A good method should not only remove the boundaries between blocks, but also recover the lost information. The current technologies are extremely hard, if not impossible, to achieve the second goal, because they are designed for generic images, and therefore, cannot utilize prior knowledge from target images. In fact, these methods make the situation even worse, because they generally smooth images, including biometric traits, to alleviate blocking artifacts. Fig. 3.2 shows resultant images from some of these methods. In addition, the difference between original (uncompressed) images and their resultant images may be even larger than that between original images and compressed

images in terms of quantized Discrete Cosine Transform (QDCT) coefficients, because many of these methods do not guarantee that the resultant and compressed images have the same QDCT coefficients.

Fig. 3.1 The effects of blocking artifacts on pigmented skin marks and veins.

(a) Uncompressed color image with a common nevus and (c) ¹V component with a clear vein pattern; (b) is the JPEG compressed image of (a) with compression ratio of 102.51 and quality factor of 25; (d) is the JPEG compressed image of (c) with compression ratio of 95.63 and quality factor of 25.

Fig. 3.2 The smoothing effect of traditional deblocking methods on skin images.

(a) Uncompressed color image with a common nevus and (e) V component with a clear vein pattern; (b) is the JPEG compressed image of (a) with compression ratio of 115.32 and quality factor of 25; (f) is the JPEG compressed image of (e) with compression ratio of 120.15 and quality factor of 25; (c) and (g) are respectively the results of (b) and (f) from a post-filtering method [37]; (d) and (h) are respectively the results of (b) and (f) from a MAP-based method [51].

In this chapter, two approaches – MAP-based and knowledge-based approaches, will be presented. We will extensively use prior knowledge from our database to solve the problem. It

¹ The JPEG algorithm transforms an image from RGB space to YUV space before performing the DCT Transform.

should be emphasized that our approaches guarantee that the resultant and compressed images have the same QDCT coefficients. In this chapter, sections 3.1 and 3.2 report the MAP-based and knowledge-based approaches and experimental results. Section 3.3 offers some summary remarks.

3.1 An MAP-Based Approach

This approach formulates skin image deblocking as an estimation problem. It embeds statistical information of skin images into a maximum a posteriori (MAP) model to perform the estimation. The resultant images and corresponding original skin images share the same statistical information. A potential function that connects the MAP model and skin statistics is proposed first. Then we will present an algorithm to optimize the potential function. Subsection 3.1.1 gives the MAP formulation. Subsection 3.1.2 presents the potential function. Subsection 3.1.3 proposes the optimization algorithm of the function. Subsection 3.1.4 reports the experimental results.

3.1.1 An MAP Formulation

Let X be an original skin image, and Y be the compressed image from X . Using an MAP approach, the problem of deblocking can be formulated as an estimation i.e.,

$$\hat{X} = \arg \max_x L(X | Y), \quad (3.1)$$

where $L(X | Y) = \log p(X | Y)$ is a log-likelihood function. According to the Bayesian rule, Eq. (3.1) can be rewritten as:

$$\begin{aligned} \hat{X} &= \arg \max_x \left\{ \log \frac{p(Y | X)p(X)}{p(Y)} \right\} \\ &= \arg \max_x \{ \log p(Y | X) + \log p(X) - \log p(Y) \} \\ &= \arg \max_x \{ \log p(Y | X) + \log p(X) \}, \end{aligned} \quad (3.2)$$

where the term, $p(Y)$, is removed, because it is a constant with respect to X . The original and compressed images have the same QDCT coefficients, and therefore, $p(Y|X)$ can be modeled by the delta function:

$$p(Y|X) = \delta\{Q[DCT(X)] - Q[DCT(Y)]\}, \quad (3.3)$$

where $DCT(\cdot)$ represents the DCT transformation, and $Q[\cdot]$ is the quantization operator. Therefore, the MAP estimation in Eq. (3.2) can be further rewritten as a constrained optimization problem,

$$\hat{X} = \arg \max_X \{\log p(X)\}, \text{ subject to } Q[DCT(X)] = Q[DCT(Y)]. \quad (3.4)$$

To construct a prior image model of X , we use the Markov Random Field (MRF), which assumes that a pixel value is conditionally dependent only on those of its neighbors [51]. The Hammersley–Clifford theorem [146] establishes that an MRF is equivalent to a Gibbs random field and the joint probability density function (pdf) of an MRF can be modeled as a Gibbs distribution:

$$p(X) = \frac{1}{Z} \exp\left\{-\sum_{c \in C} V_c(X)\right\}. \quad (3.5)$$

where c , called a clique, is a set whose elements are neighbors to each other, C is a collection of cliques, V_c is a potential function defined on pixel values in c , and Z is a normalization constant. V_c measures the consistency between X and prior information of skin images [89]-[90]. An image which is more consistent with the prior information has a smaller value of V_c . In other words, it has higher probability density. Substituting Eq. (3.5) into Eq. (3.4), we obtain

$$\hat{X} = \arg \min_X \left(\sum_{c \in C} V_c(X) \right), \text{ subject to } Q[DCT(X)] = Q[DCT(Y)]. \quad (3.6)$$

3.1.2 A Potential Function Connecting Skin Statistics and the MAP Model

Eq. (3.6) shows that the prior model of X is controlled by the potential function, V_c . An appropriate potential function is the key to the success of our MAP deblocking algorithm. Although several potential functions have been proposed to remove blocking artifacts [91]-[93], they are designed for generic images, but cannot make use of prior knowledge of target images. Here we capture statistical information of skin images from our database, and then embed this information into the potential function.

We start by studying the statistics of skin images. We select 2,500 images from our database and then compress them using the JPEG algorithm. Our deblocking algorithm is operated in the YUV space, because the JPEG compression is performed in this space. We compare the statistical characteristics of pixel values in the Y, U, and V components before and after compression. The U and V components are down-sampled before the comparison, and their deblocked results will be up-sampled to be combined with the deblocked Y component.

Let a pixel value in component m , where $m = Y, U, V$, be $p_m(x, y)$, and the pixel value of its neighbor in the k^{th} clique, $c_k = (x_k, y_k)$, where $k = 1, \dots, K$, be $p_m(x+x_k, y+y_k)$. Fig. 3.3 shows some example cliques. Their squared difference is defined as

$$d_{m_{c_k}}(x, y) = [p_m(x, y) - p_m(x + x_k, y + y_k)]^2. \quad (3.7)$$

Fig. 3.3 Illustration of cliques.

We classify d_{mc_k} into two groups: within block and across block. The within block (across block) group is generated by the cliques, whose two pixels are in the same (different) DCT block(s). For the sake of convenience, we use $d_{mc_{k,i}}$ and $d_{mc_{k,a}}$ to represent d_{mc_k} in the first group and the second group, respectively. Thus, we can use our collected skin images to estimate their probability density functions (pdf) and cumulative distribution functions (cdf).

Figs. 3.4 (a) and (b) show respectively the cdfs of $d_{Y_{c_1,i}}$ and $d_{Y_{c_1,a}}$, where $c_1 = (0, 1)$. They are estimated from the uncompressed images in our database and a pair of testing images, which is not included in the database. One testing image is an original image, while the other is the corresponding compressed image. Fig. 3.4 (a) shows that the cdf of the original testing image does not deviate much from that of the database. However, the compression process changes the distribution dramatically. Fig. 3.4 (a) shows that the compression process moves the cdf of $d_{Y_{c_1,i}}$ from the testing image towards the upper left corner. This change is due to the loss of high frequency information. The quantization parameters for high-frequency DCT coefficients are usually much larger than those for low-frequency coefficients. In other words, the number of different quantization levels for high frequency coefficients is much less than that for low frequency coefficients. In fact, many high frequency coefficients are zeros. As a result, $d_{Y_{c_1,i}}$ decreases significantly. This phenomenon is more serious in the U and V components, because of the down-sampling process, and because of the even larger quantization parameters. Consequently, many blocks in the U and V components are represented by one QDCT coefficient, which is the direct current (DC) term of the block. On the contrary, Fig. 3.4 (b) shows that the compression process moves the cdf of $d_{Y_{c_1,a}}$ from the testing image towards the right hand side. It is due to the discontinuity induced by boundaries between blocks. Fig. 3.5 illustrates the change of d_{mc_k} after compression.

Fig. 3.4 also indicates that the compression process has stronger impact on $d_{Y_{c_1,i}}$ than on $d_{Y_{c_1,a}}$. Based on this observation, we propose

$$V_{mc_k}(X) = \sum_{d_{mc_k i}} [cdf_{DB}(d_{mc_k i}) - cdf_{com}(d_{mc_k i})]^2, \quad (3.8)$$

as a potential function, where $cdf_{DB}(d_{mc_k i})$ and $cdf_{com}(d_{mc_k i})$ are the cdfs of $d_{mc_k i}$ estimated from the database and the compressed testing image, respectively. This potential function captures the statistical information of skin images through cdf_{DB} . Eq. (3.8) considers only one component and one clique format. To take into account all components and all possible cliques, we propose

$$V(X) = \sum_m \sum_{c_k \in C} \sum_{d_{mc_k i}} [cdf_{DB}(d_{mc_k i}) - cdf_{com}(d_{mc_k i})]^2, \quad (3.9)$$

as our potential function for X .

Fig. 3.4 Illustration of cdfs from different cases.

(a) $d_{Y_{c,i}}$; (b) $d_{Y_{c,i}^a}$.

Fig. 3.5 Illustration of the change of d_{mc_k} after compression.

3.1.3 Optimizing the Potential Function

We adopt the gradient descent method to minimize the potential function in Eq. (3.8). The gradient of $V_{mc_k}(X)$ with respect to the pixel value p_m is:

$$\nabla V_{mc_k}(X) = -4 \cdot [cdf_{DB}(d_{mc_k,i}) - cdf_{com}(d_{mc_k,i})] \cdot pdf_{com}(d_{mc_k,i}) \cdot [p_m(x, y) - p_m(x + x_k, y + y_k)]. \quad (3.10)$$

The gradient of all the cliques with respect to p_m is:

$$\nabla V_m(X) = \sum_{c_k \in C} \nabla V_{mc_k}(X). \quad (3.11)$$

The sign of $\nabla V_m(X)$ determines the moving direction of p_m . Let

$$e_m(t) = \sum_{c_k \in C} \sum_{d_{mc_k,i}} |cdf_{DB}(d_{mc_k,i}) - cdf_{com}(d_{mc_k,i})|, \quad (3.12)$$

represent the absolute error between the two cdfs in the current iteration, and $e_m(t-1)$ represent the error in the previous iteration. Finally, p_m is updated by the formula below:

$$p_m(t) = p_m(t-1) + \text{sign}(\nabla V_m(X)) \cdot \kappa \cdot |e_m(t) - e_m(t-1)|, \quad (3.13)$$

where κ is a positive constant.

To guarantee that compressed images and resultant images have the same QDCT coefficients, we impose the constraint, $Q[DCT(X)] = Q[DCT(Y)]$, on the optimization. In each iteration, if a new pixel satisfies this constraint, it will replace the previous one; otherwise, the update is not allowed. Fig. 3.4 (a) shows the optimized cdf of the compressed testing image. The MAP-based approach has successfully moved the cdf of the compressed testing image close to the cdf of the database. By reshaping the cdfs of different cliques, the approach recovers original information.

3.1.4 Experimental Evaluations

The performance of the MAP-based approach is evaluated on a number of skin images. Here we show the experimental results from two skin images. We use nine cliques: (0, 1), (1, 0), (1, 1), (0, 2), (2, 0), (2, 2), (0, 3), (3, 0), and (3, 3) in the following experiments. For speeding up this algorithm, we use an adaptive approach to compute the cdfs and pdfs. We compare our results with those from two other popular methods: Luo et al.'s adaptive processing method (ADPROC), which is efficient at reducing blocking artifacts in smooth regions [37], and Foi et al.'s Pointwise Shape-adaptive DCT method (SADCT) which is one of the latest deblocking techniques [94]. We do not use objective criteria such as mean square error (MSE) or peak signal-to-noise ratio (PSNR) to do evaluation, because they are not very well matched to perceived visual quality [96]. Therefore, we use subjective observation and comparison to evaluate the results of this approach.

Fig. 3.6 (a) shows a skin image with one red lesion, and Fig. 3.6 (b) is its compressed version with compression ratio of 61.38. Their corresponding Y, U, and V components are shown in the second, third and fourth rows of the first and second columns, respectively. The high compression ratio causes serious blocking artifacts. It should be noted that the U and V components have more

serious artifacts as we discuss in Section 3.1.2. From the compressed image, it is very difficult to decide the border, symmetry, and shape of the lesion. The results from ADPROC and SADCT are shown in the third and fourth columns, respectively. Both methods smooth the boundaries between blocks, but cannot recover the lost information. Our results after 10 iterations of the optimization are given in the last column. Our MAP-based approach recovers the appearance of the lesion. Furthermore, it does not introduce any smoothing effect into the image.

Fig. 3.7 (a) is another testing image with three veins. Its compression rate is 61.29 (Fig. 3.7 (b)). Subjective observation shows that the resultant images of our MAP-based approach are more close to the original images than those of ADPROC and SADCT. Take the V component as an example. ADPROC reduces the blocking artifacts to some extent, but it is still difficult to locate the veins (Fig. 3.7 (r)). SADCT removes blocks completely, but it also removes the veins (Fig. 3.7 (s)). The MAP-based approach gives us clear information about the shape, location and appearance of the three veins (Fig. 3.7 (t)). It restores the image accurately.

Because the larger quantization parameters for the U and V components make many of their neighboring pixels have the same values, the performance of the MAP-based approach in U and V is not as good as in Y. In the next section, we will present a knowledge-based approach, which will have better performance in all three components.

Fig. 3.6 Evaluation of the MAP-based approach (example 1).

(a) is an original skin image with a lesion; (f), (k) and (p) are its Y, U, and V components; (b) is a compressed version of (a) and the second column is the compressed version of the first column. The third column is the results from ADPROC [37]. The forth column is the results from SADCT [94], and the last column is the results from the proposed MAP-based approach.

Fig. 3.7 Evaluation of the MAP-based approach (example 2).

(a) is an original skin image with three veins; (f), (k) and (p) are its Y, U, and V components; (b) is a compressed version of (a) and the second column is the compressed version of the first column. The third column is the results from ADPROC [37]. The forth column is the results from SADCT [94], and the last column is the results from the proposed MAP-based approach.

3.2 A Knowledge-Based Approach

As has been mentioned, a good deblocking method should be able to recover the lost high frequency information. From this point of view, deblocking should be related to image interpretation and prediction – How can we interpret the compressed image and predict the missing information?

In the recent years, researchers have been exploring learning-based methods for image processing. Freeman et al. proposed an algorithm to learn the relationship between high resolution images and low resolution images, and applied it to single image super resolution [97]-[98]. This algorithm has also been used in shading, reflectance, and motion estimation [99]. This learning-based algorithm is especially suitable for our deblocking problem, because we target a specific type of images – human skin, which has much less variation than generic images.

In this section, we develop a knowledge-based (KB) approach to remove JPEG blocking artifacts in skin images. It extracts prior knowledge from skin images, i.e., the block relationship between compressed blocks and their original counterparts, and the neighborhood relationship among adjacent original blocks. We develop two different algorithms to make inferences based on this prior knowledge. The first one uses a Markov model and the belief propagation algorithm. This algorithm may be more accurate but requires more time. The second one is faster and is a one-pass algorithm utilizing spatial and frequency relationships simultaneously. Both algorithms require that at least one compressed block in a training set shares the same QDCT with the input compressed block. To remove this requirement, a block synthesis algorithm is proposed. We also design an indexing mechanism to deal with large datasets efficiently. Subsection 3.2.1 introduces the prior knowledge in skin images. Subsection 3.2.2 presents the KB approach including the Markov-model-based, the one-pass, and the block synthesis algorithms. We also present an indexing mechanism to improve the speed of the algorithms. Subsection 3.2.3 reports objective and subjective experimental results.

3.2.1 Prior Knowledge in Skin Images

To exploit prior knowledge in skin images, we construct a large database composed of skin images collected from different body sites, including the hand, arm, foot, leg, chest and back. We divide the database into a training set and two testing sets. We extract prior information from the training set to develop the KB approach and then use the testing sets to evaluate its function.

3.2.1.1 Database

Our database consists of two parts. The first part (Asian database) is collected in Singapore from 97 Asians with both genders and diverse ages, occupations, and body mass indices (BMI). The ethnic groups include Chinese, Malay, Indian, and Javanese. They are between 12 and 70 years old. Their occupations include students, professors, and manual workers, and their BMIs range from 18 to 40. The camera model is Nikon D70s, and the images are taken under normal daylight or fluorescent light. The second part (Caucasian database) is collected in the United States from 10 Caucasians under Institutional Review Board approval from the Los Angeles Biomedical Research Institute. The camera model is Nikon D80, and the imaging configuration such as illumination and image distance is also different from the settings in Singapore. We use images from 71 randomly selected subjects in the Asian database to form a training set. Five of them are females. Their races include Chinese, Malay, Indian and Javanese. Their ages range from 18 to 59, and BMIs range from 19 to 27. The images are taken from different body parts, including the hand, arm, thigh, leg, chest and back (see Fig. 3.8). We only use images in the Asian database to form the training set in order to demonstrate that the proposed KB approach can work well, even when input images and training images have very different characteristics (e.g. different imaging conditions and races). This experimental design is more appropriate because in legal cases, we cannot guarantee evidence images and training images will have the same characteristics. The remaining 26 subjects' images in the Asian database comprise the first testing set. The images in the Caucasian database comprise the second testing set. Each set has participants from both

genders and all age groups. These images are stored in the JPEG format with a very high quality factor of 99 and without noticeable blocking artifacts. The original images from the cameras are referred to as uncompressed images.

For the training set, because a large part of the raw images is background, we crop sub-images with 256×256 pixels containing only areas of skin. This relatively small size is chosen to reduce redundant information and improve the speed of the algorithms. Then we use the JPEG algorithm to compress the sub-images resulting in image pairs. Each pair has one original skin image and the corresponding JPEG compressed image. After cropping, the training set contains 5,662 image pairs. Fig. 3.9 illustrates a raw image (3.9 (a)) and a pair of skin sub-images obtained from it (3.9 (b) and 3.9 (f)). The KB approach is also operated in the YUV space because the JPEG compression is performed in this space. Fig. 3.9 (h) and Fig. 3.9 (i) show that the blocking artifacts in the U and V components are more significant. These artifacts are caused primarily because of the down-sampling process in the JPEG algorithm and because of larger quantization steps. By further cutting the image pairs into 8×8 pixel blocks, we have 5,797,888 block pairs in the training set. By choosing different compression quality factors we then obtain different training sets.

For the first testing set, we crop 500 skin sub-images with size of 448×512 or 512×448 pixels from the raw images of the 26 Asian subjects. We choose the larger sizes in order to cover more skin features. For the second testing set, we crop 262 skin sub-images with size of 256×256 pixels from the raw images of the 10 Caucasians. Because Caucasians have more pigmented skin marks, we use the second testing set to evaluate the performance of the KB approach on skin marks. We choose the smaller size so that skin marks can be easily observed.

Fig. 3.8 Some images in the training set.

(a) ~ (b) are hands of a female (Indian, 20 years old); (c) ~ (d) are arms of a male (Chinese, 23 years old); (e) ~ (h) are thighs of a female (Chinese, 21 years old); (i) ~ (k) are legs of a male (Chinese, 33 years old); (l) is the chest, and (m) is the back of a male (Chinese, 23 years old).

Fig. 3.9 A raw image and a pair of skin images obtained from it.

(a) is a raw image of the inner right forearm of a male (Chinese, 52 years old, BMI = 25); (b) is the skin image cropped from the red rectangle in (a); (c), (d) and (e) are the Y, U, and V components of (b); (f) is the JPEG compressed image of (b) with the compression ratio of 41.98 and quality factor of 50; (g), (h) and (i) are the Y, U, and V components of (f).

3.2.1.2 Training Blocks

The relationship between an original block and its compressed result is that they have the same QDCT coefficients. In general, only several coefficients in the upper left corner of a QDCT matrix are non-zero integers. We call them effective coefficients, and use them to form an index vector. Fig. 3.10 illustrates a QDCT coefficient matrix, the effective coefficients, and the corresponding index vector. Because the quantization is a many-to-one mapping, different original blocks can have the same QDCT coefficients and index vector. Fig. 3.11 (a) shows a compressed block in the V component. Figs. 3.11 (b) ~ (s) are some original blocks which can be compressed to Fig. 3.11 (a). All their index vectors are $[13 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$, where only the DC term is non-zero and therefore, Fig. 3.11 (a) has only one intensity value. This many-to-one

block relationship implies that only the local information inside one block is not sufficient to uniquely determine corresponding original blocks. We also consider the relationship between neighboring blocks. For each original block, we define two neighborhoods: a frequency neighborhood, which is composed by the index vectors of the 8 neighboring blocks and a spatial neighborhood, which is composed by the 36 pixels in the 8 neighboring blocks which are connected with the target block (as illustrated in Fig. 3.12).

These block and neighborhood relationships represent the prior knowledge in skin images. They are extracted from the training set to infer original blocks in evidence images. We develop two algorithms to make inference, which will be discussed in the following subsection.

Fig. 3.10 QDCT coefficient matrix, effective coefficients, and index vector.

Fig. 3.11 A compressed block and its corresponding original blocks.

(a) is a compressed block; (b)-(s) are the corresponding original blocks of (a).

Fig. 3.12 Spatial and frequency neighborhoods.

3.2.2 The Two Inference Algorithms Based on Prior Knowledge

In this subsection, we present two algorithms to make inference based on prior knowledge: one is based on a Markov model, while the other is a one-pass algorithm. We also present a block synthesis algorithm to handle the cases where the training set does not contain input compressed blocks. To speed up these algorithms, an indexing mechanism is also developed.

3.2.2.1 An Algorithm Based on a Markov Model

The first algorithm uses the traditional Markov model illustrated in Fig. 3.13, where the circles represent network nodes, and the lines indicate statistical dependencies between nodes [98]. Let compressed blocks be observation nodes, y , and original blocks with corresponding QDCT coefficients be the different states of hidden nodes, x , which we seek to estimate [97]. For this network, the probability of any choice of original blocks is proportional to the product of functions $\varphi_i(x_i, y_i)$ relating each observation to the underlying hidden states, and compatibility functions $\psi_i(x_i)$ relating the possible states of neighboring hidden nodes i.e.,

$$P(x | y) \propto \prod_{i=1}^M \varphi_i(x_i, y_i) \prod_{i=1}^M \psi_i(x_i), \quad (3.14)$$

where M is the number of blocks. The function $\varphi_i(x_i, y_i)$ can be specified by the JPEG compression algorithm:

$$\varphi_i(x_i, y_i) = \delta\{Q[DCT(x_i)] - Q[DCT(y_i)]\}, \quad (3.15)$$

where $DCT(\cdot)$ represents the DCT transform, $Q[\cdot]$ is the quantization operator, and δ is a delta function. The function $\psi_i(x_i)$ represents the spatial compatibility between neighboring original blocks. Let p_{ij}^s be pixels in the spatial neighborhood (see Fig. 3.12) of block i which overlap with the neighboring block j ($j \in N(i)$, where $N(i)$ is the 8 neighbors of block i). Let p_{ji} be pixels in the neighboring block j which overlap with the spatial neighborhood of block i . The compatibility function is defined as:

$$\psi_i(x_i) = \exp \left(- \frac{\left[\sum_{j \in N(i)} |p_{ij}^s - p_{ji}| \right]^2}{2\sigma^2} \right). \quad (3.16)$$

The optimal original blocks are the collection that maximizes the probability of the Markov network. Finding the global optimal solution is computationally intractable, so we use belief propagation to obtain a suboptimal solution. The σ in the Eq. (3.16) can be removed from the optimization formulation because φ_i can be considered as constraints and all ψ_i s are multiplied together in Eq. (3.14).

Fig. 3.14 illustrates the performance of the algorithm based on the Markov model. Fig. 3.14 (a) is the V component of an original skin sub-image where one can observe a patterned distribution of hemoglobin in the veins of the skin. However, in the compressed image, the pattern is destroyed (see Fig. 3.14 (b)). We use belief propagation to infer original blocks. At the beginning (zero iteration), a resultant image consists of the original blocks in the training set which have the same index vectors as the compressed blocks. Spatial smoothness is not considered in this initialization. Because there may be more than 10,000 candidate blocks corresponding to one compressed block, we use frequency neighborhoods (see Fig. 3.12) as a constraint to reduce the searching range. However, the result is still noisy and has many repeated blocks (see Fig. 3.14 (c)). A high compression ratio of 86.92 produces large areas with the same intensity value (see Fig. 3.14 (b)) leading to this effect. In the following iterations, more than half of these estimated blocks are subsequently corrected by belief propagation based on spatial

smoothness constraints. It quickly converges after four to five iterations and significantly improves the quality of the result (see Figs. 3.14 (d) ~ (h)). The vein pattern in the skin is successfully recovered, and the blocking artifacts are completely removed.

Most of other Markov-model-based deblocking algorithms formulate the problem as a minimization based on the Gibbs distribution [17][144], but they do not use the relationship between original and compressed blocks. Consequently, they cannot guarantee that compressed images and their resultant images have the same QDCT coefficients. The proposed KB approach uses the functions φ_i to represent the JPEG compression constraints, which ensure that deblocked images and corresponding compressed images have the same QDCT coefficients. In addition, it utilizes information from the training database extensively.

Fig. 3.13 A Markov network model.

Fig. 3.14 The performance of the Markov-model-based algorithm.

(a) is the V component of an original skin sub-image; (b) is a compressed version of (a) with the compression ratio of 86.92 and quality factor of 50; (c) ~ (h) are respectively the results after 0, 1, 3, 6, 9 and 18 iterations.

3.2.2.2 A One-Pass Algorithm

Although the belief propagation algorithm converges quickly, it still needs several iterations to produce a satisfactory result. We also develop a one-pass algorithm that infers original blocks based on combined spatial and frequency information. Assume that a compressed image is processed block by block in a raster-scan order – from left to right and from top to bottom. Thus, for a compressed block target, three upper and one left neighboring blocks have already been processed (illustrated in Fig. 3.15). The spatial information in the processed blocks (blocks 1-4 in Fig. 3.15) and the frequency information in the to-be-processed blocks (blocks 5-8 in Fig. 3.15) are used as smooth constraints to search the best original block in the training set. The pixels in the processed blocks that are connected to the target block represent the spatial neighborhood. The index vectors of the to-be-processed blocks represent the frequency neighborhood. Together these form the “hybrid neighborhood” of the target block. Thus, each original block in the training set has a hybrid neighborhood determined from the source image (as depicted in Fig. 3.16). In this way, each record in the training set contains an original block, its hybrid neighborhood, and its compressed block.

In the testing stage, for a target block w , we use its index vector to find a group of candidate original blocks, $G(w) = \{z \mid QDCT[z] = QDCT[w]\}$. Then we search an optimal candidate according to the hybrid neighborhood of the target block. This search is carried out in two steps. Firstly, we use the frequency neighborhood to narrow down the group i.e.,

$$G'(w) = \{u \in G(w) \mid F_w = F_u\}, \quad (3.17)$$

where F_w (or F_u) represents the frequency neighborhood of w (or u). If it is an empty set, we search 20 candidate blocks from $G(w)$ whose frequency neighborhoods are the nearest to F_w to form $G'(w)$. Then we use the spatial neighborhood to find the optimal original block t^* i.e.,

$$t^* = \arg \min_{t \in G'(w)} d(S_t, S_w), \quad (3.18)$$

where S_w (or S_t) represents the spatial neighborhood of w (or t), and d represents l^l -norm.

To compare the one-pass algorithm with the Markov-model-based algorithm, we use the same testing image (see Fig. 3.14 (b)) to evaluate the one-pass algorithm (see Fig. 3.17). Theoretically, the result from the Markov-model-based algorithm should be more accurate because it optimizes the probability in Eq. (3.14). However, Fig. 3.17 (c) and Fig. 3.17 (d) demonstrate that the difference between the two results is difficult to discern with normal human vision. The one-pass algorithm has similar performance but is 12 times faster than belief propagation. Thus, the one-pass algorithm is used for the following evaluations. Fig. 3.18 illustrates some blocks from Fig. 3.17. Original blocks in the first row correspond to the compressed blocks in the second row. The high compression ratio causes great information loss. The third and fourth rows are the resultant blocks from the Markov-model-based and one-pass algorithms, respectively. For the one-pass algorithm, the hybrid neighborhood is sufficient to recover the lost information.

Fig. 3.15 Illustration of hybrid neighborhood in the testing process.

Fig. 3.16 Illustration of hybrid neighborhood in training set construction.

Fig. 3.17 The performance of the one-pass algorithm.

(a) is the V component of an original skin sub-image; (b) is a compressed version of (a) with the compression ratio of 86.92 and quality factor of 50; (c) is the result from the Markov-model-based algorithm; (d) is the result from the one-pass algorithm.

Fig. 3.18 The corresponding blocks from Fig. 3.17.

In the first row are the original blocks. In the second row are the corresponding compressed blocks. In the third row are the resultant blocks obtained by the Markov-model-based algorithm. In the fourth row are the resultant blocks obtained by the one-pass algorithm.

3.2.2.3 A Block Synthesis Algorithm

Both Markov-model-based and one-pass algorithms require that at least one compressed block in the training set shares the same QDCT coefficients with the input compressed block. The variation of compressed skin blocks and the completeness of our training set are important for the KB approach. In this subsection, we first estimate the variation of compressed skin blocks and then present our block synthesis algorithm to remove this requirement.

To estimate the variation of compressed skin blocks and the completeness of our training set, we calculate the ranges of QDCT coefficients in our database and assume that all QDCT coefficients are statistically independent. Note that QDCT coefficients are integers. Using the ratios of the estimated total number of different compressed blocks to the total number of compressed blocks in our database as an index, we note that the variation of compressed blocks is very limited in U and V components, but it is very large in Y component. Under the quality factor of 25, the ratios in U and V components are 0.009% and 0.02%, respectively. Under the quality factor of 50, the ratios in U and V components are 3.21% and 5.11%, respectively. However, for Y component, the ratios under either quality factor are larger than 100%. When the quality factor is higher than 50, the loss of information is not significant, so there is no need for blocking artifact removal.

To handle the cases where input compressed blocks cannot be found in our training set, two or three training blocks are used to synthesize a new block with the same index vector as the input block. Fig. 3.19 illustrates this block synthesis algorithm. The search of low (high) frequency coefficients is from the rightmost (leftmost) element in the index vector to the left (right) until the quantized coefficients do not match. As a result, the number of overlapping coefficients is maximized to achieve the best compatibility with two blocks. If more than one block has the same number of matched coefficients, we use the hybrid neighborhood as a constraint to select the best one. After that we combine their DCT matrices and take inverse DCT transform to obtain the synthetic block. For the overlapping part, we use the DCT coefficients of the low frequency block. If two source blocks are not sufficient to cover the index vector, the third one would be used for the remaining coefficients. An iterative process is used to identify this block. In each iteration, the search is performed in three situations: one coefficient is removed from the rightmost part, the leftmost part, and each part. If no candidate is found, the coefficients are continuously removed until one block is found. In our experiment, more than 99.9% of the compressed blocks could be replaced with a source block or a synthetic block from two or three source blocks. With the increase in the database size, we can further reduce the percentage of non-replaced blocks. Note that we use 1D index vector instead of 2D plane to find coefficients because the former is easier to handle and offers higher searching speed. Fig. 3.20 shows some results of this algorithm. The first row contains original blocks and the second row contains the corresponding compressed blocks. The third row contains the synthetic blocks obtained from the source blocks shown in the fourth and fifth rows. These results demonstrate that our algorithm has the capability to synthesize new blocks whose compressed blocks are not in the training set. Therefore, the KB approach can recover skin patterns “unknown” to the database. The synthetic blocks in Fig. 3.20 are more similar to their high frequency blocks than their low frequency blocks because (1) human vision is more sensitive to high frequency information such as edges, and (2) the number of effective coefficients in the synthetic blocks from the low frequency blocks is less than that from the high frequency blocks. Note that the variation of low frequency

coefficients is larger. Whether the synthetic coefficients are from the high frequency blocks or the low frequency blocks is not an issue. Most importantly, the synthetic blocks and the input index vectors have same QDCT coefficients and the compatibility between different source blocks have been optimized.

Fig. 3.19 Illustration of the block synthesis algorithm.

Fig. 3.20 The block synthesis results.

In the first row are original blocks of Y component. In the second row are the corresponding compressed blocks with the compression ratio of 86.92 and quality factor of 50. In the third row are the synthetic blocks. In the fourth and fifth rows are source blocks from the training set.

3.2.2.4 An Indexing Mechanism to Speed up the Algorithms

To make inference based on prior knowledge, a large training set is essential. Our training set contains more than 5 million block-pairs. It would be extremely time-consuming (about 20 hours) to search the entire training set exhaustively for each testing block. To speed up the searching, we propose an indexing mechanism using a multi-dimensional structure to store the information of each original block in the training set. The number of dimensions corresponds to the length of index vectors. Each entry represents one index vector and stores the information of the corresponding original blocks. This includes their positions in the source images and their hybrid neighborhoods. For a testing block, its index vector immediately leads us to the corresponding entry in the structure. This mechanism is much faster than matching it with all the blocks in the training set.

This indexing mechanism cannot be used directly in the Y component, because its quantization steps are much smaller than those in U and V components. As a result, the number of different index vectors is too large to be stored in a multi-dimensional structure due to memory constraints. We preprocess the Y components of the original images by normalizing their intensity values to zero mean and unit variance. Then we recalculate their index vectors whose varying range decreases significantly. This normalization step clusters the index vectors in Y components into a limited number of groups allowing storage in a multi-dimensional structure. In each entry of the subsequent structure, the un-normalized index vectors are added as extra information to distinguish individual blocks. This indexing mechanism makes it possible to handle such a large training set quickly and efficiently. The one-pass algorithm with the indexing mechanism are implemented using MATLAB on a PC embedded with an Intel® Core™2 Quad processor (3.0 GHz). Only one core is used in our implementation. Removing blocking artifacts in the Y, U, and V components of a 448×512 image, it uses respectively 20, 5 and 5 minutes. Generally speaking, computational speed is not vital for forensic analysis. In the mentioned legal case, Kong and Craft had several weeks to process the evidence images.

3.2.3 Experimental Evaluations

To evaluate the KB approach, we perform extensive experiments and compare it with four other popular deblocking methods. The other methods are Sun et al.'s maximum a posteriori method based on a Field of Experts prior (FOE) which achieves higher PSNR gain [51], Foi et al.'s Pointwise Shape-adaptive DCT method (SADCT) which is one of the latest deblocking techniques [94], Luo et al.'s adaptive processing method (ADPROC) which is efficient at reducing blocking artifacts in smooth regions [37], and Chou et al.'s nonlinear filtering method (NLF) which is fast and robust for different images and quantization strategies [95]. These methods and the KB approach are analyzed using the two testing sets. The first testing set with 500 images is compressed with the JPEG quality factor of 50, and the average compression ratio is 72.55. The second testing set with 262 images is compressed with the JPEG quality factor of 25, and the average compression ratio is 126.93. As a result of this high compression ratio in the second set, most skin features are destroyed. As was mentioned in Section 3.2, the one-pass algorithm has similar performance to the Markov-model-based algorithm, but is much faster. Thus, the one-pass algorithm is used for the following evaluations. The U and V components are down-sampled before deblocking, and their results are up-sampled and combined with the deblocked Y component. The same training set from the Asian database is used to generate all the experimental results.

In Figs. 3.21 ~ 3.26, we compare the original images (a), the compressed images (b), the results from FOE (c) [51], SADCT (d) [94], ADPROC (e) [37], and NLF (f) [95] methods. Additionally, the results from the KB approach (g) are shown. These six sets of skin images are shown for visual comparison. Figs. 3.21 ~ 3.23 are respectively Y, U, and V components, and Figs. 3.24 ~ 3.26 are color images. The red circles in Figs. 3.24 ~ 3.26 denote skin marks identified by dermatology experts in the uncompressed images. The compression ratios of Figs. 3.21 ~ 3.26 are respectively 76.12, 69.50, 50.84, 115.32, 126.97, and 121.16. These figures show that FOE and SADCT methods have strong smoothing effect, which occasionally result in unrecognizable or absent vein patterns and skin marks; ADPROC and NLF methods have less

smoothing effect, but they do not change the compressed images significantly; the KB approach removes the blocking artifacts and successfully recovers lost skin information, including the vein patterns and pigmented skin marks.

To quantify these visual comparisons, we can either use image quality indices or carry out a subjective evaluation. It is well-known that the traditional point-wise comparison measures such as mean square error (MSE) or peak signal-to-noise ratio (PSNR) are not very well matched to perceived visual quality [47]. Other image quality indices, such as SSIM [96] attempt to mimic the extremely complicated human vision system (HSV). It is very difficult, if not impossible, to perfectly model HSV [99]. To avoid modeling defects, we carry out one subjective evaluation and three objective evaluations on the resultant images.

Fig. 3.21 Evaluation of the knowledge-based approach in the Y component.

(a) is the Y component of an original skin image; (b) is a compressed version of (a) with the compression ratio of 76.12 and quality factor of 50; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) is the result from the KB approach.

Fig. 3.22 Evaluation of the knowledge-based approach in the U component.

(a) is the U component of an original skin image; (b) is a compressed version of (a) with the compression ratio of 69.50 and quality factor of 50; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) is the result from the KB approach.

Fig. 3.23 Evaluation of the knowledge-based approach in the V component.

(a) is the V component of an original skin image; (b) is a compressed version of (a) with the compression ratio of 81.61 and quality factor of 50; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) is the result from the KB approach.

Fig. 3.24 Evaluation of the knowledge-based approach in color images (example 1).

(a) is an original skin sub-image. The red circle denotes a pigmented skin mark identified by dermatology experts in the uncompressed image. (b) is a compressed version of (a) with the compression ratio of 115.32 and quality factor of 25; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) The result from the KB approach.

Fig. 3.25 Evaluation of the knowledge-based approach in color images (example 2).

(a) is an original skin sub-image. The red circle denotes a pigmented skin mark identified by dermatology experts in the uncompressed image; (b) is a compressed version of (a) with the compression ratio of 126.97 and quality factor of 25; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) The result from the KB approach.

Fig. 3.26 Evaluation of the knowledge-based approach in color images (example 3).

(a) is an original skin sub-image. The red circle denotes a pigmented skin mark identified by dermatology experts in the uncompressed image. (b) is a compressed version of (a) with the compression ratio of 121.16 and quality factor of 25; (c) ~ (f) are the results from FOE [51], SADCT [94], ADPROC [37], and NLF [95] methods respectively; (g) is the result from the KB approach.

3.2.3.1 Subjective Evaluation

Twenty-three observers participate in this experiment. Three of them have dermatological knowledge (one is a board-certified dermatologist and the other two are medical students studying dermatology); twelve of them are familiar with image processing; and the rest of them have computer science background. All the testing images are prepared by one person. It should be emphasized that there are no so-called skin mark examiners, like fingerprint examiners. Law enforcement agents, including the U.S. Department of Justice, recruit board-certified dermatologists to recognize skin marks in legal cases. The experiment is carried out in Y, U, V components and color images. In each case, 25 groups of images are presented. For Y, U, and V components, 20 groups are from the first testing set, and 5 groups are from the second testing set, while for color images, 5 groups are from the first testing set, and 20 groups are from the second testing set. Totally 100 groups are evaluated. In each group, we present an original uncompressed image (as reference), the corresponding compressed image, and the 5 resultant images (4 from the other methods and 1 from the KB approach) to the unbiased observers. We ask the observers to

rate the images using a 10-point scale. For the Y, U, and V components, the observers are required to rate the resultant images according to their similarity with the references. Higher grades (i.e., 10) represent more similarity between original uncompressed images and resultant images. For color images, skin marks in reference images are highlighted. An example is given in Fig. 3.24. Observers are asked to compare skin marks. As with the Y, U, and V components, the same grading scheme is employed. Ten observers participate in the Y, U, and V evaluation, and the other ten participate in the color evaluation. The three dermatological professionals participate in all evaluations. The average scores from the dermatological professionals and from other participants are illustrated in Fig. 3.27 and Fig. 3.28, respectively. They show clearly that the KB approach provides the greatest visual quality improvement. These results pinpoint clearly that the KB approach is effective not only for generic skin images, but also for skin marks. They further confirm our visual comparison in Figs. 3.21 ~ 3.26.

Fig. 3.27 Subjective evaluation results from the three dermatological professionals.

Higher scores represent more similarity to the original image. (a) Y component; (b) U component; (c) V component; (d) color image.

Fig. 3.28 Subjective evaluation results from the other participants.

Higher scores represent more similarity to the original image. (a) Y component; (b) U component; (c) V component; (d) color image.

The application of the KB approach could improve confidence in diagnosis from digital images in real cases. The KB approach could reduce unnecessary debates in real cases because it significantly improves the visual quality of skin marks and other features for identification. The visual quality of images is important for accurate identification. Dermatology training is mostly based on “in person” physical exams and even enhanced visualization with dermatoscopes. Thus, digital diagnostic training is best performed when using uncompressed images. Although there is no absolute threshold for useable resolution, previous research demonstrates that image resolution is most important when trying to accurately identify small pigmented lesions such as those used in the test cases here [106].

3.2.3.2 Objective Evaluation Based on Phase Information

In the previous experiment, we evaluate the KB approach in terms of the visual quality. In this experiment, we use phase information to obtain an objective evaluation. Phase is extremely important information for biometric recognition. The famous IrisCode is based on coarse representation of Gabor phases [101]. In addition to IrisCode, phases are also employed for palmprint and face recognition [102][103]. Behar et al. pinpoint that images can be reconstructed from localized phase [104]. Phase also has strong relationship with edges. Thus, we evaluate phases in the resultant images. To extract pure phase information in images I and J , where I is an original image, and J is a resultant image from one of the methods, we take the 2D Fourier Transform of I and J , and remove their magnitude information in the Fourier domain. Then, we take the inverse 2D Fourier Transform. Once we obtain the phases of I and J , the function

$$S(I, J) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \sum_{i=1}^B \sum_{j=1}^B |I_p(xB+i, yB+j)| \cdot |J_p(xB+i, yB+j)|, \quad (3.19)$$

where I_p (or J_p) is the phase of I (or J), B is a block size, and N and M are numbers of blocks in the column and row, respectively, is used to evaluate the similarity between the magnitude of the pure phase information in the two images. We divide the phase images into different blocks, because JPEG is a block-based algorithm. The block size varies from 8×8 pixels to 256×256 pixels. The higher value of S indicates more similarity between phases of I and J . We also evaluate the similarity of the entire image. The average measures for the Y, U, and V components of the testing images are illustrated in Fig. 3.29. The standard deviations of the measures are illustrated in Fig. 3.30. The block synthesis algorithm is deployed for the Y component. The average percentage of synthetic blocks per image is 26.3% for the first testing set and 3.6% for the second testing set. In all settings, the results from the KB approach are clearly superior. In general, the similarity measures of the Y component are higher than those of the U and V components, due to the lower compression ratio in Y. In addition, for most of the results, the distributions of the results from the KB approach have the smallest standard deviations. This implies that the KB

approach performs more consistently on varying images. This objective evaluation demonstrates that the KB approach can successfully recover the phase information.

Fig. 3.29 The average of similarity measures.

(a), (b) and (c) are the results from the Y, U, and V components of the 1st testing set; (d), (e) and (f) are the results from the Y, U, and V components of the 2nd testing set.

Fig. 3.30 The standard deviation of similarity measures.

(a), (b) and (c) are the results from the Y, U, and V components of the 1st testing set; (d), (e) and (f) are the results from the Y, U, and V components of the 2nd testing set.

3.2.3.3 Objective Evaluation Based on Automatic Skin Mark Detection

In this subsection, an evaluation of recoverability of the KB approach for skin marks is reported. Note that skin marks are important biometric traits for forensic analysis. For this evaluation, an automatic skin mark detection algorithm is designed. Given one training image (original image) and one testing image (original image or deblocked image) from the same body part of the same person, an image patch with a skin mark is cropped from the training image. The size of the image patch is 19×19 pixels. We use the 2D Fourier Transform to remove the magnitude information and use the pure phase information for detection. We also employ the same approach to remove the magnitude information in the testing image. The cosine measure is utilized to obtain the similarity between the skin mark patch from the training image and all possible 19×19 pixel image patches in the testing image. In an ideal case, the corresponding skin mark in the testing image should have the highest similarity. Because some of the testing images have more than one skin mark, if the corresponding skin mark in the testing image has the highest or the second highest similarity, it is regarded as a correct detection. For each deblocking algorithm, 150 image pairs (one is an original image and the other is a resultant image) from the Caucasian database are examined. The image size is 256×256 pixels. Table 3.1 lists the detection rates. It is clear that the KB approach performs much better than other algorithms and its detection rate is close to that from matching two original images of the same person. These results indicate that the KB approach can effectively recover skin marks from compressed images to overcome quantization noise.

Table 3.1 Automatic skin mark detection rate from different algorithms.

Original	Compressed	FOE [51]	SADCT [94]	ADPROC [37]	NLF [95]	KB
95.3%	49.3%	82.5%	80.2%	54.6%	62.1%	90.1%

3.2.3.4 Objective Evaluation Based on a Blockiness Index

To evaluate discontinuity between blocks, we calculate the blockiness index [105] of the original, compressed, and resultant images. The range of this index is between zero and one. A higher value represents more serious discontinuity between blocks. The indices for the Y, U, and V components of the two testing sets are listed in Table 3.2. Note that the block synthesis algorithm is used for the Y component. The KB approach is always the best compared with other methods and its resultant images are comparable to the original images.

Table 3.2 Means of blockiness index from different algorithms.

	The first testing set			The second testing set		
	Y	U	V	Y	U	V
Original	0.1311	0.1893	0.1472	0.2208	0.1898	0.1997
Compressed	0.3402	0.9565	0.9443	0.6861	0.9485	0.9584
FOE [51]	0.3607	0.5253	0.4315	0.4889	0.4190	0.4772
SADCT [94]	0.2274	0.5669	0.5078	0.4906	0.5100	0.6136
ADPROC [37]	0.2501	0.7353	0.6786	0.5932	0.7037	0.7748
NLF [95]	0.3092	0.6716	0.5983	0.3739	0.5957	0.8377
KB	0.1615	0.1914	0.1336	0.3676	0.1682	0.2380

3.3 Summary

In this chapter, we have proposed two approaches to remove JPEG blocking artifacts in skin images for forensic analysis. One is a MAP-based approach and the other is a knowledge-based approach. The first approach formulates skin image deblocking as an estimation problem. It embeds statistical information of skin images into a maximum a posteriori (MAP) model to perform the estimation. The resultant images and corresponding original skin images share the same statistical information. A potential function that connects the MAP model and skin statistics was proposed and an algorithm to optimize the potential function was presented. Visual evaluations show that the MAP-based approach recovers the skin information and does not introduce smoothing effect into the image. However, its performance of in U and V components is not as good as in Y component.

The second approach extracts prior knowledge of skin images from a training set, and uses it to infer original blocks in compressed evidence images. A Markov-model-based algorithm and a one-pass algorithm were developed to implement the inference. A block synthesis algorithm was designed to handle the cases where input compressed blocks are not contained in the training set. An indexing mechanism was employed to speed up these algorithms. The KB approach was evaluated on skin images from people with different races, genders and ages, taken under different imaging configurations, and with different compression ratios. Evaluations for visual quality, phase information, automatic skin mark detection, and blockiness index demonstrate that the KB approach outperforms other deblocking methods. It successfully removes blocking artifacts and recovers lost skin mark information. Using the KB approach, the visual quality of biometric features such as pigmented skin marks can be significantly improved. The performance of the KB approach is better than the MAP approach.

Both of the approaches extensively use prior knowledge from our database to solve the problem. It should be emphasized that our approaches guarantee that the resultant and compressed images have the same QDCT coefficients.

Chapter 4

An RGB-NIR Mapping Approach to Uncover Vein Patterns

It is an extremely challenging task to uncover vein patterns from color images because the veins lie underneath the skin surface and are almost invisible to the naked eye. Currently there is no algorithm or software package especially designed for this purpose. Chapter 2 reviews some techniques related to uncovering vein patterns in color images. However, these techniques do not work well under many circumstances, because they neither target our problem, nor utilize any prior knowledge.

In this chapter, we propose an RGB-NIR mapping approach to uncover vein patterns. It extracts information from a pair of synchronized color and NIR images, and uses a neural network (NN) to map RGB values to NIR intensities. An automatic intensity adjustment scheme is designed for illumination compensation. In addition, an NN weight adjustment scheme is proposed to improve the performance.

The RGB-NIR mapping approach is composed of three parts: an RGB-NIR mapping network, an automatic intensity adjustment scheme for illumination compensation, and an automatic neural network weight adjustment scheme to improve the uncovering performance. A pair of synchronized RGB image and NIR image is used to train a neural network for mapping RGB images to NIR images so that veins become visible. However, the trained network is sensitive to illumination variation and noise in G channel. The automatic intensity adjustment scheme and the weight adjustment scheme are designed to address these problems. They modify testing RGB images before inputting them into the trained neural network. In this chapter, section 4.1 presents the RGB-NIR mapping network. Section 4.2 describes the automatic intensity

adjustment scheme. Section 4.3 presents the automatic neural network weight adjustment scheme. Section 4.4 offers experimental results and discussion. Section 4.5 provides some summary remarks.

4.1 An RGB-NIR Mapping Network

We use synchronized color and NIR images collected by the 2-CCD multi-spectral camera as training data to construct the mapping network. The camera simultaneously measures visible and NIR light spectra through a single lens using two channels. The first channel captures visible light using a Bayer mosaic color imager, while the second channel captures NIR light using a monochrome imager. The color and NIR images are perfectly registered and synchronized. We collect data from the forearms of ten Asian males, and select the pair with the best image quality shown in Fig. 4.1 for network construction. Then we segment the skin regions and cut the central parts with both veins and general skin pixels as training data. We use the RGB values as inputs and the corresponding NIR intensity as target outputs to train a three-layered feed-forward neural network, as shown in Fig. 4.2 (a). The transfer functions in the hidden and output layers are respectively tan-sigmoid and linear functions. To avoid complex network structure and the much training time, we put 5 neurons in the hidden layer. The network is trained with the Levenberg-Marquardt back-propagation algorithm [114].

Fig. 4.1 A pair of synchronized images of a male's left forearm obtained by the 2-CCD multi-spectral camera.

(a) is the RGB image and (b) is its corresponding NIR image.

Fig. 4.2 (a) A schematic diagram of the mapping network to uncover vein patterns and (b) a schematic diagram of the generic NN mapping model.

4.2 Automatic Adjustment of Image Intensity

The illumination condition can significantly influence the skin color in images. Therefore, the performance of the RGB-based mapping network deteriorates if the imaging condition in testing images is very different from that of the training image. To compensate it, we propose a scheme which automatically adjusts image intensity to improve the uncovering performance.

An intensity adjustment ratio a is introduced in the mapping i.e.,

$$N = \varphi(C/a), \quad (4.1)$$

where C is a color skin image, φ represents the neural network function mapping RGB values to NIR intensity, and N is the uncovered image. The average local variance of the uncovered image is used as an objective function to determine the optimal adjustment ratio a^* :

$$a^* = \arg \max_{a \in A} \left\{ \frac{1}{M} \sum_{i=1}^M \text{var}_{(x,y) \in S_{b_i}} [N_{b_i}(x,y)] \right\}, \quad (4.2)$$

where A is the set of different intensity adjustment ratios, N_{b_i} is the i^{th} 5×5 pixel block in N , M is the total number of blocks, (x, y) is the position of a pixel, and S_{b_i} is the set of skin pixels in N_{b_i} . Note that only skin pixels are considered in the calculation of local variance. The adjusted RGB values are input into the trained neural network and the output is the corresponding NIR intensity, as shown in Fig. 4.2 (b).

In the following parts of this chapter, the intensity adjustment scheme is applied with the mapping network. We call their combination a “generic NN mapping model”.

4.3 Automatic Adjustment of Neural Network Weight

Veins look green in a color image because the ratio of G value to the sum of RGB values in vein pixels is higher than that in generic skin pixels. Therefore, the mapping approach is more sensitive to the G value. In this section, we propose a new scheme which automatically adjusts the NN input-layer weights to further improve the robustness of the generic NN mapping model.

Before presenting the new scheme, we briefly analyze the quality of vein patterns which significantly affects the matching performance. Figs. 4.3 and 4.4 illustrate two examples of extracted vein patterns. Visual comparison of the intensity images obtained by the generic NN mapping model (Figs. 4.3 (c) and 4.4 (c)) and the corresponding enhanced images (Figs. 4.3 (d) and 4.4 (d)) shows that the image quality of the first example is better. Consequently, the extracted vein patterns of the first example (Fig. 4.3 (e)) are longer. It is expected that it will provide more accurate matching results. On the contrary, the extracted vein patterns of the second example (Fig. 4.4 (e)) are shorter. Note that the intensity adjustment scheme is included in the generic NN mapping model. We will design a new scheme based on this property for enhancing vein uncovering performance.

Fig. 4.3 Illustration of extracted vein patterns (example 1).

(a) is a color skin image; (b) is its corresponding NIR image; (c) is the intensity image obtained by the mapping model; (d) is the enhanced image; (e) is the vein pattern image skeletonized from (d); (f) shows the sampled points from the vein patterns in (e).

Fig. 4.4 Illustration of extracted vein patterns (example 2).

(a) is a color skin image; (b) is its corresponding NIR image; (c) is the intensity image obtained by the mapping model; (d) is the enhanced image; (e) is the vein pattern image skeletonized from (d); (f) shows the sampled points from the vein patterns in (e).

The mapping function is a feedforward neural network, which can be written as:

$$y = \mathbf{W}^h \cdot f^h(\mathbf{W}^i \cdot \mathbf{x} + \mathbf{b}^i) + \mathbf{b}^h, \quad (4.3)$$

where $\mathbf{x} = [R, G, B]^T$ is the input vector of color values, y is the output intensity value, f^h is the transfer function of the hidden layer, $\mathbf{W}^i$ and $\mathbf{W}^h$ are respectively the weight matrices of the input and hidden layers, and $\mathbf{b}^i$ and $\mathbf{b}^h$ are respectively the bias vectors of the input and hidden layers. $\mathbf{W}^i$ is a matrix with three columns corresponding to the input R , G , and B values:

$$\mathbf{W}^i = [\mathbf{w}_R^i \quad \mathbf{w}_G^i \quad \mathbf{w}_B^i]. \quad (4.4)$$

To reduce the sensitivity of the approach to the green channel, we introduce a parameter, b , to adjust the weights corresponding to the input G value:

$$\tilde{\mathbf{W}}^i = [\mathbf{w}_R^i \quad \mathbf{w}_G^i \cdot b \quad \mathbf{w}_B^i]. \quad (4.5)$$

The average size of vein patterns in the skeletonized image is used as an objective function to determine the optimal adjustment ratio b^* :

$$b^* = \arg \max_{b \in B} \left\{ \frac{1}{M} \sum_{i=1}^M S(v_i) \right\}, \quad (4.6)$$

where B is a set of different adjustment values, v_i is the i^{th} vein segment, which is defined as an 8-connected object in the skeletonized image (see. Fig. 4.3 (e)). M is the total number of vein

segments in the image, and S is a function calculating the size of the vein segments, which is defined as:

$$S(v_i) = \sqrt{(r_{\max} - r_{\min})^2 + (c_{\max} - c_{\min})^2}, \quad (4.7)$$

where $r_{\max}$ ($r_{\min}$) and $c_{\max}$ ($c_{\min}$) are respectively the maximum (minimum) row and column coordinates in the vein segment v_i . We use the vein pattern image in Fig. 4.3 (e) as an example to illustrate vein segments. Figs. 4.5 (b) ~ (g) illustrate the vein segments in Fig. 4.5 (a) (the same as Fig. 4.3 (e)). The average size of vein patterns in Fig. 4.3 (e) is 175.3, while that in Fig. 4.4 (e) is only 56.8, which verifies that the average size is monotonically related to the quality of vein patterns.

Figs. 4.6 ~ 4.8 show the results from the weight adjustment scheme. Compared with the results from the generic NN mapping model, after adjusting the NN weights, the enhanced vein images and its skeletonized results contain longer and clearer vein patterns. The average sizes of the vein patterns in Figs. 4.6 (e) and (h) are 78.3 and 98.6, respectively; those of the vein patterns in Figs. 4.7 (e) and (h) are 72.8 and 118.8, respectively; and those of the vein patterns in Figs. 4.8 (e) and (h) are 67.6 and 69.5, respectively. This demonstrates that using average size of vein patterns as an objective function, the automatic NN weight adjustment scheme efficiently improves the quality of vein pattern images. We call the combination of the generic NN mapping model and the automatic NN weight adjustment an “enhanced NN mapping model”.

Fig. 4.5 Illustration of vein segments.

(a) is a skeletonized image of vein patterns which contains 6 vein segments; (b) ~ (g) are the vein segments contained in (a).

Fig. 4.6 The NN weight adjustment result (example 1).

(a) is a color skin image; (b) is its corresponding NIR image; (c) is the intensity image obtained by the generic NN mapping model; (f) is the result of the enhanced NN mapping model; (d) and (g) are respectively the enhanced images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images skeletonized from (d) and (g).

Fig. 4.7 The NN weight adjustment result (example 2).

(a) is a color skin image; (b) is its corresponding NIR image; (c) is the intensity image obtained by the generic NN mapping model; (f) is the result of the enhanced NN mapping model; (d) and (g) are respectively the enhanced images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images skeletonized from (d) and (g).

Fig. 4.8 The NN weight adjustment result (example 3).

(a) is a color skin image; (b) is its corresponding NIR image; (c) is the intensity image obtained by the generic NN mapping model; (f) is the result of the enhanced NN mapping model; (d) and (g) are respectively the enhanced images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images skeletonized from (d) and (g).

4.4 Experimental Evaluations

In this section, we evaluate the RGB-NIR mapping approach. The enhanced NN mapping model is applied to the color images. Firstly, we use visual observation to evaluate the uncovering performance. Then, we evaluate the deblocking approach that we presented in chapter 3 and the uncovering approach simultaneously. Finally, we use the automatic matching algorithm to perform vein identification.

4.4.1 Evaluation by Visual Observation

We use the corresponding color images collected by the DSLR cameras and NIR (800 nm) images collected by the Hitachi and JAI cameras in our database as testing images. Because of the penetration capability of NIR, NIR images have better quality than the distribution maps. Thus, they are considered as ground truth of vein patterns. Figs. 4.9 ~ 4.11 show the uncovered images obtained from the RGB-NIR mapping approach. Visual comparison demonstrates that the uncovered results contain clear vein patterns, which are very close to those from the NIR images. To simulate images collected from crime scenes, we use images taken from different viewing angles and poses to test the approach. Figs. 4.12 ~ 4.15 show the results. Because we are not able to take the exactly corresponding NIR images of the color images, here we use the NIR images of the corresponding body position as ground truth. For these images, the problems of uneven illumination and incident angle are more serious than the relatively well-posed images. The resultant images show that the RGB-NIR mapping approach still can uncover vein patterns.

Fig. 4.9 Visual evaluation of the RGB-NIR mapping approach (example 1).

(a) is a color image of the inner left arm of a male; (b) is its skin image; (c) is its NIR image; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.10 Visual evaluation of the RGB-NIR mapping approach (example 2).

(a) is a color image of the inner left thigh of a male; (b) is its skin image; (c) is its NIR image; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.11 Visual evaluation of the RGB-NIR mapping approach (example 3).

(a) is a color image of the front thighs of a male; (b) is its skin image; (c) is its NIR image; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.12 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 1).

(a) is a color image of the left arm of a male; (b) and (c) are the related NIR images; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.13 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 2).

(a) is a color image of the chest of a male; (b) and (c) are the related NIR images; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.14 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 3).

(a) is a color image of the inner right leg of a male; (b) and (c) are the related NIR images; (d) is the uncovered image obtained from the enhanced NN mapping model.

Fig. 4.15 Visual evaluation of the RGB-NIR mapping approach for images taken from different viewing angles and poses (example 4).

(a) is a color image of the chest and left upper arm of a male; (b) and (c) are the related NIR images; (d) is the uncovered image obtained from the enhanced NN mapping model.

4.4.2 Evaluation of Deblocking and Vein Uncovering

As mentioned in chapter 3, for compressed skin images, our uncovering algorithm will provide blocky and discontinuous vein patterns, while the ICA-based method even cannot converge. Here we evaluate the deblocking approach developed in chapter 3 and the vein uncovering approach simultaneously.

Fig. 4.16 (a) is an image of the back of the left leg of a male, and Fig. 4.16 (b) shows the corresponding NIR image. Fig. 4.16 (c) is the compressed version of Fig. 4.16 (a) with the compression ratio of 76.95. Figs. 4.16 (d), (e) and (f) are respectively the resultant images from the adaptive processing (ADPROC) method [37], the Pointwise Shape-adaptive DCT (SADCT) method [94], and our deblocking approach. As mentioned in section 3.3, the knowledge-based approach has better performance. Therefore, we use it to remove the blocking artifacts in the compressed image. After that we use the RGB-NIR mapping approach to uncover vein patterns from the original, compressed, and deblocked images. Fig. 4.16 (b) shows that this skin patch contains complex vein patterns, which are successively uncovered by the RGB-NIR mapping approach from the original image (see Fig. 4.16 (g)). However, the patterns are destroyed by blocking artifacts and cannot be detected from the uncovering result of the compressed image (see Fig. 4.16 (h)). The ADPROC method [37] smoothes the boundaries of the blocks without recovering the veins (see Fig. 4.16 (i)), while the SADCT method [94] blurs the blocks and almost removes the veins (see Fig. 4.16 (j)). Our knowledge-based approach relieves the artifacts, and recovers the patterns successfully (see Fig. 4.16 (k)). Figs. 4.17 ~ 4.18 show two other examples.

These experimental results demonstrate that using our deblocking and vein uncovering approaches, we can uncover vein patterns from compressed skin images.

Fig. 4.16 Vein uncovering results from a deblocked image (example 1).

(a) is an original skin image of the back of the left leg of a male; (b) is its NIR image; (c) is the compressed version of (a) with the compression ratio of 76.95; (d) and (e) are the deblocked images using methods ADPROC [37] and SADCT [94]; (f) is the deblocked image using our knowledge-based approach; (g), (h), (i), (j), and (k) are uncovered images obtained by the enhanced NN mapping model from (a), (c), (d), (e), and (f), respectively.

Fig. 4.17 Vein uncovering results from a deblocked image (example 2).

(a) is an original skin image of the back of the left leg of a male; (b) is its NIR image; (c) is the compressed version of (a) with the compression ratio of 39.73; (d) and (e) are the deblocked images using methods ADPROC [37] and SADCT [94]; (f) is the deblocked image using our knowledge-based approach; (g), (h), (i), (j), and (k) are uncovered images obtained by the enhanced NN mapping model from (a), (c), (d), (e), and (f), respectively.

Fig. 4.18 Vein uncovering results from a deblocked image (example 3).

(a) is an original skin image of the back of the right leg of a male; (b) is its NIR image; (c) is the compressed version of (a) with the compression ratio of 60.62; (d) and (e) are the deblocked images using methods ADPROC [37] and SADCT [94]; (f) is the deblocked image using our knowledge-based approach; (g), (h), (i), (j), and (k) are uncovered images obtained by the enhanced NN mapping model from (a), (c), (d), (e), and (f), respectively.

4.4.3 Evaluation by Automatic Vein Matching

In this subsection, we use the automatic vein matching algorithm reviewed in section 2.5 to evaluate the RGB-NIR mapping uncovering approach. We perform automatic matching experiments on two testing image sets. The first one is a smaller testing set containing 300 color images and 300 NIR images from 150 different forearms. The second one is a larger testing set containing 498 color images and 498 NIR images from 249 different forearms. The smaller testing set is part of the larger testing set but has better image quality. We do experiments on these two sets to evaluate the uncovering performance on images with different qualities. All the testing images come from our skin image database. We do not use all the images in the database as testing images because some subjects have tattoos or too much hair on their skin. Fig. 4.19 shows some examples. The testing images are collected on two separate occasions, at an interval of around two weeks, and each subject in each session has one color image and one NIR image. The color images are collected by Nikon D70s and Canon 500D cameras, while the NIR images are collected by the JAI camera. Note that the color and NIR images are not captured simultaneously by the 2-CCD multi-spectral camera. In addition, the camera models and imaging conditions are different from those of the training image.

Fig. 4.19 Some images in our database which have tattoos or too much hair.

(a) is an image of a right forearm with tattoo; (b) is an image of a right forearm with a lot of hair; (c) is an image of a right forearm with both tattoo and a lot of hair. (a) and (c) are blurred for privacy protection.

The RGB-NIR mapping approach is applied to the color images. To demonstrate the effectiveness of the NN weight adjustment scheme, we evaluate both the generic NN mapping model and the enhanced NN mapping model. Vein patterns are extracted from the resultant images. Cumulative match characteristic curves generated from the matching results are used as a performance indicator. For each testing set, two groups of experiments are performed. The first group is to match vein patterns from the same type of images (i.e. color vs. color and NIR vs. NIR) taken on different occasions, i.e., 150 (or 249) vein patterns uncovered from the first session of color (NIR) images are matched with 150 (or 249) vein patterns from the second session of the color (NIR) images. The second group is to match vein patterns from different types of images (color vs. NIR) taken on different occasions. The vein patterns from the red channel are employed for comparison. The red channel is included in the experiments because it is close to the NIR channel.

Figs. 4.20 and 4.21 show the cumulative match characteristic curves from the experiments on the smaller and larger testing sets, respectively. Tables 4.1 and 4.2 list the corresponding rank-one and rank-10% identification accuracies. The term “rank-10% identification accuracy” refers to the percentage of input vein patterns whose corresponding vein patterns can be found in the database within the top 10% of the vein patterns given by the matching algorithm. The same-type-image matching results from the smaller testing set shows that the rank-one identification accuracy of the enhanced NN mapping model is 81.3%, which is very close to that of matching NIR images with NIR images, 84%. The rank-10% identification accuracy of the enhanced NN mapping model is as high as that of matching NIR images with NIR images, 92%. The different-type-image matching results are similar. Both rank-one and rank-10% identification accuracies of the enhanced NN mapping model are close to that of matching NIR images with NIR images.

The experimental results demonstrate that: (1) vein patterns can be uncovered using information extracted from corresponding color values and NIR intensities; (2) the proposed NN weight adjustment scheme is effective. It results in great improvement of the matching performance, and the final identification accuracy is comparable to that of the NIR images, which

are usually considered as ground truth of vein patterns; (3) the red channel contains vein information; (4) matching vein patterns from the same image types (e.g. color vs. color or NIR vs. NIR) performs better than matching vein patterns from different types (e.g. color vs. NIR); and (5) for both the same-type-image matching and the different-type-image matching, the rank-10% identification accuracy of the larger image set is worse than that of the smaller image set. It demonstrates that image quality has great influence on matching accuracy.

Fig. 4.20 Cumulative match characteristic curves of vein patterns uncovered by the RGB-NIR mapping approach on the smaller testing set.

(a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Fig. 4.21 Cumulative match characteristic curves of vein patterns uncovered by the RGB-NIR mapping approach on the larger testing set.

(a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Table 4.1 The rank-one and rank-10% identification accuracies from the RGB-NIR mapping approach on the smaller testing set.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	84%	92%
Red channel	46%	79.3%
The generic NN mapping model	65.3%	80.7%
The enhanced NN mapping model	81.3%	92%

(b)

	rank-one	rank-10%
Red channel	36%	58%
The generic NN mapping model	54%	72%
The enhanced NN mapping model	62.7%	80.7%

Table 4.2 The rank-one and rank-10% identification accuracies from the RGB-NIR mapping approach on the larger testing set.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	70.3%	82.3%
Red channel	45.4%	72.3%
The generic NN mapping model	61%	79.1%
The enhanced NN mapping model	67.9%	81.5%

(b)

	rank-one	rank-10%
Red channel	26.9%	45%
The generic NN mapping model	38.6%	56.6%
The enhanced NN mapping model	47.8%	66.7%

4.5 Summary

In this chapter, we have proposed an approach which extracts information from synchronized color and NIR images to uncover vein patterns from color images. An automatic intensity adjustment scheme was designed for illumination compensation. An automatic NN weight adjustment scheme was proposed to improve the performance of the approach. It is not limited to the hand or wrist and can theoretically be applied to any part of the body. Visual evaluations show that the uncovered images contain clear vein patterns. The matching experiments on two testing sets with images from 150 and 249 forearms indicate that the identification accuracy of the proposed approach is comparable to that of NIR images. We also performed simultaneous evaluation of the deblocking approach that we proposed in chapter 3 and the RGB-NIR mapping uncovering approach that we proposed in this chapter. The comparisons with other deblocking methods demonstrate that our approaches can uncover vein patterns from compressed skin images.

Chapter 5

K-M Model-Based Approach to Uncover Vein Patterns

Uncovering vein patterns from color skin images is a novel and challenging problem. In chapter 4, we proposed an RGB-NIR mapping approach to uncover vein patterns. It extracts information from a pair of synchronized color and NIR images, and uses a neural network (NN) to map RGB values to NIR intensities. The approach uses prior information from images to uncover vein patterns. However, it does not utilize any physical or biological information.

As mentioned in the previous chapters, skin is an intricate layered material whose internal structures vary with individuals, body sites, and time. Additionally, the physical process of image formation may complicate the situation. When the light from an illumination source hits the skin, some is absorbed while some is reflected and captured by the sensor in a camera. The characteristics of the illuminant and the camera also play a significant role in the skin color formation in an image. In this chapter we propose a vein uncovering approach based on the principles of optics and skin biophysics. It uses the radiative transfer equation to quantitatively describe light transport in the human skin. Then it inverses the process of skin color formation in an image, and derives corresponding biophysical parameters. Although pixels of veins and pixels of generic skin have very similar RGB values, their biophysical parameters differ significantly. Therefore, we can uncover the vein patterns from their spatial distributions. This approach is also not limited to the hand or wrist, and can theoretically be applied to any part of the body. Based on this approach we develop four optical models. In this chapter we will introduce the first and second optical models. The third and fourth optical models will be presented in chapter 6.

In the first optical model, we use the most popular and simplest solution to the radiative transfer equation – the K-M model, to approximate the transport of light in the human skin, and assume that optical properties of human skin is determined by three layers – the stratum corneum, the epidermis, and the dermis. This model has two limits: 1) it does not use skin color to optimize the biophysical parameters, and 2) it is sensitive to the power of the illuminant. To overcome these limits, we develop the second optical model which uses a color optimization scheme to derive the range of biophysical parameters. We also adopt the automatic intensity adjustment scheme that we proposed in chapter 4 to improve the robustness of the uncovering performance.

Claridge et al.’s method [85], using color images to diagnose pigmented skin lesions is perhaps the most related method to our approach. The proposed approach and their method both reveal the distributions of melanin, hemoglobin, and the depth of dermis from color images. However, they have the following differences. First, the proposed approach is based on a more complicated and accurate biophysical structure of the skin. Second, Claridge et al. collected color images in a controlled environment with white incident light, while the proposed approach works under different illumination conditions. Third, their method is used for medical applications, but our target is for personal identification. Fourth, their method is based on table lookup, while the proposed approach utilizes neural network, which has good extrapolation ability and can well approximate the nonlinear relationship between skin color and biophysical parameters.

In this chapter, section 5.1 presents the optical approach and the first optical model. Section 5.2 presents the second optical model with the schemes to improve the first model. In each of these two sections we also offer experimental evaluations and discussion. Section 5.3 offers some summary remarks.

5.1 The First Optical Model

We will present the uncovering approach and our first optical model in this section. Subsection 5.1.1 gives an overview of the approach. Subsection 5.1.2 introduces the principles of skin optics

and the process of skin color formation. Subsection 5.1.3 presents the inverse process of skin color formation and the first optical model. Subsection 5.1.4 offers experimental results and discussion.

5.1.1 Overview of the Approach

In the first three optical models, we assume that the optical properties of human skin are determined by three layers – the stratum corneum, the epidermis, and the dermis. In addition, we assume that veins are located in the dermis. To uncover vein patterns hidden in color images, a computation model is used to reveal the spatial distributions of the volume fraction (%) of the epidermis occupied by melanosomes, the volume fraction (%) of the dermis occupied by blood, and the depth of the dermis. They are respectively denoted as $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} . Intuitively, vein patterns can be found in $\mathcal{G}_p$, because they contain higher concentration of blood than other skin components. Veins with a certain size can occupy the spaces of dermis and affect the epidermis and therefore, they can also be observed from the spatial distributions of $\mathcal{G}_m$ and d^{der} .

Mathematically, we sought a function g such that given a color image and other information, it maps input variables to $\mathcal{G}_m$, $\mathcal{G}_p$ and d^{der} . More precisely, the mapping is

$$[\mathcal{G}_m, \mathcal{G}_p, d^{der}] = g(R, G, B, E), \quad (5.1)$$

where R , G , and B represent the three color components in a given image, and E represents other information. In our current model, E includes prior knowledge from cameras and illuminants. To obtain the function g , we firstly model the formation of skin color in an image i.e.,

$$[R, G, B] = f(I(\lambda), \mathcal{G}_m, \mathcal{G}_p, d^{der}, S_R(\lambda), S_G(\lambda), S_B(\lambda)), \quad (5.2)$$

where $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$ are respectively the R , G , and B spectral response functions of a camera, $I(\lambda)$ is an illuminant, and λ represents wavelengths. These components nonlinearly influence the R , G and B values. In legal cases, the spectral response functions can be known through the camera model, which can be found in headers of evidence images. Three types of

light sources are commonly encountered in a real environment: daylight, incandescent lamps, and fluorescent lamps. They correspond respectively to illuminant D65, A and F, which are standard illuminants established by The International Commission on Illumination [107]. Based on environments exposed in evidence images, the illuminant can be determined. Thus, it is not unrealistic to assume that $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ and $I(\lambda)$ are known.

Once the function f is established, R , G , and B values of the skin in an image can be obtained, when $I(\lambda)$, $\mathcal{G}_m$, $\mathcal{G}_p$, d^{der} , $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$ are given. In other words, the mapping from $[I(\lambda), \mathcal{G}_m, \mathcal{G}_p, d^{der}, S_R(\lambda), S_G(\lambda), S_B(\lambda)]$ to $[R, G, B]$ is established. If this mapping is one-to-one,

$$[I(\lambda), \mathcal{G}_m, \mathcal{G}_p, d^{der}, S_R(\lambda), S_G(\lambda), S_B(\lambda)] = f^{-1}(R, G, B), \quad (5.3)$$

would be obtained. Because $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ and $I(\lambda)$ are known, Eq. (5.3) can be simplified as $[\mathcal{G}_m, \mathcal{G}_p, d^{der}] = f^{-1}(R, G, B)$. However, f is not a one-to-one function, and therefore, f^{-1} does not exist. We construct a function $g \approx f^{-1}$, i.e.,

$$[\mathcal{G}_m, \mathcal{G}_p, d^{der}] = g(R, G, B), \quad (5.4)$$

from a neural network. It should be emphasized that the information from $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ and $I(\lambda)$ is embedded in Eq. (5.4), although they are not shown explicitly. Fig. 5.1 illustrates our problem formulation, where the left block represents the skin color formation model, f , and the right block represents the inversion of the skin color formation, g to obtain $[\mathcal{G}_m, \mathcal{G}_p, d^{der}]$ from $[R, G, B]$. Subsection 5.1.2 will present in detail the process of skin color formation, and subsection 5.1.3 will present the inversion of skin color formation.

Fig. 5.1 A schematic diagram of the optical approach to uncover vein patterns.

The left block represents the process of skin color formation and the right block represents the inversion of skin color formation to obtain $[\mathcal{G}_m, \mathcal{G}_p, \mathcal{A}^{der}]$ from $[R, G, B]$.

5.1.2 The Skin Color Formation

In this subsection, we use the K-M model and the optical properties of human skin to model skin color formation in an image. Subsection 5.1.2.1 gives an overview of the spectral reflectance of skin. Then, subsections 5.1.2.2 ~ 5.1.2.4 describe respectively the optical properties of the three skin layers: the stratum corneum, the epidermis and the dermis. Subsection 5.1.2.5 presents a model of illuminants. Subsection 5.1.2.6 presents a model of sensors in digital cameras.

5.1.2.1 An Overview of the Spectral Reflectance of Skin

According to the K-M model, the reflectance of a material is mainly determined by the absorption and scattering characteristics of each layer. In our first three optical models, we assume that the optical properties of human skin are determined by three layers – the stratum corneum, the epidermis, and the dermis. We do not consider the hypodermis at this stage.

Fig. 5.2 illustrates the process of light transport in human skin. A fraction of the incident light is reflected by the stratum corneum and the rest penetrates into the deeper layers. In the epidermal and dermal layers, the light is scattered multiple times and absorbed by melanin, hemoglobin, β -carotene, and bilirubin [112].

Fig. 5.2 Illustration of the process of light transport in human skin.

The reflectance $R_{123}(\lambda)$ and transmittance $T_{123}(\lambda)$ of the whole skin can be obtained from the recursive equations (Eqs. (2.7) ~ (2.8)):

$$R_{123}(\lambda) = R_{12}(\lambda) + \frac{T_{12}(\lambda)^2 R_3(\lambda)}{1 - R_{12}(\lambda) R_3(\lambda)}, \quad (5.5)$$

$$T_{123}(\lambda) = \frac{T_{12}(\lambda) T_3(\lambda)}{1 - R_{12}(\lambda) R_3(\lambda)}, \quad (5.6)$$

where

$$R_{12}(\lambda) = R_1(\lambda) + \frac{T_1(\lambda)^2 R_2(\lambda)}{1 - R_1(\lambda) R_2(\lambda)}, \quad (5.7)$$

$$T_{12}(\lambda) = \frac{T_1(\lambda) T_2(\lambda)}{1 - R_1(\lambda) R_2(\lambda)}. \quad (5.8)$$

The variables $R_i(\lambda)$ and $T_i(\lambda)$ ($i = 1, 2, 3$) in the equations are respectively the reflectance and transmittance of the i^{th} layer of the skin. The 1st, 2nd and 3rd layers are respectively the stratum corneum, the epidermis, and the dermis. According to the K-M model (Eqs. (2.3) ~ (2.6)), $R_i(\lambda)$ and $T_i(\lambda)$ are determined by the absorption coefficient, μ_{ia} , the scattering coefficient, μ_{is} , and the thickness of the corresponding layer. The following subsections describe the computational details of these coefficients.

5.1.2.2 The Stratum Corneum

The total absorption coefficient of the stratum corneum is:

$$\mu_{1a}(\lambda) = \mu_a^{base}(\lambda) + \mu_a^{cs}(\lambda) \quad [cm^{-1}], \quad (5.9)$$

where $\mu_a^{base}(\lambda)$ and $\mu_a^{cs}(\lambda)$ are respectively the absorption coefficients of the baseline skin and the β -carotene in the stratum corneum [108]. They are calculated as:

$$\mu_a^{base}(\lambda) = 0.244 + 85.3 \cdot \exp\left(-\frac{\lambda - 154}{66.2}\right) \quad [cm^{-1}], \quad (5.10)$$

$$\mu_a^{cs}(\lambda) = \frac{\varepsilon^{car}(\lambda)}{m^{car}} c^{cs} \quad [cm^{-1}], \quad (5.11)$$

where $\varepsilon^{car}(\lambda)$ is the extinction coefficient of β -carotene, m^{car} is its molecular weight, and c^{cs} is its concentration in the stratum corneum [108]. Fig. 5.3 shows the spectral absorption coefficient curve of the baseline skin [111]. Fig. 5.4 shows the spectral extinction coefficient curve of β -carotene [111].

Fig. 5.3 The spectral absorption coefficient curve of the baseline skin.

Fig. 5.4 The spectral extinction coefficient curve of β -carotene.

The scattering of light in the stratum corneum combines the Mie scattering by large cylindrical collagen fibers and the Rayleigh scattering by small-scale structure associated with the

collagen fibers and other cellular structures. The total scattering coefficient in the stratum corneum can be obtained by

$$\mu_{1s}(\lambda) = 2 \times 10^5 \times \lambda^{-1.5} + 2 \times 10^{12} \times \lambda^{-4} \quad [cm^{-1}], \quad (5.12)$$

where the first term is for Mie scattering and the second is for Rayleigh scattering. Fig. 5.5 shows the spectral curves of the scattering coefficients [109]. Eq. (5.12) is also applicable to calculate the scattering coefficient of the epidermis [110].

Fig. 5.5 The spectral scattering coefficient curve of the stratum corneum.

5.1.2.3 The Epidermis

The total absorption coefficient of the epidermis is:

$$\mu_{2a}(\lambda) = (\beta_m \mu_a^{em}(\lambda) + (1 - \beta_m) \mu_a^{pm}(\lambda)) \mathcal{G}_m + (\mu_a^{base}(\lambda) + \mu_a^{ce}(\lambda)) (1 - \mathcal{G}_m) \quad [cm^{-1}], \quad (5.13)$$

where $\mu_a^{em}(\lambda)$ and $\mu_a^{pm}(\lambda)$ are the absorption coefficients of eumelanin and pheomelanin, respectively. β_m controls the ratio of their amount [108]. $\mathcal{G}_m$ is the volume fraction (%) of the epidermis occupied by melanosomes, organelles containing melanin. For light-colored skin, $\mathcal{G}_m = 1.3 \sim 6.3\%$; for moderately pigmented skin, $\mathcal{G}_m = 7 \sim 16\%$, and darkly pigmented skin, $\mathcal{G}_m = 17 \sim 43\%$ [109]. $\mu_a^{ce}(\lambda)$ is the absorption coefficient of β -carotene in the epidermis. $\mu_a^{em}(\lambda)$, $\mu_a^{pm}(\lambda)$ and $\mu_a^{ce}(\lambda)$ are calculated as:

$$\mu_a^{em}(\lambda) = \frac{\varepsilon^{em}(\lambda)}{m^{em}} c^{em} \quad [cm^{-1}], \quad (5.14)$$

$$\mu_a^{pm}(\lambda) = \frac{\varepsilon^{pm}(\lambda)}{m^{pm}} c^{pm} \quad [cm^{-1}], \quad (5.15)$$

$$\mu_a^{ce}(\lambda) = \frac{\varepsilon^{car}(\lambda)}{m^{car}} c^{ce} \quad [cm^{-1}], \quad (5.16)$$

where $\varepsilon^{em}(\lambda)$ and $\varepsilon^{pm}(\lambda)$ are the extinction coefficients of eumelanin and pheomelanin, respectively, m^{em} and m^{pm} are their molecular weights, c^{em} and c^{pm} are their concentrations in melanosomes, respectively, and c^{ce} is the concentration of β -carotene in the epidermis [108]. Figs. 5.6 (a) and (b) show respectively the spectral extinction coefficient curves of eumelanin and pheomelanin [111].

Fig. 5.6 The spectral extinction coefficient curves of eumelanin and pheomelanin.

(a) eumelanin, (b) pheomelanin.

5.1.2.4 The Dermis

The total absorption coefficient of the dermis is:

$$\mu_{3a}(\lambda) = (\gamma \mu_a^{ohb}(\lambda) + (1 - \gamma) \mu_a^{dhh}(\lambda) + \mu_a^{bil}(\lambda) + \mu_a^{cd}(\lambda)) \mathcal{G}_p + \mu_a^{base}(\lambda) (1 - \mathcal{G}_p) \quad [cm^{-1}], \quad (5.17)$$

where $\mu_a^{ohb}(\lambda)$ and $\mu_a^{dhb}(\lambda)$ are the absorption coefficients of oxy-hemoglobin and deoxy-hemoglobin, respectively. γ controls the ratio of their amount [110]. $\mu_a^{bil}(\lambda)$ and $\mu_a^{cd}(\lambda)$ are the absorption coefficients of bilirubin and the β -carotene in the dermis, respectively. $\mathcal{G}_p$ is the volume fraction (%) of the dermis occupied by blood [111]. $\mu_a^{ohb}(\lambda)$, $\mu_a^{dhb}(\lambda)$, $\mu_a^{bil}(\lambda)$ and $\mu_a^{cd}(\lambda)$ are calculated as:

$$\mu_a^{ohb}(\lambda) = \frac{\varepsilon^{ohb}(\lambda)}{m^{hb}} c^{hb} \quad [cm^{-1}], \quad (5.18)$$

$$\mu_a^{dhb}(\lambda) = \frac{\varepsilon^{dhb}(\lambda)}{m^{hb}} c^{hb} \quad [cm^{-1}], \quad (5.19)$$

$$\mu_a^{bil}(\lambda) = \frac{\varepsilon^{bil}(\lambda)}{m^{bil}} c^{bil} \quad [cm^{-1}], \quad (5.20)$$

$$\mu_a^{cd}(\lambda) = \frac{\varepsilon^{car}(\lambda)}{m^{car}} c^{cd} \quad [cm^{-1}], \quad (5.21)$$

where $\varepsilon^{ohb}(\lambda)$, $\varepsilon^{dhb}(\lambda)$, and $\varepsilon^{bil}(\lambda)$ are the extinction coefficients of oxy-hemoglobin, deoxy-hemoglobin, and bilirubin, respectively. m^{hb} and m^{bil} are the molecular weights of hemoglobin and bilirubin, respectively. c^{hb} , c^{bil} , and c^{cd} are the concentrations of hemoglobin, bilirubin, and β -carotene in blood, respectively [112]. Figs. 5.7 (a) and (b) show respectively the spectral extinction coefficient curves of hemoglobin and bilirubin [111].

The deeper regions of skin are generally less scattering than the upper areas, because there is a lower concentration of scatterers [109]. The scattering coefficient of the dermis is about 50% of that of the epidermis, i.e., $\mu_{3s}(\lambda) = 0.5\mu_{2s}(\lambda)$.

Using the absorption coefficients, μ_{1a} , μ_{2a} , and μ_{3a} , the scattering coefficients, μ_{1s} , μ_{2s} , and μ_{3s} , Eqs. (2.3) ~ (2.6), and Eqs. (5.5) ~ (5.8), the reflectance of the skin, $R_{123}(\lambda)$ can be obtained.

Fig. 5.7 The spectral extinction coefficient curves of hemoglobin and bilirubin.

(a) hemoglobin, (b) bilirubin.

5.1.2.5 A Model of Illuminants

Spectral characteristics and spatial distribution of illuminants play an important role in the formation of skin color in an image. In 1931, The International Commission on Illumination (usually abbreviated as CIE for its French name) introduced six standard illuminants and named them by A, B, C, D, E, and F [107]. Illuminant A represents typical domestic tungsten-filament lighting. Illuminant B represents noon sunlight, while illuminant C represents average day light. The D series of illuminants represents natural daylight. Illuminant D65 is part of the D series that portrays standard illumination conditions at open air. Illuminant E, having a constant spectral power distribution, is always used as a theoretical reference. The F series of illuminants represents various types of fluorescent lighting. The FL1–FL6 illuminants represent “standard” fluorescent lamps [107].

Three types of light sources are commonly encountered in a real environment: daylight, incandescent lamps, and fluorescent lamps, corresponding to illuminant D65, A and F, respectively. Their spectral power distributions are shown in Fig. 5.8.

Fig. 5.8 The spectral power distributions of illuminants D65, A, and F1-F6.
(a) D65 and A, (b) F1-F6 [107].

5.1.2.6 A Model of Sensors in Digital Cameras

We usually model RGB sensors in digital cameras by their spectral response functions. Figs. 5.9 (a) and (b) show respectively the spectral response functions of the Nikon D70s and Canon 500D cameras [113] as examples.

By supplying the incident light and the skin reflectance, the color captured by a camera is then derived by convolving the reflected light with spectral response functions of a camera, i.e.,

$$R = \int_0^\infty I(\lambda) R_{123}(\lambda) S_R(\lambda) d\lambda, \quad (5.22)$$

$$G = \int_0^\infty I(\lambda) R_{123}(\lambda) S_G(\lambda) d\lambda, \quad (5.23)$$

$$B = \int_0^\infty I(\lambda) R_{123}(\lambda) S_B(\lambda) d\lambda. \quad (5.24)$$

where $I(\lambda)$ is the illuminant, $R_{123}(\lambda)$ is the reflectance of skin, $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$ are respectively the R , G , and B spectral response functions of the camera. As mentioned before, given a camera model, these spectral response functions can be obtained from manufacturers. By selecting different illuminants, camera models, and skin parameters, we can simulate different imaging configurations and types of skin.

Fig. 5.9 The spectral response functions of Nikon D70s and Canon 500D.

(a) Nikon D70s, (b) Canon 500D.

5.1.3 An Inversion of the Skin Color Formation

Using Eqs. (5.5) ~ (5.24) as a forward model, the constant skin parameters determined by a series of biophysical experiments [108], and other biological data to limit the ranges of $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} , data pairs (R, G, B) and $(\mathcal{G}_m, \mathcal{G}_p, d^{der})$ can be obtained. We change $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} exhaustively and obtain the corresponding color values. Then we use the color values as inputs and the parameters as target outputs to train a three-layered feed-forward neural network. The transfer functions in the hidden and output layers are respectively the classical tan-sigmoid and linear functions. To avoid too complex network structure and much training time, we put 5 neurons in the hidden layer. The network is trained with the Levenberg-Marquardt back-propagation algorithm [114] and converges in five iterations. It should be emphasized that the trained neural network depends on the spectral response functions of the camera and the illuminant. Thus, different cameras or different illuminants should have different neural networks. The trained neural network is in fact the function g in Eq. (5.4).

To visualize vein patterns in a testing image, we input its RGB values to the trained neural network, and collect the outputs to form three distribution maps: the spatial distributions of melanin and hemoglobin, and the variation of the depth of dermis. Because hemoglobin is mainly

found in veins which are assumed to be located in the dermis, the pixels of veins and pixels of generic skin have different parameter values. The melanin and depth of dermis also provide information of vein patterns. Because veins occupy some space, generic skin has a thicker dermis layer than skin regions with veins. For the same reason, the distribution of melanin is also influenced by veins. Consequently, they can also be seen from these distribution maps.

5.1.4 Experimental Evaluations

In this subsection, we evaluate the first optical model. Firstly, we use visual observation to evaluate the uncovering performance of the optical model. Then, we use manual matching to verify the reliability of the uncovered vein patterns. The experimental results of automatic matching are given in subsection 5.2.2.2.

5.1.4.1 Evaluation by Visual Observation

To evaluate the first optical model, we use the corresponding color and NIR skin images in our database as testing images for evaluation. We also compare our results with those obtained from Tsumura et al.'s method [74] and Claridge et al.'s method [85]. Nikon D60 and the illuminant of D65 are employed in these experiments. The values of the constant skin parameters are given in Table 5.1.

Table 5.1 Constant skin parameters.

Parameter	Descriptions	Default Value
c^{cs}	the concentration of β -carotene in the stratum corneum	2.1×10^{-4} g/L
c^{ce}	the concentration of β -carotene in the epidermis	2.1×10^{-4} g/L
c^{cd}	the concentration of β -carotene in blood	7.0×10^{-5} g/L
c^{em}	the concentration of eumelanin in melanosomes	80 g/L
c^{pm}	the concentration of pheomelanin in melanosomes	12 g/L
c^{hb}	the concentration of hemoglobin in blood	150 g/L
c^{bil}	the concentration of bilirubin in blood	0.05 g/L
m^{car}	the molecular weight of β -carotene	537 g/mole
m^{em}	the molecular weight of eumelanin	178 g/mole
m^{pm}	the molecular weight of pheomelanin	295 g/mole
m^{hb}	the molecular weight of hemoglobin	66500 g/mole
m^{bil}	the molecular weight of bilirubin	585 g/mole
β_m	the ratio between the amount of eumelanin and pheomelanin	0.2
γ	the ratio between the amount of oxy- and deoxy- hemoglobin	0.75
d^{str}	the depth of the stratum corneum	0.001 cm
d^{epi}	the depth of the epidermis	0.01 cm

Figs. 5.10 ~ 5.13 show the experimental results, where (a)s are original color skin images, (b)s are the corresponding NIR images, (c)s and (d)s are the two components obtained from Tsumura et al.'s method [74], (e)s, (f)s, and (g)s are respectively the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85], and (h)s, (i)s and (j)s are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

Fig. 5.10 (a) is an image of the front side of the left leg of a female. The skin in this image has low concentration of melanin and fat, and therefore, Tsumura et al.'s method [74] can show the veins. Claridge et al.'s method [85] provides almost nothing about veins. The first optical model shows clear vein patterns as Tsumura et al.'s method. The clarity of the vein patterns in the right part of the image is comparable to or even better than that in the NIR image. Fig. 5.11 (a) is an image of the back side of the left leg of a male. Fig. 5.12 (a) is an image of the back side of the right upper arm of a male. Fig. 5.13 (a) is an image of the front side of the left leg of a male. The skin in these images contains higher concentration of melanin or fat. It is nearly impossible to see the vein patterns in the color images. For these images, Tsumura et al.'s method [74] and Claridge et al.'s method [85] provide almost nothing about veins, but the first optical model uncovers clear vein patterns. Fig. 5.14 (a) is an image of the outer left forearm of a male. For this image, Tsumura et al.'s method [74] cannot converge, and Claridge et al.'s method [85] provides nothing, but the first optical model uncovers vein patterns. These experimental results are very encouraging and clearly demonstrate that body vein patterns can be uncovered from color skin images for forensic investigation.

Fig. 5.10 Visual evaluation of the first optical model (example 1).

(a) is a skin image of the front side of the left leg of a female; (b) is its NIR image; (c) and (d) are the first and second components obtained from Tsumura et al.'s method [74]; (e), (f), and (g) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85]; and (h), (i), and (j) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

Fig. 5.11 Visual evaluation of the first optical model (example 2).

(a) is a skin image of the back side of the left leg of a male; (b) is its NIR image; (c) and (d) are the first and second components obtained from Tsumura et al.'s method [74]; (e), (f), and (g) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85]; and (h), (i), and (j) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

Fig. 5.12 Visual evaluation of the first optical model (example 3).

(a) is a skin image of the back side of the right upper arm of a male; (b) is its NIR image; (c) and (d) are the first and second components obtained from Tsumura et al.'s method [74]; (e), (f), and (g) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85]; and (h), (i), and (j) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

Fig. 5.13 Visual evaluation of the first optical model (example 4).

(a) is a skin image of the front side of the left leg of a male; (b) is its NIR image; (c) and (d) are the first and second components obtained from Tsumura et al.'s method [74]; (e), (f), and (g) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85]; and (h), (i), and (j) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

Fig. 5.14 Visual evaluation of the first optical model (example 5).

(a) is a skin image of the outer left forearm of a male; (b) is its NIR image; (c), (d), and (e) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of papillary dermis obtained from Claridge et al.'s method [85], they provide nothing about veins; and (f), (g), and (h) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model.

5.1.4.2 Evaluation by Manual Matching

To verify the reliability of our uncovered vein patterns, we ask five “examiners” to match NIR images with uncovered distribution maps we generate. Manual matching is expected to be more accurate and reliable than automatic matching. For personal identification in legal cases, final fingerprint comparisons are still based on expert latent print examiners. Our testing set contains image pairs (NIR and color images) from 4 body parts of 150 persons: inner forearm (IF), outer forearm (OF), front thigh (FT), and outer thigh (OT). The images were taken in two sessions, with

an interval of at least around two weeks. We uncover vein patterns from the color images using the first optical model. The evaluations by the 5 examiners are performed in two testing sessions. In each session, 16 groups of images (4 groups from each body part) are presented to each of the “examiners”. Each group contains 3 distribution maps from a color skin image and 10 NIR images — one is from the same individual that the color image is from and the other nine are selected from the same body part of the other 149 persons. The examiners are asked to match the distribution maps to one of the ten NIR images. If they are not confident in their match, they are allowed to answer as such. In the first testing session, the original color image and the corresponding NIR image are taken on the same photography day, while in the second session, they are taken on different days. Thus, in the second testing session, there are more illumination and perspective variations between the NIR and the color image. Figs. 5.15 ~ 5.16 respectively give an example of manual matching in the two testing sessions. Table 5.2 gives the matching results, where C represents “correct” and N represents “NO confidence”. The percentage of “correct” is calculated by the total number of correct matches over the total number of attempted matches. The percentage of “NO confidence” is calculated by the total number of unanswered matches over the total number of attempts. In the first session, Examiners 1, 3 and 5 are 100% correct, answering all matches. In the second session, Examiners 2 and 5 are 100% correct; Examiner 2 does not answer in 25% of the trials, while Examiner 5 answers for all of the matches. These results demonstrate that different examiners have different capability to recognize the veins. Examiner 5 may have performed significantly better than others because he is familiar with fingerprint, palmprint, iris, and skin mark matching. Most importantly, these results indicate that the uncovering vein patterns can be used for personal identification. Of course, professional examiners who would be performing vein pattern matching for legal cases would have received sufficient training, as is the case for professional fingerprint examiners.

Although the experimental results are encouraging, this manual matching is impractical in cases where no suspects have been apprehended (i.e., the vein patterns hidden in evidence images

have to be compared to a large criminal database). In subsection 5.2.2.2, we will use the automatic matching algorithm to evaluate the first and second optical models together.

Fig. 5.15 An example of manual matching in the first session.

(a), (b) and (c) are the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from an inner forearm (IF) by the first optical model. (1) ~ (10) are 10 NIR images — one is from the same individual that the color image is from and the other nine are selected from the same body part of the other 149 persons. The NIR images have been enhanced for manual matching.

Fig. 5.16 An example of manual matching in the second session.

(a), (b) and (c) are the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from an outer thigh (OT) by the first optical model. (1) ~ (10) are 10 NIR images — one is from the same individual that the color image is from and the other nine are selected from the same body part of the other 149 persons. The NIR images have been enhanced for manual matching.

Table 5.2 Manual matching results (%).

		Session 1					Session 2				
		IF	OF	FT	OT	AVG	IF	OF	FT	OT	AVG
¹ E1	² C	100	100	100	100	100	100	75	100	100	93.75
	³ N	0	0	0	0	0	0	0	0	0	0
E2	C	100	75	75	100	87.5	100	100	100	100	100
	N	0	0	0	50	12.5	0	25	25	50	25
E3	C	100	100	100	100	100	100	75	100	100	93.75
	N	0	0	0	0	0	0	0	0	0	0
E4	C	100	100	75	100	93.75	100	100	25	100	81.25
	N	0	0	0	0	0	0	0	0	0	0
E5	C	100	100	100	100	100	100	100	100	100	100
	N	0	0	0	0	0	0	0	0	0	0
⁴ AVG	C	100	95	90	100	96.25	100	90	85	100	93.75
	N	0	0	0	10	2.5	0	5	5	10	5

¹E stands for Examiner; ²C represents “correct matching”, ³N represents “NO-confidence”, and ⁴AVG means average.

5.2 The Second Optical Model

The first optical model proposed in the previous section has two limits. First, it does not use skin color to optimize the biophysical parameters. When simulating the process of skin color formation to produce training data for the neural network, the parameter ranges are derived from the literature. They are obtained in ideal medical settings from a particular race and body site (e.g. hand). Therefore, the data from the literature may not be applicable to images obtained in uncontrolled environments. Second, it is sensitive to the power of the illuminant. Skin images taken under different illumination conditions have different intensity values. This can have significant influence on the vein uncovering performance. The first optical model does not take intensity variation into consideration. Fig. 5.17 shows a set of uncovering results from the first optical model, from which it is difficult to detect vein patterns.

In this section, we propose the second optical model. It has two new schemes to overcome the limits of the first model. First, we propose a color optimization scheme to derive the range of biophysical parameters to enhance the performance of the vein uncovering approach. Second, it uses the automatic intensity adjustment scheme we proposed in chapter 4 to improve its

robustness. Subsection 5.2.1 presents the color optimization scheme. Subsection 5.2.2 reports the evaluation results.

Fig. 5.17 An example of the vein uncovering results using the first optical model.

(a) is a color image of the inner right forearm of a male; (b) is its NIR image; (c), (d) and (e) are respectively the distribution maps of melanin, hemoglobin and the depth of dermis.

5.2.1 Parameter Range Optimization through Color

The second optical model is based on the same principles of optics and skin biophysics as the first model. We still use a three-layered feed-forward neural network to approximate the inversion of skin color formation. In contrast to the first model, when simulating the forward skin color formation to prepare training data, we do not take the parameter ranges directly from the literature because they are obtained in ideal medical settings from a particular race (e.g. Caucasian) and body site (e.g. hand). Therefore, the data from the literature may not be perfectly applicable to our images. Our image database mainly consists of Asian skin, so we select a typical skin image from an Asian subject and use its average RGB value as a target to optimize the parameters in the model of skin color formation. After obtaining the optimal values, we set the parameter ranges from 50% to 200% of the optimal values. Fig. 5.18 illustrates the process of parameter range optimization. Then we simulate the process of skin color formation and obtain the corresponding RGB values. Fig. 5.19 illustrates the process of training data preparation for the neural network. The trained neural network approximates the inversion of skin color formation. To visualize vein patterns in a test image, its RGB values are input to the trained neural network, and the outputs are collected to form the distribution maps of $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} .

Fig. 5.18 Illustration of parameter range optimization.

Fig. 5.19 Data preparation for training a three-layered feed-forward neural network.

The illuminant can significantly influence the skin color in images. Therefore, the performance of the RGB-based uncovering approach deteriorates if $I(\lambda)$ is not estimated accurately. Although it is true that several standard illuminants (e.g. D65) can be used to estimate $I(\lambda)$, their power can vary significantly. To avoid manually estimating the power, we use the automatic intensity adjustment scheme that we proposed in chapter 4 to improve the robustness of the uncovering approach.

Although all of the three parameter maps show vein patterns, the distribution map of the depth of dermis is the clearest. Therefore, the distribution map of the depth of dermis is used during intensity adjustment and also later during automatic matching. An intensity adjustment ratio is introduced in the mapping (see Eq. (4.1)), and the average local variance of the distribution map is used as an objective function to determine the optimal adjustment ratio (see Eq. (4.2)).

5.2.2 Experimental Evaluations

In this subsection, we evaluate the second optical model. Firstly, we visually compare the uncovering performance of the first and second optical models. Then, we use the automatic matching algorithm to perform vein identification.

5.2.2.1 Evaluation by Visual Observation

Like the evaluation of the first optical model, we use the corresponding color and NIR skin images in our database as testing images. Figs. 5.20 ~ 5.22 show the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first and second optical models. Visual comparison with NIR images demonstrates that the second optical model results in more and clearer vein patterns.

Fig. 5.20 Visual evaluation of the second optical model (example 1).

(a) is a color image of the inner right forearm of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

Fig. 5.21 Visual evaluation of the second optical model (example 2).

(a) is a color image of the inner right leg of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

Fig. 5.22 Visual evaluation of the second optical model (example 3).

(a) is a color image of the inner left thigh of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

To simulate crime scene images, we use images taken from different viewing angles and poses in our database to test the optical models. We use the NIR images of the corresponding body position as ground truth. Figs. 5.23 ~ 5.25 show the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first and second optical models. The problems of uneven illumination and incident angle in these images are more serious than the well-pose images. The resultant distribution maps show that the first optical model still can

uncover some vein patterns, but the second optical model results in more and clearer vein patterns.

Fig. 5.23 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 1).

(a) is a color image of the thighs of a male; (b) and (c) are the related NIR images of the thighs, (c) has been enhanced for visual comparison; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

Fig. 5.24 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 2).

(a) is a color image of the legs and feet of a male; (b) and (c) are the related NIR images of the legs and feet; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

Fig. 5.25 Visual evaluation of the second optical model for images taken from different viewing angles and poses (example 3).

(a) is a color image of the legs and feet of a male; (b) and (c) are the related NIR images of the legs; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model.

5.2.2.2 Evaluation by Automatic Vein Matching

In this section, we use the automatic matching algorithm to evaluate the vein patterns uncovered by the first and second optical models. We use the testing set of 150 different forearms to evaluate the extracted vein patterns. This testing set is also used for evaluation of the RGB-NIR mapping approach in chapter 4. The first and second optical models are applied to the color images. For the two models, the same illuminant D65 is used to perform uncovering. As mentioned before, the distribution map of the depth of dermis is the clearest. Therefore, it is used for automatic matching. Figs. 5.26 ~ 5.28 illustrate some extracted vein patterns obtained by the two optical models. In the first and third examples, the first optical model does not provide a lot vein patterns (see Fig. 5.26 (e) and Fig. 5.28 (e)). In the second example, the distribution map obtained by the first optical model is very noisy (see Fig. 5.27 (c)). Consequently, the extracted vein pattern image contains a lot of small bifurcations (see Fig. 5.27 (e)). In all of the examples, the second optical model provides clean and clear vein patterns.

Fig. 5.26 Comparison between the extracted vein patterns from the first and second optical models (example 1).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (f) are respectively the distribution maps of the depth of dermis obtained by the first and second optical models; (d) and (g) are respectively the enhanced vein images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images extracted from (d) and (g).

Fig. 5.27 Comparison between the extracted vein patterns from the first and second optical models (example 2).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (f) are respectively the distribution maps of the depth of dermis obtained by the first and second optical models; (d) and (g) are respectively the enhanced vein images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images extracted from (d) and (g).

Fig. 5.28 Comparison between the extracted vein patterns from the first and second optical models (example 3).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (f) are respectively the distribution maps of the depth of dermis obtained by the first and second optical models; (d) and (g) are respectively the enhanced vein images obtained from (c) and (f); (e) and (h) are respectively the vein pattern images extracted from (d) and (g).

Cumulative match characteristic curves generated from the matching results are used as a performance indicator. Two sets of experiments are performed: matching same types of images and matching different types of images. The vein patterns from the red channel are employed for comparison. Fig. 5.29 (a) shows the experimental results of matching vein patterns from the same image type. The rank-one identification accuracy of the second optical model is 72%, while those of the first optical model and the red channel are only 25% and 47%, respectively. The rank-10% identification accuracy of the second optical model is 87%, while those of the first optical model and the red channel are 60% and 76%, respectively.

Fig. 5.29 (b) shows the experimental results of matching vein patterns from different image types. As with the results in Fig. 5.29 (a), the second optical model is still superior. Its rank-one identification accuracy is 66% and the rank-10% identification accuracy is 76.6%. Compared with the results in Fig. 5.29 (a), the performance of the red channel drops significantly. The rank-one identification accuracy drops to 27%, even worse than the rank-one identification accuracy of 28% from the first optical model. However, the red channel still performs better in other regions. The rank-10% identification accuracy of the second model is 76.6%, which is 29.9% higher than that of the red channel and 38% higher than that of the first model. Table 5.3 lists the rank-one and rank-10% identification accuracies.

The experimental results in Fig. 5.29 and Table 5.3 demonstrate that (1) vein patterns can be uncovered from color images using information of optics and skin biophysics; and (2) the proposed schemes of parameter range optimization and automatic adjustment of image intensity to address limits in the first model are effective. Although the identification accuracy of the first optical model is worse than the red channel, it is difficult to conclude that the first model is not as good as the red channel, because we assume that the illuminant can be accurately estimated. Manually adjusting the illuminant for the first model is very time-consuming and impractical for large databases. Thus, the intensity adjustment scheme is vital for automatic criminal and victim identification based on vein patterns.

We also use a larger testing set containing 500 color images and 500 NIR images from 250 different forearms to evaluate the first and second optical models. The results will be presented together with the third and fourth optical models in chapter 6.

Fig. 5.29 Cumulative match characteristic curves of vein patterns uncovered by the first and second optical models.

(a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Table 5.3 The rank-one and rank-10% identification accuracies from the first and second optical models.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	84%	92%
Red channel	47%	76%
The first optical model	25%	60%
The second optical model	72%	87%

(b)

	rank-one	rank-10%
Red channel	27%	46.7%
The first optical model	28%	38.6%
The second optical model	66%	76.6%

5.3 Summary

In this chapter, we have proposed a vein uncovering approach based on the principles of optics and skin biophysics. Using this approach we developed two optical models to visualize the vein patterns hidden in color images. They model the inverse process of skin color formation in an image and derive biophysical parameters from color images. Because the biophysical parameters of veins differ significantly from those of generic skin, we can uncover the vein patterns from the distribution maps of parameters. The first model has two limits: it does not use skin color to optimize the biophysical parameters and it is sensitive to the power of the illuminant. The second model uses two schemes, the color optimization scheme and the automatic intensity adjustment scheme, to overcome the limits.

We evaluated the first optical model through visual observation and manual matching of vein patterns. Experimental results indicate that vein patterns can be uncovered from color images (evidence images) and matched to NIR images for personal identification in legal cases. The second model was examined through visual observation and automatic vein matching on a database with 300 pairs of color and NIR images collected from the forearms of 150 subjects. The experimental results are encouraging and indicate that the second model performs better than the first model and that the uncovered patterns can potentially be used for automatic criminal and victim identification.

Chapter 6

Reichman Model-Based Approach to Uncover Vein Pattern

In chapter 5, we proposed an optical approach to uncover vein patterns hidden in color images. This approach is based on the radiative transfer equation which quantitatively describes light transport in the human skin. However, at this moment, its exact analytical solution cannot be obtained for complex and multiple scattering media such as human skin [145]. Therefore, in the first two optical models, we use the most popular and simplest solution, the K-M model, to approximate the transport of light in the human skin. In addition, we assume that optical properties of the human skin are determined by the three layers – the stratum corneum, the epidermis, and the dermis, and veins are located in the dermis. The second model utilizes two additional schemes to overcome the limits of the first model.

In this chapter we present two new optical models called the third and fourth optical models to uncover vein patterns from color images. In these two optical models, another approximate solution to the radiative transfer equation is used to replace the K-M model. In the fourth model, one more layer, the hypodermis consisting of adipose and blood vessels is added to the original skin structure with three layers.

In this chapter, sections 6.1 and 6.2 present the third and the fourth optical models, respectively. In each of these three sections we also offer experimental results and discussion. Section 6.3 provides some summary remarks.

6.1 The Third Optical Model

In this section we present the third optical model. We first describe the model and then report the experimental results.

6.1.1 Model Description

In chapter 5, we proposed two optical models which computationally reverse the process of skin color formation. They utilize the K-M model to approximate the transport of light in human skin. However, a comparison study with the Monte Carlo simulation of human skin shows that the K-M model is very sensitive to the K-M coefficients k and s , which are very difficult to be measured accurately for the complexity of skin structure [118]. To overcome this problem, the solution to the radiative transfer equation based on the Schuster-Schartzchild approximation derived by Reichman is employed in the third optical model [118]. When the incidence is normal, the reflectance and transmittance of this solution can be computed by:

$$R = \frac{\omega_0}{2} \frac{3\Gamma + (1 - \Gamma^2)\exp[-(1 + \alpha)\gamma d] - 1 - \Gamma(3 - \Gamma)\exp(-2\alpha\gamma d)}{(1 - \alpha^2)[1 - \Gamma^2\exp(-2\alpha\gamma d)]}, \quad (6.1)$$

$$T = \frac{\omega_0}{2} \frac{3(1 - \Gamma^2)\exp(-\alpha\gamma d) - (3 - \Gamma)\exp(-\gamma d) + \Gamma(3\Gamma - 1)\exp[-(1 + 2\alpha)\gamma d]}{(1 - \alpha^2)[1 - \Gamma^2\exp(-2\alpha\gamma d)]} + \exp(-\gamma d), \quad (6.2)$$

where

$$\gamma = \mu_a + \mu, \quad (6.3)$$

$$\omega_0 = \frac{\mu_a}{\mu_a + \mu_s}, \quad (6.4)$$

$$\alpha = 2\sqrt{1 - \omega_0}, \quad (6.5)$$

$$\Gamma = \frac{2 - \alpha}{2 + \alpha}. \quad (6.6)$$

$\mu_a(\lambda)$ and $\mu_s(\lambda)$ are respectively the absorption and scattering coefficients of a layer at a particular wavelength and d is the thickness of the layer. The advantage of this solution over the K-M model is that it does not rely on the K-M coefficients k and s [119].

The third optical model is still based on the three-layered skin structure, i.e., the stratum corneum, the epidermis, and the dermis, given in subsection 5.1.2. For each of these three layers, we calculate the reflectance and transmittance based on the Reichman's equations (Eqs. (6.1) ~ (6.2)). For the stratum corneum, the absorption coefficient is determined by Eqs. (5.9) ~ (5.11) and the scattering coefficient is determined by Eq. (5.12). For the epidermis, the absorption coefficient is calculated by Eqs. (5.13) ~ (5.16) and the scattering coefficient is also calculated by Eq. (5.12). For the dermis, the absorption coefficient is provided in Eqs. (5.17) ~ (5.21) and the scattering coefficient is given in Eq. (5.12). Then the reflectance $R_{123}(\lambda)$ and transmittance $T_{123}(\lambda)$ of the whole skin can be computed through Eqs. (5.5) ~ (5.8). In other words, the entire computational processes of the second model and the third model are the same except that the K-M model in the second model is replaced with the Reichman's equations.

With the same models of illuminants and sensors in digital cameras discussed in chapter 5 (Eqs. (5.22) ~ (5.24)), the forward process of skin color formation based on the third optical model can be simulated. The parameters $\mathcal{G}_m$, $\mathcal{G}_p$ and d^{der} are changed exhaustively and the corresponding color values are obtained for training a neural network. The color optimization scheme proposed in chapter 5 is also used here to derive the range of the biophysical parameters. Same as the first and second models, a three-layered feed-forward neural network is used to approximate the inversion of skin color formation. To visualize vein patterns in a testing image, its RGB values are inputted to the trained neural network and the outputs form three distribution maps, $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} , where vein patterns can be observed. To compensate the illumination variation, the scheme of automatic image intensity adjustment proposed in chapter 4 is also adopted.

6.1.2 Experimental Evaluations

This subsection reports experimental results from visual comparison and automatic matching. The third optical model is also compared with the first and second optical models.

6.1.2.1 Evaluation by Visual Observation

As with the previous evaluations, the corresponding color and NIR skin images in our database is used as testing images. Figs. 6.1 ~ 6.3 show uncovering results from the first, second, and third optical models. Figs. 6.4 ~ 6.6 show color images taken from different viewing angles and poses, and the corresponding uncovering results from the optical models. NIR images of the corresponding body positions are also given as ground truth. Visual comparison demonstrates that the distribution maps uncovered by the first model contain a lot of noise. For the images with viewing angle and pose variations (Figs. 6.5 ~ 6.6), the first model almost cannot uncover any vein patterns. Comparing the performance between the second and third models, we can find that the distribution maps uncovered by the third model have higher contrast than those from the second model. Thus, it is easier for human vision to detect the vein patterns obtained by the third model.

Fig. 6.1 Visual evaluation of the third optical model (example 1).

(a) is a color image of the left leg of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

Fig. 6.2 Visual evaluation of the third optical model (example 2).

(a) is a color image of the outer left thigh of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

Fig. 6.3 Visual evaluation of the third optical model (example 3).

(a) is a color image of the inner right thigh and leg of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

Fig. 6.4 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 1).

(a) is a color image of the two legs of a male; (b) and (c) are the related NIR images of the legs and foot; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

Fig. 6.5 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 2).

(a) is a color image of the left leg and right foot of a male; (b) and (c) are the related NIR images of the leg and foot; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

Fig. 6.6 Visual evaluation of the third optical model for images taken from different viewing angles and poses (example 3).

(a) is a color image of the chest and right upper arm of a male; (b) and (c) are the related NIR images of the chest and arm; (d), (e), and (f) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model.

6.1.2.2 Evaluation by Automatic Vein Matching

This subsection reports the experimental results of the third optical model from the automatic matching algorithm. A large testing set containing 500 color images and 500 NIR images from 250 different forearms is employed in this evaluation. The images are collected on two separate occasions, at an interval of around two weeks. Figs. 6.7 ~ 6.9 show some vein patterns extracted from the distribution maps of the depth of dermis obtained by the three optical models. We can see that the vein patterns uncovered by the third optical model are more complete than the first two models. As with the experiments in chapter 4, two sets of experiments are performed. One is from matching images of same type and the other is from matching images of different types. The vein patterns from the red channel are also matched for comparison. Fig. 6.10 reports the cumulative match characteristic curves from the experiments. Table 6.1 lists the rank-one and rank-10% identification accuracies. In both experimental settings, the first optical model has the worst matching results. For the same-type-image matching, the rank-one identification accuracy of the third optical model is 61.6%, which is 1.6% higher than the second model, while the rank-10% identification accuracy of the third optical model is 0.8% lower than the second model. The different-type-image matching results are similar. The rank-one identification accuracy of the third optical model is 45.6%, which is 2% higher than the second model, while the rank-10% identification accuracy of the third optical model is 1.6% lower than the second model. Overall speaking, the matching performance of the third optical model is a little better than the second optical model.

Fig. 6.7 Comparison between the extracted vein patterns from the first, second, and third optical models (example 1).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (d), (g), and (j) are respectively the enhanced vein images obtained from (c), (f), and (i); (e), (h), and (k) are respectively the vein pattern images extracted from (d), (g), and (j).

Fig. 6.8 Comparison between the extracted vein patterns from the first, second, and third optical models (example 2).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (d), (g), and (j) are respectively the enhanced vein images obtained from (c), (f), and (i); (e), (h), and (k) are respectively the vein pattern images extracted from (d), (g), and (j).

Fig. 6.9 Comparison between the extracted vein patterns from the first, second, and third optical models (example 3).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (d), (g), and (j) are respectively the enhanced vein images obtained from (c), (f), and (i); (e), (h), and (k) are respectively the vein pattern images extracted from (d), (g), and (j).

Fig. 6.10 Cumulative match characteristic curves of vein patterns uncovered by the first, second, and third optical models.

(a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Table 6.1 The rank-one and rank-10% identification accuracies from the first, second, and third optical models.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	70.4%	82.4%
Red channel	45.2%	72.4%
The first optical model	29.6%	46.4%
The second optical model	60%	77.2%
The third optical model	61.6%	76.4%

(b)

	rank-one	rank-10%
Red channel	26.8%	45.2%
The first optical model	22.4%	35.2%
The second optical model	43.6%	60.4%
The third optical model	45.6%	58%

6.2 The Fourth Optical Model

In this section we present the fourth optical model. We first describe the model and then report the experimental results.

6.2.1 Model Description

In the first three optical models, we assume that the optical properties of human skin are determined by three layers – the stratum corneum, the epidermis, and the dermis, and veins are located in the dermis. However, skin anatomy shows that on many occasions veins are located under the dermis and inside the hypodermis [128]. In our fourth optical model, one more skin layer, the hypodermis consisting of adipose and blood vessels, is added. Thus, a new skin structure has totally four layers – the stratum corneum, the epidermis, the dermis, and the hypodermis. Using the recursive equations (Eqs. (2.7) and (2.8)) and the four-layered skin

structure, the total reflectance $R_{1234}(\lambda)$ and transmittance $T_{1234}(\lambda)$ of the new skin structure can be computed from

$$R_{1234}(\lambda) = R_{123}(\lambda) + \frac{T_{123}(\lambda)^2 R_4(\lambda)}{1 - R_{123}(\lambda) R_4(\lambda)}, \quad (6.7)$$

$$T_{1234}(\lambda) = \frac{T_{123}(\lambda) T_4(\lambda)}{1 - R_{123}(\lambda) R_4(\lambda)}, \quad (6.8)$$

where

$$R_{123}(\lambda) = R_{12}(\lambda) + \frac{T_{12}(\lambda)^2 R_3(\lambda)}{1 - R_{12}(\lambda) R_3(\lambda)}, \quad (6.9)$$

$$T_{123}(\lambda) = \frac{T_{12}(\lambda) T_3(\lambda)}{1 - R_{12}(\lambda) R_3(\lambda)}, \quad (6.10)$$

$$R_{12}(\lambda) = R_1(\lambda) + \frac{T_1(\lambda)^2 R_2(\lambda)}{1 - R_1(\lambda) R_2(\lambda)}, \quad (6.11)$$

$$T_{12}(\lambda) = \frac{T_1(\lambda) T_2(\lambda)}{1 - R_1(\lambda) R_2(\lambda)}. \quad (6.12)$$

The variables $R_i(\lambda)$ and $T_i(\lambda)$ ($i = 1, 2, 3, 4$) in the equations are respectively the reflectance and transmittance of the i^{th} layer of the skin. The 1st, 2nd, 3rd, and 4th layers are respectively the stratum corneum, the epidermis, the dermis, and the hypodermis. The Reichman's solution to the radiative transfer equation (Eqs. (6.1) ~ (6.6)) is still used to calculate the reflectance and transmittance of each layer. The calculations of the absorption and scattering coefficients of the stratum corneum, the epidermis, and the dermis are the same as the third optical model. In this model, it is assumed that the hypodermis is composed of adipose and veins, and veins are located just under the dermis. The spectra of the absorption and scattering coefficients of the subcutaneous adipose tissue are obtained from the experimental data [120]. Because the major pigments in veins are oxy-hemoglobin and deoxy-hemoglobin, Eqs. (5.18) ~ (5.19) are used to calculate the absorption

coefficient of veins. The scattering coefficient of veins is obtained from the experimental data [121].

With the same models of illuminants and sensors in digital cameras as the previous optical models, Eqs. (5.22) ~ (5.24) can be used to simulate the forward process of skin color formation for this four-layered model. In addition to $\mathcal{G}_m$, $\mathcal{G}_p$, and d^{der} , two more parameters, d^v and d^a are added to represent the depth of veins and adipose, respectively. For the sake of mathematical convenience, d^v and d^a are combined and represented by one symbol, d^{hyp} which is called the depth of the hypodermis. Let the variation range of d^v be $[0 \ a]$ and the variation range of d^a be $[0 \ b]$. The variation range of d^{hyp} is set to be $[-a \ b]$. When d^{hyp} is less than zero, its absolute value representing d^v is used to calculate the reflectance and transmittance of veins; otherwise it represents d^a and can be used directly to calculate the reflectance and transmittance of adipose. This model assumes that veins and adipose cannot co-exist in the same column in the fourth layer. $\mathcal{G}_m$, $\mathcal{G}_p$, d^{der} , and d^{hyp} are changed exhaustively and the corresponding RGB values can be obtained. The color optimization scheme is also used here to derive the range of the biophysical parameters. A three-layered feed-forward neural network is still employed to approximate the inversion of skin color formation. Fig. 6.11 shows the schematic diagram of the fourth optical model. In the resultant distribution map of d^{hyp} , pixels with lower intensity values indicate higher probability of being on veins. To compensate the illumination problem, the scheme of automatic image intensity adjustment is still adopted.

Fig. 6.11 A schematic diagram of the fourth optical model to uncover vein patterns.

The left block represents the skin color formation model and the right block represents the inversion of the skin color formation to obtain $[\mathcal{G}_m, \mathcal{G}_p, d^{der}, d^{hyp}]$ from $[R, G, B]$.

6.2.2 Experimental Evaluations

This subsection reports experimental results of the fourth optical model from visual comparison and automatic matching. The fourth optical model is also compared with the other optical models.

6.2.2.1 Evaluation by Visual Observation

As with the previous evaluations, the corresponding color and NIR skin images in our database are used as testing images. Figs. 6.12 ~ 6.15 show the results from the four optical models for visual comparison. The original color images in Figs. 6.12 ~ 6.13 are taken from standard poses, while the original images in Figs. 6.14 ~ 6.15 are taken from different viewing angles and poses. NIR images of the corresponding body positions are also given for validation. For the fourth optical model, we can observe veins from the distribution maps of the depths of dermis and hypodermis. However, the distribution maps of melanin and hemoglobin do not provide much information about vein patterns. It should be mentioned that the distribution map of hemoglobin represents the concentration of hemoglobin in the dermis. In the four-layered skin structure, veins are assumed to be located under the dermis and inside the hypodermis. Therefore, the distribution of hemoglobin in the dermis does not show the location of veins. Visual comparison demonstrates that the vein patterns uncovered by the fourth model are less noisy than the other optical models.

Fig. 6.12 Visual evaluation of the fourth optical model (example 1).

(a) is a color image of the inner right leg of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model; (m), (n), (o), and (p) are respectively the distribution maps of melanin, hemoglobin, the depth of dermis, and the depth of hypodermis obtained from the fourth optical model.

Fig. 6.13 Visual evaluation of the fourth optical model (example 2).

(a) is a color image of the outer left forearm of a male; (b) is its skin image; (c) is its NIR image; (d), (e), and (f) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model; (m), (n), (o), and (p) are respectively the distribution maps of melanin, hemoglobin, the depth of dermis, and the depth of hypodermis obtained from the fourth optical model.

Fig. 6.14 Visual evaluation of the fourth optical model for images taken from different viewing angles and poses (example 1).

(a) is a color image of the left leg and foot of a male; (b) and (c) are the corresponding NIR images; (d), (e), and (f) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h), and (i) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k), and (l) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model; (m), (n), (o), and (p) are respectively the distribution maps of melanin, hemoglobin, the depth of dermis, and the depth of hypodermis obtained from the fourth optical model.

Fig. 6.15 Visual evaluation of the fourth optical model for images taken from different viewing angles and poses (example 2).

(a) is a color image of the two legs and feet of a male; (b) and (c) are the corresponding NIR images; (d), (e) and (f) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the first optical model; (g), (h) and (i) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the second optical model; (j), (k) and (l) are respectively the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the third optical model; (m), (n), (o) and (p) are respectively the distribution maps of melanin, hemoglobin, the depth of dermis, and the depth of hypodermis obtained from the fourth optical model.

6.2.2.2 Evaluation by Automatic Vein Matching

In this subsection, the automatic matching algorithm is used to evaluate the vein patterns uncovered by the fourth optical model. As with the experimental setting for the third model, the fourth optical model is examined on the 500 color images and the 500 NIR images from the 250 different forearms. For the first three optical models, the vein patterns are extracted from the distribution map of the depth of dermis, while for the fourth optical model, the vein patterns are extracted from the distribution map of the depth of hypodermis. Figs. 6.16 ~ 6.18 show some extracted vein patterns obtained by the four models. We can see that the overall vein patterns extracted by the fourth model are similar to those from the second and the third models. However, in some skin areas, the patterns extracted by different models are supplementary to one another. Fig. 6.19 shows the cumulative match characteristics curves from the experiments. Table 6.2 lists their rank-one and rank-10% identification accuracies. For the same-type-image matching, the rank-one identification accuracy of the fourth optical model is almost the same as the third model, but higher than the second model. However, for the different-type-image matching, the rank-one and rank-10% identification accuracies of the fourth optical model are lower than the second and third models. Overall speaking, the fourth optical model has comparable performance with the second and third models. One possible reason why the fourth optical model does not make significant improvement in terms of automatic matching results is that veins can occur in both dermis and hypodermis. Therefore, our assumption about vein location may not be applicable to all skin images. Visual evaluation demonstrates that the vein patterns uncovered by the fourth optical model can be supplementary to the other models. This property will help us design the fusion method in the next chapter for further enhancing the matching performance.

Fig. 6.16 Comparison between the extracted vein patterns from the four optical models (example 1). (a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (l) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (g), (j), and (m) are respectively the enhanced vein images obtained from (c), (f), (i), and (l); (e), (h), (k), and (n) are respectively the vein pattern images extracted from (d), (g), (j), and (m).

Fig. 6.17 Comparison between the extracted vein patterns from the four optical models (example 2). (a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (l) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (g), (j), and (m) are respectively the enhanced vein images obtained from (c), (f), (i), and (l); (e), (h), (k), and (n) are respectively the vein pattern images extracted from (d), (g), (j), and (m).

Fig. 6.18 Comparison between the extracted vein patterns from the four optical models (example 3).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c), (f), and (i) are respectively the distribution maps of the depth of dermis obtained by the first, second, and third optical models; (l) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (g), (j), and (m) are respectively the enhanced vein images obtained from (c), (f), (i), and (l); (e), (h), (k), and (n) are respectively the vein pattern images extracted from (d), (g), (j), and (m).

Fig. 6.19 Cumulative match characteristic curves of vein patterns uncovered by the four optical models. (a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Table 6.2 The rank-one and rank-10% identification accuracies from the four optical models.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	70.4%	82.4%
Red channel	45.2%	72.4%
The first optical model	29.6%	46.4%
The second optical model	60%	77.2%
The third optical model	61.6%	76.4%
The fourth optical model	61.2%	76%

(b)

	rank-one	rank-10%
Red channel	26.8%	45.2%
The first optical model	22.4%	35.2%
The second optical model	43.6%	60.4%
The third optical model	45.6%	58%
The fourth optical model	38.8%	55.2%

6.3 Summary

In this chapter we developed two additional optical models, which are called the third and fourth optical models. In both of the models, the K-M model is replaced with another approximate solution to the radiative transfer equation. In the fourth model, a four-layered skin structure with a layer of hypodermis was proposed. Experimental results show that the third and fourth optical models have comparable performance with the second model. Visual evaluation demonstrates that the vein patterns uncovered by the two new models can be supplementary to the second model. This property will help us design the fusion method in the next chapter for further enhancing the matching performance.

Chapter 7

Fusion of Models to Uncover Vein Pattern

In chapters 5 and 6, we have developed four optical models to uncover vein patterns from color skin images. They are based on two different solutions to the radiative transfer equation and two different skin structure models. However, all these models simulate light transport in the skin approximately. Because different persons and different body sites have diverse skin properties, it is not expected that the single optical model approach can handle all types of skin. In this chapter we propose a multi-optical-model fusion method. To better use information provided by different optical models, the vein patterns obtained from the second, third, and fourth models are fused. As with the organization of the previous chapters, in this chapter, we will first describe the fusion method and then provide a comprehensive evaluation of all optical models and the RGB-NIR mapping approach that we proposed in this thesis.

7.1 Description of the Fusion Method

The experimental results in chapters 5 and 6 show that all the four optical models provide partial information about vein patterns. However, due to the complex skin structure and the imperfect solutions to the radiative transfer equation, none of the models can provide exact and complete vein patterns. Visual comparison in the previous chapters shows that the models can supplement one another. Therefore, the optical models should be used together. Our proposed fusion method uses the union of the sampled points as a final vein pattern for matching. The experimental results in chapters 5 and 6 show that the uncovering performance of the second, third, and fourth optical models are similar, but that of the first optical model is much worse. Thus, the proposed method only fuses the vein patterns uncovered by the second, third, and fourth optical models. Figs. 7.1 ~ 7.3 give some extracted vein patterns and their fusion results. In each of the figures, (c), (d), and

(e) are respectively the uncovered results from the second, third, and fourth optical models. More precisely, (c) and (d) are respectively the distribution maps of the depth of dermis uncovered by the second and third optical models, and (e) is the distribution map of the depth of hypodermis uncovered by the fourth optical model. (f), (g), and (h) are point representations of the vein patterns extracted from the corresponding uncovered results. The symbol “2 + 3” denotes that vein patterns from the second and third models are fused. The other symbols “2 + 4” and “2 + 3 + 4” are defined in the same way. (i), (j), and (k) are respectively the fusion results of “2 + 3”, “2 + 4”, and “2 + 3 + 4”. It is easy to note that the fusion results have more complete vein patterns than those from any of the single models, and are more close to the vein patterns in the NIR images.

Fig. 7.1 The extracted vein patterns and the fusion result (example 1).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (e) are respectively the distribution maps of the depth of dermis obtained by the second and third optical models; (g) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (f), and (h) are respectively the point maps extracted from the vein patterns contained in (c), (e), and (g); (i) is the fusion result of (d) and (f); (j) is the fusion result of (d) and (h); (k) is the fusion result of (d), (f) and (h).

Fig. 7.2 The extracted vein patterns and the fusion result (example 2).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (e) are respectively the distribution maps of the depth of dermis obtained by the second and third optical models; (g) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (f), and (h) are respectively the point maps extracted from the vein patterns contained in (c), (e) and (g); (i) is the fusion result of (d) and (f); (j) is the fusion result of (d) and (h); (k) is the fusion result of (d), (f) and (h).

Fig. 7.3 The extracted vein patterns and the fusion result (example 3).

(a) is a color skin image of a right forearm; (b) is its corresponding NIR image; (c) and (e) are respectively the distribution maps of the depth of dermis obtained by the second and third optical models; (g) is the distribution map of the depth of hypodermis obtained by the fourth optical model; (d), (f), and (h) are respectively the point maps extracted from the vein patterns contained in (c), (e) and (g); (i) is the fusion result of (d) and (f); (j) is the fusion result of (d) and (h); (k) is the fusion result of (d), (f) and (h).

7.2 Experimental Evaluations

As with the previous sections, the automatic matching algorithm is used for performance evaluation. Here we provide a comprehensive evaluation of all the uncovering models that we proposed in this thesis. We use the testing set with the 250 different forearms. Two sets of experiments are performed: same-type-image and different-type-image matching. Fig. 7.4 shows the cumulative match characteristic curves from the experiments. For clearness, the figure only shows the matching curves from the fusion method. The rank-one and rank-10% identification accuracies of the RGB-NIR mapping approach, the single optical models, and the multi-optical-model fusion method are listed in Table 7.1. It should be mentioned that the RGB-NIR mapping approach is actually evaluated on the testing set with 249 different forearms because images of one subject is used to train the neural network.

The results from the same-type-image matching show that the highest rank-one identification accuracy from the single models is 61.6%, while that from the fusion of the second and third models is 67.2%, and those from the fusions of the second and fourth models and of all three models are both 72.4%, which are better than the rank-one identification accuracy of 67.9% from the RGB-NIR mapping approach and even 2% higher than the rank-one identification of 70.4% from matching NIR images.

The rank-10% identification accuracy from the fusion of the second and third models is 80.8% and the rank-10% identification accuracy from the fusion of the second and fourth models is 84.4%. Both are higher than any of the single models and the rank-10% identification accuracy from the fusion of the second and fourth models is 2.9% higher than that from the RGB-NIR mapping algorithm. The rank-10% identification accuracy from the fusion of all three models is 87.2%, which is 4.8% higher than that from matching NIR images with NIR images. The results from different image types matching are also encouraging. The rank-one and rank-10% identification accuracies from the fusion of all models are 48% and 67.6%, respectively. Both are much higher than any single models and the RGB-NIR mapping approach.

The experimental results in Fig. 7.4 and Table 7.1 demonstrate that the fusion scheme is effective. Its automatic matching results are better than any of the single models and also the RGB-NIR mapping approach. The vein pattern information given by the fusion scheme is even more complete than that in NIR images, which are usually considered as ground truth of vein patterns. Comparing the fusion results, we can find that the fusion of all models offers the best performance, which further verifies that the uncovering results from different optical models are supplementary to one another.

Fig. 7.4 Cumulative match characteristic curves of vein patterns obtained by the multi-optical-model fusion method.

(a) from the same type of images (color vs. color and NIR vs. NIR); (b) from different types of images (color vs. NIR).

Table 7.1 The rank-one and rank-10% identification accuracies from different uncovering models.

(a) from the same-type-image matching and (b) from different-type-image matching.

(a)

	rank-one	rank-10%
NIR images	70.4%	82.4%
Red channel	45.2%	72.4%
The first optical model	29.6%	46.4%
The second optical model	60%	77.2%
The third optical model	61.6%	76.4%
The fourth optical model	61.2%	76%
RGB-NIR mapping approach	67.9%	81.5%
Fusion of models “2 + 3”	67.2%	80.8%
Fusion of models “2 + 4”	72.4%	84.4%
Fusion of models “2 + 3 + 4”	72.4%	87.2%

(b)

	rank-one	rank-10%
Red channel	26.8%	45.2%
The first optical model	22.4%	35.2%
The second optical model	43.6%	60.4%
The third optical model	45.6%	58%
The fourth optical model	38.8%	55.2%
RGB-NIR mapping approach	47.8%	66.7%
Fusion of models “2 + 3”	47.6%	58.4%
Fusion of models “2 + 4”	44.4%	58.4%
Fusion of models “2 + 3 + 4”	48%	67.6%

7.3 Summary

In this chapter we have proposed a multi-optical-model fusion method to uncover vein patterns from color skin images. In chapters 5 and 6, we proposed four optical models. All the models provide some information about vein patterns. Due to the complex skin structure and the approximate solutions to the radiative transfer equation, none of the models can provide complete vein patterns, but they supplement one another. Therefore, we developed a multi-optical-model fusion method to combine the vein patterns uncovered from different models. Because of the poor performance of the first model, it was excluded from the fusion. Visual evaluations show that the fusion results contain more complete vein patterns than any results from the single models. The proposed multi-optical-model fusion method was examined on a testing set with 500 pairs of color and NIR images collected from the forearms of 250 subjects. The automatic matching results of the fusion method were much better than those from any single optical model and also the RGB-NIR mapping approach, and even superior to the result from matching NIR images with NIR images.

Chapter 8

Conclusion

In this chapter, we will summarize the main contributions of this thesis. We will also discuss the limits of the proposed techniques and suggest possible directions for future work.

8.1 Summary of the Main Contributions

Traditional vein recognition systems rely on either hand or wrist veins in high quality IR images. In this research, we extended applications of vein patterns to forensic investigation, where no IR images are provided. We proposed a new recognition scheme illustrated in Fig. 1.4 to extract body vein patterns from color images collected from crime scenes for criminal and victim identification. In this thesis, we concentrated on developing algorithms for removing JPEG blocking artifacts especially for skin images and uncovering vein patterns from color images. We presented a set of algorithms to address the related problems. We developed two approaches for removing JPEG blocking artifacts in skin images and another two approaches for uncovering vein patterns hidden in RGB images. We also constructed a large skin image database.

To remove blocking artifacts in JPEG-compressed skin images, we developed two specific approaches. The first one is a maximum-a-posteriori (MAP)-based approach which formulates skin image deblocking as an estimation problem and embeds statistical information of skin images into a MAP model to perform the estimation. The second one is a knowledge-based approach which extracts prior knowledge of skin images from a training set and uses it to infer original blocks in compressed evidence images. Two inference schemes, a block synthesis algorithm, and an indexing mechanism were also proposed for this approach. Both of the approaches extensively use prior knowledge of skin images and guarantee that the resultant and

compressed images have the same quantized DCT coefficients. Experimental results demonstrate that the proposed approaches perform better than other traditional methods.

In addition, we proposed two approaches for uncovering vein patterns hidden in color skin images. The first approach of vein uncovering is based on RGB-NIR mapping. It extracts information from a pair of synchronized color and NIR images, and uses a neural network (NN) to map RGB values to NIR intensities. We designed an automatic intensity adjustment scheme for illumination compensation. We also proposed an NN weight adjustment scheme to improve the robustness of the approach. Using an automatic matching algorithm, we matched resultant images from the RGB-NIR mapping approach and found that its matching result is comparable to the result from matching NIR images.

The second approach for uncovering vein patterns is based on the principles of optics and skin biophysics. It models the inverse of the process of the skin color formation in an image and derives corresponding biophysical parameters. Although pixels of veins and pixels of generic skin have very similar RGB values, their biophysical parameters differ significantly. Therefore, we can observe vein patterns from their spatial distributions. Based on this approach we developed four optical models for the process of skin color formation. They are all based on the radiative transfer equation which quantitatively describes light transport in tissue (e.g. human skin). However, at this moment, its exact analytical solution has not yet been obtained for complex and multiple scattering media such as human skin. Therefore, in our first and second optical models, we used the most popular and simplest solution, the K-M model, to simulate the transport of light in human skin. These two models assume that optical properties of human skin are determined by three layers – the stratum corneum, the epidermis, and the dermis, and veins are located in the dermis. The first optical model has two limits: 1) it does not use skin color to optimize the biophysical parameters, and 2) it is sensitive to the power of the illuminant. To overcome these limits, the second optical model uses a color optimization scheme to derive the ranges of the biophysical parameters. It also adopts the automatic intensity adjustment scheme to improve the robustness of the uncovering performance.

Our third and fourth optical models give up the K-M model and use Reichman's solution to the radiative transfer equation. In our fourth optical model, we made improvement on modeling skin structure. Skin anatomy shows that on many occasions veins are located under the dermis and inside the hypodermis. In this optical model, we added one more layer, the hypodermis consisting of adipose and blood vessels to the skin structure. All the optical models provide information about vein patterns. However, due to the complex skin structure and the approximate solutions to the radiative transfer equation, none of the models can offer exact and complete vein patterns. Visual evaluations show that the models can supplement one another. Therefore, we proposed a method to fuse the uncovering vein patterns from different models. Experimental evaluations show that the fusion results are much better than any of the single optical models and also the RGB-NIR mapping approach. Its matching result is even better than matching NIR images, which are always considered as ground truth of vein patterns.

For algorithm development and evaluation, we constructed a large database containing corresponding RGB and NIR skin images. We also built a data acquisition system constituted by a digital single-lens reflex camera (Nikon D70s or Canon 500D) and a multi-spectral 2-CCD camera (JAI's AD-080 CL) and developed an image capture software package for data collection. The current database contains images contributed by participants with different BMI from diverse racial, age, and occupation groups.

8.2 Future Work

The algorithms proposed in this thesis have been proven effective in many cases. However, future research is still demanded for enhancing robustness of the algorithms and for increasing capability to handle lower quality images from crime scenes.

The proposed optical uncovering approach is based on the radiative transfer equation to simulate the formation of skin color in an image. Our four optical models use two approximate

solutions to the equation: the K-M model and Reichman's solution. Both of them assume that the incidence is normal to the skin surface, which cannot be satisfied under many circumstances. Although Reichman's solution has some improvement in terms of accuracy, it is still an approximate solution to the radiative transfer equation. More accurate methods are needed to achieve more realistic simulation of skin color formation. The Monte Carlo (MC) method [129]-[132], the auxiliary function (AF) method [133]-[136], and the discrete-ordinate (DO) method [137]-[139] are some possible choices. The MC is a statistical method, while the other two are quasi-analytic methods [140]. In addition, due to the complex structure of human skin, our current skin optical models are still very rough and many biophysical parameters are assumed to be constants. We hope to construct a more realistic and flexible optical model for skin. What's more, in this thesis, we used neural networks to approximate the inversion of skin color formation. Although neural networks have good extrapolation ability, their training results are sensitive to the initial weights. Therefore, more computational methods need to be investigated for practical solutions of the color inverse problem. In recent years, differential evolution has been applied in the field of electromagnetic inverse scattering problems [141]-[143]. It may also find applications in the color inverse problem. Furthermore, our optical models work independently for each pixel, which results in grainy and noisy appearance of some uncovered images. Local contextual information of skin images should be used to improve the approach in future work.

Illumination is a very important factor that affects skin color in an image. The proposed uncovering algorithms assume that the illumination is evenly distributed on skin. This requirement cannot be satisfied under some circumstances. Variations in the spatial and/or spectrum distributions of illumination produce a change in skin color. For images taken from different viewing angles and poses, this problem becomes more serious. Although the image intensity adjustment scheme that we proposed alleviates this problem to some extent, it does not solve the problem completely. More comprehensive methods are necessary for illumination compensation.

The proposed RGB-NIR mapping approach tackles the uncovering problem in a completely different way. It extracts prior information from synchronized RGB and NIR images and maps RGB values to NIR intensities. Currently we utilize only one pair of synchronized images to construct one mapping model. Although robust performance has been demonstrated, for better uncovering performance, more information from our skin image database should be captured. We hope to construct a batch of mapping models based on synchronized images with different characteristics. Given a testing image, the most appropriate model(s) will be found to give uncovering results.

Our knowledge-based deblocking approach searches uncompressed blocks from a skin image database to recover compressed skin images. It utilizes only local dependence within each component (Y, U, and V) to perform deblocking. To further improve its performance, we hope to use more features with long-range dependency and cross-component relationship into the deblocking procedure.

References

- [1] M. Motivans, T. Kyckelhahn, "Federal Prosecution of Child Sex Exploitation Offenders, 2006", Bureau of Justice Statistics, December, 2007
- [2] Canada's National Tipline for Reporting the Online Sexual Exploitation of Children.
- [3] BBC, International child porn ring smashed 26 March 2001, <http://news.bbc.co.uk/1/hi/world/americas/1244457.stm>.
- [4] Yahoo! News, May 29 2008, "LA Jury Convicts Ex-Marine in Sex Tourism Case", http://news.yahoo.com/s/ap/20080530/ap_on_re_us/sex_tourism_1.
- [5] A.K. Jain, J. Lee, R. Jin, "Tattoo-ID: Automatic Tattoo Image Retrieval for Suspect & Victim Identification", in *Proceedings of Pacific-Rim Conference on Multimedia*, pp. 256-265, 2007.
- [6] A. Nurhudatiana, A. Kong, K. Matinpour, S. Cho, N. Craft, "Fundamental Statistics of Relatively Permanent Pigmented or Vascular Skin Marks", in *Proceedings of 2011 International Joint Biometrics Conference*, pp. 1-6, 2011.
- [7] "Vein Recognition Biometrics", <http://www.findbiometrics.com/vein-recognition/>.
- [8] L. Wang, G. Leedham, S.-Y. Cho, "Infrared Imaging of Hand Vein Patterns for Biometric Purposes", *IET Computer Vision*, vol. 1 (3-4), pp. 113-122, 2007.
- [9] "Turkish biometric ATM plans are largest in Europe", <http://www.planetbiometrics.com/article-details/i/906/>.
- [10] T. Igarashi, K. Nishino, S.K. Nayar, "The Appearance of Human Skin: A Survey", *Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 1, pp. 1-95, 2007.
- [11] G. Poirier, "Human Skin Modeling and Rendering", Master Thesis, University of Waterloo, 2003.
- [12] C. Brandorff, "Rendering Human Skin", Project Report, Vienne University of Technology, 2006.
- [13] A. Krishnaswamy, "BioSpec: A Biophysically-Based Spectral Model of Light Interaction with Human Skin", Master Thesis, University of Waterloo, 2005.
- [14] "Skin Biology and Structure", <http://www.mydr.com.au/default.asp?Article=3718>, 2002.
- [15] "Subcutaneous tissue", http://en.wikipedia.org/wiki/Subcutaneous_tissue.
- [16] G.K. Wallace, "The JPEG Still Picture Compression Standard", *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii-xxxiv, 1992.
- [17] M.-Y. Shen, C.-C. Kuo, "Review of Postprocessing Techniques for Compression Artifact Removal", *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2-14, 1998.
- [18] T.D. Tran, J. Liang, C. Tu, "Lapped Transform via Time-Domain Pre- and Post-Filtering", *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1557-1571, 2003.
- [19] J. Liang, C. Tu, T.D. Tran, "Optimal Block Boundary Pre/postfiltering for Wavelet-based Image and Video Compression", *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2151-2158, 2005.
- [20] B. Macq, "Weighted Optimum Bit Allocation to Orthogonal Transforms for Picture Coding", *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 5, pp. 875-883, 1992.
- [21] N. Nill, "A Visual Model Weighted Cosine Transform for Image Compression and Quality Assessment", *IEEE Transactions on Communications*, vol. 33, no. 6, pp. 551-557, 1985.
- [22] A.M. Bruckstein, M. Elad, R. Kimmel, "Down-Sampling for Better Transform Compression", *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1132-1144, 2003.

- [23] C.-L. Chang, D. Makaroff, B. Girod, S. Tsai, "Direction-Adaptive Partitioned Block Transform for Color Image Coding", *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1740-1755, 2010.
- [24] K. Yamatani, N. Saito, "Improvement of DCT-Based Compression Algorithms Using Poisson's Equation", *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3672-3689, 2006.
- [25] J. Wang, J. Kim, J. Jeong, "New Adaptive Postprocessing to Remove Blocking Artifacts for DCT Coded Image", in *Proceedings of IEEE International Conference on Signal Image Technology and Internet Based Systems*, pp. 352-358, 2008.
- [26] Y.-K. Lai, J. Li, C.-C. J. Kuo, "Image Enhancement for Low Bit Rate JPEC and MPEG Coding via Postprocessing", in *Proceedings of SPIE: Visual Communication and Image Processing '96*, pp. 1484-1494, 1996.
- [27] H.C. Reeves, J.S. Lim, "Reduction of Blocking Effects in Image Coding", *Optical Engineering*, vol. 23, pp. 34-37, 1984.
- [28] B. Ramamurthi, A. Gersho, "Nonlinear Space-Variant Postprocessing of Block Coded Images", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1258-1268, 1986.
- [29] Y.-F. Hsu, Y.-C. Chen, "A New Adaptive Separable Median Filter for Removing Blocking Effects", *IEEE Transactions on Consumer Electronics*, vol. 39, no. 3, pp. 510-513, 1993.
- [30] Y.L. Lee, H.C. Kim, H.W. Park, "Blocking Effect Reduction of JPEG Images by Signal Adaptive Filtering", *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 229-234, 1998.
- [31] A.Z. Averbuch, A. Schclar, D.L. Donoho, "Deblocking of Block-Transform Compressed Images Using Weighted Sums of Symmetrically Aligned Pixels", *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 200-212, 2005.
- [32] G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, "Efficient Image Deblocking Based on Postfiltering in Shifted Windows", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 122-126, 2008.
- [33] T. Chen, H. Wu, B. Qiu, "Adaptive Postfiltering of Transform Coefficients for The Reduction of Blocking Artifacts", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 5, pp. 594-602, 2001.
- [34] B. Zeng, "Reduction of Blocking Effect in DCT-Coded Images Using Zero-Masking Techniques", *Signal Processing*, vol. 79, no. 2, pp. 205-211, 1999.
- [35] S. Singh, V. Kumar, H.K. Verma, "Reduction of Blocking Artifacts in JPEG Compressed Images", *Digital Signal Processing*, vol. 17, pp. 225-243, 2007.
- [36] Y. Chang, Y. Chen, "Design of a Deblocking Filter for Both Objective and Subjective Metrics", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E91-A, no. 8, pp. 2038-2040, 2008.
- [37] Y. Luo, R.K. Ward, "Removing The Blocking Artifacts of Block-Based DCT Compressed Images", *IEEE Transactions on Image Processing*, vol. 12, no. 7, pp. 838-842, 2003.
- [38] Y. Tsaig, M. Elad, P. Milanfar, G.H. Golub, "Variable Projection for Near-Optimal Filtering in Low Bit-Rate Block Coders", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 1, pp. 154-160, 2005.
- [39] D.T. Vo, T.Q. Nguyen, S. Yea, A. Vetro, "Adaptive Fuzzy Filtering for Artifact Reduction in Compressed Images and Videos", *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1166-1178, 2009.
- [40] K. Kwon, M. Yang, J. Ma, S. Im, D. Lim, "Application of Feedforward Neural Network for the Deblocking of Low Bit Rate Coded Images", in *Proceedings of the 13th International Conference on Artificial Intelligence and Simulation*, pp. 547-555, 2004.
- [41] B. Niss, "Prediction of AC Coefficients from the DC Values", *ISO/IEC JTC1/SC2/WG8 N745*, 1988.

- [42] G. Lakhani, "Improved Equations for JPEG's Blocking Artifacts Reduction Approach", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 6, pp. 930-934, 1997.
- [43] G. Lakhani, "Image Fitting Using Arctan for JPEG AC Coefficient Prediction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 6, pp. 791-802, 2008.
- [44] B. Macq, M. Mattavelli, O.V. Claster, E. van der Plancke, S. Comes, W. Li, "Image Visual Quality Restoration by Cancellation of the Unmasked Noise", in *the Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. V.53-V.56, 1994.
- [45] Y. Nakajima, H. Hori, T. Kanoh, "A Pel Adaptive Reduction of Coding Artifacts for MPEG Video Signals", in *the Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 928-932, 1994,
- [46] S.-W. Hong, Y.-H. Chan, W.-C. Siu, "A Practical Real-Time Postprocessing Technique for Blocking Effect Elimination", in *Proceedings of the IEEE International Conference on Image Processing*, pp. II.21-II.24, 1996.
- [47] Z. Wang, A.C. Bovik, "Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures", *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98-117, 2009.
- [48] R.L. Stevenson, "Reduction of Coding Artifacts in Transform Image Coding", in *the Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 401-404, 1993.
- [49] T.P. O'Rourke, R. Stevenson, "Improved Image Decompression for Reduced Transform Coding Artifacts", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 490-499, 1995.
- [50] J. Luo, C.W. Chen, K.J. Parker, T. Huang, "Artifacts Reduction in Low Bit Rate DCT-Based Image Compression", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 9, pp. 1363-1368, 1996.
- [51] D. Sun, W. Chan, "Postprocessing of Low Bit-Rate Block DCT Coded Images Based on a Fields of Experts Prior", *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2743-2751, 2007.
- [52] T. Özcelik, J.C. Brailean, A.K. Katsaggelos, "Image and Video Compression Algorithms Based on Recovery Techniques Using Mean Field Annealing", in *the Proceedings of the IEEE*, vol. 83, no. 2, pp. 304-315, 1995.
- [53] J. Li, C.-C.J. Kuo, "Coding Artifact Removal with Multiscale Post-Processing", in *the Proceedings of IEEE International Conference on Image Processing*, vol. 1, pp. 45-48, 1997.
- [54] R. Rosenholtz, A. Zakhor, "Iterative Procedures for Reduction of Blocking Effects in Transform Coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 2, no. 1, pp. 91-95, 1992.
- [55] Y. Yang, N. Galatsanos, A.K. Katsaggelos, "Regularized Reconstruction to Reduce Blocking Discrete Cosine Transform Compressed Images", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 3, no. 6, pp. 421-432, 1993.
- [56] C. Weerasinghe, A.W. Liew, H. Yan, "Artifact Reduction in Compressed Images Based on Region Homogeneity Constraints Using the Projection Onto Convex Sets Algorithm", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 891-897, 2002.
- [57] J. Park, D. Park, R.J. Marks, M.A. El-Sharkawi, "Recovery of Image Blocks Using the Method of Alternating Projections", *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 461-474, 2005.
- [58] L. Guo, O.C. Au, M. Ma, X. Fan, H.W. Wong, D.P. Palomar, "Image Deblocking Using Convex Optimization", in *Proceedings of the 15th IEEE International Conference on Image Processing*, pp. 3148-3151, 2008.
- [59] S.K. Naik, C.A. Murthy, "Hue-Preserving Color Image Enhancement without Gamut Problem", *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1591-1598, 2003.

- [60] E. Provenzi, C. Gatta, M. Fierro, A. Rizzi, "A Spatially Variant White-Patch and Gray-World Method for Color Image Enhancement Driven by Local Contrast", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1757-1770, 2008.
- [61] S. Maria, E. Valencia, "Laplacian Filter Based on Color Difference for Image Enhancement", *In Proceedings of the 5th Iberoamerican Meeting on Optics and 8th Latin American Meeting on Optics, Lasers, and Their Applications*, vol. 5622, pp. 1259-1264, 2004.
- [62] J. Astola, P. Haavisto, Y. Nuevo, "The Vector Median Filter", *in Proceedings of the IEEE*, vol. 78, no. 4, pp. 678-689, 1990.
- [63] K.N. Plataniotis, A.N. Venetsanopoulos, "Vector Filtering", *The Color Image Processing Handbook*, S. J. Sangwine and R. E. N. Horne, eds. Chapman and Hall, 1999.
- [64] P. Witkin, "Scale Space Filtering", *in Proceedings of International Conference on Artificial Intelligence*, pp. 1019-1023, 1983.
- [65] P. Perona, J. Malik, "Scale Space and Edge Detection Using Anisotropic Diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.
- [66] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", *in Proceedings of the 6th International Conference on Computer Vision*, pp. 839-846, 1998.
- [67] D.J. Jobson, Z. Rahman, G.A. Woodell, "Properties and Performance of a Center/Surround Retinex", *IEEE Transactions on Image Processing*, vol. 6, no. 3, pp. 451-462, 1997.
- [68] M. Ogata, T. Tsuchiya, T. Kubozono, K. Ueda, "Dynamic Range Compression Based on Illumination Compensation", *IEEE Transactions on Consumer Electronics*, vol. 47, no. 3, pp. 548-558, 2001.
- [69] R. Kimmel, M. Elad, D. Shaked, R. Keshet, I. Sobel, "A Variational Framework for Retinex", *International Journal on Computer Vision*, vol. 52, no. 1, pp. 7-23, 2003.
- [70] D.J. Jobson, Z. Rahman, G.A. Woodell, "A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Scenes", *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965-976, 1997.
- [71] J.V. Stone, "Independent Component Analysis", *Encyclopedia of Statistics in Behavioral Science*, vol. 2, pp. 907-912, 2005.
- [72] A. Oursland, J.D. Paula, N. Mahmood, "Case Studies of Independent Component Analysis", <http://www.oursland.net/tutorials/ica/ica-report.pdf>.
- [73] G.J. Klinker, S.A. Shafer, T. Kanade, "A Physical Approach to Color Image Understanding", *International Journal of Computer Vision*, vol. 4, no. 1, pp. 7-38, 1990.
- [74] N. Tsumura, H. Haneishi, Y. Miyake, "Independent Component Analysis of Skin Color Image", *Journal of Optical Society of America A*, vol. 16, no. 9, pp. 2169-2176, 1999.
- [75] N. Tsumura, H. Haneish, Y. Miyake, "Independent Component Analysis of Spectral Absorbance Image in Human Skin", *Optical Review*, vol. 7, pp. 479-482, 2000.
- [76] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, Y. Miyake, "Image-Based Skin Color and Texture Analysis/Synthesis by Extracting Hemoglobin and Melanin Information in the Skin", *in the Proceedings of the 30th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH 2003*, pp. 770-779, 2003.
- [77] N. Tsumura, K. Uetsuki, N. Ojima, Y. Miyake, "Correlation Map Analysis Between Appearance of Japanese Facial Images and Amount of Melanin and Hemoglobin Components in the Skin", *in the Proceedings of SPIE: Human Vision and Electronic Imaging 6*, vol. 4299, pp. 252-260, 2001.
- [78] N. Tsumura, M. Kawabuchi, H. Haneishi, Y. Miyake, "Mapping Pigmentation in Human Skin from Multi-Channel Visible Spectrum Image by Inverse Optical Scattering Technique", *Journal of Imaging Science and Technology*, vol. 45, no. 5, pp. 444-450, 2001.
- [79] R.R. Anderson, J.A. Parrish, "The Optics of Human Skin", *Journal of Investigative Dermatology*, vol. 77, no. 1, pp. 13-19, 1981.

-
- [80] R.R. Anderson, J. Hu, J.A. Parrish, "Optical Radiation Transfer in the Human Skin and Applications in In Vivo Remittance Spectroscopy", *In Bioengineering and the Skin*, ed. R. Marks and P.A. Payne, Chapter 28, pp. 253–265, MTP Press Limited, 1981.
 - [81] S. Wan, R.R. Anderson, J.A. Parrish, "Analytical Modeling for the Optical Properties of the Skin with In Vitro and In Vivo Applications", *Photochemistry and Photobiology*, vol. 34, no. 4, pp. 493-499, 1981.
 - [82] B. Diffey, "A Mathematical Model for Ultraviolet Optics in Skin", *Physics in Medicine and Biology*, vol. 28, no. 6, pp. 647-657, 1983.
 - [83] M.J.C. Van Gemert, S.L. Jacques, H.J.C.M. Sterenborg, W.M. Star, "Skin Optics", *IEEE Transactions on Biomedical Engineering*, vol. 36, no. 12, pp. 1146-1154, 1989.
 - [84] S. Cotton, E. Claridge, "Developing a Predictive Model of Skin Colouring", *In SPIE Vol. 2708, Medical Imaging 1996*, pp. 814-825, 1996.
 - [85] E. Claridge, S. Cotton, P. Hallc, M. Moncrieffd, "From Colour to Tissue Histology: Physics-Based Interpretation of Images of Pigmented Skin Lesions", *Medical Image Analysis*, no. 7, pp. 489-502, 2003.
 - [86] M. Doi, S. Tominaga, "Spectral Estimation of Human Skin Colour using the Kubelka-Munk Theory", *In SPIE/IS&T Electronic Imaging*, SPIE, vol. 5008, pp. 221-228, 2003.
 - [87] C. Donner, H.W. Jensen, "Light Diffusion in Multi-Layered Translucent Materials", in *SIGGRAPH' 2005*, pp. 1032-1039, 2005.
 - [88] M. Störring, "Computer Vision and Human Skin Color", PhD Thesis, Aalborg University, 2004.
 - [89] S. Geman, D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721-741, 1984.
 - [90] H. Derin, H. Elliott, "Modeling and Segmentation of Noisy and Textured Images from Gibbs Random Fields", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 39-54, 1987.
 - [91] J. Luo, C.W. Chen, K.J. Parker, T.S. Huang, "Artifact Reduction in Low Bit Rate DCT-Based Image Compression", *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1363-1368, 1996.
 - [92] S.C. Zhu, D. Mumford, "Prior Learning and Gibbs Reaction-Diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1236-1250, 1997.
 - [93] T. Meier, K.N. Ngan, G. Crebbin, "Reduction of Blocking Artifacts in Image and Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 3, pp.490-500, 1999.
 - [94] A. Foi, V. Katkovnik, K. Egiazarian, "Pointwise Shape-adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images", *IEEE Transactions on Image Processing*, vol. 16, no. 5, pp. 1395-1411, 2007.
 - [95] J. Chou, M. Crouse, K. Ramchandran, "A Simple Algorithm for Removing Blocking Artifacts in Block-Transform Coded Images", in *Proceedings of the IEEE ICIP*, vol. 1, pp. 377-380, 1998.
 - [96] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
 - [97] W.T. Freeman, E.C. Pasztor, O.T. Carmichael, "Learning Low-Level Vision", *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25-47, 2000.
 - [98] W.T. Freeman, T.R. Jones, E.C. Pasztor, "Example-Based Super-Resolution", *Computer Graphics*, vol. 22, no. 2, pp. 56-65, 2002.
 - [99] T.S. Cho, M. Butman, S. Avidan, W.T. Freeman, "The Patch Transform and Its Applications to Image Editing", in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
 - [100] Z. Wang, A.C. Bovik, L. Lu, "Why Is Image Quality Assessment So Difficult?" in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, May 2002.

- [101] J.G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148-1161, 1993.
- [102] D. Zhang, W.K. Kong, J. You, M. Wong, "On-line Palmprint Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1041-1050, 2003.
- [103] M. Savvides, B.V.K.V. Kumar, P.K. Khosla, "Eigenphases vs Eigenface", in *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 3, pp. 810-813, 2004.
- [104] J. Behar, M. Porat, Y.Y. Zeevi, "Image Reconstruction from Localized Phase", *IEEE Transaction on Signal Processing*, vol. 40, no. 4, pp. 736-743, 1992.
- [105] C. Perra, F. Massidda, D.D. Giusto, "Image Blockiness Evaluation Based on Sobel Operator", in *Proceedings of the 12th International Conference on Image Processing*, pp. I-389-392, 2005.
- [106] D.A. Vidmar, D. Cruess, P. Hsieh, Q. Dolecek, H. Pak, M. Gwynn, K. Maggio, A. Montemorano, J. Powers, D. Richards, L. Sperling, H. Wong, J. Yeager, "The Effect of Decreasing Digital Image Resolution on Teledermatology Diagnosis", *Telemedicine Journal*, vol. 5, no. 4, pp. 375-383, 1999.
- [107] "Standard illuminant", http://en.wikipedia.org/wiki/Standard_illuminant.
- [108] A. Krishnaswamy, G.V.G. Baranoski, "A Study on Skin Optics", *Technical Report*, University of Waterloo, 2004.
- [109] S.L. Jacques, "Skin Optics", Oregon Medical Laser Center, <http://omlc.ogi.edu/news/jan98/skinoptics.html>, 1998.
- [110] C. Donner, H.W. Jensen, "A Spectral BSSRDF for Shading Human Skin", in *the Proceedings of the Eurographics Symposium on Rendering*, pp. 409-417, 2006.
- [111] S. Prahl, "Optical Absorption of Hemoglobin", Oregon Medical Laser Center, <http://omlc.ogi.edu/spectra/hemoglobin/index.html>, 2001.
- [112] C. Donner, T. Weyrich, E. d'Eon, R. Ramamoorthi, S. Rusinkiewicz, "A Layered, Heterogeneous Reflectance Model for Acquiring and Rendering Human Skin", *ACM Transactions on Graphics (Proc. SIGGRAPH ASIA)*, vol. 27, no. 5, 2008.
- [113] "Spectral Response Function", <http://scien.stanford.edu/class/psych221/projects/05/joanmoh.html>.
- [114] S. Haykin, "Neural Networks: A Comprehensive Foundation", New York: Macmillan Publishing, 1994.
- [115] H. Zhang, C. Tang, A. Kong, N. Craft, "Matching Vein Patterns from Color Images for Forensic Investigation", *IEEE Conference on Biometrics: Theory, Applications and Systems*, 2012 (accepted as oral presentation).
- [116] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization", *Graphic Gems IV*. San Diego: Academic Press Professional, pp. 474-485, 1994.
- [117] A. Kong, "An Evaluation of Gabor Orientation as a Feature for Face Recognition", in *Proceedings of the 19th International Conference on Pattern Recognition*, pp. 1-4, 2008.
- [118] J. Reichman, "Determination of Absorption and Scattering Coefficients for Nonhomogeneous Media 1: Theory", *Applied Optics*, vol. 12, no. 8, pp. 1811-1815, 1973.
- [119] S. Avrillier, E. Tinet, E. Delettre, "Monte Carlo Simulation of Collimated Beam Transmission through Turbid Media", *Journal of Physics (Paris)* 51, pp. 2521-2542, 1990.
- [120] A.N. Bashkatov, É.A. Genina, V.I. Kochubey, V.V. Tuchin, "Optical Properties of the Subcutaneous Adipose Tissue in the Spectral Range 400–2500 nm", *Optics and Spectroscopy*, vol. 99, no. 5, pp. 836-842, 2005.
- [121] A. Roggan, M. Friebel, K. Dörschel, A. Hahn, G. Müller, "Optical Properties of Circulating Human Blood in the Wavelength Range 400–2500 nm", *Journal of Biomedical Optics*, vol. 4, no. 1, pp. 36-46, 1999.

- [122] A. Toet, L.J. van Ruyven, J.M. Valetton, "Merging Thermal and Visual Images by A Contrast Pyramid", *Optical Engineering*, vol. 28, no. 7, pp. 789-792, 1989.
- [123] A. Toet, "Multi-scale Contrast Enhancement with Applications to Image Fusion", *Optical Engineering*, vol. 31, no. 5, pp. 1026-1031, 1992.
- [124] International Committee for Information Technology Standards, Biometric Sample Quality Standard. Draft (Rev. 4), Document M1/06-0003 2005.
- [125] P. Grother, E. Tabassi, "Performance of biometric quality measures", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 531-543, 2007.
- [126] J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, J. Bigun, "A Comparative Study of Fingerprint Image-Quality Estimation Methods", *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 4, pp. 734-743, 2007.
- [127] R. Haralick, L. Shapiro, Computer and Robot Vision, Volume I, Addison-Wesley, pp. 28-48, 1992.
- [128] "Vein", <http://en.wikipedia.org/wiki/Vein>.
- [129] J.E. Bender, K. Vishwanath, L.K. Moore, J.Q. Brown, V. Chang, G.M. Palmer, N. Ramanujam, "A Robust Monte Carlo Model for the Extraction of Biological Absorption and Scattering In Vivo", *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 960-968, 2009.
- [130] S.V. Patwardhan, A.P. Dhawan, P.A. Relue, "Monte Carlo Simulation of Light-Tissue Interaction: Three-Dimensional Simulation for Trans-Illumination-Based Imaging of Skin Lesions", *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 7, pp. 1227-1236, 2005.
- [131] M. Shimada1, Y. Yamada, M. Itoh, T. Yatagai, "Melanin and Blood Concentration in a Human Skin Model Studied by Multiple Regression Analysis: Assessment by Monte Carlo Simulation", *Physics in Medicine and Biology*, vol. 46, pp. 2397-2406, 2001.
- [132] M. Hiraokatq, M. Firbankt, M. Essenpreist, M. Copet. S.R. Anidge, P. van der Zees, D.T. Delpy, "A Monte Carlo Investigation of Optical Pathlength in Inhomogeneous Tissue and Its Application to Near-infrared Spectroscopy", *Physics in Medicine and Biology*, vol. 38, pp. 1859-1876, 1993.
- [133] J. Bergeat, B. Rutily, "Auxiliary Functions for Radiative Transfer Problems in Plane-Parallel Geometry – I, the Infinite Medium", *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 55, no. 6, pp. 715-729, 1996.
- [134] M. Elias, G. Elias, "Radiative Transfer in Inhomogeneous Stratified Scattering Media with Use of the Auxiliary Function Method", *Journal of the Optical Society of America A*, vol. 21, no. 4, pp. 580-589, 2004.
- [135] C. Magnain, M. Elias, J.-M. Frigerio, "Skin Color Modeling Using the Radiative Transfer Equation Solved by the Auxiliary Function Method", *Journal of Optical Society of America A*, vol. 24, no. 8, pp. 2196-2205, 2007.
- [136] C. Magnain, M. Elias, J.-M. Frigerio, "Skin Color Modeling Using the Radiative Transfer Equation Solved by the Auxiliary Function Method: Inverse Problem", *Journal of Optical Society of America A*, vol. 25, no. 7, pp. 1737-1743, 2008.
- [137] K. Stamnes, S.C. Tsay, W. Wiscombe, K. Jayaweera, "Numerically Stable Algorithm for Discrete-Ordinate-Method Radiative Transfer in Multiple Scattering and Emitting Layered Media", *Applied Optics*, vol. 27, no. 12, pp. 2502-2509, 1988.
- [138] V.V. Rozanov, A.A. Kokhanovsky, "The Solution of the Vector Radiative Transfer Equation Using the Discrete Ordinates Technique: Selected Applications", *Atmospheric Research*, vol. 79, no. 3-4, pp. 241-265, 2006.
- [139] R.J.D. Spurr, T.P. Kurosu, K.V. Chance, "A Linearized Discrete Ordinate Radiative Transfer Model for Atmospheric Remote-Sensing Retrieval", *Journal of Quantitative Spectroscopy & Radiative Transfer*, vol. 68, pp. 689-735, 2001.
- [140] A. da Silva, M. Elias, C. Andraud, J. Lafait, "Comparison of the Auxiliary Function Method and the Discrete-Ordinate Method for Solving the Radiative Transfer Equation for Light Scattering", *Journal of the Optical Society of America A*, vol. 20, no. 12, pp. 2321-2329, 2003.

- [141] A. Semnani, M. Kamyab, I.T. Rekanos, "Reconstruction of One-dimensional Dielectric Scatterers Using Differential Evolution and Particle Swarm Optimization", *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 671-675, 2009.
- [142] I.T. Rekanos, "Shape Reconstruction of a Perfectly Conducting Scatterer Using Differential Evolution and Particle Swarm Optimization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 1967-1974, 2008.
- [143] A. Qing, "A Parametric Study on Differential Evolution Based on Benchmark Electromagnetic Inverse Scattering Problem", in *IEEE Congress Evolutionary Computation*, Singapore, September 25-28, pp. 1904-1909, 2007.
- [144] Z. Li, E.J. Delp, "Block Artifact Reduction Using A Transform Domain Markov Random Field Model", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1583-1593, 2005.
- [145] A. Sawetprawichkul, "A Study of the Transient Radiative Transfer within Participating Media", Ph.D. Thesis, Florida Institute of Technology, 2002.
- [146] J. Besag, "Spatial interaction and the statistical analysis of lattice systems", *Journal of the Royal Statistical Society. Series B*, vol. 36, no. 2, pp. 192-236, 1974.