Restoring blood vessel patterns and tattoos from JPEG compressed skin images for forensic analysis

Li, Xiaojie

2018

Li, X. (2018). Restoring blood vessel patterns and tattoos from JPEG compressed skin images for forensic analysis. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/74576>

<https://doi.org/10.32657/10356/74576>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**RESTORING BLOOD VESSEL PATTERNS AND
TATTOOS FROM JPEG COMPRESSED SKIN IMAGES
FOR FORENSIC ANALYSIS**

LI XIAOJIE
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
2018

**RESTORING BLOOD VESSEL PATTERNS AND
TATTOOS FROM JPEG COMPRESSED SKIN IMAGES
FOR FORENSIC ANALYSIS**

LI XIAOJIE

School of Computer Science and Engineering

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

2018

Abstract

Using digital evidence images for criminal and victim identification in some legal cases, such as child sexual abuse and masked gunman, can be very challenging, because the faces of criminals or victims are not visible. Blood vessel patterns under human skin and tattoos have been proposed as biometrics to overcome this challenge. However, these images are always taken by consumer cameras and compressed by the JPEG compression method. As a lossy compression method, the JPEG compression method seriously degrades the clarity of the blood vessel patterns and tattoos. To finally utilize blood vessels and tattoos for identifying criminals and victims or searching suspects, overcoming this challenge from the JPEG compression method is essential. Existing methods are not suitable for restoring the blood vessel patterns and tattoos because most of them are designed for generic images and their main targets are to remove the blocking artifacts and improve the visual quality. In this thesis, a skin image analysis and a compression test are performed to study the characteristics of compressed skin images. Then, based on the skin image analysis and the compression test, two skin image restoration algorithms are specially designed to restore blood vessel patterns from compressed images. Finally, an algorithm based on deep learning is proposed to recover tattoo images from their degraded versions. Skin image databases and tattoo image database are constructed for algorithm development.

To recover the blood vessel patterns from compressed skin images, a skin image analysis and a compression test are first conducted to find out the critical factors influencing blood vessel pattern quality. The skin image analysis aims to figure out how blood vessel information distributes in the luminance channel (the Y channel) and the chrominance channels (the U and V channels). The compression test is further performed to identify how the lossy operations in the JPEG compression method degrade the blood vessel pattern quality. The down-sampling operation and the quantization operation are examined separately in the test. The importance of the Discrete Cosine Transform (DCT) coefficients is also tested step by step. Based on the analysis, two restoration algorithms are designed to recover blood vessel patterns by estimating the original critical DCT coefficients.

Firstly, a global algorithm is developed to recover the critical DCT coefficients. JPEG compressed images and original images are used to construct image patch pairs. The DCT coefficients in the JPEG compressed images are extracted as features and the critical DCT coefficients from the original images are extracted as targets. Regression models are learned to predict the original DCT coefficients. The models are applied to the JPEG compressed image uniformly to restore the images. Experimental results show that the proposed algorithm restores blood vessels more effectively than the state-of-the-art deblocking methods in terms of matching accuracy.

Secondly, a multi-model restoration algorithm (MMRA) is presented to remove blocking artifacts in JPEG compressed skin images and restore the lost information. Two mathematical properties in the JPEG compression process are identified and used to design MMRA. MMRA is based on a tailor-made clustering scheme to group training data and learns a model, which predicts original DCT coefficients, from each grouped dataset. Under a minor assumption, MMRA is mathematically guaranteed that it always offers lower training error comparing with the corresponding global method without using the clustering scheme. Experimental results demonstrate that MMRA outperforms the existing deblocking methods for different scales and different compression levels and effectively restores blood vessel information.

A public skin image database and an Internet image database are constructed for blood vessel restoration algorithm development. The public skin image database has two datasets, a high resolution image dataset and its downsampled version. Images in this database are used to perform the skin image analysis, compression tests, model training and algorithm evaluation. The Internet image database also contains two datasets. Both the two datasets are used to evaluate the proposed algorithms, while images from the first dataset are used for the visual comparison and objective evaluation based on quality metrics and images from the second dataset are used for the evaluation based on blood vessel matching.

For restoring tattoo images, a tree-structured deep framework is developed to recover the original DCT coefficients in the frequency domain. All the coefficients are arranged in the order

of zigzag and divided into three groups according to their frequency, the low frequency group, the middle frequency group and the high frequency group. The proposed framework takes the image patches in the RGB domain as inputs. The DCT coefficient residuals, i.e., the differences between the original DCT coefficients and the compressed DCT coefficients are employed as the target. In the training phase, three parallel branches are trained simultaneously to recover the three groups of DCT coefficients. A constraint layer is specially designed to guarantee that the outputs of the deep framework are in proper ranges. Since the DCT coefficients are not equally important to the image quality, the weighted mean squared error function is utilized as the loss function. Tattoo images collected from the Internet are used to construct a tattoo image database for algorithm design and evaluation. Extensive experiments show that the proposed algorithm can suppress the blocking artifacts and recover the missing information effectively. The proposed algorithm outperforms the state-of-the-art deblocking methods in terms of image quality metrics.

Acknowledgments

I would like to thank all those people who contributed in some way to my Ph.D. study.

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Kong Wai-Kin Adams for his invaluable guidance, selfless support and encouragement during my Ph.D. study. His enthusiasm and rigorous attitude to research inspire me in my study and work. He spends a lot of time and effort to train me to be an independent researcher. And he shares with me his own experience and vision about research which motivate me to reach my own goals as well. Without his guidance and persistent help this thesis would not have been possible.

My sincere thanks also go to Dr. Tang Chaoying and Dr. Zhang hengyi for their generous help and supports in my work. I also want to thank all my friends and lab mates, Arfika, Yanzhu, Xingpeng, Frodo, Han, Hamid, Ruhba, Quang, Abhik, Peicong, Yi, Suman and Shitala, for providing such a lovely environment for study. They are all knowledgeable and friendly. Discussions with them help me to broaden my horizon and expand my mind.

I devote my deepest gratitude to my parents for their endless love and support throughout my life. Their company and encouragement help me to go through the hard time.

Contents

Abstract	i
Acknowledgments.....	iv
Contents	v
List of Figures	ix
List of Tables	xiii
List of Acronyms	xv
List of Publications	xvi
Chapter 1 Introduction	1
1.1 JPEG Compressed Image Restoration for Criminal and Victim Identification ...	1
1.2 Criminal and Victim Identification based on Tattoos and Blood Vessel Patterns	2
1.3 Motivation and Objectives	4
1.3.1 Blood Vessel Pattern Restoration.....	4
1.3.2 Tattoo Restoration.....	8
1.4 Challenges and Solutions	10
1.5 Organization of the Thesis	12
Chapter 2 Background and Literature Review	13
2.1 Introduction to the JPEG Compression Method	13
2.2 A Literature Review of JPEG Artifact Removal.....	16

2.2.1	The Preprocessing Methods	17
2.2.2	The Postprocessing Methods	18
2.3	A Blood Vessel Pattern Uncovering Method.....	32
2.4	Blood Vessel Pattern Extraction and Automatic Matching Methods.....	34
2.4.1	A Blood Vessel Pattern Extraction Method	34
2.4.2	A Three Step Blood Vessel Matching Method.....	37
2.5	Image Quality Assessment Metrics.....	40
2.6	Summary	41
Chapter 3	Skin Image Analysis and Compression Test for Blood Vessel Restoration	43
3.1	A Public Skin Image Database.....	44
3.2	A Skin Image Analysis.....	48
3.3	A Skin Image Compression Test.....	53
3.4	Statistics on Skin Images	61
3.5	Summary	63
Chapter 4	A Global Model for Restoring Blood Vessel Patterns from Compressed Skin Images	65
4.1	A Preliminary Study on Restoring Blood Vessel Patterns.....	67
4.1.1	The Proposed Algorithm.....	68
4.1.2	Features and Regression Models.....	68
4.1.3	Post-processing	70

4.1.4	Experimental Results	71
4.2	A Global Model	77
4.2.1	The Proposed Algorithm	77
4.2.2	Experimental Results	78
4.3	Summary	85
Chapter 5	A Multi-model Restoration Algorithm for Recovering Blood Vessels in Skin Images	87
5.1	Two Important Properties and Notations	88
5.2	A Clustering Scheme.....	89
5.3	Algorithm Design.....	92
5.4	Features and Regression Models.....	93
5.5	Experimental Results	94
5.5.1	Visual comparisons	95
5.5.2	Objective evaluations based on the SSIM index and PSNR index	101
5.5.3	Objective evaluation based on blood vessel matching.....	106
5.5.4	Discussions.....	110
5.6	Summary	112
Chapter 6	Tattoo Image Restoration based on Deep Convolutional Neural Network.....	113
6.1	The Network Architecture.....	115
6.2	An Overlapping Block Restoration Scheme	118

6.3	An Adaptive Constraint Layer	120
6.4	A Weight Learning Scheme	121
6.5	Experimental Results	125
6.5.1	Implementation Details	126
6.5.2	Visual Comparisons	127
6.5.3	Objective Evaluations based on SSIM index and PSNR index	132
6.5.4	Discussions.....	133
6.6	Summary	133
Chapter 7	Conclusion and Future Work	135
7.1	Summary	135
7.2	Future Work	138
References		140

List of Figures

Fig. 1-1 Four examples of blood vessel patterns for the skin database.....	5
Fig. 1-2 Illustration of vascular information degraded by the JPEG method.....	6
Fig. 1-3 Illustration of the distribution of the blood vessel information in the Y, U, V channels....	8
Fig. 1-4 Illustration of word tattoo images degraded by the JPEG method.	9
Fig. 2-1 An illustration of the JPEG compression method.	14
Fig. 2-2 illustration of the zigzag order and two quantization tables.	15
Fig. 2-3 Illustration of decoding procedure of JPEG compression.	15
Fig. 2-4 An example from the KB method.	30
Fig. 2-5 Illustration of parameter range optimization process.	33
Fig. 2-6 Data preparation for training a three-layered feed-forward neural network.....	33
Fig. 2-7 Blood vessel uncovering results.	34
Fig. 2-8 The Gabor filter outputs.	36
Fig. 2-9 The process of blood vessel pattern extraction.....	37
Fig. 2-10 The three step matching results.	39
Fig. 3-1 Raw sample images from the database.....	45
Fig. 3-2 Four example images from NTUFID_1.	47
Fig. 3-3 The lower resolution version of Fig. 3-2.	47
Fig. 3-4 Resultant images of the image analysis from the images shown in Fig. 3-2 (a).	49

Fig. 3-5 Lower resolution resultant images of the image analysis from the images shown in Fig. 3-3 (a).	49
Fig. 3-6 The matching results of NTUFID_1 for the image analysis.	50
Fig. 3-7 The matching results of NTUFID_2 for the image analysis.	51
Fig. 3-8 The matching results of NTUFID_1 for the image analysis.	52
Fig. 3-9 The matching results of NTUFID_2 for the image analysis.	53
Fig. 3-10 The processes of the compression test.....	54
Fig. 3-11 Quantization tables modified from Figs. 2-2 (b) and (c).	55
Fig. 3-12 Resultant images of NTUFID_1 from the compression test given in Fig. 3-10(a).	55
Fig. 3-13 Resultant images of NTUFID_2 from the compression test given in Fig. 3-10(a).	55
Fig. 3-14 Results of images in NTUFID_1 from the compression process in Fig. 3-10(b).	56
Fig. 3-15 Results of images in NTUFID_2 from the compression process in Fig. 3-10 (b).	57
Fig. 3-16 The matching results of NTUFID_1 from the compression test.	58
Fig. 3-17 The matching results of NTUFID_2 from the compression test.	59
Fig. 3-18 The matching results of NTUFID_1 from the compression test.	60
Fig. 3-19 The matching results of NTUFID_2 from the compression test.	61
Fig. 3-20 The probabilities of non-zero QDCT coefficients.....	62
Fig. 3-21 The distributions of quantized DCT coefficients of the Y, U and V channels.	63
Fig. 4-1 Four examples of blood vessel patterns from NTUFID_1.	66
Fig. 4-2 Four examples of blood vessel patterns from NTUFID_2.	66

Fig. 4-3 A schematic of the proposed algorithm.....	68
Fig. 4-4 Illustration of training data generation process for U and V channels.	70
Fig. 4-5 The restoration results from the JPEG compressed images with a quality factor of 50... 72	
Fig. 4-6 The restoration results from the JPEG compressed images with a quality factor of 75... 72	
Fig. 4-7 The matching results from compressed images with a quality factor of 50.	74
Fig. 4-8 The matching results from compressed images with a quality factor of 75.	75
Fig. 4-9 Resultant images from NTUFID_1 for comparisons.	80
Fig. 4-10 Resultant images from NTUFID_2 for comparisons.	81
Fig. 4-11 The matching results from compressed images with a quality factor of 50.	83
Fig. 4-12 The matching results from compressed images with a quality factor of 75.	84
Fig. 5-1 Illustration of the training phase of the proposed multi-model restoration algorithm.....	90
Fig. 5-2 The schematic diagram of the proposed algorithm.	93
Fig. 5-3 Raw images with background from the NTUIIS.....	95
Fig. 5-4 Resultant images from NTUFID_1 for comparisons.	97
Fig. 5-5 Resultant images from NTUFID_2 for comparisons.	98
Fig. 5-6 The restoration results from the JPEG compressed images with a quality factor of 50... 99	
Fig. 5-7 The restoration results from the JPEG compressed images with a quality factor of 75. 100	
Fig. 5-8 Illustration of the degradation effects.....	105
Fig. 5-9 The matching results from compressed images with a quality factor of 50.	107
Fig. 5-10 The matching results from compressed images with a quality factor of 75.	108

Fig. 5-11 Matching results from NTUIIS_2.	110
Fig. 5-12 Quantized DC coefficient error distributions from JPEG compression with a quality factor of 50.	111
Fig. 5-13 Quantized DC coefficient error distributions from JPEG compression with a quality factor of 75.	111
Fig. 6-1 Illustration of degradation effects caused by the JPEG compressed images.	113
Fig. 6-2 The illustration of the proposed network architecture.	117
Fig. 6-3 Illustration of the two types of blocks.	119
Fig. 6-4 The learned weight vector for the DCT coefficient residuals of the Y, U and V channels with a quality factor of 50.	123
Fig. 6-5 Some examples from the tattoo image database.	125
Fig. 6-6 The restoration results from the JPEG compressed images with a quality factor of 30.	128
Fig. 6-7 The restoration results from the JPEG compressed images with a quality factor of 40.	130
Fig. 6-8 The restoration results from the JPEG compressed images with a quality factor of 50.	131

List of Tables

Table 1-1 A brief summary of the two databases.	11
Table 2-1 The summary of the existing methods used for comparison.	32
Table 3-1 A summary of the databases	46
Table 3-2 A summary of each session in the databases	46
Table 4-1 Mean SSIM indexes from NTUFID	73
Table 4-2 Matching results from NTUFID_1	76
Table 4-3 Matching results from NTUFID_2	76
Table 4-4 Mean SSIM indexes NTUFID	82
Table 4-5 Mean PSNR indexes from NTUFID.....	82
Table 4-6 Matching results from NTUFID_1	85
Table 4-7 Matching results from NTUFID_2	85
Table 5-1 (a) Mean SSIM indexes and (b) Mean PSNR indexes from NTUIIS	101
Table 5-2 (a) Mean SSIM indexes and (b) Mean PSNR indexes from NTUFID	102
Table 5-3 (a) Mean SSIM indexes and (b) mean PSNR indexes from the Interent images in the RGB domain	103
Table 5-4 (a) Mean SSIM indexes and (b) mean PSNR indexes from NTUFID in the RGB domain	103
Table 5-5 Matching results from NTUFID_1	109
Table 5-6 Matching results from NTUFID_2	109

Table 6-1 (a) Mean SSIM indexes and (b) mean PSNR indexes from the compressed tattoo
images 132

List of Acronyms

AC	Alternative Current
AKR	Adaptive Kernel Regression
ARCNN	Artifact Reduction Convolutional Neural Network
CEFAS	Contrast Enhancement Framework Artifact Suppression method
CMC	Cumulative Match Characteristic
DC	Direct Current
DCT	Discrete Cosine Transform
DnCNN	Denoising Convolutional Neural Network
DT	The preliminary Decision Tree based algorithm
GM	Global Model based method
FOE	Field of Experts
HVS	Human Visual System
IDCT	Inverted Discrete Cosine Transform
JPEG	Joint Photographic Experts Group
JPEG50	JPEG compression with a quality factor of 50
JPEG75	JPEG compression with a quality factor of 75
KB	Knowledge-based method
LF	Left Forearm dataset
LF1	The first session of left forearm dataset
LF2	The second session of left forearm dataset
LT	Left Thigh dataset
LT1	The first session of left thigh dataset
LT2	The second session of left thigh dataset
MAP	Maximum a Posteriori
MMRA	Multi-model Restoration Algorithm
MRF	Markov Random Fields
MSE	Mean Square Error
NTUFID	NTU Forensic Image Database
NTUFID_1	Original Scale Images
NTUFID_2	Lower Resolution version of NTUFID_1
NTUIIS	NTU Internet Image Set
NTUIIS_1	Internet images for visual comparison and quality metric based evaluation
NTUIIS_2	Internet images for objective evaluation based on blood vessel matching
OBTCE	Overlapped-block Transform Coefficient Estimation method
POCS	Projection onto Convex Sets
PSNR	Peak Signal-to-noise Ratio
QAC	Quantized Alternative Current
QDC	Quantized Direct Current
QDCT	Quantized Discrete Cosine Transform
ReLU	Rectified Linear Units
RF	Right Forearm dataset
RF1	The first session of right forearm dataset
RF2	The second session right forearm dataset
RT	Right Thigh dataset
RT1	The first session of right thigh dataset
RT2	The second session of right thigh dataset
SADCT	Shape-adaptive DCT method
SGD	Stochastic Gradient Descent
SSIM	Structural Similarity
TNRD	Trainable Nonlinear Reaction Diffusion framework

List of Publications

Journal Papers

1. Frodo Kin Sun Chan, **Xiaojie Li** and Adams Kong, “A Study of Distinctiveness of Skin Texture for Forensic Applications through Comparison with Blood Vessels”, *IEEE Transactions on Information Forensics and Security*, vol. 12, no 8, pp. 1900-1915, 2017.
2. **Xiaojie Li** and Adams Kong, “A Multi-model Restoration Algorithm for Recovering Blood Vessels in Skin Images”, *Image and Vision Computing*, vol. 61, pp. 40-53, 2017.

Conference Papers

1. Yanzhu Liu, **Xiaojie Li**, Adams Kong and Chi Keong Goh, “Learning from small data: a pairwise approach for ordinal regression”, *IEEE Symposium Series on Computational Intelligence*, 2016.
2. **Xiaojie Li** and Adams Kong, “Identifying Critical Factors Influencing Quality of Blood Vessel Information in JPEG Compressed Skin Images”, *IEEE Symposium Series on Computational Intelligence*, 2015.
3. Yanzhu Liu, **Xiaojie Li** and Adams Kong, “Speeding up the knowledge-based deblocking method for efficient forensic analysis”, *IEEE Symposium Series on Computational Intelligence*, 2014
4. Hengyi Zhang, Chaoying Tang, **Xiaojie Li** and Adams Kong, “A study of similarity between genetically identical body vein patterns”, *IEEE Symposium Series on Computational Intelligence*, pp. 151 - 159, 2014.
5. **Xiaojie Li** and Adams Kong, “Restoring blood vessel patterns from JPEG compressed skin images for forensic analysis”, *IEEE International Workshop on Information Forensics and Security*, pp. 19-24, 2013. (acceptance rate=35%, oral presentation)

Awards

Winner of SSCI Travel Grant for attending the IEEE Symposium Series on Computational Intelligence, \$US 800, only one from Singapore.

Chapter 1 Introduction

1.1 JPEG Compressed Image Restoration for Criminal and Victim Identification

Digital evidence images play an important role in forensic analysis. Biometric traits, e.g. faces, tattoos, skin marks and vein patterns, can be extracted from these images and used for criminal and victim identification. The quality of the images is critical because low quality images may destroy the biometric information and make the identification impossible. The proliferation of digital imaging devices (e.g. digital cameras or electronic products with integrated cameras) has led to an exponential increase in the number of digital evidence images. However, most of these images are compressed with the JPEG compressed images, which is the most common compression method and widely installed in the digital imaging devices. As a lossy compression method, the JPEG compression method generates blocking artifacts and causes information loss at the same time. As a result, the image quality is seriously degraded. The degradation may destroy, blur or even remove the biometric traits, which makes criminal and victim identification challenging. The situation becomes even worse when the compression ratio is higher. Thus, it is necessary to restore these compressed images before further operations.

Restoring images from their compressed versions, also called image deblocking [1], has been studied for decades. Many image deblocking methods have been proposed for restoring generic images. Since human visual system is sensitive to the blocking effects, these methods aim to remove the blocking artifacts and improve the visual quality. They are not suitable for restoring evidence images, because forensic analysis requires restoring specific biometric information. In this thesis, the problem of restoring skin biometric information from JPEG compressed images is studied. Two biometric traits, blood vessel patterns and tattoo patterns, are considered.

1.2 Criminal and Victim Identification based on Tattoos and Blood Vessel Patterns

The convenience of the Internet has promoted a mass increase in child sexual abuse images (also known as child pornography). According to a statistic from the Internet Watch Foundation (IWF), there is a huge increase in reports of child sexual abuse images [2]. The number of reports soared from 39, 211 in 2012 to 51, 186 in 2013 and the rate of rise was 31% [3]. The U.S. Department of Justice reported that there were more than 20 million IP addresses nationwide offering child sexual abuse images and the demand for these images was responsible for the increase in child sexual abuse cases [4]. Criminals in some of these images, e.g., pedophiles, are hard to be identified because they always hide their faces and show only non-facial skin to avoid identification. This identification problem is not limited to child sexual abuse cases, but also includes masked gunmen, riots and terrorist attacks. Using these images as evidence in legal cases can be very challenging, because most of the traditional biometric traits widely used in legal cases for criminal and victim identification, including face, fingerprint, DNA and palmprint are not applicable to these evidence images.

Skin marks, scars, tattoos and androgenic hair patterns provide valuable information for identifying the victims and criminals [5-8]. Among these soft biometric traits, tattoos have attracted a lot of attention from the forensic community due to their prevalence. A survey by The Harris Poll reports that one in five U.S. adults have at least one tattoo (21%) in 2012, which is up from 14% when the same survey was conducted in 2008 [9]. The data shows that there has been a huge increase in popularity of tattoos among U.S. adults. For victim and criminal identification, tattoos have been used to narrow down suspect lists, especially for those cases when hard biometric traits, such as face and fingerprint, are not available. For example, tattoos were used to assist in identifying victims of the 9/11 terror attacks in 2001 [10].

As a means of victim and criminal identification, tattoo detection and recognition has received increasing attention in the past decades. Many automatic detection and recognition

approaches have been proposed to detect and match tattoo images collected from crime scenes [7, 11-16]. To detect or recognize the tattoo patterns in the evidence images, proper descriptors and matching methods are desirable. Many image features with high representation capability, e.g. the shape features, the local patterns and learned features using deep neural network, are utilized to represent the tattoo information in these methods. Content-based image retrieval (CBIR) techniques [11, 12, 17], bag-of-word models [14] and deep learning framework [16] are employed to perform the tattoo searching and matching. With these methods, tattoo information can be extracted from the digital skin evidence images and matched for personal identification.

However, in some legal cases, skin marks, scars, tattoos and androgenic hair patterns are ineffective because not everyone has a unique scar, tattoo, skin mark pattern or androgenic hair pattern on a particular body site in the images for identification. For these cases, researchers have proposed to extract blood vessel patterns for identification. Blood vessels under human skin form large and robust patterns providing rich information for personal identification. They are generally regarded unique and stable over time, except for the cases of injury, disease (e.g. varicose veins and spider veins), and surgery. Even though identical twins share the same DNA sequence, their blood vessel patterns are different [18]. In addition, blood vessel patterns are difficult to be modified and therefore they are resistant to forgery. Thus, they are generally considered as a hard biometric trait and using blood vessel patterns for criminal and victim identification is desirable.

Blood vessel pattern extraction has been an active research topic for a decade [19-21]. Generally, infrared (IR) and laser imaging systems are required to capture this vascular information under skin, because they have high penetration ability, comparing to visible light. Nowadays, with IR imaging technologies, finger and hand blood vessel recognition systems have become mature. For examples, hand veins, palm veins and finger veins have been widely studied for development of commercial biometric systems in which they can be captured in contactless environments [22, 23]. These systems depend on highly controlled near infrared (NIR) imaging environments and devices to capture blood vessel pattern images. However, in forensic analysis, neither the controlled imaging environments nor the devices are applicable. Thus, it is a challenging problem to utilize blood vessel patterns for criminal and victim identification.

Recently, some researchers attempt to visualize blood vessels hidden in color images for commercial biometric, healthcare and forensic applications [24-27]. Based on the principles of optics and skin biophysics, Tang and Zhang et al. [26, 28] proposed methods to uncover blood vessel patterns hidden in color skin images for criminal and victim identification. These methods are designed for criminal and victim identification based on partial skin in evidence images, where faces of criminals and/or victims are not observable. These cases can be child pornography and masked terrorists. These methods were evaluated on a public skin database and the experimental results were encouraging, which show the usage of blood vessel for criminal and victim identification possible.

1.3 Motivation and Objectives

1.3.1 Blood Vessel Pattern Restoration

The blood vessel uncovering research results demonstrate the potential of using blood vessel patterns for criminal and victim identification. With newly developed uncovering methods, blood vessel patterns under human skin are possible to be visualized from evidence images (e.g., child sexual abuse images). The vascular information under skin is weak and sensitive. Thus, the quality of blood vessel patterns generated by the uncovering methods highly depends on the quality of input images. Low quality images will degrade the clarity of the blood vessel patterns. However, most of evidence images in the cases are taken by consumer cameras and compressed by the JPEG method [29]. Vascular information hidden in images can be seriously degraded by the JPEG method. As a consequence, the visualization methods may fail to visualize the blood vessel patterns. Fig.1-1 and Fig. 1-2 show two sets of examples to illustrate the problem.

Fig. 1-1 shows four original images (the first row), their uncovered blood vessels (the second row), their JPEG compressed versions (the third row) and the uncovered blood vessels from the JPEG compressed versions with a quality factor of 50 (the fourth row), and the JPEG

compressed versions (the fifth row) and the uncovered blood vessels from the JPEG compressed versions with a quality factor of 75 (the last row). All the images are from a public skin database¹.

Fig. 1-1 Four examples of blood vessel patterns for the skin database. Each column shows one example. The first row is four original images, which are a left forearm image, a right forearm image, a left thigh image and a right thigh image, respectively. The second row shows the corresponding blood vessel patterns from the original color images [26]. The third row is the JPEG compressed versions of the original images in the first row with a quality factor of 50. The fourth row shows the blood vessel patterns from images in the third row. The last two rows show the JPEG compressed versions of the original images in the first row with a quality factor of 75 and the corresponding blood vessel patterns.

Fig. 1-2 shows three images collected from the Internet as examples. The first column shows the original color images². The second column shows the corresponding visualized blood vessels [26]. The third and the last columns show their visualized blood vessels from the JPEG compressed versions with compression factors of 75 and 50 respectively.

¹ The introduction of the public skin database is given in Chapter 3.

² The images were downloaded from the Internet and were originally compressed by JPEG method. To show the degradation effects, original (uncompressed) images were required. Thus, the downloaded images were resized to remove the compression effects and referred to as references. The details of the Internet dataset are given in Chapter 5.

Fig. 1-2 Illustration of vascular information degraded by the JPEG method. The first column is three original images collected from the Internet. The second column is the corresponding blood vessel patterns visualized by Zhang et al.'s method [26]. The third and the last columns show their visualized blood vessels from the JPEG compressed versions with compression factors of 75 and 50 respectively.

From the figures, it can be seen that blood vessels from the original images are clear, but they are seriously degraded by the JPEG method. Thus, to finally utilize blood vessels for identifying criminals and victims or searching suspects, vascular information should be restored before applying the visualization methods.

Many deblocking methods have been proposed for generic images and their main concern is to improve the visual quality of the images. To improve the visual quality, these methods attempt to remove blocking artifacts, but original information is not guaranteed to be restored [30]. Some of them even cannot guarantee that the recovered images and the JPEG compressed images have the same QDCT coefficients and the difference between the uncompressed images and the recovered images may be even larger than that between uncompressed images and compressed images in terms of QDCT coefficients. Some priors are imposed on some of these methods to recover the lost information. However, they are not related to biometric information

[30]. In fact, many of them further deteriorate the existing information, which is already weakened by the JPEG method. Many of the methods only operate in the Y channel because the Y channel carries most of the visual information and human visual system is not sensitive to the U and V channels. However, it is not applicable to the blood vessel information. Fig. 1-3 illustrates the distribution of the blood vessel information. Fig. 1-3 shows a color image, the Y, U and V channels and the blood vessel patterns uncovering from the skin part in the color image. More analysis on the distribution of the biometric information in the three channels can be found in Chapter 3. From this figure, it can be seen that it is hardly to observe the blood vessel patterns in the Y channel, but the blood vessel patterns in the U and the V channels are clear. It indicates that the distribution of the visual information is different from that of the blood vessel information. Thus, these deblocking methods designing for improving the visual quality of the generic images are not suitable to restore vascular information. It is necessary to develop an effective restoration method to recover the original blood vessel information and suppress the blocking artifacts as well. A knowledge-based (KB) [30, 31] method is specially proposed to remove blocking artifacts and recover the skin marks from compressed skin images. It employs a brute-force search scheme to search for optimal uncompressed blocks in a large skin database and replaces the compressed blocks. It is not suitable for blood vessel restoration because it does not take the cross-block and cross channel information in to consideration.

Fig. 1-3 Illustration of the distribution of the blood vessel information in the Y, U, V channels. The figure shows the original color image, the Y, U and V channels and the blood vessels from the original image.

1.3.2 Tattoo Restoration

As biometric traits, tattoos have been applied in criminal and victim identification by law enforcement agencies for many years. Researchers have conducted some studies on tattoo detection and recognition [7, 11, 12, 15, 16]. An important factor affecting the accuracy of these methods is the quality of the input images. However, tattoo images taken by the consumer cameras are likely compressed by the JPEG compressed method. The compression process has caused information loss and blocking artifacts. As a result, the image quality is degraded and makes the tattoo images less expressive. Thus, to restore the lost information and remove the blocking artifacts are necessary for fully utilize information in tattoos. The designs of tattoos include but not limited to words, symbols, portraits and some objects, such as flowers or animals.

Fig. 1-4 Illustration of word tattoo images degraded by the JPEG method. (a), (c), (e) and (g) are four original tattoo images collected from the Internet. (b), (d), (f) and (h) are the compressed versions with a quality factor of 50 respectively.

Fig. 1-4 shows four word tattoo images collected from the Internet as examples. From the figure, it can be seen that the word tattoo patterns in the original images (Fig. 1-4 (a), (c), (e) and (g)) are smooth, clear and easy to recognize. However, the blocking effects are obvious in the JPEG compressed versions and seriously degrade the visual effect of the word tattoo images (Fig. 1-4(b), (d), (f) and (h)). Thus, it is necessary to restore the loss information and remove the blocking artifacts from the JPEG compressed images for tattoo recognition.

Existing methods designed for the generic images mainly targets at removing blocking artifacts and sharpening the edges in the compressed images [1]. Some of them deal with the restoration problem based on statistic models and assumptions, e.g. Gaussian models are generally used to model the noises and Markov random fields are used to capture the dependence between neighborhoods. The resultant images from these methos usually suffer from blurry effects [32]. Some machince learning techniques, e.g. sparse coding, deep neural network, are used to extract prior knowledge from a generic database for image restoration. Some of them, which restore images in the spatial domain, can not make sure the resultant images and the JPEG compressed correspondants have the same quantized Discrete Cosine Transform (QDCT) coefficients. Thus, they are not suitable for tattoo image restoration. The KB [31] method is not suitable for tattoo image restoration because the tattoo images are more diverse than skin images

and it is challenging to search for the optimal uncompressed blocks for a block from a tattoo image.

1.4 Challenges and Solutions

In this thesis, we mainly concentrate on developing algorithms for removing JPEG blocking artifacts and restoring biometric information, including blood vessel patterns and word tattoos. The JPEG compression method [29] is a lossy compression scheme because the down-sampling operation reduces information in the U and V channels and the quantization process removes high frequency information. Furthermore, the block-wise operation destroys the continuity between neighbour blocks and generates many horizontal and vertical blocking artifacts along the block boundaries. Even worse, the JPEG compression method applies quantization operation with larger values to reduce chrominance information in the U and V channels for further file size reduction. As a result, the blood vessel information and the tattoo information in the color skin images are degraded by the JPEG compression method (see Figs. 1-1, 1-2 and 1-4).

It is challenging to recover the specific biometric traits from the JPEG compressed images for forensic analysis because it is required to restore the original information. The current methods designed for generic images cannot achieve this target because they aim to improve the visual quality of the images and cannot guarantee to recover the lost information.

A large number of application-specific image processing methods have been developed to solve various specific problems. Some of their target applications are to increase resolutions of face images [33, 34] and car plate images [35], enhance latent fingerprints on banknotes [36], and segment latent fingerprints [37]. In general, these methods perform significantly better than methods designed for generic images because they can make use of properties of their target objects (e.g., faces and fingerprints) and extract information from large databases.

In this thesis, two application-specific algorithms for blood vessel recovery from JPEG compressed skin images are proposed: one is a global blood vessel restoration algorithm and the

other is a multi-model restoration algorithm. An algorithm for word tattoo restoration utilizing a deep learning framework is also proposed.

Since the JPEG compression method operates in the DCT domain and the image quality degradation of JPEG compressed images is mainly caused by the quantization operation in the DCT domain, it is straightforward to carry out the restoration process in the DCT domain. In addition, the JPEG compression process indicates that the recovered images should have the same QDCT coefficients as the compressed images. Restoring the original DCT coefficients in the DCT domain can guarantee to fulfill this constraint, while the restoration in the spatial domain cannot. Thus, the proposed algorithms in this thesis are designed to restore the original DCT coefficients in the DCT domain.

A public skin image database (NTU Forensic Image Database (NTUFID)) and an Internet image set (NTU Internet Image Set (NTUIIS)) are constructed to design algorithms for blood vessel pattern restoration. NTUFID has two datasets: NTUFID_1 and NTUFID_2. Images in NTUFID_2 are the lower resolution version of images in NTUFID_1. NTUIIS also contains two parts: NTUIIS_1 and NTUIIS_2, while images in NTUIIS_1 are used for the visual comparison and objective evaluation based on quality metrics and images in NTUIIS_2 are used to perform the blood vessel matching experiments. Table 1-1 summarizes the brief information on the two databases.

Table 1-1 A brief summary of the two databases.

Datasets	Number of images	Source	Components	Compression quality factor	Use case
NTUFID_1	1894	Collected in a non-well controlled lab environment	Four limb datasets and two images are taken in two sessions for each limb	Uncompressed	Chapters 3, 4 and 5
NTUFID_2	1894	The lower resolution version of NTUFID_1	Four limb datasets and two images are taken in two sessions for each limb	Uncompressed	Chapters 3, 4 and 5
NTUIIS_1	84	Internet	Random images taken for various body parts	Uncompressed	Chapter 5
NTUIIS_2	318	Internet	A left forearm dataset and a right forearm dataset and two images are collected for each limb	Compression factors range from 50 to 90	Chapter 5

Considering the properties of skin images and JPEG compression method, a skin image analysis and compression tests are performed on NTUFID to obtain a better understanding of the degradation effects imported by the compression process. Based on the skin image analysis and the compression tests, algorithms are specially designed to restore the blood vessel patterns from the compressed skin images.

To restore the tattoo images, a tree-structured deep neural network is designed to estimate the residuals of DCT coefficients in the frequency domain. Considering the DCT coefficients from one block carry different frequency information, they are divided into three groups according to their frequency. Deep features are specially learned for each group of DCT coefficients. A constraint layer is designed to fulfill the condition that the QDCT coefficients of the restored images should be the same as those from the compressed images. Tattoo images collected from Internet are used for algorithm design and evaluation.

1.5 Organization of the Thesis

The rest of this thesis is organized as follows. In Chapter 2, a brief introduction of JPEG compression method and the review of various artifact removal methods are given. A blood vessel uncovering method and a three step blood vessel matching method, which is used to evaluate the effectiveness of the proposed blood vessel pattern restoration algorithms, are also introduced. Two objective image quality metrics, which are used to measure the quality of the resultant images, are reviewed. A skin image analyses and a compression test on JPEG compression method are reported in Chapter 3. Chapter 4 presents the proposed global blood vessel restoration algorithm and the experimental results and comparisons. Chapter 5 reports the multi-model restoration algorithm (MMRA) and some analyses on blood vessel images. The extensive experimental results and comparisons are also included. The tattoo image restoration algorithm and the experimental results and comparisons are given in Chapter 6. Finally, the conclusion and future work are given in chapter 7.

Chapter 2 Background and Literature Review

In this chapter, a brief introduction to JPEG compression method is given first. Then, a literature review of JPEG artifact removal methods is listed. A blood vessel uncovering method and blood vessel pattern extraction and matching methods, which are used to visualize and extract blood vessel patterns hidden in color images and measure the similarity/dissimilarity between two blood vessel patterns, are introduced. These methods are used to evaluate the effectiveness of the proposed blood vessel restoration algorithms. In addition to the matching accuracy, two more image quality assessment metrics are employed to evaluate the performance of the proposed algorithms. The introduction of these two metrics is given in this chapter.

2.1 Introduction to the JPEG Compression Method

JPEG, standing for the Joint Photographic Experts Group (the name of the committee that created the JPEG standard), is the most widely used compression method in recent years [29]. The JPEG compression method is used in a number of applications to reduce communication bandwidth or storage space. It is the most common image format used by digital cameras and other imaging devices. As a lossy compression method, it reduces image fidelity and generates blocking artifacts, which degrade image quality. The JPEG compression process includes two parts: the encoding procedure and the decoding procedure. Uncompressed images are processed through the encoding procedure to obtain compressed data and the compressed images are reconstructed by the decoding procedure.

The main steps of the JPEG method are illustrated in Fig. 2-1. Images are first transformed from the RGB color space into the YUV color space. The three channels are processed separately in the rest of the operations. As human vision is less sensitive to the chrominance channels, the U and V channels are down-sampled. Then, the three channels are divided into 8 by 8 blocks and each block is processed by the discrete cosine transform (DCT). The quantized DCT (QDCT) coefficients are computed by,

$$x_{zk}^j = [y_{zk}^j / q_{zk}], \quad (2-1)$$

where y_{zk}^j denotes the k^{th} DCT coefficient in the j^{th} block of the z channel; q_{zk} is a value from a predefined quantization table, Q_z ; $k \in \{1, 2, \dots, 64\}$; $z \in \{Y, U, V\}$ and $[]$ is a round operator. Finally, the QDCT coefficients are encoded to increase compression ratio without further information loss. As shown in Fig. 2-2(a), the DCT coefficients are arranged in a zigzag order in which a lower frequency coefficient has a lower order. The first element is a direct current (DC) coefficient and the others are alternative current (AC) coefficients. According to the JPEG standard, the U and V channels share the same quantization table, which is different from the one for the Y channel. Figs. 2-2 (b) and (c) are two quantization tables for generating images with a quality factor of 50 [29], where Q_y is for processing the Y channel and Q_{uv} is for processing the U and V channels. Note that they are different. Most of the elements in Q_{uv} have larger values than the corresponding elements in Q_y . As a result, the compressed U and V channels have more noticeable blocking artifacts than the compressed Y channel.

In the JPEG compression process, the down-sampling operation removes information in the U and V channels; the block-based operations destroy data correlation between adjacent blocks, and the coarse quantization procedure eliminates high frequency information. These operations produce blocking artifacts and seriously deteriorate image quality, which makes blood vessel patterns unobservable.

Fig. 2-1 An illustration of the JPEG compression method.

1	2	6	7	15	16	28	29
3	5	8	14	17	27	30	43
4	9	13	18	26	31	42	44
10	12	19	25	32	41	45	54
11	20	24	33	40	46	53	55
21	23	34	39	47	52	56	61
22	35	38	48	51	57	60	62
36	37	49	50	58	59	63	64

(a)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	103	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(b)

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(c)

Fig. 2-2 illustration of the zigzag order and two quantization tables.

(a) is the zigzag order. (b) is the quantization table Q_y for the Y channel with a quality factor of 50 and (c) is the quantization table Q_{uv} for the U and V channels with a quality factor of 50.

Fig. 2-3 Illustration of decoding procedure of JPEG compression.

Fig. 2-3 shows the decoding procedure of JPEG compression method. When the compressed data are decoded, the DCT coefficients are obtained through the de-quantization, which is the inverse operation of quantization:

$$\ddot{y}_{zk}^j = x_{zk}^j * q_{zk}, \quad (2-2)$$

where $\ddot{y}_{zk}^j$ is the decoded DCT coefficient from QDCT. As a QDCT coefficient, x_{zk}^j is an integer. The decimal part has been rounded off, thus, the product of the quantization value and the decimal part cannot be recovered any more. This error is regarded as quantization error. The range of quantization error is computed as

$$e_{zk}^j = [q_{zk}/2], \quad (2-3)$$

where e_{zk}^j is the range of the difference between the original DCT coefficients and the DCT coefficients obtained from equation 2.2. After compression, most of the coefficients are reduced to zero except a few numbers in the left upper corner (Fig. 2-2 (a)). Especially for the U and V channels, only several lower frequency coefficients still have non-zero values.

In this section, the main steps of the JPEG compression method, including the encoding procedure and the decoding procedure, are introduced. More details about the lossy operations, especially the quantization procedure, are given. From the JPEG compression process, it can be figured out that the recovered images should have the same QDCT coefficients as the original images.

2.2 A Literature Review of JPEG Artifact Removal

The critical steps of the standard JPEG compression method, which cause image degradation, have been reviewed in chapter 2.1. Information loss occurs in the procedure of the down-sampling operation and the block-based quantization operation, which leads to annoying blocking artifacts along block boundaries. As a result, JPEG compressed images suffer from image quality degradation. Over the last few decades, a great deal of research has been done to remove the artifacts and improve the quality of JPEG compression images. Generally speaking, there are two types of commonly utilized techniques according to the solution approach in these works [1]. One is known as the preprocessing technique, which tries to address the problem at the encoder end. The other is the postprocessing technique, which is applied at the decoder end. For image deblocking, most of the proposed methods adopt the postprocessing technique because it solves the problem without changing the original data or modifying the encoding scheme. These two types of methods are introduced in the following subsections.

2.2.1 The Preprocessing Methods

The preprocessing methods are developed to reduce the blocking effects by modifying the encoder end. Two strategies, prefiltering and perceptual-based coding, are commonly used in the preprocessing methods [1].

2.2.1.1 Prefiltering Methods

The prefiltering methods enhance the quality of reconstructed images by removing the unnoticeable details in source images [1]. Tran et al. [38] added time-domain pre-filter and post-filter modules between two neighboring DCT operators, which were used to smooth data and reduce blocking artifacts. Bruckstein et al. [39] proposed to apply down-sampling operation before the JPEG compression procedure and the final compressed images were generated by up-scaling the lower resolution compressed versions. Experimental results demonstrated that the performance was better in terms of PSNR index. Tsaig *et al.* [40] proposed to use optimal filters in the decimation and interpolation procedures to achieve better performance in terms of both visual quality and quantitative quality. They formulated the problem of finding the optimal filters as a separable least squares problem in two sets of variables and solved it using the variable projection method.

2.2.1.2 Perceptual-based coding Methods

The perceptual-based coding methods modify the encoder based on a human visual system (HVS) model so that the image quality degradation is less observable [1]. Belkasim [41] and Derado proposed a zigzag line preprocessing scheme to reduce blocking artifacts. The original images were transformed into one dimensional arrays and the one dimensional DCT was employed to transform data. Considering image orientation features in a local region, a directional DCT-like transform scheme was derived to improve the compression performance [42, 43]. A direction-adaptive partitioned block transform was designed to alleviate blocking effects. The directional transform modes and non-directional transform modes were adaptively selected according to the

direction information in the block [44]. Nill [45] presented a new method to improve the quality of compressed images by combining the cosine transform with human visual system model [45]. Distortion was intentionally added in the encoding procedure in an iteratively manner to reduce the blocking artifacts [46]. Macq [47] modified the quantization scheme by introducing a quality criterion based on the frequency-weighted SNR to improve the compression performance. Yamatani and Saito [48] developed a polyharmonic local cosine transform (PHLCT), which modified both the encoder and decoder parts of the original JPEG framework, to achieve better visual effects and better PSNR. Eude *et al.* [49] proposed a DCT-based compression framework to remove blocking artifacts. Instead of a universal quantization table, the proposed method employed adaptive quantization tables which were deduced from a distribution model of DCT coefficients.

Since the preprocessing technique operates at the encoder end, it needs to access the original data or control the encoding process. It is not applicable to evidence images, because the images collected from crime scenes are compressed with the JPEG compression method already. It is impossible to control the imaging procedure in crime scenes.

2.2.2 The Postprocessing Methods

The postprocessing technique designed to improve the quality of JPEG compressed images has been studied extensively. Roughly speaking, the existing postprocessing methods are derived from two aspects: image enhancement and image restoration [1]. Considering the properties of the artifacts and human visual system, image enhancement based methods aim to improve the visual quality without optimizing analytical objective functions. While the image restoration based methods try to reconstruct images by using prior knowledge of the JPEG compression method and the compressed images.

2.2.2.1 Image Enhancement based Methods

The image enhancement methods consider blocking artifacts as annoying irregularities, which degrades the perceptual quality. Thereby the image enhancement methods are designed to reduce these artifacts in terms of visual quality improvement.

Postfiltering Methods

Since blocking artifacts are high frequency information, it is straightforward to remove the high frequency components by applying postfiltering methods. Consequently, improvement in image quality is achieved by smoothing the noticeable blocking artifacts. Postfiltering methods can be carried out in the spatial domain and in the frequency domain. Some methods operate in both two domains.

Spatial Filtering Methods: Many spatial postfiltering methods have been proposed to remove the blocking effects and smooth the images. A global low-pass filter is designed to reduce the artifacts along the block boundaries [50]. The linear, space-invariant filters smoothed the images in a consistent manner. As a result, the resultant images were blurry and the original edge components were wiped out as well. To overcome this drawback, adaptive filtering methods were developed. Ramamurthi and Gersho. [51] proposed a nonlinear, space variant filtering method to exploit the local characteristic of the image, so that the artifacts can be smoothed out adaptively. A median filter and a low-pass filter were applied to the images adaptively to reduce blocking artifacts and preserve edges at the same time [52]. Meier et al. [53] developed a region-based filtering method to prevent blurring the edges by applying the low-pass filter to the segmented regions. To improve image quality, the blocking boundaries edges were first distinguished from the natural edges and then removed by a one-time nonlinear smoothing [54]. Lee et al. [55] classified the blocking artifacts into three types of noises and employed a three-stage signal adaptive filtering method to alleviate those blocking effects significantly. To preserve the image details and reduce blocking artifacts, Al-Fohoum et al. [56] designed an adaptive filter using the high-frequency components of the images and applied to the images. The resultant images were further processed to remove blocking artifacts and ensure smoothness without blurring.

Considering the masking effect of the HVS, an adaptive filtering method was proposed in [57]. According to the different characteristics of local regions, the filter involving three different deblocking modes was applied to the decoded images adaptively. Zhai et al. [58] presented a three-step deblocking method to improve the image quality. A local AC coefficient regularization in the DCT domain was employed to reduce blocking effects by modifying the AC coefficients in monotone regions. Then, a block-wise shape adaptive filtering was applied to further suppress artifacts in the spatial domain. Finally, the quantization constraint set was imposed on the previous results to avoid over-smoothing. The method proposed in [59] applied postfiltering in shifted windows (PSW) of image blocks to achieve visual effect improvement and peak signal-to-noise ratio gain. The smoothing operation was only applied to the selected regions, which were thresholded with the MSE between the central block and the image blocks in shifted windows. The optimal MSE threshold was calculated from the compression quality factor and locally adjusted to prevent over-smoothing of image details. Zhai et al. [60] presented an image deblocking method involving integration of quadtree (QT) decomposition followed by the block-shift filtering. The QT decomposition was employed to locate the uniform regions and determine the corresponding suitable block sizes. The variable block sizes reduced the computational cost and large block based shift filtering produced better deblocking results.

Frequency Filtering Methods: Since the JPEG compression process is applied in the frequency domain, it is straightforward to perform filtering operations in the frequency domain. A zero-masking technique was applied to the DCT coefficients of the compressed images to reduce blocking artifacts [61]. Considering the masking effect in the human visual system (HVS), the proposed method in [62] integrated an adaptive weighting mechanism into the postfiltering. Consequently, the blocking artifacts in the low-activity areas were reduced and image details in the high-activity areas were preserved as well. Wang et al. proposed a blind blocking artifact measurement based on the characteristics of HVS, which was used to guide the process of deblocking. To avoid blurring effects, adaptive filter and compensatory coefficient matrices were employed to alleviate blocking effects [63]. A postfiltering method based on contourlet transform was proposed in [64]. By using a scale space edge detector, the blocking map and edge map were

extracted and protected simultaneously. Adaptive filtering was applied to the compressed images in the contourlet domain according to the blocking map. Foi et al. [65] proposed an image filtering method based on the shape-adaptive discrete cosine transform (SADCT) for image deblocking. A local estimate of the signal within the adaptive-shape support was first reconstructed using the hard-thresholded SA-DCT coefficients. By averaging the local estimates adaptively, the estimate of the whole image was generated, which was used as reference for an empirical Wiener filter in the SA-DCT domain. The proposed method reduced blocking artifacts and preserved sharp detail effectively. Singh [66] et al. modelled the artifacts in the frequency domain as 2-D step functions between two neighboring blocks. The smooth and non-smooth regions were identified based on HVS and block statistics. A few DCT coefficients in smooth regions were modified appropriately to remove the blocking artifacts, while an edge-preserving smoothing filter was applied to the non-smooth regions to preserve the genuine edges. The adaptive deblocking method proposed in [67] decomposed images to low-frequency (LF) vector and high frequency (HF) vector. Only the LF vector was processed to remove blocking artifacts because the HF vector contained image details. An adaptive filter was further applied to selected regions to remove blocking artifacts and preserve the original edges. Chang and Chen applied a post-processing filter in the DCT frequency domain. They adopted a shifted window within four adjacent DCT blocks to reduce computational complexity. Several DCT coefficients in the shifted window were modified to eliminate blocking artifacts [68]. An adaptive method was designed in [69] to reduce blocking artifacts based on the correction of a few QDCT coefficients. For the flat region, smoothing operation was applied to the three lowest frequency DCT coefficients for blocking artifact reduction. Then, a block based 3-D filtering was employed to further reduce blocking artifacts in the edge regions.

Hybrid Filtering Methods: Many methods were developed to apply filtering operation in both the spatial and frequency domains. Lai et al. [70] classified the blocking artifacts based on their characteristics and smoothed out the artifacts adaptively. Luo and Ward [71] presented an adaptive approach reducing blocking artifacts in both the spatial domain and the DCT domain. Regions were first classified into smooth regions and texture regions. Subsequently, the

discontinuity of pixels across boundaries was reduced for the smooth regions. And for the texture and edge regions, an edge-preserving filter was employed to avoid over-smoothing. To reduce blocking effects and preserve original edge components, an edge-preserving filter and a low-pass filter were applied to the decoded images simultaneously. The resultant images were generated by combining the two filtered results softly. The proposed method avoided the artifacts introduced by the hard segmentation of the images [72]. The method proposed in [73] improved both visual quality and PSNR index by using an adaptive fuzzy filter. To reduce the artifacts and retain the sharpness of genuine edges, the directional information between pixels and pixel's activity were taken into consideration.

Three types of postfiltering methods are reviewed in this section. The main target of the postfiltering methods is to remove the blocky effects and improve the visual quality of the images. To avoid over-smoothing of the images, various adaptive schemes are adopted to preserve more image details. However, the adaptation to the local statistics is usually sensitive to compression ratio and not robust. Poor estimation and adaptation may cause image quality degradation [1].

Other Image Enhancement based Methods

Besides the postfiltering methods, image enhancement based methods employing other techniques were also proposed to improve the quality of JPEG compressed images.

Paek et al. [74] proposed a DCT-based adaptive method to alleviate the blocking artifacts based on the assumption that the original image was highly correlated. They assumed that the high-frequency components in the global characteristics of the decoded images, which were not found in the local ones, were caused by blocking artifacts and should be removed. A deblocking method applying weighted sums on pixel quartets was proposed in [75]. To prevent the ghosting effect, a weight adaptation by a grading (WABG) scheme was designed. The proposed scheme was applied iteratively to improve the image quality. Choi and Kang proposed a deblocking method using anisotropic diffusion derived from ALM diffusion model. The rate control parameter was utilized to control the diffusion rate in the normal direction of the edges and speed parameter was used to prevent over-smoothing of the texture region [76]. Nosratinia [77]

proposed an image enhancement method to reduce coding artifacts and enhance visual quality. The proposed method recompressed the shifted versions of the JPEG compressed images and generated the final result by averaging these compressed images. Considering human perceptual characteristics, Wong and Bishop [78] proposed a deblocking method based on a phase adaptive, shifted thresholding technique. The original images were estimated as the weighted sum of the shifted versions of the decompressed image, which were generated by an integer transform. The weights were obtained adaptively based on the underlying image content. A structure-texture decomposition based method was integrated into a contrast enhancement framework to suppress the artifacts boosted by image enhancement operations [79].

The main concern of designing image enhancement methods is to match the perception of the human visual system. Thus, the target of the image enhancement methods is to remove the noticeable blocking artifacts rather than restore the original images [1]. Therefore, image enhancement methods are not suitable for restoring JPEG compressed images for forensic analysis.

2.2.2.2 Image Restoration based methods

Some researchers consider JPEG compressed image deblocking as an image restoration problem. With these methods, the blocking artifact removal problem is regarded as an image recovery process based on prior knowledge. A lot of image restoration methods have been designed for removing JPEG compression artifacts. Normally, these methods use some constraints or criteria as the prior knowledge.

Machine Learning Methods

Learning based methods were also used to reduce blocking artifacts in compressed images. Jung et al. [80] developed a deblocking method based on sparse representation. K-singular value decomposition (K-SVD) method was employed to construct a dictionary from the training data. Orthogonal matching pursuit (OMP) method was used to reduce blocking artifacts. Chang et al. [81] proposed a two-step artifact reduction approach based on sparse representations over a

dictionary. First, an over-complete dictionary was learned by using K-SVD method. Then, the total variation (TV) model was employed as a regularizer in the objective function. The function was optimized iteratively by using orthonormal matching pursuit (OMP) and SVD alternatively. An adaptive learning technique was employed to reduce blocking artifacts in [82]. The activity measure, which was defined as the fusing of local entropy and dynamic range, was used to classify the pixels into different groups. A training set comprising the original images and the corresponding compressed counterparts was constructed to train the model. The optimal model was obtained by using the least mean square optimization technique. A learning based method was proposed in [83] to suppress blocking artifacts using neural networks. The inter-block slopes of the compressed image and the difference between the uncompressed image and the compressed image were employed as the input and output respectively. An adaptive kernel regression method was employed in both the frequency and spatial domains to remove blocking artifacts [84].

DCT Coefficient Prediction Methods

Another group of methods solved the deblocking problem by estimating the DCT coefficients directly. Lakhani proposed a smoothing method to reduce blocking artifacts by predicting the DCT coefficients. A set of equations were derived to calculate the DCT coefficients considering the dependence between neighbourhood blocks [85]. Based on the local statistical information, Choy et al. [86] estimated the transform coefficients using a weighted least square technique. The quantization constraint was imposed to guarantee that the final results had the same quantized transform coefficients as the compressed ones. To estimate the DCT coefficients, Lakhani et al. [87] took the definition of the mean squared difference of slopes and derived a set of prediction equations to solve the optimization problem. Triantafyllidis et al. [88] proposed a two-stage blocking artifacts reduction method involving an AC coefficient prediction followed by an adaptive filtering operation. The prediction of AC coefficients was based on the observed probability distribution. A region classification operation was performed prior to the filtering procedure.

MMSE-based Methods

Based on the analysis of the correlation between the pixel values along the block boundary, a new blocking effect measure, which was called Mean Squared Difference of Slope (MSDS), was proposed in [89]. The theoretical and empirical observation demonstrated that the quantization operation increased the expected value of the MSDS. Thus, blocking artefact reduction method was designed to minimize the proposed metric. The blocking effects were removed by solving a constrained quadratic problem. Some researchers proposed deblocking methods based on MSDS or its enhanced versions. Jeon and Jeong [90] designed a blocking artefact reduction method by minimizing the proposed block boundary discontinuity criterion, which was defined as the sum of the squared differences of boundary pixel values in two neighboring blocks. Triantafyllidis et al. [88] proposed a blocking artifact detection and reduction method in the DCT domain. The DCT coefficients were modelled by a Laplacian probability function and the artifacts were detected by calculating the estimated relative quantization error. As the optimization criterion, the mean squared difference of slope was redefined in an enhanced form. For the blocks with artifacts, the DCT coefficients were estimated by minimizing the criterion. Lee et al. [91] presented a two-step non-iterative method to reduce the blocking artifacts. A low-pass filter was first performed to smooth the images. After that, a broken line regression model, modified from the simple linear regression model, was used to construct a set of predictors. According to the local statistical properties, the predictors were applied adaptively to predict the original images from the filtered image. Zhang et al. [92] designed a FIR Wiener filter with an efficient parameter estimation method to restore each quantized DCT coefficient. The dependence of adjacent DCT coefficients was exploited based on the linear minimum mean-square-error (LMMSE) criterion. The designed filter was applied twice in a dual adaptive filtering structure.

Statistical Methods

From the view point of statistical inference, statistical model is a common technique to remove blocking artifacts and restore the original images.

Image reconstruction methods based on maximum a posteriori (MAP) technique were proposed in [93, 94] to improve image quality. They designed a well-posed reconstruction algorithm by using stochastic regularization technique. The Huber-Markov random field model was employed to model image prior. The final optimization problem was formulated as a constrained convex optimization problem and solved iteratively. Luo et al. developed a two-step approach to reduce blocking artifacts based on MAP technique. Based on gradient continuity constraints, a block-wise DC calibration was firstly performed over the block boundaries. Then, the image reconstruction problem was formulated as a convex constrained problem by incorporating a modified Huber-Markov random field model. Two statistical models were employed to deal with pixels in different positions with respect to the block. Localized spatial constraints were imposed by utilizing the iterative conditional model [95]. Li and Kuo [96] proposed an artifact reduction method using a multiscale scheme, in which the processing in the coarse scale removed long range artifacts and the processing scheme in the fine scale kept the sharpness of genuine edges. The Huber-Markov random field model was also used in the same problem formulation. Li et al. [97] investigated a transform domain Markov random field (MRF) model to address the blocking artifact removal problem. The generalized Gaussian MRF (GGMRF) model was used to describe the image prior. The restoration problem was solved in the MAP formulation. A hierarchical Bayesian framework involving two stages was developed to reduce blocking artifacts in the block-based compressed images. First, the dependences of the noise and the reconstructed image models on the hyper-parameters were built. Then, the hyper-prior on the hyper-parameters was defined. After formulating the model in the form of global probability, the parameter estimation problem and image reconstruction problem were solved simultaneously [98]. Based on the prior knowledge that the reconstructed images should have the same quantized coefficients as the original ones, ALTER et al. [99] proposed a deblocking method by minimizing the total variation. Sun et al. [32] addressed the artifact reduction problem by using MAP technique. The artifacts were described as additive, spatially correlated Gaussian noise. A high order Markov random field based on the field of experts (FOE) framework was used to model the image prior. A local MAP estimation method is employed to restore the original image in [100]. The independent and identically distributed Gaussian model and a Gaussian MRF

are built to model the errors between an original image and its compressed image and an original image respectively. The artifact reduction method proposed in [101] formulated the problem as a MAP estimation problem. The prior knowledge of quantization noise and block similarity was exploited to predict the DCT coefficients. The final result was obtained by adaptively combining two predicted results according to their reliabilities.

The Project on Convex Sets (POCS) based Methods

A number of blocking artifact reduction methods were proposed based on the theory of projection onto convex sets (POCS). The basic idea of these methods was to restore the artifact free images by imposing a set of constraints on the coded image. These constraints can be designed based on the properties of the quantization procedure or the statistical analysis of images. The quantization constraint and the image smoothness constraint were two common used constraints in the algorithm design. The quantization constraint was used to ensure the DCT coefficients of the recovered images remained in the original quantization intervals. And the image smoothness constraint was used to filter out the artifacts. Considering different image characteristics, different image smoothness constraints were defined.

Zakhor [102] proposed a POCS based method to reduce blocking artifacts. A filter constraint was derived to alleviate the discontinuities along the boundary. The quantization constraint was also involved. The proposed POCS based artifact reduction method in [103] exploited the prior knowledge about the smoothness of the original images and imposed the smoothness constraint to recover the original images. POCS based adaptive method was developed to restore the images in [104]. Based on the human perceptual characteristics and the local image statistics, a new spatially adaptive smoothness constraint set was proposed. Based on the assumption that the original images are highly correlated, Paek et al. [105] considered the high frequency components in the global characteristics of a distorted image, which do not exist in the local blocks, as blocking artifacts. Thus, a novel smoothness constraint set was imposed to alleviate artifacts based on the theory of POCS in the DCT domain. Weerasinghe et al. [106] proposed a POCS method to suppress blocking artifacts for images containing homogeneous

regions. A soft segmentation method was used to classify the image regions. The convex region smoothness constraint sets, which allowed different degrees of smoothing in different regions, were designed based on the uniformity characteristics of image regions. Different from the previous methods, Zou et al. [107] proposed a POCS-based blocking reduction method by using locally adaptive constraint sets. The characteristics of HVS and relaxed projections were taken into consideration. For the graphic images with smooth regions, the POCS based deblocking method proposed in [108] modelled the smooth regions in the images using a spatially adaptive thin-plate spline. The smoothness constraint set was derived from the explicit model. The smoothness constraint was imposed onto a region other than only the block boundaries. [8] Park *et al.* [109] proposed a POCS based method to recover images from distorted images. A line in Hilbert space was defined by the best matching spatial information from the neighboring blocks. Convex constraint sets were formulated to ensure the continuity between adjacent blocks. Recovery vectors containing known and missing pixels were alternately projected onto the lines and the convex constraint sets [109].

Deep Learning based Method

Recently, some methods based on deep neural networks are proposed to recover the JPEG compressed images. An end-to-end deep convolutional network is used to reduce the compression artifacts. Feature enhancement layers are followed the first layer, which is used to remove the noise. Experimental results demonstrated that the proposed algorithm can suppress blocking artifacts and retain edge patterns and sharp details effectively [110].

A deep convolutional neural network was proposed for image denoising, single image super-resolution and image deblocking. The proposed algorithm utilized residual learning and batch normalization to increase the training speed and improve restoration performance. Experimental results showed that the proposed method perform well not only in terms of image quality but also the running time [111].

A trainable nonlinear reaction diffusion model was proposed for various image restoration problems. With the proposed framework, all the parameters, including the filters and the influence

functions, can be simultaneously learned from training data. Experimental results demonstrated the effectiveness of the proposed algorithm [112].

Other Image Restoration based Methods

There are still many image deblocking methods based on other techniques.

Lee et al. proposed an image recovery method based on fuzzy logic technique. The low frequency information and the high frequency information of the blocks were recovered by a modified surface fitting approach and the fuzzy logic reasoning respectively. And the final restored images were generated by combining the two results carefully [113]. A modified classified vector quantization approach was employed to improve image quality in [114]. The proposed method comprised an encoding phase and a decoding phase. A codebook was used in the encoding phase to transform the coded image into a set of indices. The indices were used to enhance the image quality using another codebook in the decoding phase. Motivated by the independent quantization in the JPEG compression method, the blocking artifact suppression method proposed in [115] modelled the quantization noise as a colored Gaussian process in the spatial domain. A nonstationary image prior model was incorporated to reduce artifacts and preserve the edge components simultaneously. To fulfil the convex requirement, a regularization function was employed. Guo et al. [116] addressed the blocking artefact removal problem by exploiting the quantization constraints and the boundary constraints. The problem was formulated as a convex optimization problem and solved.

The above image restoration methods are developed to restore the original images by importing general constraints or prior knowledge from the original images [1]. However, they are designed for restoring generic images and none of them takes the special prosperities of skin images into consideration. Thus, they are not effective to restore the biometric traits from JPEG compressed skin images.

A Knowledge-based (KB) Deblocking Method for Skin Images

The JPEG deblocking methods mentioned before are developed for generic images. Researchers show that they are ineffective to recover skin features, such as skin marks for forensic analysis [30]. To our best knowledge, the knowledge-based (KB) [30, 117] method is the only one that is specially designed for removing blocking artifacts in skin images. The KB method searches optimal uncompressed blocks in a large skin image database to replace the compressed blocks in input images. The database stores uncompressed skin images and the corresponding compressed skin images from a particular pair of JPEG quantization tables. The KB method operates in the Y, U, and V channels separately. The combined spatial and frequency information from the target block and its neighbor blocks are employed as the searching criterion. The frequency information from the target block guarantees that the resultant images and the compressed images have the same QDCT coefficients and the frequency and spatial information is used as smooth constraints to search the optimal original block in the training set.

Experimental results demonstrate that this method can effectively restore images with skin marks and outperform other deblocking methods designed for generic images [32, 54, 119]. Figs. 2-4 (a) and (b) show respectively an original image and the corresponding JPEG compressed image. Fig. 2-4 (c) shows the resultant image from the KB method. The figure indicates that the KB method can restore color image with skin marks effectively.

Fig. 2-4 An example from the KB method.

(a) is an original image, where the red circle highlights a skin mark, (b) is the corresponding JPEG compressed image of (a) and (c) is the resultant image from the KB method.

Current methods for JPEG image deblocking have been reviewed. Although researchers have shown that their methods are effective in terms of mean-square error (MSE) or HVS-model-based quality metric, they, except for the KB method, are not designed to recover skin features e.g. blood vessel patterns. The KB method is effective for skin marks, which have been tested

extensively. However, its recovery capability for blood vessel patterns hidden in color images has not been tested. Our experimental results (see Chapters 4 and 5) indicate that it cannot recover blood vessel patterns from JPEG compressed images because it operates on the three channels separately and uses very limited information across blocks. In other words, it does not take the dependences between channels and between blocks into consideration.

The state-of-the-art deblocking methods, including Foi *et al.*'s pointwise shape-adaptive DCT method (SADCT) [65], which improves image quality both in terms of objective criteria and visual appearance, Sun *et al.*'s maximum a posteriori method based on a field of experts prior (FOE) [32], which achieves higher PSNR gain and good visual quality, Tang *et al.*'s knowledge-based method (KB) [31], which is specially designed for removing blocking artifacts and restoring biometric traits (e.g., skin marks) in skin images, Zhang *et al.*'s overlapped-block transform coefficient estimation (OBTCE) [101] with block similarity method, which is one of the latest deblocking methods, Florentín-Núñez *et al.*'s adaptive kernel regression (AKR) [84] based method, which operates in the both frequency and spatial domains, Li *et al.*'s blocking artifact suppression method, which is specially designed for contrast enhancement framework (CEFAS) [79], Dong *et al.*'s artifact reduction convolutional neural network (ARCNN) [110], which uses a feature enhancement layer to perform artifact reduction, are used for comparisons. Another two latest methods based on deep neural network, including Zhang *et al.*'s end-to-end deep denoising convolutional neural network (DnCNN) [111] and Chen *et al.*'s trainable nonlinear reaction diffusion framework (TNRD) [112], which are designed for denoising and can be applied for deblocking, are utilized for comparisons in this thesis. Table 2-1 summarizes the existing methods which are used for comparison in this thesis.

Table 2-1 The summary of the existing methods used for comparison.

	Operating domain	Technique	Application-based	Use case
FOE [32]	Spatial	Statistical model	No	Blood vessel pattern
SADCT [65]	Frequency	Postfiltering methods	No	Blood vessel pattern
KB [31]	Frequency	Learning based method	Yes	Blood vessel pattern
AKR [84]	Spatial & Frequency	Learning based method	No	Blood vessel pattern
OBTCE [101]	Frequency	MAP	No	Blood vessel pattern and tattoo image
CEFAS [79]	Spatial	Image enhancement method	No	Blood vessel pattern
ARCNN [110]	Spatial	Deep learning	No	Tattoo image
DnCNN [111]	Spatial	Deep learning	No	Tattoo image
TNRD [112]	Spatial & Frequency	Deep learning	No	Tattoo image

2.3 A Blood Vessel Pattern Uncovering Method

In the latter chapters, to evaluate the performance of the proposed image restoration method, we will employ a blood vessel uncovering method to visualize the blood vessel patterns from color skin images. Tang et al. designed an uncovering method to visualize blood vessel patterns hidden in color images [25, 26]. They used the Kubelka-Munk theory to model the inverse process of skin color image formation and derive spatial distributions of biophysical parameters from color images. They assumed that the optical properties of human skin were determined by three layers - the stratum corneum, the epidermis, and the dermis and blood vessels were located in the dermis. The idea of uncovering blood vessel patterns hidden in color images is to compute the spatial distributions of the volume fraction (%) of the epidermis occupied by melanosomes (denoted as ϑ_m), the volume fraction (%) of the dermis occupied by blood (denoted as ϑ_p), and the depth of the dermis (denoted as d^{der}). Blood vessel patterns can be observed from the spatial distributions of these parameters.

Based on the principles of optics and skin biophysics, the blood vessel uncovering method built a model to simulate the formulation of skin color. To visualize blood vessel patterns, it computed the inverse of the skin color formulation process. A three-layered feed-forward neural network was trained to approximate this inverse process. The inputs of the model were the RGB

values from a test image, and the outputs were used to form the distribution maps of θ_m, θ_p and d^{der} . The model depended on a set of biophysical parameters, which were originally found in medical papers. These parameters were measured from ideal medical settings. To handle images from uncontrolled environment and avoid over-dependence on these parameters, Tang et al. developed a color optimization scheme. Considering the characteristics of the skin image database, they selected a typical skin image and used its average RGB value as a target. Then they obtained the optimal parameters in the model of skin color formation with this target value. They adjusted the parameter ranges from 50% to 200% of the optimal values in actual computation. Fig. 2-5 shows the parameter range optimization process.

Fig. 2-5 Illustration of parameter range optimization process.

Then they obtained the corresponding RGB values of these parameters by simulating the process of skin color formation. The data pairs were used as training data to train the three-layered feed-forward neural network. Fig. 2-6 illustrates the process of training data preparation for the neural network.

Fig. 2-6 Data preparation for training a three-layered feed-forward neural network.

Fig. 2-7 Blood vessel uncovering results.
(a) is a color image of the right forearm; (b), (c), and (d) are, respectively, the distribution maps of melanin, hemoglobin, and the depth of dermis obtained from the blood vessel uncovering method.

Fig. 2-7 gives resultant images from the blood vessel uncovering method. Among the three distribution maps, the distribution map of the depth of dermis is the clearest. Therefore, the distribution map of the depth of dermis is used for blood vessel pattern matching.

2.4 Blood Vessel Pattern Extraction and Automatic Matching

Methods

In this subsection, we will introduce a blood vessel pattern extraction method and a three step automatic blood vessel matching method, which were developed by Zhang et al. [26]. Later, we will use these two methods to evaluate the proposed algorithm. The images that this method deals with are the results generated by the blood vessel uncovering method introduced in last subsection.

2.4.1 A Blood Vessel Pattern Extraction Method

The blood vessel pattern extraction method includes four parts: preprocessing, local information estimation, blood vessel enhancement, and blood vessel representation.

2.4.1.1 Preprocessing and Local Information Estimation

Before extracting blood vessel patterns, uncovered images are involved in a histogram equalization method to normalize the contrast of images. Then, Gabor filter is employed to estimate the local directional information of blood vessels, and two information maps are generated. A Gabor filter bank made up of the real parts of 16 Gabor filters with different scales and orientations is applied to the blood vessel images. A real part of a Gabor filter in the spatial domain is defined as

$$G(x, y, \lambda_{mk}, \theta_k, \sigma_m, \gamma) = \frac{\gamma}{2\pi\sigma_m^2} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma_m^2}\right) \cos\left(2\pi \frac{x'}{\lambda_{mk}}\right), \quad (2-4)$$

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

$$F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y) = -G_{DC}(x, y, \lambda_{mk}, \theta_k, \sigma_m, \gamma) * Img(x, y), \quad (2-5)$$

where $*$ represents an operation of a two-dimensional convolution.

For each input image, 16 filtered images are generated by the filter bank. The orientation and scale of the blood vessel can be respectively represented by the direction and scale of the Gabor filter giving the maximum response. Since the filter response depends on image contrast, a local image power around the point (x, y) , Q_m , is used. Q_m is defined as the square root of the summation of pixel value square around the point (x, y) . Thus, $F_{\lambda_{mk}, \theta_k, \sigma_m, \gamma}(x, y)/Q_m$ serves as a local information estimator [120]. The orientation of blood vessels can be obtained from

$$O(x, y) = \underset{\theta_k}{\operatorname{argmax}}_{m, k} F_{\gamma, \sigma_m, f_{mk}, \theta_k}(x, y), \quad (2-6)$$

which is defined as the orientation map. In addition,

$$E(x, y) = \max_{m, k} [F_{\lambda_{mk}, \theta_k, \sigma_m \gamma}(x, y) / Q_m(x, y)], \quad (2-7)$$

where $\lfloor \cdot \rfloor$ represents a floor operation, is defined as the response map. It indicates the quality of local information. Figs. 2-8 (b) and (c) illustrate the Gabor filter outputs. Blood vessel patterns can be observed from the response and orientation maps.

Fig. 2-8 The Gabor filter outputs.

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

2.4.1.2 Enhancement and Representation

Since response maps generated from uncovered images might be very noisy, structural information of blood vessels is extracted from orientation maps to enhance the response maps. Zhang et al. use skin image characteristics to design an enhancement scheme. Pixels with the similar orientation are grouped into one component and connected through the connected component labelling method. A threshold is applied to the response map to distinguish the potential skin pixels and potential blood vessel components. Then the potential blood vessel components in two groups are further connected if their angular distance is shorter than or equal to $\pi/8$ and their minimum distance is shorter than a threshold. All component pairs are searched and assigned the same label to connect blood vessel components iteratively to generate a map of grouped labels Ψ . $g(i) = \{(x, y) | \Psi(x, y) = i\}$ represents pixels with the same label i and $|g(i)|$

is its size in terms of number of pixels. A weighting function $W(x, y)$ based on the size of the groups is defined as

$$W(x, y) = \kappa \cdot \left\{ 1 + \exp \left(-\frac{|g(i)|}{\mu} \right) \right\}^{-1}, \quad (2-8)$$

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 point-wisely applied to the response map. The Otsu's method [121] is used to binarize the enhanced blood vessel patterns and the resultant image is further skeletonized. Finally, the blood vessel patterns and boundaries of body parts are uniformly sampled and represented by a set of points. Fig. 2-9 illustrates the blood vessel pattern extraction process.

Fig. 2-9 The process of blood vessel pattern extraction.

(a) is an uncovered image; (b) is the contrast-limited adaptive histogram equalization (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 blood vessel patterns in (g), and (i) shows the sampled points from both the blood vessel patterns and the body boundaries.

2.4.2 A Three Step Blood Vessel Matching Method

A three step blood vessel matching method was designed to measure the dissimilarity of two blood vessel patterns. Given two sets of points in $\mathbb{R}^2$ representing two blood vessel patterns, 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. M_{BD} with size of N_{mbd} by 2 and D_{BD} with size of N_{dbd} by 2 are used to denote the corresponding body boundaries. $TR(\cdot, \tau R)$ and $TN(\cdot, \tau N)$ respectively represent a rigid transformation and a non-rigid transformation controlled by parametric vectors τR and τN . These

transformations are used to align the two point sets. The first step removes 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 sets.

To match two point sets, they should be first aligned roughly. The shape information from boundaries of body parts are used for rough alignment. A rigid transformation TR is applied to the two sets of boundary points M_{BD} and D_{BD} to align the two boundaries. The alignment parameter vector τ_1 is obtained through the rigid Coherent Point Drift (CPD) method [122]. The blood vessel pattern point set M is transformed into $MT_1 = TR(M, \tau_1)$. Then the overlapping area B of MT_1 and D is determined. Points located outside B are considered as outliers and removed. The pruned blood vessel pattern point sets are denoted as M_1 and D_1 . Two sets are considered as two different blood vessel patterns, if the numbers of points in these two sets are less than a threshold

The second step is to separately remove outliers inside M_1 and D_1 . In the previous step, M_1 and D_1 are roughly aligned. M_1 is regarded as a model set in this alignment. The resultant parameter vector τ_{21} and the correspondence c_{21} between M_1 and D_1 are determined by the rigid CPD method. Based on the alignment of blood vessel patterns $TR(M_1, \tau_{21})$ and D_1 , the distances between corresponding points in $TR(M_1, \tau_{21})$ and D_1 are calculated. The corresponding points with distances longer than a defined threshold in $TR(M_1, \tau_{21})$ are considered as outliers. All the points in $TR(M_1, \tau_{21})$ are sorted in ascending order according to their distances. The top 80% of points in $TR(M_1, \tau_{21})$ are retained and the others are considered as outliers and removed. The remaining points are denoted as M_2 . The same process is repeated to remove outliers in D_1 . The rigid CPD method is applied to align D_1 and M_2 , while D_1 is considered as a model set. With the same procedure, based on the ascending sort of the distances between corresponding points in $TR(D_1, \tau_{22})$ and M_2 , only the top 80% of points in D_1 are chosen. The remaining points in D_1 are denoted as D_2 .

The third step is to match M_2 and D_2 with a non-rigid transformation. The non-rigid CPD method with a Gaussian kernel is used to match M_2 and D_2 in this step. The non-rigid transformation TN and corresponding parameters τ_3 are generated by the CPD method,

and $TN(M_2, \tau_3)$ and D_2 are aligned optimally. As the same 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 considered as outliers.

Finally, the correspondent vector c_3 generated in the third step is used to calculate the dissimilarity of two blood vessel patterns. Let (x_{mi}, y_{mi}) in $TN(M_2, \tau_3)$ and (x_{di}, y_{di}) in D_2 be correspondences in c_3 . Their Euclidean distance denoted as e_i and r is a set containing the smallest 80% of e_i . The dissimilarity function is defined as $s = |r|^{-1} \times \sum_{e_i \in r} e_i$, where $|r|$ represents the number of elements in the set r . Fig. 2-10 illustrates the process of the three step blood vessel matching method.

Fig. 2-10 The three step matching results.

(a) and (b) are blood vessel patterns uncovered from color images from same person collected at different occasions; (c) is the points sampled from two blood vessel 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.

The blood vessel uncovering method and the blood vessel pattern extraction and matching methods are reviewed in sections 2.3 and 2.4. The blood vessel uncovering method is proposed to directly visualize the blood vessels from the skin color images based on the principles of optics and skin biophysics and the blood vessel pattern extraction method is used to extract the blood

vessel patterns from the uncovered results. The followed automatic blood vessel matching method is developed to measure the similarity/dissimilarity of two blood vessel patterns. These methods are utilized in chapters 4 and 5 to evaluate the performance of the proposed image restoration algorithms.

2.5 Image Quality Assessment Metrics

Objective image quality assessment metrics are designed to quantify image quality based on mathematical or perceptual models. In this thesis, two full reference image quality metrics, structural similarity (SSIM) index [123] and peak signal-to-noise ratio (PSNR) index [124], are employed to evaluate the visual comparisons objectively. The introduction to these two metrics is given in this section.

PSNR index is the most common used metric for measuring image and video quality because of its lower computation complexity and higher computation speed. It is obtained by calculating the logarithmic ratio of maximum possible power of pixel value and mean square error (MSE) [125], which is defined as,

$$PSNR = 10 \log_{10} \left(\frac{peakval^2}{MSE} \right) \quad (2-9)$$

where *peakval* is possible peak value of the pixel and *MSE* is defined as,

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i,j) - y(i,j)]^2 \quad (2-10)$$

where *x* is a given image with size of $m \times n$, and *y* is the distorted version of *x*.

The range of PSNR index is 0 to 100. A higher PSNR index value indicates that the image is of higher quality and PSNR index value of 100 means that the test image is identical to the reference image.

Since the PSNR index is computed based on the pixel by pixel difference, it is not well matched to the HVS model. To mimic HVS model better, Wang et al. proposed the SSIM index

based on different properties of HVS model [123]. SSIM is a perception-based model and it calculates the image quality degradation in terms of the perceived change in luminance, contrast and structure information. Comparing to the PSNR index, the computation of SSIM index is more complex. For two given images, x and its distorted version y , the SSIM index between the two images is defined as,

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma, \quad (2-11)$$

where $\alpha > 0$, $\beta > 0$, $\gamma > 0$ and $l(x, y)$, $c(x, y)$ and $s(x, y)$ are the luminance comparison, contrast comparison and the structure comparison. They are respectively defined as,

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (2-12)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (2-13)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (2-14)$$

where μ_x , μ_y , σ_x and σ_y are the local means, standard deviations and cross-covariance for image x and y . C_1 , C_2 and C_3 are predefined constants.

$$C_1 = (K_1L)^2, C_2 = (K_2L)^2, C_3 = C_2/2, \quad (2-15)$$

where L is dynamic range of pixel values and $K_1, K_2 \ll 1$.

Test image without distortion has the SSIM index value of 1. Greater values of the SSIM index indicate that the test image and its reference image are more similar.

In this section, the key points of the PSNR index and the SSIM index are given.

2.6 Summary

In this chapter, the critical steps of JPEG compression method are introduced. The existing deblocking methods are reviewed. Most of them are designed for generic images, except for the KB method, which is specially designed for removing the blocking artifacts and restoring skin

marks in JPEG compressed skin images. A brief introduction to the blood vessel pattern uncovering and the blood vessel pattern extraction and matching methods, which are used to evaluate the effectiveness of the proposed algorithms later, is given. Two objective image quality metrics, which are used to measure the performance of the proposed algorithms, are listed.

Chapter 3 Skin Image Analysis and Compression Test for Blood Vessel Restoration

Image deblocking has been an active research topic for several decades. The target of image deblocking for the generic images is to remove the blocking artifacts and improve the visual quality of the images. A lot of studies and analysis have been conducted to fulfil the tasks. Based on the well understanding of the problem, many techniques are proposed to alleviate the blocking artifacts and generate images with higher visual quality. Existing deblocking methods for generic images are reviewed in Chapter 2. However, they are not suitable for recovering the blood vessel information from the JPEG compressed skin images, because they are designed for generic images and for different purpose.

As a specific type of image, skin images are supposed to have some properties different from the generic images. Comparing to the generic images, skin image textures are more homogeneous and skin image color information concentrate in a relatively fixed range. Some of existing methods only process the Y channel based on the assumption that human visual system is more sensitive to the Y channel. But, this assumption is ineffective for blood vessel information restoration. The blood vessel information existing in the skin images are invisible in the original skin color images, which are different from the visual information. By using some properties of skin images, the KB method is specially for recovering the skin marks from JPEG compressed skin images. But it is not effective for our purpose, because it does not take the blood vessel information into consideration.

To develop an algorithm for blood vessel information restoration in the skin images, the properties of the skin images and blood vessel information should be studied. In addition, it should be figured out how the JPEG compression method degrades the quality of blood vessel information. The aim of this chapter is to systematically study the properties of skin images and identify critical factors influencing quality of blood vessel information in JPEG compressed images. A public skin image database used in this study is first introduced. Then, a skin analysis

is conducted to find out how the blood vessel information distributes in different channels. After that, a compression test is performed to find out the critical steps in the JPEG compression method degrading the quality of the blood vessel information. To measure the quality of the blood vessel information, blood vessel uncovering and matching methods reviewed in Chapter 2 are used to uncover the blood vessel patterns and generate matching results in the skin image analysis and compression test. The cumulative match characteristic (CMC) curves are used as matching performance index. Simulations using different compression ratios are performed on skin images with different resolutions. A statistical analysis of QDCT coefficients is also conducted to analyze the existing blood vessel information and the distribution of the QDCT coefficients in the compressed skin images. The findings are used to design the algorithms in Chapters 4 and 5 to restore the lost blood vessel information.

In this chapter, Section 3.1 introduces a public skin image database used for the skin image analysis and compression test. Section 3.2 presents the skin image analysis. Section 3.3 gives the skin image compression test. Statistics on skin images is introduced in Section 3.4. Section 3.5 lists some concluding remarks.

3.1 A Public Skin Image Database

A public skin image database (NTU Forensic Image Database (NTUFID), including NTU Inner Forearm Image Database v1 and NTU Inner Thigh Image Database v1), [126] is used to perform the skin image analysis and the compression tests. To test images with different compression levels, the compression quality factors of 50 and 75 were used in the experiments. These two quality factors were selected because some skin images collected from the Internet had quality factors lower than 75 and they are commonly employed to evaluate deblocking algorithms for generic images. The method reported in [26] was used to visualize blood vessels in these chapter.

Fig. 3-1 Raw sample images from the database.
(a) is from the forearm datasets and (b) is from the thigh datasets.

NTUFID has four parts: a left forearm dataset, a right forearm dataset, a left thigh dataset and a right thigh dataset. Each limb in these datasets has two images, collected in two sessions with an interval of 11 days. Images in the left and right forearm datasets were, respectively, collected from 239 and 250 subjects. Images in the left and right thigh datasets were, respectively, collected from 230 and 228 subjects. In total, the database contains 1,894 images. The imaging configuration, such as the illumination and the background, was changed in the data collection. The camera model is Nikon D70s, and the images are taken under normal daylight or fluorescent light. The poses of the subjects were not strictly restricted. The subjects were different in races and ages. Their races include Chinese, Malay, Indian and Japanese. The collected images were saved in the JPEG format with a high quality factor 99 and without noticeable quality degradation. To construct testing and training databases for algorithm development and evaluation, the limbs were extracted from the raw images and aligned. Then, the limbs were

further resized to remove the compression effects and referred to as the original (uncompressed) images. Fig. 3-1 gives some raw images from the forearm datasets and the thigh datasets, respectively. The raw images show that the data collection process was not strictly controlled. It should be emphasized that evidences in cases of child sexual abuse and other sexual offenses are often high resolution close-up images.

To obtain images with different resolutions, skin images in NTUFID are further downsampled to generate a lower resolution image database. To differentiate them, the database with the original scale is denoted as NTUFID_1 and the lower resolution one is denoted as NTUFID_2. In this thesis, the images in NTUFID_1 and NTUFID_2 are compressed with a quality factor of 50 and 75, respectively. Table 3-1 summarizes the databases. More details about images in each session in the databases are listed in Table 3-2.

Table 3-1 A summary of the databases

Datasets	Abbreviation	Number of limbs	NTUFID_1		NTUFID_2	
			<i>Length (pixels)</i>	<i>Width (pixels)</i>	<i>Length (pixels)</i>	<i>Width (pixels)</i>
Left forearm dataset	LF	239	958	393	480	185
Right forearm dataset	RF	250	1018	371	480	163
Left thigh dataset	LT	230	1127	651	512	288
Right thigh dataset	RT	228	1093	620	512	286

Table 3-2 A summary of each session in the databases

Datasets	Abbreviation	Number of images	NTUFID_1		NTUFID_2	
			<i>Length (pixels)</i>	<i>Width (pixels)</i>	<i>Length (pixels)</i>	<i>Width (pixels)</i>
The first session of LF	LF1	239	960	396	480	186
The second session of LF	LF2	239	955	390	480	184
The first session of RF	RF1	250	1017	368	480	162
The second session RF	RF2	250	1019	374	480	163
The first session of LT	LT1	230	1124	649	512	289
The second session of LT	LT2	230	1129	653	512	287
The first session of RT	RT1	228	1082	628	512	291
The second session of RT	RT2	228	1103	612	512	281

Fig. 3-2 gives four uncompressed images from NTUFID_1 as examples. They are from the data sets LF, RF, LT and RT respectively. Fig. 3-2 shows the original color images (the first row), their blood vessels given by the visualization method (the second row) [26], their JPEG compressed versions with a quality factor of 50 (the third row) and the blood vessels from the JPEG compressed versions (the last row). Fig. 3-3 shows the lower resolution version of the

images in Fig. 3-2 from NTUFID_2. Fig. 3-3 (a) is the original images and Fig. 3-3 (b) gives the corresponding blood vessel images. Fig. 3-3 (c) shows the JPEG compressed images of Fig. 3-3 (a) with a quality factor of 75. Fig. 3-3 (d) shows the corresponding blood vessel images. From these figures, it can be seen that the blood vessels from the original images are clear, but they are seriously degraded by the JPEG compression method.

Fig. 3-2 Four example images from NTUFID_1.

(a) is the original images. (b) is the corresponding blood vessel images from (a) [26]. (c) is the JPEG compressed versions of (a) with a quality factor of 50. (d) shows the blood vessel images from (c).

Fig. 3-3 The lower resolution version of Fig. 3-2.

(a) is the original images. (b) is the corresponding blood vessel images from (a) [26]. (c) is the JPEG compressed versions of (a) with a quality factor of 75. (d) shows the blood vessel images from (c).

3.2 A Skin Image Analysis

Since the JPEG method processes images in the YUV color space, intuitively, the recovery process should be operated in this space. Many of the existing methods, which are designed for recovering generic images, only process the luminance channel (the Y channel) to improve the image quality, because the Y channel carries most of the visual information and human visual system is less sensitive to the information loss and blocking artifacts in the chrominance channels (the U and V channels). However, to recover the Y channel only may not be enough to recover the blood vessel information. The blood vessel information is not visible in the color images, which is different from the visual information, obviously. Thus, it is necessary to know how blood vessel information distributes in the Y channel and the U and V channels. To perform this analysis, the uncovering and matching methods described in [26] are used to evaluate blood vessel information in different channels. In the first test, the information in the chrominance channels is removed, but the information in the luminance channel is retained. In the second test, the information in the chrominance channels is retained, but the information in the luminance channel is removed. By comparing the matching results in these two tests, it can be known which channels contain more blood vessel information. Eq. 3-1 is the transformation equation from the YUV color space to the RGB color space [127] in a range of [0, 255],

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0 \end{bmatrix} \begin{bmatrix} Y - 16 \\ U - 128 \\ V - 128 \end{bmatrix}. \quad (3-1)$$

If both the U and V values are 128, Eq. 2 can be rewritten as,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0 \end{bmatrix} \begin{bmatrix} Y - 16 \\ 0 \\ 0 \end{bmatrix}. \quad (3-2)$$

Eq. 3-2 indicates that the information in the U and V channels are eliminated and only the Y channel contributes to the RGB values. In other words, if skin images are processed by Eq. 3-2, the uncovered blood vessels will be only from the Y channel. Thus, in the first test, images in the database are first converted into the YUV space. Then the pixel values in the U and V channels are set to 128 and the images are converted back to the RGB space. It is worth to mention that the

blood vessel uncovering method used in this analysis and the latter experiments takes RGB images as inputs [26]. Similarly, in the second test, the pixel values in the Y channel are set to 16, but the pixel values in the U and V channels are retained.

Fig. 3-4 Resultant images of the image analysis from the images shown in Fig. 3-2 (a). (a) the results from the analysis on the Y channel and (b) shows the results from the analysis on the U and V channels.

Fig. 3-5 Lower resolution resultant images of the image analysis from the images shown in Fig. 3-3 (a). (a) the results from the analysis on the Y channel and (b) shows the results from the analysis on the U and V channels.

After generating the testing images, the blood vessel uncovering method [26] is applied. Fig. 3-4 shows the results from the images given in Fig. 3-2 (a). Only a few blood vessels from the Y channel (Fig. 3-4 (a)) are observable, while the blood vessels (Fig. 3-4 (b)) from the U and V channels are close to the ones from the original images (Fig. 3-2 (b)), except for some details. It indicates that the three channels all contain blood vessel information, but the Y channel carries less information than the U and V channels. The same conclusion can be drawn from the lower resolution images which are shown in Fig. 3-5.

Fig. 3-6 The matching results of NTUFID_1 for the image analysis.

(a)-(d) are CMCs from LF1, RF1, LT1 and RT1 in NTUFID_1, respectively. In the legends, JPEG50 means JPEG images with a quality factor of 50, onlyYChannel means the images with only the Y channel information, onlyUVChannel means the images with only the U and V channel information.

To determine the quality of the blood vessel patterns in different channels, instead of subjective visual observation, a three-step blood vessel matching method [26] measuring dissimilarity between two patterns is used. Cumulative match curves (CMCs) are employed as a performance index. In this analysis, the original images in NTUFID_1 and NTUFID_2 collected in the first session are, respectively, matched with their JPEG compressed versions with quality factors of 50 and 75 as references. The images with only Y channel information and the images with only U and V channel information are also matched with the original images. In these matching test, the original images are used as gallery sets and the other images are used as probe sets.

Fig. 3-7 The matching results of NTUFID_2 for the image analysis.

(a)-(d) are CMCs from LF1, RF1, LT1 and RT1 in NTUFID_2, respectively. In the legends, JPEG75 means JPEG images with a quality factor of 75, onlyYChannel means the images with only the Y channel information, onlyUVChannel means the images with only the U and V channel information.

Fig. 3-6 and Fig. 3-7 show the CMCs generating from the testing images in the first session of NTUFID_1 and NTUFID_2, respectively. The results indicate three important points. The JPEG method seriously deteriorates the blood vessel quality. The images with only the U and V channel information provide much higher matching accuracy than those with only the Y channel information. The U and V channels contain much more information than the Y channel, but they all contain blood vessel information, which agrees with the previous visual comparison. Thus, to recover the original blood vessel information, restoration operations should be applied to all the three channels.

Fig. 3-8 The matching results of NTUFID_1 for the image analysis.

(a)-(d) are CMCs from LF2, RF2, LT2 and RT2 in NTUFID_1, respectively. In the legends, JPEG50 means JPEG images with a quality factor of 50, onlyYChannel means the images with only the Y channel information, onlyUVChannel means the images with only the U and V channel information.

The images in the second session of NTUFID_1 and NTUFID_2 are also matched in the same manner. The results are shown in Fig. 3-8 and Fig. 3-9, which indicate the same three important points. Cross session matching (e.g. the first session vs. the second session) is not performed because the aim of these tests is to study the distributions of blood vessel information in the three channels, but not to evaluate the visualization and matching algorithms.

Fig. 3-9 The matching results of NTUFID_2 for the image analysis. (a)-(d) are CMCs from LF2, RF2, LT2 and RT2 in NTUFID_2, respectively. In the legends, JPEG75 means JPEG images with a quality factor of 75, onlyYChannel means the images with only the Y channel information, onlyUVChannel means the images with only the U and V channel information.

3.3 A Skin Image Compression Test

A compression test is further conducted to identify how the lossy operations (which are the down-sampling operation and the quantization operation highlighted by the red dash rectangles in Fig. 2-1) in the JPEG method degrade the blood vessel pattern quality. In this test, the down-sampling operation and the quantization operation are examined separately.

Fig. 3-10 shows the compression processes in the test. The compression process given in Fig. 3-10 (a) only applies the down-sampling procedure to the U and V channels and the original Y channel is preserved, so that its effect can be observed. The compression process in Fig. 3-10 (b) is designed to figure out how the quantization operation affects the blood vessel pattern quality. It shows that the Y, U and V channels are quantized with the quantization tables Q_{yi} and

Q_{uvi} after the down-sampling operation in the U and V channels. To identify critical DCT coefficients, the standard quantization tables are modified to generate three pairs of quantization tables. Fig. 3-11 gives the quantization tables Q_{yi} for the Y channel and Q_{uvi} for the U and V channels, where $i \in \{1,2,3\}$. They are modified from Q_y and Q_{uv} in Figs. 2-2 (b) and (c) which are used to generate compressed images with a quality factor of 50. Comparing with the original Q_y and Q_{uv} , it can be noted that several original values are modified to '1's, which means that the corresponding DCT coefficients only involve in the round operation. For each pair of Q_{yi} and Q_{uvi} , the number and the position of '1' are same. The DCT coefficients are organized in a zigzag order (Fig. 2-2 (a)) according to their frequency. The modified values in the quantization tables are in the same order. The same modification scheme is applied to the original quantization tables of a quality factor of 75 to generate three pairs of quantization tables for the images from NTUFID_2.

Fig. 3-10 The processes of the compression test.

(a) is to evaluate the down-sampling procedure in the U and V channels. (b) is to evaluate the quantization operation in the Y, U and V channels.

1	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	103	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(a)

1	1	10	16	24	40	51	61
1	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	103	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(b)

1	1	1	16	24	40	51	61
1	1	14	19	26	58	60	55
1	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	103	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(c)

1	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(d)

1	1	24	47	99	99	99	99
1	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(e)

1	1	1	47	99	99	99	99
1	1	26	66	99	99	99	99
1	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

(f)

Fig. 3-11 Quantization tables modified from Figs. 2-2 (b) and (c).

(a)-(c) are the quantization tables for the Y channel, which are respectively denoted as Q_{y1} , Q_{y2} and Q_{y3} , and (d)-(f) are the quantization tables for the U and V channels, which are respectively denoted as Q_{uv1} , Q_{uv2} and Q_{uv3} .

Fig. 3-12 Resultant images of NTUFID_1 from the compression test given in Fig. 3-10(a).

(a) is the resultant color images of NTUFID_1 from the compression process given in Fig. 3-10(a) and their original images are given in Fig. 3-2(a). (b) shows the corresponding blood vessel patterns from the compressed images.

Fig. 3-13 Resultant images of NTUFID_2 from the compression test given in Fig. 3-10(a).

(a) is the resultant color images of NTUFID_2 from the compression process given in Fig. 3-10(a) and their original images are given in Fig. 3-3(a). (b) shows the corresponding blood vessel patterns from the compressed images.

The two databases are used in this test and the two compression processes in Fig. 3-10 are applied to the images. Figs. 3-12 and 3-13 give the testing results of NTUFID_1 and NTUFID_2 from the compression process in Fig. 3-10 (a). Fig. 3-12 shows the resultant color images, whose original images are given in Figs. 3-2 (a), and the corresponding blood vessel patterns. Fig. 3-13 shows the resultant color images, whose original images are given in Fig. 3-3 (a), and the corresponding blood vessel patterns. Figs. 3-12 (b) and 3-13 (b) indicate that the blood vessel patterns from the compressed images are similar to those from the original images (Figs. 3-2 (b) and 3-3 (b)). There is no obvious degradation comparing to Figs. 3-2 (b) and 3-3 (b). From this test, it can be seen that the down-sampling operation in the U and V channels does not affect the visual quality of blood vessels significantly.

Fig. 3-14 Results of images in NTUFID_1 from the compression process in Fig. 3-10(b). (a) is generated by $\{Q_{y1}, Q_{uv1}\}$, (c) is generated by $\{Q_{y2}, Q_{uv2}\}$ and (e) is generated by $\{Q_{y3}, Q_{uv3}\}$. (b), (d) and (f) are respectively blood vessel patterns.

Figs. 3-14 (a), (c) and (e) show the results of images from NTUFID_1 generated by the compression process in Fig. 3-10 (b). They are corresponding to the three modified quantization table pairs, $\{Q_{y1}, Q_{uv1}\}$, $\{Q_{y2}, Q_{uv2}\}$ and $\{Q_{y3}, Q_{uv3}\}$ in Fig. 3-11. Figs. 3-14 (b), (d) and (f) are the corresponding blood vessel patterns. The results indicate that the more ‘1’s in the quantization tables, the clearer blood vessel patterns are. There are obvious blocking artifacts in Fig. 3-14 (b), which are generated by the quantization tables Q_{y1} and Q_{uv1} . The blocking artifacts are weaker in the results from the quantization tables Q_{y2} and Q_{uv2} , which are shown in Fig. 3-14 (d). The results (Fig. 3-14 (f)) generated by the quantization tables Q_{y3} and Q_{uv3} are very close to the blood vessel patterns (Figs. 3-2 (b)) generated from the original images. The same compression processes in Fig. 3-10 are applied to the images in NTUFID_2. The three pairs of quantization tables generated from the quantization tables of a quality factor of 75 are used in the compression process in Fig. 3-10 (b). The compressed results have the same visual effects which are shown in Figs. 3-15.

Fig. 3-15 Results of images in NTUFID_2 from the compression process in Fig. 3-10 (b). (a), (c) and (e) are generated by the three modified quantization tables. (b), (d) and (f) are respectively blood vessel patterns.

To objectively compare the quality of blood vessels generated by the compression test, blood vessel patterns are extracted from the compressed images and matched with the blood vessel patterns from the original images. As with the previous test, the original images in the one session are used as gallery sets and the other images are used as probe sets. The matching results are given in Figs. 3-16 - 3-19. The blood vessel patterns from the JPEG compressed images are also matched with the original images as references.

Fig. 3-16 The matching results of NTUFID_1 from the compression test. (a)-(d) are CMC curves from LF1, RF1, LT1 and RT1, respectively. In the legends, JPEG50 means the JPEG compressed images with a quality factor of 50; Downsampled means that only the down-sampling operation is applied to the U and V channels but the original Y channel is preserved and YUV i means that the Y, U and V channels are quantized by $\{Q_{yi}, Q_{uvi}\}$, where $i \in \{1, 2, 3\}$.

The results show that the down-sampling operation in the U and V channels does not affect the quality of the blood vessels. The results also show that the first six DCT coefficients in the Y, U and V channels contain enough information for matching. The skin image analysis and the compression test point out that the quantization degrades the blood vessel quality and the first six coefficients in the three channels are critical. To recover the original blood vessel patterns, the first six coefficients in the Y, U and V channels have to be restored.

The CMCs in Fig. 3-16 and Fig. 3-17 also show that the reduction of the matching accuracy of the results generated from the images in NTUFID_2 is lower than that of the results generated from images in NTUFID_1. It should be noted that the quality factor for the NTUFID_1 is even lower. This observation indicates that the compression process can lead to higher quality degradation on the lower resolution images. The same conclusion can be drawn from the comparing of the CMCs in Fig. 3-18 and Fig. 3-19.

Fig. 3-17 The matching results of NTUFID_2 from the compression test. (a)-(d) are CMC curves from LF1, RF1, LT1 and RT1, respectively. In the legends, JPEG75 means the JPEG compressed images with a quality factor of 75; Downsampled means that only the down-sampling operation is applied to the U and V channels but the original Y channel is preserved and YUVi means that the Y, U and V channels are quantized by the three modified quantization tables, respectively.

Fig. 3-18 The matching results of NTUFID_1 from the compression test.
(a)-(d) are CMC curves from LF2, RF2, LT2 and RT2, respectively. In the legends, JPEG50 means the JPEG compressed images with a quality factor of 50; Downsampled means that only the down-sampling operation is applied to the U and V channels but the original Y channel is preserved and YUV*i* means that the Y, U and V channels are quantized by $\{Q_{yi}, Q_{uvi}\}$, where $i \in \{1, 2, 3\}$.

Fig. 3-19 The matching results of NTUFID_2 from the compression test. (a)-(d) are CMC curves from LF2, RF2, LT2 and RT2, respectively. In the legends, JPEG75 means the JPEG compressed images with a quality factor of 75; Downsampled means that only the down-sampling operation is applied to the U and V channels but the original Y channel is preserved and YUV*i* means that the Y, U and V channels are quantized by the three modified quantization tables, respectively.

3.4 Statistics on Skin Images

For algorithm design, a statistical analysis on skin images was performed. Fig. 3-20 shows the probability of non-zero QDCT coefficients estimated from NTUFID_1. The quality factor of 50 was used in this statistical analysis. The y-axes are the probabilities of non-zero QDCT coefficients and the x-axes are their orders, given in Fig. 2-2(a). Most of the QDCT coefficients in the three channels are zero. The first DCT coefficients (the DC coefficients) in all of the channels have the highest probability of non-zero QDCT coefficients, i.e., almost one. Several other QDCT coefficients in the Y channel also have great probabilities of non-zero values. However, for the U and V channels, the probabilities of other non-zero QDCT coefficients are nearly zero.

(a)

(b)

(c)

Fig. 3-20 The probabilities of non-zero QDCT coefficients.
(a)-(c) are the probabilities of non-zero QDCT coefficients in the Y, U and V channels respectively.

Figs. 3-21(a)-(c) show the distributions of the quantized DC (QDC) coefficient values of the Y, U and V channels, respectively. Figs. 3-21(a)-(c) point out that the QDC coefficient values in all of the channels are not uniformly distributed and the probability density in the tails is very

low. Fig. 3-21(d) shows the distributions of the first five quantized AC (QAC) coefficients in the Y channel. As with the QDC coefficients, the QAC coefficients are not uniformly distributed and the probability density in the tails is also very low. Fig. 3-20 shows that the probabilities of non-zero QAC coefficients in the U and V channels are almost zero.

Fig. 3-21 The distributions of quantized DCT coefficients of the Y, U and V channels. (a)-(c) are the QDC coefficients in the Y, U and V channels respectively. (d) is the first five QAC coefficients in the Y channel.

3.5 Summary

In this chapter, a skin image analysis and a compression test are performed to identify the factors that influence the quality of blood vessel information. An analysis on the QDCT coefficients of compressed skin images is conducted. A public skin image database, which is used in the simulations, is introduction. JPEG compressed skin images with different resolutions and different compression ratios are analyzed. The results indicate that the low frequency DCT coefficients in all the channels play a vital role and information in one block and one channel is not enough to restore

the original block. In the later chapters, the findings are used to develop algorithms to recover the blood vessel information from the JPEG compressed images.

Chapter 4 A Global Model for Restoring Blood Vessel Patterns from Compressed Skin Images

Because of the recent advances of imaging technology, the number of evidence images, e.g., child pornographic images, is increasing dramatically [2]. Sometimes, identifying criminals in these images is challenging because only parts of their body skins are shown in evidence images. These legal cases may include child sexual abuse cases, masked gunmen, riots and terrorist attacks.

Most of the traditional biometric traits, including face, fingerprint, DNA and palmprint are not applicable in these cases. It is not guaranteed that every criminal has a unique scar, skin mark or androgenic hair pattern on a particular body site exposed in the images. Thus, skin marks, scars, and androgenic hair patterns are insufficient to address the problem. Blood vessels under skin surface have been used as a biometric trait for many years [21]. Traditionally, they are used only in commercial and governmental applications because infrared images are required to capture high quality blood vessels. Recent research results demonstrate that blood vessels can be extracted directly from color images potentially for forensic applications [25, 26]. However, color images taken by consumer cameras are likely compressed by the JPEG compression method. The blood vessel information under skin is weak and sensitive to the JPEG compression method. As a result, the quality of the color images is seriously degraded, which makes the blood vessels difficult to be visualized.

Fig. 4-1 and Fig. 4-2 show two sets of examples. Images in Fig. 4-1 are from NTUFID_1 and images in Fig. 4-2 are lower resolution images from NTUFID_2. They are compressed with quality factors of 50 and 75, respectively. In both of the two figures, the four rows show the original skin images, the blood vessel patterns from the original, the JPEG compressed images and the corresponding uncovering blood vessel patterns from the compressed images.

From the figures, it can be seen that the resultant blood vessel patterns from [26] are highly influenced by the JPEG method. The blocking artefact effects are obvious and the original blood vessel patterns are destroyed. To use these images for forensic analysis, it is necessary to

design an algorithm to remove the blocking artifacts and restore the original blood vessel patterns before applying the vein uncovering algorithms.

Fig. 4-1 Four examples of blood vessel patterns from NTUFID_1. Each column shows one example. (a) is the original images which are from LF, RF, LT and RT, respectively. (b) shows the corresponding blood vessel patterns from (a) [26]. (c) is the JPEG compressed versions of (a) with a quality factor of 50. (d) shows the blood vessel patterns from (c).

Fig. 4-2 Four examples of blood vessel patterns from NTUFID_2. Each column shows one example. (a) is the original images which are from LF, RF, LT and RT, respectively. (b) shows the corresponding blood vessel patterns from (a) [26]. (c) is the JPEG compressed versions of (a) with a quality factor of 75. (d) shows the blood vessel patterns from (c).

The existing deblocking methods have been reviewed in Chapter 2. Most of them, except of the KB method, are not ideal for recovering skin features, such as skin marks and blood vessels for forensic analysis because they are designed for generic images and attempt to alleviate the

blocking artifacts for visual effect. The original blood vessel information is not guaranteed to be restored. Many of these methods even do not guarantee that the resultant and compressed images have the same QDCT coefficients. Although some priors are imposed in some methods, they are not related to biometric information. In fact, the existing information in the JPEG compressed images is further deteriorated by some deblocking methods. The knowledge based (KB) method is the only one that is specially designed for removing blocking artifacts in skin images [31]. Experimental results demonstrate that this method can effectively restore images with skin marks and outperform other deblocking methods designed for generic images [32, 65, 71, 128]. However, the KB method has not been examined for restoring blood vessel patterns systematically. Our experiments indicate that it is ineffective to recover blood vessel patterns because it operates on the three channels separately and uses very limited information across blocks. In other words, it does not take the dependences between channels and between blocks into consideration. An effective restoration method should recover the original blood vessel information and suppress the blocking artifacts as well.

In this chapter, two learning based algorithms, a preliminary algorithm and a global model, will be presented. Prior knowledge from the skin image database is employed to recover the missing information in the compressed images. The organization of this chapter is as follows. Section 4.1 reports the preliminary algorithm and experiment results. Section 4.2 presents the global model and experimental results. Section 4.3 gives some summary remarks.

4.1 A Preliminary Study on Restoring Blood Vessel Patterns

The compression test in chapter 3 indicates that the low frequency coefficients play a critical role in blood vessel restoration. To restore blood vessels, an algorithm is demanded to recover these critical DCT coefficients. An intuitive approach is to use the QDCT coefficients in JPEG images as features and regression methods to predict the original values of the critical DCT coefficients. The skin image analysis in chapter 3 shows that the U and V channels carry most of the blood vessel information. Based on the analysis, the proposed preliminary algorithm restores the six lowest

frequency coefficients in the U channel and the six lowest frequency coefficients in the V channel. Their corresponding orders are 1-6 given in Fig. 2-2(a).

4.1.1 The Proposed Algorithm

The proposed algorithm includes a training procedure and a restoration procedure. Fig. 4-3 gives a schematic diagram of the proposed algorithm. As shown in the figure, a training database is constructed in the training procedure. Features from the compressed images and targets from the uncompressed images are respectively extracted for model training. In the restoration procedure, the trained models are applied to restore the compressed images.

Fig. 4-3 A schematic of the proposed algorithm

4.1.2 Features and Regression Models

To construct a training database, skin regions from uncompressed skin images are cropped into the same size. Using target compression quality factor, the images are compressed and therefore, each of the uncompressed images has one compressed counterpart in the training database.

After constructing the training database, features and targets are extracted to form the training data. The training data generation process for a block from U or V channel, which includes the target value generation from the uncompressed images and the feature generation from the compressed images, is illustrated in Fig. 4-4.

The lower part of Fig. 4-4 shows the target value generation process. Since the model is trained to restore the original critical DCT coefficients, extracting the original critical DCT coefficients from the uncompressed images as target values is straightforward. The DCT coefficients can be obtained by applying discrete cosine transform to the images in the YUV color space (Fig. 2-1). The obtained DCT coefficients are arranged in the zigzag order (Fig. 2-2(a)), and the first six critical DCT coefficients are extracted from the target block as target values.

Since the quantization operation is a many-to-one mapping, many different uncompressed blocks can be mapped to the same compressed block. Using information in one block and in one channel alone is not enough to provide accurate recovery. Thus, it is necessary to explore cross channel and cross neighbour information. The statistics given in Fig. 3-20 show that a lot of QDCT coefficients are zero and therefore, for each block, six DCT coefficients (i.e., $x_{zk}^j \times q_{zk}$) whose orders are 1-6 (Fig. 2-2(a)) in the Y channel and three DCT coefficients whose orders are 1-3 (Fig. 2-2(a)) in the U and V channels are used to form the feature vectors for restoration (Fig. 4-4).

For recovering the blocks in the U and V channels, 5 by 5 blocks in the U and V channels are used to generate features. The number of neighbor blocks is determined by cross validation. Because of the down-sampling operation, each block in these two channels corresponds to four blocks in the Y channel. Thus, DCT coefficients from 100 blocks in the Y channel are used. Features from each block in the Y, U and V channels are concatenated to form the final feature vector with a size of 750 (Fig. 4-4).

The proposed preliminary algorithm utilizes the treebagger method as a baseline method to train the regression models to restore the critical DCT coefficients. The treebagger method [129] creates an ensemble of bagged decision trees. Each tree produces a prediction and the final result is an average of the predictions. For each target DCT coefficient in the U and V channels, one regression model is trained. The trained regression models are applied to restore the JPEG compressed images to generate the predicted DCT coefficients.

Fig. 4-4 Illustration of training data generation process for U and V channels.

4.1.3 Post-processing

The restored images and the compressed images should have the same QDCT coefficients. Therefore, the recovered DCT coefficients should subject to a quantization constraint which is defined as

$$[\bar{y}_{zk}^j / q_{zk}] = x_{zk}^j, \quad (4-1)$$

where $\bar{y}_{zk}^j$ is the final restored DCT coefficient, and x_{zk}^j and q_{zk} are defined in Eq. 2-1. The predicted DCT coefficients, $\hat{y}_{zk}^j$ are further processed by,

$$\bar{y}_{zk}^j = \begin{cases} x_{zk}^j \times q_{zk} + q_{zk}/2, & \text{if } e_{zk}^j > q_{zk}/2, \\ x_{zk}^j \times q_{zk} - q_{zk}/2, & \text{if } e_{zk}^j < -q_{zk}/2, \\ \hat{y}_{zk}^j, & \text{otherwise,} \end{cases} \quad (4-2)$$

where $e_{zk}^j = \hat{y}_{zk}^j - x_{zk}^j \times q_{zk}$.

Once all DCT coefficients are obtained, the inversed DCT is applied to obtain a restored image.

4.1.4 Experimental Results

Skin images from the two databases, NTUFID_1 and NTUFID_2, which are introduced in Chapter 3, are used in the experiments. To construct the training database, skin areas are cropped with the same size for the skin images in NTUFID_1. Their sizes are 256×192 pixels for the forearm datasets and 320×256 pixels for the thigh datasets. Due to the size limit of the skin images in NTUFID_2, they are directly used for feature generation without cropping. Then, images are compressed to a predefined quality factor, so that there was one compressed counterpart for each original image. Features are equally sampled from the compressed skin images and target values are extracted from the corresponding original images. To test the proposed algorithm with different compression levels, the quality factors 50 and 75 are applied to the images from NTUFID_1 and NTUFID_2, respectively. The proposed algorithm is applied to these compressed images. Images in the left limb dataset are used to generate training data for the right limb dataset and vice versa. The blood vessel uncovering method in [26] was used in these experiments. The KB [31] method, is also applied for comparison. Comparisons of other deblocking methods will be given in Section 4.2 and Chapter 5.

4.1.4.1 Visual Comparisons

Resultant images generated from the compressed images shown in Figs. 4-1 and 4-2 are given in Figs. 4-5 and 4-6, where the 1st - 5th rows are respectively the original images, the blood vessel patterns from the original images, the blood vessel patterns from the compressed images, the blood vessel patterns from the KB [31] method, the blood vessel patterns from the proposed algorithm.

Fig. 4-5 The restoration results from the JPEG compressed images with a quality factor of 50. (a)-(e) are the original images, the blood vessel patterns from the original images, the blood vessel patterns from the compressed images with a quality factor of 50, the blood vessel patterns from the KB [31] method, the blood vessel patterns from the proposed algorithm.

Fig. 4-6 The restoration results from the JPEG compressed images with a quality factor of 75. (a)-(e) are the original images, the blood vessel patterns from the original images, the blood vessel patterns from the compressed images with a quality factor of 75, the blood vessel patterns from the KB [31] method, the blood vessel patterns from the proposed algorithm.

Figs. 4-5 and 4-6 demonstrate that the proposed algorithm not only alleviates the blocking artifacts, but also restores the blood vessel patterns. The KB [31] method removes the blocking artifacts, but it is not effective to restore the blood vessel patterns. Even worse, some blood vessels are clear in the patterns from the compressed images (Figs. 4-5 and 4-6 (c)), but they are destroyed by the KB [31] method (Figs. 4-5 and 4-6 (d)).

4.1.4.2 Objective Evaluation based on the SSIM Index

To measure the effectiveness of the proposed algorithm objectively, evaluations based on the SSIM [123] index are reported in this subsection. In this evaluation, the vascular images from the uncompressed images are regarded as references. The vascular images from the JPEG compressed images and the images restored by the proposed preliminary decision tree based algorithm (DT) and the KB [31] method are regarded as testing images.

Table 4-1 lists the results from NTUFID. The compression factor of 50 is evaluated on the four limb datasets in NTUFID_1 and the compression factor of 75 is evaluated on the the four limb datasets in NTUFID_2. The results show that SSIM indexes from the proposed algorithm, which is referred to as DT, are higher than those from the JPEG compressed images, which means than the image quality is improved after the restoration. It also indicates that the proposed algorithm performs better than the KB method in terms of SSIM index. For some of the datasets, the SSIM indexes from the KB modtod is lower than those from JPEG images without any processing. Since the KB method does not take the dependences across channels and blocks, the left information in the JPEG compressed images are destroyed after restoration.

Table 4-1Mean SSIM indexes from NTUFID										
	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	0.2765	0.2769	0.2608	0.2540	0.2671	0.3340	0.3423	0.2893	0.2908	0.3141
KB	<u>0.2528</u>	<u>0.2486</u>	<u>0.2478</u>	<u>0.2369</u>	<u>0.2465</u>	<u>0.3145</u>	<u>0.3206</u>	0.2942	0.2983	<u>0.3069</u>
DT	0.3118	0.3210	0.2803	0.2768	0.2975	0.3392	0.3777	0.3167	0.3110	0.3362

The best method for each dataset is highlighted. The methods with mean SSIM indexes lower than those of the JPEG images are underlined.

4.1.4.3 Objective Evaluation based on Blood Vessel Matching

In this subsection, an evaluation based on matching accuracy is reported. The three-step blood vessel matching method [26] was utilized to compute cumulative matching characteristic (CMC) curve as a performance index. The matching evaluation is performed on NTUFID_1 and NTUFID_2. Two-fold cross validation is adopted in the matching experiments. In the first set of experiments, the datasets collected in the first session are employed as the probe sets and the datasets collected in the second session are employed as the gallery sets. In the second set of experiments, the probe sets and the gallery sets are swapped. The average accuracy of the two folds is regarded as the final results. The matching results from NTUFID_1 and NTUFID_2 are shown in Figs. 4-7 and 4-8. The matching results from the original images and the JPEG compressed images are given as references. The results generated by the KB method is given for comparison.

Fig. 4-7 The matching results from compressed images with a quality factor of 50. (a)-(d) respectively show the results from LF, RF, LT and RT. The corresponding method for each curve is shown in the legend.

Fig. 4-7 and Fig. 4-8 show that the proposed algorithm gives higher matching accuracy comparing to the results from the JPEG compressed images. The improvement is more obvious for the forearm dataset than for the thigh dataset. The proposed algorithm performs better than the KB method for all the forearm datasets. For the thigh datasets, the proposed algorithm and the KB method give similar performance. The KB method designed for skin images can restore some blood vessels for the thigh datasets, but it performs even worse than the JPEG images without any restoration for the forearm datasets. The matching results indicate that the blood vessel information can be restored using the proposed preliminary algorithm.

Fig. 4-8 The matching results from compressed images with a quality factor of 75. (a)-(d) respectively show the results from LF, RF, LT and RT. The corresponding method for each curve is shown in the legend.

Table 4-2 and Table 4-3 give the matching results in terms of rank-1 and rank-10 identification accuracy from NTUFID_1 and NTUFID_2 respectively. The term “rank-1 (rank-10) identification accuracy” refers to the percentage of input vascular patterns whose corresponding patterns in the database can be found within the top 1 (10) of the patterns given by the algorithm. Table 4-2 shows that the proposed algorithm gives the highest matching accuracy for all the four

datasets, except for the results from the original images. Comparing with the results generated from the JPEG images without restoration, the improvements for the four limb datasets in terms of rank-10 accuracy are 6.08%, 15%, 2.16%, and 5.04%. It indicates the effectiveness of the proposed algorithm. For the results generated from the NTUFID_2, Table 4-3 shows that the proposed algorithm improves the matching accuracy of 2.7%, 4%, 4.57%, 1.09% in terms of rank-10 accuracy for the four limb datasets comparing to the raw JPEG images. For the thigh datasets in NTUFID_2, the improvements from the proposed algorithm are not as good as the KB [31] method. One reason is that only restoring the U and V channels is not enough for recovering the compressed images in the thigh datasets in NTUFID_2. The matching accuracy generated from the KB [31] method is even lower than the results generated from the JPEG compressed images for the two forearm datasets, which indicates that the performance of the KB [31] method is not stable. The proposed algorithm achieves improvements in all the four limb datasets and gives the best performance in the two forearm datasets, except for the results from the original images.

Table 4-2 Matching results from NTUFID_1

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	51.88	69.04	68.60	82.00	46.09	62.39	46.93	61.84
JPEG50	15.06	32.00	21.80	36.40	9.57	23.48	10.53	25.00
KB [31]	11.92	26.15	20.40	36.80	9.13	24.13	12.28	29.82
DT	18.62	38.08	31.60	51.40	12.83	26.09	13.16	30.04

The highest matching accuracy, except for that from the original images, is highlighted.

Table 4-3 Matching results from NTUFID_2

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	46.86	64.44	62.60	80.20	46.74	62.82	49.56	64.25
JPEG75	16.11	29.08	13.80	30.20	3.91	13.04	4.17	16.67
KB [31]	5.44	19.87	8.40	18.60	5.00	19.13	8.55	24.34
DT	18.20	31.78	18.40	34.20	4.57	17.61	7.46	17.76

The highest matching accuracy, except for that from the original images, is highlighted.

4.2 A Global Model

4.2.1 The Proposed Algorithm

The skin image analysis and the compression test in Chapter 3 indicate that all the Y, U and V channels contain blood vessel information, and the first six low frequency DCT coefficients are critical to the blood vessel quality. To recover the blood vessel information for forensic analysis, the critical DCT coefficients in the Y, U and V channels should be restored. The previous preliminary algorithm demonstrates the effectiveness of using the QDCT coefficients in the JPEG compressed images as features to restore the original critical DCT coefficients in the U and V channels. In this subsection, a global model applying the same scheme in the three channels is proposed. Different from the previous preliminary algorithm, which only restores the critical DCT coefficients in the U and V channels, the proposed algorithm restores the blood vessel information in all the three channels. In addition, the gradient boosting method, which approximates the target in the form of a weighted sum of functions, is adopted as the regression method.

The same feature generation process in Fig. 4-4 is applied to generate features and targets for the U and V channels. Because the down-sampling operation is applied to the U and V channels, the feature generation process for the U and V channels is different from that for the Y channel. In the feature generation process for the Y channel, DCT is directly applied to the compressed U and V channels without the down-sampling operation and therefore, each block in the Y channel has one corresponding block in the U and V channels. For each target block in the Y channel, features are extracted from 9 by 9 blocks in the Y, U and V channels and the final feature size is 972.

The proposed global model (GM) employs the gradient boosting algorithm [130] as the baseline regression method φ , and the error function $\varepsilon(\cdot, \cdot)$ is defined as mean squared error. Gradient boosting is an ensemble technique to construct a strong model through combining many weak models and decision trees are the most widely used weak model. Gradient boosting solves

the regression problem in an iterative manner and a model is fit to minimize the gap between the predicted value and the true value in each step.

Given a training set, $\{F, Y\}$, where $F = \{f_1, \dots, f_n\}$, $Y = \{y_1, \dots, y_n\}$, respectively denoting the features and the targets, the mean squared error function $\varepsilon(\cdot, \cdot)$ and the number of iterations T , gradient boosting optimizes φ through the process below.

1. The model $\varphi_0(f_i, \theta_0) = \omega_0, \forall f_i \in F, i = 1, 2 \dots n$, where ω_0 is computed through

$$\omega_0 = \arg \min_{\omega} \sum_{i=1}^n \varepsilon(y_i, \omega) = \arg \min_{\omega} \sum_{i=1}^n \|y_i - \omega\|^2. \quad (4-3)$$

2. For $t = 1$ to T ,

- a. The residuals are computed as, $R_t = \{r_{t,i} | r_{t,i} = y_i - \varphi_{t-1}(f_i, \theta_{t-1}), i = 1, 2 \dots n\}$.
- b. A regression tree $h_t(\cdot, \lambda_t)$, where λ_t is a parameter vector, is learned to fit the residuals R_t using F as features.
- c. The model is updated through,

$$\varphi_t(\cdot, \theta_t) = \varphi_{t-1}(\cdot, \theta_{t-1}) + \gamma h_t(\cdot, \lambda_t), \quad 0 < \gamma < 1, \quad (4-4)$$

where γ is a predefined shrinkage parameter to prevent over fitting.

3. The final model is $\varphi(\cdot, \theta_{\{F, Y\}}) = \varphi_T(\cdot, \theta_T)$.

The models are applied to the JPEG compressed images, and post-processing operation in Eq. 4-1 and Eq. 4-2 are applied to the predicted DCT coefficients to get the final results.

4.2.2 Experimental Results

The two databases described in Chapter 3 are employed to examine the proposed algorithm. The left forearm (thigh) dataset is used as a training dataset when the right forearm (thigh) dataset is used as a testing dataset, and vice versa. The blood vessel uncovering method reported in [26] is used to uncover the blood vessel patterns in these experiments. The training data preparation introduced in Section 4.1.2 is used. In the experiment, the global method is trained using as many

data as possible within the memory constraint. The number of features was no less than 20,000 in each training feature set.

For the proposed algorithm, the three channels are processed separately. The Y channel is deblocked directly. The U and V channels are first down-sampled before applying the proposed algorithm and then the deblocked results are up-sampled and combine with the Y channel results as the final restored images. In addition to the previous preliminary algorithm, which employs the decision tree regression method (referred to as DT), deblocking methods, including SADCT [65], FOE [32], KB [31] and OBTCE [101] are implemented for comparisons.

4.2.2.1 Visual comparisons

Figs. 4-9 and 4-10 give the resultant images from different deblocking methods for visual comparisons. Fig. 4-9 shows the resultant images from NTUFID_1 and Fig. 4-10 shows the resultant images from NTUFID_2. The 1st-9th rows in the two figures show respectively the original images, the blood vessel patterns from the original images, the blood vessel patterns from the compressed images, blood vessel patterns from SADCT [65], FOE [32], KB [31], OBTCE [101], DT and the proposed algorithm.

Figs. 4-9(g) and 4-10(g) show that the OBTCE [101] gives the similar visual effects to the blood vessel patterns from the JPEG compressed images (Figs. 4-9(c) and 4-10(c)). From Figs. 4-9(d-e) and 4-10(d-e), it can be seen that blocking artifacts in the images are smoothed by the SADCT [65] method and the FOE [32] method. However, the blood vessel patterns are blurred at the same time. The KB [31] method (Figs. 4-9(f) and 4-10(f)) also can reduce the blocking effects, but some of the blood vessel patterns existing in the JPEG compressed images (Figs. 4-9(c) and 4-10(c)) are further destroyed. Figs. 4-9(i) and 4-10(i) show that the proposed algorithm can restore blood vessel pattern effectively and outperform the previous deblocking methods. Comparing to the preliminary algorithm (Figs. 4-9(h) and 4-10(h)), the proposed algorithm can restore more details.

Fig. 4-9 Resultant images from NTUFID_1 for comparisons.

(a) the original images, (b) the blood vessel patterns from the original images, (c) the blood vessel patterns from the compressed images, the blood vessel patterns from the restoration results from (d) SADCT [65], (e) FOE [32], (f) KB [31], (g) OBTCE [101], (h) DT and (i) the proposed algorithm.

Fig. 4-10 Resultant images from NTUFID_2 for comparisons.

(a) the original images, (b) the blood vessel patterns from the original images, (c) the blood vessel patterns from the compressed images, the blood vessel patterns from the restoration results from (d) SADCT [65], (e) FoE [32], (f) KB [31], (g) OBTCE [101], (h) DT and (i) the proposed algorithm.

4.2.2.2 Objective evaluations based on SSIM index and PSNR index

In addition to the SSIM [123] index, evaluations based on the PSNR index is also calculated to evaluate the effectiveness of the proposed algorithm objectively. In this evaluation, the vascular images from the uncompressed images are regarded as references. The vascular images from the

JPEG compressed images and the images restored by the deblocking methods are regarded as testing images.

Table 4-5 lists the mean SSIM indexes and the mean PSNR indexes from the restored images generated by the deblocking methods, including SADCT [65], FOE [32], KB [31], OBTCE [101], the preliminary algorithm and the proposed global algorithm. The results from the compressed images with quality factors of 50 and 75 are given as references. The results indicate that the proposed algorithm performs better than the JPEG method in all the testing datasets. The proposed algorithm outperforms all other deblocking methods in terms of the SSIM indexes. For the PSNR indexes, the proposed algorithm performs better than the other methods, expect the KB method for the thigh dataset of NTUFID_2. The previous methods, including FOE [32], SADCT [65], KB [31] and OBTCE [101] perform even worse for some limb datasets than JPEG compressed images without any processing.

Table 4-4 Mean SSIM indexes NTUFID

	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	0.2765	0.2769	0.2608	0.2540	0.2671	0.3340	0.3423	0.2893	0.2908	0.3141
FOE [32]	<u>0.2681</u>	<u>0.2640</u>	0.2675	0.2620	<u>0.2654</u>	<u>0.3313</u>	0.3451	0.2936	0.2950	0.3163
SADCT [65]	<u>0.2602</u>	<u>0.2562</u>	0.2615	0.2561	<u>0.2585</u>	<u>0.2962</u>	<u>0.3045</u>	<u>0.2794</u>	<u>0.2802</u>	<u>0.2901</u>
KB [31]	<u>0.2528</u>	<u>0.2486</u>	<u>0.2478</u>	<u>0.2369</u>	<u>0.2465</u>	<u>0.3145</u>	<u>0.3206</u>	0.2942	0.2983	<u>0.3069</u>
OBTCE [101]	0.2793	0.2772	0.2669	0.2623	0.2714	0.3457	0.3597	0.2991	0.2999	0.3261
DT	0.3118	0.3210	0.2803	0.2768	0.2975	0.3392	0.3777	0.3167	0.3110	0.3362
GM	0.3532	0.3470	0.3318	0.3187	0.3377	0.3898	0.4077	0.3561	0.3567	0.3776

The best method for each dataset is highlighted. The methods with mean SSIM indexes lower than those of the JPEG images are underlined.

Table 4-5 Mean PSNR indexes from NTUFID

	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	14.92	15.33	14.57	14.39	14.82	14.58	14.80	12.89	12.85	13.81
FOE [32]	<u>12.26</u>	<u>12.48</u>	<u>12.94</u>	<u>12.92</u>	<u>12.64</u>	<u>13.13</u>	<u>13.63</u>	<u>12.21</u>	<u>12.16</u>	<u>12.81</u>
SADCT [65]	<u>12.57</u>	<u>12.75</u>	<u>12.70</u>	<u>12.61</u>	<u>12.66</u>	<u>12.69</u>	<u>13.11</u>	<u>11.94</u>	<u>11.88</u>	<u>12.42</u>
KB [31]	19.21	18.92	19.71	19.13	19.24	17.91	17.87	17.11	17.77	17.67
OBTCE [101]	<u>14.49</u>	<u>14.77</u>	<u>14.05</u>	<u>14.04</u>	<u>14.35</u>	14.63	15.16	<u>12.87</u>	<u>12.80</u>	13.90
DT	16.51	17.78	15.45	15.44	16.33	14.76	16.71	14.31	14.10	15.01
GM	20.64	21.03	19.08	18.73	19.90	18.16	19.24	16.89	17.12	17.89

The best method for each dataset is highlighted. The methods with mean PSNR indexes lower than those of the JPEG images are underlined.

4.2.2.3 Objective evaluation based on blood vessel matching

The blood vessel matching results are given in Figs. 4-11 and 4-12. The CMCs show clearly that the proposed algorithm improves the matching accuracy effectively and performs better than the JPEG images without any processing and all the other methods for all the four datasets with different resolutions and different quality factors. The matching accuracy generated by some of the deblocking methods, including SADCT [65], FOE [32] and OBTCE [101] are lower than or similar to the that generated by the JPEG compressed images. The KB [31] method improves the matching accuracy for the thigh datasets, but degrades the matching accuracy for the forearm datasets.

Fig. 4-11 The matching results from compressed images with a quality factor of 50.

(a)-(d) respectively show the results from LF, RF, LT and RT. The corresponding method for each curve is shown in the legend.

Fig. 4-12 The matching results from compressed images with a quality factor of 75. (a)-(d) respectively show the results from LF, RF, LT and RT. The corresponding method for each curve is shown in the legend.

A summary of the matching results in terms of rank-1 identification accuracy and rank-10 identification accuracy are given in Table 4-6 and Table 4-7. Comparing with the JPEG images with a quality factor of 50 in Table 4-6, the proposed algorithm improves the rank-10 identification accuracy at 1654%, 22.6%, 13.17% and 17.11% for the left forearm dataset, the right forearm dataset, the left thigh dataset and the right thigh dataset respectively. Table 4-7 shows that the proposed algorithm provides improvements of 11.92%, 17.40%, 13.26% and 19.73% in terms of the rank-10 identification accuracy for the left forearm dataset, right forearm dataset, left thigh dataset and right thigh dataset comparing with the JPEG compressed images with a quality factor of 75 without any processing. The experimental results show that proposed global model performs better than the other deblocking methods and effectively restores blood vessels.

Table 4-6 Matching results from NTUFID_1

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	51.88	69.04	68.60	82.00	46.09	62.39	46.93	61.84
JPEG50	15.06	32.00	21.80	36.40	9.57	23.48	10.53	25.00
FOE [32]	12.55	29.08	17.20	30.40	6.74	21.30	7.02	21.93
SADCT [65]	11.72	25.73	14.40	28.40	7.17	17.17	5.48	20.39
KB [31]	11.92	26.15	20.40	36.80	9.13	24.13	12.28	29.82
OBTCE [101]	13.60	30.96	18.80	34.80	8.48	22.39	9.87	25.66
DT	18.62	38.08	31.60	51.40	12.83	26.09	13.16	30.04
GM	28.24	48.54	39.40	59.00	20.00	35.65	23.25	42.11

The highest matching accuracy, except for that from the original images, is highlighted.

Table 4-7 Matching results from NTUFID_2

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	46.86	64.44	62.60	80.20	46.74	62.82	49.56	64.25
JPEG75	16.11	29.08	13.80	30.20	3.91	13.04	4.17	16.67
FOE [32]	14.23	28.66	14.80	28.80	3.26	13.04	3.29	14.47
SADCT [65]	5.23	14.64	2.60	10.80	1.96	7.39	1.32	9.21
KB [31]	5.44	19.87	8.40	18.60	5.00	19.13	8.55	24.34
OBTCE [101]	15.69	29.50	13.80	30.00	3.26	11.96	3.07	15.57
DT	18.20	31.78	18.40	34.20	4.57	17.61	7.46	17.76
GM	22.80	41.00	27.40	47.60	11.96	26.30	17.54	36.40

The highest matching accuracy, except for that from the original images, is highlighted.

4.3 Summary

In this chapter, two application-specific restoration algorithms are presented to recover blood vessel patterns in the JPEG compressed images. In Chapter 3, the skin image analysis and the compression test point out that the blood vessel quality is controlled by the critical coefficients. Based on the results, the two proposed algorithms restore blood vessel information by restoring the original critical DCT coefficients block by block in the DCT domain. Both the two proposed algorithms adopt a learning-based scheme and prior knowledge on blood vessel patterns are utilized for restoration. Some existing methods reviewed in Chapter 2 remove the blocking artifacts by predicting the DCT coefficients. However, they are not suitable for restoring the blood vessel information because none of them utilizes the prior knowledge from the skin images.

Since the skin image analysis shows that the U and V channels carry most of the blood vessel information, the first preliminary algorithm restores the blood vessel information in the U

and V channels. According to the skin image statistics, selected DCT coefficients in the compressed images are extracted as features. To capture the dependence between blocks and between channels, information in neighbour blocks and channels are also taken into account. The original critical DCT coefficients from the original images are extracted as targets. Regression models are trained to predict the original critical DCT coefficients using a baseline regression method. Images from NTUFID_1 and NTUFID_2 are used to evaluate the proposed algorithms. Quality factors of 50 and 75 are used to generate the JPEG compressed images. Visual evaluations and matching results show that the proposed preliminary algorithm can restore the original blood vessel information from the JPEG compressed images significantly.

In the second algorithm, a global model is proposed to restore the original critical DCT coefficients in the three channels. Similarly, cross-block and cross-channels information is extracted to restore the first six critical DCT coefficients in the three channels. A gradient boosting method is used to train the regression models. Images from NTUFID_1 and NTUFID_2 are used for algorithm development and evaluation. Two quality factors, i.e., 50 and 75 are used to generate the JPEG compressed images. State-of-the-art methods are implemented for comparisons. Experimental results show that the proposed global model effectively restores blood vessel patterns and outperforms the preliminary algorithm which only restores DCT coefficients in the U and V channels. The proposed global model gives better performance than other deblocking methods including the KB method, which is specially designed for restoring skin images, in terms of both quality metrics and matching accuracy.

Chapter 5 A Multi-model Restoration Algorithm for Recovering Blood Vessels in Skin Images

In the previous chapter, the global method using the QDCT coefficients in the compressed images as features to train regression methods and predict the original critical DCT coefficients is presented. Experimental results show that it is effective to restore the original blood vessel information and outperforms other deblocking methods. However, it does not take any mathematical properties of the JPEG compression process into consideration. To overcome the weaknesses of the previous methods, in this chapter, two mathematical properties in the JPEG compression process are identified and used to design restoration algorithms.

In this chapter, a multi-model restoration algorithm (MMRA), which extracts cross block and cross channel information from large databases and makes use of the mathematical properties of the JPEG compression process, is developed to remove blocking artifacts in JPEG compressed images and restore the lost information. The JPEG compression method illustrated in Fig. 2-1 implies two mathematical properties. From Eq. 2-1, it can be known that the range of the original DCT coefficients can be calculated using the QDCT coefficients and the quantization tables, which can be obtained from the JPEG file header. In addition, the range of the original DCT coefficients for different QDCT coefficients is not overlapping, which means that the DCT coefficients can be clustered according to their QDCT coefficients without error. By using the mathematical properties, a tailor-made clustering scheme is proposed to group training data. A regression model is learned from each grouped dataset to predict original DCT coefficients for a given QDCT coefficient. It can be proven that under a minor assumption, this multi-model regression approach always offers lower training error than the corresponding intuitive regression approach.

Skin images from the NTUFID are used to evaluate the performance of MMRA. In addition, a set of diverse skin images collected from the Internet are utilized to demonstrate the effectiveness of the proposed algorithm. Different resolutions and different compression factors

are examined. The blood vessel uncovering and matching methods introduced in Chapter 2 are applied to visualize the blood vessel patterns and generate the matching results. The SSIM index and PSNR index are used to measure the blood vessel image quality. In addition, CMC curves are also used as performance index. Experimental results show clearly that MMRA restores blood vessels more effectively than the state-of-the-art deblocking methods in terms of both the image quality indexes and matching accuracy.

In this chapter, section 5.1 presents two important properties in the JPEG method, which lead to a multi-model regression approach. Section 5.2 gives a clustering scheme distributing different QDCT coefficients to different regression models. This scheme is designed based on statistics of QDCT coefficients for achieving low generalization error. Section 5.3 presents the algorithm design. Section 5.4 discusses the features, the regression models and a post-processing operation used in the experiments. Experimental results and some discussions are presented in Section 5.5 and Section 5.6 offers some conclusive remarks.

5.1 Two Important Properties and Notations

The quantization process described by Eq. 2-1 implies that given a QDCT coefficient x_{zk}^j and a quantization value q_{zk} , the range of the DCT coefficient y_{zk}^j can be determined. It is reminded that k, j and z are respectively indexes for DCT coefficients, blocks, and channels. y_{zk}^j is a real number between $(x_{zk}^j - 1/2) \times q_{zk}$ and $(x_{zk}^j + 1/2) \times q_{zk}$. It should be highlighted that the quantization tables can be obtained from the JPEG image headers. In other words, x_{zk}^j and q_{zk} are always known. In addition, when $x_{zk}^j \neq x_{zk}^i$, the ranges of their DCT coefficients, y_{zk}^j and y_{zk}^i , are not overlapping. These two properties imply that there is no point in using a regression method, which mixes training data with different values of x_{zk}^j together to learn a regression model for estimating y_{zk}^j . In the previous chapter, the global model in fact employs this scheme. To make use of these two mathematical properties, a multi-model restoration algorithm is proposed.

For purposes of presentation clarity, a set of notations is presented. Let (f_{zk}, y_{zk}) be a training datum for restoring a QDCT coefficient at location $k \in \{1, 2, \dots, 64\}$ in channel $z \in \{Y, U, V\}$, where f_{zk} is a column feature vector and y_{zk} is a target DCT coefficient from uncompressed images. Because all of the target values y_{zk} discussed below are from the same location k and the same channel z , the subscripts k and z are removed to simplify the notations. Let also the entire feature set and target set, respectively, be $F = \{f_1, \dots, f_n\}$ and $Y = \{y_1, \dots, y_n\}$, where n is the total number of training data for restoring a QDCT coefficient at location k in channel z . The feature set F and the target set Y are respectively organized as $\{F_1, \dots, F_M\}$ and $\{Y_1, \dots, Y_M\}$, where all $y_p \in Y_i$ have the same quantized value, i.e., $\lfloor y_p/q \rfloor = C_i$, $f_p \in F_i$ are the corresponding features and M is the number of different C_i in the training set. $C_i < C_j$ if $i < j$. $\varphi_i(\cdot, \theta)$ is a regression model controlled by a parameter vector θ and trained by $\{F_i, Y_i\}$. According to the definition of F_i , it is easily noted that $F_i \cap F_j = \emptyset$, where $\emptyset$ denotes an empty set if $i \neq j$ and the QDCT coefficients of the blocks to be restored, C_i and C_j , are some of their feature elements. Moreover, given a feature vector f and its QDCT coefficient value C_i , f can be assigned to one of F_i without any classification error. These mathematical properties play an important role in designing MMRA.

5.2 A Clustering Scheme

The proposed algorithm uses multiple $\varphi_i(\cdot, \theta)$, which are trained on different datasets, to perform the restoration. Thus it is named multi-model restoration algorithm (MMRA). The original training set $\{F, Y\}$ for restoring a QDCT coefficient at location $k \in \{1, 2, \dots, 64\}$ in channel $z \in \{Y, U, V\}$ is organized as $\{F_L, F_a, \dots, F_{a+m}, F_R\}$ and $\{Y_L, Y_a, \dots, Y_{a+m}, Y_R\}$, where $F_L = F_1 \cup \dots \cup F_{a-1}$ and $F_R = F_{a+m+1} \cup \dots \cup F_M$ are, respectively, training features for the left and right tails of the DCT distributions and $Y_L = Y_1 \cup \dots \cup Y_{a-1}$ and $Y_R = Y_{a+m+1} \cup \dots \cup Y_M$ are their corresponding target DCT coefficient values. Note that $F_i \cap F_j = \emptyset$, $\forall i \neq j$; $F = F_L \cup F_a \cup \dots \cup F_{a+m} \cup F_R$ and $Y = Y_L \cup Y_a \cup \dots \cup Y_{a+m} \cup Y_R$. Fig. 5-1 illustrates the training phase of the proposed algorithm. Each of the baseline regression models φ_i is trained by a dataset (F_i, Y_i) whose target range is the

same. More precisely, $(C_i - 1/2) \times q \leq y_p \leq (C_i + 1/2) \times q$, $\forall y_p \in Y_i$ and $a \leq i \leq a + m$; $(C_1 - 1/2) \times q \leq y_p \leq (C_{a-1} + 1/2) \times q$, $\forall y_p \in Y_L$ and $(C_{a+m+1} - 1/2) \times q \leq y_p \leq (C_M + 1/2) \times q$, $\forall y_p \in Y_R$. Note that given a testing datum with its QDCT coefficient to be restored, these properties can be used to select the corresponding trained model without any error.

Fig. 5-1 Illustration of the training phase of the proposed multi-model restoration algorithm. The clustered training data in each set share the same predication range, e.g., $(C_i - 1/2) \times q \leq y_p \leq (C_i + 1/2) \times q$, $\forall y_p \in Y_i$ and $a \leq i \leq a + m$. Except for (F_L, Y_L) and (F_R, Y_R) , each clustered dataset (F_i, Y_i) is defined as $Y_i = \{y_p | [y_p/q] = C_i\}$, and $F_i = \{f_p | (f_p, y_p), y_p \in Y_i\}$, where $i = a, a + 1, \dots, a + m$, and $p = 1, 2, \dots, n$. $F_L = F_1 \cup \dots \cup F_{a-1}$ and $F_R = F_{a+m+1} \cup \dots \cup F_M$.

Assume that the regression model $\varphi(\cdot, \theta)$ is controlled by a parameter vector θ to predict y_i and it always achieves a global minimum training error for a given training set $\{F, Y\}$ and a predefined error function $\varepsilon(\cdot, \cdot)$, i.e., $\sum_{i=1}^n \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}})) \leq \sum_{i=1}^n \varepsilon(y_i, \varphi(f_i, \theta))$, $\forall \theta$, where $y_i \in Y$, $f_i \in F$, n is the total number of data in F and $\theta_{\{F, Y\}}$ is the optimal parameter vector for the training set $\{F, Y\}$. It can be proven that the proposed MMRA outperforms the proposed global method in Chapter 4.2 in terms of training error defined by ε . The proof is as follows. The training error for $\{F, Y\}$ is $\sum_{i=1}^n \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}}))$ and can be rewritten as

$$\sum_{i=1}^n \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}})) = \sum_{j \in E} \sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}})) \quad (5-1)$$

where E denotes the set $\{L, a, \dots, a + m, R\}$ and φ is regarded as a baseline regression method. Using the assumption and the properties given above,

$$\sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}})) \geq \sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F_j, Y_j\}})), \quad (5-2)$$

where $\theta_{\{F_j, Y_j\}}$ is the optimal parameter vector for the training set $\{F_j, Y_j\}$, can be obtained and therefore,

$$\sum_{i=1}^n \varepsilon(y_i, \varphi(f_i, \theta_{\{F, Y\}})) \geq \sum_{j \in E} \sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F_j, Y_j\}})). \quad (5-3)$$

Eq. 5-3 suggests that we should train φ using the clustered datasets, $\{F_L, Y_L\}, \{F_a, Y_a\}, \dots, \{F_{a+m}, Y_{a+m}\}$ and $\{F_R, Y_R\}$, to obtain $m+3$ parameter vectors, $\theta_{\{F_L, Y_L\}}, \theta_{\{F_a, Y_a\}}, \dots, \theta_{\{F_{a+m}, Y_{a+m}\}}, \theta_{\{F_R, Y_R\}}$, instead of using the entire dataset to train φ . This implies that the multi-model approach, which can be trained in parallel, should be employed to achieve lower training error. In addition to training error, Eq. 5-3 also provides a way to use more data to improve accuracy if $\{F, Y\}$ is too large for training φ because of memory constraints. Mathematically, if the training set is organized as $\{F_1, \dots, F_M\}$ and $\{Y_1, \dots, Y_M\}$, lower training error i.e.,

$$\sum_{j \in E} \sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F_j, Y_j\}})) \geq \sum_{j=1}^M \sum_{y_i \in Y_j, f_i \in F_j} \varepsilon(y_i, \varphi(f_i, \theta_{\{F_j, Y_j\}})), \quad (5-4)$$

can be achieved. The proof is similar. The probabilistic upper bound provided by Vapnik–Chervonenkis dimension (VC dimension) [131] indicates that the generalization error can be well approximated by the training error when the training set is large enough. Training error does not necessarily represent generalization error, especially when the training set is small. To obtain low generalization error, which is well approximated by the training error, the size of each $\{F_i, Y_i\}$ should be taken into consideration. Fig. 3-21 shows that the probability density in the tails of the QDCT distributions is very low. In other words, the corresponding $\{F_i, Y_i\}$ do not have enough data to achieve low generalization error. Thus, the training sets for the tails are grouped as $F_L = \{F_1 \cup \dots \cup F_{a-1}\}$, $F_R = \{F_{a+m+1} \cup \dots \cup F_M\}$, $Y_L = \{Y_1 \cup \dots \cup Y_{a-1}\}$ and $Y_R = \{Y_{a+m+1} \cup \dots \cup Y_M\}$.

The skin image analysis and compression test in chapter 3 indicate that the first six coefficients in the Y, U and V channels play a critical role in blood vessel restoration. Their corresponding orders are 1-6 given in Fig. 2-2(a). Thus, the proposed algorithm restores these

critical DCT coefficients. The grouping scheme is applicable to all of these coefficient values except for non-zero quantized AC coefficient values in the U and V channels. Figs. 3-20 (b) and (c) show that the probabilities of non-zero QAC coefficients in the U and V channels are almost zero. Even if all non-zero quantized AC coefficient values in the U and V channels are included to form F_R and F_L , they are still too small. Our training database is already very large. The baseline regression method employed in the experiment cannot even use up all of the training data because their memory usage reaches the system limit. Thus, it is impossible to use all data directly to train the baseline regression method for restoring the non-zero quantized AC coefficients in the U and V channels. To handle this limit, the feature set F_{ac} sharing the same quantized DC values is used to train the baseline regression method for predicting the non-zero quantized AC coefficients in the U and V channels. More clearly, F_{ac} and the block being restored share the same quantized DC values.

5.3 Algorithm Design

Fig. 5-2 summarizes the proposed algorithm. It includes a training procedure and a testing procedure. In the training procedure, skin images are first extracted from the raw images and resized to remove the compression effects. These images are employed as original (uncompressed) images. After JPEG compression, image pairs, including uncompressed images and their compressed counterparts are obtained to form the training database. Then, feature vectors and target values are extracted from them separately. Before training the regression models, the data is reorganized according to the clustering rules discussed previously. Then, the regression models are trained using the corresponding datasets, e.g., $\{F_j, Y_j\}$. In the testing procedure, the same feature extraction operation is applied to the input testing image to generate the feature vectors. Then, regression models are selected to restore the DCT coefficients according to the QDCT coefficient values of the target block. Note that because of the properties discussed before, there is no error for the model selection.

Fig. 5-2 The schematic diagram of the proposed algorithm.

5.4 Features and Regression Models

The same training data generation process for the global model in the chapter 4 is applied to generate the features and target. The above clustering scheme is applied to group all the training data according to the QDCT coefficients. Gradient boosting is employed to train the regression models. The training process is as follow.

Given a training set, $\{F_j, Y_j\}$, where $F_j = \{f_1, \dots, f_n\}$, $Y_j = \{y_1, \dots, y_n\}$, and $j \in \{L, a, \dots, a + m, R\}$, the mean squared error function $\varepsilon(\cdot, \cdot)$ and the number of iterations T , gradient boosting optimizes φ through the process below.

1. The model $\varphi_0(f_i, \theta_0) = \omega_0, \forall f_i \in F_j, i = 1, 2 \dots n$, where ω_0 is computed through

$$\omega_0 = \arg \min_{\omega} \sum_{i=1}^n \varepsilon(y_i, \omega) = \arg \min_{\omega} \sum_{i=1}^n \|y_i - \omega\|^2. \quad (5-5)$$

2. For $t = 1$ to T ,

- a. The residuals are computed as, $R_{j,t} = \{r_{t,i} | r_{t,i} = y_i - \varphi_{t-1}(f_i, \theta_{t-1}), i = 1, 2 \dots n\}$.

- b. A regression tree $h_t(\cdot, \lambda_t)$, where λ_t is a parameter vector, is learned to fit the residuals $R_{j,t}$ using F_j as features.
- c. The model is updated through,

$$\varphi_t(\cdot, \theta_t) = \varphi_{t-1}(\cdot, \theta_{t-1}) + \gamma h_t(\cdot, \lambda_t), \quad 0 < \gamma < 1, \quad (5-6)$$

where γ is a predefined shrinkage parameter to prevent over fitting.

3. The final model is $\varphi(\cdot, \theta_{\{F_j, Y_j\}}) = \varphi_T(\cdot, \theta_T)$.

The post-processing operation in Eq. 4-1 and Eq. 4-2 is also applied to the predicted DCT coefficients to guarantee that the restored images and the compressed images should have the same QDCT coefficients.

5.5 Experimental Results

Images in NTUFID and a set of images collected from the Internet (NTU Internet Image Set v1 (NTUIIS³)) are used to evaluate the performance of the proposed algorithm. To test images with different compression levels, the compression quality factors of 50 and 75 were used in the experiments. These two quality factors were selected because some skin images collected from the Internet had quality factors lower than 75 and they are commonly employed to evaluate deblocking algorithms for generic images. Images in NTUFID were used to generate training data for evaluating MMRA on both the NTUFID and the Internet image set. More clearly, images in the left limb datasets in NTUFID_1 and NTUFID_2 are respectively used to generate training data for the right limb datasets in NTUFID_1 and NTUFID_2 and vice versa. Skin images from both NTUFID_1 and NTUFID_2 are employed to generate training data for the Internet image set. The method reported in [26] was used to visualize blood vessels in these experiments.

³ The dataset is available at <http://forensics.sce.ntu.edu.sg/>. If the website is down, please email the author at xli16@e.ntu.edu.sg.

The Internet images in NTUIIS were taken by different persons using different cameras and compressed by the JPEG compression method with different quality factors. NTUIIS contains two parts: NTUIIS_1 and NTUIIS_2. Images in NTUIIS_1 were employed for visual comparison and objective evaluation based on quality indexes. To generate original (uncompressed) images for comparisons, they were resized to remove the compression artifacts and referred to as references. To show the restoration effects clearly, images were downsampled to a low resolution so that blood vessels were clear in the original images and JPEG compression artifacts were observable in the compressed images with the compression factors of 50 and 75. There are 44 images and 40 images in NTUIIS_1 for testing JPEG compression factors of 50 and 75, respectively. For each image in NTUIIS_2, there is one corresponding image from the same body part of the same subject. They were collected for objective evaluation based on blood vessel matching. Fig. 5-3 shows sample images in NTUIIS. It shows that the Internet images vary in terms of background, lighting, resolution, pose and body part, etc. In addition, it can be seen that the Internet images are very different from the images in NTUFID.

Fig. 5-3 Raw images with background from the NTUIIS.

5.5.1 Visual comparisons

Figs. 5-4 and 5-5 show the restoration results from some Internet images in NTUFID. Figs. 5-6 and 5-7 show the restoration results from some Internet images in NTUIIS_1. The vascular patterns

from the original images and the JPEG compressed images are shown as references. The state-of-the-art deblocking methods, including SADCT [65], FOE [32], KB [31], OBTCE [101], AKR [84] and CEFAS [79], were used for comparison. The skin prior learning algorithm given by Laurens van der Maaten [132] is used in FOE [32], except explicit indication. The state-of-the-art deblocking methods are selected for comparison because no other research team attempted to restore blood vessels in JPEG images and the most related works are deblocking methods for generic images. Fig. 5-4 and Fig. 5-6 show the results restored from JPEG compressed images with a quality factor of 50 and Fig. 5-5 Fig. 5-6 shows the results from JPEG compressed images with a quality factor of 75.

The results show that the SADCT [65] method and the FOE [32] method smooth the blocking artifacts effectively. However, they blur the blood vessels at the same time and some details are wiped away. The blurry effect is even more serious in the low resolution images. The OBTCE [101] method does not improve the image quality significantly comparing with the blood vessels from the JPEG compressed images. The blocking artifacts are suppressed by the AKR [84] and CEFAS [79] methods, but some details of the blood vessels are missing. The KB [31] method also can reduce the blocking artifacts, but it is not effective to restore the blood vessels. Some blood vessels are clear in the results from the compressed images but they are destroyed by the KB [31] method. The proposed algorithm gives better performance than the existing deblocking methods in terms of visual effect. It reduces the blocking artifacts effectively and recovers the deteriorated blood vessels in the Internet images, which are very different from the training images from NTUFID.

Fig. 5-4 Resultant images from NTUFID_1 for comparisons.

(a) the original images, (b) the blood vessel patterns from the original images, (c) the blood vessel patterns from the compressed images, the blood vessel patterns from the restoration results from (d) SADCT [65], (e) FOE [32], (f) KB [31], (g) OBTCE [101], (h) AKR [84], (i) CEFAS [79] and (j) the proposed algorithm.

Fig. 5-5 Resultant images from NTUFID_2 for comparisons.

(a) the original images, (b) the blood vessel patterns from the original images, (c) the blood vessel patterns from the compressed images, the blood vessel patterns from the restoration results from (d) SADCT [65], (e) FOE [32], (f) KB [31], (g) OBTCE [101], (h) AKR [84], (i) CEFAS [79] and (j) the proposed algorithm.

Fig. 5-6 The restoration results from the JPEG compressed images with a quality factor of 50. (a)-(c) show three examples. The first row of each example shows the original color image, the uncovered vein patterns from the skin regions, the vein patterns from the JPEG compressed images with a quality factor of 50 and the restoration results from FOE [32] and SADCT [65] methods. The second row of each example shows the results from KB [31], AKR [84], OBTCE [101], CEFAS [79] and MMRA. 'JPEG50' means the JPEG compression with a quality factor of 50.

Fig. 5-7 The restoration results from the JPEG compressed images with a quality factor of 75. (a)-(c) show three examples. The first row of each example shows the original color image, the uncovered vein patterns from the skin regions, the vein patterns from the JPEG compressed images with a quality factor of 75 and the restoration results from FOE [32] and SADCT [65] methods. The second row of each example shows the results from KB [31], AKR [84], OBTCE [101], CEFAS [79] and MMRA. 'JPEG75' means the JPEG compression with a quality factor of 75.

5.5.2 Objective evaluations based on the SSIM index and PSNR index

In this subsection, objective evaluations based on the SSIM [123] index and PSNR index are reported. The mean SSIM indexes and the mean PSNR indexes from the compressed Internet images in NTUIIS_1 and the compressed NTUFID images with quality factors of 50 and 75 are given in Tables 5-1 and 5-2. The proposed MMRA performs the best in all the experiments in terms of PSNR index. For NTUFID images, it also outperforms all the other methods in terms of SSIM. For the Internet images, it is the second best in terms of SSIM. Though CEFAS [79] obtains the highest SSIM for the Internet images, it is not stable. It performs even worse than JPEG images in terms of both SSIM and PSNR for NTUFID. For the Internet images with quality factor of 50, the PSNR index from CEFAS [79] is also lower than that from JPEG images without any processing. All the other methods also have a similar situation. In some experiments, they perform even worse than the JPEG images. The improvements from the proposed MMRA are more obvious for NTUFID than that from the Internet images because the Internet images are different from the training data.

Table 5-1 (a) Mean SSIM indexes and (b) Mean PSNR indexes from NTUIIS

(a)								
	JPEG	#FOE [32]	SADCT [65]	KB [31]	AKR [84]	OBTCE [101]	CEFAS [79]	Proposed MMRA
JPEG50	0.4377	0.4490/0.4460	0.4458	<u>0.3271</u>	0.4541	0.4439	0.4621	0.4589
JPEG75	0.5099	0.5133/0.5120	<u>0.4736</u>	<u>0.3899</u>	0.5181	0.5138	0.5195	0.5164

(b)								
	JPEG	#FOE [32]	SADCT [65]	KB [31]	AKR [84]	OBTCE [101]	CEFAS [79]	Proposed MMRA
JPEG50	24.13	<u>21.64/23.76</u>	<u>22.93</u>	<u>22.37</u>	<u>23.60</u>	<u>24.08</u>	<u>23.72</u>	24.21
JPEG75	25.28	<u>23.46/24.83</u>	<u>23.66</u>	<u>24.17</u>	<u>24.57</u>	25.35	25.44	25.81

The best method for each quality factor is highlighted. The methods with the mean SSIM or PSNR lower than those of the JPEG images are underlined.

Two FOE values are given. The first one is obtained from a prior learned through generic images and the other is obtained from a prior learned through NTUFID.

Table 5-2 (a) Mean SSIM indexes and (b) Mean PSNR indexes from NTUFID

(a)

	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	0.2765	0.2769	0.2608	0.2540	0.2671	0.3340	0.3423	0.2893	0.2908	0.3141
FOE [32]	<u>0.2681</u>	<u>0.2640</u>	0.2675	0.2620	<u>0.2654</u>	<u>0.3313</u>	0.3451	0.2936	0.2950	0.3163
SADCT [65]	<u>0.2602</u>	<u>0.2562</u>	0.2615	0.2561	<u>0.2585</u>	<u>0.2962</u>	<u>0.3045</u>	<u>0.2794</u>	<u>0.2802</u>	<u>0.2901</u>
KB [31]	<u>0.2528</u>	<u>0.2486</u>	<u>0.2478</u>	<u>0.2369</u>	<u>0.2465</u>	<u>0.3145</u>	<u>0.3206</u>	0.2942	0.2983	<u>0.3069</u>
OBTCE [101]	0.2793	0.2772	0.2669	0.2623	0.2714	0.3457	0.3597	0.2991	0.2999	0.3261
AKR [84]	0.3235	0.3305	0.3080	0.2964	0.3151	0.3732	0.3981	0.3371	0.3335	0.3618
CEFAS [79]	<u>0.2680</u>	<u>0.2636</u>	0.2629	0.2591	<u>0.2635</u>	<u>0.3184</u>	<u>0.3270</u>	<u>0.2838</u>	<u>0.2856</u>	<u>0.3044</u>
MMRA	0.3575	0.3496	0.3380	0.3260	0.3431	0.3990	0.4118	0.3687	0.3649	0.3868

The best method for each dataset is highlighted. The methods with mean SSIM indexes lower than those of the JPEG images are underlined.

(b)

	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	14.92	15.33	14.57	14.39	14.82	14.58	14.80	12.89	12.85	13.81
FOE [32]	<u>12.26</u>	<u>12.48</u>	<u>12.94</u>	<u>12.92</u>	<u>12.64</u>	<u>13.13</u>	<u>13.63</u>	<u>12.21</u>	<u>12.16</u>	<u>12.81</u>
SADCT [65]	<u>12.57</u>	<u>12.75</u>	<u>12.70</u>	<u>12.61</u>	<u>12.66</u>	<u>12.69</u>	<u>13.11</u>	<u>11.94</u>	<u>11.88</u>	<u>12.42</u>
KB [31]	19.21	18.92	19.71	19.13	19.24	17.91	17.87	17.11	17.77	17.67
OBTCE [101]	<u>14.49</u>	<u>14.77</u>	<u>14.05</u>	<u>14.04</u>	<u>14.35</u>	14.63	15.16	<u>12.87</u>	<u>12.80</u>	13.90
AKR [84]	17.27	18.32	16.66	16.29	17.16	15.41	16.55	14.41	14.44	15.23
CEFAS [79]	12.66	12.99	12.90	12.88	<u>12.86</u>	13.45	13.77	12.32	12.28	<u>12.98</u>
MMRA	21.27	21.23	19.64	19.46	20.43	18.93	19.76	17.82	17.87	18.62

The best method for each dataset is highlighted. The methods with mean PSNR indexes lower than those of the JPEG images are underlined.

Tables 5-3 and 5-4 list the mean SSIM indexes and mean PSNR indexes from the Internet images and NTUFID images. They are measured in the RGB domain. Tables 5-3 and 5-4 indicate that the compressed images and original images in the RGB domain are very similar in terms of quality metrics. All their mean SSIM indexes and PSNR indexes are over 0.98 and 38, respectively. The similarity in terms of quality metrics is consistent with human observation. AKR [84] performs slightly better than the other methods for the Internet images. For NTUFID, the proposed MMRA performs the best in terms of SSIM. In terms of PSNR, SADCT [65] performs the best. It should be emphasized that blood vessels are hardly observed in the RGB domain and the SSIM index and PSNR index in the RGB domain do not reflect the quality of

blood vessels. Comparing with generic images, skin texture images are not sharp and more homogeneous, so the degradation from JPEG compression is less observable to human vision and less detectable for the quality indexes designed to mimic human vision in the RGB domain.

Table 5-3 (a) Mean SSIM indexes and (b) mean PSNR indexes from the Interent images in the RGB domain

(a)								
	JPEG	#FOE [32]	SADCT [65]	KB [31]	OBTCE [101]	AKR [84]	CEFAS [79]	Proposed MMRA
JPEG50	0.9928	0.9921/0.9937	0.9932	0.9821	0.9933	0.9939	0.9933	0.9933
JPEG75	0.9951	0.9948/0.9955	0.9950	0.9879	0.9952	0.9957	0.9952	0.9952

(b)								
	JPEG	#FOE [32]	SADCT [65]	KB [31]	OBTCE [101]	AKR [84]	CEFAS [79]	Proposed MMRA
JPEG50	39.22	39.03/39.81	39.45	35.45	39.25	40.00	39.37	39.57
JPEG75	40.97	40.78/41.36	40.94	36.59	40.78	41.62	41.04	41.10

Two FOE values are given. The first one is obtained from a prior learned through generic images and the other is obtained from a prior learned through NTU skin images.

Table 5-4 (a) Mean SSIM indexes and (b) mean PSNR indexes from NTUFID in the RGB domain

(a)										
	JPEG50 for NTUFID_1					JPEG75 for NTUFID_2				
	LF	RF	LT	RT	Mean	LF	RF	LT	RT	Mean
JPEG	0.9844	0.9843	0.9856	0.9853	0.9847	0.9860	0.9853	0.9862	0.9860	0.9859
FOE [32]	<u>0.9722</u>	<u>0.9704</u>	<u>0.9768</u>	<u>0.9761</u>	<u>0.9738</u>	<u>0.9795</u>	<u>0.9766</u>	<u>0.9791</u>	<u>0.9786</u>	<u>0.9784</u>
SADCT [65]	0.9895	0.9885	0.9902	0.9900	0.9895	0.9904	0.9900	0.9915	0.9914	0.9908
KB [31]	<u>0.9606</u>	<u>0.9587</u>	<u>0.9679</u>	<u>0.9665</u>	<u>0.9633</u>	<u>0.9539</u>	<u>0.9496</u>	<u>0.9560</u>	<u>0.9530</u>	<u>0.9531</u>
OBTCE [101]	0.9871	0.9862	0.9879	0.9876	0.9872	0.9885	0.9879	0.9891	0.9890	0.9886
AKR [84]	0.9863	0.9851	0.9871	0.9869	0.9863	0.9897	0.9892	0.9899	0.9898	0.9896
CEFAS [79]	0.9844	0.9835	0.9862	0.9859	0.9850	0.9862	0.9854	0.9869	0.9867	0.9863
MMRA	0.9917	0.9906	0.9917	0.9914	0.9913	0.9917	0.9912	0.9920	0.9917	0.9916

The best method for each dataset is highlighted. The methods with mean PSNR indexes lower than those of the JPEG images are underlined.

(b)

	JPEG50 for NTUFID 1					JPEG75 for NTUFID 2				
	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>	<i>LF</i>	<i>RF</i>	<i>LT</i>	<i>RT</i>	<i>Mean</i>
JPEG	38.31	37.90	38.20	38.10	38.12	39.68	39.52	40.17	40.07	39.85
FOE [32]	<u>34.06</u>	<u>33.33</u>	<u>34.96</u>	<u>34.79</u>	<u>34.26</u>	<u>36.52</u>	<u>36.01</u>	<u>36.90</u>	<u>36.75</u>	<u>36.51</u>
SADCT [65]	39.36	38.91	39.31	39.24	39.20	40.61	40.48	41.27	41.18	40.88
KB [31]	<u>31.86</u>	<u>31.42</u>	<u>32.72</u>	<u>32.43</u>	<u>32.09</u>	<u>31.11</u>	<u>30.77</u>	<u>31.50</u>	<u>31.13</u>	<u>31.12</u>
OBTCE [101]	38.58	38.18	38.46	38.35	38.39	39.81	39.67	40.41	40.32	40.04
AKR [84]	38.78	38.25	38.80	38.68	38.62	40.19	39.99	40.82	40.72	40.41
CEFAS [79]	38.77	38.32	38.63	38.53	38.56	39.95	39.78	40.45	40.35	40.12
MMRA	38.80	38.32	38.66	38.50	38.57	39.91	39.76	40.41	40.29	40.09

The best method for each dataset is highlighted. The methods with mean PSNR indexes lower than those of the JPEG images are underlined.

Images in Fig. 5-8 have been highly compressed but it is hard for human vision to observe their JPEG artifacts. In the JPEG compression process, the U and V channels are compressed more seriously than the Y channel because human vision is less sensitive to the U and V channels, which leads to the more serious degradation in the U and V channels. The analysis of blood vessel quality in Chapter 3 shows that most of the vascular information exists in the U and V channels, while the Y channel only contains little vascular information. The degradation in the blood vessel images is more noticeable than that in the RGB domain because they rely on information in the U and V channels. To show clearly the compression artifacts in the RGB domain and different channels, a skin image with clear blood vessels in the U and V channels is selected.

Fig. 5-8 shows the visual effect of the degradation caused by JPEG compression in the RGB domain, the Y, U and V channels, and the vascular image. The original color image, the Y, U and V channels and the blood vessels from the original image are shown as references in the first row. The second row shows the corresponding compressed image with a quality factor of 50, the Y, U and V channels and the blood vessels from the compressed image. The figure shows that some blood vessels are visible in the U and V channels. It also can be seen that the degradation in the Y channel is not observable, while the blocking artifacts in the U and V channels and the vascular image are obvious.

Fig. 5-8 Illustration of the degradation effects.

Illustration of the degradation effects caused by JPEG compression method in the RGB domain, the Y, U, V channels and the blood vessels. The first column shows the original color image, the Y, U and V channels and the blood vessels from the original image. The second column shows the corresponding compressed color image with a quality factor of 50, the Y, U and V channels and the blood vessels from the compressed image.

5.5.3 Objective evaluation based on blood vessel matching

In this subsection, an evaluation based on matching accuracy is reported. The three-step blood vessel matching method [26] was utilized to compute cumulative matching characteristic (CMC) curve as a performance index. The matching evaluation was performed on NTUFID_1 and NTUFID_2. Two-fold cross validation was adopted in the matching experiments. In the first set of experiments, the datasets collected in the first session were employed as the probe sets and the datasets collected in the second session were employed as the gallery sets. In the second set of experiments, the probe sets and the gallery sets were swapped. The average accuracy of the two folds is regarded as the final results. The matching results from NTUFID_1 and NTUFID_2 are shown in Figs. 5-9 and 5-10. The matching results from the original images and the JPEG compressed images are given as references. Besides the proposed MMRA, the KB [31], FOE [32], SADCT [65], OBTCE [101], CEFAS [79], AKR [84], GM and DT methods are given for comparisons. GM stands for a global method without using the clustering scheme proposed in this Chapter and DT refers to the preliminary approach using decision tree.

Fig. 5-9 shows that MMRA provides the best performance comparing other deblocking methods and the results from the JPEG compressed images for all the four limb datasets. It pinpoints that in most of the cases, the results given by the OBTCE [101], FOE [32], CEFAS [79] and SADCT [65] methods are lower than the JPEG images without any processing because they were in fact designed for generic images. The KB [31] method designed for skin images can restore some blood vessels for all the datasets, except for the left forearm dataset. Fig. 5-10 shows that the proposed algorithm significantly improves the matching accuracy comparing the results from the JPEG compressed images. It also outperforms the other deblocking methods. The matching results show that MMRA is effective for restoring the degraded vascular information from JPEG compressed images.

Fig. 5-9 The matching results from compressed images with a quality factor of 50. (a)-(d) respectively show the results from LF, RF, LT and RT.

Fig. 5-10 The matching results from compressed images with a quality factor of 75. (a)-(d) respectively show the results from LF, RF, LT and RT.

Table 5-5 and Table 5-6 give the matching results in terms of rank-1 and rank-10 identification accuracy from NTUFID_1 and NTUFID_2 respectively. The term “rank-1 (rank-10) identification accuracy” refers to the percentage of input vascular patterns whose corresponding patterns in the database can be found within the top 1 (10) of the patterns given by the algorithm. Table 4 shows that in terms of the rank-10 identification accuracy, MMRA offers improvements of 17.79%, 24.20%, 15.65% and 19.74% for respectively the left forearm dataset, the right forearm dataset, the left thigh dataset and the right thigh dataset comparing with the JPEG compressed images with a quality factor of 50. Table 5-5 also shows that MMRA always outperforms the corresponding global regression method using the same baseline regression method. It indicates the effectiveness of the clustering scheme. Comparing the JPEG images without any processing in Table 5-6, MMRA offers, respectively, improvements of 13.39%, 19.20%, 15.66% and 20.61% in terms of the rank-10 identification accuracy for the left forearm dataset, right forearm dataset, left thigh dataset and right thigh dataset. The experimental results show that MMRA performs better than the other deblocking methods and effectively restores blood vessels. It offers the highest rank-

10 matching accuracy in all the evaluations. In terms of rank-1 matching accuracy, GM gives the similar results. Note that the clustering scheme guarantees a lower error in terms of l_2 , which is not equivalent to the matching accuracy.

Table 5-5 Matching results from NTUFID_1

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	51.88	69.04	68.60	82.00	46.09	62.39	46.93	61.84
JPEG50	15.06	32.00	21.80	36.40	9.57	23.48	10.53	25.00
FOE [32]	12.55	29.08	17.20	30.40	6.74	21.30	7.02	21.93
SADCT [65]	11.72	25.73	14.40	28.40	7.17	17.17	5.48	20.39
KB [31]	11.92	26.15	20.40	36.80	9.13	24.13	12.28	29.82
AKR [84]	20.08	41.84	31.60	50.80	16.30	34.78	18.64	35.75
OBTCE [101]	13.60	30.96	18.80	34.80	8.48	22.39	9.87	25.66
CEFAS [79]	8.79	21.55	13.40	25.00	6.96	18.26	6.58	19.74
DT	18.62	38.08	31.60	51.40	12.83	26.09	13.16	30.04
GM	28.24	48.54	39.40	59.00	20.00	35.65	23.25	42.11
Proposed MMRA	28.87	49.79	41.00	60.60	21.30	39.13	25.44	44.74

The highest matching accuracy, except for that from the original images, is highlighted.

Table 5-6 Matching results from NTUFID_2

	LF		RF		LT		RT	
	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>	<i>Rank-1</i>	<i>Rank-10</i>
Original	46.86	64.44	62.60	80.20	46.74	62.82	49.56	64.25
JPEG75	16.11	29.08	13.80	30.20	3.91	13.04	4.17	16.67
FOE [32]	14.23	28.66	14.80	28.80	3.26	13.04	3.29	14.47
SADCT [65]	5.23	14.64	2.60	10.80	1.96	7.39	1.32	9.21
KB [31]	5.44	19.87	8.40	18.60	5.00	19.13	8.55	24.34
AKR [84]	19.46	36.82	20.20	38.60	6.52	21.74	12.94	26.75
OBTCE [101]	15.69	29.50	13.80	30.00	3.26	11.96	3.07	15.57
CEFAS [79]	13.60	25.31	10.40	25.00	2.39	10.00	2.19	12.06
DT	18.20	31.78	18.40	34.20	4.57	17.61	7.46	17.76
GM	22.80	41.00	27.40	47.60	11.96	26.30	17.54	36.40
Proposed MMRA	22.59	42.47	27.00	49.40	11.74	28.70	17.98	37.28

The highest matching accuracy, except for that from the original images, is highlighted.

To match images collected from the Internet, NTUIS_2, which contains image pairs from the same parts of the same subjects, was constructed. The forearms are extracted and formed two datasets: a left forearm dataset and a right forearm dataset. Their sizes are 156 and 162 images from 78 and 81 subjects respectively. Their quality factors range from 50 to 90. For each subject, two different images were collected so that image pairs can be constructed for matching. Note that two images from the same subject may have different compression quality factors and these images may include both left and right forearms. In the experiments, all the methods used for comparison and MMRA were directly applied to the Internet images without any preprocessing. More clearly, the methods are applied to remove the block artifacts which are not generated by any of our preprocessing but are originally in them. After applying the methods, the forearm skin regions were

extracted and further processed for matching. For matching the left (right) forearm images, one image from each left (right) forearm is put into a probe set and the rest of the left (right) forearm images are mixed up with the first session images in NTUFID_2 to form a gallery set. The sizes of the left forearm probe and gallery sets are 78 and 317 images and the sizes of the right forearm probe and gallery sets are 81 and 331 images. Fig. 5-11(a) shows that the proposed MMRA outperforms other deblocking methods in all the ranks and most of them perform even worse than the JPEG images without any processing when matching the left forearm images. MMRA offers 5% improvement comparing to JPEG images in terms of rank-10 matching accuracy. Fig. 5-11(b) shows the proposed MMRA, FOE [32], AKR [84], OBTCE [101], and CEFAS [79] perform similarly in the first seven ranks, but the proposed MMRA outperforms the others in the rest of ranks. These experimental results indicate that MMRA can restore JPEG images taken in uncontrolled environment and improve their matching performance.

⁴Fig. 5-11 Matching results from NTUIIS_2.

(a) is the matching result from the left forearm set and (b) is the matching result from the right forearm set.

5.5.4 Discussions

Some may believe that the quantization error is uniformly distributed and therefore, estimating the original DCT coefficients is impossible. If only one quantized DCT was used as an input feature and the quantization error was uniformly distributed, it would be no way to estimate the original DCT coefficients. The proposed algorithm can restore skin images because the input features are formed by the quantized DCT coefficients from one block and its neighbouring blocks and

⁴ The matching accuracy in Fig.5-11 is higher than the previous experiments because the images have higher quality factors ranging from 50 to 90.

channels. Though the compression is performed block-by-block and channel-by-channel, the information in compressed images is across channels and across blocks. More clearly, blood vessels pass through many blocks and their information can be found in Y, U and V channels. To demonstrate that MMRA can reduce the QDCT coefficient error, Figs. 5-12 and 5-13 give the error distributions estimated from NTUFID before and after processing. They show that JPEG quantization errors in skin images are not uniformly distributed and MMRA can effectively reduce the errors.

Fig. 5-12 Quantized DC coefficient error distributions from JPEG compression with a quality factor of 50. (a)-(c) The error distributions from respectively the Y, U and V channels.

Fig. 5-13 Quantized DC coefficient error distributions from JPEG compression with a quality factor of 75. (a)-(c) The error distributions from respectively the Y, U and V channels.

The proposed algorithm has two potential applications. The first one is to obtain clearer blood vessels for manual comparison as supporting evidence in courts. Note that soft biometrics, e.g., height and skin color, is always used as supporting evidence even though they are not unique. Also note that fingerprint experts still compare latent prints manually. If clearer blood vessels are restored, they can provide high image quality for manual comparison as supporting evidence. In fact, clear blood vessel patterns have more information than many soft biometrics, e.g., skin color. The experimental results have shown that MMRA can improve visual quality. The second application is to shorten suspect lists. The experimental results show that MMRA can improve matching performance. It implies that it has potential for shortening suspect lists. Using hidden

vascular information in JPEG images for criminal and victim identification is a very new research direction, and further research is necessary.

5.6 Summary

In this chapter, a multi-model restoration algorithm (MMRA) is presented for restoring JPEG skin images to visualize clear blood vessels. Two mathematical properties in the JPEG compression method and the statistics of skin images are used to derive MMRA. It has been proven that under a minor assumption, the data clustering scheme used in MMRA always offers lower training error than the corresponding global regression method. To extend this mathematical result for achieving lower generalization error, the size of training sets is taken into consideration and the training sets are re-organized. MMRA is evaluated on a public database with 1894 images from 947 limbs and also images collected from the Internet. To examine its effectiveness, images with different scales and different compression factors are tested. Matching accuracy, PSNR index and SSIM index are used as performance indexes. The experimental results show clearly that MMRA outperforms the existing deblocking methods for different scales and different compression levels and effectively restores blood vessel information.

Chapter 6 Tattoo Image Restoration based on Deep Convolutional Neural Network

Because of the lossy operations, the JPEG compression method produces blocking artifacts and leads to information loss in the compressed images. As a result, information in tattoo images is degraded. Fig. 6-1 gives four pairs of images to illustrate the degradation problem. Fig. 6-1(a) shows the original images and Figs. 6-1(b) – (d) show the corresponding JPEG compressed image with quality factors of 50, 40 and 30, respectively.

Fig. 6-1 Illustration of degradation effects caused by the JPEG compressed images. (a) shows four uncompressed tattoo images and (b) - (d) gives the corresponding JPEG compressed images with quality factors of 50, 40 and 30, respectively.

From the figure, it can be seen that the blocking artifacts in the JPEG compressed images (Fig. 6-1(b)) are obvious. In addition, some words and tattoo patterns, which are clear in the uncompressed images (Fig. 6-1(a)), are degraded in the JPEG compressed images (Fig. 6-1(b)). The figure also indicates that the degradation becomes more serious when the compression ratio is higher. To use tattoos as a biometric trait for personal identification, the blocking artifacts have to be removed and the lost information has to be recovered.

Current deblocking methods are reviewed in Chapter 2. Most of them are developed for generic images. Though the KB [31] method is designed to restore JPEG skin images, it is not effective for tattoo images because it restores skin images only based on local dependence.

The proposed deblocking algorithms in Chapter 4 and Chapter 5 are designed based on the results of the skin image analysis and compression tests in Chapter 3, which are conducted specially for blood vessel information restoration. These algorithms only recover the DCT coefficients which are critical to the blood vessel quality. Thus, they are not suitable for restoring tattoo images. In addition, the clustering scheme used in MMRA is not applicable to tattoo image restoration because the tattoo images are more diverse than skin images.

In this chapter, a restoration algorithm based on deep convolutional neural network is proposed to recover the lost tattoo information in the JPEG compressed images. The JPEG compression process reviewed in Chapter 2 implies that the QDCT coefficients from the uncompressed images should be the same as those from the JPEG compressed images. To fulfil this requirement, the proposed algorithm recovers the original DCT coefficients in the frequency domain. This is achieved by directly estimating the residual of the DCT coefficients, i.e. the difference between the DCT coefficients from the original images and that from the JPEG compressed versions, in a block based manner. The block based operations may lead to blocking artifacts along the boundaries. To overcome this weakness, an overlapping block restoration scheme is also proposed. Explicit constraint is imposed to guarantee that the resultant images from the proposed algorithm and the corresponding JPEG compressed images have the same

QDCT coefficients. A tattoo database with images collected from the Internet is constructed for the algorithm development and evaluation.

Some image restoration methods based on deep neural networks, which can be applied to JPEG compressed images, have been developed [110-112]. They are different from the proposed algorithm. First, they are mainly developed for different purposes, e.g. image denoising and image deblurring, and are applied to image deblocking problem. The proposed algorithm is specially developed for restoring JPEG images and the proposed deep architecture is designed based on the properties of the DCT coefficients. Second, some of these methods operate in the spatial domain and therefore they cannot guarantee that the recovered images have the same QDCT coefficients as the JPEG compressed images. The proposed algorithm employs a constraint layer to make sure that the recovered DCT coefficients are in the target range (Eq. 2-3).

In this chapter, Section 6.1 introduces the network architecture of the proposed algorithm. Section 6.2 presents an overlapping restoration scheme. Section 6.3 gives a constraint layer specially designed for DCT coefficient restoration. Section 6.4 reports a scheme to estimate the weight for each DCT coefficient in the objective function. Section 6.5 provides experimental results and discussion. Section 6.6 offers some concluding remarks.

6.1 The Network Architecture

In this Section, an algorithm utilizing deep neural net framework is proposed to remove the blocking artifacts and restore the lost tattoo information in the JPEG compressed tattoo images.

The proposed deep framework takes image patches in the RGB domain as input and the residual, i.e., the difference between the DCT coefficients from the original images and that from the JPEG compressed versions, as output. In other words, the proposed model implicitly transforms the images from the spatial domain into the frequency domain through the operations in the hidden layers. The residual learning strategy based on deep framework is first proposed in [133], which is utilized to solve the training performance degradation problem in very deep architectures. This proposal is based on the assumption that the residual mapping is easier to be

learned than the original mapping. The usage of the residual learning strategy can significantly boost the performance in the DCT coefficient estimation because the original DCT coefficient values distribute in a large range, but their residuals are in a fix range (Eq. 2-3). The batch normalization [134] technique is utilized to stabilize the training process and enhance the performance of the deep model. An adaptive constraint layer, which is explicitly carried out by a piecewise function, is embedded before the loss layer to make sure that the predicted residuals for the JPEG blocks are in the target range.

According to the JPEG compression process introduced in Chapter 2, images in the Y, U and V channels are divided into 8 by 8 blocks and the DCT coefficients are obtained by applying the discrete cosine transform. The quantization operation shown in Eq. 2-1 is further utilized to generate the QDCT coefficients. As shown in Fig 2-2 (a), the QDCT coefficients are arranged in a zigzag order according to their frequency. Each DCT coefficient carries specific frequency information. Figs 2-2 (b) and (c) indicate that the quantization values are higher for these DCT coefficients with higher frequency, which leads to higher residual ranges in terms of DCT coefficient values. Intuitively, restoring DCT coefficients with different frequency should rely on different information extracted from training data. Thus, it is improper to train one estimator for all DCT coefficients. To deal with this problem, the DCT coefficients are divided into three frequency groups, i.e. the low frequency group, the middle frequency group and the high frequency group. According to the zigzag order shown in Fig 2-2 (a), the first ten DCT coefficients are grouped into the low frequency group; the middle frequency group consists of the 11th- 30th DCT coefficients, and the last 34 DCT coefficients are grouped into the high frequency group. To restore the DCT coefficients in these three groups, a tree-structured deep neural network is designed. Three branches are used to predict the DCT residuals of the three groups in a parallel manner. In the first layer of the network, the three branches share the parameters. After that the three branches are learned independently.

Fig. 6-2 The illustration of the proposed network architecture.

Fig. 6-2 illustrates the proposed deep neural network architecture with depth $D+1$. As shown in the figure, three types of layers, which are highlighted with different colors, are used. The input data is fed into the first layer, which consists of a set of convolutional filters (Conv) for feature map generation, batch normalization operation (Batch Norm) for internal covariance reduction and Rectified Linear Units (ReLU) for nonlinearity mapping. From the second layer to the D^{th} layer, the network is divided into three branches and the three branches have the same structures. The structures of the layer 2 to the layer $D-2$ in each branch are same as those of the first layer. A convolutional layer is used in layer $D-1$. Before the loss calculation, a specially designed constraint layer is utilized. The outputs of the three branches are the low, middle and high frequency DCT coefficient residuals and form the final predictions. Because DCT coefficients with different frequency have different importance to the images quality, a weighted l_2 loss function is utilized as an objective function. More details about the parameter setting of the network are introduced in section 6.4.1. The proposed tree-structured deep network is specially for estimating the DCT coefficient with different frequencies. The three branches, i.e. including layers from layer 2 to the layer D , are trained independently so that specific information for different frequency can be learned. The parameter sharing in the first layer helps to maintain the relationship among all the DCT coefficients.

6.2 An Overlapping Block Restoration Scheme

In the JPEG compression process, the images are divided into 8 by 8 blocks without overlapping and the DCT coefficients are obtained by applying the block based discrete cosine transform to these blocks. In this chapter, these blocks are referred to as JPEG blocks. Blocks across the boundaries of the JPEG blocks with 8 by 8 pixels are referred to as non-JPEG blocks. Let $\mathbf{I}$ denote an given image with size of H by W and $\mathbf{I}(i, j)$ denote an image pixel located at (i, j) , where i and j are the vertical and horizontal coordinates, respectively. All the blocks in image $\mathbf{I}$ are defined as,

$$\mathbf{B} = \{\mathbf{I}(m:m+7, n:n+7) | 1 \leq m \leq H-7, 1 \leq n \leq W-7\}, \quad (6-1)$$

where $\mathbf{B}$ is the set of blocks in image $\mathbf{I}$, and $\mathbf{I}(m:m+7, n:n+7)$ represents a block in image $\mathbf{I}$ with its upper left pixel at (m, n) and lower right pixel at $(m+7, n+7)$. The sets of the JPEG blocks and the non-JPEG blocks in image $\mathbf{I}$, $\mathbf{B}_{JPEG}$ and $\mathbf{B}_{NJPEG}$, are defined as,

$$\begin{aligned} \mathbf{B}_{JPEG} &= \{\mathbf{I}(m:m+7, n:n+7) | 1 \leq m \leq H-7, 1 \leq n \leq W-7, \text{mod}(m, 8) \\ &= 1, \text{mod}(n, 8) = 1\}, \end{aligned} \quad (6-2)$$

$$\mathbf{B}_{NJPEG} = \mathbf{B} \setminus \mathbf{B}_{JPEG}. \quad (6-3)$$

The above definitions imply that there are 8 by 8 division methods to divide the images into blocks without overlapping. Among them, only one division method generates the JPEG blocks, while the others generate non-JPEG blocks. Fig. 6-3 gives examples of the two types of blocks. As shown in the figure, blocks with red boundaries are JPEG blocks. The figure also shows two sets of non-JPEG blocks generated with two different divisions, which are highlighted with green and blue colors. The figure also indicates that the blocks are overlapped.

Fig. 6-3 Illustration of the two types of blocks.

Blocks with red color boundaries are the JPEG blocks. Blocks with green and blue color boundaries are the non-JPEG blocks.

If only the JPEG blocks without overlapping are recovered, the blocking artifacts may exist along the block boundaries. To alleviate the blocking effects, an overlapping block restoration scheme is used. In the training phase, the JPEG block and the non-JPEG blocks are fed into the network for model training. In the testing phase, DCT coefficient residuals from both types of blocks in the compressed images are estimated. The predicted residuals are added to the DCT coefficients in the corresponding blocks. Inverted discrete cosine transform is applied to generate images in the spatial domain. Inspired by the back-projection method for image super-resolution [135, 136], the recovered images from the non-JPEG blocks are projected to the recovered images from the JPEG blocks to ensure that the DCT coefficients of the recovered images are in the proper range. The back projection process is given in Algorithm 6-1.

Algorithm 6-1 The image projection process for generating the final resultant images.

Input: Images recovered from the JPEG blocks: I_{JPEG} ; the number of images recovered from the non-JPEG blocks: T ; image set recovered from the non-JPEG blocks: $I_{NJPEG} = \{I_{NJPEG,1}, I_{NJPEG,2}, \dots, I_{NJPEG,T}\}$;

Output: The final result: I_{Re} ;

Initialization: $I_{Re,0} = I_{JPEG}$;

-
1. For $t=1$ to T
 2. Calculate the difference between the current recovered image $\mathbf{I}_{Re,t-1}$ and the t^{th} image recovered from the non-JPEG blocks $\mathbf{I}_{NJPEG,t}$, $\mathbf{E}_t = \mathbf{I}_{NJPEG,t} - \mathbf{I}_{Re,t-1}$;
 3. Apply Gaussian smoothing filter with kernel size 3 by 3 to generate $\hat{\mathbf{E}}_t$, i.e., $\hat{\mathbf{E}}_t = G(\mathbf{E}_t)$;
 4. Calculate $\hat{\mathbf{I}}_{Re,t}$ through $\hat{\mathbf{I}}_{Re,t} = \mathbf{I}_{Re,t-1} + \hat{\mathbf{E}}_t$;
 5. Obtain $\mathbf{I}_{Re,t}$ by applying the post-processing introduced in Chapter 4 to make sure the $\mathbf{I}_{Re,t}$ has the same QDCT coefficients as $\mathbf{I}_{JPEG}$;
 6. End
 7. The final resultant image is obtained by $\mathbf{I}_{Re} = \mathbf{I}_{Re,T}$.
-

6.3 An Adaptive Constraint Layer

According to the JPEG compression method, the QDCT coefficients from the Y, U and V channels and the quantization tables can be directly obtained from the JPEG image headers. As described in Eq. 2-1, the quantization process in the JPEG compression method implies that given a QDCT coefficient x_{zk}^j and a quantization value q_{zk} , the range of the DCT coefficient residual e_{zk}^j can be determined, which is calculated in Eq. 2-3. In other words, the DCT coefficient residual should be a real number between $-q_{zk}/2$ and $q_{zk}/2$. Here, k, j and z are respectively indexes for DCT coefficients, blocks, and channels and q_{zk} is a quantization value from the quantization table for channel z in position k . It should be noted that this range constraint is only applicable to the DCT coefficient residuals from the JPEG blocks. To adaptively impose the range constraint to the predicted outputs, a constraint layer is designed and embedded before the loss layer in the deep network. The adaptive constraint function in the constraint layer is defined as,

$$e_{zk}^j = \begin{cases} -q_{zk}/2, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j < -q_{zk}/2, \\ \hat{e}_{zk}^j, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j \in \left[-\frac{q_{zk}}{2}, \frac{q_{zk}}{2}\right], \\ q_{zk}/2, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j > q_{zk}/2, \\ \hat{e}_{zk}^j, & \text{if block } j \text{ is a non-JPEG block,} \end{cases} \quad (6-4)$$

where $\hat{e}_{zk}^j$ is the output of layer $D-1$. In the back propagation process, the derivative of $\hat{e}_{zk}^j$, Der_{zk}^j , is computed as,

$$Der_{zk}^j = \begin{cases} 0, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j < -q_{zk}/2, \\ 1, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j \in \left[-\frac{q_{zk}}{2}, \frac{q_{zk}}{2}\right], \\ 0, & \text{if block } j \text{ is a JPEG block and } \hat{e}_{zk}^j > q_{zk}/2, \\ 1, & \text{if block } j \text{ is a non-JPEG block.} \end{cases} \quad (6-5)$$

With the proposed constraint layer, the outputs from the JPEG blocks are guaranteed to be in the constraint range.

6.4 A Weight Learning Scheme

As shown in Fig. 6-2, the outputs of the constraint layer are fed into the loss layer to calculate the loss. It is straightforward to utilize mean square error (MSE) to calculate the loss between the predictions and the targets. However, it neglects that different DCT coefficients have different importance to image quality because different quantization values, as shown in the quantization tables in Figs. 2-2 (b) and (c), are applied to different DCT coefficients. It implies that residuals with the same values from different DCT coefficients lead to different level of image quality degradation. Thus, to calculate the loss, the weights of the residuals should be determined. In this subsection, a learning scheme is proposed to estimate the weight of the residual of each DCT coefficient to the image quality. A model is built to capture the relationship between the DCT coefficient residuals and the image quality degradation. SSIM [123] index is adopted as an image quality metric to measure the image quality degradation in the weight learning process.

To calculate the residuals of the DCT coefficients, tattoo image pairs consisting of uncompressed image and the corresponding compressed counterparts are constructed. Uncompressed tattoo images are cropped into patches with size of 80 by 80 pixels. After that, the JPEG compression method is applied to generate the compressed image patches. Both the uncompressed and compressed images are converted into the DCT domain, and the DCT coefficient residuals from the Y, U and V channels are computed. The mean DCT coefficient

residuals from all the blocks are further calculated for the three channels. Since the image quality index is calculated based on the three channels, the mean DCT coefficient residuals from the three channels are concatenated to form a vector with size of 192 and denoted as R . SSIM scores, which is denoted as Q , are calculated for each pair of images. SSIM index is a metric measuring the similarity between a given image and its reference, and the larger value indicates better quality. The range of SSIM index is $(0, 1]$. The value $1 - Q$ is used to measure the image degradation. The weight vector, $W = \{w_1, w_2, \dots, w_{192}\}$, is computed by optimizing the objective function defined as

$$W^* = \underset{W}{\operatorname{argmin}} \sum_{j=1}^N \left\| \sum_{i=1}^{192} w_i R_{i,j} - (1 - Q_j) \right\|^2, \text{ s.t. } w_i > 0, \quad (6-6)$$

where N is the number of image patch, Q_j is the SSIM index from the j^{th} image pairs, $R_{i,j}$ is i^{th} the DCT coefficient residuals from the j^{th} image pairs.

The weight vector W learned through Eq. 6-6 consists of the weight vectors for the Y, U and V channels. Fig. 6-4 gives the three weight vectors for the DCT coefficient residuals of the Y, U and V channels with a quality factor of 50. From the figure, it can be seen that the weights for the low frequency DCT coefficients are higher than those for the high frequency DCT coefficients in all the three channels, which is consistent with the fact that the low frequency DCT coefficients are more important to the image quality.

(a)

(b)

(c)

Fig. 6-4 The learned weight vector for the DCT coefficient residuals of the Y, U and V channels with a quality factor of 50.

The x-axis shows the DCT coefficient index (see Fig. 2-2(a)) and the y-axis shows the weight. (a)-(c) gives the weight vector for the Y, U and V channels, respectively.

With the learned weight vector, the weighted l_2 loss function used in the deep network is defined as

$$l(\theta) = 1/N \sum_{i=1}^N W_z^T ((\mathcal{F}(B_i, \theta) - (D_{o,i} - D_{c,i})) \odot (\mathcal{F}(B_i, \theta) - (D_{o,i} - D_{c,i}))) \quad (6-7)$$

where W_z denotes the weight column vector for channel z , $z = \{Y, U, V\}$; $\mathcal{F}$ denotes the mapping function with parameters θ ; B_i is the input blocks in the RGB domain; $D_{o,i}$ and $D_{c,i}$ are the DCT coefficients from the corresponding blocks in the original images and the JPEG compressed images, respectively; $\odot$ is an element-wise multiplication operator; N is the number of blocks and T denotes the transpose operation. $D_{o,i}$ and $D_{c,i}$ are in a vector form in the order of zigzag (Fig. 2-2(a)).

The proposed restoration algorithm is summarized in Algorithm 6-2.

Algorithm 6-2 A summary of the proposed algorithm

I. The training phase.

Input: Training data consisting of uncompressed image patches and JPEG compressed counterparts, a predefined deep neural network implementing the mapping function $\mathcal{F}$, the learned weight vector W ;

Output: Optimal parameter set θ^* ;

1. Calculate the residual of the DCT coefficients from the training data;
2. Feed the residuals and the corresponding JPEG compressed patches into the deep neural network;
3. Train the deep model illustrated in Fig. 6-2 using the back propagation method;
4. The optimal parameter set is obtained by minimize the objective function in Eq. 6-7, i.e., $\theta^* = \operatorname{argmin}_{\theta} l(\theta)$.

II. The testing phase.

Input: A given compressed image I_c and the learned mapping function $\mathcal{F}(\cdot, \theta^*)$;

Output: Resultant image I_{re} ;

1. Divide the image I_c into JPEG blocks I_{JPEG} and non-JPEG blocks I_{NJPEG} using the scheme introduced in Section 6.2;
-

-
2. Feed both the two sets of blocks into the mapping function $\mathcal{F}(\cdot, \theta^*)$;
 3. Merge the images recovered images from JPEG blocks I_{JPEG} and non-JPEG blocks I_{NJPEG} using Algorithm 6-1 to generate the resultant image I_{re} .
-

6.5 Experimental Results

Tattoo images are collected from the Internet to construct a tattoo image database. It includes a training dataset and a testing dataset. There are 10, 309 tattoo images of different sizes in the training dataset and 1188 tattoo images in the testing dataset, respectively. Fig. 6-5 shows some examples from the database. Fig. 6-5 (a) gives tattoo images from the training dataset and Fig. 6-5 (b) gives tattoo images from the testing dataset. The tattoo images in the database are taken in uncontrolled environments. They are diverse in terms of illumination configuration, scales and body part and so on. It also can be seen that the collected images cover various content, e.g. words, flowers, animals. Images shown in Fig. 6-1 are also from this database.

Fig. 6-5 Some examples from the tattoo image database.
(a) tattoo images from the training dataset. (b) tattoo images from the testing dataset.

Images collected from the Internet are compressed with various quality factors. To generate the uncompressed images for algorithm development, they are downsampled to remove the compressed effects and referred to as uncompressed images. To generate more images for testing, images in the testing dataaet are further downscale into different sizes. In total, the number of testing images is 8649. In this experiment, JPEG compression factor of 30, 40 and 50 are tested. The JPEG compression method is applied to the uncompressed images to generate the compressed counterparts.

6.5.1 Implementation Details

Three deep models are trained independently to recover the DCT coefficients from the Y, U and V channels. To capture the dependence between the neighbour blocks, more than one block is recovered at one time. The input of the deep network is image patches with a size of 32 by 32 pixels, i.e. 4 by 4 blocks, in the RGB domain. The difference between the DCT coefficients from the uncompressed image patches and those from the corresponding compressed image patches are calculated and taken as the targets. It should be noted that the down-sampling operation is applied to the U and V channels in the JPEG compression process. Therefore, the size of the output from the model trained for the Y channel is 4 by 4 blocks, while the size of the output from the model trained for the U and V channels is 2 by 2 blocks. The DCT coefficient residuals are rearranged into a 64-dimensional vector in the order of zigzag (Fig. 2-1(a)). In other words, the size of the output from the model trained for the Y channel is $4 \times 4 \times 64$, and that for the U and V channels is $2 \times 2 \times 64$.

For all the convolutional layers, 64 convolutional filters with a kernel size of $3 \times 3 \times c$ are used to generate 64 feature maps. Here, c refers to the number of channels. Since the input image patches is in the RGB domain, 3 channels are used in the first layer. From layer 2 to layer $D-2$, 64 channels are used. As the last convolutional layer, there is no further dimension decreasing after layer $D-1$. Therefore, in layer $D-1$, c is set to 10, 20 and 34 for the three branches, respectively, which are consistent with the number of DCT coefficients in the low, middle and high frequency groups. There are 16 ($D=15$ in Fig. 6-2) layers and 17 ($D=16$ in Fig. 6-2) layers in the deep networks trained for the Y and the U and V channels respectively. It should be emphasized that the structures in layer 1 to layer $D-2$ is considered as one layer to simplify the presentation.

The MatConvNet package [137] is employed to train the proposed deep models. Stochastic gradient descent (SGD) method is adopted to perform the weight updating. The weight decay is 0.001, and momentum is 0.9. The mini-batch size is fixed as 512. A number of 150

epochs is set as a maximum. The learning rate is decayed exponentially from 1e-2 to 1e-4 after every 50 epochs.

6.5.2 Visual Comparisons

Figs. 6-6 to 6-8 show three sets of restoration results from JPEG compressed images with quality factors of 30, 40 and 50 in the testing dataset. Each figure gives three examples. The JPEG compressed images with quality factors of 30, 40 and 50 and the uncompressed images are shown as references. The state-of-the-art deblocking methods, including OBTCE [101] and ARCNN [110] and another two latest methods based on deep neural network, including DnCNN [111] and TNRD [112] are utilized for comparisons. Among them, the ARCNN [110] method and the DnCNN [111] method perform the restoration in the spatial domain.

Fig. 6-6 The restoration results from the JPEG compressed images with a quality factor of 30. (a)-(c) show three examples. The first row of each example shows the original color image, the JPEG compressed images with a quality factor of 30, the restoration results from OBTC [101] and ARCNN [110], and the second row shows the restoration results from DnCNN [111], TNRD [112] and the proposed algorithm.

Original

JPEG

OBTCE

ARCNN

DnCNN

TNRD

The proposed algorithm

(a)

Original

JPEG

OBTCE

ARCNN

DnCNN

TNRD

The proposed algorithm

(b)

Fig. 6-7 The restoration results from the JPEG compressed images with a quality factor of 40. (a)-(c) show three examples. The first row of each example shows the original color image, the JPEG compressed images with a quality factor of 40, the restoration results from OBTCE [101] and ARCNN [110], and the second row shows the restoration results from DnCNN [111], TNRD [112] and the proposed algorithm.

Fig. 6-8 The restoration results from the JPEG compressed images with a quality factor of 50. (a)-(c) show three examples. The first row of each example shows the original color image, the JPEG compressed images with a quality factor of 50, the restoration results from OBTC [101] and ARCNN [110], and the second row shows the restoration results from DnCNN [111], TNRD [112] and the proposed algorithm.

From the figures, it can be seen that all the deblocking methods used for comparisons can effectively suppress the blocking artifacts. However, the resultant images generated with these methods suffer from blurry effects. The original edge information in the uncompressed images, e.g. the boundaries of the word tattoos (Fig. 6-7(a)), is blurred at the same time. The proposed framework can remove the blocking artifacts effectively and keep the original sharp

information. The recovered images generated with the proposed framework are more natural and closer to the original images.

6.5.3 Objective Evaluations based on SSIM index and PSNR index

To evaluate the proposed algorithm, objective evaluations based on SSIM [123] index and PSNR index are presented. The mean SSIM indexes and the mean PSNR indexes calculating from the tattoo images in the testing dataset are given in Table 6-1. The uncompressed images are used as the reference images and the resultant images from the deblocking methods are used as the test images. The mean SSIM indexes and the mean PSNR indexes from the JPEG compressed images with quality factors of 30, 40 and 50 are also calculated for references.

Table 6-1 (a) Mean SSIM indexes and (b) mean PSNR indexes from the compressed tattoo images

(a) the Mean SSIM Indexes						
	JPEG	OBTCE [101]	ARCNN [110]	DnCNN [111]	TNRD [112]	The proposed algorithm
JPE G 30	0.9467/0.9757	0.9524/0.9781	<u>0.9447</u> /0.9793	0.9556/0.9804	0.9578 /0.9808	0.9572/ 0.9816
JPE G 40	0.9546/0.9796	0.9600/0.9816	0.9517/0.9823	0.9617/0.9836	0.9641 /0.9838	0.9639/ 0.9845
JPE G 50	0.9615/0.9823	0.9650/0.9841	<u>0.9567</u> /0.9844	0.9676/0.9857	0.9686/0.9859	0.9687/0.9866

The best method is highlighted. The methods with mean SSIM indexes lower than those of the JPEG images are underlined.

(b) The Mean PSNR Indexes						
	JPEG	OBTCE [101]	ARCNN [110]	DnCNN [111]	TNRD [112]	The proposed algorithm
JPEG 30	30.85/33.03	31.44/33.62	31.67/33.91	31.89/34.10	32.03/34.16	32.13/34.51
JPEG 40	31.71/33.86	32.26/34.42	32.42/34.65	32.75/34.96	32.84/34.97	32.93/35.29
JPEG 50	32.39/34.53	32.91/35.08	33.03/35.25	33.41/35.62	33.49/35.63	33.59/35.96

The best method is highlighted.

Note that, two values are provided for each method and for each quality metric. The former values are calculated from the whole tattoo images, while the latter values are computed only from the tattoo regions in the images. From Table 6-1(a), it can be seen the proposed deep framework gives the best performance in terms of SSIM indexes from the tattoo regions. The TNRD [112] method and the proposed framework provide the similar performance and outperform the other methods in terms of SSIM indexes from the whole images. Table 6-1(b)

indicates that all the methods can improve the image quality in terms of PSNR indexes and the proposed deep framework performs the best in all the experiments.

6.5.4 Discussions

Some existing deblocking methods based on the deep neural network restore JPEG compressed images in an end-to-end manner, i.e. both the input and the output are images in the spatial domain [110, 111]. The proposed algorithm takes the image patches in the spatial domain as input and produces the DCT coefficient residuals in the frequency domain. With the proposed architecture, the range constraint of the DCT coefficient residuals is explicitly imposed to the predicted results and the constraint operation is also involved in the training process.

After applying the convolution layer without padding, the output size of a deep neural network with convolution layers is generally smaller than the input size. In other words, to generate an image with the same size as the input, pixels outside of the image boundary, i.e. paddings, are demanded. It leads to a problem that the pixels close to the boundary of the images cannot be processed properly. By rearranging the output into a vector, the proposed algorithm can avoid this problem well. The proposed deep framework takes the residuals of DCT coefficients in the frequency domain as the output. It is not necessary to keep the spatial arrangement for the residuals of DCT coefficients in the frequency domain.

6.6 Summary

In this chapter, a deep convolutional neural network is proposed to restore tattoo information from the JPEG compressed tattoo images. This is achieved by estimating the residuals of the DCT coefficients, i.e. the difference between the original DCT coefficients and the compressed DCT coefficients, in the DCT domain. The image patches in the RGB domain are taken as inputs and the corresponding residuals of the DCT coefficients from the same image patches are employed as the targets. To deal with the DCT coefficients with different frequency, a deep neural network with tree-structure is constructed. The DCT coefficients are divided into three groups according to

their frequency and the three branches of the neural network are used to restore the three groups. Batch normalization technique is applied to speed up the training process and improve the restoration performance. To alleviate the blocking effects, an overlapping block restoration scheme is proposed. Blocks directly obtained from the JPEG compressed image files are referred to as JPEG blocks and blocks across the JPEG blocks are referred to as non-JPEG blocks. In the proposed algorithm, both types of blocks are fed into the network for training. In the testing procedure, the JPEG blocks and the non-JPEG blocks from the testing images are recovered. The recovered images from the JPEG blocks and those from the non-JPEG blocks are merged properly to make sure that the resultant images and the JPEG compressed images have the same QDCT coefficients. An adaptive constraint layer is designed to guarantee that the predicted residuals from the JPEG blocks are in the target range. Tattoo images collecting from the Internet are used for algorithm development and evaluation. Experimental results show that the proposed algorithm is effective to remove the blocking artifacts and restore the lost tattoo information and outperforms the existing deblocking methods.

Chapter 7 Conclusion and Future Work

In this chapter, the main findings and contributions of this study are summarized and the limits of the current work are pointed out as well. Some discussions on the future work are also presented.

7.1 Summary

With the rapidly growing popularity of the digital imaging devices, digital evidence images play a more important role in forensic analysis. In some legal cases, such as terrorist attacks and child sexual abuse, only non-facial body parts are shown in the digital evidence images. With these images, skin biometric traits, e.g. tattoos, skin marks or blood vessel patterns, can be extracted and used for criminal and victim identification. However, the images are always compressed with the JPEG compression method, which is the most common used compression method and widely installed in the digital imaging devices. Since the JPEG compression method is a lossy compression technique, the quality of the digital evidence images is degraded. As a result, the skin biometric traits are destroyed or even removed, which makes the criminal and victim identification challenging. Thus, it is necessary to restore the JPEG compressed images before extracting the biometric information.

The existing restoration methods designed for generic images are not always suitable for the biometric information restoration, because they target at removing blocking artifacts and improving the visual quality of the images. It is desirable to develop algorithms for restoring the biometric information from the JPEG compressed skin images. In this research, the problems of restoring blood vessel patterns and tattoos from the JPEG compressed skin images are studied. Skin image analysis and compression test were performed to find out the factors degrading the quality of blood vessel patterns. Three algorithms specially designed for restoring blood vessel patterns were proposed. A deep learning framework was also developed for tattoo image restoration.

To recover the lost blood vessel information in the JPEG compressed skin images, skin image databases were constructed to generate the training data and evaluate the proposed algorithms. A skin image database containing 1894 limb images and its lower resolution version were employed to generate the training data for the development of the blood vessel pattern restoration models. To demonstrate the effectiveness of the trained models, the two skin image databases and a set of diverse skin images collected from the Internet were utilized.

To restore the lost blood vessel information in the JPEG compressed skin images, a skin image analysis, which aimed to figure out how the blood vessel information distributes in the Y, U and V channels, was conducted. A set of compression tests were also performed to find out the critical operations affecting the quality of blood vessel patterns. Statistics of the compressed skin images were studied. The analysis indicated that all the three channels contain blood vessel information but U and V contain more than Y. The test showed that the quantization operation in the JPEG compression method highly degrades the quality of the blood vessel patterns and the first six coefficients in the three channels control the quality of the blood vessel patterns.

Based on the analysis and test, three tailor-made algorithms were developed to restore the missing blood vessel information. First, a preliminary algorithm was proposed to restore the first six DCT coefficients of the U and V channels. It employed the QDCT from cross block and cross channel in the compressed skin images as the features and the original DCT coefficients from the target block as the labels. The decision tree method was trained as the regression models. The trained models were applied to the JPEG compressed skin images to predict the original critical DCT coefficients. Post-processing was applied to guarantee that the restored images had the same QDCT coefficients as the corresponding JPEG compressed versions. The blood vessel uncovering and matching methods were used to generate the blood vessel images and matching results. In addition, SSIM index and PSNR index are employed as image quality indexes. Experimental results demonstrated that the regression scheme in the preliminary algorithm was effective to restore the original blood vessel information. Then, a global algorithm was proposed to restore the first six original DCT coefficients in the Y, U and V channels using the same regression scheme. Gradient boosting algorithm was applied to train the regression models. Experimental results

showed that the global algorithm can restore the blood vessel information effectively and outperform the existing deblocking methods. To make use of the mathematical properties of the JPEG compression process, a multi-model restoration algorithm (MMRA) was developed to remove the blocking artifacts in JPEG compressed images and restore the lost information. Two mathematical properties in the JPEG compression process were identified and utilized to design MMRA. A tailor-made clustering scheme was used to group the training data and regression models were trained to predict the original DCT coefficients for each group. Based on the mathematical properties, it was proven that under a minor assumption, MMRA always offers lower training error than the global algorithm. Evaluations on skin images with different resolutions and different compression factors were performed. Experimental results show clearly that MMRA restores blood vessels significantly and perform better the state-of-the-art deblocking methods in terms of image quality metrics and matching accuracy.

To restore the tattoo information in the JPEG compressed images, a deep net framework was designed. To guarantee that the recovered images have the same QDCT coefficients as the JPEG compressed images, the proposed framework took images in the spatial domain as input and estimated the residuals between the original DCT coefficients and the compressed DCT coefficients in the frequency domain. To deal with the DCT coefficients with different frequency, the DCT coefficients were divided into three groups according to their frequency. The three groups of the DCT coefficients only shared the first layer of the net and were predicted in a parallel manner. To alleviate the blocking artifacts, patches across blocks were also processed. A constraint layer was added to make sure that the predicted residuals for these compressed blocks were in the range determined by the quantization table. A weighted loss function was applied to measure the loss. Because DCT coefficients with different frequency have different impact on the image quality, a weight vector measuring the importance of the DCT coefficients was learned. Tattoo images were collected from the Internet for training data generation and algorithm evaluation in the tattoo information restoration algorithm design. Experimental results showed that the proposed framework can reduce the blocking artifacts effectively and restore the lost information simultaneously

7.2 Future Work

The effectiveness of the proposed algorithms in this study has been demonstrated in many cases. Future work is still demanded for improving their performance robustness.

The proposed blood vessel pattern restoration algorithms directly employ the DCT coefficients from cross neighbours and cross channels in the JPEG compressed images as features for model training. DCT coefficients contain the raw frequency information of the images and they can be obtained by applying discrete cosine transform. Although the trained models are effective to estimate the original DCT coefficients, the feature generation process does not take the properties of blood vessel patterns into account. To further improve the performance of the proposed algorithms, more effective features should be specially designed for blood vessel information restoration. Recently, the learned features utilizing deep neural networks have achieved good performance in many applications [138-144]. With this technique, features can be learned for a specific task from raw data automatically. Deep features also can be learned to restore the blood vessel patterns.

Considering the structure of the blood vessel patterns, the long-term dependence is critical to restore the original information. Meanwhile, the local features are useful to restore more details. Thus, a good feature model should not only catch the long-term dependence, but also extract the local features. Huang *et al.* [145] proposed a densely connected convolutional network, which can integrate the low level and high level features within one deep framework. The experimental results generated from the model are encouraging. This framework can be utilized to learn the features for blood vessel pattern restoration.

In addition, the proposed algorithms process the compressed images in a block based manner. The images are restored block by block and each block is processed independently. As a result, the information across neighbour blocks is not restored well and some blocking artifacts still exist, although they are much weaker than those in images restored by other methods. To further alleviate this problem, the dependence between the restored neighbour blocks should be

captured. Some graphic models, e.g. Markov random fields (MRF) and Conditioned random fields (CRF) can be used to model the dependence between pixels or image patches. These techniques can be used to further improve the proposed algorithms.

A weighted l2 loss function is employed as an objective function. It is simple and easy to calculate, but it may not be the most ideal objective for DCT coefficient restoration. Since it cannot measure the image quality degradation well, a more comprehensive loss function should be designed in the future work.

The proposed framework estimates the original DCT coefficients in the DCT domain. To finally obtain the color images, the inverted discrete cosine transform (IDCT) is applied to convert the DCT coefficients into the spatial domain for the Y, U and V channels. For the U and V channels, an upsampling operation is further applied to generate images with the original size. There is no information loss in the Y channel in this process. However, interpolation is used for upsampling the U and V channels, which can be replaced with super-resolution methods for further improvement. The challenge is to design an algorithm that jointly optimizes the deblocking and super-resolution schemes such that the improvement can be maximized.

Digital evidence images collected from crime scenes are taken in uncontrolled environments with diverse illumination, pose and viewpoints. It is challenging because the performance of the algorithms will be degraded when there is a large difference between training images and testing images. To improve the capability and robustness of the models, more skin images taken under various configurations should be collected in the future work.

References

- [1] M.-Y. Shen, and C.-C. J. Kuo, "Review of postprocessing techniques for compression artifact removal", *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2-14, 1998.
- [2] <http://www.thetimes.co.uk/tto/technology/internet/article4049580.ece>.
- [3] <http://www.heraldscotland.com/news/crime-courts/pictures-of-child-sex-abuse-are-hidden-in-uk-sites.23881400>.
- [4] U. S. D. o. Justice, *National strategy for child exploitation prevention and interdiction: a report to congress*, 2010.
- [5] A. Nurhudatiana, A. W. K. Kong, K. Matinpour, D. Chon, L. Altieri, S.-Y. Cho, and N. Craft, "The individuality of relatively permanent pigmented or vascular skin marks (RPPVSM) in independently and uniformly distributed patterns", *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 998-1012, 2013.
- [6] H. Su, and A. W. K. Kong, "A study on low resolution androgenic hair patterns for criminal and victim identification", *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3-4, pp. 666-680, 2014.
- [7] J.-E. Lee, A. K. Jain, and R. Jin, "Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification", in *BSYM*, 2008, pp. 1-8.
- [8] N. Srinivas, P. J. Flynn, and R. W. Vorder Bruegge, "Human Identification Using Automatic and Semi - Automatically Detected Facial Marks" , *Journal of Forensics Science*, 2015.
- [9] http://www.theharrispoll.com/health-and-life/One_in_Five_U_S_Adults_Now_Has_a_Tattoo.html.
- [10] J.-P. Beauthier, P. Lefevre, and E. De Valck, "Autopsy and identification techniques", *The Tsunami Threat-Research and Technology*: InTech, 2011.
- [11] A. Jain, J.-E. Lee, and R. Jin, "Tattoo-ID: Automatic tattoo image retrieval for suspect and victim identification", *Advances in Multimedia Information Processing*, pp. 256-265, 2007.
- [12] S. T. Acton, and A. Rossi, "Matching and retrieval of tattoo images: Active contour cbir and glocal image features", in *IEEE Southwest Symposium on Image Analysis and Interpretation*, 2008, pp. 21-24.
- [13] B. Heflin, W. Scheirer, and T. E. Boulton, "Detecting and classifying scars, marks, and tattoos found in the wild", in *IEEE International Conference on Biometrics: Theory, Applications and Systems 2012*, pp. 31-38.
- [14] M. J. Wilber, E. Rudd, B. Heflin, Y.-M. Lui, and T. E. Boulton, "Exemplar codes for facial attributes and tattoo recognition", in *IEEE Winter Conference on Applications of Computer Vision 2014*, pp. 205-212.
- [15] H. Han, and A. K. Jain, "Tattoo based identification: Sketch to image matching", in *IEEE International Conference on Biometrics*, 2013, pp. 1-8.

- [16] Q. Xu, S. Ghosh, X. Xu, Y. Huang, and A. W. K. Kong, "Tattoo detection based on CNN and remarks on the NIST database", in *IEEE International Conference on Biometrics*, 2016, pp. 1-7.
- [17] J.-E. Lee, A. K. Jain, and R. Jin, "Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification", in *Biometrics Symposium*, 2008, pp. 1-8.
- [18] C. Roberts, "Biometric technologies-palm and hand ", *Link: <http://www.ccip.govt.nz/newsroom/information-notes/2006/biometricstechnologies-palmhand.pdf>*, 2006.
- [19] L. Wang, and G. Leedham, "Near-and far-infrared imaging for vein pattern biometrics", in *IEEE International Conference on Video and Signal Based Surveillance*, 2006, pp. 52-52.
- [20] A. Kumar, and K. V. Prathyusha, "Personal authentication using hand vein triangulation and knuckle shape", *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2127-2136, 2009.
- [21] A. K. Jain, and A. Ross, "Bridging the gap: from biometrics to forensics", *Philosophical Transactions of the Royal Society B*, vol. 370, no. 1674, pp. 20140254, 2015.
- [22] L. Wang, G. Leedham, and D. S.-Y. Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics", *Pattern recognition*, vol. 41, no. 3, pp. 920-929, 2008.
- [23] J.-G. Wang, W.-Y. Yau, A. Suwandy, and E. Sung, "Fusion of palmprint and palm vein images for person recognition based on" Laplacianpalm" feature", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
- [24] J. H. Song, C. Kim, and Y. Yoo, "Vein Visualization Using a Smart Phone With Multispectral Wiener Estimation for Point-of-Care Applications", *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 773-778, 2015.
- [25] C. Tang, H. Zhang, and A. W.-K. Kong, "Using multiple models to uncover blood vessel patterns in color images for forensic analysis", *Information Fusion*, 2015.
- [26] H. Zhang, C. Tang, A. W. K. Kong, and N. Craft, "Matching vein patterns from color images for forensic investigation", in *IEEE International Conference on Biometrics*, 2012, pp. 77-84.
- [27] T. Watanabe, and T. Tanaka, "Vein authentication using color information and image matching with high performance on natural light", in *ICCAS-SICE*, 2009, pp. 3625-3629.
- [28] C. Tang, A. W. K. Kong, and N. Craft, "Uncovering vein patterns from color skin images for forensic analysis", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 665-672.
- [29] G. K. Wallace, "The JPEG still picture compression standard", *ACM Commun.* , vol. 34, no. 4, pp. 30-44, 1991.
- [30] C. Tang, A.-K. Kong, and N. Craft, "Using a Knowledge-Based Approach to Remove Blocking Artifacts in Skin Images for Forensic Analysis", *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 3, pp. 1038-1049, 2011.
- [31] C. Tang, A. W. K. Kong, and N. 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.

- [32] D. Sun, and W.-K. Cham, "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.
- [33] Y. Li, C. Cai, G. Qiu, and K.-M. Lam, "Face hallucination based on sparse local-pixel structure", *Pattern Recognition*, vol. 47, no. 3, pp. 1261-1270, 2014.
- [34] K. Jia, and S. Gong, "Hallucinating multiple occluded face images of different resolutions", *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1768-1775, 2006.
- [35] X. Yang, G.-D. Su, J. Chen, and Y.-S. Moon, "Restoration of low resolution car plate images using PCA based image super-resolution", in *IEEE International Conference on Image Processing*, 2010, pp. 2789-2792.
- [36] X. Si, J. Feng, and J. Zhou, "Enhancing latent fingerprints on banknotes", in *IEEE International Joint Conference on Biometrics* 2014, pp. 1-8.
- [37] H. Choi, M. Boaventura, I. A. Boaventura, and A. K. Jain, "Automatic segmentation of latent fingerprints", in *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 2012, pp. 303-310.
- [38] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre-and post-filtering", *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1557-1571, 2003.
- [39] A. M. Bruckstein, M. Elad, and R. Kimmel, "Down-scaling for better transform compression", *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1132-1144, 2003.
- [40] Y. Tsaig, M. Elad, P. Milanfar, and 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.
- [41] S. Belkasim, and G. Derado, "Zigzag line discrete cosine transform for blocking artifact removal", in *IEEE 46th Midwest Symposium on Circuits and Systems*, 2003, pp. 540-543.
- [42] H. Xu, J. Xu, and F. Wu, "Lifting-based directional DCT-like transform for image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1325-1335, 2007.
- [43] B. Zeng, and J. Fu, "Directional discrete cosine transforms—a new framework for image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 3, pp. 305-313, 2008.
- [44] C.-L. Chang, M. Makar, S. S. Tsai, and B. Girod, "Direction-adaptive partitioned block transform for color image coding", *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1740-1755, 2010.
- [45] 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.
- [46] Y.-H. Fok, O. C. Au, and C. Chang, "Bit rate and blocking artifact reduction by iterative pre-distortion", in *IEEE International Conference on Image Processing*, , 1996, pp. 13-16.
- [47] B. Macq, "Weighted optimum bit allocations to orthogonal transforms for picture coding", *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 5, pp. 875-883, 1992.

- [48] K. Yamatani, and 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.
- [49] T. Eude, H. Cherifi, and R. Grisel, "Statistical distribution of DCT coefficients and their application to an adaptive compression algorithm", in *IEEE Region 10's Ninth Annual International Conference. Theme: Frontiers of Computer Technology*, 1994, pp. 427-430.
- [50] H. C. Reeve III, and J. S. Lim, "Reduction of blocking effects in image coding", *Optical Engineering*, vol. 23, no. 1, pp. 230134-230134-, 1984.
- [51] B. Ramamurthi, and A. Gersho, "Nonlinear space-variant postprocessing of block coded images", *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 34, no. 5, pp. 1258-1268, 1986.
- [52] Y.-F. Hsu, and 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.
- [53] T. Meier, K. N. Ngan, and G. Crebbin, "A region-based algorithm for enhancement of images degraded by blocking effects", in *IEEE TENCON. Digital Signal Processing Applications*, 1996, pp. 405-408.
- [54] J. Chou, M. Crouse, and K. Ramchandran, "A simple algorithm for removing blocking artifacts in block-transform coded images", *IEEE Signal Processing Letters*, vol. 5, no. 2, pp. 33-35, 1998.
- [55] Y. Lee, H. Kim, and H. Park, "Blocking effect reduction of JPEG images by signal adaptive filtering", *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 229-234, 1998.
- [56] A. Al-Fohoum, and A. M. Reza, "Combined edge crispiness and statistical differencing for deblocking JPEG compressed images", *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1288-1298, 2001.
- [57] Y.-Y. Chen, Y.-W. Chang, and W.-C. Yen, "Design a deblocking filter with three separate modes in DCT-based coding", *Journal of Visual Communication and Image Representation*, vol. 19, no. 4, pp. 231-244, 2008.
- [58] G. Zhai, W. Zhang, X. Yang, W. Lin, and Y. Xu, "Efficient deblocking with coefficient regularization, shape-adaptive filtering, and quantization constraint", *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 735-745, 2008.
- [59] G. Zhai, W. Zhang, X. Yang, W. Lin, and 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.
- [60] G. Zhai, W. Lin, J. Cai, X. Yang, and W. Zhang, "Efficient quadtree based block-shift filtering for deblocking and deringing", *Journal of Visual Communication and Image Representation*, vol. 20, no. 8, pp. 595-607, 2009.
- [61] B. Zeng, "Reduction of blocking effect in DCT-coded images using zero-masking techniques", *Signal processing*, vol. 79, no. 2, pp. 205-211, 1999.
- [62] T. Chen, H. R. Wu, and 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.

- [63] C. Wang, W.-J. Zhang, and X.-Z. Fang, "Adaptive reduction of blocking artifacts in DCT domain for highly compressed images", *IEEE Transactions on Consumer Electronics*, vol. 50, no. 2, pp. 647-654, 2004.
- [64] H. Song, S. Yu, C. Wang, L. Song, and H. Xiong, "A new deblocking algorithm based on adjusted contourlet transform", in *IEEE International Conference on Multimedia and Expo*, 2006, pp. 449-452.
- [65] A. Foi, V. Katkovnik, and 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.
- [66] S. Singh, V. Kumar, and H. Verma, "Reduction of blocking artifacts in JPEG compressed images", *Digital Signal Processing*, vol. 17, no. 1, pp. 225-243, 2007.
- [67] T. Lim, J. Ryu, J. Kim, and J. Jeong, "Adaptive deblocking method using a transform table of different dimension DCT", *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, 2008.
- [68] Y.-W. Chang, and Y.-Y. Chen, "Design of a deblocking filter for both objective and subjective metrics", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 91, no. 8, pp. 2038-2040, 2008.
- [69] S. B. Yoo, K. Choi, and J. B. Ra, "Post processing for blocking artifact reduction", in *IEEE International Conference on Image Processing*, 2011, pp. 1509-1512.
- [70] Y.-K. Lai, J. Li, and C.-C. J. Kuo, "Image enhancement for low-bit-rate JPEG and MPEG coding via postprocessing", in *Visual Communications and Image Processing'96*, 1996, pp. 1484-1495.
- [71] Y. Luo, and 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.
- [72] T. Q. Pham, and L. J. Van Vliet, "Blocking artifacts removal by a hybrid filter method", in *ASCI*, 2005, pp. 372-377.
- [73] D. T. Vo, T. Q. Nguyen, S. Yea, and 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.
- [74] H. Paek, R.-C. Kim, and S.-U. Lee, "A DCT-based spatially adaptive post-processing technique to reduce the blocking artifacts in transform coded images", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 36-41, 2000.
- [75] A. Z. Averbuch, A. Schclar, and 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.
- [76] E. Choi, and M. G. Kang, "Deblocking algorithm for DCT-based compressed images using anisotropic diffusion", in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003, pp. III-717.
- [77] A. Nosratinia, "Enhancement of JPEG-compressed images by re-application of JPEG", *The Journal of VLSI Signal Processing*, vol. 27, no. 1, pp. 69-79, 2001.

- [78] A. Wong, and W. Bishop, "Deblocking of block-transform compressed images using phase-adaptive shifted thresholding", in *IEEE International Symposium on Multimedia*, 2008, pp. 97-103.
- [79] Y. Li, F. Guo, R. T. Tan, and M. S. Brown, "A contrast enhancement framework with JPEG artifacts suppression", in *European Conference on Computer Vision* 2014, pp. 174-188.
- [80] C. Jung, L. Jiao, H. Qi, and T. Sun, "Image deblocking via sparse representation", *Signal Processing: Image Communication*, vol. 27, no. 6, pp. 663-677, 2012.
- [81] H. Chang, M. K. Ng, and T. Zeng, "Reducing artifacts in JPEG decompression via a learned dictionary", *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 718-728, 2014.
- [82] L. Shao, "Unified compression artifacts removal based on adaptive learning on activity measure", *Digital Signal Processing*, vol. 17, no. 6, pp. 1065-1070, 2007.
- [83] G. Qiu, "MLP for adaptive postprocessing block-coded images", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1450-1454, 2000.
- [84] M. N. Florentín-Núñez, E. López-Rubio, and F. J. López-Rubio, "Adaptive kernel regression and probabilistic self-organizing maps for JPEG image deblocking", *Neurocomputing*, vol. 121, pp. 32-39, 2013.
- [85] 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.
- [86] S. S. Choy, Y.-H. Chan, and W.-C. Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients", *IEEE Signal Processing Letters*, vol. 4, no. 1, pp. 5-7, 1997.
- [87] G. Lakhani, and N. Zhong, "Derivation of prediction equations for blocking effect reduction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 3, pp. 415-418, 1999.
- [88] G. A. Triantafyllidis, D. Tzovaras, and M. G. Strintzis, "Blocking artifact detection and reduction in compressed data", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 877-890, 2002.
- [89] S. Minami, and A. Zakhor, "An optimization approach for removing blocking effects in transform coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 2, pp. 74-82, 1995.
- [90] B. Jeon, and J. Jeong, "Blocking artifacts reduction in image compression with block boundary discontinuity criterion", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 345-357, 1998.
- [91] K. Lee, D. S. Kim, and T. Kim, "Regression-based prediction for blocking artifact reduction in JPEG-compressed images", *IEEE Transactions on Image Processing*, vol. 14, no. 1, pp. 36-48, 2005.
- [92] R. Zhang, W. Ouyang, and W.-K. Cham, "Image deblocking using dual adaptive FIR Wiener filter in the DCT transform domain", in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 1181-1184.

- [93] R. L. Stevenson, "Reduction of coding artifacts in transform image coding", in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1993, pp. 401-404.
- [94] T. P. O'Rourke, and R. L. 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.
- [95] J. Luo, C. W. Chen, K. J. Parker, and 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.
- [96] J. Li, and C.-C. Kuo, "Coding artifact removal with multiscale postprocessing", in *International Conference on Image Processing*, 1997, pp. 45-48.
- [97] Z. Li, and 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.
- [98] J. Mateos, A. K. Katsaggelos, and R. Molina, "A Bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts", *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1200-1215, 2000.
- [99] F. Alter, S. Durand, and J. Froment, "Adapted total variation for artifact free decompression of JPEG images", *Journal of Mathematical Imaging and Vision*, vol. 23, no. 2, pp. 199-211, 2005.
- [100] H. Noda, and M. Niimi, "Local MAP estimation for quality improvement of compressed color images", *Pattern Recognition*, vol. 44, no. 4, pp. 788-793, 2011.
- [101] X. Zhang, R. Xiong, X. Fan, S. Ma, and W. Gao, "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity", *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4613-4626, 2013.
- [102] A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 2, no. 1, pp. 91-95, 1992.
- [103] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 3, no. 6, pp. 421-432, 1993.
- [104] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images", *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 896-908, 1995.
- [105] H. Paek, R.-C. Kim, and S.-U. Lee, "On the POCS-based postprocessing technique to reduce the blocking artifacts in transform coded images", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 358-367, 1998.
- [106] C. Weerasinghe, A.-C. Liew, and 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.

- [107] J. J. Zou, and H. Yan, "A deblocking method for BDCT compressed images based on adaptive projections", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 430-435, 2005.
- [108] A.-C. Liew, H. Yan, and N.-F. Law, "POCS-based blocking artifacts suppression using a smoothness constraint set with explicit region modeling", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 795-800, 2005.
- [109] J. Park, D.-C. Park, R. J. Marks, and 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.
- [110] C. Dong, Y. Deng, C. Change Loy, and X. Tang, "Compression artifacts reduction by a deep convolutional network", in *IEEE International Conference on Computer Vision*, 2015, pp. 576-584.
- [111] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising", *IEEE Transactions on Image Processing*, 2017.
- [112] Y. Chen, and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration", *IEEE Transactions on pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256-1272, 2017.
- [113] X. Lee, Y.-Q. Zhang, and A. Leon-Garcia, "Information loss recovery for block-based image coding techniques-a fuzzy logic approach", *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 259-273, 1995.
- [114] Y.-C. Liaw, W. Lo, and J. Z. Lai, "Image restoration of compressed image using classified vector quantization", *Pattern Recognition*, vol. 35, no. 2, pp. 329-340, 2002.
- [115] S. C. Park, M. G. Kang, C. A. Segall, and A. K. Katsaggelos, "Spatially adaptive high-resolution image reconstruction of DCT-based compressed images", *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 573-585, 2004.
- [116] L. Guo, O. C. Au, M. Ma, X. Fan, P. H. Wong, and D. P. Palomar, "Image deblocking using convex optimization", in *IEEE International Conference on Image Processing*, 2008, pp. 3148-3151.
- [117] C. Tang, A.-K. Kong, and N. Craft, "A knowledge-based algorithm to remove blocking artifacts in skin images for forensic analysis", in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 1928-1931.
- [118] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images", *Image Processing, IEEE Transactions on*, vol. 16, no. 5, pp. 1395-1411, 2007.
- [119] Y. Luo, and R. K. Ward, "Removing the blocking artifacts of block-based DCT compressed images", *Image Processing, IEEE Transactions on*, vol. 12, no. 7, pp. 838-842, 2003.
- [120] A. Kong, "An evaluation of Gabor orientation as a feature for face recognition", in *IEEE International Conference on Pattern Recognition*, 2008, pp. 1-4.
- [121] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

- [122] A. Myronenko, and X. Song, "Point set registration: Coherent point drift", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2262-2275, 2010.
- [123] Z. Wang, A. C. Bovik, H. R. Sheikh, and 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.
- [124] "PSNR"; https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.
- [125] <https://www.mathworks.com/help/images/ref/psnr.html>.
- [126] NTU Forensic Skin Database. <http://forensics.sce.ntu.edu.sg/>.
- [127] C. A. Poynton, *A technical introduction to digital video*, p.^pp. 175-176: John Wiley & Sons, Inc., 1996.
- [128] J. Chou, M. Crouse, and K. Ramchandran, "A simple algorithm for removing blocking artifacts in block-transform coded images", *IEEE Singal Processing Letter*, vol. 5, no. 2, pp. 33-35, 1998.
- [129] "Treebagger"; <https://www.mathworks.com/help/stats/treebagger.html>.
- [130] J. H. Friedman, "Greedy function approximation: a gradient boosting machine", *Annals of Statistics*, pp. 1189-1232, 2001.
- [131] V. Vapnik, "The nature of statistical learning theory. 2000", *Red Bank: Springer*, vol. 2.
- [132] FOE. <https://lvdmaaten.github.io/software/>.
- [133] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.
- [134] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", in *International Conference on Machine Learning*, 2015, pp. 448-456.
- [135] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1-8.
- [136] M. Irani, and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency", *Journal of Visual Communication and Image Representation*, vol. 4, no. 4, pp. 324-335, 1993.
- [137] A. Vedaldi, and K. Lenc, "Matconvnet: Convolutional neural networks for matlab", in *ACM international conference on Multimedia*, 2015, pp. 689-692.
- [138] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition", in *European Conference on Computer Vision*, 2016, pp. 499-515.
- [139] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, "Learning deep representation for face alignment with auxiliary attributes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 918-930, 2016.
- [140] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition", in *International conference on machine learning*, 2014, pp. 647-655.

- [141] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, "Learning deep features for scene recognition using places database", in *Advances in neural information processing systems*, 2014, pp. 487-495.
- [142] H. Lee, P. Pham, Y. Largman, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks", in *Advances in neural information processing systems*, 2009, pp. 1096-1104.
- [143] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition", in *British Machine Vision Conference 2015*, pp. 6.
- [144] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection", *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864-879, 2015.
- [145] G. Huang, Z. Liu, K. Q. Weinberger, and L. van der Maaten, "Densely connected convolutional networks", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3.