



Image Denoising and Its Applications in Biomedical and Biometric Images

Author:

Xu, Jinwei

Publication Date:

2015

DOI:

<https://doi.org/10.26190/unsworks/18157>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/54348> in <https://unsworks.unsw.edu.au> on 2024-05-02

Image Denoising and Its Applications in Biomedical and Biometric Images

Jinwei Xu

A dissertation submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy



School of Engineering and Information Technology
Canberra Campus
The University of New South Wales

4 May 2015

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed

Date

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed

Date

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

Date

Abstract

Image denoising techniques are important to cope with different types of noise in biomedical and biometric images. Not only random noise but also general irrelevant components in image spatial domains have been treated as undesired information. To reduce or eliminate the adverse effects of random noise and irrelevant components, this thesis works on following four real-life problems encountered in biomedical and biometric image applications.

Firstly, microscopy images are often degraded by random noise from readout procedures and image data acquisition systems, devices or equipment. This thesis proposes an integration of trend surface mapping, Q-Q plot, bootstrapping, and Gaussian spatial kernel for removing Gaussian-like noise in microscopy images. Furthermore, the proposed approach can be extended to handle Poisson noise. Experiments on synthetic and real noise datasets demonstrate the advantages of the proposed method.

Secondly, medical image classification is challenged by concurrent occurrence of image rotation change, scale variation and noise corruption. This thesis introduces two image features, named particle potential motion entropy histogram (PPMEH) and its updated version PPMEH-FT, incorporated with discrete Fourier transform (DFT), to deal with the multiple effects of rotation, scaling and noise for classifying medical images. The experiments on computed tomography (CT) and magnetic resonance imaging (MRI) datasets show that the proposed image features outperform state-of-the-art methods.

Thirdly, latent fingerprints are usually small-sized, blurred, and overlapped with irrelevant image components. Segmentation of latent fingerprint is very challenging under complex and poor quality image conditions. Furthermore, subsequent latent fingerprint matching is another difficulty. This thesis developed a fully automated latent matcher embedded with a robust latent segmentation module and experiments with a latent fingerprint database demonstrate superiority of the proposed multi-module latent matcher.

This thesis targets three real-life problems in biomedical and biometric image applications. Performance evaluation and comparison with current state-of-the-art approaches validate that the proposed techniques are effective solutions to such problems.

Acknowledgements

First and foremost, I would like to convey my deepest appreciation and gratitude to my supervisors, Prof. Jiankun Hu and Dr. Xiuping Jia, for their invaluable time, patience, fruitful discussion as well as offering amiable and incisive guidance, support, comment, and help during my candidature. Prof. Jiankun Hu has been very encouraging, inspiring and has given me a lot of freedom in conducting my research. I also would like to thank my co-supervisor, Dr. Xiuping Jia, for her unceasing assistance, encouragement and valuable advice at times of need to guide me through my candidature when Prof. Tuan D. Pham was relocated to University of Aizu, Japan. I would like to acknowledge the first year PhD supervision provided by Prof. Tuan D. Pham when he was at The University of New South Wales, Australia.

Next, my special grateful acknowledgement and thank goes to Ms. Wei Zhou and Mr. Wencheng Yang, for their valuable help regarding fingerprint minutiae extraction via VeriFinger software. Furthermore, I would like to thank Dr. Kai Xi in Multi-model Australian Sciences Imaging and Visualization Environment (MASSIVE) for his valuable assistance regarding the guidance of using high performance computing facility for latent fingerprint matching experiment. Also, the MASSIVE is greatly appreciated for its wonderful parallel computing environment and professional directions to facilitate the completion of my PhD thesis.

Publications

Journal Publications:

J. Xu, J. Hu, X. Jia, “A multi-staged automatic restoration of noisy microscopy cell images”, *IEEE Journal of Biomedical and Health Informatics* (retitled from *IEEE Transactions on Information Technology in Biomedicine*), vol. PP, no. 99, pp. 1-10, 2014.

T.D. Pham, D.T. Le, **J. Xu** and et al., “Personalized identification of abdominal wall hernia meshes on computed tomography”, *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 153-161, 2014.

J. Xu, J. Hu, X. Jia, “Particle potential motion entropy histogram: an image feature for medical image classification”, *Computer Methods and Programs in Biomedicine*, under review.

J. Xu, J. Hu, X. Jia, “A fully automated latent fingerprint matcher with embedded self-learning segmentation module”, *IEEE Transactions on Information Forensics and Security*, under review.

Conference Proceedings:

J. Xu, J. Hu, X. Jia, “Genetic algorithm for distorted point set matching”, *Proceedings of 2013 International Congress on Image and Signal Processing*, vol. 3, pp. 1724-1729, 2013.

J. Xu, T.D. Pham, “Robust impulse-noise filtering for biomedical images using numerical interpolation”, *Image Analysis and Recognition, Lecture Notes in Computer Science*, Springer, vol. 7325, pp. 146-155, 2012.

J. Xu, T.D. Pham, X. Zhou, “A double thresholding method for cancer stem cell detection”, *Proceedings of 2011 International Symposium on Image and Signal Processing and Analysis*, pp. 695-699, 2011.

Contents

Abstract	vii
Acknowledgements	ix
Publications	xi
List of Figures	xvii
List of Tables	xxi
1 Introduction	1
1.1 Overview	1
1.2 Research Problem Statement	2
1.3 Contributions	4
1.3.1 Automatic Restoration for Noisy Microscopy Cell Images	5
1.3.2 Multi-robust Image Feature for Medical Image Classification	6
1.3.3 Robust Segmentation Module for Fully Automated Latent Finger- print Matcher	7
1.4 Thesis Outline	9
2 Literature Review	11
2.1 Overview	11
2.2 Survey of Random Noise Cancellation Methods	12
2.2.1 Adaptive Random Noise Filters	12
2.2.1.1 Wiener Filter	12
2.2.1.2 Conventional Linear Filters	14
2.2.1.3 Non-linear Filters	14
2.2.1.4 Bilateral Filter	15
2.2.2 Patch-based Methods	15
2.2.3 PDE and Variational Methods	16
2.2.3.1 Anisotropic Diffusion Filter	17
2.2.3.2 Total Variation Filter	18
2.2.4 Transform Domain-based Methods	19
2.2.4.1 BLS-GSM	19
2.2.4.2 BM3D	20
2.2.4.3 LPG-PCA	22

2.2.5	Dictionary Learning-based Methods	23
2.2.5.1	K-SVD	23
2.2.5.2	LSSC	25
2.2.5.3	CSR	28
2.3	Survey of Irrelevant Component Removal Approaches	29
2.3.1	Spatial Point-based Approaches	30
2.3.1.1	Histogram Thresholding Approach	30
2.3.1.2	Clustering-based Method	30
2.3.2	Spatial Neighborhood-dependent Approaches	31
2.3.2.1	Region Updating Technique	31
2.3.2.2	Evolving Contour Model	33
2.3.2.3	Markov Random Field Model	36
2.3.2.4	Graph Theory-based Technique	38
2.3.2.5	Watershed Approach	39
2.4	Summary	40
3	Automatic Restoration for Noisy Microscopy Cell Images	43
3.1	Introduction	43
3.2	Proposed Integration Approach	46
3.2.1	Image Noise Model and Trend-Surface Mapping	46
3.2.2	Identification of Degraded Pixels by Q-Q Plot	49
3.2.3	Estimation of Noise Statistics by Bootstrapping	51
3.2.4	Restoration of Degraded Pixels using Gaussian Filter	52
3.3	Experimental Results	54
3.3.1	Experiment 1: Removal of Synthetic Gaussian Noise on Yeast Cell Database	54
3.3.2	Experiment 2: Removal Synthetic Poisson Noise on Yeast Cell Database	62
3.3.3	Experiment 3: Segmentation for Real Noisy Phagocytosing Bud-ded Yeast Cell Database	63
3.4	Conclusion	64
4	Multi-robust Feature for Medical Image Classification	69
4.1	Introduction	69
4.1.1	Existing Works on Robustness Against Image Rotation	70
4.1.2	Existing Works on Robustness Against Image Rotation and Scaling	73
4.1.3	The Strategies to Cope with Image Rotation and Scaling	74
4.1.4	Contributions in This Chapter	75
4.2	Proposed Image Feature	76
4.2.1	Adaptive and Noise-robust Local Gradient Estimation	77
4.2.2	Image Gradient Force	81
4.2.3	Image Particle Potential Motion Probability	82
4.2.4	Particle Potential Motion Entropy: An Uncertainty Measurement of Particle Potential Motion And Self Information: A Local Structure Detector	89
4.2.4.1	Particle Potential Motion Entropy	89
4.2.4.2	Self Information	91

4.2.5	Particle Potential Motion Entropy Histogram: A Global Image Feature	93
4.3	Experiment	95
4.3.1	Data Preparation	95
4.3.2	Experiment 1: Local Structure Detection Performance Evaluation	96
4.3.3	Experiment 2: Categorization of Given CT Slices to Correct Colon Lesion Group	99
4.3.4	Experiment 3: Categorization of Given MRI Samples to Correct Head Scan Group	103
4.3.5	Experiment Results Analysis	104
4.4	Conclusion	107
5	Robust Segmentation Module for Fully Automated Latent Fingerprint Matcher	109
5.1	Introduction	109
5.2	Proposed Latent Fingerprint Matching System	116
5.2.1	Dictionary Learning-Based ROI Segmentation Module	116
5.2.1.1	Query Latent Image-Based Dictionary Learning	116
5.2.1.2	Ridge-Valley Atom Identification	119
5.2.1.3	Sparse Coefficient-Based ROI Segmentation	123
5.2.2	Genetic Algorithm-Based Latent Matching Unit	127
5.2.2.1	ROI-Based Minutiae Extraction	127
5.2.2.2	Problem Formulation for Minutiae Set-Based Fingerprint Matching	128
5.2.2.3	Minutiae Set-Based Matching Problem Solved by GA	129
5.3	Experimental Results	132
5.3.1	Data Preparation	133
5.3.2	Experiment 1: ROI-Based Minutiae Extraction	133
5.3.3	Experiment 2: Latent Fingerprint Matching	137
5.4	Conclusion	141
6	Conclusions and Future Studies	143
6.1	Conclusions	143
6.2	Future Studies	145
	Bibliography	147

List of Figures

3.1	Q-Q plot of two samples coming from the same type of distribution (Gaussian) but with different parameters	50
3.2	Q-Q plot of two samples coming from different distribution families	50
3.3	Four DIC yeast cell images used in synthetic noise experiment: (a) Cell 1; (b) Cell 2; (c) Cell 3; and (d) Cell 4	54
3.4	Trend surfaces and Q-Q plots of Figure 3.3(a): (a) Original Figure 3.3(a); (b) Corrupted Figure 3.3(a) by synthetic Gaussian noise ($\sigma_n = 30$); (c) and (d) are trend surfaces for (a) and (b); and (e) and (f) are Q-Q plots corresponding to (a) and (b) respectively	56
3.5	Restoration of Figure 3.3 (a) degraded by synthetic Gaussian noise ($\sigma_n = 30$): (a) BF; (b) WF; (c) NLM; (d) TV; (e) PDE; (f) FUZZY; and (g) MAR2	57
3.6	The comparison of noise level estimation between MAR2 and WF for Figure 3.3(a)	61
3.7	Segmentation results given by morphological filtering based on real noisy cell image and restored images: (a) ground truth obtained by manual segmentation; (b) real noisy image; (c)-(i) restoration results given by 7 approaches: BF, WF, NLM, TV, PDE, FUZZY and MAR2, respectively; (j)-(q) corresponding binary segmentation results, where the optimal size of “disk” structure element r is assigned respectively as follows: (j) $r = 3$, (k) $r = 6$, (l) $r = 6$, (m) $r = 10$, (n) $r = 7$ or 8 , (o) $r = 7$, (p) $r = 4$ and (q) $r = 8$	65
4.1	The rotation issues encountered in colon CT and head MRI datasets: (a) face-up posture; (b) face-down posture; (c) patient body is inclining to the left; (d) patient body is inclining to the right; (e) patient head is inclining to the left; and (f) patient head is inclining to the right	71
4.2	The scaling issues encountered in colon CT and head MRI datasets: (a) the obesity of the patient abdomen; and (b) the thinness of the patient abdomen expressed in different colon CT slices	72
4.3	The noise issues encountered in head MRI dataset	72
4.4	The illustration of the three binary indicator maps to highlight the local structure pixels: (a) original image (head CT scan); (b) the binary indicator map obtained based on the sequential pattern generated by the VFDD with fine frequency resolution $\pi/4$ and K-means clustering ($k = 2$); (c) the binary indicator map obtained based on the single PPME value by K-means clustering ($k = 2$); and (d) the binary indicator map obtained based on the sequential pattern generated by the VFDD with coarse frequency resolution $\pi/2$ and K-means clustering ($k = 2$)	88

- 4.5 The PPME calculated at each pixel location is illustrated: (a) original image (head CT scan); and (b) the PPME map generated at each pixel location 91
- 4.6 The illustration of potential probability distribution, self-information and DFT-based shift invariant feature in regard to local structure pixel (LSP) and non-structure pixel (NSP), respectively: (a) Potential probability distribution for an LSP centrally located in the 3×3 neighbor: [128, 157, 160; 124, 166, 165; 111, 154, 152]; (b) and (c) are the corresponding $\{h_{\theta_k}^{LSP}\}$ and $\left|DFT\left(\{h_{\theta_k}^{LSP}\}\right)\right|$, respectively; (d) potential probability distribution for an NSP centrally located in the 3×3 neighbor: [7, 0, 0; 0, 5, 0; 5, 0, 1]; (e) and (f) are the corresponding $\{h_{\theta_k}^{NSP}\}$ and $\left|DFT\left(\{h_{\theta_k}^{NSP}\}\right)\right|$ 92
- 4.7 The illustration of image rotation robustness for the proposed PPMEH-FT: (a) original image; (b) and (c) correspond to PPMEH and PPMEH-FT, respectively; (d) 90° anticlockwise rotation happens to (a); (e) and (f) are PPMEH and PPMEH-FT corresponding to (d). The difference between (b) and (e) exhibits that the image rotation directly results in the translation of the histogram pattern. (c) and (f) demonstrate that very similar patterns can be achieved after using DFT and histogram normalization 94
- 4.8 The illustration of image scaling adaptation for the proposed PPMEH-FT: (a) the image with scaling factor $s = 0.8$; (b) and (c) are PPMEH and PPMEH-FT corresponding to (a); (d), (e) and (f) are original image (scaling factor $s = 1$), PPMEH and PPMEH-FT, respectively; (g) the image with scaling factor $s = 1.2$; (h) and (i) are corresponding PPMEH and PPMEH-FT, respectively. The differences between (b), (e) and (h) demonstrate that the image scaling-up or scaling-down leads to the enlargement or shrinkage of the histogram bin amplitude. However, the histogram patterns expressed by (b), (e) and (h) still look similar and the magnitude spectrum features extracted by DFT are also similar in (c), (f) and (i), respectively 95
- 4.9 The binary local structure indicators obtained depending on diverse features and K-means clustering ($k = 2$ is applied to all the features): (a) ground-truth local structure image; (b) noisy local structure image degraded by synthetic Gaussian noise ($\sigma_n = 30$); (c) pixel intensity \mathbf{I} ; (d) directly computed gradient magnitude $\sqrt{z_{x_1}^2 + z_{x_2}^2}$; (e) two directly computed first derivatives $[z_{x_1}, z_{x_2}]$; (f) adaptively estimated local gradient magnitude $|\nabla\mathbf{I}|$; (g) two first derivatives obtained depending on local gradient $\nabla\mathbf{I}$ and its direction $\theta_{\mathbf{G}}$ [$\nabla\mathbf{I}_{x_1}, \nabla\mathbf{I}_{x_2}$] ($\nabla\mathbf{I}_{x_1} = |\nabla\mathbf{I}| \cdot \cos(\theta_{\mathbf{G}})$ and $\nabla\mathbf{I}_{x_2} = |\nabla\mathbf{I}| \cdot \sin(\theta_{\mathbf{G}})$); (h) single PPME value; (i) sparse coding of $\nabla\mathbf{I}$ conducted by OMP [263]; (j) sequential pattern $|DFT(\{h_{\theta_k}\})|$ via the VFDD with coarse frequency resolution $\pi/2$; and (k) sequential pattern $|DFT(\{h_{\theta_k}\})|$ via the VFDD with fine frequency resolution $\pi/4$ 98
- 4.10 The illustration of sliding window size selection to cope with CT image random scaling in Scenario 4 and Scenario 5, respectively, where the classification accuracy of PPMEH is adopted as the selection criterion 101
- 4.11 The illustration of sliding window size selection to cope with MRI image random scaling in Scenario 4 and Scenario 5, respectively, where the classification accuracy of PPMEH is adopted as the selection criterion 103

4.12	The illustration of the noise robustness of the proposed local structure detector: (a) the colon CT slice in Figure 4.1c degraded by the synthetic Gaussian noise ($\sigma_n = 30$); (c) the heads MRI scan in Figure 4.3a and (e) Figure 4.3b carry real noise; (b), (d) and (f) are the corresponding binary indicators obtained by the proposed local structure detector	108
5.1	Three types of fingerprint: (a) the rolled print; (b) the plain print; and (c) the latent print	110
5.2	Challenges involved in latent print images: (a) small area in ROI; (b) overlapping with a variety of structured noise; and (c) blur	111
5.3	The dictionary \mathbf{D} is learned directly from the given query latent image by ODL: (a) original query latent image (U224 in NIST SD27); and (b) its learned dictionary \mathbf{D} ($N_a = 100$ and $K = 2$)	118
5.4	The illustration of the details during the automated “ridge-valley” atom identification procedure: (a) the “ridge-valley” atom patch ($p_{d_{66}}$ in Figure 5.3b); (b) the spectral magnitude map $ DFT(p_{d_{66}}) $; (c) the frequency bandpass filter corresponding to the ridge period range $[3 \text{ pixels}, 20 \text{ pixels}]$; (d) the constrained spectral magnitude map depending on (c); (e) the calculated orientation $o_{d_{66}}$; (f) the reconstructed “ridge-valley” pattern $u_{d_{66}}$; and (g) the 1D sinusoidal-shaped wave modeled based on the pixel intensities along the green dashed line indicated in (f)	120
5.5	The examples of atom-based identification for various atoms: (a) the “ridge-valley” atom ($p_{d_{66}}$ in Figure 5.3b); (b) and (c) are the reconstructed $u_{d_{66}}$, the calculated $xcorr_{p_{d_{66}}, u_{d_{66}}}$ and $v_{d_{66}}$ for (a), respectively; (d) the line-like atom ($p_{d_{45}}$ in Figure 5.3b); (e) and (f) are the reconstructed $u_{d_{45}}$, the calculated $xcorr_{p_{d_{45}}, u_{d_{45}}}$ and $v_{d_{45}}$ for (d), respectively; (g) the periodical speckle atom ($p_{d_{57}}$ in Figure 5.3b); (h) and (i) are the reconstructed $u_{d_{57}}$, the calculated $xcorr_{p_{d_{57}}, u_{d_{57}}}$ and $v_{d_{57}}$ for (g), respectively	121
5.6	The example of the identification results obtained by using the proposed automated “ridge-valley” atom identification procedure: (a) the learned dictionary D in Figure 5.3b; and (b) the identified “ridge-valley” atoms. The white patches indicate that the original atoms with the same patch-wise coordinates in (a) are “ridge-valley” ones, while the black patches label of the corresponding atoms in (a) are “non-ridge-valley” ones	123
5.7	The illustration of the details during the automated sparse coefficient-based ROI segmentation procedure: (a) the two types of image block in the latent print image – the foreground block (marked by a pink square) and the background block (marked by a green square); (b) the marked foreground block; (c) seeking the specific atom corresponding to the $\max(\gamma_f^{(*)})$ for (b) (the found atom is identified as the “ridge-valley” atom in the preceding phase); (d) the marked background block; and (e) seeking the specific atom corresponding to the $\max(\gamma_b^{(*)})$ for (d) (the found atom is labelled as the “non-ridge-valley” atom in the preceding phase)	125

5.8	The example of the ROI obtained by using the proposed sparse coefficient-based ROI segmentation procedure: (a) the image \mathbf{M} obtained after all the signal vectors have been processed; (b) the binary image obtained by using the normalization and Otsu's adaptive threshold-based segmentation method; and (c) the polygonal ROI obtained via the mathematical morphology operation and the maximal-area convex polygon extraction (marked by a pink convex polygon over the original latent print image U224 in NIST SD27)	127
5.9	The example of successful matching of a genuine "rolled-latent" pair is by the proposed GA-based matching unit. In this case, seven minutiae pairs are found. The paired minutiae between the rolled and latent print are connected by green lines and marked by red solid points, respectively. (a) the rolled print image (left) corresponds to G009 in NIST SD27 (right); the automatically extracted minutiae in the rolled print image are marked by yellow hollow points and the ones in the latent print image are marked by red hollow points; the ROI is highlighted by a pink convex polygon and the automatically extracted minutiae within the ROI are marked by green solid points; and (b) the augment of the paired minutiae number with the iteration of GA	132
5.10	The example of three scenarios on U267 in NIST SD27: (a) the ground-truth minutiae set provided by NIST SD27; (b) Minutiae Extraction Scenario 1 - four genuine minutiae can be correctly extracted, $GMPR = 1$, $FMAR = 1$, and $AUC = 0.5$; (c) Minutiae Extraction Scenario 2 - four genuine minutiae can be correctly extracted, $GMPR = 1$, $FMAR = 0.3167$, and $AUC = 0.8416$; and (d) Minutiae Extraction Scenario 3 - only one genuine minutia can be correctly detected, $GMPR = 0.25$, $FMAR = 0.0667$, and $AUC = 0.5916$	136
5.11	Different \overline{CMC} obtained by the three scenarios.	138
5.12	\overline{CMC} obtained in Matching Scenario 1 in regard to the three types of latent image: Good, Bad and Ugly.	140
5.13	\overline{CMC} obtained in Matching Scenario 2 in regard to the three types of latent image: Good, Bad and Ugly.	140
5.14	\overline{CMC} obtained in Matching Scenario 3 in regard to the three types of latent image: Good, Bad and Ugly.	141

List of Tables

3.1	Selection of polynomial order p according to the criterion of “minimal MSE” for different cell images degraded by synthetic Gaussian noise with varying intensity levels	55
3.2	Comparison of restoration results from 8 approaches for cell image 1 degraded by varying intensities of synthetic Gaussian noise	60
3.3	Comparison of restoration results from 8 approaches for cell image 2 degraded by varying intensities of synthetic Gaussian noise	60
3.4	Comparison of restoration results from 8 approaches for cell image 3 degraded by varying intensities of synthetic Gaussian noise	60
3.5	Comparison of restoration results from 8 approaches for cell image 4 degraded by varying intensities of synthetic Gaussian noise	61
3.6	Comparison of computational time (seconds) for single cell image with the size of 512×512	61
3.7	Comparison of restoration results from 7 approaches for cell images 1-4 degraded by synthetic Poisson noise	63
3.8	Statistical comparison for segmentation performance (evaluated by average mean ROC attributes: \overline{AUC} , \overline{TPR} and \overline{FPR}) based on 7 filtering approaches	64
3.9	The selection of optimal size of “disk” structure element according to the criterion of “higher TPR and lower FPR”, the term in brackets stands for (optimal structure element size)	67
4.1	The Local Structure Detection Performance Evaluated by ROC Attributes	97
4.2	Image Feature Comparison for Image Random Rotation and Scaling in Colon CT Images Classification Experiment (Classification Rate: $\mu \pm \sigma$) .	102
4.3	Image Feature Comparison for Image Random Rotation and Scaling in Head MRI Images Classification Experiment (Classification Rate: $\mu \pm \sigma$) .	104
5.1	The comparison of mean GMPR (\overline{GMPR}), mean FMAR (\overline{FMAR}) and mean AUC (\overline{AUC}) among three scenarios	135
5.2	Different $\overline{p\bar{r}}$ obtained by the three scenarios	139
5.3	$\overline{p\bar{r}}$ obtained in three scenarios according to the three categories of latent image: Good, Bad and Ugly	140

Chapter 1

Introduction

1.1 Overview

Image denoising has been an intensively studied research area in the image processing community, and continues to serve as an important preprocessing procedure prior to many image applications. With the evolution of image denoising technologies and the advancement of state-of-the-art methodologies, a series of subsequent image applications such as object representation, object recognition, and object classification are greatly facilitated. As a result of image denoising, the resultant images are expected to be in good quality, and accordingly, the following image analysis tasks become easy to handle.

Image noise can be categorized into the following two types: one is random noise [1–10], and the other is irrelevant components [11–20], according to the visual information quality and the content of interest.

- **Random Noise:** image random noise generally originates from the readout procedures and image data acquisition systems, devices, and equipments. As random noise is caused by imaging defects, it is considered as the principal factor in the degrading of image quality and the breaching of visual information. Due to the deterioration of image quality and the lack of reliable visual information, the subsequent image applications could be more challenging to tackle; that is, the visual information reflected by images is mostly unreliable in the presence of random

noise, therefore, the subsequent image applications are most likely to be confronted with straitened circumstances and the subsequent image processing and analysis tasks are probably to be affected.

- **Irrelevant Component:** irrelevant image component is determined by the user's interest; that is, the user interest regulates which image component is the target, and which image component is irrelevant. Subject to the user's interest, the target image component is regarded as object, whereas the irrelevant image component is treated as noise. Since the existence of irrelevant components, the objects in images are not salient to be captured but overlapped with redundant visual information. As a consequence, those object-based image applications, in terms of object representation, object recognition, and object classification, could become more difficult to handle.

Both random noise and irrelevant components have been extensively observed in biomedical and biometric images [4, 5, 8, 10, 13, 18–20]. To be explicit, the biomedical images, including biological images and medical images, are usually contaminated by random image noise, due to the measuring of defects during the imaging procedures. For example, the microscopy images, one category of typical biological images, are quite often corrupted by Gaussian and Poisson noise. The MRI scans, belonging to the medical images and providing the anatomic details of the human body in clinical diagnosis, are probably degraded by Gaussian noise. Being different from the biomedical images, the biometric images are interfered with by the effect of irrelevant image components. As an example, latent fingerprint images, one type of representative biometric images, are unavoidably overlapped with irrelevant structural components.

1.2 Research Problem Statement

The denoising techniques, not only for random image noise but also for irrelevant image components, have been intensively studied and greatly improved during the last decade. However, the following research topics are still open and need to be further investigated to meet real-life demands.

- Microscopy imaging allows scientists to study biological cells at the level where a single cell can be observed, and it provides a powerful tool for examination of both living and fixed cells in cell culture. Microscopy imaging technology has multiple advantages in biological applications, although such images are often degraded by random noise from readout procedures and image data acquisition systems, devices or equipment. Such image degradation imposes an adverse impact on the quality of microscopy imaging and the subsequent automated cell segmentation. In order to effectively restore the degraded microscopy images, several computerized restoration approaches have been developed during the past years. However, microscopy image restoration problem is still open; that is, in biological practice, not only the restored image quality needs to be improved but also the restored images are expected to be beneficial for the subsequent cell segmentation.
- Medical imaging allows a radiologist to visually analyze the anatomic details about the normality and abnormality of tissues inside the human body. However, the random noise contamination, which affects medical images, is pervasive. The mechanisms of random noise formation in computed tomography (CT) and magnetic resonance imaging (MRI) modalities have been extensively studied, and a large number of medical image denoising methods have been proposed in recent years. Nevertheless, few works on image descriptor's tolerance for random noise corruption are reported in medical image classification community; that is, the research gap, in terms of developing a robust medical image descriptor, is still wide. Furthermore, apart from random noise issues, the coexistence of rotation and scale variation in medical image data is also usually encountered. Medical image rotation is caused by the difficulty of perfectly positioning a patient's body in CT or MRI scan modality, and the tissue scale variation is inevitable, due to the varying patient body size. Therefore, there is a particular need to develop an image feature tolerant to the concurrent occurrences of random noise, image rotation, and scale variation in medical image classification tasks.
- A latent fingerprint, inadvertently left by a person at a crime scene is used to identify the criminal or to exclude a suspect. Latent fingerprint images are usually small-sized, blurred, and overlapped with irrelevant image components, therefore latent fingerprint images are generally of poor quality and latent fingerprint segmentation is indeed a challenging task. During the past years, several techniques have been

presented to address the latent fingerprint segmentation problem. However, such conventional methods have not achieved satisfactory segmentation performance to effectively support the subsequent minutiae extraction and latent fingerprint matching; that is, as the result of latent segmentation, the regions of interest (ROIs) are applied to extract the minutiae. Since the resultant ROIs obtained by conventional methods contain more irrelevant image components, such as character, stain and structural graphic patterns, a large number of imposter minutiae are falsely detected while few genuine ones can be correctly extracted. Depending on such unreliable minutiae extraction, the subsequent latent fingerprint matching becomes extremely difficult and the matching performance accordingly deteriorates. The overlapping irrelevant components are unavoidable in latent fingerprint images. Therefore, for reliable minutiae extraction and subsequent latent fingerprint matching, it is essential to develop a robust latent fingerprint segmentation method, not only to effectively capture the valid fingerprint ridge regions but also to significantly remove the irrelevant image components.

In summary, we identify the following three research targets in this thesis to tackle the above-mentioned issues:

- Propose an automatic restoration scheme for noisy microscopy cell images, not only to significantly suppress the random noise but also to preserve cellular structures well for subsequent cell segmentation.
- Introduce a multi-robust medical image feature against random noise, rotation, and scale variation for CT and MRI scans classification.
- Develop a fully automated latent fingerprint image segmentation module to be tolerant against the variety of irrelevant components such as stain, character and graphic pattern.

1.3 Contributions

The primary contributions in addressing the above three research problems are shown as follows.

1.3.1 Automatic Restoration for Noisy Microscopy Cell Images

We have proposed a new scheme through trend surface mapping, Q-Q plot, bootstrapping, and Gaussian spatial kernel for removing Gaussian-like noise in microscopy cell images. The integrated approach performs as an adaptive image restoration which achieves better results compared with state-of-the-art restoration methods. One advantage of the proposed approach is that its methodology is flexible and therefore can be extended to handle Poisson noise. Besides, the potential improvement is the appropriate selection of noisy pixel candidates, according to the linear pattern along a reference line in a Q-Q plot. Such selection could improve the reliability of noise estimation by bootstrapping. Above all, the proposed multi-staged automatic restoration method provides a superior denoising scheme, and could be a highly valuable step prior to cell segmentation for numerous generated and widely used microscopy cell images. The distinctive advantages of the proposed automatic restoration scheme are summarized as follows:

- Brightness trend is one special characteristic in microscopy cell images. Therefore, trend surface mapping, inspired by geostatistics, is employed to model the trend of image brightness (it is regarded as the signal trend).
- The Q-Q plot is utilized to identify highly corrupted pixels by comparing the pixel intensities versus known probability distribution (here, the Gaussian noise model is assumed, therefore Gaussian distribution is assigned as the known distribution). Furthermore, noise intensity level estimation is achieved by bootstrapping only, depending on the identified corrupted pixels.
- The smoothing parameter of the Gaussian spatial kernel is automatically determined according to the estimation of bootstrapping, which effectively reduces under-smoothing or over-smoothing.
- With the integration of Variance Stability Transform (VST), the proposed multi-staged restoration scheme can be generalized to deal with Poisson noise, which is another commonly encountered noise in microscopy cell images.

1.3.2 Multi-robust Image Feature for Medical Image Classification

A series of new concepts, such as image gradient force, particle potential motion probability, particle potential motion entropy and self-information are introduced, defined and interpreted, respectively. Based on such concepts, two robust image features, PPMEH and PPMEH-FT, are proposed to cope with the multi-effect of image rotation, scaling and noise in medical image classification tasks. The experimental results on colon CT and head MRI image datasets demonstrate that the presented PPMEH and PPMEH-FT provide generally satisfactory classification performance in contrast to other state-of-the-art image features. Besides, the experimental results suggest that the proposed PPMEH and PPMEH-FT can be applied to a medical image classification and retrieval system with large data-bases. Detailed contributions of the proposed multi-robust image feature are summarized as follows:

- To detect local structures with various orientations and scales, we introduce an adaptive estimation strategy to simultaneously identify local dominant orientation and scale. In addition, for estimating local gradient (local gradient is used to represent the dominant local pixel intensity variation), a Gaussian first-derivative kernel is adaptively constructed, based on the identified local orientation and scale.
- To further resist the noise effect and better preserve the local structure, the anisotropic kernel (elliptical shape) instead of the isotropic kernel (circular shape) is adopted, not only to suppress the noise, but also to preserve the local salient structure along its dominant orientation.
- To quantify the isotropy or anisotropy of pixel-wise particle potential motion, a new feature entitled particle potential motion entropy (PPME) is presented.
- To discriminate between local structure and non-structure pixels, a new sequential feature, named self-information, is proposed.
- For image representation, a new histogram-based global image feature, entitled particle potential motion entropy histogram (PPMEH), is introduced.
- To achieve image rotation robustness, an upgraded version of PPMEH (named PPMEH-FT), modified by discrete Fourier transform (DFT), is proposed.

1.3.3 Robust Segmentation Module for Fully Automated Latent Fingerprint Matcher

We have proposed a fully automated minutiae set-based matcher embedded with a robust segmentation module against irrelevant components in latent fingerprint images. The proposed latent matcher integrates the following two schemes: (i) dictionary learning (DL)-based ROI segmentation scheme, robust against image irrelevant components; and (ii) GA-based minutiae set matching unit, robust against spurious minutiae. For the DL-based ROI segmentation scheme, the dictionary is firstly learned from the query latent fingerprint image. Then, based on the learned dictionary, “ridge-valley” pattern elements (dictionary atoms) can be automatically identified. Further, the sparse representation for the original latent image patches is performed. Finally, depending on the presence or absence of the sparse coefficients that correspond to the identified “ridge-valley” atoms, the foreground (fingerprint region) is segmented. In the GA-based minutiae-level matching unit, the two minutiae sets, one from the segmented ROI in query latent image (obtained via segmentation module) and the other one, from the print currently being compared, are extracted beforehand, through a normal automated minutiae extraction program like Verifinger SDK, which is widely available to the public. Then, according to the affine transformation parameters estimated by GA, the minutiae set alignment between the query latent and the compared print is performed. Furthermore, after aligning the two sets of minutiae, the correspondence between the two sets needs to be found. Accordingly, the corresponding minutiae points between the query latent and the compared print can be paired. Finally, the number of matched minutiae is obtained and simply regarded as the matching score. The distinctive advantages of the proposed robust segmentation module and the introduced latent fingerprint matcher are summarized as follows:

- One commercial off-the-shelf (COTS3) matcher has been assigned as a latent matcher for latent fingerprint matching, however, such matcher is unavailable in public domain and unaccessible to the researchers [253, 257]. Except COTS3, there is no fully-automated latent matcher. Therefore, in this thesis, we introduce a multi-module matcher to cope with the latent fingerprint matching problem. The proposed system is performed in a fully automatic mode. Experimental results, based on NIST SD27, demonstrate that the proposed matcher with the proposed

segmentation module (SM) (say, proposed matcher + proposed SM) can achieve 33.692% penetration rate (in biometric community, penetration rate is defined as average number of results returned by a pre-matching or indexing algorithm as proportion of total database size). Comparative experiments have been conducted to evaluate the effect of the different SMs by using the proposed matcher with and without SM. By designating the proposed matcher without SM as the baseline (say, proposed matcher only), such benchmark penetration rate is 38.728%. Based on the benchmark, the proposed matcher with a state-of-the-art SM only achieves 38.321% penetration rate. The relative penetration rate enhancement percentage for “proposed matcher + proposed SM” ($13.00\% = \frac{|33.692\% - 38.728\%|}{38.728\%}$) is much better than that of “proposed matcher + SM [253]” ($1.05\% = \frac{|38.321\% - 38.728\%|}{38.728\%}$).

- The fully automated ROI segmentation module is plugged into the proposed latent matcher and is performed as the pre-processing stage of the subsequent matching task. The proposed SM consists of the following phases: (i) image structure dictionary learning; (ii) “ridge-valley” atom identification; and (iii) sparse coding and ROI segmentation. The existing method requires the establishment of a dictionary from the high-quality rolled image patches in advance [253]. Different from such conventional methods, we propose to build up the structure dictionary, directly learned from the query latent image. A learned dictionary based on the target image can better adapt to the target image. Therefore, the dictionary obtained in the proposed SM is not based on good quality rolled image patches but is dependent on query latent images.
- The robust latent matching unit consists of the following stages: (i) ROI-based minutiae extraction; (ii) GA-based minutiae set alignment; and (iii) counting of paired minutiae. The existing method demands the yielding of the local minutiae descriptors during the iteration of GA optimization. Different from such conventional methods, the global topology of the entire minutiae set is directly adopted in the proposed matching unit instead of the local minutiae structure. Accordingly, the proposed matching unit is more robust in the presence of spurious minutiae and more efficient in terms of matching.

1.4 Thesis Outline

The remaining parts of this thesis are organized as follows:

- **Chapter 2: Literature Review**

This chapter reviews the related works in the literature. This chapter covers the following two streams: (i) the methodologies to cancel random noise; and (ii) the technologies to remove irrelevant components.

- **Chapter 3: A Multi-staged Automatic Restoration of Noisy Microscopy Cell Images**

This chapter consists of the following four sections: (i) at the beginning, recently developed approaches on noisy microscopy image restoration are reviewed, and consequently, the research challenge and motivation are presented; (ii) then, the proposed restoration scheme, integrating with trend surface mapping, quantile-quantile plot, bootstrapping and Gaussian spatial filtering kernel, is introduced; (iii) performance comparison among state-of-the-art methods in synthetic noise experiments, and the availability test on real noisy microscopy data are conducted, and consequently, the experimental results are discussed; and (iv) finally, some concluding remarks are given.

- **Chapter 4: Multi-robust Image Feature for Medical Image Classification**

This chapter includes the following four parts: (i) in the first place, the tolerance capability of existing methods against image rotation, scaling, and random noise are studied, and then the research problem and motivation, especially for medical image analysis, are stated; (ii) secondly, the proposed image feature, incorporating local gradient estimation, image gradient force calculation, particle potential motion probability, particle potential motion entropy, and self-information, is presented; (iii) thirdly, local structure detection performance evaluation on synthetic structure dataset, and the image representation performance comparison among state-of-the-art features on colon CT and head MRI scans are conducted and analyzed; and (iv) at the end, conclusion is provided.

- **Chapter 5: Robust Segmentation Module for Fully Automated Latent Fingerprint Matcher**

This chapter covers the following four units: (i) first of all, the recently reported matching and segmentation techniques on latent fingerprint are reviewed, and then the research gap and motivation are stated; (ii) afterwards, the proposed latent fingerprint matching system, which consists of the dictionary learning-based ROI segmentation phase and the genetic algorithm-based minutiae set matching phase, is introduced; (iii) later, based on NIST SD27 dataset, the introduced matcher is evaluated by ROI-based minutiae extraction and latent-to-rolled fingerprint matching, respectively; and (iv) lastly, the conclusive statement is summarized.

- **Chapter 6: Conclusions and Future Studies**

This chapter generally summarizes the findings achieved in this thesis, and briefly indicates the future studies which would be pursued as the extension of the current thesis work.

Chapter 2

Literature Review

2.1 Overview

This chapter reviews the recently developed works related to random noise cancelation and irrelevant component removal. The remainder of this chapter is structured as follows:

- **Section 2.2: Survey of Random Noise Cancelation Methods**

This section consists of the following five Subsections from [2.2.1](#) to [2.2.5](#). These five subsections primarily provide a broad review regarding the methods, which were presented in recent years and regarded as the state-of-the-art techniques in the random noise filtering field (they are: adaptive filters, patch-based methods, PDE and variational methods, transform domain-based methods, and dictionary learning-based methods); after that, their advantages and disadvantages are analyzed in details.

- **Section 2.3: Survey of Irrelevant Component Removal Approaches**

This section includes the following two Subsections [2.3.1](#) and [2.3.2](#). In Subsection [2.3.1](#), the extensive techniques in regard to “spatial point operation” are reviewed and commented on; then, in Subsection [2.3.2](#), shifting to “spatial neighborhood operation”, the related methods are discussed.

- **Section 2.4: Summary**

This section provides the general statement for the above-mentioned two streams of methodologies when canceling the random noise for noisy images and removing

the irrelevant image components, respectively. Furthermore, the related challenges in biomedical and biometric image processing and analysis are highlighted.

2.2 Survey of Random Noise Cancellation Methods

Random noise, denoting the random distortions which occur in pixel intensities, is routinely confronted when images are being acquired and transmitted. Denoising, as the process of suppressing the effect of such random distortions and preserving the image structures and details, is conducted after random noise contamination occurs. To design a new random noise cancellation method, the technical challenges are not only to smooth the degraded images, but also to preserve the image features as accurately as possible. As an important research field in the image processing community, in recent years, the various random noise cancellation approaches have been developed and symbolized as the state-of-art techniques. Such state-of-the-art methods can be categorized into the following forms, depending on the diverse denoising mechanisms.

2.2.1 Adaptive Random Noise Filters

Adaptive filtering is the simplest strategy to suppress random noise in degraded images. It relies on the energy distribution of target images in the frequency domain. To be specific, the spectrum of noise-free signal is usually concentrated in the lower and intermediate frequency domain. Whereas, white noise, which keeps a constant power spectral density and follows a specific probability distribution (e.g. Gaussian white noise), is often encountered in the higher frequency domain. Therefore, a filtering conducted in higher frequency domain would be feasible to improve the signal-to-noise ratio (SNR). In the remainder of this section, the most popular adaptive filters are discussed in the following subsections.

2.2.1.1 Wiener Filter

The wiener filter is one of the classical approaches for tackling image restoration problems in historical accounts of image denoising [21]. The observed noisy image \mathbf{y} is given as follows

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (2.1)$$

where \mathbf{x} denotes the noise-free image and \mathbf{n} stands for the additive random noise. To be explicit, the image restoration problem is regarded as the estimation of real \mathbf{x} based on its degraded version \mathbf{y} by implementing $\hat{\mathbf{x}} = \mathbf{y} \circ \mathbf{w}$ (here \mathbf{w} is the discussed Wiener filter and \circ denotes the entrywise product between \mathbf{y} and \mathbf{w}). In order to obtain the Wiener filter, the subsequent minimization problem needs to be solved

$$\hat{\mathbf{w}} = \arg \min \|\mathbf{y} \circ \mathbf{w} - \mathbf{x}\|_2^2 \quad (2.2)$$

The solution of Equation 2.2 is the Wiener filter [22] and its frequency response is given as follows

$$\mathbf{W}(\omega) = \frac{\mathbf{S}_{\mathbf{x}}(\omega)}{\mathbf{S}_{\mathbf{x}}(\omega) + \mathbf{S}_{\mathbf{n}}(\omega)} \quad (2.3)$$

where $\mathbf{S}_{\mathbf{x}}(\omega)$ is the power spectrum for the noise-free image \mathbf{x} and $\mathbf{S}_{\mathbf{n}}(\omega)$ is the power spectrum for the additive random noise \mathbf{n} . Assuming noise-free images can be well approximated by various Gaussian distributions, when the Wiener filter is applied to noisy images in the real-world, the following form in the local image spatial domain is adopted [22]

$$\hat{x} = \frac{\sigma_L^2 - \sigma_n^2}{\sigma_L^2} \cdot (y - \mu_L) + \mu_L \quad (2.4)$$

$$\sigma_n^2 = \frac{\sum \sigma_L^2}{H \cdot W} \quad (2.5)$$

where σ_L^2 and σ_n^2 are the variance of noisy pixel intensities in the local spatial domain L , and the random noise variance, respectively. σ_n^2 can be approximated by using the mean of local noisy pixel intensity variances according to Equation 2.5 (H and W are image height and width, respectively). μ_L denotes the mean of noisy pixel intensities in the local spatial domain. The Wiener filter is performed like a pointwise linear estimator in a local image spatial domain. Therefore, based on the locally computed variance and

mean of pixel intensities within a fixed-size sliding window over images, such a locally adaptive version has been broadly utilized for smoothing the random noise existing in real-world images. Furthermore, this adaptive version of the Wiener filter results in significant improvements, in terms of SNR, MSE and visual quality. The Wiener filter is able to achieve better denoising results under mild noise contamination except when the images are seriously degraded by random noise.

2.2.1.2 Conventional Linear Filters

Without prior knowledge and understanding of the statistics of noise-free images as well as random noise, another two classical linear filters could be used: (i) one is the moving average filter and (ii) the other is the Gaussian filter [23]. For achieving the self-adaptation to local image structures, the following two measures are utilized: (i) the filter spatial size of a moving average filter is dynamically adjusted to the local signal structures; and (ii) the scale parameter (or the filter smoothing parameter) of the Gaussian filter is adaptively determined by the local signal structures. To be explicit, the filter spatial size could be tuned to smaller for preserving edge when image edges turn up; instead, the larger spatial size is selected for smoothing more random noise when encountering flat regions in images. Similarly, the scale parameter of the Gaussian filter is also adaptively tuned according to the variation of the local image structures. Adaptive linear filters can effectively preserve edges but they heavily rely on the self-adjustment module to adapt to image edges.

2.2.1.3 Non-linear Filters

One of the most popular non-linear filters is the median filter [23]. The median filter targets random impulse noise like salt-and-pepper noise, and cancels the noise effect by setting the central pixel intensity via the median value of neighboring pixel intensities. Except median filter, some order-statistics filters are recently developed, not only to cope with impulse noise but also to handle mixed noise (e.g. Gaussian and impulse noise) [25, 26]. Furthermore, the mathematical morphology operator-based non-linear filters are adopted, not only for impulse noise cancelation (e.g. salt-and-pepper noise) but also for other types of random noise suppression (e.g. additive Gaussian noise and spackle noise etc.) [24]. Although the median filter works well for removing random impulse

noise, its applicability is constrained to impulse noise only. Mathematical morphology operator-based non-linear filters can handle various categories of random noise, however, the selection of mathematical morphology operator, structure element type and its size is subjective and user experience-dependent. There is no rigorous methods for an optimal selection for these operators, structure elements and relative parameters.

2.2.1.4 Bilateral Filter

As the combination of spatial domain filtering and pixel intensity range filtering, bilateral filtering is proposed by Tomasi [27]. In the bilateral filter, the spatial and photometric (pixel intensity) distances between two pixels are considered in a separable fashion as follows

$$\mathbf{b}(u_i, u_j, y_i, y_j) = \exp \left\{ \frac{-\|u_i - u_j\|_2^2}{h_{\mathbf{u}}^2} + \frac{-(y_i - y_j)^2}{h_{\mathbf{y}}^2} \right\} \quad (2.6)$$

where $u_i = [u_i^{(x_1)}, u_i^{(x_2)}]^T$ and $u_j = [u_j^{(x_1)}, u_j^{(x_2)}]^T$ stand for the coordinates of pixels u_i and u_j in a horizontal direction x_1 and a vertical direction x_2 , respectively. y_i and y_j denote the intensities of pixels u_i and u_j , respectively. The spatial domain filtering parameter $h_{\mathbf{u}}$ and the pixel intensity range filtering parameter $h_{\mathbf{y}}$ are empirically tuned to regulate the overall filtering performance.

As formulated in the exponent part in Equation 2.6, the similarity between vectors (u_i, y_i) and (u_j, y_j) not only includes weighted Euclidean distance but also involves weighted photometric distance. In addition, the bilateral filter is easy to implement and computationally efficient in adapting to the local image structures. Such a mechanism theoretically guarantees that image edges could be well preserved, however, two parameters manipulating the filtering performance are too difficult to be optimally tuned, and the approaches for optimal parameter selection are not available.

2.2.2 Patch-based Methods

The mechanism of patch-based random noise cancelation is to take advantage of the redundancy of various image structures existing in real-world images. Instead of the point-to-point similarity measuring strategy in the bilateral filter, patch-based methods

exploit the patch-to-patch similarity measuring principle. As one of the most popular patch-based denoising approaches, the non-local mean filter has drawn a great deal of attention in the image denoising community during past years [28–31]. In contrast to the bilateral filter, the following two differences are clearly observed: (i) the pointwise photometric similarity term in the bilateral filter is substituted with the patchwise one in the non-local mean filter; and (ii) the point-to-point geometric similarity computation, based on neighboring pixels in the bilateral filter is replaced by the one based on the patches, which may not be spatially close to the pixel of interest; that is, not neighboring patches but rather non-local ones are taken into consideration for random noise cancelation. To summarize, the non-local mean filter is formulated as follows

$$\mathbf{NLM}(u_i, u_j, p_{u_i}, p_{u_j}) = \exp \left\{ \frac{-\|u_i - u_j\|_2^2}{h_{\mathbf{u}}^2} + \frac{-\|p_{u_i} - p_{u_j}\|_2^2}{h_{\mathbf{p}}^2} \right\} \quad (2.7)$$

where p_{u_i} and p_{u_j} stand for the patches centered at pixels u_i and u_j , respectively. In practice, the following two details need to be noticed in the implementation: (i) the target patch p_{u_i} rather than other patches should be assigned with higher weighting; and (ii) it is computationally impractical to compare all the patches $\{p_{u_j}\}$ to the target patch p_{u_i} , therefore, in practice, the patch search range is restricted to a reasonable spatial neighborhood of the target patch p_{u_i} . In this sense, the non-local mean filter is not entirely “non-local” but more or less “local”, and accordingly its denoising performance could be somewhat affected.

2.2.3 PDE and Variational Methods

The random noise removal via PDE and variational methods is an iterative process, where the pursuit of denoised image is progressively attained. After a number of iterations, the obtained images are regarded as the “noise-free” versions of contaminated images. The anisotropic diffusion filter and the total variation filter, two broadly used partial differential-based random noise cancelation techniques, are introduced, respectively, in the remainder of this section.

2.2.3.1 Anisotropic Diffusion Filter

The goal of diffusion filtering is to constrain the diffusion process to contiguous homogeneous regions instead of crossing region boundaries [32, 33]. As one classical diffusion process, anisotropic diffusion is an instance of the heat equation as a partial differential equation (PDE) and its formula is given as follows

$$\frac{\partial}{\partial t} I(x, y, t) = \text{div} \{c(x, y, t) \nabla I(x, y, t)\} \quad (2.8)$$

where the image under different diffusion degree t is represented by $I(x, y, t)$ and $I(x, y, 0)$ (when $t = 0$) stands for the noisy image as initial condition. $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]^T$ denotes the image spatial gradient operator. $c(x, y, t)$ represents the diffusion conductance to regulate the diffusion process. To be specific, when $c(x, y, t)$ is a constant C , the heat diffusion process is conducted with isotropic and homogenous diffusivity. Such a diffusion process leads to a Gaussian-like denoising result for the given noisy image. Therefore, the boundaries among flat regions are blurred; on the contrary, as suggested by Perona and Malik, a signal-dependent diffusivity $c(x, y, t) = f(|\nabla I(x, y, t)|)$ is able to guide the diffusion process to be anisotropic rather than isotropic. Taking the advantages of anisotropic diffusion conductance, the boundaries such as edges and contours can be well preserved while the random noise existing in homogeneous regions can be effectively smoothed. In practice, the following two types of signal-dependent diffusivity $c(x, y, t)$ are widely used in the anisotropic diffusion filter

$$c(x, y, t) = \exp \left\{ - \left[\frac{|\nabla I(x, y, t)|}{K} \right]^2 \right\} \quad (2.9)$$

$$c(x, y, t) = \frac{1}{1 + \left(\frac{|\nabla I(x, y, t)|}{K} \right)^2} \quad (2.10)$$

where $|\nabla I(x, y, t)|$ is defined as follows

$$|\nabla I(x, y, t)| = \sqrt{\left[\frac{\partial I(x, y, t)}{\partial x} \right]^2 + \left[\frac{\partial I(x, y, t)}{\partial y} \right]^2} \quad (2.11)$$

The intricate issue for adopting anisotropic diffusion-based filtering approaches is the choice of an appropriate diffusivity function. Even though diverse analytical expressions for the diffusivity function have been proposed, the users still have to fix at least two parameters: the diffusion speed and the number of iterations. Accordingly, the selection of such parameters becomes a bottleneck when exploiting anisotropic diffusion-based filters, and the rigorous methodologies to tune the optimal parameters for given noisy images are still unavailable.

2.2.3.2 Total Variation Filter

Total variation (TV) denoising is an approach for random noise reduction to preserve sharp edges in an underlying signal [34, 35]. In contrast to conventional denoising methods, TV-based approaches convert the random noise cancelation problem into a constrained minimization problem. After solving such a constrained minimization problem, the denoising result is obtained. To be explicit, a TV-based image filtering framework is formulated as follows

$$\hat{\mathbf{x}} = \arg \min \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_{TV} \right\} \quad (2.12)$$

where \mathbf{x} and \mathbf{y} stand for the noise-free image and the observed noisy image, respectively. λ is the regularization parameter. $\|\mathbf{x}\|_{TV}$ is the image total variation term based on a spatial gradient operator, which is defined as follows

$$\|\mathbf{x}\|_{TV} = \sum_{x_1, x_2} \sqrt{\left(\frac{\partial \mathbf{x}}{\partial x_1}\right)^2 + \left(\frac{\partial \mathbf{x}}{\partial x_2}\right)^2} \quad (2.13)$$

Equation 2.12 is able to be solved by using numerical computing techniques. Specifically, since Equation 2.12 is actually a convex regularized cost function, some iterative algorithms, based on convex optimization technology, can be used to achieve solution $\hat{\mathbf{x}}$.

During the iteration, the regularization parameter λ plays a dominant role in the random noise cancelation process. In particular, when $\lambda = 0$, the filtering process would not be conducted, and the obtained image is still the same as the given noisy image. When $\lambda \rightarrow \infty$, the total variation term (defined in Equation 2.13) plays a progressively dominant

role, which reinforces the filtering strength and manipulates the resultant image to have a smaller total variation. Consequently, the resultant image could be over-smoothed and the salient structures in image are most likely to be removed, even though the noise effect has been largely suppressed. Hence, the choice of λ is crucial to balance the local structure preservation and random noise removal. Although λ can dynamically decay with the progression of TV filtering, only the customized parameter selection methods, instead of the unified ones, are applicable. Thus, the denoising results would be unsatisfactory if λ is not appropriately tuned.

2.2.4 Transform Domain-based Methods

Transform domain-based methods have been intensively studied in the recent decade [36]. Although there are a large number of variants belonging to this category, the wavelet-based methods are still the dominant ones [37–44]. To be specific, the overall procedure of wavelet-based methods is summarized as follows: (i) in the first place, the noisy image is transformed into multiple sub-bands, depending on various wavelet bases with different orientations and scales. Hence, the coefficients in low-frequency sub-bands represent the essential signal components, whereas the coefficients in high-frequency sub-bands correspond to the detailed signal structures; (ii) then, threshold-based filtering strategies can be applied to selectively adjust such high-frequency sub-band coefficients. For example, according to the predefined threshold (hard threshold) or adaptively adjusted threshold (soft threshold), the high-frequency sub-band coefficients corresponding to random noise are tuned to be zero for noise effect suppression, while the detailed structure-related coefficients are maintained or even enlarged for signal structure preservation; (iii) finally, the denoised image is reconstructed by adopting inverse transformation from the wavelet domain back to the spatial domain. In the remainder of this section, the recently developed state-of-the-art random noise cancelation methods, based on the transform domain, are discussed in the following subsections.

2.2.4.1 BLS-GSM

BLS-GSM proposes an overcomplete wavelet-based denoising framework to prevent the visual distortion caused by the critically sampled wavelet basis. The sub-band coefficient modification is conducted depending on Bayes least square estimation and Gaussian

scale mixture model [45]. To be specific, the BLS-GSM procedure for random noise cancelation is introduced as follows

- *Step 1:* The observed noisy image is transformed into the wavelet domain;
- *Step 2:* The neighborhood in the wavelet domain, which includes adjacent wavelet coefficients crossing contiguous scales, is focused. Subsequently, the Gaussian scale mixture model, through neighboring scales, is applied to model the target neighborhood;
- *Step 3:* Based on the created Gaussian scale mixture model, the centroid of the target neighborhood is estimated according to the Bayesian least squares estimation. Furthermore, the estimated centroid is assigned as the wavelet coefficient centered on the target neighborhood;
- *Step 4:* After all the wavelet coefficients are adjusted, the denoised image is obtained by inversely transforming the modified wavelet sub-bands back to the spatial domain.

In the BLS-GSM scheme, the resultant updated wavelet coefficients not only consider the relation of neighboring coefficients in the same scale sub-band (intrascale correlation), but also involve the relation from the adjacent scale sub-bands (interscale correlation). Taking the benefit of the crossing-scale Gaussian mixture modeling, the robustness for random noise is accordingly enhanced and the computation efficiency is also improved. However, the BLS-GSM algorithm relies on the Gaussian assumption for wavelet coefficients. If the coefficients in transform domain fail to comply with Gaussian-like distribution, neither the denoising performance nor the result can be guaranteed.

2.2.4.2 BM3D

BM3D incorporates the non-local patch grouping and the overcomplete representation to achieve a remarkable denoising performance [46]. The patch grouping is conducted based on patch structure similarity, which is not restricted to the smaller spatial domain but concerns a larger searching area. Furthermore, the overcomplete representation is realized, since the wavelet coefficients in decomposed sub-bands are redundant. To be explicit, the BM3D framework for random noise cancelation is presented as follows

- *Step 1:* The observed noisy image is divided into overlapping patches;
- *Step 2:* The noisy image patches are categorized into diverse clusters, according to the similarity of local structures;
- *Step 3:* A 3D transformation, including 2D wavelet transformation, based on DCT within a cluster and 1D wavelet transformation, based on Haar wavelet across clusters, is implemented;
- *Step 4:* After the 3D transformation of image patch groups, the shrinkage of the transformed spectrum is conducted. To be specific, the following two steps for shrinking the transformed spectrum are employed, respectively: (i) as an initial step, a hard threshold strategy is used; then (ii) an adaptive threshold strategy, derived by Wiener filtering, is adopted;
- *Step 5:* Depending on the modified spectrum in the transformation domain, an inverse 3D transformation is subsequently performed and as a consequence, the denoised image is obtained.

Being different from a conventional NLM filter, BM3D not only involves more reliable patch candidates for restoring the target noisy patch (each cluster could hold a large quantity of image patches with similar detailed structures), but also considers the sparsity between grouped image patches (multi-wavelet sub-bands could provide redundant information for the description of the image local structure). Therefore, BM3D promptly became one of the state-of-the-art denoising methods after it was published.

BM3D is capable of achieving satisfactory filtering results when the noise contamination is not serious. However, when the noise corruption is severe, the denoising performance could dramatically drop due to the difficulty of image patch grouping under serious noise effects. In addition, 2D transformation, based on DCT, could lead to periodic artifacts in denoised results; that is, in contrast to the learned dictionary atoms, the fixed wavelet basis such as the predefined discrete cosine wavelet is not adaptive to extensive image local structures in the real world.

2.2.4.3 LPG-PCA

LPG-PCA explores a new mechanism to create an adaptive basis instead of predefined ones [47]. In the LPG-PCA denoising scheme, the principal component analysis (PCA) is exploited to generate an adaptive basis. Compared with BM3D, their major difference is summarized as follow: the bases used in image transformation are different; that is, in BM3D denoising procedure, the predetermined basis such as DCT basis is adopted. Instead, in the LPG-PCA filtering scheme, the adaptive basis generated, based on the locally data-driven PCA, is exploited. In contrast to the predefined wavelet basis, the adaptive one is more fitting to the locally geometric structures. To be specific, the LPG-PCA scheme for random noise cancelation is given as follows

- *Step 1:* The observed noisy image is divided into overlapping patches;
- *Step 2:* The non-overlapping image patches are grouped into various clusters, depending on the similarity of the patch local structures;
- *Step 3:* For each cluster, the image patches belonging to such a cluster are transformed by PCA. The PCA-based transformation is implemented for every cluster, until all the clusters are processed. As a consequence, the diverse clusters will result in different PCA domains;
- *Step 4:* For each PCA domain (corresponding to its own image patch cluster), the noise effect is suppressed via coefficient shrinkage. The coefficient shrinkage-based noise suppression is conducted for all PCA domains obtained in *Step 3*;
- *Step 5:* Depending on the updated PCA domains, inverse PCA transformations are consequently performed. Finally, the denoised image is gained.

Being different from BM3D, LPG-PCA depends on a PCA-based local data-driven basis. Therefore, the denoising performance of LPG-PCA is better than that of BM3D, not only due to its effectiveness for random noise removal but also because of its superiority for fine structure preservation.

The filtering procedure of LPG-PCA is usually conducted twice, since the second-time repetition is able to further boost the denoising performance, based on the first-round result. However, the filtered result obtained in the first stage may contain errors. In

particular, under the serious noise contamination, the first-round filtering error caused by incorrect image patch grouping is not avoidable. Hence, the error could be propagated into the subsequent filtering stage. As a result, the repeated LPG-PCA denoising procedure, under such a progressive error, could degrade the final filtering result.

2.2.5 Dictionary Learning-based Methods

With the rapid development of the overcomplete dictionary learning algorithm, the signal sparse modeling and representation based on learned dictionary has been intensively explored during the past decade [48–56]. Especially when the signal is corrupted by random noise, the signal sparse representation, based on the learned overcomplete dictionary practically demonstrates its superiority: not only to resist noise effects but also to capture intrinsic signal patterns. Therefore, the random noise cancelation problem in the image denoising community can be effectively addressed by employing dictionary learning-based image denoising methods.

2.2.5.1 K-SVD

K-SVD achieves satisfactory denoising results depending on the learned dictionary, which is adapted to target image underlying structures [49]. Even under random noise conditions, K-SVD procedure is still able to retrieve salient image structures through an iterative and alternative optimization scheme. To be explicit, the alternative optimization process within each iteration is summarized as follows

- **Sparse Coding Step:** given a dictionary (this dictionary could be an initial dictionary like Fourier basis and DCT basis at the beginning of iteration, or an updated dictionary during the iteration), the image patches attached to the target noisy image are projected on to such given dictionary. Then, each patch can be represented by a few atoms from the given dictionary. Therefore, the sparse coding corresponding to each patch is obtained. In order to achieve image patch sparse coding, the following optimization problem needs to be solved

$$\hat{\alpha}_i = \arg \min_{\alpha_i} \left\{ \|\mathbf{p}_i - \mathbf{D}\alpha_i\|_2^2 + \lambda \|\alpha_i\|_0 \right\} \quad (2.14)$$

where \mathbf{p}_i denotes the vectorized image patch cropped from the observed noisy image. \mathbf{D} is the given dictionary. α_i stands for the sparse projection coefficient vector (sparse coding) regarding patch \mathbf{p}_i , which needs to be solved via Equation 2.14. λ is the Lagrange multiplier to keep the trade-off between L_2 , the norm-based data fitting term $\|\mathbf{p}_i - \mathbf{D}\alpha_i\|_2^2$ and L_0 , the norm-based sparsity regulator $\|\alpha_i\|_0$. Several algorithms are available in the public domain to calculate the sparse coding vector α_i , and one of the most popular algorithms is the approximation pursuit method such as matching pursuit (MP) and orthogonal matching pursuit (OMP).

- **Dictionary Update Step:** in this step, the obtained sparse projection coefficient vector $\hat{\alpha}_i$ in **Sparse Coding Step** is used. Under these circumstances, sparse coding $\hat{\alpha}_i$ is fixed and then dictionary \mathbf{D} can be updated to improve the accuracy of the sparse projection for the subsequent iteration. Given a known sparse coding vector $\hat{\alpha}_i$, dictionary \mathbf{D} fitting the locally salient signal structures is upgraded by solving the following optimization problem

$$\hat{\mathbf{D}} = \arg \min_{\mathbf{D}} \left\| \mathbf{P} - \mathbf{D}\hat{\mathbf{A}} \right\|_2^2 \quad (2.15)$$

where dictionary \mathbf{D} consists of K atoms $\mathbf{D} = \{d_j | j = 1, 2, \dots, K\}$. $\mathbf{P} = \{\mathbf{p}_i | i = 1, 2, \dots, N\}$ stands for the image patch set including N patches cropped from the observed noisy image. $\hat{\mathbf{A}} = \{\hat{\alpha}_i | i = 1, 2, \dots, N\}$ denotes the set of sparse coding vectors corresponding to N image patches. However, simultaneously updating the entire dictionary \mathbf{D} is extremely difficult, therefore, single column d_j is upgraded each time. Accordingly, the right-side term in Equation 2.15 is rewritten as follows

$$\left\| \mathbf{P} - \mathbf{D}\hat{\mathbf{A}} \right\|_2^2 = \left\| \mathbf{P} - \sum_{k=1}^K d_k \hat{\alpha}_T^k \right\|_2^2 \quad (2.16)$$

$$= \left\| \left(\mathbf{P} - \sum_{k \neq j} d_k \hat{\alpha}_T^k \right) - d_j \hat{\alpha}_T^j \right\|_2^2 \quad (2.17)$$

$$= \left\| \mathbf{R}_j - d_j \hat{\alpha}_T^j \right\|_2^2 \quad (2.18)$$

where $\hat{\alpha}_T^k$ and $\hat{\alpha}_T^j$ denote the k^{th} and j^{th} rows in $\hat{\mathbf{A}}$, respectively. In \mathbf{R}_j , all of its $K - 1$ terms are assumed to be fixed while the remaining j^{th} term is still

unknown. Furthermore, singular value decomposition (SVD) is utilized to obtain $\mathbf{R}_j = \mathbf{U}_j \mathbf{S}_j \mathbf{V}_j^T$. Then, the first column of \mathbf{U} is assigned to d_j as its solution. The same procedure is applied to the remaining atoms, until all the atoms d_j ($j = 1 \rightarrow K$) are updated, column by column. Finally, the updated dictionary \mathbf{D} is achieved.

Depending on the learned dictionary, the process to project a single noisy patch on to several dictionary atoms is regarded as the approximation process; that is, the target patch is able to be reconstructed by using the selected dictionary atoms meanwhile the random noise existing in the target patch can be effectively removed. Taking advantage of the patch sparse projection, followed by the approximation (or reconstruction), the single denoised patch can be obtained. In addition, all the obtained denoised patches are synthesized into a whole image by the weighting average method.

The K-SVD framework has achieved success in the image random noise cancelation field, however, the following two drawbacks restrict its further applications: (i) the K-SVD scheme involves massive computation during its iteration-based sparse coding and atom update procedure, therefore, it is not practical to adopt K-SVD for huge-size image denoising tasks; and (ii) the dictionary update process within the K-SVD framework is sensitive to the slight variation which occurs on the input image patch set; that is, we assume that the image to be handled is the same, but the input patch sets used for dictionary learning are different (e.g. the strategies to divide the entire image into overlapping patches are different). As a consequence, the atoms of learned dictionaries could be quite different.

2.2.5.2 LSSC

LSSC integrates non-local architecture and simultaneous sparse coding strategy, not only to boost the computational efficiency of dictionary learning but also to diversify the intrinsic signal structures by grouping the image patches with similar patterns [48]. Due to the efficiency of the learning process and atom diversity of learned dictionary, LSSC is regarded as a significant improvement on K-SVD procedure. To be explicit, the LSSC procedure is introduced as follows

- **Patch Clustering Step:** the image patches with similar structures are categorized together as one group for sub-dictionary learning purposes. Within each group, the image patches are probably from the diverse spatial locations, therefore, such a patch grouping process is conducted in a non-local (NL) manner. To be specific, the patch grouping process is implemented by k-means clustering

$$\|\mathbf{c}_i - \mathbf{p}_j\|_2^2 \leq \xi_i \quad (2.19)$$

where \mathbf{p}_j is the image patch vector needed to be grouped. \mathbf{c}_i denotes the centroid vector of the i^{th} patch group (cluster). ξ_i is the error parameter to select the eligible image patches for the i^{th} cluster. When k-means clustering is completed, $\mathbf{S}_i = \{\mathbf{p}_j | j = 1, 2, \dots, N_i\}$ is obtained where totally N_i image patches are selected to form the i^{th} training patch subset \mathbf{S}_i ($i = 1, 2, \dots, K$). The motivations behind patch grouping are: (i) the image patches with similar structures most likely have similar sparse projection coefficients (sparse codings) on an overcomplete basis (learned dictionary). Thus, such patch clustering prior to the sparse coding procedure is able to speed up the learning procedure; and (ii) the sparsity within an individual cluster (similar structure-based patch subset) can be further exploited; that is, being different from the unified sparse coding method over the whole training patch set in K-SVD, LSSC focuses on the specific training subset to update dictionary atoms.

- **Sub-dictionary Learning and Sparse Coding Step:** LSSC can not only retrieve more signal detailed structures but also prevents the retrieved structure atoms from being ruined. Since the obtained image patch clusters in **Patch Clustering Step** are isolated, the interference caused by non-similar signal phenotypes from other clusters is effectively removed. Accordingly, the learned sub-dictionary corresponding to the specific patch subset can provide more specific and structural atoms, meanwhile the resultant sparse codings for the same patch subset are not varied but consistent. The simultaneous learning and sparse coding mechanism can be formulated as follows

$$\left(\hat{\mathbf{D}}_i, \hat{\mathbf{A}}_i\right) = \arg \min_{\mathbf{D}_i, \mathbf{A}_i} \left\{ \|\mathbf{S}_i - \mathbf{D}_i \mathbf{A}_i\|_2^2 + \lambda \left(\frac{\|\mathbf{A}_i\|_{1,2}}{|\beta_i|} \right) \right\} \quad (2.20)$$

where $\hat{\mathbf{A}}_i = \{\alpha_i^{(j)} | j = 1, 2, \dots, N_i\}$ stands for the sparse codings regarding the i^{th} image patch cluster \mathbf{S}_i . In detail, $\alpha_i^{(j)}$ and β_i denote the sparse coding vectors for image patch \mathbf{p}_j belonging to \mathbf{S}_i and cluster centroid \mathbf{c}_i , respectively. The learning and coding procedure is repeated until all the patch clusters ($\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \dots, \mathbf{S}_K$) are handled. Hence, the learned sub-dictionary set $\hat{\mathbf{D}} = \{\hat{\mathbf{D}}_i | i = 1, 2, \dots, K\}$ is achieved.

Based on the learned sub-dictionaries, the denoised image patch is calculated by weighting average in each patch cluster as follows

$$\hat{\mathbf{p}}_j = \sum_{n=1}^{N_i} \mathbf{w}_{n \rightarrow j} \hat{\mathbf{D}}_i \hat{\alpha}_i^{(n)} \quad (2.21)$$

where $\hat{\mathbf{p}}_j$ is the denoised version for noisy patch \mathbf{p}_j . $\hat{\mathbf{D}}_i$ and $\hat{\alpha}_i^{(n)}$ stand for the learned sub-dictionary corresponding to the patch training subset \mathbf{S}_i and the sparse coding vector for patch \mathbf{p}_n ($\mathbf{p}_n \in \mathbf{S}_i$ or even $\mathbf{p}_n = \mathbf{p}_j$), respectively. $\mathbf{w}_{n \rightarrow j}$ denotes the weighting of patch \mathbf{p}_n related to target patch \mathbf{p}_j , and it can be calculated as follows

$$\mathbf{w}_{n \rightarrow j} = \frac{\exp\left(\frac{-\|\hat{\alpha}_i^{(n)} - \hat{\alpha}_i^{(j)}\|_2^2}{h_\alpha^2}\right)}{\sum_{n=1}^{N_i} \exp\left(\frac{-\|\hat{\alpha}_i^{(n)} - \hat{\alpha}_i^{(j)}\|_2^2}{h_\alpha^2}\right)} \quad (2.22)$$

Equation 2.22 provides a weighting calculation approach based on the difference between sparse coding vectors $\hat{\alpha}_i^{(n)}$ and $\hat{\alpha}_i^{(j)}$. Also $\mathbf{w}_{n \rightarrow j}$ is scaled to $[0, 1]$ and $\sum_{n=1}^{N_i} \mathbf{w}_{n \rightarrow j} = 1$. Depending on the denoised patches, the restored image can be obtained by synthesizing all denoised patches.

LSSC slightly outperforms K-SVD in terms of computation efficiency and denoising performance. Besides, LSSC achieves better denoised results in contrast to BM3D, because it is capable of retrieving more detailed image structures under sub-dictionary learning architecture. Even though LSSC has several superiorities when performing image denoising tasks, its framework still results in the unstructured atoms which is also encountered in the K-SVD scheme. Consequently, due to the adverse impact of unstructured atoms, the artifacts in reconstructed images can be visually observed.

2.2.5.3 CSR

CSR further explores the clustering-based sparse representation and establishes a connection between clustering and sparsity [57]. The general framework of CSR is similar to that of LSSC, however, there are still some differences between both schemes: (i) instead of adopting any initial dictionary as the starting point for subsequent dictionary learning (e.g. DCT basis is used as initial dictionary in the K-SVD scheme, and a dictionary learned at the preliminary learning stage is applied as initial dictionary in the LSSC framework), CSR does not rely on initial dictionary. Therefore, CSR is performed in a complete on-line learning mode; and (ii) L_1 norm-based regulator is used in the CSR framework to exploit non-local sparsity, while L_2 norm is employed in the LSSC scheme for the same purpose. In that respect, the sparsity of L_1 norm in CSR can be well guaranteed compared with the L_2 norm in LSSC. To be explicit, the optimization process of CSR is described as follows

$$\left(\hat{\mathbf{D}}_i, \hat{\mathbf{A}}_i \right) = \arg \min_{\mathbf{D}_i, \mathbf{A}_i} \left\{ \|\mathbf{S}_i - \mathbf{D}_i \mathbf{A}_i\|_2^2 + \lambda_1 \|\mathbf{A}_i\|_1 + \lambda_2 \sum_{j=1}^{N_i} \left| \alpha_i^{(j)} - \beta_i \right| \right\} \quad (2.23)$$

where Equation 2.23 is a typical multi-constraint optimization problem, which involves the following variables: $\hat{\mathbf{A}}_i = \left\{ \alpha_i^{(j)} \mid j = 1, 2, \dots, N_i \right\}$ stands for the sparse coding vectors corresponding to i^{th} image patch cluster \mathbf{S}_i . $\alpha_i^{(j)}$ and β_i denote the sparse coding vectors for image patch \mathbf{p}_j ($\mathbf{p}_j \in \mathbf{S}_i$) and cluster centroid, respectively. Similar to the K-SVD optimization process, dictionary \mathbf{D}_i and sparse coding matrix \mathbf{A}_i are alternatively updated by k-means clustering and principle component analysis (PCA) during multi-iteration.

As the modified version of LSSC, CSR still adopts similar architecture to that of LSSC but its computation complexity is significantly reduced; that is, by exploiting double-level sparsity: first-level sparsity on image patch group centroid and second-level sparsity on image patch representation, the process to refine and extract signal detailed structures (dictionary atoms) becomes more effective and efficient. Although the denoising performance of CSR is evidenced by several published works, the unstructured atoms in learned dictionary still exists. Furthermore, because of such unstructured atoms, the visual distortions and artifacts in reconstructed images are also unavoidable.

2.3 Survey of Irrelevant Component Removal Approaches

In the context of imagery, the demarcation of irrelevant components is usually regarded as an ill-defined problem; that is, no unique and objective standard but multiple and subjective criteria are involved [12]. The multiplicity and subjectivity to define irrelevant components have been frequently confronted in object recognition-based image processing and analysis tasks. Targeting the regions or even pixels attached to objects of interest, end-users consider the rest of the regions and pixels as redundant. Because the information expressed by the remaining image components is not related to the target object and is not required for subsequent processing and analysis, the components, apart from the interesting object, are deemed irrelevant.

Irrelevant component removal is actually the complementary process of object detection [23]. To be specific, we assume that the irrelevant components are background, while the object is treated as foreground. The removal of irrelevant components (background) can be realized by successfully positioning and extracting the expected object (foreground). Therefore, the challenge for effective removal of irrelevant image components is converted into the difficulty of object separation from the intricate context of imagery.

In order to address the object separation problem, image segmentation techniques are extensively exploited. Image segmentation is defined as the procedure to partition a whole image into different segments. Consequently, the meaningful segment which interests users is treated as object (foreground), whereas the remaining ones are reckoned as irrelevant components (background) since users pay less attention to them [58–60]. As is the case with segmentation output, the region of interest object manifests the unified or nearly unified attributes in terms of color, texture, geometric structure, and even spatial location; that is, all pixels in the object region should obey certain criteria, however, the same principle is not universally complied with by the pixels from irrelevant component regions.

Depending on whether the spatial relationship of pixels is taken into consideration in image partitioning, the enormous image segmentation techniques can be categorized into the following two streams: (i) spatial point-based approaches; and (ii) spatial neighborhood-dependent methods.

2.3.1 Spatial Point-based Approaches

Spatial point-based approaches, as the name suggests, are techniques which are able to perform without involving a pixel spatial relationship. In other words, neither local spatial interaction nor global spatial relevance is considered when tackling image segmentation tasks. As revealed by copious published works, the following two popular segmentation methods can be classified as spatial point-based ones.

2.3.1.1 Histogram Thresholding Approach

The histogram-based threshold approach is an automatic technique primarily based on the principle: the segmentation of an image can be conducted by delineating peaks, valleys, and shapes in a target image intensity histogram. Histogram thresholding-based algorithms require minimal manual effort to actualize an image segmentation purpose. Therefore, due to its simplicity, histogram thresholding becomes a classical segmentation technique in the image processing community [61–65]. However, during its development and application, it is reported that histogram thresholding can not work well for low-contrast images because the salient peaks in low-contrast image histograms can not be easily detected. Besides, for color images, the multi-dimension color space poses another technical challenge to the histogram thresholding method; that is, based on multi-histogram corresponding to multi-color channels, searching multiple thresholds, then fusing them and finally yielding a segmentation result is a difficult task.

2.3.1.2 Clustering-based Method

The clustering-based method is a typically unsupervised technique to group the image features within a feature space into several clusters. By assuming that each cluster corresponds to a meaningful region in the target image, the image region partitioning is performed by implementing image feature clustering [66–79]. To be explicit, from the viewpoint of clustering, image features are treated as the feature points scattering in the feature space. The feature discussed here could be color, texture, and descriptor on both a pixel and region basis. The essence of the clustering framework is to analyze the scattering feature points and categorize them by adopting a predefined objective function to demarcate meaningful point groupings (also namely, clusters). Furthermore,

the clustering process is terminated, when the feature points within the same cluster possess high affinity while the points from different clusters have low similarity. For tackling image segmentation tasks, clustering-based technologies are extensively utilized owing to their simplicity and ease of implementation. However, as the most important parameter in clustering, the cluster number is still not objectively tuned but subjectively determined. Also, if more image attributes are involved to form the feature vector, then the dimensionality of the feature space could be increased exponentially. Such high-dimensionality feature space also incurs a difficult technical issue for the clustering algorithm, and accordingly, its convergency may not be guaranteed either.

The clustering techniques under review are typically classified into two forms: (i) hard clustering approaches [66–74]; and (ii) soft clustering methods [75–79]. In particular, hard clustering approaches generate unambiguous boundaries for the cluster-dependent feature space division, since each feature point in the feature space has a unique cluster label. In contrast to hard clustering methods, soft clustering algorithms (e.g. fuzzy clustering techniques) result in overlapping clusters with fuzzy boundaries, because each feature point holds several different degrees of membership with respect to diverse clusters. Therefore, clustering-based image segmentation can be realized by adopting either hard clustering approaches or soft ones.

2.3.2 Spatial Neighborhood-dependent Approaches

In contrast to spatial point-based approaches, spatial neighborhood-dependent methods do not tune off a pixel spatial correlation in their frameworks, but heavily rely on image contextual property to realize image partitioning. The following five popular segmentation technologies are grouped into a spatial neighborhood-dependent type.

2.3.2.1 Region Updating Technique

The region updating technique provides an adaptive region updating mechanism depending on the spatial relationship of pixels for segmentation. The region updating segmentation techniques progressively group pixels according to the criteria of compactness and homogeneity in the spatial domain [80–89]. To be explicit, region updating

approaches primarily employ growing, splitting and merging in either an individual way or an integrated framework. Three such elementary modules are summarized as follows

- **Region Growing:** starts from individual pixels or small-size regions (namely, seeds) to iteratively accumulate neighboring pixels according to a certain homogeneity criteria. The process of region growing terminates when the pixels meeting such homogeneity criterion can not be found anymore. As a consequence, the region growing strategy produces spatially connected and homogeneous regions [80–82, 86]. In the engineering application of region growing, the sequential growing procedure, instead of a parallel one, is usually adopted due to the ease of implementation, therefore, the growing process is performed on a pixel-by-pixel basis. Such sequential procedure leads to the following issues: (i) both computational load and memory requirement could be intensively increased; (ii) over-segmented results are probably encountered because of the presence of color shade variation, texture non-uniformity, and illumination variation; and (iii) the segmentation quality is primarily dependent on the processing priority of initialized seeds.
- **Region Splitting:** initiates from an inhomogeneous region, where the specific homogeneity criterion is not satisfied. Then, such an inhomogeneous region is continuously smashed to generate more and more segments until each segment satisfies the specific homogeneity criterion, at which point, the region splitting procedure is terminated. Region splitting can be achieved by various techniques, and quadrature tree decomposition is broadly used as a state-of-the-art method [84, 87, 88]. The region splitting procedure generally results in satisfactory performance when handling simple images with well-defined homogeneous regions. However, under complex circumstances such as multiple inhomogeneous objects, the region splitting algorithm is not able to yield complete objects but produces a quantity of fragmentary regions.
- **Region Merging:** integrates a number of subregions into a whole one. Such an integrated segment contains spatially semantic content and corresponds to an object in an image domain. Region merging is usually performed as a subsequent measure after region growing or splitting [83, 85, 86, 89]. To be specific, for reasonably complex images, region growing and splitting techniques most likely result in over-segmented fragmentary subregions. In order to achieve compact

and uniform segments with respect to spatially meaningful image objects, the region merging scheme is employed as the compensation for over-segmentation. Consequently, the fragmentary subregions are combined together via certain region merging strategies to further facilitate the final segmentation performance.

2.3.2.2 Evolving Contour Model

The evolving contour model transforms the image segmentation problem into a partial differential equation-based (PDE) minimization problem [90–107]. To be explicit, an image object can be effectively segmented via the following strategy: firstly, to locate a sealed contour which circles around the target object; furthermore, to extract the target object which is surrounded by the located contour; as a result, the target object is obtained as the segmentation outcome. Such a sealed contour to distinguish the target object from background can be represented by a parametric curve. Therefore, the procedure to obtain a parametric contour is realized by solving a PDE-based minimization problem; that is, the contour can be deformable and dynamically evolved with the minimization of a PDE-based objective function. According to the differences in the PDE-based objective function, evolving contour models can be categorized as follows

- **Snake Model:** defines a parametric curve $\mathbf{v}(s) = [x(s), y(s)]$ called “snake”, which continuously evolves depending on a formulated objective function $\mathbf{E}(\mathbf{v})$. When $\mathbf{E}(\mathbf{v})$ generates a curve that perfectly fits the contour of the target object, the curve evolution process is terminated and accordingly, such obtained curve is regarded as an “optimal” contour to encircle the target object [90–97]. To be explicit, the objective function $\mathbf{E}(\mathbf{v})$ is formulated as follows

$$\mathbf{E}(\mathbf{v}) = \int_0^1 \{\mathbf{E}_i[\mathbf{v}(s)] + \mathbf{E}_e[\mathbf{v}(s)]\} ds \quad (2.24)$$

$$= \frac{1}{2} \int_0^1 \left[\alpha(s) \left| \frac{d\mathbf{v}(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2\mathbf{v}(s)}{ds^2} \right|^2 \right] ds + \int_0^1 \mathbf{E}_e[\mathbf{v}(s)] ds \quad (2.25)$$

where $\mathbf{E}_i[\mathbf{v}(s)]$ stands for internal energy to manipulate the elasticity and rigidity of the evolving curve simultaneously. Therefore, the minimization of $\mathbf{E}_i[\mathbf{v}(s)]$

regarding the trade-off between elasticity and rigidity, can be attained when parameters $\alpha(s)$ (to control elasticity, represented by a first-order derivative) and $\beta(s)$ (to regulate rigidity, depicted by a second-order derivative) are appropriately tuned. $\mathbf{E}_e[\mathbf{v}(s)]$ denotes external energy to guide the deformable curve approaching towards the real contour of the target object in the image domain. The minimization of $\mathbf{E}_e[\mathbf{v}(s)]$ is reached when the evolving curve is spatially overlapping with the object contour. During the past years, the snake model has achieved success in effectively capturing the contours and boundaries for distinguishing homogeneous regions from background, due to the following superiorities: (i) the snake model is able, automatically and self-adaptively, to seek a global energy minimum; (ii) the external energy term $\mathbf{E}_e[\mathbf{v}(s)]$ can be customized according to user demands; and (iii) the curve evolution, to attract object contours, is theoretically guaranteed to be unbiased. Despite possessing several merits, the snake model still suffers from the following two deficiencies: (i) it is sensitive to the placement of the initial contour, and it may not converge if the initialized contour is far away from the real object boundary; and (ii) it is not good at handling concave object boundaries, in contrast to convex ones.

- **Level Set Model:** transforms curve evolution in a 2D spatial domain into a 3D surface deformation framework [98–103]. In order to capture the desired object contour, the level set model starts at a manually initialized level set function $\Phi_0 = \Phi(x, y, t)|_{t=0}$ (its corresponding zero level set function $\Phi_0 = 0$ is actually a 2D curve in an image domain) at time $t = 0$. Then, as time t elapses, the level set function Φ_t iteratively propagates along an orthogonal direction, depending on its former version Φ_{t-1} at time $t - 1$. The propagation speed is dominated by speed function F , which is a higher dimensional scalar function to depict image characteristics (e.g. grey gradient). Furthermore, the level set model stops working when the following criterion is satisfied

$$\Phi_t + \|\nabla\Phi_t\| F = 0 \quad (2.26)$$

Solving Equation 2.26 can result in the well-formed level set function $\Phi_T = \Phi(x, y, t)|_{t=T}$. Finally, the 2D evolving curve to capture the target object boundary is attained by solving the zero level set function $\Phi_T = 0$ at time $t = T$.

The level set model, due to its stability and irrelevancy with respect to image topology, exhibits prominent technical advantages when coping with image object segmentation tasks. Its merits are summarized as follows: (i) the level set model is robust to topology variation and structure scale change, since its intrinsic 2D zero level set function adapts to region splitting and merging during iteration; and (ii) it is easier to implement, in contrast to the snake model, because of its non-parametric nature. However, one drawback is encountered when partitioning regions without salient contours or boundaries; that is, the stopping criterion, based on the image gradient, fails to be reached at the location of the low gradient contrast contours and boundaries. As a result, the evolving curve can not stop at such weak contours and boundaries but ignores them in its continuing evolution.

- **Mumford-Shah Model:** copes with image segmentation tasks based on a total variation (TV) framework [104–107]. To be specific, the Mumford-Shah model assumes that the entire image domain Ω can be segmented into two homogeneous segments: one is defined as the target object $\Omega_{\mathbf{C}}$, which is encircled by a closed curve \mathbf{C} ; and the other is denoted as $\Omega_{\bar{\mathbf{C}}}$, which represents image background except for the target object $\Omega_{\mathbf{C}}$. The partition of image Ω into such homogeneous domains is performed by searching for dual targets: appropriate approximation of the original image \mathbf{g} (original image is denoted as \mathbf{f}), and the ideal evolving curve \mathbf{C} . In order to obtain \mathbf{g} and \mathbf{C} , the following minimization problem needs to be solved

$$\left(\hat{\mathbf{g}}, \hat{\mathbf{C}}\right) = \arg \min_{\mathbf{g}, \mathbf{C}} \left\{ \alpha \iint_{\Omega} |\mathbf{g} - \mathbf{f}|^2 dx dy + \beta \iint_{\Omega_{\bar{\mathbf{C}}}} |\nabla \mathbf{g}|^2 dx dy + \gamma \oint_{\mathbf{C}} ds \right\} \quad (2.27)$$

where α , β and γ are predefined parameters to regulate the following three terms in Equation 2.27: (i) **image approximation term** - the good image approximation \mathbf{g} is ensured by minimizing the difference between \mathbf{g} and \mathbf{f} over the entire image domain Ω ($\min\{\iint_{\Omega} |\mathbf{g} - \mathbf{f}|^2 dx dy\}$); (ii) **background smoothness term** - the general smoothness over the background domain $\Omega_{\bar{\mathbf{C}}}$ is guaranteed by minimizing TV norm over the same domain ($\min\{\iint_{\Omega_{\bar{\mathbf{C}}}} |\nabla \mathbf{g}|^2 dx dy\}$); and (iii) **curve length term** - the minimal length of the deformable curve \mathbf{C} can be attained by minimizing the curve integral along the deformable curve path ($\min\{\oint_{\mathbf{C}} ds\}$).

It is obvious that the minimum in Equation 2.27 can be achieved when the evolving curve exactly lies on the boundary between the target object $\Omega_{\mathbf{C}}$ and the background $\Omega_{\bar{\mathbf{C}}}$.

In contrast to the snake model and the level set model, the Mumford-Shah model makes use of globally spatial domain information instead of only relying on local contours and boundaries to facilitate the process of image segmentation. Accordingly, better segmentation performance could be expected by employing the Mumford-Shah model rather than the snake or the level set model. In practice, however, the selection of parameters α , β and γ for balancing the three terms in Equation 2.27 becomes a technical issue, and the method to tune such parameters is not yet self-adaptive but still manually experience-dependent.

2.3.2.3 Markov Random Field Model

The Markov random field model establishes a connection between image characteristics and the spatial neighboring relationship and synthesizes both into an integrated model. With the backup of well-established probability theory, the Markov random field (MRF) based image model has been widely used in the community of image segmentation during the past decade [108–114]. Within the MRF framework, the image segmentation problem is transformed into a pixel labeling problem; that is, according to the Bayes inference principle, a label either “foreground” or “background”, is assigned to a pixel. To be explicit, the MRF image segmentation procedure is mathematically denoted as follows

$$\hat{\mathbf{f}}_{MAP} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{v})\} \quad (2.28)$$

Following the maximum posteriori probability (MAP) criterion, Equation 2.28 seeks the most probable pixel labeling $\hat{\mathbf{f}}_{MAP}$. $\mathbf{f} = \{f_1, f_2, f_3, \dots, f_M\}$ stands for the labeling set from the pixel location set $\mathbf{S} = \{1, 2, 3, \dots, M\}$ to pixel tag set $\mathbf{L} = \{l_1, l_2, l_3, \dots, l_M\}$ (M is the total pixel number in the target image). Also, the labeling set \mathbf{f} can be regarded as the mapping from the pixel location set \mathbf{S} to the pixel tag set \mathbf{L} ($\mathbf{f} : \mathbf{S} \rightarrow \mathbf{L}$). For example, $f_i : i \rightarrow foreground$ means pixel i ($i \in \mathbf{S}$) is recognized as “foreground” under mapping f_i , while $f_j : j \rightarrow background$ means pixel j ($j \in \mathbf{S}$ but $j \neq i$) is classified as “background” under mapping f_j . $p(\mathbf{f}|\mathbf{v})$ denotes posterior probability consisting

of the following two variables from diverse fields: (i) \mathbf{f} is from the labeling field; and (ii) $\mathbf{v} = \{v_1, v_2, v_3, \dots, v_M\}$ stands for the pixel-based image feature set (which can be obtained via different image feature extraction methods) in the observation field. Furthermore, posterior probability $p(\mathbf{f}|\mathbf{v})$ can be rewritten according to Bayes rule as follows

$$p(\mathbf{f}|\mathbf{v}) \propto p(\mathbf{v}|\mathbf{f}) \cdot p(\mathbf{f}) \quad (2.29)$$

where $p(\mathbf{v}|\mathbf{f})$ and $p(\mathbf{f})$ are conditional probability (or, say, likelihood) and priori probability, respectively. Conditional probability $p(\mathbf{v}|\mathbf{f})$ characterizes different underlying feature distributions corresponding to various labels, based on training feature sets. To be specific, given the foreground training feature set \mathbf{v}_F and the background training feature set \mathbf{v}_B , conditional probabilities $p(\mathbf{v}_F|\mathbf{f}_F)$ and $p(\mathbf{v}_B|\mathbf{f}_B)$ are able to be modeled by diverse Gaussian distributions, respectively. In contrast, priori probability $p(\mathbf{f})$ is calculated as follows

$$p(\mathbf{f}) = \prod_{i \in \mathbf{S}} p(f_i) \quad (2.30)$$

For pixel i , probability $p(f_i)$ not only depends on pixel i 's location but also relies on its surrounding labels. Mathematically speaking, in order to obtain $p(f_i)$, the following conditional probability needs to be calculated

$$p(f_i) = p(f_i | \{f_{\bar{i}}\}) \quad (2.31)$$

$$= p(f_i | \{f_{N_i}\}) \quad (2.32)$$

where \bar{i} stands for the pixel set in which pixel i is excluded. $\{f_{\bar{i}}\}$ denotes the set of remaining labels except label f_i . $\{f_{N_i}\}$ stands for the labels within the neighborhood of pixel i . Employing $p(f_i | \{f_{N_i}\})$ to approximate $p(f_i | \{f_{\bar{i}}\})$ is guaranteed by the Markov property inside the spatial neighboring region. Furthermore, the conditional probability $p(f_i | \{f_{N_i}\})$ can be calculated as follows

$$p(f_i | \{f_{N_i}\}) = \exp[-U(f_i)] \quad (2.33)$$

where $U(f_i)$ is the energy term with respect to label f_i and it is defined as follows

$$U(f_i) = \sum_{c \in \mathbf{C}} V_c(f_i) \quad (2.34)$$

$$= V_1(f_i) + \sum_{j \in N_i(j \neq i)} V_2(f_i, f_j) \quad (2.35)$$

where $V_c(f_i)$ denotes the clique potential for pixel i over the clique type set \mathbf{C} . Depending on the clique type set \mathbf{C} and the second-order spatial neighborhood system for pixel i , two types of clique potentials: singleton potential $V_1(f_i)$ ($c = 1$) and doubleton potential $V_2(f_i, f_j)$ ($c = 2$) are usually used for calculating the energy term $U(f_i)$.

The MRF model integrates image features (represented by conditional probability $p(\mathbf{v} | \mathbf{f})$) and the spatial neighboring relationship (represented by priori probability $p(\mathbf{f})$) to actualize image segmentation. During the MRF modeling process, several optimization techniques such as iterative conditional modes (ICM), highest confidence first (HCF), and simulated annealing (SA), as well as parameter estimation approaches such as expectation maximization (EM) and maximum likelihood estimation (MLE) are utilized to yield the most probable pixel labels for the target image. Taking account of the benefits from such a synthesized framework, the images with non-deterministic textural regions can be better processed in contrast to other traditional segmentation technologies. Besides, as the extension of the MRF model, another two probabilistic models: the conditional random field model (CRF) [115, 116] and the Gaussian mixture model (GMM) [117–119] are exploited and further improved for handling image segmentation tasks.

2.3.2.4 Graph Theory-based Technique

The graph theory-based technique provides another spatial context description for images under the architecture of graph theory [120–131]. To be explicit, the target image can be depicted by an undirected graph model $\mathbf{G} = (\mathbf{V}, \mathbf{E})$. In such a graph model \mathbf{G} ,

$v_i \in \mathbf{V}$ denotes the node corresponding to a single pixel i in the image. $e_{ij} \in \mathbf{E}$ stands for the edge to connect node v_i and v_j in model \mathbf{G} , which represents the spatial relationship between pixel i and j . Such spatial relationship is quantified by weighting w_{ij} depending on relevance between pixel i 's and j 's attributes (e.g. pixel intensity, or pixel neighbor-based features like texture). Therefore, from the viewpoint of graph theory, the image segmentation problem can be converted into a graph division problem; that is, some edges are removed from the graph model, according to certain criteria. As a consequence, the whole graph model is decomposed into several disconnected subgraphs. Each subgraph corresponds to an image component, which could be either an expected foreground object or irrelevant background content. One principal advantage of graph theory-based segmentation techniques is the flexibility to define various criteria based on human knowledge for guiding the graph decomposition procedure. However, the graph modeling and partitioning process also leads to significant computational complexity. In order to slash high computation loads in practice, the reduction and simplification of graph division criteria is usually adopted for tackling real-life segmentation problems. Accordingly, segmentation results could be obtained more quickly, but at the cost of segmentation accuracy.

2.3.2.5 Watershed Approach

The watershed approach treats the target image as a 3D topographic surface, which involves 2D spatial pixel coordinates versus pixel intensities in the third dimension. According to such topographic surfaces, image segmentation is actualized by seeking watershed lines and then partitioning the whole image, based on the obtained watershed lines [132–135]. To be specific, the basic idea of watershed segmentation methods is introduced as follows: (i) firstly, it is supposed that several holes are punched in every regional minimum; (ii) then, the entire topography is flooded from below by letting water rise up through the punched holes; (iii) when the rising water from different basins is about to merge, a dam is built to prevent the merging; (iv) after that, the rising water will eventually arrive at a stage, where only the tops of the built dams can be observed above the water line; (v) further, continuous watershed lines, corresponding to the observed dam tops, are marked; (vi) finally, the whole target image is divided into several segments, based on such marked watershed lines. Due to their easy implementation and low computational complexity, watershed approaches have been broadly employed to

cope with image segmentation tasks. However, the results achieved by watershed-based techniques are usually over-segmented. Therefore, a post-processing measure such as region merging is indispensable to yield better segmentation results.

2.4 Summary

This chapter presents a literature review regarding the following two fields: (i) random noise cancelation for noisy images; and (ii) irrelevant image component removal. For random noise cancelation techniques, the review commences with state-of-the-art works by summarizing their mechanisms. Furthermore, the advantages and disadvantages of these filtering methods are analyzed. For irrelevant component removal, image segmentation approaches for partitioning foreground objects are concentrated on, since irrelevant component removal can be implemented through the detection of foreground objects yielded by segmentation methods. This review focuses on popular segmentation methods proposed in the recent decade, then their working principles as well as their superiorities / shortcomings are stated.

With the extensive applications of random noise cancelation approaches and irrelevant component removal methods, however, additional concerns arise and also draw scholars' attention. Under real-life circumstances, for example, the following realistic issues are the subject of concern for end users in the biomedical and biometric image processing and analysis community:

- **Can random noise be conspicuously suppressed while image features are effectively preserved, for contributing to cell segmentation in microscopy images?**
- **Can random noise suppression and image feature extraction be conducted simultaneously in an integrated scheme to generate an image descriptor in favor of medical image classification?**
- **Ridge structure identification in latent fingerprint images is a challenging task. Is there any segmentation strategy which is able to partition ridge structure regions in an automated way?**

Unfortunately, existing literatures can not provide sufficient and appropriate solutions to address such problems. In the following chapters, we develop new methodologies to tackle the above-mentioned issues.

Chapter 3

Automatic Restoration for Noisy Microscopy Cell Images

3.1 Introduction

Microscopy imaging allows scientists to study biological cells at the level where a single cell can be observed. It provides an essential tool for examination of both living and fixed cells in cell culture. The screening of modern high-content high-throughput image-based cells helps biologists discover distinctive phenotypes of individual cell populations, which can be used to identify related genes for targeted therapeutic treatment and new drug discovery [136–138]. For example, high-content high-throughput microscopy produces high resolution images for inspection of normal and abnormal cell division. Since cancer is regarded as uncontrolled abnormal cell division, chemotherapies are often implemented to aim at the cells during mitosis. Identification of new genes, highly correlated with the specific cancer during cell mitosis, provides a possible solution for the development of new anti-cancer drugs. Because of the rapid expansion of cellular imaging data, computerized information processing tools such as image analysis and pattern recognition are increasingly in demand [139–143]. Even though microscopy imaging technology has multiple advantages in medical and biological applications, microscopy images are often degraded by noise from readout procedures and image data acquisition systems, devices or equipment [144–146]. As noise is caused by defects in intensity measurement, it is the main factor which adversely affects the quality of microscopy

imaging and the task of automated segmentation. Due to the statistical and cumulative nature of photon detection, noise may be reduced most effectively by increasing the length of the image acquisition time. However, it has been pointed out that this approach is not practical for microscopy imaging because of the problems of live cell movements, photobleaching, and radiation damage [147].

The corruption of microscopy images is caused by several types of noise, such as Gaussian noise (additive), Poisson noise (signal intensity dependent), salt and pepper noise (pixel intensity on and off), and speckle noise (multiplicative) [146]. Such degradation of microscopy images adversely affects subsequent cell image processing and analysis tasks. To be more explicit, under noise contamination, many image pixels are corrupted, leading to a decline in signal-to-noise ratio (SNR) and inaccuracy in automated cell segmentation. In particular, subcellular organelle is not easy to distinguish from noisy images, so the task of restoring microscopy images from noise degradation is one of the most important steps prior to segmentation [140, 148]. Furthermore, the segmentation result of cellular regions is necessary for cell counting [139, 154] and cell morphology analysis [152, 153]. The accuracy of such segmentation-based cell analysis relies heavily on the quality of the microscopy images under study, so the computerized restoration of such images is considered to be the most appropriate tool for the quality improvement of degraded microscopy images.

Computerized microscopy image restoration methods have been the subject of much research during the past decade. One of the simplest approaches to suppress noise in microscopy images is the mean filter, but whilst this removes noise, it blurs image detail. Bilateral filter (BF) and its variants have been developed to preserve cell edges, but their performance in edge preservation depends on the intensity range filter response and such filtering response is easily affected by moderate and high intensity noise [27, 155, 156]. [28, 29] propose an adaptive image restoration scheme for restoring degraded images by comparing the similarity of patches inside the non-local searching region. Although these studies made no assumption about the properties of image signal and noise and their computational complexity is high, they were empirically observed to perform well with additive Gaussian noise. Total variation (TV) filtering, based on the reduction of noisy signal total variation, is another technique which is often utilized in microscopy image restoration [34, 148]. These TV-based noise removal techniques effectively smooth noise while preserving edges. However, they also lead to artifacts such as the blocky effects

especially for low SNR cell images. Partial differential equation-based (PDE) restoration algorithms have also been developed for the noise cancelation of corrupted microscopy cell images, which smooth the noise in flat sub-regions except at strong edges [32, 157]. However, these PDE-based filters over-smooth cell internal micro-structures whose pixel intensities do not vary distinctively. Fuzzy-based (FUZZY) restoration, based on the fuzzy membership function and fuzzy rule, is a well-established technique applied to general image filtering [158–161]. These fuzzy-based noise cancelation approaches perform well to remove the noise and preserve the image details. However, fuzzy-based methods rely heavily on the manual selection of membership function as well as the subjective definition of the fuzzy rule. Therefore, their restoration performance is restricted to the artificial experience.

Even though the development of computerized restoration approaches for corrupted microscopy images has drawn much attention and many methods have been proposed during the past years, the restoration problem still remains, because it is an ill-posed inverse problem, and necessary information required about the degraded microscopy image to reconstruct the original version is inherently imprecise. In practice, we not only focus on the image quality achieved by the restoration method, but we are also concerned about whether the restored images are expected to be beneficial for cell segmentation. Based on this motivation, we present in this chapter a novel Gaussian noise restoration scheme, which integrates trend surface mapping, probability distribution plotting known as quantile-quantile plot, bootstrapping and the Gaussian spatial kernel for restoration of microscopy cell images. Some distinctive advantages of the proposed approach are summarized as follows:

- Brightness trend is one special characteristic in microscopy cell images. Therefore, trend surface mapping, inspired by geostatistics, is employed to model the trend of image brightness (it is regarded as the signal trend according to [162]);
- Q-Q plot is utilized to identify highly corrupted pixels by comparing the pixel intensities versus known probability distribution (here the Gaussian noise model is assumed according to [145], therefore Gaussian distribution is assigned as the known distribution). Furthermore, noise intensity level estimation is achieved by bootstrapping, depending only on the identified corrupted pixels;

- The smoothing parameter of the Gaussian spatial kernel is automatically determined according to the estimation of bootstrapping, which effectively reduces under-smoothing or over-smoothing;
- With the integration of Variance Stability Transform (VST), the proposed multi-staged restoration scheme can be more generalized to deal with Poisson noise, which is another commonly encountered noise in microscopy cell images.

The rest of this chapter is organized as follows: in Section 3.2, we present the details of proposed restoration approach; in Section 3.3, synthetic Gaussian noise, synthetic Poisson noise and real noise experiments are implemented, respectively. In the synthetic noise experiments, the proposed method achieves higher denoising performance compared with other state-of-the-art restoration techniques in terms of peak signal-to-noise ratio (PSNR) and structure similarity (SSIM). In the real noise experiment, the effectiveness of the proposed method is further verified as it contributes to improving the performance of automated segmentation; in Section 3.4, conclusions and remarks are presented.

3.2 Proposed Integration Approach

In this section, we first apply trend surface mapping to model the gradually varied brightness trend, then we apply Q-Q plot to identify corrupted pixels according to the linear pattern between pixel intensities and standard Gaussian distribution. In order to restore these degraded pixels, we apply bootstrapping to estimate the noise intensity level, which is assigned as the smoothing parameter of the Gaussian spatial kernel for image restoration. The following subsections describe the detail of the proposed approach, respectively.

3.2.1 Image Noise Model and Trend-Surface Mapping

Inspired by the multi-component signal decomposition proposed in [162], we propose a decomposable model to describe the degraded image $\mathbf{g}(x, y)$ corrupted by additive noise $\mathbf{n}(x, y)$ as follow

$$\mathbf{g}(x, y) = \mathbf{l}(x, y) + \mathbf{m}(x, y) + \mathbf{n}(x, y) \quad (3.1)$$

where $\mathbf{l}(x, y)$ represents the trend of image brightness (signal trend), and $\mathbf{m}(x, y)$ is the image structure (signal structure).

The original signal consists of two components: signal trend $\mathbf{l}(x, y)$ and signal structure $\mathbf{m}(x, y)$. The purpose of image restoration is to provide as accurate an approximate estimate $\hat{\mathbf{l}}(x, y) + \hat{\mathbf{m}}(x, y)$ to the original signal $\mathbf{l}(x, y) + \mathbf{m}(x, y)$ as is possible. Here $\hat{\mathbf{l}}(x, y)$ and $\hat{\mathbf{m}}(x, y)$ are the estimation for $\mathbf{l}(x, y)$ and $\mathbf{m}(x, y)$, respectively.

The idea here is to use the global trend-based estimation technique, trend surface mapping [163, 164], to generate a smooth approximation $\hat{\mathbf{l}}(x, y)$ for the gradually varied brightness trend $\mathbf{l}(x, y)$. Consequently, the residuals, also known as the noisy image structure $\mathbf{m}(x, y) + \mathbf{n}(x, y)$, can be obtained. Trend surface mapping is a mathematical technique used in environmental sciences for manipulating map data to produce a smooth approximation of a 3D surface, which is a method based on low-order polynomials of spatial coordinates for estimating a regular grid of points from neighbor observations. The suitability of applying trend surface mapping to cell image data can be summarized as follows: (i) most cells (targets) contained in microscopy images are small-sized or medium-sized; and (ii) the brightness trend of microscopy cell images usually varies gradually. Such image circumstances facilitate good mapping of the trend surface equation, based on low-order linear polynomials. Two outputs of the trend surface mapping of a cell image can be represented as a trend map and a residual map, where the trend map can be used to define a geometric model of the 3D shape as a brightness trend $\hat{\mathbf{l}}(x, y)$ and the residual map $\mathbf{m}(x, y) + \mathbf{n}(x, y)$ is intended to be filtered subsequently.

The p -order trend-surface mapping of a pixel of a cell image of size $M \times N$ can be defined as

$$\hat{\mathbf{l}}(x, y) = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + \dots + b_ky^p \quad (3.2)$$

where $x = 1, 2, 3, \dots, M$, $y = 1, 2, 3, \dots, N$ and $k + 1$ ($k = p \cdot (p + 3)/2$) is the number of trend surface mapping coefficients for vector $\mathbf{b} = [b_0, b_1, b_2, b_3, b_4, b_5, \dots, b_k]^T$. The trend surface function will produce a tilted plane surface for $p = 1$ (linear equation), a

simple hill or valley for $p = 2$ (quadratic equation), or a cubic surface for $p = 3$ (cubic equation).

The trend surface mapping coefficients of a degraded image $\mathbf{g}(x, y)$ of size $M \times N$ can be optimally obtained by minimizing the following objective function, using the principle of least squares

$$\sum_{x=1}^M \sum_{y=1}^N [\mathbf{g}(x, y) - \hat{\mathbf{1}}(x, y)]^2 \quad (3.3)$$

The minimization of Equation 3.3 leads to the trend surface equation expressed in matrix notation as

$$\mathbf{X}^T (\mathbf{g} - \mathbf{X}\mathbf{b}) = 0 \quad (3.4)$$

where

$$\mathbf{g} = [\mathbf{g}(1, 1), \mathbf{g}(1, 2), \mathbf{g}(1, 3), \dots, \mathbf{g}(M, N)]^T \quad (3.5)$$

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1^2 & 1 & 1^2 & \dots & 1^p \\ 1 & 1 & 2 & 1^2 & 2 & 2^2 & \dots & 2^p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x & y & x^2 & xy & y^2 & \dots & y^p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & M & N & M^2 & MN & N^2 & \dots & N^p \end{bmatrix} \quad (3.6)$$

If the inverse of $\mathbf{X}^T \mathbf{X}$ exists, the trend surface mapping coefficient vector \mathbf{b} can be obtained as

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{g} \quad (3.7)$$

Therefore, the trend surface mapping is given by

$$\hat{\mathbf{1}} = \mathbf{H}\mathbf{g} \quad (3.8)$$

where $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is the Hessian matrix.

Finally, the trend surface residual $\Delta(x, y)$ can be obtained by

$$\Delta(x, y) = \mathbf{g}(x, y) - \hat{\mathbf{I}}(x, y) \quad (3.9)$$

where $\Delta(x, y) = \mathbf{m}(x, y) + \mathbf{n}(x, y)$.

3.2.2 Identification of Degraded Pixels by Q-Q Plot

In statistics, a Q-Q plot [165] is a technique for determining whether two datasets come from populations with the same type of distribution. It is a graphical display of a quantile-quantile plot of the sample quantiles versus theoretical quantiles. For example, assuming that theoretical distribution follows Gaussian distribution and if the distribution of the dataset also follows Gaussian distribution, the plot should be close to forming a linear pattern. Here “quantile” means the point at which a given fraction (or percentage) of points lies below. That is, the 25% quantile is the point at which 25% percent of the data fall below. The advantages of using a Q-Q plot are that the sample sizes do not need to be equal, and many distribution characteristics can be simultaneously tested. In other words, when there are two data samples, it is often desirable to know whether the assumption of the same distribution is justified. If so, the assumption that two samples are from the same type of distribution is verified. If two samples differ, it is also useful to gain some understanding of the differences. Figure 3.1 shows a Q-Q plot of two Gaussian distributions, and Figure 3.2 shows a Q-Q plot of a Gaussian distribution (\mathbf{X}) versus an exponential distribution (\mathbf{Y}), where the dotted lines are the reference lines (fitted straight lines). The linear pattern shown in Figure 3.1 indicates that the two samples come from the same distribution, while the non-linear pattern between the two samples shown in Figure 3.2 clearly illustrates that they are not from the same distribution family.

The idea here is to construct a Q-Q plot for the histogram of trend surface residual $\Delta(x, y)$, which is contaminated by noise and the noise distribution is assumed as Gaussian, according to [145]. The degraded pixels can be identified via the linear pattern closely matching the reference line in the Q-Q plot, because if two samples come from

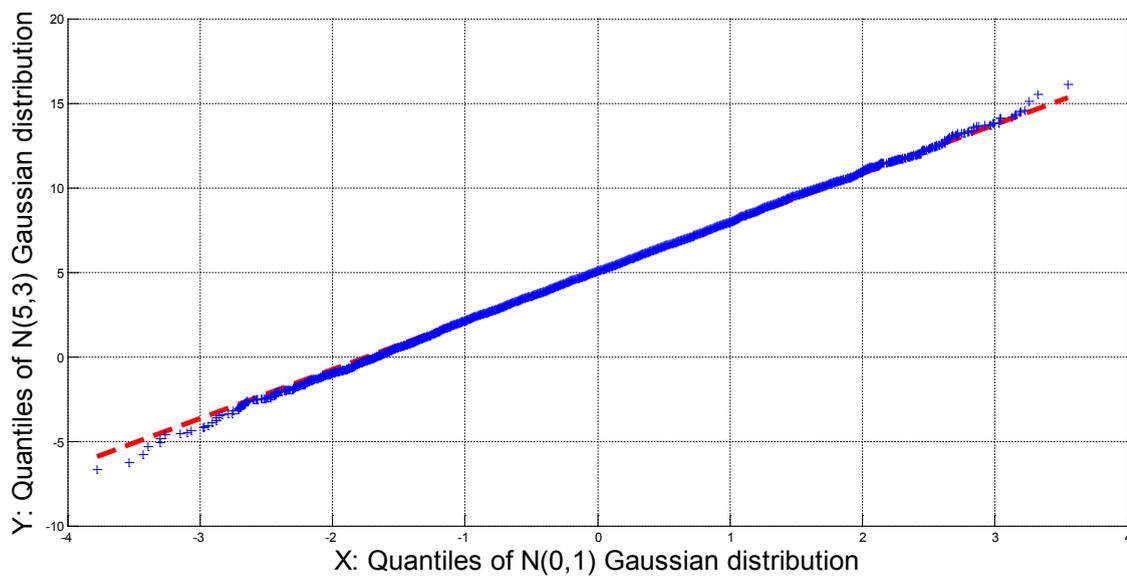


Figure 3.1: Q-Q plot of two samples coming from the same type of distribution (Gaussian) but with different parameters

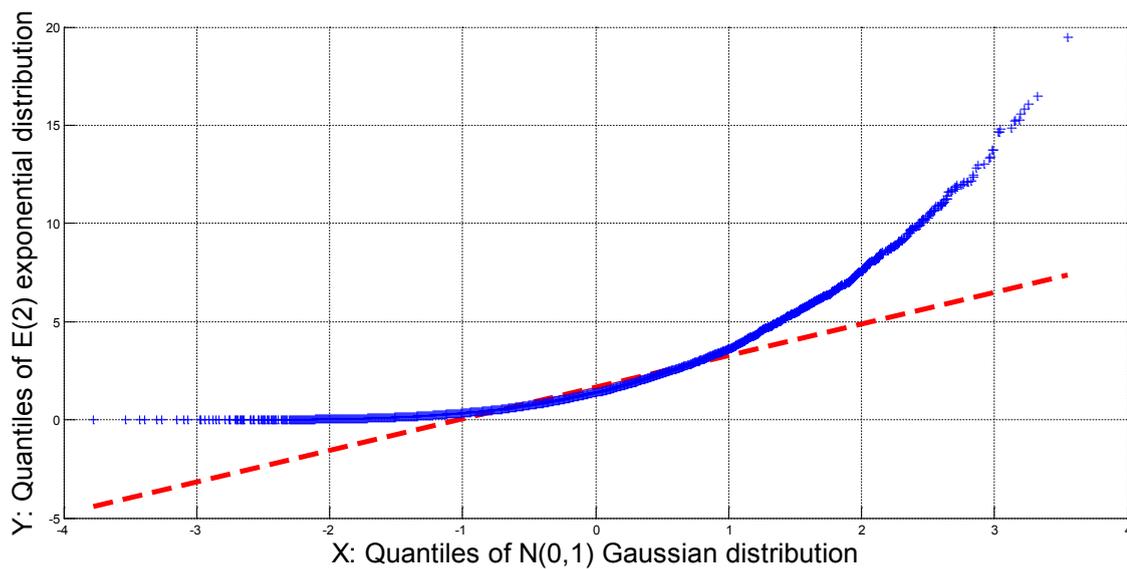


Figure 3.2: Q-Q plot of two samples coming from different distribution families

the same distribution family, the points should appear in a linear pattern and locate closely along the reference line. Otherwise, the greater the departure from the reference line, the greater the evidence demonstrates that the two samples come from different distributions. Therefore, $\Delta_{lp} = \{\Delta(x, y) \in \mathbf{linear\ pattern}\}$ can be recognized as the corrupted pixels. Then the identified pixel set Δ_{lp} is used for the estimation of noise standard deviation by means of the bootstrapping method.

3.2.3 Estimation of Noise Statistics by Bootstrapping

Bootstrapping [166] provides a practical procedure for estimating statistics such as the median, mean and standard deviation by measuring those properties depending on resampling from an approximated distribution. One standard choice for an approximated distribution is the empirical distribution of the observed data. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples with equal size from the observed data. Each resample is obtained by random sampling with replacement from the original observed data. An advantage of bootstrapping is that it provides an easy procedure for estimating the statistics and corresponding confidence interval of complex parameters regarding the empirical distribution of the observed data. In general, bootstrapping can be useful when the following conditions are satisfied: (i) the theoretical distribution of one statistic is unknown; and (ii) the sample size is insufficient for straightforward statistical inference.

In this study, we therefore use bootstrapping to estimate the noise intensity level σ_n according to the corrupted pixels identified by the Q-Q plot. The sample of identified pixel intensities can be regarded as the “only one available sample”, and σ_n is one parameter for such population. In fact, σ_n cannot be known directly because the real noise intensities are unknown in the cell image, therefore we have to estimate $\hat{\sigma}_n$ of the real values, depending on the identified sample from the population. Taking the main idea of bootstrapping into consideration, many samples from the same population can be efficiently replaced by resampling from “only one available sample”. Using mathematical notations, let θ be the mean based on one random bootstrapping sample of size m (m equals total number of pixels in set Δ_{lp}). The estimated noise intensity level $\hat{\sigma}_n$

is obtained via selecting the median value of $(\theta_1, \theta_2, \theta_3, \dots, \theta_z)$ based on z times bootstrapping sampling (each sampling size is m). We use the following formula to calculate different θ_i . Based on the calculated $\{\theta_i | i = 1, 2, 3, \dots, z\}$, the $\hat{\sigma}_n$ can be obtained as follows

$$\theta_i = \frac{1}{m} \sum_{j=1}^m \left| \delta_j^{(i)} \right| \quad (3.10)$$

$$\hat{\sigma}_n = \text{med}(\theta_1, \theta_2, \theta_3, \dots, \theta_z) \quad (3.11)$$

where $\left| \delta_j^{(i)} \right|$ is the absolute difference value regarding the j^{th} pixel in the i^{th} bootstrapping sample.

The estimation of $\hat{\sigma}_n$ by bootstrapping in this study is further discussed in the experimental section. $\hat{\sigma}_n$ will be used to define the mask of a Gaussian filter for the restoration over trend residual map, which is discussed in the subsequent subsection.

3.2.4 Restoration of Degraded Pixels using Gaussian Filter

The Gaussian kernel $h_g(n_1, n_2)$, which produces a 2D Gaussian mask in the image spatial domain, can be defined properly based on appropriate smoothing parameter σ . Here, we assign the estimated $\hat{\sigma}_n$ as the smoothing parameter of the Gaussian kernel within a local spatial domain L of size $N_1 \times N_2$. Then the Gaussian spatial mask $h(n_1, n_2)$ is defined as follows

$$\mathbf{h}_{\mathbf{g}}(n_1, n_2; \hat{\sigma}_n) = \exp\left(-\frac{n_1^2 + n_2^2}{2\hat{\sigma}_n^2}\right) \quad (n_1, n_2) \in L \quad (3.12)$$

$$\mathbf{h}(n_1, n_2; \hat{\sigma}_n) = \frac{\mathbf{h}_{\mathbf{g}}(n_1, n_2; \hat{\sigma}_n)}{\sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \mathbf{h}_{\mathbf{g}}(n_1, n_2; \hat{\sigma}_n)} \quad (n_1, n_2) \in L \quad (3.13)$$

In order to restore the trend residual map $\Delta(x, y)$ in Equation 3.9, the above-defined Gaussian spatial mask is used as follow

$$\hat{\mathbf{m}}(x, y) = \sum_{n_1=-(N_1-1)/2}^{(N_1-1)/2} \sum_{n_2=-(N_2-1)/2}^{(N_2-1)/2} \mathbf{h}(n_1, n_2; \hat{\sigma}_n) \Delta(x - n_1, y - n_2) \quad (3.14)$$

Finally, the restoration result is obtained by adding the restored residual map $\hat{\mathbf{m}}(x, y)$ to the modeled trend surface $\hat{\mathbf{l}}(x, y)$.

The restoration result presented in Equation 3.15 is similar to the adaptive Wiener filter in Equation 3.16.

$$\mathbf{MAR}(x, y) = \hat{\mathbf{l}}(x, y) + \mathbf{h}(x, y; \hat{\sigma}_n) [\mathbf{g}(x, y) - \hat{\mathbf{l}}(x, y)] \quad (3.15)$$

$$\mathbf{WF}(x, y) = \mathbf{m}_f(x, y) + \frac{\sigma_f^2(x, y)}{\sigma_f^2(x, y) + \sigma_n^2} [\mathbf{g}(x, y) - \mathbf{m}_f(x, y)] \quad (3.16)$$

Equation 3.16 assumes that pixel intensity is stationary within the local spatial domain L , then $\mathbf{g}(x, y)$ can be modeled by its local mean $\mathbf{m}_f(x, y)$ and the local structure response $\frac{\sigma_f^2(x, y)}{\sigma_f^2(x, y) + \sigma_n^2}$ (σ_f is the local standard deviation in the spatial domain L). Based on this assumption, the Wiener filter can produce the restored image $\mathbf{WF}(x, y)$ as summarized in [168]. Even though both formulas in Equation 3.15 and Equation 3.16 have similar terms, their differences are still evident and can be stated as follows

- $\hat{\mathbf{l}}(x, y)$ represents the global brightness trend which is more noise-robust than $\mathbf{m}_f(x, y)$, because $\mathbf{m}_f(x, y)$ is calculated based on the local support region and its noise robustness is limited;
- $\mathbf{h}(n_1, n_2; \hat{\sigma}_n)$ involves $\hat{\sigma}_n$, which is not calculated based on the local support region but automatically estimated, depending on the corrupted pixels from the whole image. By comparison with the global estimation, the Wiener filter locally estimates σ_n .

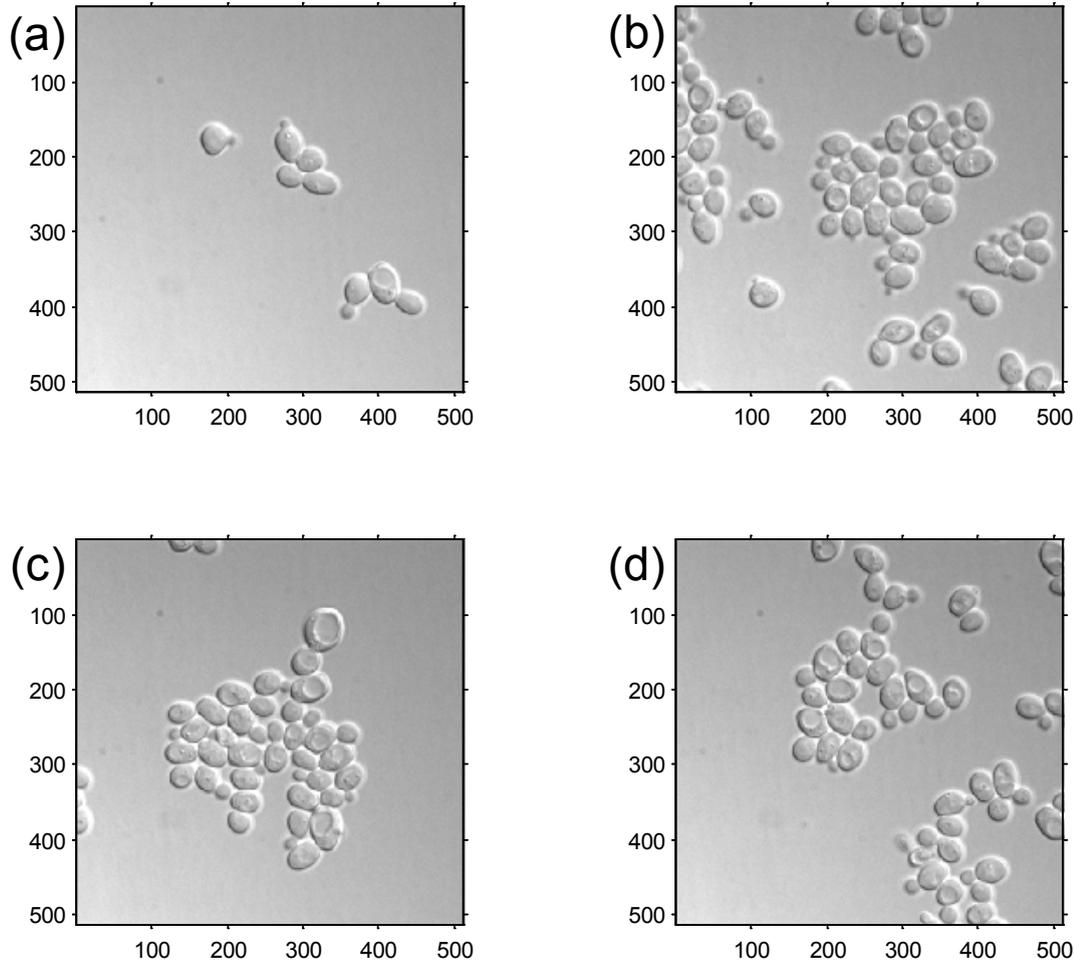


Figure 3.3: Four DIC yeast cell images used in synthetic noise experiment: (a) Cell 1; (b) Cell 2; (c) Cell 3; and (d) Cell 4

3.3 Experimental Results

3.3.1 Experiment 1: Removal of Synthetic Gaussian Noise on Yeast Cell Database

In this study, we tested the proposed restoration method by using DIC yeast cell images provided in [169]. Four datasets are used for the synthetic Gaussian noise experiment and shown in Figure 3.3. Such datasets are largely noiseless and have the following imaging settings: pixel size $0.12758 \mu m$, objective $100\times$, image size 512×512 , and intensity range is scaled to $[0, 255]$.

In the trend surface analysis, the polynomial order p plays a vital role in brightness trend modeling. In order to obtain reasonable trend modeling results, we select the appropriate one via assigning different p over a range of trials. For the quantitative

evaluation of modeling performance and the proper selection of p , mean square error (MSE) is utilized and the criterion of “minimal MSE” is adopted; that is, the less MSE is, the better modeling performance is achieved. $p = 3$ is empirically tuned for brightness trend modeling. The selection process is demonstrated in Table 3.1.

Table 3.1: Selection of polynomial order p according to the criterion of “minimal MSE” for different cell images degraded by synthetic Gaussian noise with varying intensity levels

	$\sigma_n = 10$	$\sigma_n = 20$	$\sigma_n = 30$	$\sigma_n = 40$	$\sigma_n = 50$
Cell 1					
$p = 1$	14.407	22.207	30.614	38.626	46.141
$p = 2$	14.222	22.098	30.549	38.588	46.110
$p = 3$	13.399	21.586	30.189	38.322	45.914
$p = 4$	16.533	23.657	31.703	39.515	46.903
$p = 5$	39.388	42.855	47.871	53.277	58.896
$p = 6$	56.065	58.565	62.340	66.501	70.958
Cell 2					
$p = 1$	20.526	26.641	34.050	41.710	49.136
$p = 2$	20.434	26.565	34.000	41.677	49.108
$p = 3$	20.011	26.236	33.749	41.482	48.947
$p = 4$	21.961	27.742	34.935	42.450	49.755
$p = 5$	40.092	43.482	48.388	54.117	59.990
$p = 6$	54.993	57.489	61.274	65.861	70.674
Cell 3					
$p = 1$	19.329	25.614	33.318	41.216	48.508
$p = 2$	19.176	25.500	33.238	41.166	48.469
$p = 3$	18.724	25.163	32.983	40.987	48.322
$p = 4$	20.811	26.760	34.204	41.972	49.151
$p = 5$	39.108	42.576	47.690	53.530	59.317
$p = 6$	54.282	56.822	60.769	65.390	70.126
Cell 4					
$p = 1$	22.034	27.720	34.702	42.103	49.361
$p = 2$	21.978	27.676	34.674	42.079	49.348
$p = 3$	21.581	27.354	34.433	41.880	49.195
$p = 4$	23.588	28.961	35.727	42.935	50.111
$p = 5$	42.188	45.350	49.951	55.138	60.969
$p = 6$	56.947	59.305	62.883	66.936	71.754

The bootstrapping in this study is implemented as follows: (i) $z = 100$ samples are repeatedly drawn from the Q-Q plot identified pixel set Δ_{lp} ; (ii) the mean θ is calculated for each bootstrapping sample according to Equation 3.10; (iii) the estimated noise intensity level $\hat{\sigma}_n$ is obtained when selecting the median value of $(\theta_1, \theta_2, \theta_3, \dots, \theta_z)$ according to Equation 3.11; (iv) the $(1 - \alpha)$ confidence interval of the estimated $\hat{\sigma}_n$ is $[\hat{\sigma}_{z(\alpha/2)}, \hat{\sigma}_{z(1-\alpha/2)}]$. For example, when $z = 100$ and $\alpha = 0.05$, 3^{th} and 97^{th} ascending ordered elements are marked as lower and upper bounds of the confidence interval, and 50^{th} ordered element (median value) is selected as $\hat{\sigma}_n$.

Subfigures (a), (c) and (e) in Figure 3.4 show the trend surface and the Q-Q plot regarding Figure 3.3 (a). Then subfigures (b), (d) and (f) in Figure 3.4 show the corrupted cell image by additive Gaussian noise ($\sigma_n = 30$), their corresponding trend surface and

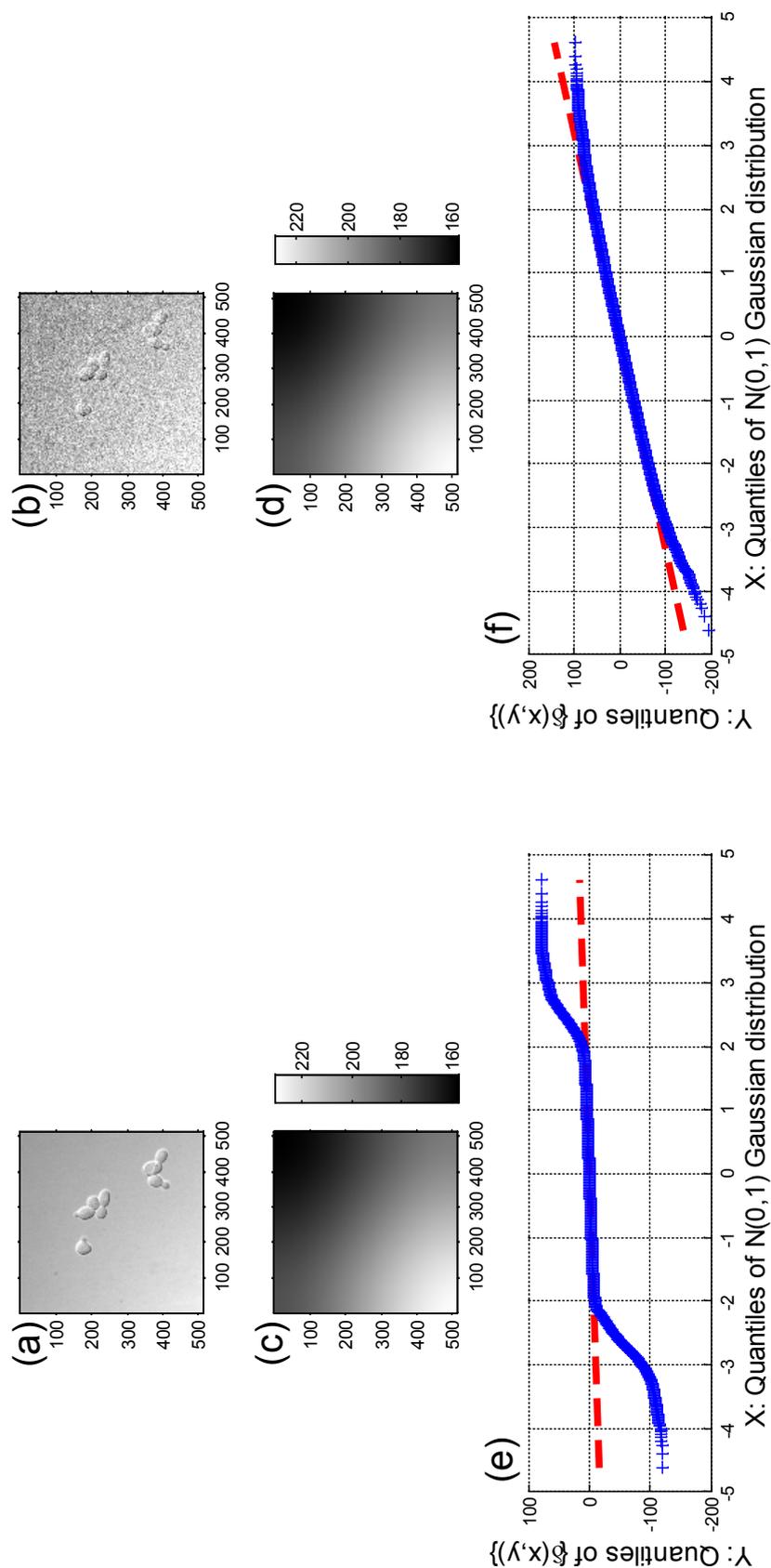


Figure 3.4: Trend surfaces and Q-Q plots of Figure 3.3(a): (a) Original Figure 3.3(a); (b) Corrupted Figure 3.3(a) by synthetic Gaussian noise ($\sigma_n = 30$); (c) and (d) are trend surfaces for (a) and (b); and (e) and (f) are Q-Q plots corresponding to (a) and (b) respectively

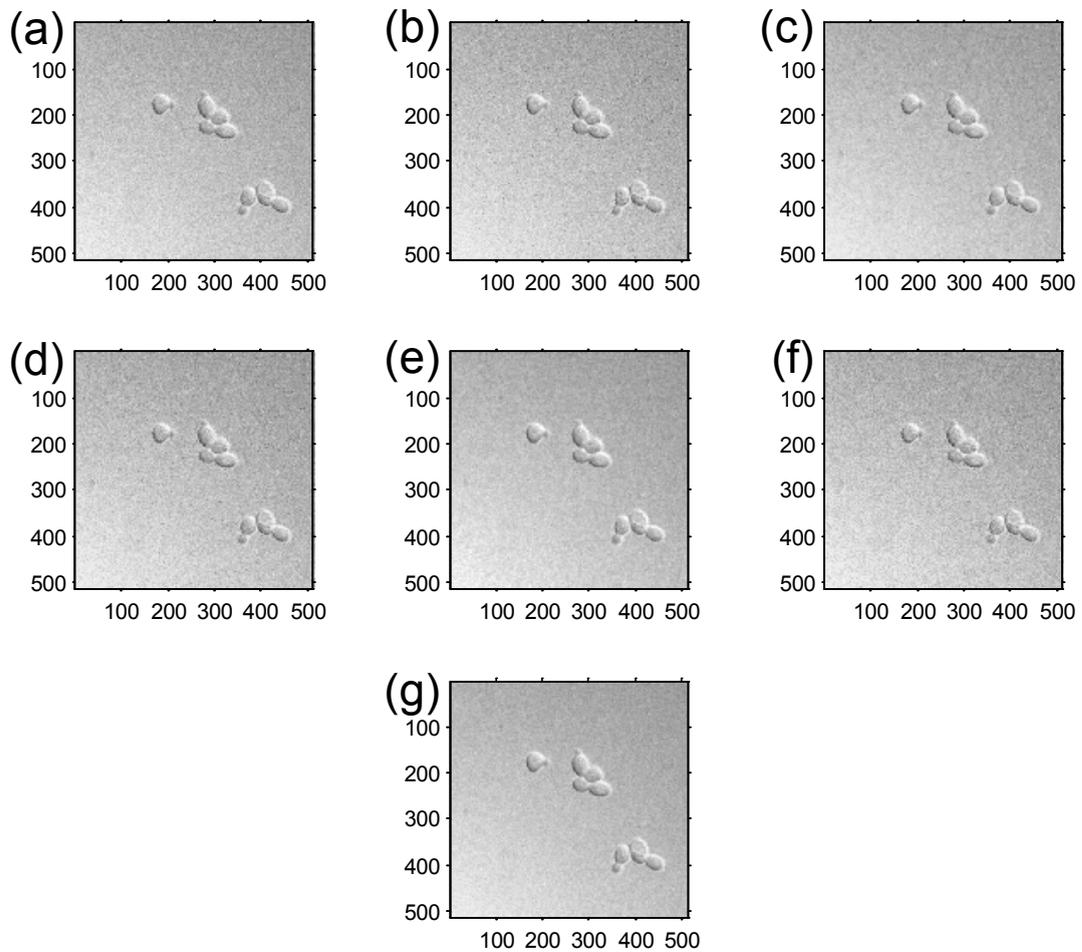


Figure 3.5: Restoration of Figure 3.3 (a) degraded by synthetic Gaussian noise ($\sigma_n = 30$): (a) BF; (b) WF; (c) NLM; (d) TV; (e) PDE; (f) FUZZY; and (g) MAR2

Q-Q plot, respectively. Figure 3.5 shows the restored versions of Figure 3.3 (a) (degraded by synthetic Gaussian noise with $\sigma_n = 30$) by 7 filtering approaches. To be explicit, the degraded images are restored via the following 7 approaches: bilateral filter (BF) [27], adaptive Wiener filter (AWF) [168], non-local mean filter (NLM) [28], total variation filter (TV) [34], non-linear fourth order PDE and relaxed median filter (PDE) [157], fuzzy-based filter (FUZZY) [161], and proposed multi-staged automatic restoration scheme (MAR). All filter parameters are tuned as follows: window size 3×3 , $\sigma_{spatial} = 1$ and $\sigma_{range} = 30$ for BF; window size 3×3 for AWF; searching window size 7×7 , similarity window size 3×3 and smoothing parameter $h = \sigma_n$ for NLM; theta value 16 for TV; iteration time 3 for PDE; window size 5×5 for FUZZY; in order to verify the necessity of brightness trend modeling in the proposed restoration scheme, two different combinations: MAR1 and MAR2 are compared. Here MAR1 stands for “QQ + Boot + GF” while MAR2 represents “Trend + QQ + Boot + GF”, respectively. Obviously,

the procedure of trend surface mapping is ignored in MAR1 in contrast to MAR2. We assign filter mask size 3×3 for MAR1, and mask size 3×3 as well as polynomial order $p = 3$ for MAR2, respectively.

Two restoration quality criteria: peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to evaluate the performance of these 6 approaches. The mathematical definition of PSNR is given by Equations 3.17 and 3.18, respectively.

$$MSE = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W [I(i, j) - \hat{I}(i, j)]^2 \quad (3.17)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (3.18)$$

where I is noise-free image with height H and width W . \hat{I} denotes restored image. MAX_I stands for maximal pixel intensity of image I . The mathematical definition of SSIM is given by Equation 3.19.

$$SSIM(I, \hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + c_1)(2\sigma_{I\hat{I}} + c_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + c_2)} \quad (3.19)$$

where μ_I and $\mu_{\hat{I}}$ denote average pixel intensities of image I and \hat{I} , respectively. σ_I^2 and $\sigma_{\hat{I}}^2$ stand for pixel intensity variances of image I and \hat{I} , respectively. $\sigma_{I\hat{I}}$ is covariance between I and \hat{I} . $c_1 = (k_1 \cdot L)^2$ and $c_2 = (k_2 \cdot L)^2$ are two parameters to stabilize the division with weak denominator ($k_1 = 0.01$ and $k_2 = 0.03$ are usually tuned). L is pixel intensity range (e.g. for a 8 bit image, its pixel intensity range $L = 2^8 - 1 = 255$).

The definitions of PSNR and SSIM reveal: the higher the PSNR and SSIM, the better the restoration. In particular, the SSIM index was proposed to predict human preferences in evaluating image quality. The introduction of the SSIM index was based on the motivation that although PSNR is among the most popular measures of image quality, it does not account for both intensity variations and geometric distortions. The structural information of an image expressed by the SSIM index is defined as attributes that represent the visual structures of the objects, apart from the mean intensity and contrast. The SSIM index has been reported as a good image quality assessment for a wide variety

of image distortions [170]. Thus, the use of SSIM can be useful for quantifying the overall structure of a restored image which is reconstructed to its original version.

Tables 3.2, 3.3, 3.4 and 3.5 present the comparison of restoration results provided by 7 methods for 4 yeast cell datasets degraded by varying intensities of Gaussian noise ($\sigma_n = 10, 20, 30, 40$ and 50 , respectively). The restoration performance achieved by MAR2 is better than that of MAR1. Besides, MAR2 outputs the best restoration results in terms of PSNR and SSIM at moderate and high noise intensities of $\sigma_n = 20, 30, 40$ and 50 . At lower noise intensity of $\sigma_n = 10$, MAR2 does not achieve the highest PSNR and SSIM but it is still competitive with NLM, TV and FUZZY. In the proposed MAR2 scheme, the brightness trend modeling and the Q-Q plot contribute to the effective identification of corrupted pixels. According to Figure 3.4 (c) and (d), the modeled brightness trends are quite similar to each other before and after noise corruption, due to the advantage of noise robustness provided by trend surface fitting in geostatistics; that is, the involvement of measurement error in geographic surveys probably results in noisy measured data. Trend surface mapping captures global trends of measured data instead of the local patterns, therefore, the adverse effects of noise could be significantly reduced. In the procedure of brightness trend modeling, we test different polynomial orders and find that the use of higher orders (e.g. $p = 4, 5$ and 6) quite often leads to distortion around the image boundary and results in high MSE. This issue has been discussed in [163, 164] and demonstrated in Table 3.1. Furthermore, from top right to bottom left in studied images, the image brightness gradually changes from dark to bright in Figure 3.3 (a), and this gradually varied trend is captured by trend surface mapping in Figure 3.4 (c) and (d), respectively. As evidenced in Figure 3.4 (f), the Q-Q plot reveals that the distribution of noisy pixel intensities could evolve to Gaussian-like distribution due to the effect of synthetic Gaussian noise; that is, the linear pattern along the reference line in the Q-Q plot highlights the corrupted pixels. In noiseless cases, images carry little noise, therefore the linear pattern along the reference line in the Q-Q plot is not obvious (shown in Figure 3.4 (e)), however, for severe noise corruption, the linear pattern could be more distinct because the images suffer much more noise (illustrated in Figure 3.4 (f)). In the proposed MAR2 scheme, only the highly corrupted pixels, identified by linear pattern, are involved in bootstrapping for noise estimation. In order to verify the efficiency of noise level estimation based on the proposed “Trend + QQ + Boot” scheme, synthetic Gaussian noise with varying intensities ($\sigma_n = 10, 20, 30, 40, 50, 60, 70$ and 80 ,

respectively) is added up to four studied cell images. Then the estimation is compared with the ground truth and WF scheme. Figure 3.6 demonstrates the efficiency of the proposed “Trend + QQ + Boot” noise estimation scheme.

Table 3.2: Comparison of restoration results from 8 approaches for cell image 1 degraded by varying intensities of synthetic Gaussian noise

Method	$\sigma_n = 10$		$\sigma_n = 20$		$\sigma_n = 30$		$\sigma_n = 40$		$\sigma_n = 50$	
	PSNR	SSIM								
BF	36.048	0.839	30.854	0.627	27.612	0.456	25.362	0.345	23.679	0.274
WF	35.903	0.839	29.932	0.589	26.358	0.402	24.012	0.294	22.286	0.229
NLM	37.409	0.875	32.625	0.715	29.620	0.569	27.651	0.472	26.213	0.409
TV	37.811	0.902	31.939	0.704	26.841	0.418	23.176	0.239	20.497	0.148
PDE	36.731	0.888	32.554	0.770	29.522	0.644	27.217	0.531	25.396	0.441
FUZZY	37.794	0.883	33.428	0.736	29.341	0.506	25.141	0.271	21.6518	0.139
MAR1	32.231	0.897	30.638	0.794	28.939	0.683	27.331	0.582	25.885	0.503
MAR2	37.545	0.897	33.648	0.795	30.795	0.683	28.591	0.582	26.783	0.504

Table 3.3: Comparison of restoration results from 8 approaches for cell image 2 degraded by varying intensities of synthetic Gaussian noise

Method	$\sigma_n = 10$		$\sigma_n = 20$		$\sigma_n = 30$		$\sigma_n = 40$		$\sigma_n = 50$	
	PSNR	SSIM								
BF	35.386	0.863	30.631	0.680	27.471	0.525	25.231	0.416	23.547	0.341
WF	35.184	0.864	29.741	0.652	26.244	0.477	23.875	0.364	22.162	0.294
NLM	36.260	0.888	31.724	0.745	28.849	0.610	26.871	0.509	25.605	0.448
TV	36.022	0.910	31.262	0.739	26.555	0.483	22.927	0.302	20.311	0.206
PDE	34.762	0.894	31.373	0.789	28.824	0.678	26.681	0.575	25.076	0.491
FUZZY	36.296	0.892	32.111	0.768	28.845	0.591	25.467	0.381	22.341	0.223
MAR1	31.715	0.905	30.203	0.817	28.501	0.712	26.977	0.619	25.593	0.543
MAR2	35.390	0.906	32.486	0.818	29.947	0.712	27.985	0.619	26.333	0.543

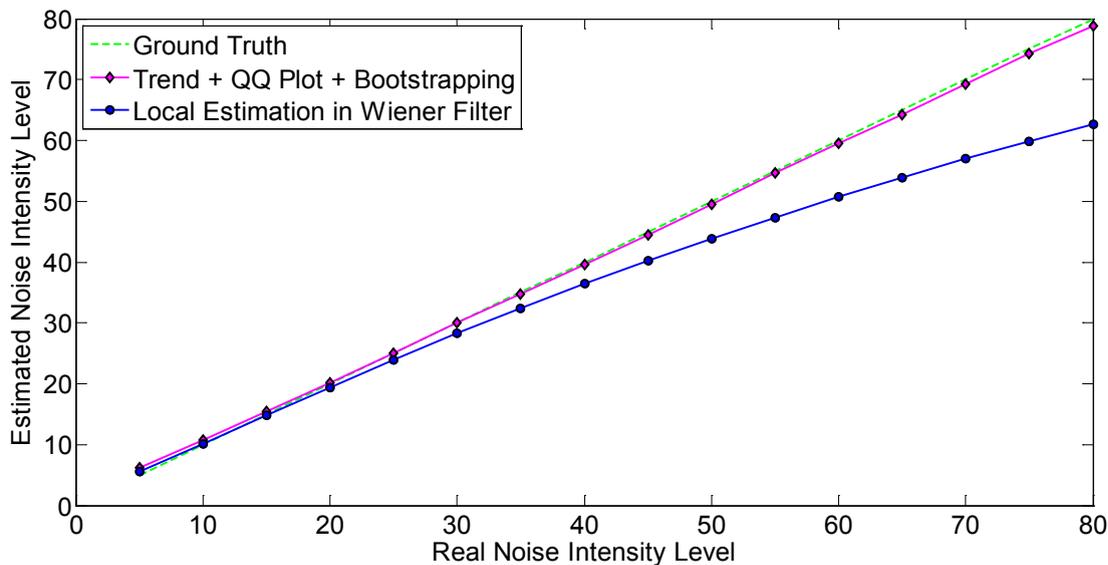
Table 3.4: Comparison of restoration results from 8 approaches for cell image 3 degraded by varying intensities of synthetic Gaussian noise

Method	$\sigma_n = 10$		$\sigma_n = 20$		$\sigma_n = 30$		$\sigma_n = 40$		$\sigma_n = 50$	
	PSNR	SSIM								
BF	35.319	0.856	30.587	0.666	27.435	0.506	25.214	0.398	23.522	0.324
WF	35.222	0.858	29.667	0.635	26.196	0.455	23.852	0.345	22.142	0.278
NLM	36.404	0.884	31.806	0.735	28.900	0.596	27.008	0.496	25.560	0.430
TV	36.110	0.907	31.262	0.728	26.540	0.463	22.943	0.287	20.285	0.192
PDE	34.939	0.891	31.463	0.784	28.813	0.666	26.795	0.564	25.070	0.476
FUZZY	36.616	0.891	32.427	0.768	28.932	0.575	25.563	0.366	22.458	0.211
MAR1	31.818	0.902	30.214	0.807	28.582	0.703	26.990	0.606	25.597	0.525
MAR2	35.475	0.902	32.476	0.808	30.016	0.703	27.965	0.606	26.329	0.525

Table 3.6 illustrates a comparison of computation time taken by 7 restoration methods for denoising a single image whose size is 512×512 . According to Table 3.6, the ranking of different filters in an increasing order is sorted as follows (1 and 7 represent the least and most processing time, respectively): (1) WF, (2) TV, (3) PDE, (4) MAR2, (5) BF, (6) FUZZY and (7) NLM. Both WF and TV require little time for image denoising and can be regarded as the fastest filters in this experiment. PDE needs to iteratively

Table 3.5: Comparison of restoration results from 8 approaches for cell image 4 degraded by varying intensities of synthetic Gaussian noise

Method	$\sigma_n = 10$		$\sigma_n = 20$		$\sigma_n = 30$		$\sigma_n = 40$		$\sigma_n = 50$	
	PSNR	SSIM								
BF	35.502	0.874	30.652	0.701	27.477	0.551	25.276	0.445	23.543	0.368
WF	35.183	0.874	29.705	0.671	26.237	0.501	23.909	0.392	22.126	0.316
NLM	36.007	0.892	31.451	0.753	28.704	0.626	26.860	0.534	25.480	0.467
TV	35.934	0.914	31.203	0.753	26.573	0.508	23.025	0.332	20.351	0.226
PDE	34.739	0.897	31.285	0.797	28.764	0.691	26.787	0.598	25.050	0.512
FUZZY	35.950	0.893	31.757	0.771	28.231	0.579	25.129	0.382	22.019	0.223
MAR1	31.727	0.909	30.169	0.826	28.480	0.729	26.924	0.637	25.573	0.563
MAR2	35.573	0.910	32.487	0.826	29.975	0.729	27.920	0.637	26.314	0.563

**Figure 3.6:** The comparison of noise level estimation between MAR2 and WF for Figure 3.3(a)

update pixel intensities while MAR2 integrates three different approaches to work as a whole. Both could be comparably slower. BF locally detects the edges depending on the intensity range kernel sliding over the whole image domain, thus, more processing time has to be used. FUZZY calculates the local difference by sliding window and obtains the pixel-wise uncertainty, therefore, the running-time is long. The computational time of NLM is significantly higher than other techniques, and this filter would take even more time for image restoration if the searching or similarity window size is enlarged.

Table 3.6: Comparison of computational time (seconds) for single cell image with the size of 512×512

Method	BF	WF	NLM	TV	PDE	FUZZY	MAR2
Run Time (seconds)	4.12	0.03	58.38	0.22	1.16	8.94	2.02

3.3.2 Experiment 2: Removal Synthetic Poisson Noise on Yeast Cell Database

This experiment extends the proposed approach to remove the Poisson noise, which is easily generated in the microscopy system [146, 149]. Unlike Gaussian-distributed noise, removing Poisson-distributed noise in microscopy images is not an easy task, due to the parameter of Poisson distribution being a function of the underlying image pixel intensity. Consequently, the proposed approach cannot be utilized in a straightforward way to handle Poisson noise. However, a transformation-based strategy is adopted to cope with Poisson noise by integrating Variance Stability Transform (VST) with the proposed multi-staged restoration scheme; that is, VST is applied to the input image, and then, through the VST transformation, the Poisson distributed noise is converted to an approximately Gaussian distributed noise. According to [151, 202], the VST is mathematically denoted as follows

$$S = 2\sqrt{Z + 3/8} \quad (3.20)$$

where S and Z are the transformed image (after VST) and the input image (before VST), respectively. After using VST, our proposed algorithm can be used to denoise the transformed image with the same multi-staged restoration scheme proposed in the manuscript, and the denoised result can be transformed back to the restored image by using an inverse transformation (iVST), denoted as follows

$$\hat{Z} = (1/4)\hat{S}^2 - (3/8) \quad (3.21)$$

where \hat{Z} represents the final restored image (after iVST), and \hat{S} represents the result obtained by the proposed multi-staged restoration scheme (before iVST). This integrated strategy (VST + MAR2 + iVST) is easy to implement. Similarly, we also apply the same integrated strategy to BF, WF, NLM, TV, PDE and FUZZY for Poisson noise removal, respectively. All filter parameters are maintained in the same way as those in experiment 1. Table 3.7 demonstrates that the proposed method outperforms other methods, in terms of PSNR and SSIM.

Table 3.7: Comparison of restoration results from 7 approaches for cell images 1-4 degraded by synthetic Poisson noise

Image	Noisy image		BF		WF		NLM	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
cell 1	25.353	0.302	33.689	0.756	33.251	0.744	35.127	0.809
cell 2	25.624	0.397	33.470	0.800	33.028	0.795	33.647	0.853
cell 3	25.618	0.378	33.458	0.792	33.038	0.786	33.618	0.847
cell 4	25.602	0.439	33.527	0.815	32.982	0.807	33.602	0.842
Image	TV		PDE		FUZZY		MAR2	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
cell 1	35.601	0.847	34.961	0.845	35.651	0.821	35.874	0.860
cell 2	34.293	0.869	33.485	0.860	34.227	0.869	34.331	0.877
cell 3	34.271	0.865	33.687	0.857	34.276	0.871	34.376	0.873
cell 4	34.365	0.875	33.471	0.864	34.269	0.870	34.431	0.884

3.3.3 Experiment 3: Segmentation for Real Noisy Phagocytosing Budded Yeast Cell Database

In this experiment, we use a real noisy DIC microscopy cell dataset provided in [171] to verify whether the proposed restoration scheme would improve the performance of cell segmentation. The used phagocytosing budded yeast cell database is a time series whose imaging settings are: image quantity 182, pixel size $0.14 \mu m$, NA DIC objective 63×1.4 , time spacing 3.94 seconds, image size 154×126 and intensity range is scaled to $[0, 255]$. The above 7 restoration approaches: BF, WF, NLM, TV, PDE, FUZZY and MAR2 are applied to restore this real noisy dataset, respectively, and subsequent cell segmentation is implemented based on the obtained restoration results.

The segmentation method used in this experiment is one of the typical morphological filtering approaches [154]. By combining different morphological operators such as erosion, dilation, opening and closing, the cellular regions could be effectively marked and segmented [172]. Here “disk” structure element is assigned to implement morphological filtering. In order to obtain correctly segmented cellular regions, we select the appropriate element size based on the segmentation performance obtained over a range of trials. For the quantitative evaluation of segmentation results, ground truth is given by manual segmentation. Here true positive rate (TPR: sensitivity) is defined as the percentage of pixels belonging to cellular regions which are correctly segmented from the background, and false positive rate (FPR: specificity) is defined as the percentage of pixels belonging to the background but wrongly recognized as cellular pixels. The proper selection of structure element size is demonstrated in Table 3.9, where the criterion of “the higher TPR and the lower FPR” is adopted. For the images denoised by different restoration

Table 3.8: Statistical comparison for segmentation performance (evaluated by average mean ROC attributes: \overline{AUC} , \overline{TPR} and \overline{FPR}) based on 7 filtering approaches

	Original	BF	WF	NLM
\overline{AUC}	0.8433	0.9555	0.8238	0.7453
\overline{TPR}	0.6932	0.9188	0.9033	0.6947
\overline{FPR}	0.0065	0.0067	0.6647	0.1141
	TV	PDE	FUZZY	MAR2
\overline{AUC}	0.8467	0.9194	0.9670	0.9860
\overline{TPR}	0.9774	0.9711	0.9416	0.9803
\overline{FPR}	0.6948	0.1324	0.0051	0.0083

methods, the appropriate structure element sizes are used and the corresponding segmentation results are shown in Figure 3.7. In general, the segmentation result based on MAR2 demonstrates the correctly segmented cellular regions in contrast to the other 6 methods. The segmentation result dependent on FUZZY is competitive with that of MAR2, even though small areas inside the cell are missing (shown in Figure 3.7 (p)). The segmentation results based on BF and WF appear in the incomplete cellular regions (shown in Figure 3.7 (k) and (l), respectively). For WF-, NLM-, TV- and PDE-based segmentation, the wrongly recognized pixels can be clearly observed in Figure 3.7 (l), (m), (n) and (o), respectively. The direct segmentation based on real noisy data produces a poor result. Because a large number of cellular pixels are falsely segmented as background ones, they are missing in the binary result. To yield statistics for filtering-based segmentation evaluation metrics, 182 yeast cell images are used for testing. Table 3.8 shows the resulting statistics of ROC attributes.

3.4 Conclusion

Segmentation-based image analysis for the individual cell from microscopy images is promising. However, this task encounters a number of challenges [173–175]: (i) the intensities of the fluorophores and their spatial distribution may change over time during an experiment, leading to spot stain in each image; (ii) the images suffer low signal-to-noise ratio (SNR), because the laser light intensity needed to excite fluorescent probes is often reduced to as a low level as possible to avoid cell damage; (iii) the cells are often touching and occluded, when cells are aggregated and interact with each other; and (iv) the cell morphology change could continuously happen during the cell cycle, and consequently, the cell morphology may be diverse at the different cell phases. Therefore,

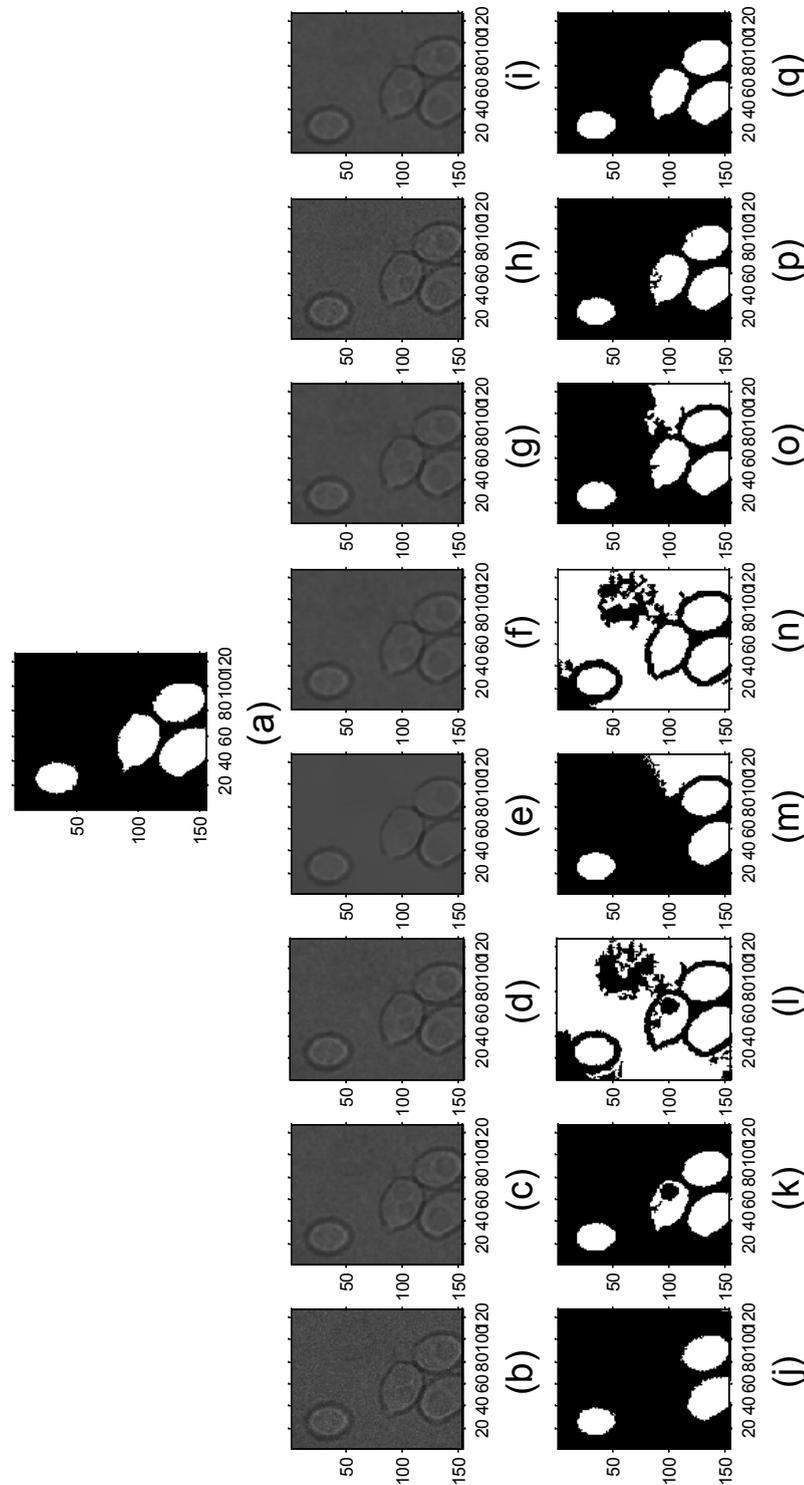


Figure 3.7: Segmentation results given by morphological filtering based on real noisy cell image and restored images: (a) ground truth obtained by manual segmentation; (b) real noisy image; (c)-(i) restoration results given by 7 approaches: BF, WF, NLM, TV, PDE, FUZZY and MAR2, respectively; (j)-(q) corresponding binary segmentation results, where the optimal size of “disk” structure element r is assigned respectively as follows: (j) $r = 3$, (k) $r = 6$, (l) $r = 6$, (m) $r = 10$, (n) $r = 7$ or 8 , (o) $r = 7$, (p) $r = 4$ and (q) $r = 8$

the influence of noise can cause difficulty in automated cell segmentation and feature extraction of cell images such as cell counting, texture representation and morphological characteristic extraction.

Subcellular organelle segmentation is another promising direction with the high-resolution microscopy [176, 177]. The subcellular components and organelle structures can be clearly observed and well investigated via high-resolution microscopy. In contrast to lower resolution microscopy, however, higher resolution microscopy images are easily corrupted by higher intensity noise, because of the high-resolution light microscopy approaching the resolution limit [146]. Prior to subcellular organelle segmentation, the following assumptions need to be made:

- The noise corruption needs to be assumed as mild and moderate. Otherwise, the restored image would contain lots of noise once serious noise corruption happened. As a consequence, the subcellular organelle segmentation, based on such not well restored images, would result in a large number of missing or false subcellular organelle;
- For the real-world images generated from microscopy, the noise needs to be assumed as Gaussian-distributed. If so, the proposed method would be confident of removing the Gaussian-distributed noise. Otherwise, the proposed algorithm could not be applied directly. For the assumption of Poisson noise corruption, VST could be integrated with the proposed algorithm to effectively remove Poisson noise.

We have discussed a novel integration of trend surface mapping, a Q-Q plot, bootstrapping, and Gaussian spatial kernel for removing Gaussian-like noise in microscopy cell images. The integrated approach performs as an adaptive image restoration which achieves better results compared with some other state-of-the-art restoration methods. One advantage of the proposed approach is that its methodology is flexible and therefore can be extended to handle Poisson noise. Besides, the potential improvement is the appropriate selection of noisy pixel candidates according to the linear pattern along a reference line in a Q-Q plot. Such selection could improve the reliability of noise estimation by bootstrapping. Above all, the proposed multi-staged automatic restoration method provides a superior denoising scheme, and could be a highly valuable step prior to cell segmentation for numerous generated and widely used microscopy cell images.

Table 3.9: The selection of optimal size of “disk” structure element according to the criterion of “higher TPR and lower FPR”, the term in brackets stands for (optimal structure element size)

r	Original($r = 3$)		BF($r = 6$)		WF($r = 6$)		NLM($r = 10$)		TV($r = 7ors$)		PDE($r = 7$)		FUZZY($r = 4$)		MAR2($r = 8$)	
	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
1	0.0218	0.0195	0.0952	0.0729	0.0765	0.0791	0.0195	0.0060	0.0683	0.0432	0.0899	0.0756	0.0580	0.0492	0.0862	0.0717
2	0.1580	0.0013	0.0108	0.0308	0.0184	0.0317	0.0001	0.0026	0.0050	0.0288	0.0313	0.0529	0	0.0083	0.0181	0.0573
3	0.6932	0.0065	0.1640	0.0070	0.1630	0.0132	0.4143	0.0017	0.0000	0.0187	0.1580	0.0309	0.3943	0.0067	0.1627	0.0455
4	0.4519	0.0078	0.4411	0.0035	0.4072	0.0312	0.6333	0.0013	0.4272	0.0010	0.4033	0.0136	0.9484	0.0050	0.5906	0.0475
5	0.4511	0.8691	0.6993	0.0048	0.6982	0.0051	0.6853	0.1140	0.6843	0.0027	0.6564	0.0122	0.7042	0.0080	0.6819	0.0027
6	0	1	0.9159	0.0076	0.9017	0.6596	0.6922	0.1147	0.6961	0.0043	0.8909	0.1297	0.5654	0.8077	0.6966	0.0039
7	0	1	0.8189	0.7979	0.2792	0.7919	0.6979	0.1153	0.9777	0.6927	0.9719	0.1322	0.7386	0.9037	0.9342	0.0071
8	0	1	0.8199	0.7988	0.2792	0.7919	0.6979	0.1153	0.9777	0.6927	0.7027	0.0083	0.7386	0.9037	0.9797	0.0085
9	0	1	0.7098	0.8931	0.4527	0.8887	0.7003	0.1167	0.7295	0.7975	0.4495	0.0056	0.7113	0.9043	0.4518	0.0047
10	0	1	0.7392	0.8978	0.7379	0.8957	0.7024	0.1172	0.7308	0.7980	0.7298	0.7795	0.7113	0.9043	0.7300	0.7985
11	0	1	0.7392	0.8978	0.7379	0.8957	0.4479	0.1149	0.7308	0.7980	0.7324	0.7887	0.7113	0.9043	0.7315	0.7992
12	0	1	0.7397	0.9042	0.7379	0.8957	0.4493	0.1153	0.7319	0.7992	0.7324	0.7887	0	1	0.7318	0.7999
13	0	1	0.7397	0.9042	0.4527	0.8887	0.2839	0.1138	0.5631	0.7967	0.5639	0.7850	0	1	0.5649	0.8202
14	0	1	0.4543	0.8959	0.4535	0.8969	0.2852	0.0050	0.5631	0.8002	0.2813	0.8019	0	1	0.5657	0.8244
15	0	1	0.4543	0.8959	0.7121	0.9049	0.2852	0.0050	0.2797	0.8173	0.2813	0.8019	0	1	0.2804	0.8173
16	0	1	0.4550	0.9002	0.7121	0.9049	0	1	0.2797	0.8173	0.2813	0.8019	0	1	0.2804	0.8173
17	0	1	0.4550	0.9002	0.7121	0.9049	0	1	0.4529	0.8896	0.2813	0.8019	0	1	0.4532	0.8892
18	0	1	0.4550	0.9002	0.7121	0.9049	0	1	0.4529	0.8896	0.2813	0.8019	0	1	0.4532	0.8892
19	0	1	0.7135	0.9086	0.7121	0.9049	0	1	0.4529	0.8896	0.7087	0.8921	0	1	0.4532	0.8892
20	0	1	0.7135	0.9086	0.7121	0.9049	0	1	0.4540	0.8946	0.7087	0.8921	0	1	0.4532	0.8892

Chapter 4

Multi-robust Feature for Medical Image Classification

4.1 Introduction

Medical imaging is a powerful tool for radiologists and surgeons making complex disease diagnoses in clinics [178]. With the extensive application of a medical imaging system, medical images are rapidly generated. In order to efficiently archive such explosively generated image data, medical image classification is performed as an essential step in a picture archive and communication system (PACS) [179]. Therefore, an effective medical image classification technique is in urgent demand to satisfy the requirement of convenient management of massive medical image data [180, 181]. Classifier-based and feature-based medical image classification methods have become the research focus for such a purpose in the recent decade. To be specific, classifier-based medical image classification approaches, such as fuzzy logic [182], support vector machine (SVM) [183], multiple kernel learning [184], hierarchical fusion strategy [185] and subspace ensemble learning [186] are proposed. For feature-based medical image classification techniques, texture analysis [187, 191], global feature [188], enhanced feature extraction [189] and spatial uncertainty feature [190] are introduced, respectively.

Object rotation, scale variation and noise degradation can often be simultaneously encountered in computed tomography (CT) and magnetic resonance imaging (MRI) systems. To be explicit, the medical image rotation issue is not easily avoided. Such an

issue is mainly caused by the following reasons: (i) difference of body postures like face-up and face-down poses when scanning patient bodies via medical imaging systems (shown in Figures 4.1a and 4.1b); and (ii) difficulty of positioning a patient body perfectly against the reference axis of medical scan modalities [221] (shown in Figures 4.1c and 4.1d). Patient body size (e.g. obesity or thinness) is the key factor leading to scale diversity in the generated medical images [222, 223] (shown in Figure 4.2). The mechanisms of noise formation in CT and MRI modality were analyzed further in [224, 225], respectively. Their stated conclusions are that noise effect in medical images is inevitable (shown in Figure 4.3). Therefore, there is a particular need to develop a multi-robustness image feature to cope with the occurrences of image rotation, scale variation and noise corruption encountered in medical image analysis tasks. Although multi-robustness may be achieved via classifier-based techniques, feature-based multiple robustness is our focus. Hence, in the remainder of this chapter, our proposed method is restricted to feature-based medical image classification, unless stated otherwise.

Medical images consist of numerous local structures such as edges, ridges, contours and texture. All of these local structures are regarded as subregions with varying pixel intensities. A local structure is supported by its surrounding neighbors. Accordingly, the detection of a local structure has to involve its corresponding neighboring region. A set of characteristics extracted from local structure neighboring regions can be integrated as the features for image representation [192]. How to make use of local structure for robust medical image feature generation has drawn much attention. Some ideas and works can be referred in the general image analysis field: utilization of local structures to represent the whole image has become very popular in recent years, particularly in general image classification tasks [193–200] [202–211]. These works can be categorized into two streams in the following subsections, based on their robustness when only handling image rotation or simultaneously handling concurrent occurrence of image rotation and scaling.

4.1.1 Existing Works on Robustness Against Image Rotation

Zhang et al. [193] proposed two types of microstructure detector with continuous rotation invariance and used a texon library-based vector quantization approach to represent texture images. Good performance on classifying rotated textures was demonstrated. However, the proposed detector is not robust in the face of noise degradation. Zhu et al.

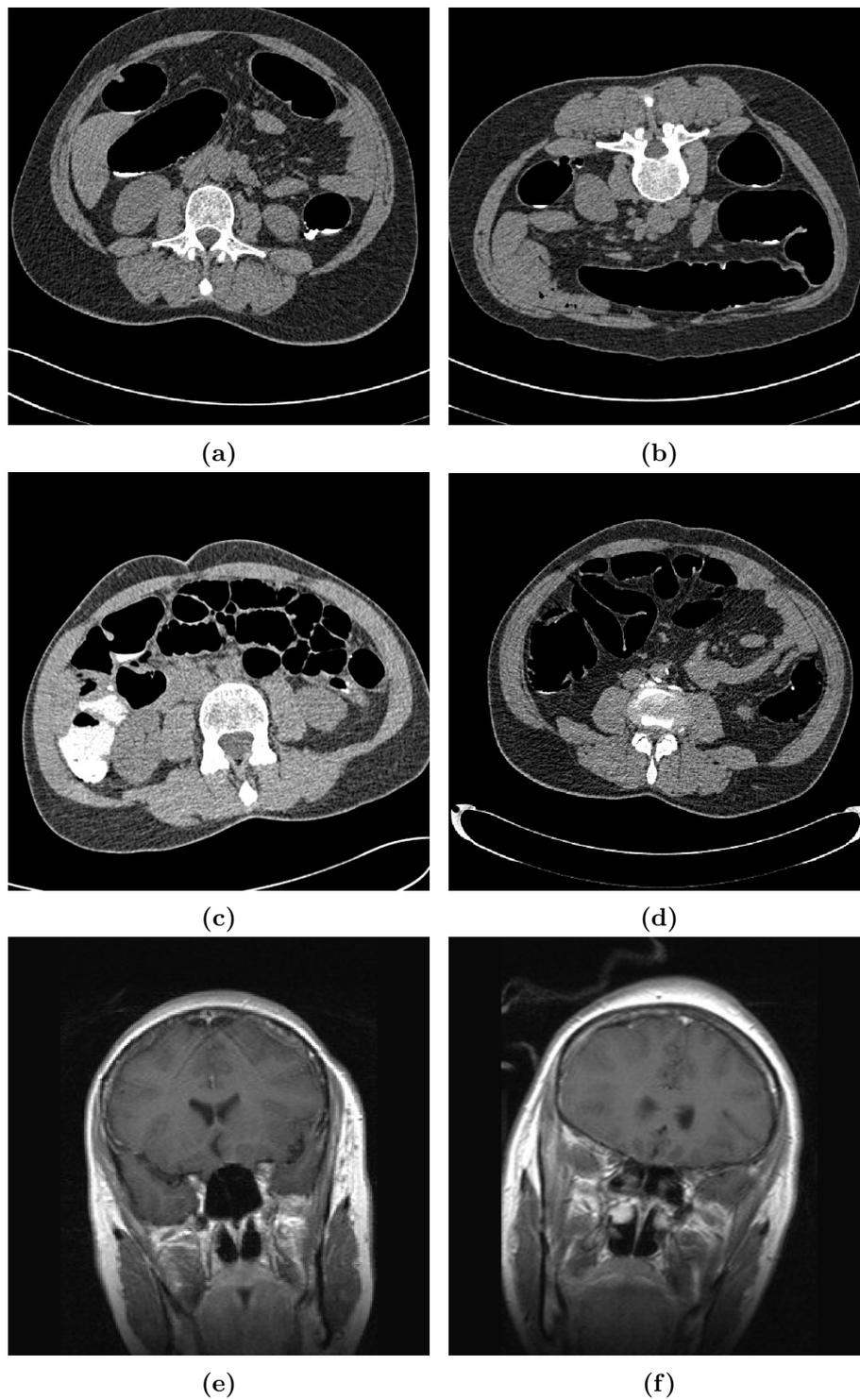


Figure 4.1: The rotation issues encountered in colon CT and head MRI datasets: (a) face-up posture; (b) face-down posture; (c) patient body is inclining to the left; (d) patient body is inclining to the right; (e) patient head is inclining to the left; and (f) patient head is inclining to the right

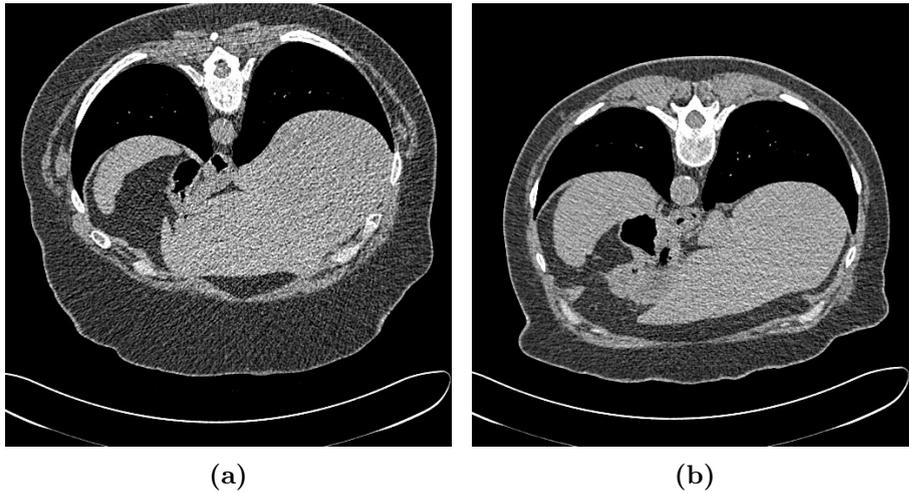


Figure 4.2: The scaling issues encountered in colon CT and head MRI datasets: (a) the obesity of the patient abdomen; and (b) the thinness of the patient abdomen expressed in different colon CT slices

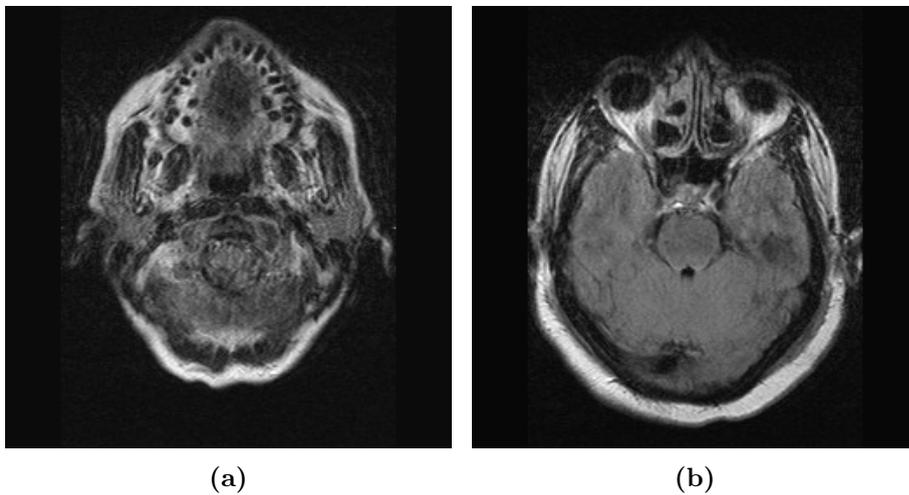


Figure 4.3: The noise issues encountered in head MRI dataset

[194] developed an image representation strategy based on combined feature sets, which involved contourlet coefficients, moments and GLCM-based features. High accuracy for rotated image retrieval was obtained. However, after the contourlet transformation, high-frequency sub-band coefficients could carry noise components that coexisted with other local structure details and degraded the retrieval performance for noisy images. Zhao et al. [195] applied the Fourier transform to LBP-based histogram to obtain rotation invariant features for static texture description and their experiments showed that their method outperforms earlier versions of LBP for rotated texture classification. The weakness of this method is that the LBP itself is sensitive to noise and therefore, the performance for classifying noise-corrupted images would not be satisfactory. Varma et

al. [196] proposed a straightforward raw image pixel-based feature from a local image patch to represent the texture image. Such a method with the property of local rotation invariance outperformed the conventional filter bank-based feature such as a maximum response (MR) filter multi-orientation feature. However, the raw image pixel-based local feature is vulnerable to noise, because the noise can randomly change the pixel intensity within the local image patch. Krishnamoorthi and Sathiya [200] introduced an orthogonal polynomial model to extract low-level textured information and model coefficients are reorganized into multi-resolution sub-bands to construct statistical and perceptual features for texture representation. Their experiments on image retrieval demonstrated the effectiveness of the proposed method. Unfortunately, the orthogonal polynomial model which is similar to edge or dot detectors, was not only sensitive to line or dot structures but also susceptible to noise effects. In many real-world applications, image degradation by noise is a very pervasive issue [201]. Hence, it is essential to consider the property of noise robustness in the development of rotation-invariant or rotation-robust image features.

4.1.2 Existing Works on Robustness Against Image Rotation and Scaling

Li et al. [202] employed a Laplacian of a Gaussian-based scale determination scheme to estimate the optimal scale for each pixel to adaptively select the LBP texton. The resultant LBP feature achieved scale invariance for rescaled texture representation. Although the scale variation was addressed, the orientation of local structures could not be determined adaptively and the LBP still relied on pre-quantified patterns for tackling image rotation. Li et al. [203] proposed a multilayer strategy incorporating texture pre-classification, rotation variant reduction and scale matching to create favorable conditions for diverse rotated and rescaled texture classification. However, this scheme relied primarily on pre-classification output and on preliminary orientation and scale adjustment prior to defining the Gabor filter bank. In addition, the Fourier spectrum analysis involved in the proposed strategy was easily affected by noise and scale-matching based on the prefixed discrete scale factor not being adaptive to the variation of real local spatial scale. Xie et al. [204] utilized predefined Gabor kernels to obtain a filtering impulse response and generate global description features. As circular orientation normalization and elastic cross-frequency searching were used to cope with the effect

of rotation and scaling, respectively, the parameters corresponding to the direction and central frequency of the Gabor kernels were predefined. Similarly, the predefined Gabor filter bank with fixed discrete orientations and scales was employed to achieve rotation and scale robustness for texture classification in [205].

4.1.3 The Strategies to Cope with Image Rotation and Scaling

Image rotation or scaling directly changes the orientation or scale of local structure. Therefore, the estimation of orientation and scale for local structure under conditions of image rotation and scaling is still a challenge because the necessary information required about the local image patch to infer the real local orientation and scale is inherently imprecise. In addition, such a challenge becomes even more difficult under noise degradation.

In order to cope with the local orientation variation caused by the image rotation, the LBP depends on quantified patterns to cover possible local rotation types and Gabor features rely on filter responses generated by a predefined multi-orientation filter bank. Being different from the previous quantization and predefined multi-orientation strategy, the rotation invariant filter-based approaches are introduced. To be explicit, the circularly shaped Laplacian of Gaussian filters and Gaussian filters are predefined in Leung-Malik (LM) [199] and maximum response (MR) filter banks [197] to achieve the rotation invariant features. Similarly, the predefined rotationally symmetric Schmid (S) filter bank [198] is also to yield the rotation invariant features.

In order to deal with the local scale variation caused by the image scaling, a multi-scale strategy is introduced to reduce the effects of image scaling [206]. This straightforward approach consists of detecting local features over a range of fixed discrete scales from coarseness to fineness, and then unifying all these extracted features from the specified scales to construct an image feature. Similar work can be found in [207–211]. To be specific, Ojala et al. [207] replicated the LBP extraction procedure through the whole image with different sized sliding windows (e.g. 3×3 , 5×5 and 7×7) and concatenated the resultant statistical histograms into a whole one for texture representation in order to resist local scale variation. The multi-scale strategy is very popular and is widely used to reinforce the robustness against image scaling such as the multi-scale complete LBP [208], multi-scale invariant texture feature [209], multi-scale robust WLD [210] and

multi-scale GLCM [211]. Being different from the multi-scale methods, an automatic local scale selection scheme is proposed in [213], based on the scale space theory [212] [214]. Then, a scale invariant feature transform (SIFT) is introduced to extract the key points with varying local scales [220]. Also, another adaptive local scale estimation, based on the local support region in the image spatial domain, is proposed in [234, 235].

Even though development of robust strategies to cope with image rotation and scaling have drawn much attention and several methods have been reported during the past years, more work still remains to be done, in particular, concerning variable random noise in medical images. Few studies can simultaneously address multiple effects – rotation, scaling and noise. The SIFT scheme is robust enough to cope with image rotation and scaling, however, such a scheme cannot be immediately adopted in medical image classification; that is, the key point descriptors yielded in the SIFT scheme are not able to describe the whole image but only to represent the extracted key points themselves. Accordingly, such descriptors cannot be scattered in feature space for categorizing or distinguishing images. Furthermore, the SIFT scheme works on an image-to-image matching basis. Until all pairwise matchings are completed, image classification cannot proceed. Such a scheme is inefficient especially for handling medical image classification tasks involving large databases. Similar to the key point descriptor in the SIFT scheme, conventional key point-based local structure descriptors, such as SURF [215], CORNER [216, 217], FAST [218] and BRISK [219], still cannot be directly scattered into feature space to represent various images.

4.1.4 Contributions in This Chapter

Concerning the particular characteristics and tangible demands in medical images classification tasks, the contributions of this chapter are summarized as follows:

- To detect local structures with various orientations and scales, we introduce an adaptive estimation strategy to simultaneously identify local dominant orientation and scale. Furthermore, for estimating local gradient (local gradient is used to represent the dominant local pixel intensity variation), a Gaussian first derivative kernel is adaptively constructed based on the identified local orientation and scale;

- To further resist the noise effect and better preserve the local structure, the anisotropic kernel (elliptical shape) instead of the isotropic kernel (circular shape) is adopted, not only to suppress the noise perpendicular to the local dominant orientation but also to preserve the local salient structure along its dominant orientation;
- To quantify the isotropy or anisotropy of pixel-wise particle potential motion, a new feature named particle potential motion entropy (PPME) is introduced;
- To discriminate local structure and non-structure pixels, a new sequential feature named self-information is proposed;
- To describe a whole image, a new histogram-based image feature named particle potential motion entropy histogram (PPMEH) is constructed;
- To achieve image rotation robustness, an updated version of PPMEH (named PPMEH-FT) modified by discrete Fourier transform (DFT) is yielded.

The remainder of this chapter is organized as follows: in Section 4.2, we present the details of the proposed image feature. In Section 4.3, the experiment on the synthetic structure image dataset demonstrates the effectiveness of the proposed scheme when detecting image local structures. Furthermore, the image classification experiments on colon CT slices and head MRI scans are conducted to demonstrate the effectiveness of the proposed image feature. The comparison with other state-of-the-art methods is also given to show the superiority of the proposed image feature. Finally, the conclusion is stated in Section 4.4.

4.2 Proposed Image Feature

In this section, the procedure to generate the proposed image feature is summarized as follows: (i) firstly, an adaptive and noise-robust local gradient estimation is conducted; (ii) then, image gradient force for individual pixel is calculated; (iii) depending on the pixel-wise gradient force, the image potential particle motion probability (PPMP) distribution for each pixel is computed; (iv) furthermore, potential particle motion entropy (PPME) to quantify pixel-wise uncertainty and self-information sequential pattern to

characterize local structure pixel are achieved, respectively; and (v) finally, by integration of pixel-wise PPME and orientation, a potential particle motion entropy histogram (PPMEH) is constructed to represent the whole image.

4.2.1 Adaptive and Noise-robust Local Gradient Estimation

A local structure stands for the subregion whose pixel intensities vary. For quantifying the variation of pixel intensities in a local structure region, the local gradient is a very useful property [192]. We estimate a local gradient by using a Gaussian first-derivative kernel, which is a powerful tool to extract image gradient information. In consideration of the diversity of orientations and scales for different local structures and the sensitivity to noise effect, a straightforward estimation of gradient by first derivatives along the horizontal and vertical directions is impractical. Therefore, to extract local structure information and resist noise effect, an anisotropic Gaussian first-derivative kernel with an elliptical shape is preferable to an isotropic kernel with a circular shape. The resistance to noise corruption and local feature preservation for an anisotropic Gaussian first-derivative kernel have been theoretically and experimentally verified in [226]. In consideration of the superiorities of the anisotropic Gaussian first-derivative kernel, one question arises here: how can we generate such a kernel to adapt to image local structure? Some works with respect to affine Gaussian kernel have been reported in [227, 228], however, this question still remains open and more work needs to be done to specifically overcome noise effect when adapting to noisy local structures.

We utilize a local singular value decomposition (SVD)-based scheme to adaptively tune the kernel parameters and then yield the kernel. This scheme is embedded into our system as an internal module. For implementation, the SVD-based module is plugged into the kernel generation scheme. We estimate the local dominant orientation and scale according to [234, 235]. Herein, the local gradient matrix for image \mathbf{I} is denoted as follows

$$\Lambda_i = \begin{bmatrix} \vdots & \vdots \\ z_{x_1}(u_j) & z_{x_2}(u_j) \\ \vdots & \vdots \end{bmatrix} \quad (4.1)$$

$$= \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T \quad (4.2)$$

where $u_j = [x_j, y_j]^T$ denotes the coordinate of pixel j belonging to local window \mathbf{w}_i ($u_j \in \mathbf{w}_i$). $z_{x_1}(\cdot)$ and $z_{x_2}(\cdot)$ are the first derivatives along horizontal x_1 and vertical x_2 directions, respectively. The local window \mathbf{w}_i centered at pixel i supports SVD analysis. $\mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T$ denotes singular value decomposition of the gradient matrix Λ_i : (i) \mathbf{U}_i is the unitary matrix; (ii) \mathbf{S}_i is a diagonal matrix with singular values placed along the diagonal and reflects the energy along the dominant orientation of the local structure; and (iii) unitary matrix \mathbf{V}_i^T denotes the transposing of the unitary matrix \mathbf{V}_i . Here we use first column $[\nu_{11}^i, \nu_{21}^i]^T$ of matrix \mathbf{V}_i to calculate local dominant orientation θ_i as follows

$$\theta_i = \arctan\left(\frac{\nu_{11}^i}{\nu_{21}^i}\right) \quad (4.3)$$

where the singular vector $v_1^i = [\nu_{11}^i, \nu_{21}^i]^T$ corresponding to the largest non-negative singular value s_{11}^i in \mathbf{S}_i represents the dominant orientation of the local structure. Furthermore, the elongation parameter e_i for kernel generation can be subsequently tuned according to the following formula

$$e_i = \frac{s_{11}^i}{s_{22}^i} \quad (4.4)$$

where s_{11}^i and s_{22}^i are the singular values along the diagonal of matrix \mathbf{S}_i , in descending order. Equation 4.4 signifies that the kernel shape in a non-structure region (with mild pixel intensity variation) approximates to circle ($s_{11} \approx s_{22} \approx 0$), but deforms to ellipse when encountering a salient local structure (edge and texture) due to the distinct variation of pixel intensities in the salient structure region ($s_{11} \gg s_{22}$).

$$\sigma_i = \sqrt{\frac{\sigma_{\mathbf{w}_i}^2 + \sigma_n^2}{\sigma_{\mathbf{w}_i}^2}} \quad (4.5)$$

where

$$\sigma_{\mathbf{w}_i}^2 = \frac{\sum_{u_j \in \mathbf{w}_i} (u_j - \bar{u}_{\mathbf{w}_i})^2}{M} \quad (4.6)$$

and

$$\sigma_n^2 = \frac{\sum_i \sigma_{\mathbf{w}_i}^2}{HW} \quad (4.7)$$

smoothing parameter σ_i manipulates the suppression of the noise (the more noise carried in an image, the larger σ_i is obtained and the more smoothing is conducted). σ_i is regarded as relative local standard deviation and computed according to Equation 4.5. $\sigma_{\mathbf{w}_i}^2$ is the image local variance computed over the local window \mathbf{w}_i . M is the total number of pixels in the local window \mathbf{w}_i and $\bar{u}_{\mathbf{w}_i}$ is the average pixel intensity. The image noise σ_n^2 can be approximated by means of all local variance in the image via Equation 4.7 (H and W are image height and width, respectively). The scale parameter γ_i for kernel generation can be calculated according to the following formula

$$\gamma_i = \left(\frac{s_{11}^i s_{22}^i}{M} \right)^\alpha \quad (4.8)$$

where the product of two singular values in Equation 4.8 indicates the energy of the local structure: the more salient the local structure is, the larger γ_i that can be gained; otherwise, in a non-structure region, a smaller γ_i is obtained. In the proposed adaptive kernel generation scheme, the larger γ_i makes the kernel smaller in the local structure region, whereas the smaller γ_i results in the larger kernel shape in the non-structure region. α is a noise suppression parameter ($0 < \alpha \leq 0.5$), and smaller α is favorable when the image carries more noise.

A 2D anisotropic Gaussian kernel is denoted as follows

$$\mathbf{g}(u_i) = \frac{1}{2\pi\sigma_i^2} \exp \left(-\frac{1}{2\sigma_i^2} \gamma_i u_i^T \mathbf{R}_{\theta_i}^T \begin{bmatrix} e_i & 0 \\ 0 & 1/e_i \end{bmatrix} \mathbf{R}_{\theta_i} u_i \right) \quad (4.9)$$

where

$$\mathbf{R}_{\theta_i} = \begin{bmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{bmatrix} \quad (4.10)$$

$u_i = [x_i, y_i]^T$ denotes pixel i coordinate in image \mathbf{I} . Based on Equation 4.9, the Gaussian first-derivative kernel along the local dominant orientation θ_i can be formulated as follows

$$\mathbf{g}'(u_i) = \frac{\partial [\mathbf{g}(u_i)]}{\partial \theta_i} \quad (4.11)$$

$$= -\frac{e_i \cdot [\cos \theta_i, \sin \theta_i] \cdot u_i}{\sigma_i^2} \mathbf{g}(u_i) \quad (4.12)$$

Masking local window \mathbf{w}_i with kernel $\mathbf{g}'(u_i)$ results in the following response

$$|\nabla \mathbf{I}_i| = \frac{\partial [\mathbf{I}(u_i)]}{\partial \theta_i} \quad (4.13)$$

$$= \mathbf{w}_i \odot \mathbf{g}'(u_i) \quad (4.14)$$

where \odot is the dot-product operator and the resultant response is the maximal directional derivative perpendicular to the local dominant orientation θ_i . Based on the relationship between the gradient and the maximal directional derivative, gradient magnitude $|\nabla \mathbf{I}_i|$ at pixel i is theoretically regarded as the maximal response of kernel $\mathbf{g}'(u_i)$ and the gradient direction $\theta_{\mathbf{G}_i} = \arctan(\nu_{12}^i / \nu_{22}^i)$ perpendicular to θ_i (singular vector $v_2^i = [\nu_{12}^i, \nu_{22}^i]^T$ corresponds to the smallest non-negative singular value s_{22}^i in \mathbf{S}_i and represents the orientation of the local gradient).

The procedure for estimating the local gradient is summarized as follows

- *Step 1:* SVD is utilized to estimate the local dominant orientation θ_i , elongation factor e_i , smoothing parameter σ_i and local scale γ_i , respectively;
- *Step 2:* Gaussian first-derivative kernel $\mathbf{g}'(u_i)$ is generated by means of the parameters adaptively tuned in *Step 1*;
- *Step 3:* Mask the local window \mathbf{w}_i with the generated kernel $\mathbf{g}'(u_i)$;

- *Step 4:* The maximal response $\frac{\partial[\mathbf{I}(u_i)]}{\partial\theta_i}$ within the local window \mathbf{w}_i is obtained, then the local gradient magnitude $|\nabla\mathbf{I}_i|$ is achieved (gradient direction $\theta_{\mathbf{G}_i} = \arctan\left(\frac{v_{12}^i}{v_{22}^i}\right)$ is computed);
- *Step 5:* The reduplication from *Step 1* to *Step 4* is performed until the local gradient information for all pixels is obtained.

4.2.2 Image Gradient Force

Image gradient force is a new concept, and it is firstly presented in this thesis. Such force is caused by image gradient, which is analogical to the air pressure gradient force in the atmosphere; that is, air pressure gradient force is driven by the air pressure difference, to push air to move from a high-pressure zone towards a low-pressure zone and accordingly result in wind. Its direction is perpendicular to the air pressure isobar, and its effect is proportional to the pressure difference within the unit distance [236]. The formula to define air pressure gradient force is given as follows

$$\frac{\mathbf{F}_p}{m} = \frac{1}{\rho} \frac{\Delta\mathbf{P}}{\Delta\mathbf{d}} \quad (4.15)$$

where \mathbf{F}_p is the air pressure gradient force imposing on the unit mass air m . ρ is air density which is a constant. $\Delta\mathbf{P}$ denotes the difference in air pressure. $\Delta\mathbf{d}$ is the distance in the direction from the high-pressure zone to the low-pressure zone. We rewrite Equation 4.15 as follows

$$\mathbf{F}_p = \frac{m}{\rho} \frac{\Delta\mathbf{P}}{\Delta\mathbf{d}} \quad (4.16)$$

$$= K \frac{\Delta\mathbf{P}}{\Delta\mathbf{d}} \quad (4.17)$$

$$\stackrel{\Delta\mathbf{d} \rightarrow 0}{=} K \nabla\mathbf{P} \quad (4.18)$$

where K is a constant item including unit air weight m and density ρ . $\nabla\mathbf{P}$ is the pressure gradient representing a rapid pressure change along the direction perpendicular to the pressure isobar. In consideration of the analogy between the image pixel intensity field

and the air pressure field (both are scalar fields), we define the gradient force driven by the image pixel intensity gradient as follows

$$\mathbf{F}_G = K_1 \frac{\Delta \mathbf{I}}{\Delta \mathbf{d}} \quad (4.19)$$

$$\stackrel{\Delta \mathbf{d} \rightarrow 0}{=} K_1 \nabla \mathbf{I} \quad (4.20)$$

where K_1 is a constant item. $\Delta \mathbf{I}$ and $\Delta \mathbf{d}$ denote the difference of pixel intensity between pairwise pixels and their distance, respectively. Image gradient $\nabla \mathbf{I}$ can be obtained when $\Delta \mathbf{d} \rightarrow 0$. \mathbf{F}_G has the same direction as the gradient $\nabla \mathbf{I}$.

The local gradient estimation is the foundation for the gradient force definition. Since the local gradient magnitude and orientation for each pixel can be obtained in the previous section, we adopt the obtained local gradient information to define pixel-wise gradient force according to Equation 4.20.

4.2.3 Image Particle Potential Motion Probability

Image particle potential motion probability, another new concept, is introduced in this section based on the effect of image gradient force. Such a probability quantifies the chance of particle potential motion happening in the specific direction of a biased random walk. Furthermore, such a probability not only involves the intrinsic potential motion probability due to the unbiased random walk, but also contains the extrinsic potential motion probability driven by the external force. In order to calculate the potential motion probability in the specific direction, therefore, the external force which happened in such a direction needs to be calculated in advance. To be explicit, we firstly assume the sole particle located at each pixel location. Then, we utilize a mechanism model to analyze the pixel-wise gradient force acting on the pixel-wise particle. According to the theory of Newton classical mechanics, one force can be decomposed into several component forces, in an orthogonal or even in a non-orthogonal coordinate system. Because the gradient force is a local concept restricted to its pixel neighbor, the pixel neighbor with fixed size 3×3 is adopted and analyzed in the following sections (3×3 neighbor is the basic setting for the image local analysis). Therefore, the gradient force decomposition is conducted within the 3×3 local region. As a result, the eight-component

force $\mathbf{F}_{G_{\theta_k}}$ ($\theta_k = 0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ$ and 315°) in eight directions within the 3×3 neighbor can be obtained by following the force decomposition principle. Even though \mathbf{F}_G could be decomposed in any arbitrary orientation (e.g. $32^\circ, 1.5^\circ, 300^\circ$), we only consider eight specific directions because the particle centrally located in the 3×3 local region only has eight neighbors. Herein, the eight surrounding neighbors impose their own force on the central particle separately. Thus, the gradient force \mathbf{F}_G on the central particle can be regarded as the combined effect of the eight component forces caused by its eight neighbors.

In an unbiased random walk restricted to the 3×3 local region, the central particle has equal potential probabilities in eight directions (isotropic potential motion case). On the contrary, with the effect of the gradient force, the random walk within the 3×3 domain becomes a biased one. The potential probabilities in eight directions would be different (anisotropic potential motion case). Furthermore, how to calculate the potential probability in eight directions within the 3×3 local region is illustrated by the following cases.

Case 1 - Isotropic Particle Potential Motion Case: we assume all the pixel intensities within the 3×3 local region are equivalent (no pixel intensity variation), therefore $\mathbf{F}_G = K_1 \nabla \mathbf{I}$ equals to zero everywhere ($\nabla \mathbf{I} = 0$). Accordingly, we have the following potential motion probability in eight directions

$$p_{\theta_1} = p_{\theta_2} = p_{\theta_3} = \dots = p_{\theta_k} = \dots = p_{\theta_8} = 1/8 \quad (4.21)$$

Case 2 - Anisotropic Particle Potential Motion Case: we assume the pixel intensities within the 3×3 local region vary (fluctuant intensity surface), therefore \mathbf{F}_G exists and we have

$$\hat{p}_{\theta_k} = p_{\theta_k} + \Delta p_{\theta_k} \quad (4.22)$$

$$= 1/8 + \Delta p_{\theta_k} \quad (4.23)$$

$$\Delta p_{\theta_k} = K_2 \mathbf{F}_{G_{\theta_k}} \quad (4.24)$$

where Δp_{θ_k} denotes the probability increment caused by the component force $\mathbf{F}_{G_{\theta_k}}$ in the direction of θ_k . In order to calculate the component force $\mathbf{F}_{G_{\theta_k}}$ (if $\mathbf{F}_{G_{\theta_k}}$ could be calculated, Δp_{θ_k} would be obtained subsequently), the component force $\mathbf{F}_{G_{\theta_k}}$ calculation is conducted through the decomposition of the image gradient force \mathbf{F}_G based on the overcomplete basis [238].

For a set of overcomplete basis vectors, the number of basis vectors is obviously greater than the dimensionality of the input space. Consequently, the representation of an input is not a unique linear combination over such overcomplete basis vector sets. Overcomplete representation has been theoretically proved and practically validated to be superior to complete representation, because the overcomplete representation not only has more flexibility on fitting the data structure in the input space but also has more robustness when handling the noisy data [54, 239]. To be explicit, the input in this case denotes the given image gradient force \mathbf{F}_G . Then, the dimension of input space is two, because \mathbf{F}_G can be completely represented by the following two component forces: $\mathbf{F}_G^{(x)} = \mathbf{F}_G \cdot \cos(\theta_G)$ along x axis (0° direction) and $\mathbf{F}_G^{(y)} = \mathbf{F}_G \cdot \sin(\theta_G)$ along y axis (90° direction). Also, the directional vectors $\mathbf{C}_0 = e^{i0}$ and $\mathbf{C}_{\pi/2} = e^{i\frac{\pi}{2}}$ are used as the complete basis vectors for complete representation of \mathbf{F}_G .

Being different from the complete representation of \mathbf{F}_G , the overcomplete representation of \mathbf{F}_G is conducted based on a series of overcomplete basis vectors. Therefore, the construction of overcomplete basis vectors plays a crucial role in overcomplete representation. In practice, the methods to yield the overcomplete basis vectors are multiple, however, one simple strategy can be adopted to produce overcomplete basis vectors; that is, we add some extra directional vectors to an existing complete basis set (which has already contained $\mathbf{C}_0 = e^{i0}$ and $\mathbf{C}_{\pi/2} = e^{i\frac{\pi}{2}}$) to obtain an overcomplete basis vector set. Accordingly, the following six directional vectors are added: $\mathbf{C}_{\pi/4} = e^{i\frac{\pi}{4}}$, $\mathbf{C}_{3\pi/4} = e^{i\frac{3\pi}{4}}$, $\mathbf{C}_\pi = e^{i\pi}$, $\mathbf{C}_{5\pi/4} = e^{i\frac{5\pi}{4}}$, $\mathbf{C}_{3\pi/2} = e^{i\frac{3\pi}{2}}$ and $\mathbf{C}_{7\pi/4} = e^{i\frac{7\pi}{4}}$. As a consequence, the yielded overcomplete basis vector set is illustrated as follows

- Along 0° direction: $\mathbf{C}_0 = e^{i0} = (1, 0)$;
- Along 45° direction: $\mathbf{C}_{\pi/4} = e^{i\frac{\pi}{4}} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$;
- Along 90° direction: $\mathbf{C}_{\pi/2} = e^{i\frac{\pi}{2}} = (0, 1)$;
- Along 135° direction: $\mathbf{C}_{3\pi/4} = e^{i\frac{3\pi}{4}} = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$;

- Along 180° direction: $\mathbf{C}_\pi = e^{i\pi} = (-1, 0)$;
- Along 225° direction: $\mathbf{C}_{5\pi/4} = e^{i\frac{5\pi}{4}} = \left(\frac{-\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right)$;
- Along 270° direction: $\mathbf{C}_{3\pi/2} = e^{i\frac{3\pi}{2}} = (0, -1)$;
- Along 315° direction: $\mathbf{C}_{7\pi/4} = e^{i\frac{7\pi}{4}} = \left(\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right)$.

However, because of the redundancy caused by the proposed overcomplete basis vector set, a given image gradient force could have non-unique linear representation over such an overcomplete set. In order to address this issue, we develop a new transformation: vector frequency domain decomposition (VFDD) for \mathbf{F}_G overcomplete representation and $\mathbf{F}_{G_{\theta_k}}$ calculation. The motivation to utilize the obtained component forces $\{\mathbf{F}_{G_{\theta_k}}\}$ rather than using the single gradient force \mathbf{F}_G is: the sequential pattern represented by $\{\mathbf{F}_{G_{\theta_k}}\}$ is more useful for distinguishing the local structure pixels from the non-structure ones compared with the single variable \mathbf{F}_G (shown in Figure 4.4). To be explicit, the vector \mathbf{F}_G can be decomposed by a set of vectors \mathbf{C}_n with varying frequency f_n . Then we have the following formula

$$\mathbf{F}_G = \sum_{n=1}^{\infty} |\mathbf{C}_n| e^{i\theta_n} \quad (4.25)$$

$$= \sum_{n=1}^{\infty} |\mathbf{C}_n| e^{i(2\pi f_n)} \quad (4.26)$$

where vector \mathbf{C}_n is also denoted as $|\mathbf{C}_n| e^{i\theta_n}$. We define $f_n = \frac{n-1}{N}$ where n and N are frequency and resolution control parameters, respectively. Consequently, Equation 4.26 can be rewritten as follows

$$\mathbf{F}_G = \sum_{n=1}^{\infty} |\mathbf{C}_n| e^{i[(n-1)\frac{2\pi}{N}]} \quad (4.27)$$

The meaning behind Equation 4.27 is: \mathbf{F}_G can be projected on to a set of vectors whose frequencies are diverse. With the augment of n , the frequency of the vector increases progressively. Here, a small n corresponds to a low-frequency vector while a large n corresponds to a high-frequency vector. Furthermore, the frequency resolution for the

vector is manipulated by N . That is, one revolution ($0 \rightarrow 2\pi$) is intended to be divided into only a few sectors (each sector has a large range) by a small N (corresponding to coarse frequency resolution); or divided into several sectors (each sector has a small range) by a large N (corresponding to fine frequency resolution). Therefore, the original vector could be projected on to either coarse or fine resolution vector sets. To achieve high-resolution vector decomposition, the vector set with fine resolution is preferable.

In order to decompose \mathbf{F}_G into $\mathbf{F}_{G\theta_k}$, the coefficients $|\mathbf{C}_n|$ for $n = 1, 2, 3, \dots, \infty$ need to be calculated in advance. Here, we propose the following equations for $|\mathbf{C}_n|$ calculation

1. **when** $n = 1$

$$|\mathbf{C}_1| = |\mathbf{F}_G| \cos(\theta_G) \quad \theta_G \in [0, 2\pi] \quad (4.28)$$

$$\theta_{\mathbf{R}_1} = \pi/2 \quad (4.29)$$

$$|\mathbf{R}_1| = |\mathbf{F}_G| \sin(\theta_G) \quad \theta_G \in [0, 2\pi] \quad (4.30)$$

2. **when** $n \geq 2$

$$|\mathbf{C}_n| = |\mathbf{R}_{(n-1)}| \cos \left[\theta_{\mathbf{R}_{(n-1)}} - \frac{(n-1)2\pi}{N} \right] \quad (4.31)$$

$$\theta_{\mathbf{R}_n} = \pi/2 + \frac{(n-1)2\pi}{N} \quad (4.32)$$

$$|\mathbf{R}_n| = |\mathbf{R}_{(n-1)}| \sin \left[\theta_{\mathbf{R}_{(n-1)}} - \frac{(n-1)2\pi}{N} \right] \quad (4.33)$$

In the 3×3 neighbor, velocity angles along the eight directions are spanned by $\pi/4$. Therefore, we choose $N = 8$ to obtain $\pi/4$ frequency resolution for VFDD. We rewrite Equation 4.27 as follows

$$\mathbf{F}_G = \sum_{n=1}^P |\mathbf{C}_n| e^{i[(n-1)\frac{2\pi}{8}]} + |\mathbf{R}_{(P+1)}| e^{i[\theta_{\mathbf{R}_{(P+1)}}]} \quad (4.34)$$

where $n = P$ corresponds to the highest frequency vector and $\mathbf{R}_{(P+1)} = |\mathbf{R}_{(P+1)}| e^{i[\theta_{\mathbf{R}_{(P+1)}}]}$ stands for the residual vector. As $\forall \varepsilon > 0, \exists n = P \in N^+, s.t. |\mathbf{R}_{(P+1)}| < \varepsilon$, then we have

$$\mathbf{F}_G \doteq \sum_{n=1}^P |\mathbf{C}_n| e^{i[(n-1)\frac{2\pi}{8}]} \quad (4.35)$$

Under such conditions, the component force $\mathbf{F}_{G_{\theta_k}}$ can be obtained by using the following formula

$$\mathbf{F}_{G_{\theta_k}} = \sum_{n \in \Omega_k} |\mathbf{C}_n| e^{i[(n-1)\frac{2\pi}{8}]} \quad (4.36)$$

where $\Omega_k = \left\{ n \left| g \left[\frac{(n-1)2\pi}{8} \right] = \theta_k, n = 1, 2, \dots, P \right. \right\}$ and $g(*)$ is the function to restrict the angle radian to the interval of $[0, 2\pi)$ (e.g. $g \left[\frac{(10-1)2\pi}{8} \right] = \pi/4$).

VFDD has the characteristics to actualize multi-resolution analysis in the frequency domain. In particular, the conventional complete decomposition is a special case under the VFDD framework; that is, with the frequency resolution $\pi/2$ (when $N = 4$), two basis vectors $\mathbf{C}_0 = e^{i0}$ and $\mathbf{C}_{\pi/2} = e^{i\frac{\pi}{2}}$ are selected for the VFDD. As a result, the complete decomposition only results in two frequency responses in two directions: 0° (x axis) and 90° (y axis); On the contrary, the VFDD with the frequency resolution $\pi/4$ (when $N = 8$) is capable of providing eight frequency responses in eight directions: $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ$ and 315° , respectively. The VFDD with fine frequency resolution $\pi/4$ (the case of overcomplete decomposition) outperforms VFDD with coarse frequency resolution $\pi/2$ (the case of conventional complete decomposition) when distinguishing the local structure pixels from the non-structure ones (shown in Figure 4.4). The white pixels indicate the local structure pixels identified by the proposed local structure pixel detection procedure, whereas the black pixels indicate the non-structure pixels. The visual comparison between Figure 4.4b and Figure 4.4c demonstrates that the sequential pattern is more effective than the single PPME value for distinguishing the local structure pixels from the non-structure ones. The visual comparison between Figure 4.4b and Figure 4.4d demonstrates that the VFDD with fine frequency resolution is more capable of producing the multi-directions information and benefit the local structure pixel identification than the VFDD with the coarse frequency resolution.

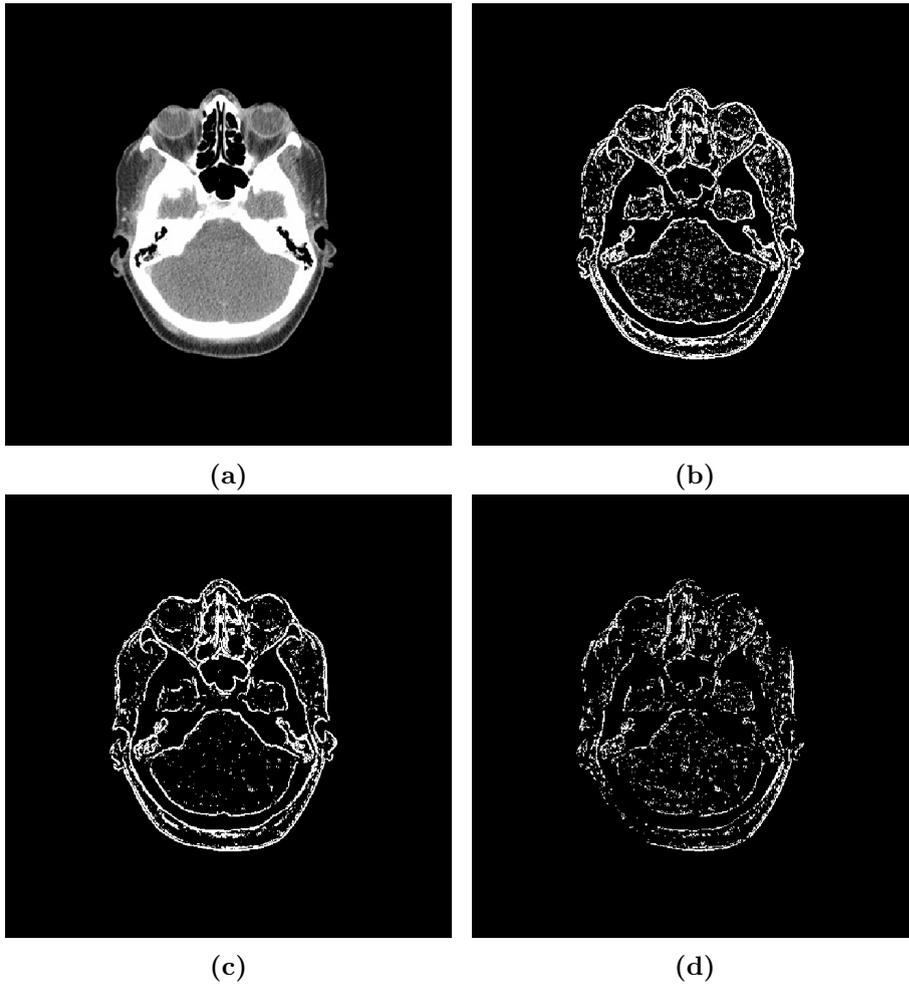


Figure 4.4: The illustration of the three binary indicator maps to highlight the local structure pixels: (a) original image (head CT scan); (b) the binary indicator map obtained based on the sequential pattern generated by the VFDD with fine frequency resolution $\pi/4$ and K-means clustering ($k = 2$); (c) the binary indicator map obtained based on the single PPME value by K-means clustering ($k = 2$); and (d) the binary indicator map obtained based on the sequential pattern generated by the VFDD with coarse frequency resolution $\pi/2$ and K-means clustering ($k = 2$)

Overall, the procedure to calculate the particle potential motion probability \hat{p}_{θ_k} is summarized as follows

- *Step 1:* The component force $\mathbf{F}_{G_{\theta_k}}$ is computed according to Equation 4.36;
- *Step 2:* Calculate the probability increment Δp_{θ_k} according to Equation 4.24 (K_2 is a constant and it can be removed in Equation 4.41);
- *Step 3:* The potential probability \hat{p}_{θ_k} is calculated according to Equation 4.22 ($\hat{p}_{\theta_k}^i$ may not satisfy the probability axioms in the current stage, therefore it needs to be updated according to Equation 4.41. The detail is given in the following section;

- *Step 4*: The reduplication from *Step 1* to *Step 3* is performed until the potential probability distributions $\{\hat{p}_{\theta_k}^i | k = 1, 2, 3, \dots, 8\}$ in the 3×3 neighbor for all pixels in the image domain are obtained.

4.2.4 Particle Potential Motion Entropy: An Uncertainty Measurement of Particle Potential Motion And Self Information: A Local Structure Detector

Based on the obtained image particle potential motion probability, particle potential motion entropy (PPME) and self-information, the two new concepts, are presented in this section.

4.2.4.1 Particle Potential Motion Entropy

Entropy is a measure of overall uncertainty with respect to all states of a random variable [237]. We extend the definition of Shannon information entropy to describe the pixel-wise uncertainty of potential motion. Therefore, PPME is defined as the uncertainty measurement to characterize the isotropic or anisotropic particle potential motion at each pixel location. The mathematical notation for PPME at pixel i is defined as follows

$$\mathbf{E}_i = - \sum_{k=1}^n \hat{p}_{\theta_k}^i \ln(\hat{p}_{\theta_k}^i) \quad (4.37)$$

where n denotes the total number of directions to be considered ($n = 8$). For the particle located at the central pixel in the 3×3 local region, the following cases are shown to illustrate how to calculate PPME.

Case 1 - Isotropic Particle Potential Motion Case: we assume all the pixel intensities within the 3×3 local region are equivalent (no pixel intensity variation), therefore $\mathbf{F}_G = K_1 \nabla \mathbf{I}$ equals to zero everywhere ($\nabla \mathbf{I} = 0$). Accordingly, the particle centrally located in the flat-intensity 3×3 local region behaves as in the unbiased random walk. The PPME is computed as follows

$$\mathbf{E}_i = - \sum_{k=1}^8 p_{\theta_k}^i \ln(p_{\theta_k}^i) = 2.0794 \quad (4.38)$$

Case 2 - Anisotropic Particle Potential Motion Case: we assume the pixel intensities within the 3×3 local region vary (fluctuant intensity surface), therefore \mathbf{F}_G exists and the particle centrally located in the varying-intensity 3×3 local region behaves as in the biased random walk. According to Equations 4.22, 4.24 and 4.36, we derive the formula for calculating the PPME as follows

$$\hat{p}_{\theta_k}^i = p_{\theta_k}^i + K_2 \mathbf{F}_{G_{\theta_k}}^i \quad (4.39)$$

$$= 1/8 + K_2 \mathbf{F}_{G_{\theta_k}}^i \quad (4.40)$$

where $\hat{p}_{\theta_k}^i$ may not satisfy the probability axioms in the current stage ($\hat{p}_{\theta_k}^i$ may be a negative decimal and its absolute value may be larger than 1). Following the probability axioms, the probability should be a non-negative decimal and in the interval of $[0, 1]$; besides, the summation of potential probabilities in eight directions should equal to 1. In order to comply with the probability axioms, we normalize $\hat{p}_{\theta_k}^i$ according to the following formula

$$\hat{p}_{\theta_k}^{i*} = \left(\frac{\hat{p}_{\theta_k}^i - \delta_i}{\eta_i - \delta_i} \right) / \left(\sum_k \frac{\hat{p}_{\theta_k}^i - \delta_i}{\eta_i - \delta_i} \right) \quad (4.41)$$

where $\hat{p}_{\theta_k}^{i*}$ is the normalized value of $\hat{p}_{\theta_k}^i$. δ_i and η_i are defined as follows

$$\delta_i = \min_k (\hat{p}_{\theta_k}^i) \quad (4.42)$$

$$\eta_i = \max_k (\hat{p}_{\theta_k}^i) \quad (4.43)$$

Equations 4.42 and 4.43 signify that δ_i and η_i need to be determined by eight directions $k = 1, 2, 3, \dots, 8$ in the 3×3 local region prior to the normalization. We also notice that the constant terms K_1 in Equation 4.20 and K_2 in Equation 4.39 can be removed simultaneously in the normalization conducted by Equation 4.41. Consequently, K_1 and K_2 could be tuned to any positive constant during the calculation procedures.

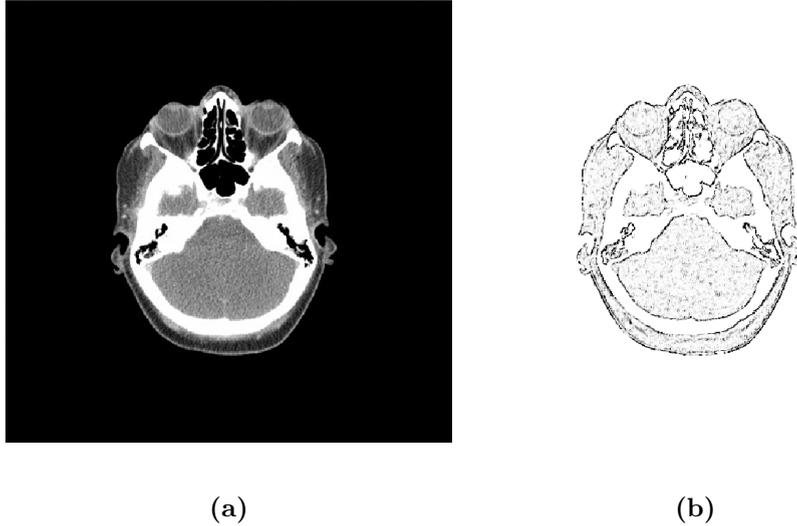


Figure 4.5: The PPME calculated at each pixel location is illustrated: (a) original image (head CT scan); and (b) the PPME map generated at each pixel location

$$\mathbf{E}_i = - \sum_{k=1}^8 (\hat{p}_{\theta_k}^i)^* \cdot \ln \left(\hat{p}_{\theta_k}^i \right) \quad (4.44)$$

A pixel-wise PPME map can be generated by following the computation process presented in **Case 2**. As an example, Figure 4.5 illustrates the pixel-wise PPME map obtained by following the computation process presented in **Case 2** for a given medical image. To be explicit, the dark pixels in the PPME map highlight the local structures whose anisotropy is dominant; otherwise, for the non-structure pixels, isotropy overwhelms anisotropy and the pixels appear bright.

4.2.4.2 Self Information

Self-information, being different from the PPME, is the quantity of information contained in a probabilistic potential motion in the specific direction and only depends on the potential probability in such a direction. The self-information sequence derived from the potential probability distribution $\left\{ \hat{p}_{\theta_k}^i \mid k = 1, 2, 3, \dots, 8 \right\}$ is also capable of characterizing the isotropy or anisotropy at pixel i . To discriminate two different types of pixel (local structure pixels and non-structure pixels), the self-information at pixel i in direction θ_k is defined according to [240]

$$h_{\theta_k}^i = - \ln \left(\hat{p}_{\theta_k}^i \right) \quad (4.45)$$

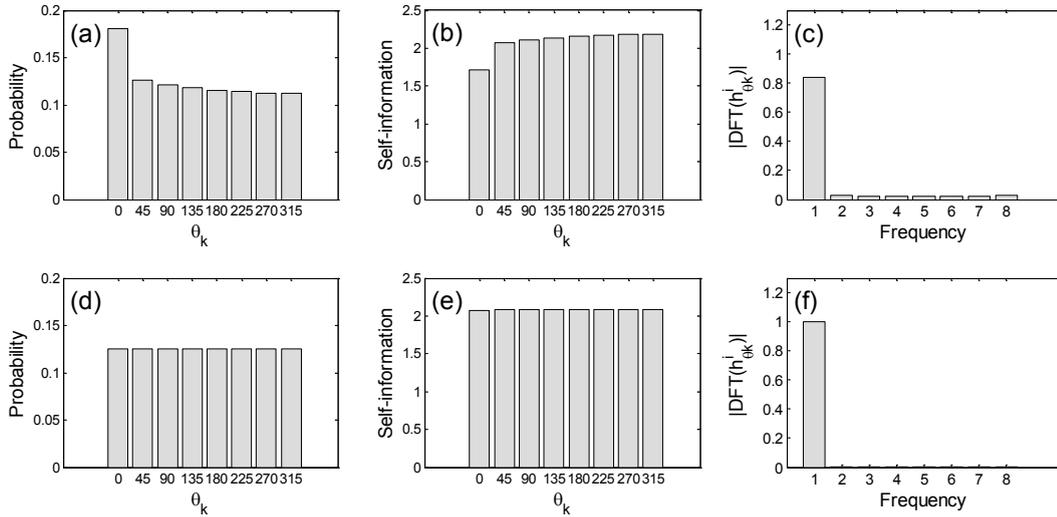


Figure 4.6: The illustration of potential probability distribution, self-information and DFT-based shift invariant feature in regard to local structure pixel (LSP) and non-structure pixel (NSP), respectively: (a) Potential probability distribution for an LSP centrally located in the 3×3 neighbor: [128, 157, 160; 124, 166, 165; 111, 154, 152]; (b) and (c) are the corresponding $\{h_{\theta_k}^{LSP}\}$ and $|DFT(\{h_{\theta_k}^{LSP}\})|$, respectively; (d) potential probability distribution for an NSP centrally located in the 3×3 neighbor: [7, 0, 0; 0, 5, 0; 5, 0, 1]; (e) and (f) are the corresponding $\{h_{\theta_k}^{NSP}\}$ and $|DFT(\{h_{\theta_k}^{NSP}\})|$

The sequential pattern reflected by sequence $\{h_{\theta_k}^i\}$ is more informative than the single PPME value for discriminating two different types of pixel, because $\{h_{\theta_k}^i\}$ provides multi-orientation information to characterize pixel-wise isotropic or anisotropic property. Figure 4.6 illustrates two types of pixel: local structure pixel (with anisotropic property), non-structure pixel (with isotropic property) and their corresponding self-information sequences. As shown in Figure 4.6, either the fluctuant or the flat pattern expressed by a self-information sequence is a distinctive feature to discriminate two such types of pixel. In other words, depending on fluctuant or flat sequential pattern, the pixel can be categorized into a local structure pixel or a non-structure pixel. However, the occurrences of “ups” and “downs” in a fluctuant self-information sequence are different from that of other fluctuant sequences. Therefore, the fluctuation patterns reflected by self-information sequences are correlated with direction θ_k . In order to avoid the correlation with θ_k , when comparing with different self-information sequences, the discrete Fourier transform (DFT) is applied to extract the spectrum information from the raw $\{h_{\theta_k}^i\}$ series, because of the shift invariance of DFT in 1D signal analysis. The local structure pixel detection, based on the integration of $\{h_{\theta_k}^i\}$ and DFT, is summarized as follows

- *Step 1:* Obtain the self-information sequence $\{h_{\theta_k}^i\}$ according to Equation 4.45 (shown in Figures 4.6b and 4.6e);
- *Step 2:* DFT is applied to $\{h_{\theta_k}^i\}$ and the magnitude spectrum $\left|DFT\left(\{h_{\theta_k}^i\}\right)\right|$ is used as a shift-invariant feature for local structure pixel identification. The magnitude spectrums $\left|DFT\left(\{h_{\theta_k}^i\}\right)\right|$ are illustrated in Figures 4.6c and 4.6f;
- *Step 3:* K-means clustering ($k = 2$) is utilized to classify the two types of pixel (local structure pixels and non-structure pixels) based on $\left|DFT\left(\{h_{\theta_k}^i\}\right)\right|$;
- *Step 4:* The pixel-wise clustering results are converted into a binary image as the pixel label indicator. In such a binary image, we set "one" for the local structure pixels and set "zero" for the non-structure pixels (shown in Figure 4.4b).

4.2.5 Particle Potential Motion Entropy Histogram: A Global Image Feature

Based on the locally estimated orientation and precalculated pixel-wise PPME in preceding sections, a particle potential motion entropy histogram (PPMEH) is proposed as the new image feature. The procedure to yield PPMEH is summarized as follows

- *Step 1:* Estimate the local dominant orientation θ_i according to Equation 4.3;
- *Step 2:* Calculate \mathbf{E}_i according to Equations 4.38 and 4.44 (the resultant PPME map is shown in Figure 4.5);
- *Step 3:* After the local dominant orientation estimation and the PPME computation for all pixels, the PPMEH is constructed according to the following stages: (i) for a specific integer angle $\bar{\beta} \in [1^\circ, 180^\circ]$, the image pixels whose dominant orientations are approximated to such $\bar{\beta}$ are selected as $\{u_i | \theta_i \doteq \bar{\beta}\}$; (ii) the PPME values corresponding to the selected pixels $\{u_i | \theta_i \doteq \bar{\beta}\}$ are summated to form a single histogram bin at angle $\bar{\beta}$; and (iii) with the change of $\bar{\beta}$ from 1° to 180° , the corresponding histogram bin is formed. When all the $\bar{\beta}$ (integer value) in the range of $[1^\circ, 180^\circ]$ are considered, the complete histogram with the fixed number of bins (in total 180) is formed (the resultant PPMEH is shown in Figure 4.7b).

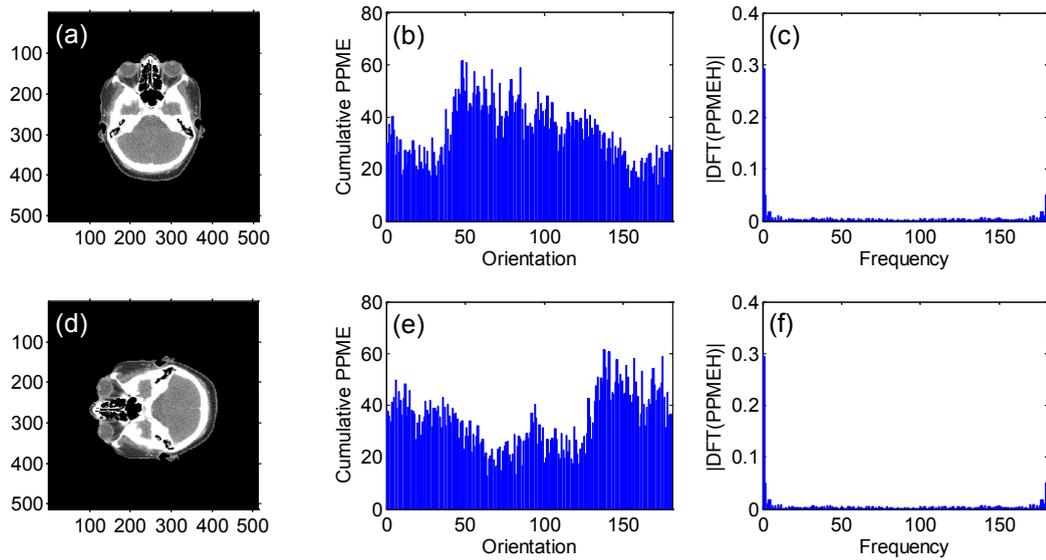


Figure 4.7: The illustration of image rotation robustness for the proposed PPMEH-FT: (a) original image; (b) and (c) correspond to PPMEH and PPMEH-FT, respectively; (d) 90° anticlockwise rotation happens to (a); (e) and (f) are PPMEH and PPMEH-FT corresponding to (d). The difference between (b) and (e) exhibits that the image rotation directly results in the translation of the histogram pattern. (c) and (f) demonstrate that very similar patterns can be achieved after using DFT and histogram normalization

A reliable image feature should be robust in the face of image rotation and scaling in practice. However, the basic PPMEH is correlated with the angle $\bar{\beta}$. Thus, it would be vulnerable to image rotation. To address the image rotation issue, DFT is still applied to the basic PPMEH for extracting the shift invariance magnitude spectrum $|DFT(PPMEH)|$ (as shown in Figure 4.7c). Consequently, the obtained magnitude spectrum is adopted as the image feature. For the image scaling issue, the image scaling directly changes the local structure scales [241, 242]. To be more explicit, the image scaling-up enlarges the local structure while the image scaling-down leads to the shrinkage of the local structure. In this sense, the capability of the image feature to adaptively cope with the varying local structure scales is critical. Fortunately, the self-adaptation to local structure scale variation has been achieved in Section 4.2.1, where the local scale adaptive estimation scheme has been embedded.

In summary, the basic PPMEH is only robust to the scale variation of the image local structures. Furthermore, with the modification of DFT, the upgraded version of PPMEH: PPMEH-FT is not only robust against the local structure scale variation but also robust in terms of image rotation (the rotation robustness is illustrated in Figure 4.7 and the scaling robustness is shown in Figure 4.8).

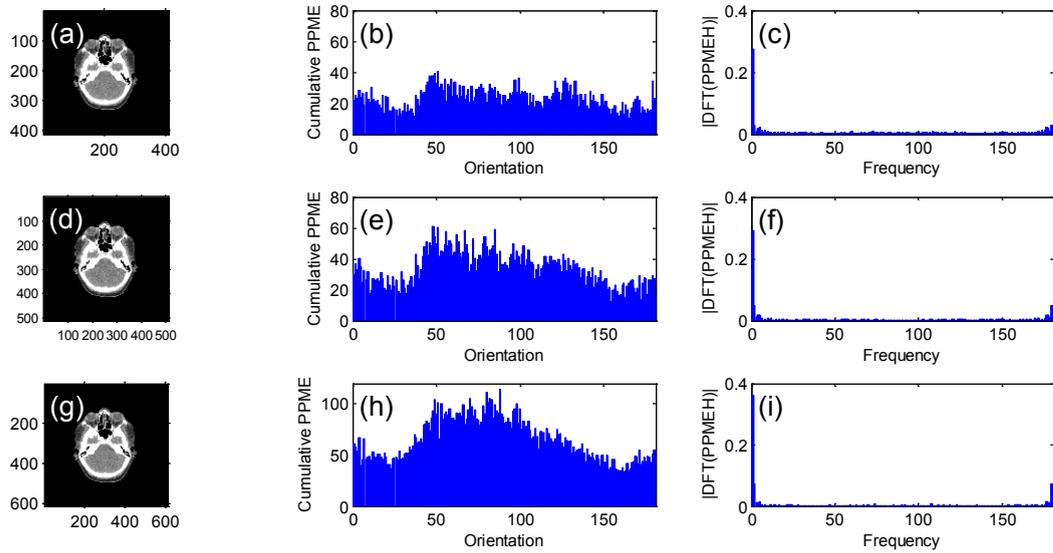


Figure 4.8: The illustration of image scaling adaptation for the proposed PPMEH-FT: (a) the image with scaling factor $s = 0.8$; (b) and (c) are PPMEH and PPMEH-FT corresponding to (a); (d), (e) and (f) are original image (scaling factor $s = 1$), PPMEH and PPMEH-FT, respectively; (g) the image with scaling factor $s = 1.2$; (h) and (i) are corresponding PPMEH and PPMEH-FT, respectively. The differences between (b), (e) and (h) demonstrate that the image scaling-up or scaling-down leads to the enlargement or shrinkage of the histogram bin amplitude. However, the histogram patterns expressed by (b), (e) and (h) still look similar and the magnitude spectrum features extracted by DFT are also similar in (c), (f) and (i), respectively

4.3 Experiment

In this section, the proposed local structure detector and image features are tested by the following three image datasets.

4.3.1 Data Preparation

- **Dataset 1 - Local Structure Image Data:** multi-orientation and multi-scale local structures are included in this image (shown in Figure 4.9a). We implement the proposed local structure detector to distinguish the local structure pixels from the non-structure pixels.
- **Dataset 2 - Colon CT Image Data:** 178 colon CT slices are acquired from nine patients who are suffering different colon lesions. Even though the CT scanning only targets the patient abdominal areas, the images obtained by the medical imaging system have visible differences due to the different types of colon lesion. In the colon CT scans, for example, the colon cancer usually accompanies the

hydatoncus characterized by the region with uneven texture [243]. Considering the anatomic detail that can be provided by colon CT imaging, we implement the proposed image features to represent the image content for the colon lesion identification.

- **Dataset 3 - Head MRI Image Data:** 312 head MRI scans are collected from five categories of patient samples: three groups of slices with different scanning layer depth along the axial direction, one group of the sagittal view and one group of the coronal view. Brain tumors are the abnormal tissues with distinctive features which are quite different from the other normal tissues and can mostly grow inside the brain [244]. For the given MRI slices with the fixed scanning plane and scanning depth, the radiologists prefer to manually select the other relevant slices from a background database for reference and comparison by means of similar image detailed content. Because MRI imaging procedure can effectively detect the signals emitted from normal and abnormal tissue and provide clear images with rich anatomic detail, we adopt the proposed image features to represent the head MRI scan samples.

To verify the effectiveness of the proposed local structure detector and image features PPMEH and PPMEH-FT, the following three experiments are performed.

4.3.2 Experiment 1: Local Structure Detection Performance Evaluation

The local structure detection is a typical binary classification problem; that is, the pixels belonging to the local structures are supposed to be distinguished from the non-structure ones. In this experiment, an image containing synthetic Gaussian noise and the local structures with diverse orientations and scales is used as the testing dataset. Furthermore, the following nine features are exploited for local structure pixel identification: (i) pixel intensity \mathbf{I} ; (ii) directly computed gradient magnitude $\sqrt{z_{x_1}^2 + z_{x_2}^2}$; (iii) two directly computed first derivatives $[z_{x_1}, z_{x_2}]$; (iv) adaptively estimated local gradient magnitude $|\nabla\mathbf{I}|$; (v) two first derivatives $[\nabla\mathbf{I}_{x_1}, \nabla\mathbf{I}_{x_2}]$ obtained based on local gradient $\nabla\mathbf{I}$ and its direction $\theta_{\mathbf{G}}$ (here $\nabla\mathbf{I}_{x_1} = |\nabla\mathbf{I}| \cdot \cos(\theta_{\mathbf{G}})$ and $\nabla\mathbf{I}_{x_2} = |\nabla\mathbf{I}| \cdot \sin(\theta_{\mathbf{G}})$); (vi) pixel-wise PPME value; (vii) sparse coding of $\nabla\mathbf{I}$ conducted by OMP [263]; (viii) sequential pattern

Table 4.1: The Local Structure Detection Performance Evaluated by ROC Attributes

	\mathbf{I}	$\sqrt{z_{x_1}^2 + z_{x_2}^2}$	$[z_{x_1}, z_{x_2}]$
TPR	0.4646	0.8342	0.8357
FPR	0.1790	0.2384	0.2435
AUC	0.6428	0.7978	0.7961
	$ \nabla\mathbf{I} $	$[\nabla\mathbf{I}_{x_1}, \nabla\mathbf{I}_{x_2}]$	PPME
TPR	0.5402	0.4818	0.5515
FPR	0.0037	0.0025	0.0039
AUC	0.7682	0.7396	0.7737
	$OMP(\nabla\mathbf{I})$	$ DFT(\{h_{\theta_k}\}) $ by VFDD ($\pi/2$)	$ DFT(\{h_{\theta_k}\}) $ by VFDD ($\pi/4$)
TPR	0.4562	0.3743	0.6424
FPR	0.0002	0.0004	0.0050
AUC	0.7280	0.6869	0.8187

$|DFT(\{h_{\theta_k}\})|$ via the VFDD with coarse frequency resolution $\pi/2$; and (ix) sequential pattern $|DFT(\{h_{\theta_k}\})|$ via the VFDD with fine frequency resolution $\pi/4$. After that, binary clustering (K-means clustering with $k = 2$) is performed on all above-mentioned features. As a result of binary clustering, the local structure pixels are labeled as the “white”, whereas the non-structure pixels are marked as the “black” (as shown in Figure 4.9).

For the quantitative evaluation of the local structure detection results, ground truth is given (as shown in Figure 4.9a) and Receiver Operating Characteristic (ROC) attributes are adopted. To be explicit, true positive rate (TPR: sensitivity) is defined as the percentage of pixels belonging to local structures which are correctly identified. False positive rate (FPR: specificity) is defined as the percentage of pixels belonging to non-structure but wrongly recognized as the local structure pixels. For the performance comparison among these seven features, Area Under Curve (AUC) in ROC analysis is utilized. The definition of AUC is given by Equation 4.46.

$$AUC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \left(\frac{TPR_{i+1} + TPR_i}{2} \right) \quad (4.46)$$

The definition of AUC reveals: the higher the AUC value is, the more reliably the feature works. The ROC attributes for seven features are demonstrated in Table 4.1, respectively.

As shown in Figures 4.9c, 4.9d and 4.9e, the pixel intensity \mathbf{I} , the directly computed gradient magnitude $\sqrt{z_{x_1}^2 + z_{x_2}^2}$ and the directly computed first derivatives $[z_{x_1}, z_{x_2}]$

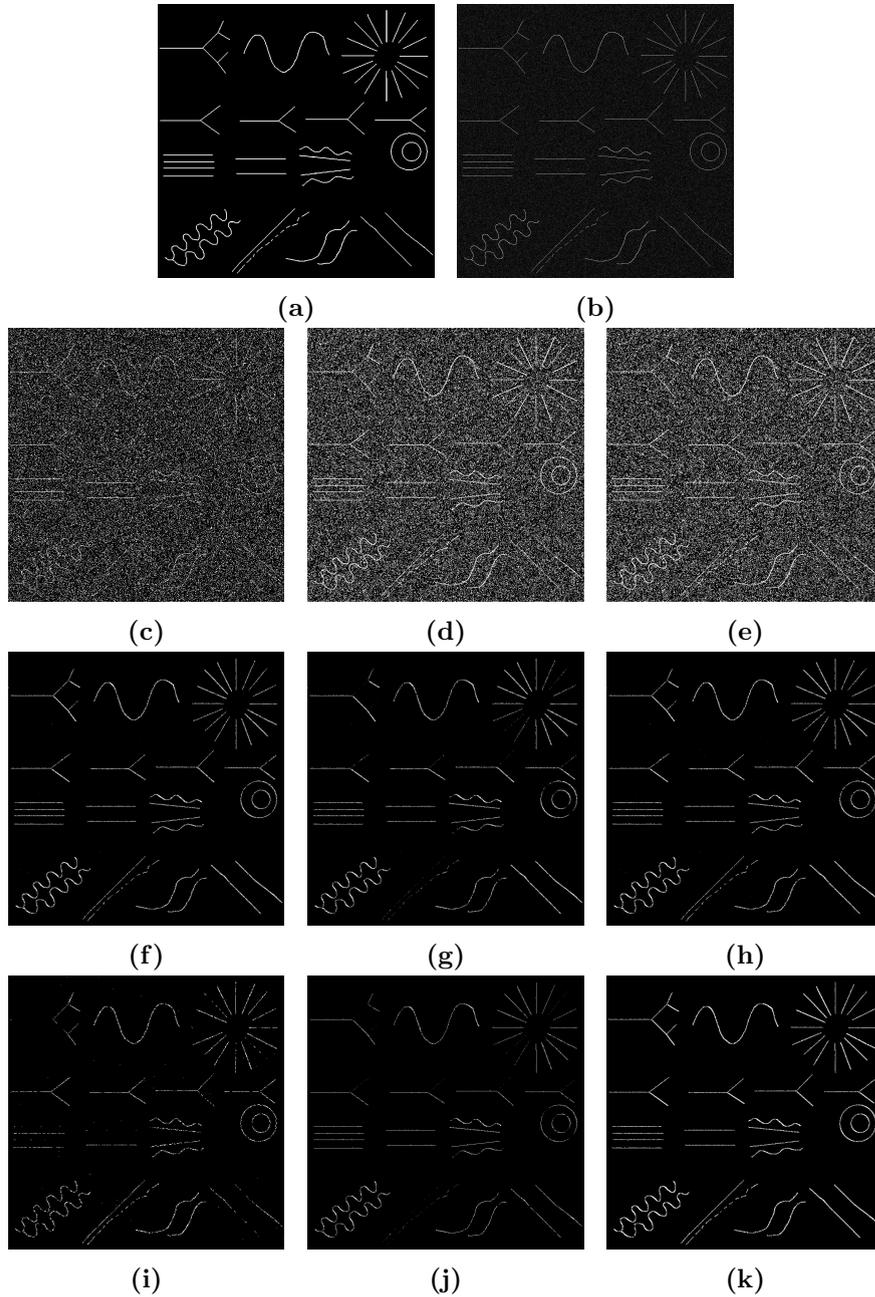


Figure 4.9: The binary local structure indicators obtained depending on diverse features and K-means clustering ($k = 2$ is applied to all the features): (a) ground-truth local structure image; (b) noisy local structure image degraded by synthetic Gaussian noise ($\sigma_n = 30$); (c) pixel intensity \mathbf{I} ; (d) directly computed gradient magnitude $\sqrt{z_{x_1}^2 + z_{x_2}^2}$; (e) two directly computed first derivatives $[z_{x_1}, z_{x_2}]$; (f) adaptively estimated local gradient magnitude $|\nabla \mathbf{I}|$; (g) two first derivatives obtained depending on local gradient $\nabla \mathbf{I}$ and its direction $\theta_{\mathbf{G}}$ $[\nabla \mathbf{I}_{x_1}, \nabla \mathbf{I}_{x_2}]$ ($\nabla \mathbf{I}_{x_1} = |\nabla \mathbf{I}| \cdot \cos(\theta_{\mathbf{G}})$ and $\nabla \mathbf{I}_{x_2} = |\nabla \mathbf{I}| \cdot \sin(\theta_{\mathbf{G}})$); (h) single PPME value; (i) sparse coding of $\nabla \mathbf{I}$ conducted by OMP [263]; (j) sequential pattern $|DFT(\{h_{\theta_k}\})|$ via the VFDD with coarse frequency resolution $\pi/2$; and (k) sequential pattern $|DFT(\{h_{\theta_k}\})|$ via the VFDD with fine frequency resolution $\pi/4$

are sensitive to random image noise. On the contrary, the adaptively estimated local gradient-based features are generally robust to image noise contamination. As illustrated in Figures 4.9f, 4.9g, 4.9h, 4.9i, 4.9j and 4.9k, the most local structure pixels are correctly identified while the few non-structure pixels are falsely detected. Additionally, the loss of local structure pixels is maintained at lower level, especially for those local structures with varying orientations and scales. In general, the fine frequency resolution VFDD scheme is able to yield the more informative and discriminating features rather than the coarse frequency resolution version. As visually compared between Figure 4.9j and Figure 4.9k, the local structure preservation of fine frequency resolution VFDD scheme is much better than that of the coarse frequency resolution VFDD scheme. This superiority could facilitate the subsequent PPMEH-based medical image classification tasks.

4.3.3 Experiment 2: Categorization of Given CT Slices to Correct Colon Lesion Group

The categorization of the given colon CT slice to a correct colon lesion group is indeed a multi-classification problem. For a given colon CT slice, the classification system needs to categorize the query sample into the correct colon lesion type. This experiment proceeded as follows

- *Step 1:* The proposed image features PPMEH and PPMEH-FT are extracted for colon CT images;
- *Step 2:* The extracted PPMEHs and PPMEH-FTs are divided into two portions: (i) some are used as background samples and (ii) the others are treated as query samples;
- *Step 3:* The PPMEHs and PPMEH-FTs in regard to background samples are scattered into the feature space;
- *Step 4:* Given the PPMEHs and PPMEH-FTs corresponding to query samples, apply K-nearest neighbor classifier to yield the category labels (colon lesion group labels) for query samples.

Next, five state-of-the-art features are selected for comparison. Local binary pattern (LBP) is a very popular feature in texture classification. Here, a complete local binary

pattern (CLBP), and an upgraded version of a conventional LBP are compared. Three LBP histograms $S_{(1,8)}^{riu2}/M_{(1,8)}^{riu2}/C$, $S_{(2,16)}^{riu2}/M_{(2,16)}^{riu2}/C$ and $S_{(3,24)}^{riu2}/M_{(3,24)}^{riu2}/C$ are concatenated to form the entire histogram as the image feature [208]. The Gabor filter (GF) has been broadly used in image classification tasks. Gabor kernels with multi-orientations and scales are predefined. Here, eight orientations $\varphi_k = k\pi/8$ for $k = 0 \rightarrow 7$ and five scales $f_l = (\sqrt{2})^{-l}$ for $l = 0 \rightarrow 4$ are adopted to generate a Gabor filter bank. The mean and standard deviation for each Gabor filter response are calculated, and then all the means and standard deviations for a total of forty filter responses are combined to yield the GF feature [205]. Grey level co-occurrence matrix (GLCM) is a conventional texture feature extractor. It has become a benchmark feature on many image classification applications. Here, we specify sixteen combinations to cover the varying directions and lags: four directions 0° , 45° , 90° and 135° and in every direction, four lags $h = 1 \rightarrow 4$ are selected. Therefore, all sixteen GLCMs are generated to quantify a given image. Based on the obtained GLCMs, four types of statistics, such as contrast, correlation, energy and homogeneity, are computed to create the GLCM feature [211]. Semi-variogram (SV) is one kind of spatial autocorrelation function based on spatial statistics, which is similar to GLCM and has been used for texture classification in the recent decade [230, 231]. In this experiment, a 7×7 sized window is used and a total of twenty-seven pixel lags $h = 1 \rightarrow 27$ are used to create the SV feature. Histogram of gradient (HOG) is a simple but very powerful image descriptor, which has been regarded as a state-of-the-art technique in a vision identification system [229, 232, 233]. For the formation of the proposed PPMEH and PPMEH-FT, the local structure parameters estimation is conducted based on the sliding window. Since the window size is the single parameter which needs to be tuned, this plays a critical role in influencing classification performance of the proposed image feature. As illustrated in Figure 4.10, the 11×11 sliding window is selected as an optima. Such window size is suitable to capture the large-scale image structures and also robust in terms of the image scale variation.

The image rotation and scaling issues are usually encountered during the medical imaging process [221–223]. In order to verify the robustness of the proposed PPMEH and PPMEH-FT features, the following scenarios are simulated

- **Scenario 1:** Only random rotation within $[-180^\circ, 180^\circ]$;
- **Scenario 2:** Only random rotation within $[-20^\circ, 20^\circ]$;

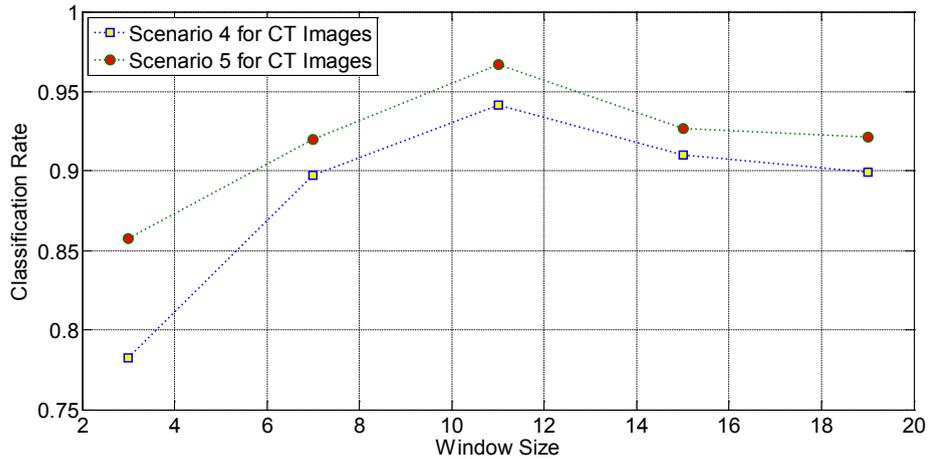


Figure 4.10: The illustration of sliding window size selection to cope with CT image random scaling in Scenario 4 and Scenario 5, respectively, where the classification accuracy of PPMEH is adopted as the selection criterion

- **Scenario 3:** Only random rotation within $[-10^\circ, 10^\circ]$;
- **Scenario 4:** Only random scaling within $[0.8, 1.2]$;
- **Scenario 5:** Only random scaling within $[0.9, 1.1]$;
- **Scenario 6:** Random rotation within $[-180^\circ, 180^\circ]$ in the conjunction with random scaling within $[0.8, 1.2]$;
- **Scenario 7:** Random rotation within $[-20^\circ, 20^\circ]$ in the conjunction with random scaling within $[0.8, 1.2]$;
- **Scenario 8:** Random rotation within $[-10^\circ, 10^\circ]$ in the conjunction with random scaling within $[0.9, 1.1]$.

In this classification experiment, all the colon CT slices belonging to each colon lesion group are randomly and evenly split into background and query samples (half background samples and half query samples). K-nearest neighbor ($k = 1$), the simple classifier, and Euclidean distance, the simple distance metric, are exploited to evaluate the classification performance. In addition, in order to ensure the statistics of significance, the classification experiment is independently and repeatedly conducted 50 times. Accordingly, the classification accuracies during the 50 trials are recorded respectively. The obtained statistics, in terms of the mean and standard deviation depending on 50-trial classification accuracies, are shown in Table 4.2.

Table 4.2: Image Feature Comparison for Image Random Rotation and Scaling in Colon CT Images Classification Experiment (Classification Rate: $\mu \pm \sigma$)

	Scenario 1 Rotation Only [$-180^\circ, 180^\circ$]	Scenario 2 Rotation Only [$-20^\circ, 20^\circ$]	Scenario 3 Rotation Only [$-10^\circ, 10^\circ$]	Scenario 4 Scaling Only [0.8, 1.2]
LBP[208]	31.24% \pm 4.08%	76.41% \pm 4.01%	83.22% \pm 2.86%	64.50% \pm 4.89%
GF[205]	20.59% \pm 3.35%	50.11% \pm 3.86%	58.63% \pm 3.68%	96.89% \pm 1.19%
GLCM[211]	45.09% \pm 3.54%	61.65% \pm 3.67%	75.93% \pm 3.82%	93.00% \pm 2.25%
SV[231]	59.02% \pm 4.35%	62.76% \pm 4.29%	71.37% \pm 3.84%	76.72% \pm 2.97%
HOG[229]	38.47% \pm 2.52%	79.13% \pm 5.17%	84.34% \pm 1.40%	58.69% \pm 3.85%
PPMEH	59.78% \pm 4.83%	79.37% \pm 4.28%	85.73% \pm 3.02%	94.15% \pm 2.51%
PPMEH-FT	62.17% \pm 3.51%	80.11% \pm 2.35%	86.35% \pm 4.35%	91.20% \pm 2.43%
	Scenario 5 Scaling Only [0.9, 1.1]	Scenario 6 Rotation [$-180^\circ, 180^\circ$] + Scaling [0.8, 1.2]	Scenario 7 Rotation [$-20^\circ, 20^\circ$] + Scaling [0.8, 1.2]	Scenario 8 Rotation [$-10^\circ, 10^\circ$] + Scaling [0.9, 1.1]
LBP[208]	88.57% \pm 2.81%	21.96% \pm 3.08%	58.15% \pm 4.18%	80.15% \pm 3.43%
GF[205]	97.02% \pm 0.94%	16.63% \pm 3.39%	65.85% \pm 4.64%	66.00% \pm 3.77%
GLCM[211]	98.11% \pm 1.37%	45.57% \pm 3.90%	50.09% \pm 4.36%	66.11% \pm 4.05%
SV[231]	87.78% \pm 2.81%	51.28% \pm 4.27%	52.04% \pm 3.74%	66.65% \pm 3.99%
HOG[229]	74.13% \pm 3.63%	18.91% \pm 3.45%	41.95% \pm 2.26%	54.78% \pm 4.48%
PPMEH	96.67% \pm 2.50%	51.13% \pm 4.47%	74.26% \pm 3.48%	81.23% \pm 3.72%
PPMEH-FT	94.02% \pm 2.72%	52.84% \pm 2.61%	74.65% \pm 2.56%	83.37% \pm 2.99%

Generally speaking, the proposed PPMEH-FT achieves comparably good classification performance in terms of the mean and standard deviation of classification accuracies after 50 independent and repeated trials. The PPMEH-FT is robust in the face of image random rotation, random scaling and particularly, the conjunction of random rotation and scaling. For image random rotation issues (Scenario 1, Scenario 2 and Scenario 3), the PPMEH-FT achieves the best performance, which evidences that the incorporation of DFT is capable of enhancing the robustness against image random rotation, in contrast to the basic PPMEH. For image random scaling issues (Scenario 4 and Scenario 5), GF is the best for Scenario 4 and GLCM is the best for Scenario 5. Although PPMEH and PPMEH-FT cannot outperform the GF and GLCM, PPMEH still achieves a competitive result. For the concurrent occurrence of image random rotation and scaling (Scenario 6, Scenario 7 and Scenario 8), the PPMEH-FT achieves the best performance. The simultaneous occurrence of image rotation and scaling is a more challenging circumstance than others. Such a hybrid issue is quite often encountered in medical image classification practice. As demonstrated in Table 4.2, the proposed PPMEH-FT is more robust than the basic PPMEH when dealing with the concurrent occurrence of image rotation and scaling.

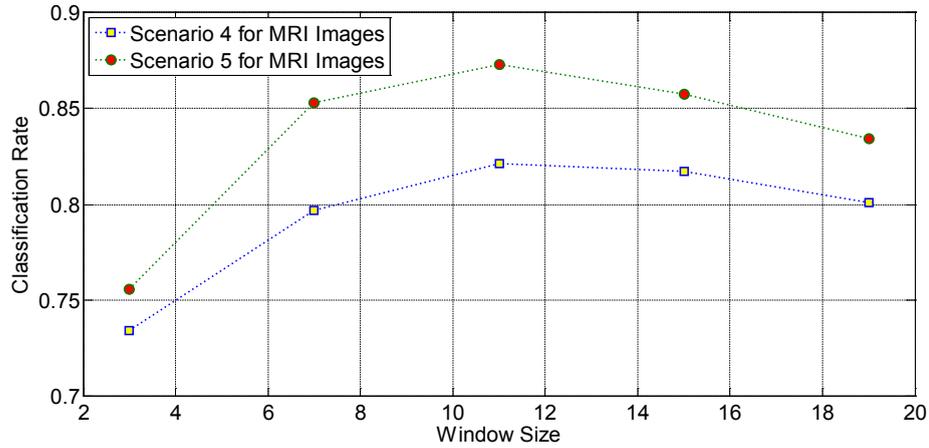


Figure 4.11: The illustration of sliding window size selection to cope with MRI image random scaling in Scenario 4 and Scenario 5, respectively, where the classification accuracy of PPMEH is adopted as the selection criterion

4.3.4 Experiment 3: Categorization of Given MRI Samples to Correct Head Scan Group

In medical practice, given a patient head MRI scan, several samples stored in a large-scale gallery need to be retrieved for case study and comparison by a radiologist. Such retrieval is conducted depending on image content similarity. In this experiment, the image retrieval problem for head MRI scans can be regarded as a classification problem. To be explicit, the features (PPMEH and PPMEH-FT) corresponding to MRI scans in background database are generated and scattered in the feature space. Furthermore, the given query MRI scan is categorized into the correct head MRI group (such group also contains the given query scan) via K-nearest neighbor, according to feature-based similarity. Then, the assigned group label is verified, since the real category label for the given query MRI scan is already known. This experiment is performed by following the same setting presented in Subsection 4.3.3; that is, the parameters to yield the proposed image features, the scenarios to simulate diverse imaging circumstances, and the experiment setup are all maintained in the same way (as shown in Figure 4.11, sliding window size is tuned to 11×11). The obtained statistics regarding the classification performance and the comparison among the different state-of-the-art approaches are demonstrated in Table 4.3.

Similar to the results obtained in the previous experiment, the PPMEH-FT achieves the best performance in terms of random rotation issues (Scenario 1, Scenario 2 and Scenario 3), and hybrid random rotation and scaling issues (Scenario 6, Scenario 7 and

Table 4.3: Image Feature Comparison for Image Random Rotation and Scaling in Head MRI Images Classification Experiment (Classification Rate: $\mu \pm \sigma$)

	Scenario 1 Rotation Only [$-180^\circ, 180^\circ$]	Scenario 2 Rotation Only [$-20^\circ, 20^\circ$]	Scenario 3 Rotation Only [$-10^\circ, 10^\circ$]	Scenario 4 Scaling Only [0.8, 1.2]
LBP[208]	37.95% \pm 5.00%	41.93% \pm 5.22%	53.00% \pm 5.48%	53.00% \pm 5.92%
GF[205]	24.74% \pm 3.88%	40.45% \pm 5.10%	47.50% \pm 5.58%	76.45% \pm 5.34%
GLCM[211]	32.62% \pm 4.29%	37.90% \pm 4.60%	40.62% \pm 5.18%	40.71% \pm 3.86%
SV[231]	24.86% \pm 3.08%	28.79% \pm 3.85%	30.57% \pm 3.40%	36.95% \pm 4.76%
HOG[229]	40.23% \pm 4.98%	47.14% \pm 5.60%	48.57% \pm 3.56%	49.76% \pm 5.14%
PPMEH	54.83% \pm 4.72%	66.57% \pm 4.85%	75.24% \pm 3.11%	82.14% \pm 4.94%
PPMEH-FT	60.23% \pm 3.73%	75.00% \pm 4.12%	80.35% \pm 3.77%	81.66% \pm 3.08%
	Scenario 5 Scaling Only [0.9, 1.1]	Scenario 6 Rotation [$-180^\circ, 180^\circ$] + Scaling [0.8, 1.2]	Scenario 7 Rotation [$-20^\circ, 20^\circ$] + Scaling [0.8, 1.2]	Scenario 8 Rotation [$-10^\circ, 10^\circ$] + Scaling [0.9, 1.1]
LBP[208]	64.14% \pm 5.14%	29.62% \pm 4.74%	27.33% \pm 4.25%	44.02% \pm 4.78%
GF[205]	79.43% \pm 5.06%	22.95% \pm 4.25%	38.98% \pm 4.75%	46.38% \pm 5.47%
GLCM[211]	46.86% \pm 5.27%	30.74% \pm 4.34%	31.76% \pm 4.33%	37.12% \pm 4.25%
SV[231]	42.93% \pm 5.17%	25.07% \pm 4.08%	33.64% \pm 4.48%	38.00% \pm 4.91%
HOG[229]	49.28% \pm 3.26%	35.71% \pm 3.24%	37.14% \pm 3.53%	45.23% \pm 3.58%
PPMEH	87.31% \pm 3.37%	42.45% \pm 4.42%	57.41% \pm 4.66%	70.48% \pm 4.04%
PPMEH-FT	84.64% \pm 0.0378	44.16% \pm 0.0439	59.52% \pm 0.0495	75.47% \pm 0.0508

Scenario 8). For the image random scaling issues (Scenario 4 and Scenario 5), PPMEH achieves the best performance while PPMEH-FT and GF are ranked second and third, respectively.

4.3.5 Experiment Results Analysis

The LBP represents an image by statistics of the local pixel intensity variations in a binary string framework. These local pixel intensity variations correspond to the image local structures such as lines, spots, ridge, weak edge and even noise. However, the LBP operator is unable to distinguish the salient local structures from the trivial ones. Consequently, the LBP carries out an undifferentiated treatment to all local patterns including salient and trivial ones. Therefore, the redundant information corresponding to the trivial local patterns is still preserved and involved in the LBP histogram, which leads to the degradation of the classification performance. On the contrary, the PPMEH and PPMEH-FT focus on the salient local structures instead of the trivial ones; that is, the trivial details are filtered out in the proposed local structure detection phase.

The Gabor filter bank is predefined based on the multi-orientations and scales. Compared with the PPMEH and PPMEH-FT, the GF kernel bank only depends on the limited directional filters to produce the limited directional filter responses. Therefore, the GF cannot compete with PPMEH and PPMEH-FT in the case of image random rotation, because PPMEH and PPMEH-FT involve more spatial orientation information than the GF. For the image scaling issue, the robustness of PPMEH and PPMEH-FT is equivalent to that of the GF. Although different from the conventionally predefined multi-scale strategy, the local self-adaptation mechanism in the PPMEH and PPMEH-FT framework is still capable of coping with the local structure variation caused by image scaling.

The GLCM is a tabulation of how often the different combinations of pairwise pixel intensities are presented in a given image. In order to represent image texture, several statistics such as contrast, correlation, energy and homogeneity are calculated based on the obtained GLCM. The GLCM usually covers four fixed directions 0° , 45° , 90° and 135° , which provide less orientation information than the PPMEH and PPMEH-FT schemes. Therefore, the GLCM is not able to compete with the PPMEH and PPMEH-FT in respect of image random rotation. For the image random scaling issue, the GLCM involves four predefined lags in every direction to deal with the image local scale variation. This is a typical multi-scale strategy and cannot work well if the predefined lags do not appropriately match the real local scales, or the predefined lags are limited.

The SV is a spatial autocorrelation function to represent image spatial correlation. Either image rotation or scaling can change SV functions. Accordingly, the discrimination of SV features is degraded. Even under minor rotation or minor scaling, SV features are significantly varied. This is evidenced by the poor classification performance in Scenario 3, Scenario 5 and Scenario 8 in Table 4.2 and Table 4.3.

The HOG adopts gradient information to depict local image structures, which lies on the same foundation as PPMEH and PPMEH-FT. Being different from PPMEH and PPMEH-FT, however, HOG suffers from the following two technical defects: (i) similar to LBP, HOG involves some trivial spatial information and accordingly, the formed histogram is over-informative. For medical image classification practice, involvement of complete spatial attributes is not always helpful to reinforce the feature discrimination,

but the salient ones are undoubtedly worthy of serious consideration. Hence, the information redundancy leads to the degradation of the HOG classification performance; (ii) gradient is very sensitive to random noise. Since the majority of MRI images contain random noise, HOG performs comparably poorly in the MRI classification experiment.

Most of the head MRI scans are corrupted by real random noise (shown in Figure 4.3). The image features involved in this study have different noise robustness. The LBP treats noise as a microstructure pattern and involves it in the binary string-based histogram statistics. Accordingly, its performance in such a noisy head MRI scan classification experiment is unsatisfactory. Gabor filter yields intensive response for local structures whose spatial frequency fittingly matches the predefined filter scale. However, the noise can be regarded as a microstructure with high spatial frequency. Therefore, the Gabor filter with fine scale generates prominent filter response for noise, and the classification performance consequently drops. The GLCM and SV are pairwise pixel-based statistical features. Under noise corruption, the image pixel intensities are randomly altered and such statistical features are accordingly changed. As a consequence, the classification performance is influenced. Regarding the proposed PPMEH and PPMEH-FT, their noise robustness can be ensured because of the following measures adopted in the local structure detection procedure: (i) the adoption of the noise-robust anisotropic Gaussian first-derivative kernel, not only for the local feature preservation but also for the noise resistance; and (ii) the utilization of the smoothing parameter for the noise suppression. As examples of noisy medical images, the superiority of the proposed robust local structure detection scheme is illustrated in Figure 4.12.

Table 4.2 and Table 4.3 demonstrate that there are classification performance disparities between PPMEH and PPMEH-FT: (i) for the scenarios involving image rotation, PPMEH-FT is always better than PPMEH; and (ii) for the scenarios only concerning image scaling, PPMEH outperforms PPMEH-FT. Such performance disparities can be attributed to the adoption of 1D DFT on PPMEH (PPMEH is treated as a 1D signal). DFT magnitude spectrum is able to capture the general pattern of original PPMEH signal on frequency domain, even in the presence of signal shift (PPMEH signal shift is caused by image rotation). Hence, PPMEH-FT can be more robust and achieve better classification performance than its original version PPMEH, when image rotation happens. However, DFT magnitude spectrum also lose some frequency details of original PPMEH signal, since finite Fourier frequency components are used for DFT. For

scaling-only cases, PPMEH which can preserve more signal details, consequently yield higher classification rate than PPMEH-FT.

4.4 Conclusion

In this chapter, a series of new concepts: image gradient force, particle potential motion probability, particle potential motion entropy and self-information are introduced, defined and interpreted, respectively. Based on such concepts, two robust image features, PPMEH and PPMEH-FT, are proposed to cope with the multi-effect of image rotation, scaling and noise in medical image classification tasks. The experimental results on colon CT and head MRI image datasets demonstrate that the presented PPMEH and PPMEH-FT provide generally satisfactory classification performance in contrast to other state-of-the-art image features. Besides, the experimental results suggest that the proposed PPMEH and PPMEH-FT can be applied to a medical image classification and retrieval system with large databases.

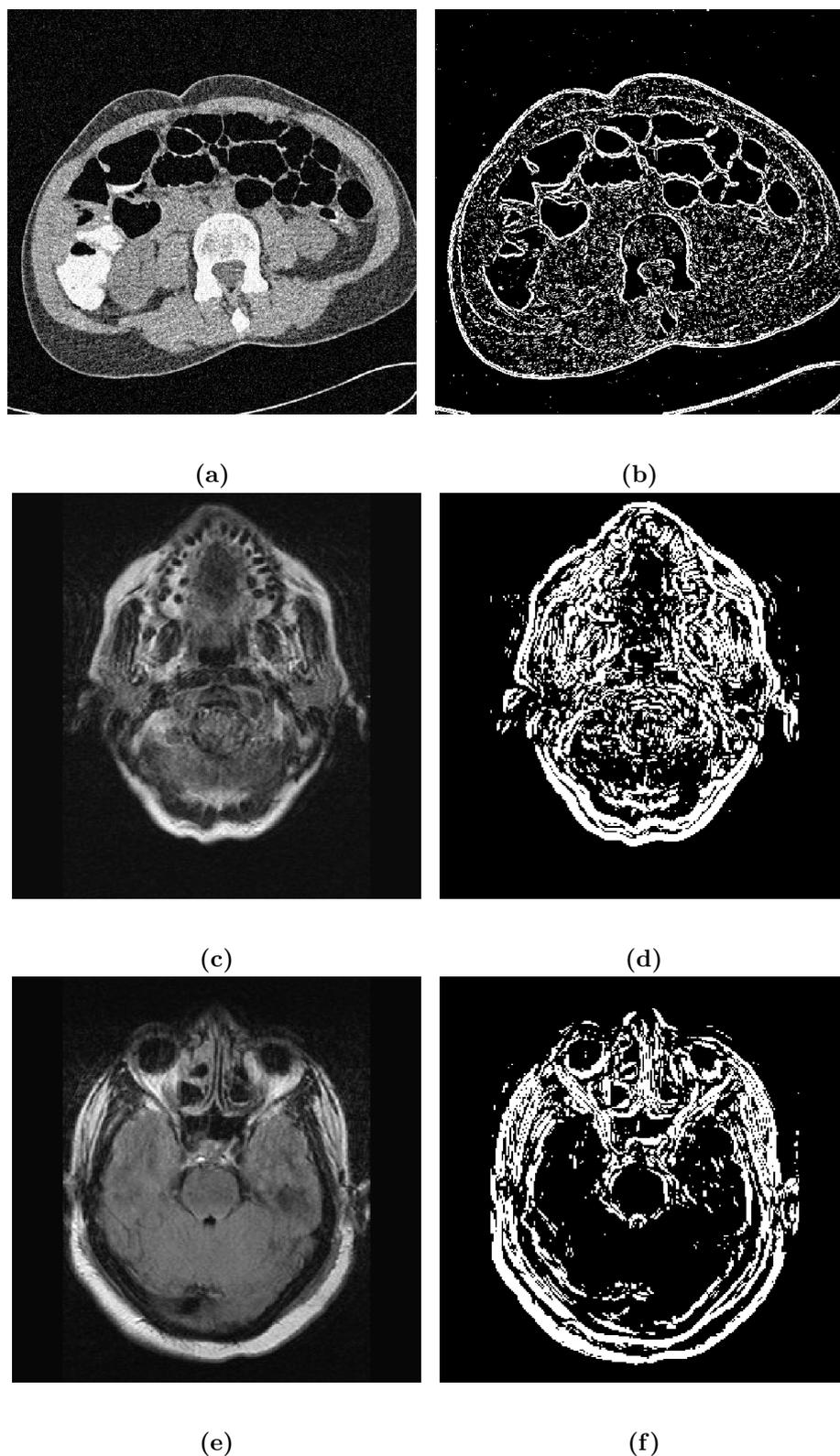


Figure 4.12: The illustration of the noise robustness of the proposed local structure detector: (a) the colon CT slice in Figure 4.1c degraded by the synthetic Gaussian noise ($\sigma_n = 30$); (c) the heads MRI scan in Figure 4.3a and (e) Figure 4.3b carry real noise; (b), (d) and (f) are the corresponding binary indicators obtained by the proposed local structure detector

Chapter 5

Robust Segmentation Module for Fully Automated Latent Fingerprint Matcher

5.1 Introduction

The pioneering study of fingerprint identification for distinguishing criminals can be traced back to [245]. The fingerprint inadvertently left by a person at a crime scene is capable of identifying a criminal or excluding a suspect. Fingerprints collected from crime scenes are compared to the fingerprints collected from suspects so that the fingerprints belonging to criminals can be identified [246]. As a consequence, an Automated Fingerprint Identification Systems (AFIS) was established and developed to satisfy such urgent need [247]. One important function of the AFIS system is to identify suspects against a large fingerprint database from an unknown fingerprint. There are two basic ways of searching. One approach is fingerprint indexing including classification, where the query fingerprint is mapped into a cluster with similar characteristics and such clusters will become the candidates for further inspection. The second approach is to perform matching on a one-to-one basis against the whole database. In principal, the first approach is more efficient. However, even though some progress has been made on the partial fingerprint indexing [248], few literatures have been found on latent fingerprint indexing. Existing latent fingerprint identification work is virtually on a one-to-one

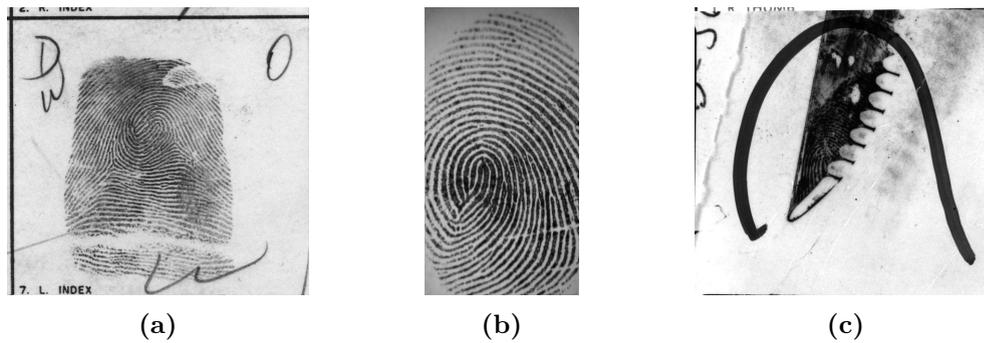


Figure 5.1: Three types of fingerprint: (a) the rolled print; (b) the plain print; and (c) the latent print

basis which is also our focus in this thesis. Therefore in the remainder of the chapter, our AFIS discussion is restricted to the category of one-to-one matching unless stated otherwise. AFIS is widely used to identify three main types of fingerprint: the rolled, the plain and the latent. The rolled fingerprint is a print which is obtained by rolling the finger from one side of the nail to the other side of the nail (namely, nail-to-nail) on a card or inside a platen scanner (shown in Figure 5.1a). The plain is a print collected by pressing the finger down on a card or placing the finger flat on a scanner (shown in Figure 5.1b). The latent fingerprint is acquired from a crime scene where the print is not intentionally left by a suspect or criminal (shown in Figure 5.1c). Both rolled and plain fingerprints are acquired in a controlled mode; that is, they are typically of good quality and are rich in reliable detailed features (e.g. minutiae). Consequently, AFIS is able to handle the rolled and plain identification cases in fully automatic mode. In contrast, the fingerprint in latent images is usually small-sized, overlapped with other image components and blurred due to the following possible causes:

- Small area: most fingerprints collected from crime scenes are not complete but partial (shown in Figure 5.2a);
- Overlapping with other structured components: fingerprints usually overlap with other structured noise such as arch, line, character, stain, and graphic pattern (shown in Figure 5.2b);
- Blur: most fingerprints acquired from crime scenes have large distortions due to the pressure variations when the fingers touch or press down on the object surface (shown in Figure 5.2c).

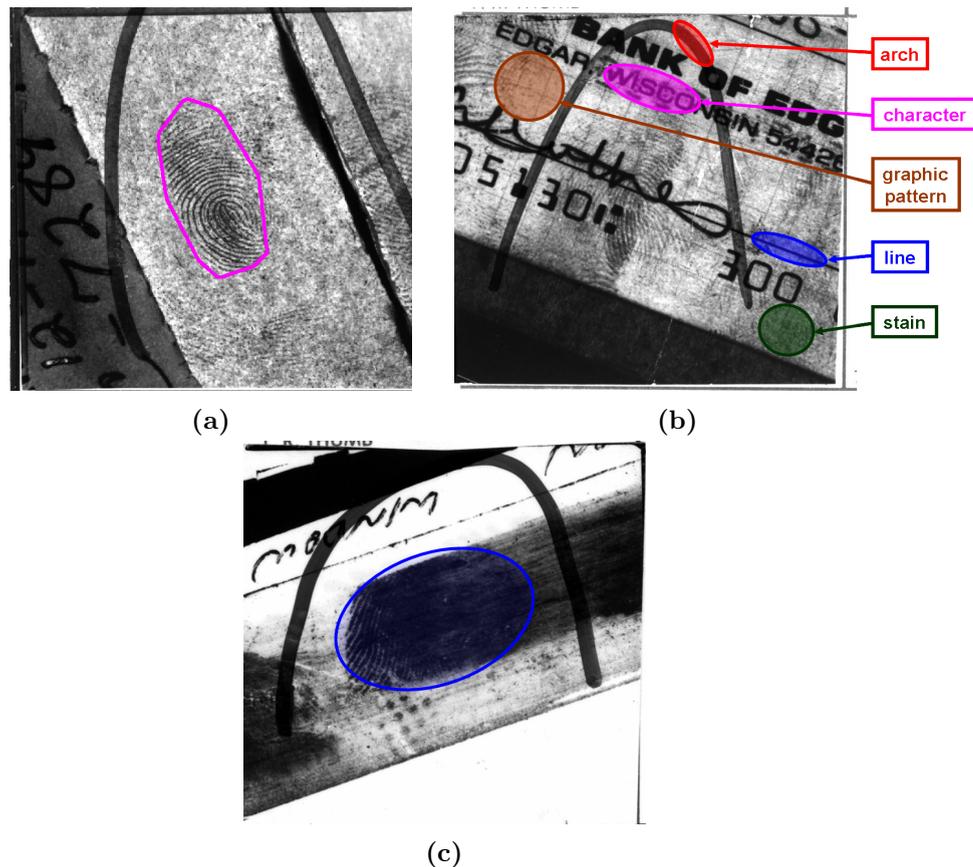


Figure 5.2: Challenges involved in latent print images: (a) small area in ROI; (b) overlapping with a variety of structured noise; and (c) blur

All of the above-mentioned adverse effects are not solely encountered but concurrently confronted, therefore latent fingerprint images are generally of poor quality. Considering that AFIS primarily depends on sufficient and reliable features, the latent fingerprint identification based on automatically extracted features is inaccurate; that is, the poor image quality imposes difficulties on the automated feature extraction so that the extracted features are limited and most of them are not reliable. Consequently, the fully automated matching conducted by AFIS based on the limited and unreliable features would lead to an inaccurate result. In order to ensure the reliability and accuracy of latent fingerprint identification, the good-quality region of interest (ROI) as well as the reliable features are often manually marked instead of automatically extracted. With human intervention involved, for latent fingerprint identification, the semi-automatic mode rather than the fully automatic mode is adopted.

Semi-automatic latent fingerprint identification procedure consists of the following four stages: (i) the ROI in latent images is manually labeled; (ii) based on the labeled ROI,

the features such as minutiae, singularity, ridge quality map, orientation field, ridge wavelength map, and skeleton are manually extracted [249]; (iii) the marked features are uploaded to a latent fingerprint matcher, then are automatically matched against the features derived from the rolled / plain fingerprints in the background database; and (vi) according to the matching scores, the candidate rolled / plain prints are retrieved, and the candidates are visually verified by latent examiners. After visual verification, the most likely archived fingerprint in the background database might be found. If not found, a new search routine with more cautious ROI labeling, feature extraction, and visual verification would be conducted to obtain the most likely candidate. For semi-automatic mode in latent fingerprint identification, the resultant accuracy is satisfactory [250].

In order to reduce the cost of expert work, a fully automated latent fingerprint matching and identification system is needed [253]. This is also desirable for the advancement of the technology [251, 252]. We propose a latent matching algorithm, without involving many manually marked features, which only uses the minutiae information provided by latent experts. Given that [249, 250] the most elaborately marked features (e.g. singularity, ridge quality map, orientation field, ridge wavelength map, and skeleton) made by humans are not considered and used in the proposed matching system, human intervention is remarkably reduced. However, the minutiae extraction still proceeds manually. In order to avoid the human intervention and achieve a high degree of automation, the development of a ROI segmentation method and minutiae set-based latent matcher could be another possible solution. Recently, the ROI detector and the minutiae level matching algorithm have been the subject of several studies.

Several approaches have been proposed to address the problem of ROI segmentation. An automated ROI segmentation technique is presented in [256]. For the proposed technique, however, the local orientation and spatial frequency are estimated by using a local pixel intensity projection which is sensitive to the variation of pixel intensity caused by structured noise. [258] generates an ideal “ridge-valley” pattern template and then uses the cross-correlation between a local image patch and the generated template to evaluate the local fingerprint quality. The frequency of an ideal “ridge-valley” pattern template is predefined according to the fixed empirical value, therefore, the generated template is not adaptive to the real local spatial frequency. A total variation (TV) model-based approach is proposed to handle the latent fingerprint segmentation task.

In this approach, the latent fingerprint image is decomposed into cartoon and texture layers and the ROI is detected based on the texture layer by using traditional segmentation methods [254, 255]. However, the proposed TV model incorporates the orientation field which is directly calculated from the original poor-quality latent image. Thus, the directional information used in the proposed TV model is not reliable. [257] presents an automated method based on orientation tensor and local ridge frequency to concurrently localize the ROI in latent images. However, the local ridge frequency directly estimated from local Fourier analysis on the original latent patch is sensitive to the presence of structured noise. [253] proposes a dictionary learning-based segmentation and enhancement method, where the multilayer ridge structure dictionaries from the coarse level to fine level are separately established by using a dictionary learning algorithm. Such an approach heavily relies on the learned dictionaries and the training patches being preselected from the good-quality rolled fingerprint images. It is demonstrated that a dictionary learned from the target image is preferable (target image means the image currently being processed), since such dictionary can be more adaptive to the target image [49]. However, the ridge structure dictionaries are not learned from the query latent images but from the rolled ones. Therefore, the potentially useful “ridge-valley” pattern in latent fingerprint images are not utilized but ignored in [253].

For the reported latent fingerprint matching algorithms, the manual markup of minutiae is regarded as a common practice in latent fingerprint identification cases. This is not only to ensure the matching accuracy but also to keep the latent matcher in proper working condition. [251, 252] only consider the manually marked minutiae as the sole input for the hough transform (HT)-based matcher. Due to the involvement of manually marked minutiae (ground truth), the proposed matcher could achieve a satisfactory matching result. However, the straightforward adoption of the minutiae extracted by automated computer programs (e.g. Verifinger SDK) on low-quality latent images is most likely to result in a poor matching performance. Because of poor quality and overlapping structured noise in latent images, a fair amount of spurious minutiae is possibly yielded via a fully automatic procedure. Consequently, the proposed latent matcher is not robust but vulnerable to the corruption caused by spurious minutiae. [249, 250] also propose the baseline matching algorithm which takes only the manually marked minutiae as the matcher input. Similar to [251, 252], the proposed matching approach is also sensitive to the presence of spurious minutiae. Apparently, the robustness and

tolerance of the minutiae-level matcher for spurious minutiae is therefore an important property and plays a critical role when fulfilling poor-quality fingerprint matching duties.

The work on a robust fingerprint matcher for rolled prints has been proposed [259]. In [259], a fingerprint matcher based on the genetic algorithm (GA) is developed to deal with the significant occlusion and clutter of minutiae caused by low-quality prints. This method achieves good performance when handling low-quality rolled prints. However, the matching performance for latent prints is still unknown. In addition, the proposed matcher heavily depends on the local minutiae triangle-based fitness function. Since the local triangle generation based on each triplet of minutiae is computationally intensive, such a type of fitness function is too inefficient for GA-based optimization to solve the large-size minutiae set matching problem.

The state-of-the-art latent fingerprint matcher is unavailable in public domain. Although one commercial off-the-shelf (COTS3) matcher has been assigned as a latent matcher for latent print matching task as well as performance evaluation, such matcher is not accessible to the researchers [253, 257]. Except COTS3, there is no fully-automated latent matcher which is publicly available. In this chapter, we propose a robust minutiae set-based matcher embedded with a self-learning module for ROI identification in latent fingerprint images. The proposed latent matcher integrates the following two modules: (i) the dictionary learning (DL)-based ROI segmentation scheme; and (ii) the GA-based minutiae set matching unit. For the DL-based ROI segmentation scheme, the dictionary is firstly learned from the query latent fingerprint image. Then, based on the learned dictionary, the “ridge-valley” pattern elements (dictionary atoms) can be automatically identified. Furthermore, the sparse representation for the original latent image patches is performed. Finally, depending on the presence or absence of the sparse coefficients that correspond to the identified “ridge-valley” atoms, the foreground (fingerprint region) is segmented. In the GA-based minutiae-level matching unit, the two minutiae sets, one from the segmented ROI in the query latent image (obtained via segmentation module) and the other from the print currently being compared, are extracted beforehand through a normal automated minutiae extraction program like Verifinger SDK, which is widely available to the public. Then, according to the affine transformation parameters estimated by GA, the minutiae set alignment between the query latent and the compared print is performed. Furthermore, after aligning two sets of minutiae, the correspondence between the two sets needs to be found. Accordingly, the corresponding

minutiae points between the query latent and the compared print can be paired. Finally, the number of matched minutiae is obtained and simply regarded as the matching score.

The main contributions of this chapter are summarized as follows:

- In this chapter, we introduce a multi-module matcher to cope with the latent fingerprint matching problem. The proposed system is performed in a fully automatic mode. Experimental results based on NIST SD27 demonstrate that the proposed matcher, with the proposed segmentation module (SM) (say, proposed matcher + proposed SM), can achieve a 33.692% penetration rate. Comparative experiments have been conducted to evaluate the effect of the different SMs by using the proposed matcher with and without SM. By designating the proposed matcher without SM as the baseline (say, proposed matcher only), the benchmark penetration rate is 38.728%. Based on the benchmark, the proposed matcher with the state-of-the-art SM such as [253] (say, proposed matcher + SM [253]) only achieves a 38.321% penetration rate. The relative penetration rate enhancement percentage for “proposed matcher + proposed SM” ($13.00\% = \frac{|33.692\% - 38.728\%|}{38.728\%}$) is at least tenfold better than that of “proposed matcher + SM [253]” ($1.05\% = \frac{|38.321\% - 38.728\%|}{38.728\%}$).
- The fully automated ROI segmentation module is plugged into the proposed latent matcher and is performed as the preprocessing procedure for the subsequent matching task. The proposed SM consists of the following phases: (i) the image structure dictionary learning; (ii) the “ridge-valley” atom identification; and (iii) the sparse coding and ROI segmentation. The existing method requires the establishment, in advance, of a dictionary from the high-quality rolled image patches [253]. In contrast to such conventional methods, we propose building up the structure dictionary directly learned from the query latent image. As demonstrated in [49], a learned dictionary based on the target image can better adapt to the target image. Therefore, the dictionary obtained in the proposed SM is not good-quality rolled image patches-determined, but rather, query latent image-oriented.
- The robust latent matching unit consists of the following stages: (i) the ROI-based minutiae extraction; (ii) the GA-based minutiae set alignment; and (iii) the counting of paired minutiae. Existing method demands yielding of local minutiae descriptors during iteration of GA optimization [259]. In contrast to such conventional methods, the global topology of the entire minutiae set is directly adopted in

the proposed matching unit instead of the local minutiae structure. Accordingly, the proposed matching unit is more robust in the presence of spurious minutiae and more efficient in the matching.

The rest of this chapter is organized as follows: in Section 5.2, the details of the proposed matching system are introduced; in Section 5.3, ROI-based minutiae extraction, and latent fingerprint matching experiments are implemented, respectively. In the ROI-based minutiae extraction experiment, by adopting the obtained ROI in the automated segmentation module, the reduction of spurious minutiae points as well as the preservation of genuine ones are assessed. In the latent fingerprint matching experiment, the matching performance of the introduced latent matcher is evaluated; in Section 5.4, the conclusion is presented.

5.2 Proposed Latent Fingerprint Matching System

In this section, the proposed multi-module matching system for latent fingerprint is proposed, which consists of the two following modules: (i) a dictionary learning-based ROI segmentation scheme; and (ii) a genetic algorithm (GA)-based minutiae set matching unit.

5.2.1 Dictionary Learning-Based ROI Segmentation Module

5.2.1.1 Query Latent Image-Based Dictionary Learning

The dictionary learning procedure is performed based on the query latent fingerprint image instead of the good-quality rolled or plain print images. As suggested in [49], a learned dictionary based on the target image can better adapt to the target image and more specifically represent the intrinsic signal structure in the target image. Compared with other image-based learned dictionary, the target image-based dictionary can more effectively keep the scale consistency for the real signal structures; that is, the signal structure scales in other images might be more, or less, different from those in the target image. Therefore, the other image-based dictionary atoms often do not fittingly model the real-scale structures in the target image. Accordingly, the image restoration

tasks such as denoising and inpainting are conducted according to the target image-based dictionary rather than the other image-based dictionary or the predefined analytic dictionary (e.g. wavelets, curvelets and DCT basis) [49]. Inspired by the advantage of the target image-based dictionary, the dictionary learned from the query latent image is beneficial to the following “ridge-valley” atom identification phase.

The dictionary learning for the query latent image can be mathematically formulated as follows: let $\mathbf{S} = \{s_i | i = 1, 2, 3, \dots, N\}$ as the training set, where s_i is the vector obtained after the vectorization of the latent image block p_i with size $w \times w$ and N is the number of selected image blocks in the latent fingerprint image. In this study, the image blocks centered at all pixels are vectorized to form training set \mathbf{S} ($N = H \times W$, where H and W denote image height and width, respectively). The purpose of dictionary learning is to establish a numerical dictionary \mathbf{D} with size $N_s \times N_a$, based on the training set \mathbf{S} (N_s , denoting the atom vector dimension, where $N_s = w \times w$ and N_a stands for the atom number). Such established dictionary \mathbf{D} can effectively represent the vectorized image block s_i in a sparse way. In order to obtain dictionary \mathbf{D} , the following optimization problem needs to be solved

$$\min_{\mathbf{D}, \Gamma} \|\mathbf{S} - \mathbf{D}\Gamma\|_F^2 \quad s.t. \quad \forall i, \quad \|\gamma_i\|_0 \leq K \quad (5.1)$$

where Γ is the sparse coefficient matrix with size $N_a \times N$. γ_i is the i^{th} column vector in sparse coefficient matrix Γ and also corresponds to the i^{th} vectorized image block s_i . $\|\cdot\|_F$ and $\|\cdot\|_0$ denote the Frobenius norm and the l_0 norm, respectively. $\|\gamma_i\|_0$ is equal to the number of non-zero elements in vector γ_i . Therefore, in order to obtain the sparse vector γ_i , the regulator $\|\gamma_i\|_0 \leq K$ ensures that the non-zero elements in vector γ_i are less than the sparsity parameter K . Although l_0 is the limit of l_p when p approaches zero, it is not a true norm (unlike the l_1 norm which has all the properties of a true norm) and also leads to the NP-hard problem. In order to avoid the NP-hard problem, the above l_0 norm-based regulator $\|\gamma_i\|_0 \leq K$ in Equation 5.1 can be replaced by the convex l_1 norm-based regulator $\min(\|\gamma_i\|_1)$. Accordingly, Equation 5.1 is updated in the following approximate form

$$\min_{\mathbf{D}, \Gamma} \left(\|\mathbf{S} - \mathbf{D}\Gamma\|_2^2 + \lambda \sum_{i=1}^N \|\gamma_i\|_1 \right) \quad (5.2)$$

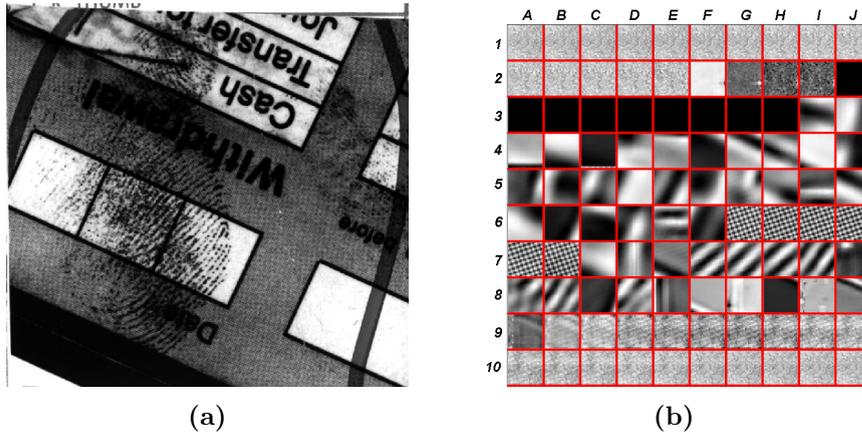


Figure 5.3: The dictionary \mathbf{D} is learned directly from the given query latent image by ODL: (a) original query latent image (U224 in NIST SD27); and (b) its learned dictionary \mathbf{D} ($N_a = 100$ and $K = 2$)

where Frobenius norm $\|\cdot\|_F$ has been embodied by the l_2 norm and λ is the Lagrange coefficient to balance the data fitting and sparsity level. For solving the optimization problem in Equation 5.2, the sparse coefficient matrix Γ and the dictionary \mathbf{D} are alternatively updated as follows: (i) keeping \mathbf{D} fixed, compute the sparse coefficient matrix Γ ; then (ii) keeping Γ fixed, update the dictionary \mathbf{D} . The above two steps are iteratively repeated until convergence is reached. In order to learn a dictionary \mathbf{D} , one can apply any available dictionary learning technique such as MOD [260], KSVD [49] or ODL [262]. As reported in [262], an alternate optimization of sparse coding and dictionary update is performed based on a subset of the training data. Such subset continues to be augmented with a new training sample. Based on the outcome of the previous iteration, the same alternate optimization is executed again for new training data. The ODL repeats until all training samples have been used. As demonstrated by the experiments in [262], ODL is faster than MOD or KSVD. Because the large size of the training set is produced based on the dense training set pick-up strategy and the whole training set has to be used by MOD and KSVD at each iteration, both MOD and KSVD are computationally expensive. Therefore, instead of MOD and KSVD, ODL is applied to learn the dictionary for query latent fingerprint image due to its more efficient learning mechanism and lower computational complexity. The dictionary \mathbf{D} learned from the given query latent image by using ODL is shown in Figure 5.3.

5.2.1.2 Ridge-Valley Atom Identification

The atoms with a “ridge-valley” pattern need to be distinguished from the ones which are not structured as a “ridge-valley” pattern. Detection or modeling of the “ridge-valley” pattern is not a new subject and the previous works have been reported in [256, 258], however, it is still a challenging task. Although the “ridge-valley” pattern is one type of intrinsic signal structure in latent fingerprint image, such a pattern is usually mingled with other types of structured noise (e.g. arch, line, character, stain, speckle and motif) and even vague or disconnected due to the wet or dry prints. Accordingly, the direct detection or straightforward modeling for such a pattern based on the original latent image patch is inaccurate. In contrast to the original patch-level detection, atom-level detection is easier and more reliable; that is, the dictionary learning procedure is not only for signal structural decomposition but also for signal structure refinement. After completion of dictionary learning, the intrinsic signal structures at atom level become more salient and easier to be distinguished. Because the identified “ridge-valley” atoms play a crucial role in the subsequent sparse coefficient-based ROI segmentation phase, development of a fully automated identification approach is necessary.

In this study, a local Fourier analysis and cross-correlation-based method is proposed, which consists of the following steps:

- *Step 1:* Given a learned dictionary $D = \{d_k | k = 1, 2, \dots, N_a\}$, each single atom vector d_k with dimension N_s needs to be converted to the atom patch p_{d_k} with size $w \times w$ ($N_s = w \times w$) (shown in Figure 5.4a);
- *Step 2:* Input an atom patch p_{d_k} , the 2D Discrete Fourier Transform (2D DFT) is applied to p_{d_k} , then the spectrum $DFT(p_{d_k})$ in the frequency domain is obtained (shown in Figure 5.4b);
- *Step 3:* In the spectral magnitude map $|DFT(p_{d_k})|$ (rearranged by moving the zero-frequency component to the center), the paired highest magnitude points are detected within the frequency range corresponding to the ridge period range [3 pixels, 20 pixels]. Such ridge period range can cover most image underlying structures from low spatial frequency to high spatial frequency. The bandpass constraint in the frequency domain corresponding to the ridge period range [3 pixels,

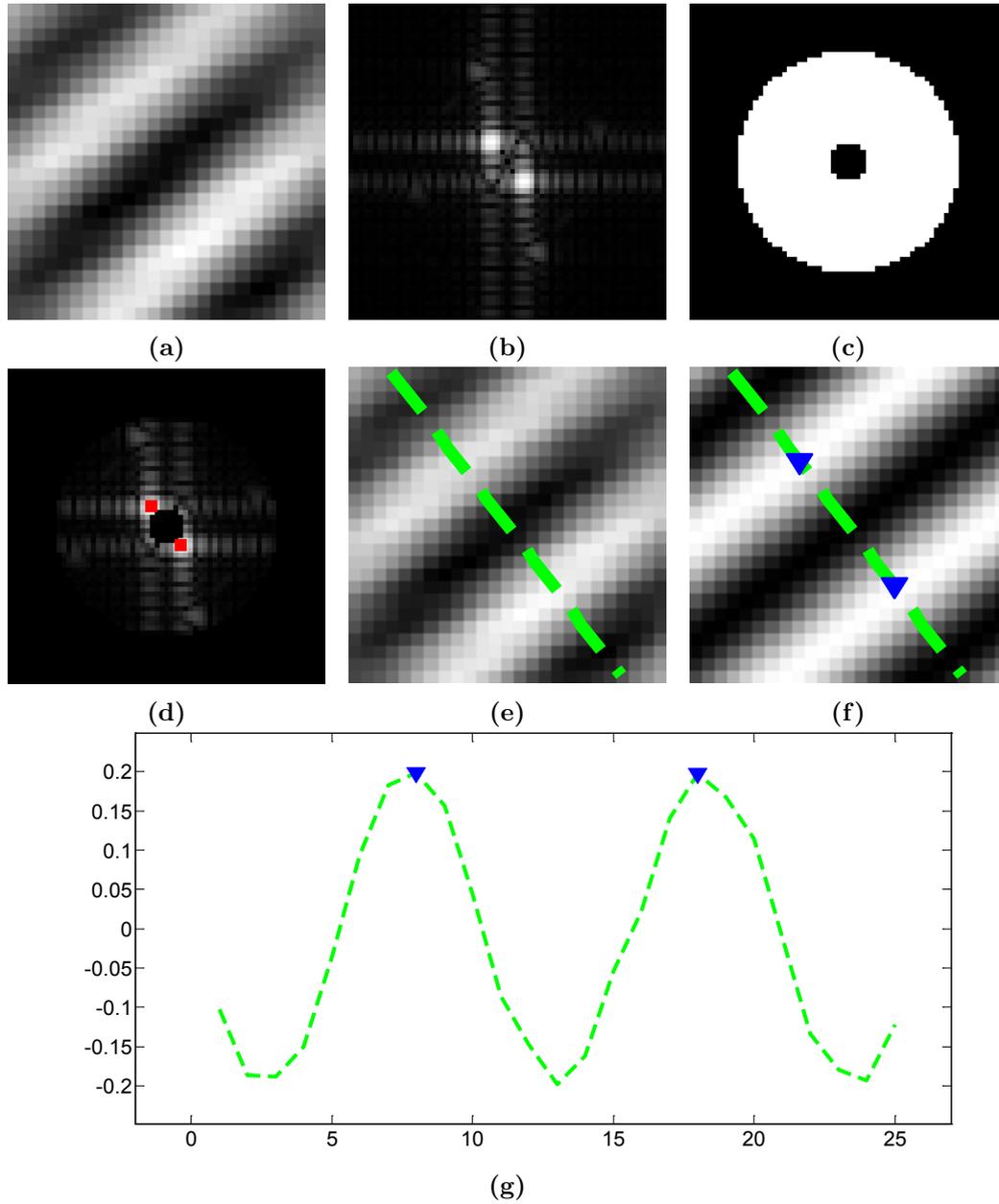


Figure 5.4: The illustration of the details during the automated “ridge-valley” atom identification procedure: (a) the “ridge-valley” atom patch ($p_{d_{66}}$ in Figure 5.3b); (b) the spectral magnitude map $|DFT(p_{d_{66}})|$; (c) the frequency bandpass filter corresponding to the ridge period range $[3 \text{ pixels}, 20 \text{ pixels}]$; (d) the constrained spectral magnitude map depending on (c); (e) the calculated orientation $o_{d_{66}}$; (f) the reconstructed “ridge-valley” pattern $u_{d_{66}}$; and (g) the 1D sinusoidal-shaped wave modeled based on the pixel intensities along the green dashed line indicated in (f)

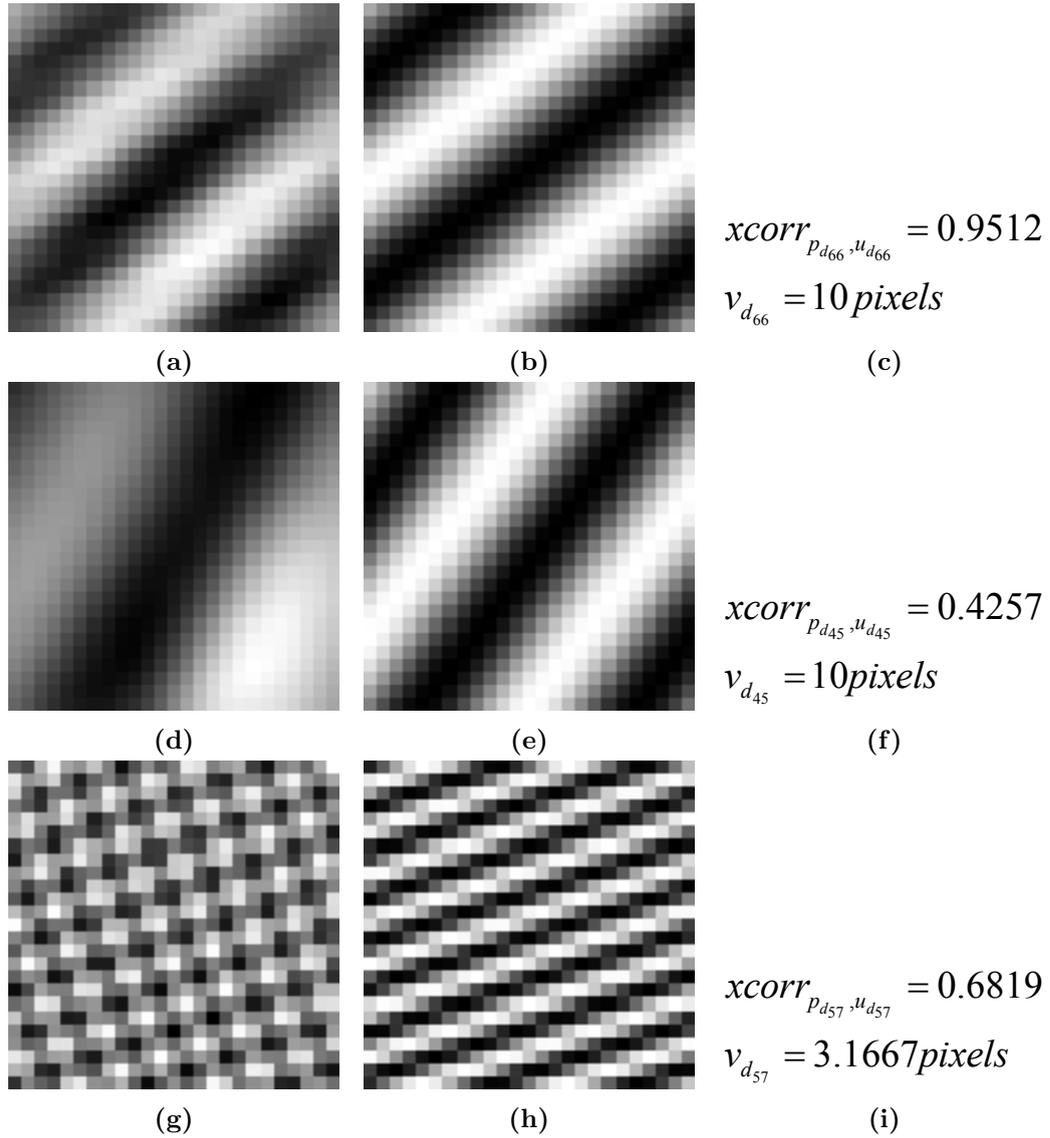


Figure 5.5: The examples of atom-based identification for various atoms: (a) the “ridge-valley” atom ($p_{d_{66}}$ in Figure 5.3b); (b) and (c) are the reconstructed $u_{d_{66}}$, the calculated $xcorr_{p_{d_{66}}, u_{d_{66}}}$ and $v_{d_{66}}$ for (a), respectively; (d) the line-like atom ($p_{d_{45}}$ in Figure 5.3b); (e) and (f) are the reconstructed $u_{d_{45}}$, the calculated $xcorr_{p_{d_{45}}, u_{d_{45}}}$ and $v_{d_{45}}$ for (d), respectively; (g) the periodical speckle atom ($p_{d_{57}}$ in Figure 5.3b); (h) and (i) are the reconstructed $u_{d_{57}}$, the calculated $xcorr_{p_{d_{57}}, u_{d_{57}}}$ and $v_{d_{57}}$ for (g), respectively

20 pixels] in the spatial domain is shown in Figure 5.4c. Such a frequency band-pass filter masks the spectral magnitude map to yield the constrained spectral magnitude map, and the highest magnitude points can be detected and shown in Figure 5.4d (marked by red squares).

- *Step 4:* Based on the coordinate of paired points detected in *Step 3*, the orientation o_{d_k} for the currently processed atom patch p_{d_k} is calculated (for “ridge-valley”

atom, the orientation o_{d_k} is not along but across the ridge, as indicated by the green dashed line in Figure 5.4e);

- *Step 5:* Reconstruct the “ridge-valley” pattern u_{d_k} according to the magnitude and phase corresponding to the detected points in the constrained spectral magnitude map (the reconstructed u_{d_k} is illustrated in Figure 5.4f), then calculate the cross-correlation value $xcorr_{p_{d_k}, u_{d_k}}$ between the atom patch p_{d_k} and the reconstructed “ridge-valley” pattern u_{d_k} , according to Equation 5.3.

$$xcorr_{p_{d_k}, u_{d_k}} = \frac{\sum_{x,y} (\alpha \cdot \beta)}{\sqrt{\sum_{x,y} (\alpha^2) \cdot \sum_{x,y} (\beta^2)}} \quad (5.3)$$

where $\alpha = f_{p_{d_k}}(x, y) - \bar{f}_{p_{d_k}, u_{d_k}}$ and $\beta = f_{u_{d_k}}(x - a, y - b) - \bar{f}_{u_{d_k}}$. $f_{p_{d_k}}(x, y)$ and $f_{u_{d_k}}(x, y)$ denote the atom image p_{d_k} and the reconstructed pattern u_{d_k} , respectively. $\bar{f}_{p_{d_k}, u_{d_k}}$ stands for the mean of $f_{p_{d_k}}(x, y)$ when overlapping with $f_{u_{d_k}}(x, y)$. $\bar{f}_{u_{d_k}}$ is the mean of the entire reconstructed pattern image u_{d_k} . a and b denote the offsets along the x - and the y - axis, respectively. Equation 5.3 indicates the pattern similarity between the atom patch p_{d_k} and the “ridge-valley” pattern u_{d_k} . The higher cross-correlation value is that the atom patch p_{d_k} is more likely to represent the “ridge-valley” pattern.

- *Step 6:* For the reconstructed “ridge-valley” pattern u_{d_k} , the pixel intensities along the orientation o_{d_k} are acquired to model a 1D sinusoidal-shaped wave, then the wave peaks are detected to calculate the spatial ridge period v_{d_k} . The pixels along o_{d_k} are highlighted by the green dashed line in Figure 5.4f, and the modeled 1D sinusoidal-shaped wave is shown in Figure 5.4g where the blue triangles indicate the detected wave peaks;
- *Step 7:* Check whether the following two criteria are concurrently satisfied: (i) $xcorr_{p_{d_k}, u_{d_k}} \geq Th_{xcorr}$ (Th_{xcorr} is the threshold to determine the pattern similarity and $Th_{xcorr} = 0.6$ is empirically tuned after several trials); and (ii) $v_{d_k} \in [5.3 \text{ pixels}, 12.8 \text{ pixels}]$ ($[5.3 \text{ pixels}, 12.8 \text{ pixels}]$ is suggested in [257]. In the proposed atom identification procedure, the broad range $[3 \text{ pixels}, 20 \text{ pixels}]$ are firstly adopted to involve more potential candidate atoms, then the narrow range $[5.3 \text{ pixels}, 12.8 \text{ pixels}]$ are applied to filter out the unqualified atoms). If both are satisfied, the currently processed atom patch p_{d_k} is regarded as the “ridge-valley”

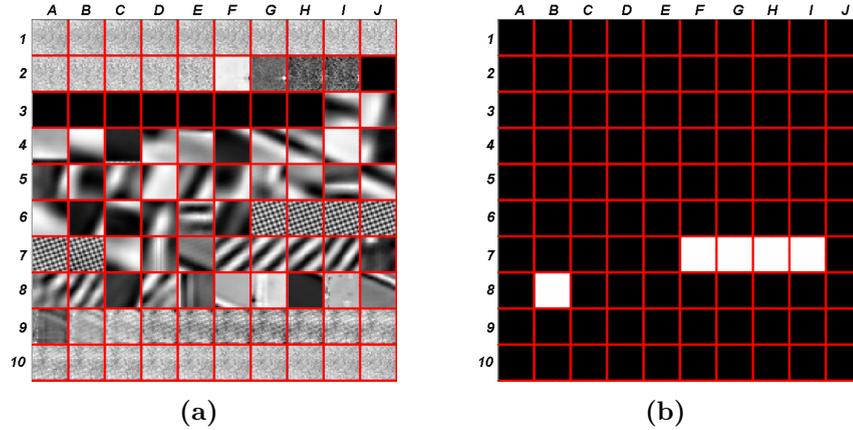


Figure 5.6: The example of the identification results obtained by using the proposed automated “ridge-valley” atom identification procedure: (a) the learned dictionary D in Figure 5.3b; and (b) the identified “ridge-valley” atoms. The white patches indicate that the original atoms with the same patch-wise coordinates in (a) are “ridge-valley” ones, while the black patches label of the corresponding atoms in (a) are “non-ridge-valley” ones

atom. Otherwise, label p_{d_k} is the “non ridge-valley” atom. Judgement depending on a single criterion is too limited to make a correct decision on atom level, therefore, both criteria need to be simultaneously satisfied. As demonstrated in Figure 5.5, the “non ridge-valley” atoms either satisfy $xcorr_{p_{d_k}, u_{d_k}} \geq Th_{xcorr}$ (here $Th_{xcorr} = 0.6$) or meet $v_{d_k} \in [5.3 \text{ pixels}, 12.8 \text{ pixels}]$, however, they are not the true “ridge-valley” atoms;

- *Step 8:* Check whether all the atoms in D have been judged. If not, go back to *Step 2*. Otherwise, terminate the “ridge-valley” atom identification procedure and output all the atom labels. An example of the “ridge-valley” atoms identification for all the atoms in the learned dictionary is illustrated in Figure 5.6. In Figure 5.6b, the white patches indicate that the original atoms with the same patch-wise coordinates in Figure 5.6a are “ridge-valley” ones, while the black patches label that the corresponding atoms in Figure 5.6a are “non-ridge-valley” ones.

5.2.1.3 Sparse Coefficient-Based ROI Segmentation

The sparse coefficients, by projecting the image blocks in the query latent image on to the learned dictionary, are utilized to determine whether the original image blocks belong to the foreground. To be specific, the sparse projection from every single latent image block to the learned dictionary yields the sparse coefficient vector. Given a single

latent image block, the elements inside its sparse coefficient vector are quite different; that is, most of the elements are zero while a few are non-zero. The non-zero elements in the sparse coefficient vector measures the similarity between such a latent image block and the corresponding atoms. Among the non-zero elements, the higher one indicates that the corresponding atom is more similar to such an image block, while the lower one indicates lower similarity. As a latent image block with a “ridge-valley” pattern is considered as one part of a real fingerprint, the latent image block whose highest non-zero element in the sparse coefficient vector corresponds to the “ridge-valley” atom is regarded as the subset of the foreground. Otherwise, the image block whose highest non-zero element in the sparse coefficient vector corresponds to the “non-ridge-valley” atom is regarded as the background. Motivated by the presence or absence of the highest non-zero sparse coefficients corresponding to the identified “ridge-valley” atoms in the preceding stage, the ROI segmentation phase can proceed smoothly.

The generation of sparse coefficients can be mathematically formulated as the following equation

$$\tilde{s}_i = \gamma_i^{(1)} d_1 + \gamma_i^{(2)} d_2 + \gamma_i^{(3)} d_3 + \dots + \gamma_i^{(N_a)} d_{N_a} = \sum_{k=1}^{N_a} \gamma_i^{(k)} d_k \quad (5.4)$$

where \tilde{s}_i is the approximation of the given signal vector s_i (s_i is obtained after the vectorization for the latent image block p_i). $\gamma_i^{(k)}$ is the element inside the sparse coefficient vector $\gamma_i = [\gamma_i^{(1)}, \gamma_i^{(2)}, \gamma_i^{(3)}, \dots, \gamma_i^{(k)}, \dots, \gamma_i^{(N_a)}]^T$. Due to the sparsity constraint in Equation 5.1, most atoms in dictionary \mathbf{D} are not selected for the representation of s_i , but only a small number of atoms are adopted. Accordingly, most elements in vector γ_i are zero while a few are non-zero.

In this study, those non-zero elements inside the sparse coefficient vector γ_i are used for ROI segmentation. Such a sparse coefficient-based ROI segmentation stage consists of the following steps:

- *Step 1:* Initialize a vacant image \mathbf{M} (all the pixel intensities are zero) with the same size as the original latent fingerprint image, then divide the latent image into the overlapping blocks (block size $w \times w$, as shown in Figure 5.7b and Figure 5.7d) and rearrange the obtained image blocks into the column vectors (here the training

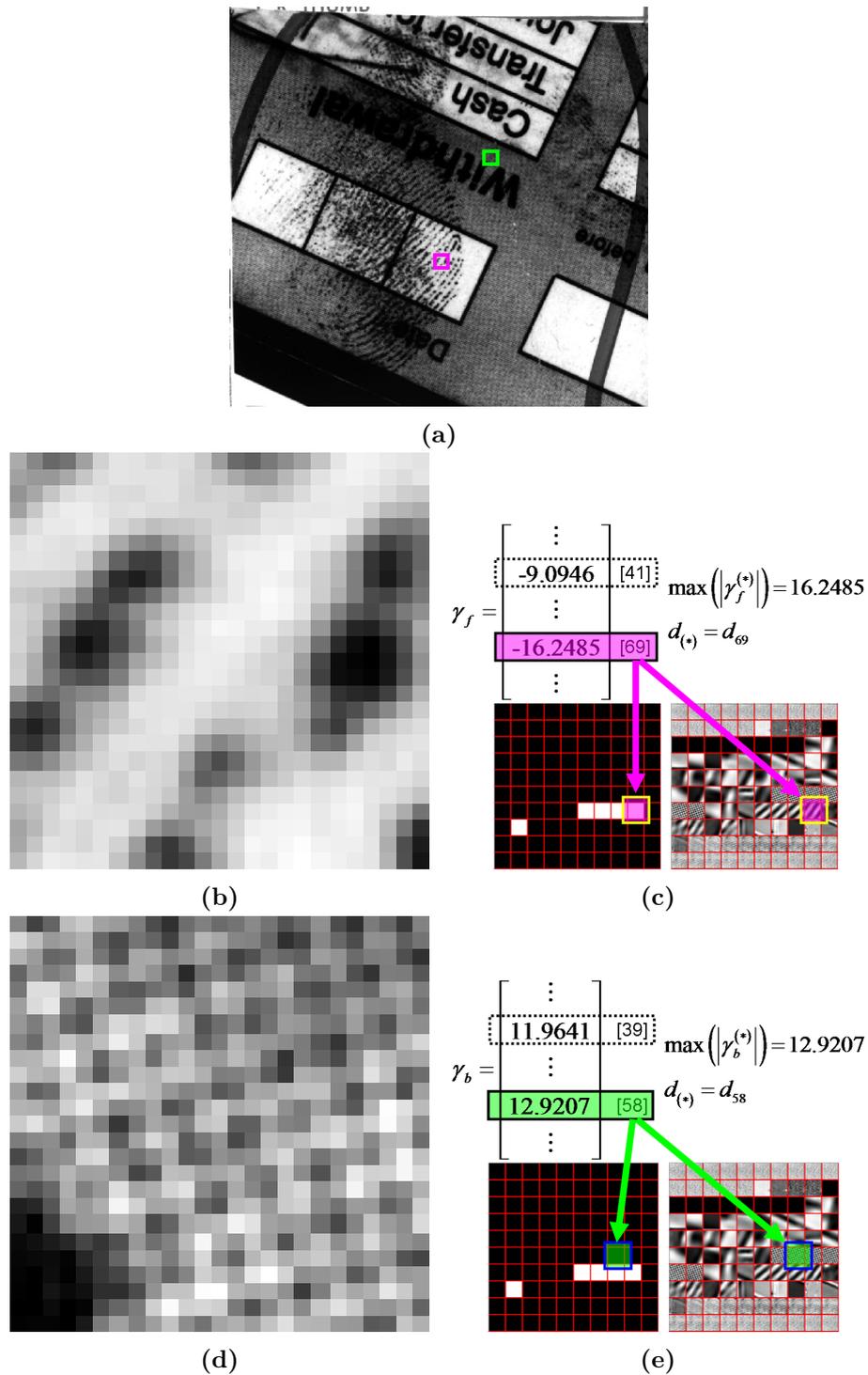


Figure 5.7: The illustration of the details during the automated sparse coefficient-based ROI segmentation procedure: (a) the two types of image block in the latent print image – the foreground block (marked by a pink square) and the background block (marked by a green square); (b) the marked foreground block; (c) seeking the specific atom corresponding to the $\max(|\gamma_f^{(*)}|)$ for (b) (the found atom is identified as the “ridge-valley” atom in the preceding phase); (d) the marked background block; and (e) seeking the specific atom corresponding to the $\max(|\gamma_b^{(*)}|)$ for (d) (the found atom is labelled as the “non-ridge-valley” atom in the preceding phase)

set $\mathbf{S} = \{s_i | i = 1, 2, 3, \dots, N\}$ for dictionary learning is directly used, where s_i is obtained after the vectorization for the image block p_i ;

- *Step 2:* Given a signal vector s_i , orthogonal matching pursuit (OMP) [263] is applied to the given vector s_i for selecting a few atoms to sparsely approximate s_i , then the sparse coefficient vector γ_i is obtained;
- *Step 3:* Find the highest non-zero absolute value $|\gamma_i^{(*)}|$ in the sparse coefficient vector γ_i and the corresponding atom $d_{(*)}$ in the dictionary \mathbf{D} (shown in Figure 5.7c and Figure 5.7e);
- *Step 4:* Check whether the found atom $d_{(*)}$ is the identified “ridge-valley” atom (shown in Figure 5.7c and Figure 5.7e). Then, the pixel intensities inside the block of \mathbf{M} , whose size and block-wise coordinate are the same as the currently processed block p_i in the latent image, are tuned according to Equation 5.5.

$$M_{p_i}(x, y) = \begin{cases} M_{p_i}(x, y) + 1, & d_{(*)} \text{ is "ridge - valley" atom} \\ M_{p_i}(x, y), & d_{(*)} \text{ is not "ridge - valley" atom} \end{cases} \quad (5.5)$$

where $M_{p_i}(x, y)$ denotes the pixel intensities inside the block of \mathbf{M} , whose size and block-wise coordinate are the same as the currently processed block p_i in the latent image.

- *Step 5:* Check whether all the signal vectors have been processed. If not, go back to *Step 2*. Otherwise, terminate the calculation of the sparse coefficient for the signal vector and output the image \mathbf{M} (shown in Figure 5.8a);
- *Step 6:* All the pixel intensities in image \mathbf{M} need to be normalized to the range $[0, 1]$, then the normalized image is binarized by Otsu’s adaptive threshold-based segmentation method [268] (shown in Figure 5.8b);
- *Step 7:* A set of mathematical morphology operators such as close, open, hole-filling, and small block removal are applied to the obtained binary image, and finally, the maximal-area convex polygon containing the foreground regions is achieved as the ROI (shown in Figure 5.8c).

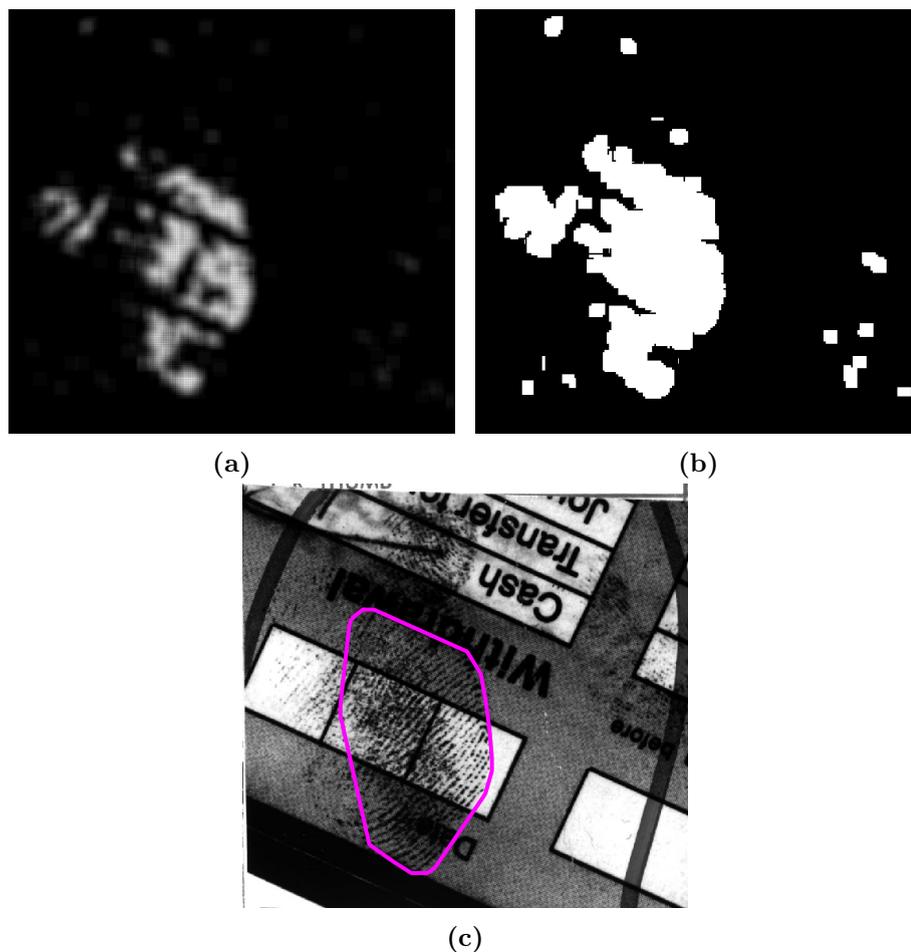


Figure 5.8: The example of the ROI obtained by using the proposed sparse coefficient-based ROI segmentation procedure: (a) the image M obtained after all the signal vectors have been processed; (b) the binary image obtained by using the normalization and Otsu's adaptive threshold-based segmentation method; and (c) the polygonal ROI obtained via the mathematical morphology operation and the maximal-area convex polygon extraction (marked by a pink convex polygon over the original latent print image U224 in NIST SD27)

5.2.2 Genetic Algorithm-Based Latent Matching Unit

5.2.2.1 ROI-Based Minutiae Extraction

The minutiae sets for the latent fingerprint images are extracted from the ROI obtained in the segmentation module. As demonstrated in [249–252], the reliability of minutiae plays a crucial role in the performance of the latent matcher. Obviously, the involvement of spurious minutiae leads to the deterioration of the matching performance; that is, the more spurious minutiae involved, the poorer the matcher behaves. However, how to effectively reduce or even avoid the spurious minutiae in fully automatic mode becomes a significant issue. As the outcome of the segmentation scheme, the resultant

ROI is used for automated minutiae extraction which can not only effectively preserve the genuine minutiae but also significantly reduce the false ones; that is, the genuine minutiae are most likely to be detected within the ROI while the spurious ones caused by the structured noise in the background are effectively eliminated by the ROI mask. Therefore, the minutiae set obtained from the ROI in the latent image is supposed to be reliable.

5.2.2.2 Problem Formulation for Minutiae Set-Based Fingerprint Matching

The minutiae set-based fingerprint matching can be mathematically formulated as follows: let $\mathbf{C} = \{c_i | i = 1, 2, \dots, m\}$ and $\mathbf{L} = \{l_j | j = 1, 2, \dots, n\}$ be the two-point set in \mathbf{R}^2 space (\mathbf{C} stands for the minutiae set extracted from the currently compared print and \mathbf{L} denotes the minutiae set extracted from the ROI in the query latent print, respectively). Both sets determine whether there is a specific affine transformation $\mathbf{T} = (\theta, s, t_x, t_y)$ (θ denotes the rotation angle, s represents the scaling factor, and t_x and t_y are the offsets along the x - and the y - axis, respectively) that maps set \mathbf{C} on to or close to set \mathbf{L} to indicate a correspondence. Therefore, seeking the affine transformation as well as the correspondences (an exact one-to-one correspondence, or an approximate correspondence) between both point sets \mathbf{C} and \mathbf{L} are coupled point matching problems. Here, let (c_i, l_j) be one of the corresponding pairs under \mathbf{T} , and denote $c_i = (x_{c_i}, y_{c_i}, o_{c_i}, pt_{c_i})^T$ and $l_j = (x_{l_j}, y_{l_j}, o_{l_j}, pt_{l_j})^T$. (x_{c_i}, y_{c_i}) and (x_{l_j}, y_{l_j}) are corresponding coordinates, o_{c_i} , o_{l_j} , pt_{c_i} and pt_{l_j} denote the point orientation and type of c_i and l_j , respectively. The formula of affine transformation is denoted as follows

$$(x_{l_j}, y_{l_j}) = \mathbf{T} [(x_{c_i}, y_{c_i})] \quad (5.6)$$

$$\begin{bmatrix} x_{l_j} \\ y_{l_j} \end{bmatrix} = s \cdot \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x_{c_i} \\ y_{c_i} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (5.7)$$

Equation 5.7 indicates that the coupled point matching problem can be regarded as the parameter tuning problem; that is, with the optimal selection of the parameter, the corresponding affine transformation is generated. Given the optimal affine transformation,

set \mathbf{C} is mapped on to set \mathbf{L} . After transformation, most of the points in set \mathbf{C} being close to the points in set \mathbf{L} are regarded as matched points.

5.2.2.3 Minutiae Set-Based Matching Problem Solved by GA

Seeking the appropriate affine transformation parameters is important for the alignment between the latent minutiae set and the compared print minutiae set. Such suitable parameters ensure the largest overlap of global topology between the two point sets. Considering that the affine transformation parameters need to be elaborately tuned, the parameter optimization technique is necessary. In contrast to the deterministic optimization algorithm, like the greedy searching method proposed in [249, 250], the evolutionary optimization approaches have the advantage of computational efficiency and the capability to effectively avoid the local optimum. As a consequence, GA, one of the typical evolutionary optimization approaches, is applied to solve the minutiae set-based matching problem.

To be specific, GA begins with the randomly initialized chromosomes which represent the solution of the problem. In subsequent iteration, the updated chromosomes are obtained by using various genetic operators. According to the fitness function, the previous chromosomes are substituted by the updated ones when the updated ones are judged to be fitter individuals than the previous ones. With the continuous iteration, the chromosomes are motivated to evolve to the fittest individuals until the termination of the algorithm. Because GA is not easy to trap into the local optimal and its searching manner is potentially parallel, it has been broadly utilized to solve the point set matching problem [265–267].

In order to use GA, the affine transformation parameters, namely rotation angle θ , scaling factor s , and translation offsets t_x and t_y are coded as chromosome. The chromosome vector is denoted as follows

$$\mathbf{Chrom} = (\theta, s, t_x, t_y)^T \quad (5.8)$$

where each parameter value is a random real number and is restricted in an appropriate range. Compared with binary coding, real number-based coding has the following advantages: effectively avert the Hamming cliff and avoid the decimal digits assignment.

Each encoded chromosome corresponds to one parameter set of affine transformation to align the point set. During the iterations of GA, the chromosomes are constrained in the same range as they are randomly initialized.

To motivate the evolution of GA, the fitness of each chromosome needs to be evaluated. In order to define the fitness function, coordinate distance e_d and point orientation difference e_o need to be calculated in advance

$$e_d = \sqrt{[\mathbf{T}(x_{c_i}) - x_{l_j}]^2 + [\mathbf{T}(y_{c_i}) - y_{l_j}]^2} \quad (5.9)$$

$$e_o = |\mathbf{T}(o_{c_i}) - o_{l_j}| \quad (5.10)$$

where $\mathbf{T}(o_{c_i})$ is point c_i 's orientation after rotation caused by transformation \mathbf{T} . Given the resultant e_d and e_o , the fitness function is defined as follows

$$num^t = \begin{cases} num^{t-1} + 1, & \text{if } e_d \leq \delta_d, e_o \leq \delta_o, pt_{c_i} = pt_{l_j} \\ num^{t-1}, & \text{otherwise} \end{cases} \quad (5.11)$$

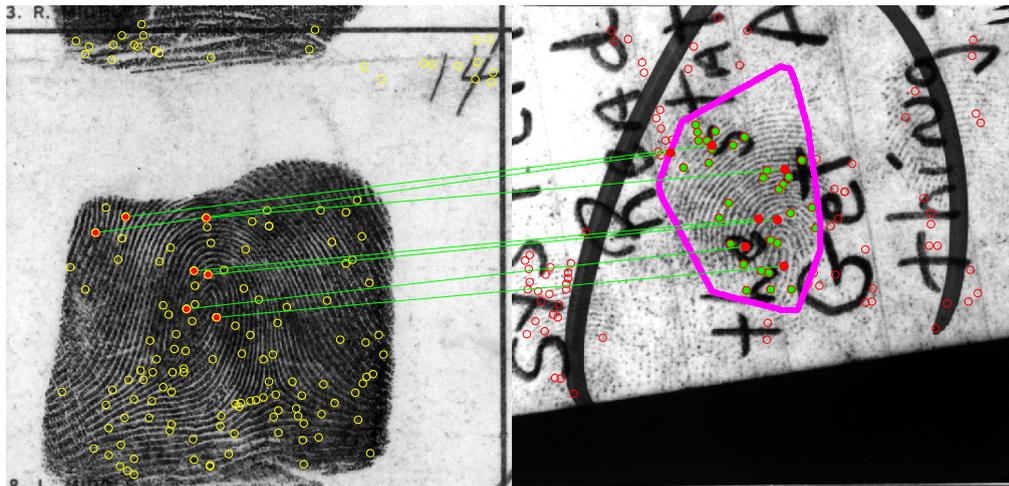
where num^t and num^{t-1} are the numbers of matched point pairs in the t^{th} and $(t-1)^{th}$ iteration, respectively. Therein, the number of paired points is directly assigned as the fitness function value. δ_d and δ_o are the tolerances for e_d and e_o , respectively. Equation 5.11 indicates that for every iteration all the matched point pairs need to be found and counted when the following three criteria are simultaneously satisfied: (i) $e_d \leq \delta_d$; (ii) $e_o \leq \delta_o$; and (iii) $pt_{c_i} = pt_{l_j}$. Therefore, a chromosome producing a larger num (higher fitness) is considered to be superior to the other chromosomes with a smaller num (lower fitness). As such, the fittest chromosomes can be determined according to the fitness values of all chromosomes. In addition, Equations 5.9, 5.10 and 5.11 not only consider the global topology of the point set but also involve the point property such as point orientation and type during the iterative evolution of GA.

To boost the evolution of GA, the GA operators such as selection, crossover and mutation are also important. To be more specific, the selection operator is used to select the chromosomes with higher fitness, which preserve and inherit the information of these

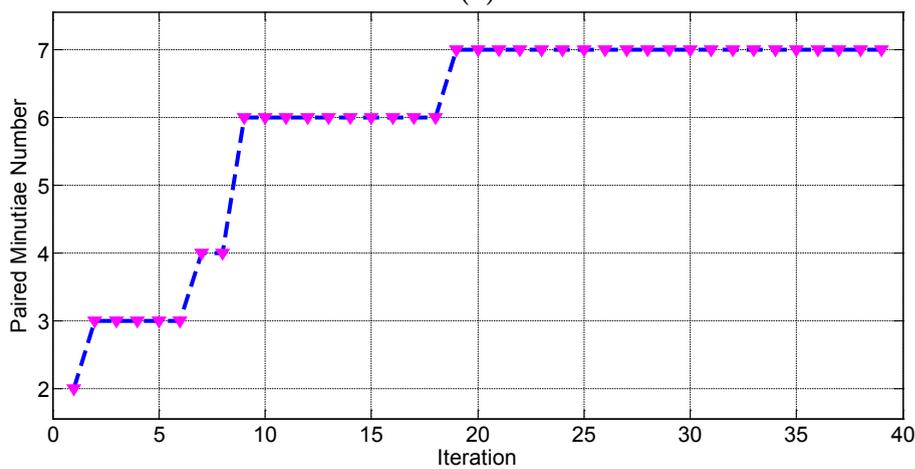
fitter chromosomes into the next generation. Therein, the widely used Roulette Wheel selection scheme is applied and the selection probability of each chromosome is proportional to its own fitness value. Besides, the creation of new offspring to enhance the diversity of the chromosomes is necessary because the inheritance of fitter chromosomes in iterative evolution alone is not sufficient. Therefore, the crossover operator produces new chromosomes through combining partial segments of two parent chromosomes. In this approach, the multi-point crossover operator is adopted. Moreover, in order to further enhance chromosome diversity, the uniform mutation operator is performed to randomly shift the value of the chromosome vector with a small probability.

In this study, the GA-based minutiae set matching algorithm is summarized as follows:

- *Step 1:* Extract the minutiae set \mathbf{C} from the currently compared print and the minutiae set \mathbf{L} from the ROI in the query latent print (shown in Figure 5.9a), then set the population size of chromosomes S_{chrom} , crossover probability p_c , mutation probability p_m , the maximal iterations g_{max} , point coordinate distance tolerance δ_d , point orientation difference tolerance δ_o and the value range for the chromosome vector;
- *Step 2:* Randomly generate chromosomes within the value range as the initial generation;
- *Step 3:* Compute fitness values for all chromosomes in the current generation and then select the fitter ones by selection operator;
- *Step 4:* Apply the crossover operator and mutation operator to the fitter chromosomes, then create the new chromosomes for the next generation;
- *Step 5:* Check whether the maximal iterations are reached, or check whether the highest fitness values are maintained during several iterations. If not reached or maintained, go back to *Step 3*. Otherwise, terminate GA and output the fittest chromosome as the estimated optimal parameter for affine transformation;
- *Step 6:* Align minutiae set \mathbf{C} versus minutiae set \mathbf{L} , based on the obtained transformation parameter, then seek the corresponding minutiae point pairs, according to the criteria in Equation 5.11 (shown in Figure 5.9a);



(a)



(b)

Figure 5.9: The example of successful matching of a genuine “rolled-latent” pair is by the proposed GA-based matching unit. In this case, seven minutiae pairs are found. The paired minutiae between the rolled and latent print are connected by green lines and marked by red solid points, respectively. (a) the rolled print image (left) corresponds to G009 in NIST SD27 (right); the automatically extracted minutiae in the rolled print image are marked by yellow hollow points and the ones in the latent print image are marked by red hollow points; the ROI is highlighted by a pink convex polygon and the automatically extracted minutiae within the ROI are marked by green solid points; and (b) the augment of the paired minutiae number with the iteration of GA

- *Step 7:* Output the number of paired minutiae between \mathbf{C} and \mathbf{L} as the matching score when GA has already converged (shown in Figure 5.9b).

5.3 Experimental Results

In this section, the introduced latent fingerprint matcher has been evaluated by the following two experiments: (i) the minutiae extraction based on the ROI; and (ii) the

latent fingerprint matching.

5.3.1 Data Preparation

All the experiments are conducted on the latent fingerprint database NIST SD27, which is available in the public domain. Such a database includes the 258 latent print images and their corresponding rolled print images, where these 258 latent images are grouped by the latent examiners into the following three categories: “Good”, “Bad” and “Ugly”. The number of latent images involved in the “Good”, “Bad” and “Ugly” categories are 88, 85 and 85, respectively.

5.3.2 Experiment 1: ROI-Based Minutiae Extraction

In this experiment, the reliability of automated minutiae extraction based on the ROI in the latent print is evaluated. All the minutiae within the ROI are automatically extracted by VeriFinger SDK. Given the query latent print image, the following three scenarios are compared.

- **Minutiae Extraction Scenario 1 - Whole Image-Based Automated Minutiae Extraction:** instead of the segmentation of the foreground, the whole latent image is directly used for the automated minutiae extraction.
- **Minutiae Extraction Scenario 2 - Proposed ROI Segmentation Module-Based Automated Minutiae Extraction:** the ROI is obtained by the proposed dictionary learning-based ROI segmentation module and then used for the automated minutiae extraction. The parameters used in the proposed ROI segmentation module are empirically tuned as follows: the patch size for training sample selection $w = 32$, the atom number $N_a = 100$, the sparsity parameter $K = 2$, and the cross-correlation threshold to judge the pattern similarity $Th_{xcorr} = 0.6$.
- **Minutiae Extraction Scenario 3 - Cao’s ROI Segmentation Approach-Based Automated Minutiae Extraction:** the ROI is obtained by [253] and then used for the automated minutiae extraction. The parameters used in this method are tuned as the suggested ones.

In order to assess the performance of the ROI-based automated minutiae extraction, the manually marked minutiae provided by NIST SD27 for the query latent fingerprint are used as the ground truth to label the genuine minutiae. To be more specific, \mathbf{MS}_1 is the ground-truth minutiae set (provided by NIST SD27 and marked by the FBI latent fingerprint examiners); \mathbf{MS}_2 is the minutiae set automatically extracted from the whole latent image by VeriFinger SDK; and \mathbf{MS}_3 is the minutiae set automatically extracted from the ROI by VeriFinger SDK. Accordingly, two metrics such as genuine minutiae preservation rate (GMPR) and false minutiae acceptance rate (FMAR) are defined as follows:

$$GMPR = \frac{\#\{\mathbf{MS}_1 \cap \mathbf{MS}_3\}}{\#\{\mathbf{MS}_1 \cap \mathbf{MS}_2\}} \quad (5.12)$$

$$FMAR = \frac{\#\{\mathbf{MS}_3 - \mathbf{MS}_1 \cap \mathbf{MS}_3\}}{\#\{\mathbf{MS}_2 - \mathbf{MS}_1 \cap \mathbf{MS}_2\}} \quad (5.13)$$

where GMPR is defined as the percentage of the minutiae belonging to the genuine minutiae set which are also correctly extracted by computer program. FMAR is defined as the percentage of the minutiae belonging to the structured noise in the latent image which are wrongly detected as the genuine ones by computer program. $\#\{*\}$ denotes the number of minutiae included in set $\{*\}$. Under the condition of fully automated minutiae extraction, the effect of ROI is evaluated depending on the benchmark where no ROI mask is adopted but the whole query latent image is used for automated minutiae extraction. Therefore, in Equations 5.12 and 5.13, not the manually marked genuine minutiae $\#\{\mathbf{MS}_1\}$ but the automatically extracted genuine minutiae $\#\{\mathbf{MS}_1 \cap \mathbf{MS}_2\}$ and $\#\{\mathbf{MS}_2 - \mathbf{MS}_1 \cap \mathbf{MS}_2\}$ are used as the baseline. For **Minutiae Extraction Scenario 1**, $\mathbf{MS}_1 - \mathbf{MS}_1 \cap \mathbf{MS}_2$ stands for the missing genuine minutiae, even if the entire latent image is imported into the automated minutiae extractor; $\mathbf{MS}_1 \cap \mathbf{MS}_2$ represents the genuine minutiae which are correctly detected by the computer program; and $\mathbf{MS}_2 - \mathbf{MS}_1 \cap \mathbf{MS}_2$ are the spurious minutiae falsely extracted by the computer program. For **Minutiae Extraction Scenario 2 and 3**, $\mathbf{MS}_1 - \mathbf{MS}_1 \cap \mathbf{MS}_3$ denotes the missing genuine minutiae after the adoption of ROI; $\mathbf{MS}_1 \cap \mathbf{MS}_3$ stands for the genuine minutiae correctly detected by the computer program inside the ROI domain; and

Table 5.1: The comparison of mean GMPR (\overline{GMPR}), mean FMAR (\overline{FMAR}) and mean AUC (\overline{AUC}) among three scenarios

	\overline{GMPR}	\overline{FMAR}	\overline{AUC}
Minutiae Extraction Scenario 1	0.9689	1	0.4845
Minutiae Extraction Scenario 2	0.7772	0.2711	0.7531
Minutiae Extraction Scenario 3	0.6854	0.2642	0.7106

$MS_3 - MS_1 \cap MS_3$ represents the false minutiae wrongly extracted by the computer program within the ROI.

The minutiae extracted by the automated computer program could not be exactly the genuine ones and unavoidably contain fakes; that is, image-based minutiae detection and extraction is primarily dependent on the locally salient image structures such as ridge ending or ridge bifurcation in the image domain. For the latent image, such typical structures are not distinct or are even lost due to the low clarity of the “ridge-valley” pattern in the fingerprint region, and consequently lead to the loss of genuine minutiae; some other image components in the background also have similar ridge ending or bifurcation structures, and accordingly yield spurious minutiae. Therefore, the missing detection of the genuine minutiae and the false detection of the imposter ones by the automated computer program (e.g. VeriFinger SDK) are usually unavoidable. As a consequence, for **Minutiae Extraction Scenario 1**, the sets $MS_1 - MS_1 \cap MS_2$ and $MS_2 - MS_1 \cap MS_2$ are not empty. Similarly, for **Minutiae Extraction Scenario 2 and 3**, $MS_1 - MS_1 \cap MS_3$ and $MS_3 - MS_1 \cap MS_3$ are not empty either.

Based on the three scenarios, their corresponding GMPRs and FMARs on the 258 latent images in NIST SD27 are calculated according to the Equations 5.12 and 5.13, respectively. Taking Figure 5.10 as an example, the cyan squared points in Figure 5.10a stand for the genuine minutiae labeled by ground-truth. The red and yellow hollow points in Figures 5.10b, 5.10c and 5.10d stand for the automatically extracted minutiae. The green solid points in Figures 5.10b, 5.10c and 5.10d stand for the genuine minutiae which are correctly detected via automated program. The ROIs (highlighted by pink convex polygons) in Figure 5.10c and Figure 5.10d are obtained by proposed segmentation module and method in [253] respectively. Furthermore, the mean GMPR (\overline{GMPR}) and the mean FMAR (\overline{FMAR}) are summarized in Table 5.1. For the performance comparison among these three scenarios, the area under the curve (AUC) in

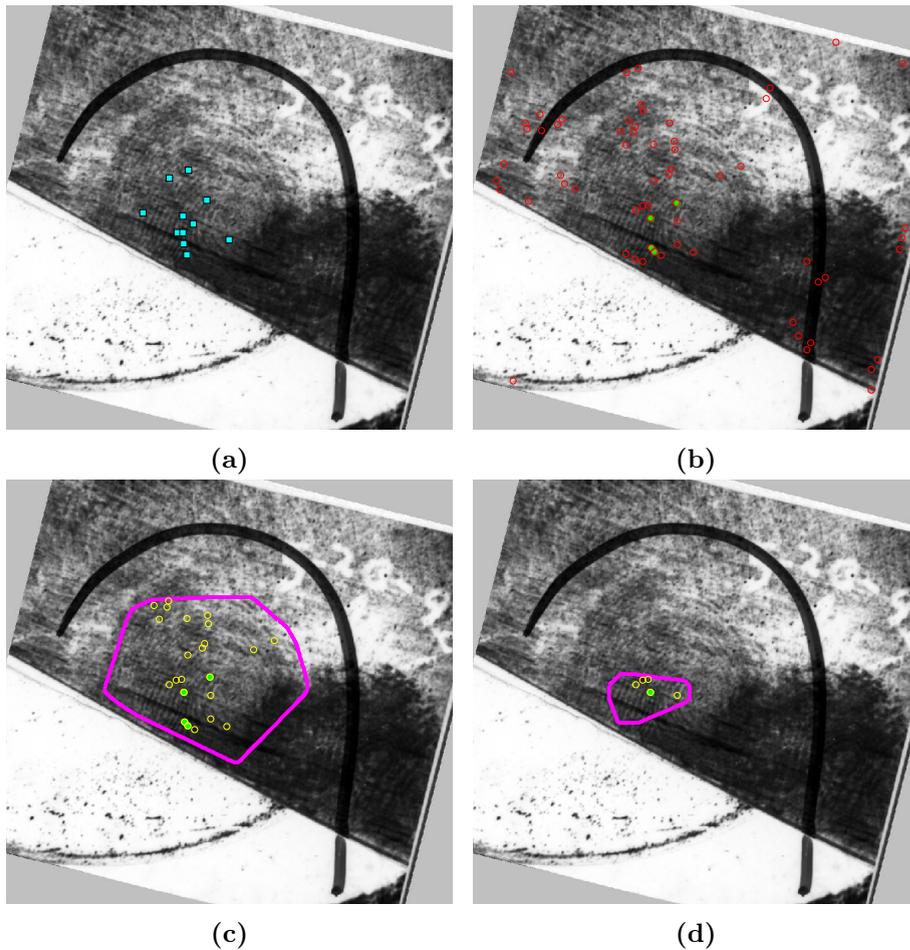


Figure 5.10: The example of three scenarios on U267 in NIST SD27: (a) the ground-truth minutiae set provided by NIST SD27; (b) Minutiae Extraction Scenario 1 - four genuine minutiae can be correctly extracted, $GMPR = 1$, $FMAR = 1$, and $AUC = 0.5$; (c) Minutiae Extraction Scenario 2 - four genuine minutiae can be correctly extracted, $GMPR = 1$, $FMAR = 0.3167$, and $AUC = 0.8416$; and (d) Minutiae Extraction Scenario 3 - only one genuine minutia can be correctly detected, $GMPR = 0.25$, $FMAR = 0.0667$, and $AUC = 0.5916$

the receiver operating characteristics analysis (ROC analysis) is adopted. Because the GMPR and the FMAR have the equivalent principle as the true positive rate (TPR: sensitivity) and the false positive rate (FPR: specificity) in ROC analysis, the AUC is used to evaluate the impact of ROI-based automated minutiae extraction; that is, the higher the AUC, the more reliable the ROI-based automated minutiae extraction is. For the statistics of the performance regarding the three scenarios, their corresponding mean AUCs (\overline{AUC}) on the overall 258 latent images are calculated. As shown in Table 5.1, the automated minutiae extraction, based on the ROI obtained by the proposed segmentation module, achieves the best performance, where the genuine minutiae are effectively preserved while the imposter ones are significantly reduced. In addition, the

automated minutiae extractor cannot detect even one genuine minutia for some latent images, although the images in such failed cases are completely imported into the automated minutiae extractor. Accordingly, such failed cases where $\mathbf{MS}_1 \cap \mathbf{MS}_2 = \emptyset$ have been evidenced by $\overline{GM\overline{PR}} = 0.9689 < 1$ in Table 5.1.

5.3.3 Experiment 2: Latent Fingerprint Matching

In this session, the matching experiment is implemented to verify whether the proposed dictionary learning-based ROI segmentation module and the GA-based matching unit are able to boost the matching performance. As the ultimate goal of the fully automated latent fingerprint matching, matching accuracy plays a vital role in demonstrating the improvement caused by the proposed ROI segmentation module and the GA-based matching unit.

In this experiment, the 258 latent images provided by NIST SD27 are used to form a background database. The size of the background database is 258. Each query latent print is matched against such a background database. The minutiae for the rolled prints and the ROI in the latent prints are automatically extracted by VeriFinger SDK 4.0. The parameters used in the proposed GA-based matching platform are empirically tuned as follows: the size of the chromosome population $S = 400$, crossover probability $p_c = 0.2$, mutation probability $p_m = 0.05$, minutiae coordinate distance tolerance $\delta_d = 15$, and minutiae orientation difference tolerance $\delta_o = 20^\circ$. Also, the value ranges for the rigid transformation parameters are set as follows: the orientation range $0 \leq \theta \leq 359^\circ$, the scaling range $0.8 \leq s \leq 1.2$, the horizontal shift range $-400 \leq t_x \leq 400$, and the vertical shift range $-400 \leq t_y \leq 400$. Due to the intensive computation caused by GA's evolution, this matching experiment is implemented on a high performance computing facility. Also, a parallel strategy to simultaneously launch the different latent matchings is adopted. For one CPU on such large-scale computing resource, an one-to-one matching (one latent print against one template) averagely needs 30 seconds.

In order to evaluate and compare the matching performance, the cumulative match characteristic curve (CMC) is utilized. Furthermore, for the statistics of matching performance, the mean CMC (\overline{CMC}) and mean penetration rate (\overline{pr}) on overall 258 latent

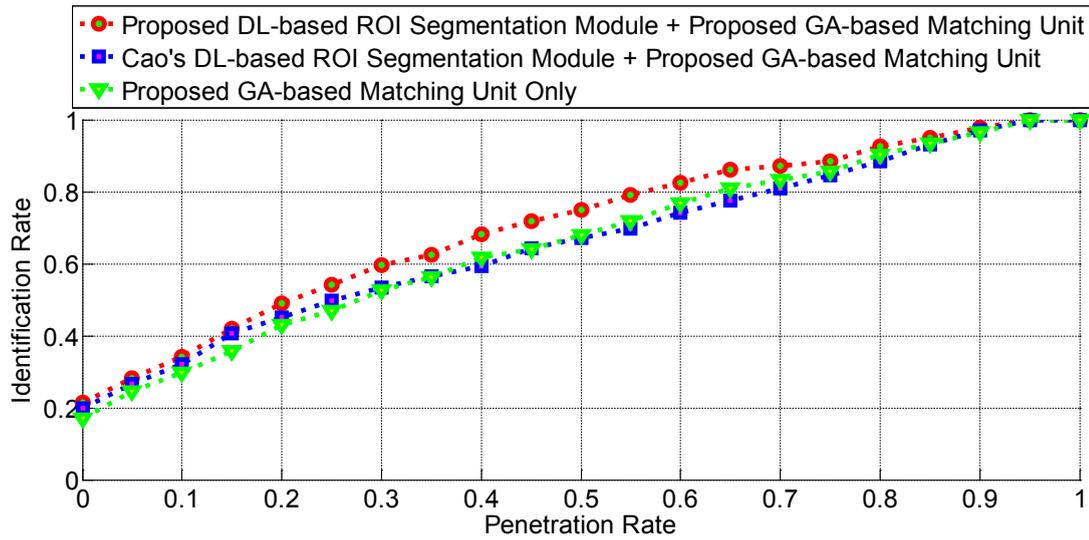


Figure 5.11: Different \overline{CMC} obtained by the three scenarios.

prints during several independent trials are obtained. For the following three scenarios, their corresponding \overline{CMC} and \overline{pr} are demonstrated in Figure 5.11 and Table 5.2, respectively.

- **Matching Scenario 1 - Proposed GA-Based Matching Unit Only:** without the segmentation module, the minutiae automatically extracted from the whole query latent image are directly imported into the proposed GA-based matching unit.
- **Matching Scenario 2 - Proposed ROI Segmentation Module + Proposed GA-Based Matching Unit:** with the preprocessing procedure performed by the proposed ROI segmentation module, the ROI is obtained; then the minutiae automatically extracted from the obtained ROI are imported into the proposed GA-based matching unit.
- **Matching Scenario 3 - Cao's ROI Segmentation Method + Proposed GA-Based Matching Unit:** by adopting the approach introduced in [253], the ROI is obtained; then the minutiae automatically extracted from the obtained ROI are imported into the proposed GA-based matching unit.

As illustrated in Figure 5.11, the proposed GA-based matching unit is performed as the baseline latent matcher in **Matching Scenario 1**. Based on such a baseline matcher, further improvement is expected to be achieved via the ROI segmentation method.

Table 5.2: Different \overline{pr} obtained by the three scenarios

	Matching Scenario 1	Matching Scenario 2	Matching Scenario 3
\overline{pr}	38.728%	33.692%	38.321%

Matching Scenario 2 and 3 provide the two different segmentation modules for the ROI identification in the latent images, respectively. As demonstrated in the preceding experiment, the reliability of automated minutiae extraction based on the ROI segmentation module in **Matching Scenario 2** is better than that of **Matching Scenario 3**. As a result, the matching performance achieved by **Matching Scenario 2** is consequently better than that of **Matching Scenario 3**. The significant enhancement of the mean penetration rate \overline{pr} in Table 5.2 also demonstrates the effectiveness of the proposed multi-module latent matcher in **Matching Scenario 2**.

The matching performances in the three categories of latent images are further evaluated by **Matching Scenario 1**, **Matching Scenario 2** and **Matching Scenario 3**, respectively. The results in Figures 5.12, 5.13 and 5.14 and Table 5.3 demonstrate that the \overline{CMC} and the \overline{pr} achieved on the basis of “Good” latent images are more satisfying than the “Bad” and “Ugly” versions. Also, the matching results indicate that the “Bad” cases are more challenging than the “Ugly” ones, even though the image quality of the “Bad” type is supposed to be better than that of the “Ugly” type, in terms of visual comparison. To be specific, the “Good” penetration rates are further decreased from 30.065% (in **Matching Scenario 1**) to 27.493% (in **Matching Scenario 2**) and 26.498% (in **Matching Scenario 3**), respectively, due to the utilization of ROI masks yielded by the segmentation modules. Particularly, the “Ugly” penetration rate of GA-based benchmark 42.339% is significantly reduced to 29.717% via the ROI segmentation scheme in **Matching Scenario 2**, while the reduced “Ugly” penetration rate regarding **Matching Scenario 3** reaches 33.767% and it is still higher than that of **Matching Scenario 2**. The “Bad” penetration rate is difficult to reduce even by exploiting the ROI segmentation technique. On the contrary, the adoption of the ROI mask obtained by **Matching Scenario 3** degrades the matching performance of the “Bad” cases (lifting from 44.086% to 55.116%). In contrast to **Matching Scenario 3**, **Matching Scenario 2** does not raise the “Bad” penetration rate but maintains it at the same level, 44.086%, as that in **Matching Scenario 1**.

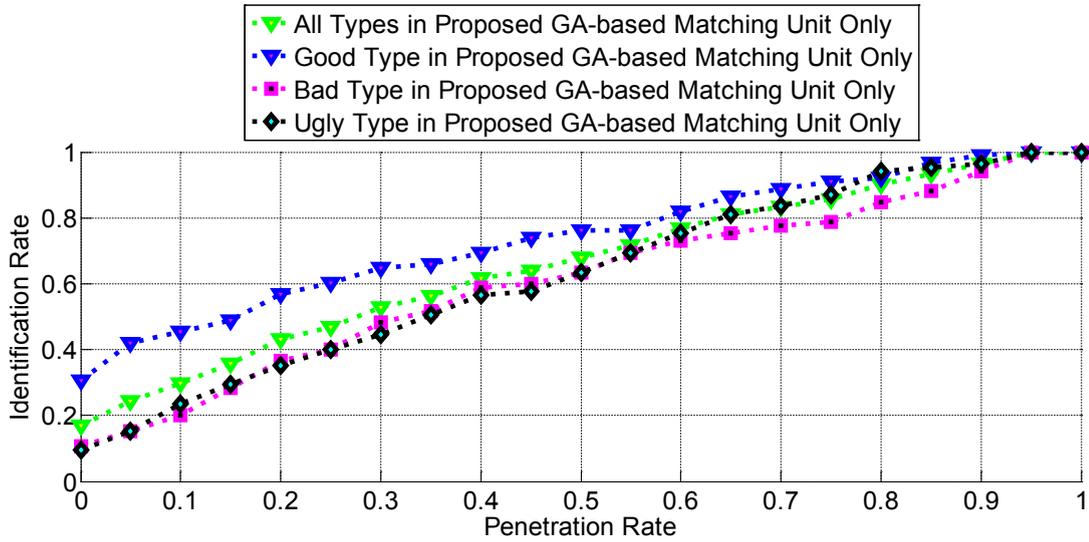


Figure 5.12: \overline{CMC} obtained in Matching Scenario 1 in regard to the three types of latent image: Good, Bad and Ugly.

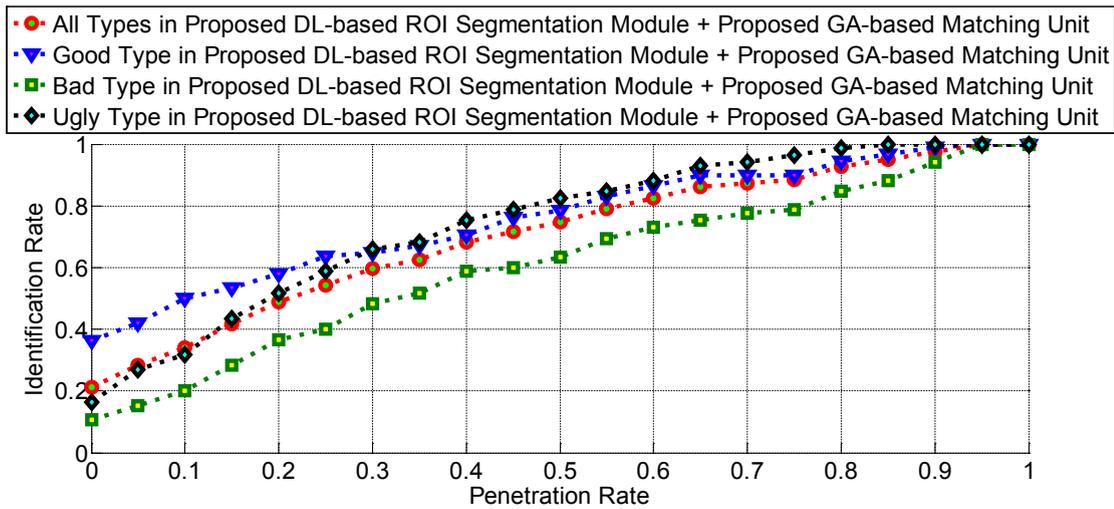


Figure 5.13: \overline{CMC} obtained in Matching Scenario 2 in regard to the three types of latent image: Good, Bad and Ugly.

Table 5.3: \overline{pr} obtained in three scenarios according to the three categories of latent image: Good, Bad and Ugly

	Matching Scenario 1	Matching Scenario 2	Matching Scenario 3
\overline{pr}_G	30.065%	27.493%	26.498%
\overline{pr}_B	44.086%	44.086%	55.116%
\overline{pr}_U	42.339%	29.717%	33.767%

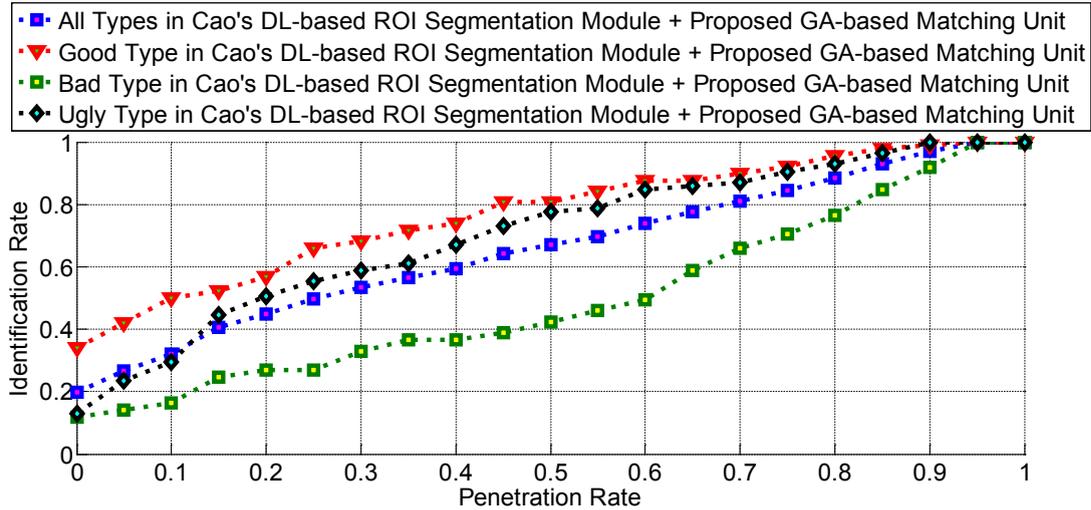


Figure 5.14: \overline{CMC} obtained in Matching Scenario 3 in regard to the three types of latent image: Good, Bad and Ugly.

5.4 Conclusion

Although the impressive matching performance has been achieved in the rolled / plain print identification tasks, the matching of the latent fingerprint is still a challenging problem because of the following issues: (i) the poor-quality image and the low-clarity “ridge-valley” pattern in the ROI; and (ii) the corruption with the extra structured image components. Such issues impose adverse effects on the fully automated ROI segmentation, minutiae extraction, and minutiae-based matching. In order to deal with the challenges in the latent fingerprint matching task, a multi-module latent matching system is introduced in this chapter. The proposed latent matcher consists of the following two modules: (i) the dictionary learning-based ROI segmentation scheme; and (ii) the genetic algorithm-based minutiae set matching unit. Experimental results on the NIST SD27 latent image database demonstrate the effectiveness of the proposed multi-module latent matching system. In consideration of its practical value, the proposed matching system can be available in the public domain for latent print identification, and can also be smoothly extended and incorporated with other preprocessing or post-processing modules.

Chapter 6

Conclusions and Future Studies

6.1 Conclusions

Advancement in image denoising technology enables massive biomedical and biometric images to be processed in a comprehensible form for human beings. Not only random noise, but also irrelevant components are regarded as targets and included in the denoising framework. Consequently, the research topic of image denoising is not only restricted to random noise cancellation, but also is extended to irrelevant image content removal. Due to extensive application demands, an increasing number of challenges have been of concern. In order to solve technical difficulties and satisfy tangible demand in biomedical and biometric image processing, three specific problems have been addressed in this thesis.

Firstly, we propose a multi-staged method based on a novel integration of trend surface analysis, quantile-quantile plot, bootstrapping, and Gaussian spatial kernel for restoration of noisy microscopy cell images. We show this multi-staged approach achieves higher performance compared with other state-of-the-art restoration techniques in terms of peak signal-to-noise ratio (PSNR) and structure similarity (SSIM) in synthetic noise experiments. We also report an experiment on real noisy microscopy data which demonstrated the advantages of the proposed restoration method for improving cell segmentation performance.

Secondly, particle potential motion entropy, a new uncertainty measure is developed to characterize pixel-wise particle potential motion. Self-information, a discriminative feature, is introduced to detect image local structures. Based on obtained particle potential motion entropy and identified local structure, we create a histogram-based descriptor named particle potential motion entropy histogram (PPMEH) to represent the whole image. The proposed PPMEH has the following two advantages: (i) it synthesizes the adaptively estimated local structure information into a global descriptor to depict the whole image; and (ii) it is tolerant to several image distortion and degradation, such as large-range image rotation (from -180° to 180°), medium-scale image scaling (from 0.8 to 1.2) and medium-level random noise corruption ($\sigma_n = 30$). The experiment on synthetic structure image dataset demonstrates the feasibility and effectiveness of the proposed scheme when detecting image local structures. Also, the proposed PPMEH is tested under diverse simulated scenarios involving image rotation, scaling and noise based on colon computed tomography (CT) and head magnetic resonance imaging (MRI) datasets. The comparison with state-of-the-art image descriptors such as the local binary pattern (LBP), the Gabor filter feature (GF), the grey level co-occurrence matrix feature (GLCM), the semi-variogram feature (SV) and the histogram of gradient (HOG) is also performed. The experimental results demonstrate that the proposed PPMEH outperforms other state-of-the-art image descriptors.

Thirdly, we develop a latent fingerprint matcher to remove irrelevant structural patterns over the partial and poor-quality latent fingerprint image, and to find a match against a background database. Currently, there is no fully automated latent fingerprint matcher available to the public and most literature reports have utilized a specialized latent fingerprint matcher, COTS3, which is not accessible to the public. In this thesis, we target the development of a fully automated latent matcher for adaptive detection of the region of interest and robust matching of latent prints. Unlike the manually conducted matching procedure, the proposed latent matcher can run like a sealed black box, without any manual intervention. This matcher consists of the following two modules: (i) a dictionary learning-based region of interest (ROI) segmentation scheme; and (ii) a genetic algorithm-based minutiae set matching unit. Experimental results on the NIST SD27 latent fingerprint database demonstrate the superiority of the proposed fingerprint segmentation scheme and matching unit.

6.2 Future Studies

In this thesis, several random noise cancelation and irrelevant component removal algorithms have been developed to address the challenges in biomedical and biometric image processing and interpretation. However, there are more technical issues to investigate further in future studies. Recommendations for on-going research are briefly stated as follows

- **Incorporating random noise type identification module in the proposed microscopic denoising algorithm:** the Q-Q plot plays an important role in the proposed microscopic denoising algorithm, where it is able to identify unknown distribution by comparing with given reference distribution. In future study, a random noise type identification module based on the Q-Q plot could be further developed. To be explicit, various random noise distributions with fixed parameters can be established and retained as reference in advance. Given a noisy image carrying certain random noise, the pixel intensity samples are collected from such an image for comparison with reference distribution via the Q-Q plot. According to linear or non-linear patterns presented in the Q-Q plot, the unknown random noise could be categorized or excluded.
- **Extending 2D PPMEH to 3D PPMEH:** the currently proposed PPMEH and PPMEH-FT features are unable to characterize 3D volumes from stereo imaging systems, but only represent 2D slices generated by tomographic imaging modalities (e.g. CT, MRI and X-ray). In order to effectively represent stereo volumes, a 3D PPMEH based on its 2D version could be further developed to characterize 3D objects. Accordingly, the application of PPMEH will not be only restricted to 2D scans retrieval but also is extended to 3D targets indexing.
- **Integrating the “ridge-valley” structure enhancement module in the proposed latent fingerprint matching system:** The currently proposed latent fingerprint matcher is still not able to cope with very low-quality latent images. Since the clarity of the “ridge-valley” pattern in the foreground is poor, the minutiae within such a region may not be extracted in a fully automatic mode and consequently, searching against the background database would not be successful. Therefore, the development and integration of an enhancement module for

the low-clarity “ridge-valley” structure in the foreground is essential in ongoing research.

- **Reducing computation load caused by iterative GA:** The intensive computation caused by the GA iterative procedure needs to be reduced. To be specific, defining a simpler fitness function, or exploring the utilization of a large-scale parallel / distributed computing system can help increase computational efficiency.

Bibliography

- [1] M.C. Motwani, M.C. Gadiya, R.C. Motwani and et al., Survey of image denoising techniques, *Proceeding of Global Signal Processing Expo and Conference*, 2004.
- [2] P. Milanfar, A tour of modern image filtering: new insights and methods, both practical and theoretical, *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 106-128, 2013.
- [3] L. Shao, R. Yan, X. Li and et al., From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms, *IEEE Transactions on Cybernetics*, vol. 44, no. 7, pp. 1001-1013, 2014.
- [4] J. Mohan, V. Krishnaveni, Y. Guo, A survey on the magnetic resonance image denoising methods, *Biomedical Signal Processing and Control*, vol. 9, no. 1, pp. 56-69, 2014.
- [5] G. Chen, R. Chen, Z. Huang and et al., Iterative regularization denoising method based on OSV model for biomedical image denoising, *Journal of Physics: Conference Series*, vol. 277, pp. 012005, 2011.
- [6] D. Mary Sugantharathnam, D. Manimegalai, B. Ganapathy Ram, Wavelets and ridgelets for biomedical image denoising, *WSEAS Transactions on Systems*, vol. 12, no. 3, pp. 165-178, 2013.
- [7] P. Qiu, P.S. Mukherjee, Edge structure preserving 3D image denoising by local surface approximation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1457-1468, 2012.
- [8] B. Kumar, M. Savvides, K. Venkataramani and et al., Spatial frequency domain image processing for biometric recognition, *IEEE International Conference on Image Processing*, vol. 1, pp. 53-56, 2002.

-
- [9] T.L. Huang, T.L. Hwang, C.W. Liu and et al., A modern algorithm for fingerprint image enhancement, *WSEAS Transactions on Information Science and Applications*, vol. 4, no. 6, pp. 1195-1199, 2007.
- [10] B. Tan, S. Schuckers, New approach for liveness detection in fingerprint scanners based on valley noise analysis, *Journal of Electronic Imaging*, vol. 17, no. 1, pp. 011009:1-9, 2008.
- [11] B. Peng, L. Zhang, D. Zhang, A survey of graph theoretical approaches to image segmentation, *Pattern Recognition*, vol. 46, no. 3, pp. 1020-1038, 2013.
- [12] S.R. Vantaram, E. Saber, Survey of contemporary trends in color image segmentation, *Journal of Electronic Imaging*, vol. 21, no. 4, pp. 040901:1-28, 2012.
- [13] N. Gordillo, E. Montseny, P. Sobrevilla, State of the art survey on MRI brain tumor segmentation, *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426-1438, 2013.
- [14] S. Bauer, R. Wiest, L.P. Nolte and et al., A survey of MRI-based medical image analysis for brain tumor studies, *Physics in Medicine and Biology*, vol. 58, no. 13, pp. R97-R129, 2013.
- [15] S. Wang, R.M. Summers, Machine learning and radiology, *Medical Image Analysis*, vol. 16, no. 5, pp. 933-951, 2012.
- [16] A.M. Mharib, A.R. Ramli, S. Mashohor and et al., Survey on liver CT image segmentation methods, *Artificial Intelligence Review*, vol. 37, no. 2, pp. 83-95, 2012.
- [17] K. Suzuki, Machine learning in computer-aided diagnosis of the thorax and colon in CT: a survey, *IEICE Transactions on Information and Systems*, vol. E96-D, no. 4, pp. 772-783, 2013.
- [18] S. Ghose, A. Oliver, R. Marti and et al., A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images, *Computer Methods and Programs in Biomedicine*, vol. 108, no. 1, pp. 262-287, 2012.

- [19] R.D. Labati, A. Genovese, V. Piuri and et al., Touchless fingerprint biometrics: a survey on 2D and 3D technologies, *Journal of Internet Technology*, vol. 15, no. 3, pp. 325-332, 2014.
- [20] H. Hasan, S. Abdul-Kareem, Fingerprint image enhancement and recognition algorithms: a survey, *Neural Computing and Applications*, vol. 23, no. 6, pp. 1605-1610, 2013.
- [21] N. Wiener, Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications, *Technology Press of the Massachusetts Institute of Technology*, 1949.
- [22] J.S. Lee, Digital image enhancement and noise filtering by use of local statistics, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 165-168, 1980.
- [23] R.C. Gonzalez, R.E. Woods, Digital image processing, *Pearson Education*, 2011.
- [24] M.H. Sedaaghi, Morphological operators, *IEEE Electronics Letters*, vol. 38, no. 22, pp. 1333-1335, 2002.
- [25] Z. Ma, H. Wu, D. Feng, Partition-based vector filtering technique for suppression of noise in digital color images, *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2324-2342, 2006.
- [26] Z. Ma, H. Wu, D. Feng, Fuzzy vector partition filtering technique for color image restoration, *Computer Vision and Image Understanding*, vol. 107, no. 1-2, pp. 26-37, 2007.
- [27] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, *Proceedings of 1998 6th International Conference on Computer Vision*, pp. 839-846, 1998.
- [28] A. Buades, B. Coll and J.M. Morel, A non-local algorithm for image denoising, *Proceeding of 2005 IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 60-65, 2005.
- [29] A. Buades, B. Coll and J.M. Morel, Image denoising by non-local averaging, *Proceeding of 2005 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 25-28, 2005.

- [30] C. Kervrann, J. Boulanger, Optimal spatial adaptation for patch-based image denoising, *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 2866-2878, 2006.
- [31] B. Wohlberg, Inpainting by joint optimization of linear combinations of exemplars, *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 75-78, 2011.
- [32] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, 1990.
- [33] J. Weickert, Anisotropic diffusion in image processing, *Stuttgart: Teubner Verlag*, 1998.
- [34] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, vol. 60, pp. 259-268, 1992.
- [35] T. Chan, S. Osher, J. Shen, The digital TV filter and nonlinear denoising, *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 231-241, 2001.
- [36] S. Mallat, A wavelet tour of signal processing, *Academic*, 2008.
- [37] E. Pennec, S. Mallat, Sparse geometric image representations with bandelets, *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 423-438, 2005.
- [38] D. Donoho, Wedgelets: nearly-minimax estimation of edges, *Annals of Statistics*, vol. 27, no. 3, pp. 859-897, 1999.
- [39] E. Candes, D.L. Donoho, Recovering edges in ill-posed inverse problems: optimality of curvelet frames, *Annals of Statistics*, vol. 30, no. 3, pp. 784-842, 2002.
- [40] E. Candes, D.L. Donoho, New tight frames of curvelets and the problem of approximating piecewise c_2 image with piecewise c_2 edges, *Communications on Pure and Applied Mathematics*, vol. 57, pp. 219-266, 2003.
- [41] S. Mallat, E.L. Pennec, Bandelet image approximation and compression, *Multi-scale Modeling and Simulation: A SIAM Interdisciplinary Journal*, vol. 4, no. 3, pp. 992-1039, 2005.
- [42] M. Do, M. Vetterli, Frame pyramids, *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2329-2342, 2003.

- [43] S. Mallat, G. Peyre, Orthogonal bandlet bases for geometric images approximation, *Communications on Pure and Applied Mathematics*, vol. 61, no. 9, pp. 1173-1212, 2008.
- [44] W.T. Freeman, E.H. Adelson, The design and the use of steerable filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891-906, 1991.
- [45] J. Portilla, V. Strela, M.J. Wainwright and et al., Image denoising using scale mixtures of Gaussians in the wavelet domain, *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338-1351, 2003.
- [46] K. Dabov, A. Foi, V. Katkovnik and et al., Image denoising by sparse 3D transform-domain collaborative filtering, *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080-2095, 2007.
- [47] L. Zhang, W. Dong, D. Zhang and et al., Two-stage image denoising by principal component analysis with local pixel grouping, *Pattern Recognition*, vol. 43, no. 4, pp. 1531-1549, 2010.
- [48] J. Mairal, F. Bach, J. Ponce and et al., Non-local sparse models for image restoration, *Proceeding of 2009 IEEE International Conference on Computer Vision*, pp. 2272-2279, 2009.
- [49] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736-3745, 2006.
- [50] B.A. Olshausen, D.J. Field, Emergence of simplecell receptive field properties by learning a sparse code for natural images, *Nature*, vol. 381, no. 6583, pp. 607-609, 1996.
- [51] B.A. Olshausen, D.J. Field, Sparse coding with an overcomplete basis set: a strategy employed by v1? *Vision Research*, vol. 37, no. 23, pp. 311-325, 1996.
- [52] K. Engan, S. Aase, J. Hakon-Husoy, Method of optimal directions for frame design, *Proceeding of 1999 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 2443-2446, 1999.

-
- [53] K. Kreutz-Delgado, J. Murray, B. Rao and et al., Dictionary learning algorithms for sparse representation, *Neural Computation*, vol. 15, no. 2, pp. 349-396, 2003.
- [54] M.S. Lewicki, T.J. Sejnowski, Learning overcomplete representations, *Neural Computation*, vol. 12, no. 2, pp. 337-365, 2000.
- [55] L. Lesage, R. Gribonval, F. Bimbot and et al., Learning unions of orthonormal bases with thresholded singular value decomposition, *Proceeding of 2005 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2272-2279, 2005.
- [56] J. Mairal, M. Elad, G. Sapiro, Sparse representation for color image restoration, *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53-69, 2008.
- [57] W. Dong, X. Li, L. Zhang and et al., Sparsity-based image denoising via dictionary learning and structural clustering, *Proceeding of 2011 IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 457-464, 2011.
- [58] K.S. Fu, J.K. Mui, A survey on image segmentation, *Pattern Recognition*, vol. 13, no. 1, pp. 3-16, 1981.
- [59] N.R. Pal, S.K. Pal, A review on image segmentation techniques, *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.
- [60] H.D. Cheng, Color image segmentation: advances and prospects, *Pattern Recognition*, vol. 34, no. 12, pp. 2259-2281, 2001.
- [61] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146-165, 2004.
- [62] F. Kurugollu, B. Sankur, A.E. Harmanci, Color image segmentation using histogram multithresholding and fusion, *Image and Vision Computing*, vol. 19, no. 13, pp. 915-928, 2001.
- [63] H.D. Cheng, X.H. Jiang, J. Wang, Color image segmentation based on homogram thresholding and region merging, *Pattern Recognition*, vol. 35, no. 2, pp. 373-393, 2002.

-
- [64] M.M. Mushrif, A.K. Ray, Color image segmentation: rough-set theoretic approach, *Pattern Recognition Letters*, vol. 29, no. 4, pp. 483-493, 2008.
- [65] M.M. Mushrif, A.K. Ray, A-IFS histon based multithresholding algorithm for color image segmentation, *IEEE Signal Processing Letters*, vol. 16, no. 3, pp. 168-171, 2009.
- [66] J. Chen, Adaptive perceptual color-texture image segmentation, *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1524-1536, 2005.
- [67] M. Mignotte, Segmentation by fusion of histogram-based k-means clusters in different color spaces, *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 780-787, 2008.
- [68] M. Mignotte, A de-texturing and spatially constrained k-means approach for image segmentation, *Pattern Recognition Letters*, vol. 32, no. 2, pp. 359-367, 2011.
- [69] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603-619, 2002.
- [70] Y. Hong, J. Yi, D. Zhao, Improved mean shift segmentation approach for natural images, *Applied Mathematics and Computation*, vol. 185, no. 2, pp. 940-952, 2007.
- [71] M. Ozden, E. Polat, A color image segmentation approach for content-based image retrieval, *Pattern Recognition*, vol. 40, no. 4, pp. 1318-1325, 2007.
- [72] N.C. Yeo, Colour image segmentation using the self-organizing map and adaptive resonance theory, *Image and Vision Computing*, vol. 23, no. 12, pp. 1060-1079, 2005.
- [73] A.R.F. Araujo, D.C. Costa, Local adaptive receptive field selforganizing map for image color segmentation, *Image and Vision Computing*, vol. 27, no. 9, pp. 1229-1239, 2009.
- [74] D.E. Ilea, P.F. Whelan, CTex-an adaptive unsupervised segmentation algorithm based on color-texture coherence, *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1926-1939, 2008.

-
- [75] J.F. Yang, S.S. Hao, P.C. Chung, Color image segmentation using fuzzy C-means and eigenspace projections, *Signal Processing*, vol. 82, no. 3, pp. 461-472, 2002.
- [76] A.W.C. Liew, H. Yan, N.F. Law, Image segmentation based on adaptive cluster prototype estimation, *IEEE Transactions on Fuzzy System*, vol. 13, no. 4, pp. 444-453, 2005.
- [77] W.L. Hung, M.S. Yang, D.H. Chen, Bootstrapping approach to feature-weight selection in fuzzy C-means algorithms with an application in color image segmentation, *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1317-1325, 2008.
- [78] S. Krinidis, V. Chatzis, A robust fuzzy local information C-means clustering algorithm, *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1328-1337, 2010.
- [79] Z. Yu, An adaptive unsupervised approach toward pixel clustering and color image segmentation, *Pattern Recognition*, vol. 43, no. 5, pp. 1889-1906, 2010.
- [80] J. Fan, Automatic image segmentation by integrating color-edge extraction and seeded region growing, *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1454-1466, 2001.
- [81] S.Y. Wan, W.E. Higgins, Symmetric region growing, *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1007-1015, 2003.
- [82] A.K. Qin, D.A. Clausi, Multivariate image segmentation using semantic region growing with adaptive edge penalty, *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 2157-2170, 2010.
- [83] R. Nock, F. Nielsen, Statistical region merging, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452-1458, 2004.
- [84] Z. Liu, L. Shen, Z. Zhang, Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction, *Signal Processing*, vol. 91, no. 2, pp. 290-299, 2011.
- [85] F. Calderero, F. Marques, Region merging techniques using information theory statistical measures, *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1567-1586, 2010.

- [86] L.G. Ugarriza, Automatic image segmentation by dynamic region growth and multiresolution merging, *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2275-2288, 2009.
- [87] T. Celik, T. Tjahjadi, Unsupervised colour image segmentation using dual-tree complex wavelet transform, *Computer Vision and Image Understanding*, vol. 114, no. 7, pp. 813-826, 2010.
- [88] C. Panagiotakis, I. Grinias, G. Tziritas, Natural image segmentation based on tree equipartition, bayesian flooding and region merging, *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2276-2287, 2011.
- [89] A.C. Sobieranski, E. Comunello, A.V. Wangenheim, Learning a nonlinear distance metric for supervised region-merging image segmentation, *Computer Vision and Image Understanding*, vol. 115, no. 2, pp. 127-139, 2011.
- [90] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, *International Journal on Computer Vision*, vol. 1, no. 4, pp. 321-331, 1988.
- [91] A. Dumitras, A.N. Venetsanopoulos, Angular map-driven snakes with application to object shape description in color images, *IEEE Transactions on Image Processing*, vol. 10, no. 12, pp. 1851-1859, 2001.
- [92] F. Precioso, Robust real-time segmentation of images and videos using a smooth-spline snake-based algorithm, *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 910-924, 2005.
- [93] U. Ozertem, D. Erdogmus, Nonparametric snakes, *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2361-2368, 2007.
- [94] S. Lankton, A. Tannenbaum, Localizing region-based active contours, *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2029-2039, 2008.
- [95] X. Xie, M. Mirmehdi, MAC: magnetostatic active contour model, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 632-646, 2008.
- [96] R. Goldenberg, Fast geodesic active contours, *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1467-1475, 2001.

- [97] C. Sagiv, N.A. Sochen, Y.Y. Zeevi, Integrated active-contours for texture segmentation, *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1633-1646, 2006.
- [98] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape modeling with front propagation: a level set approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158-175, 1995.
- [99] T. Brox, J. Weickert, Level set segmentation with multiple regions, *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3213-3218, 2006.
- [100] I.B. Ayed, A. Mitiche, A region merging prior for variational level set image segmentation, *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2301-2311, 2008.
- [101] X. Xie, Active contouring based on gradient vector interaction and constrained level set diffusion, *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 154-164, 2010.
- [102] C. Li, Distance regularized level set evolution and its application to image segmentation, *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3243-3254, 2010.
- [103] M.B. Salah, A. Mitiche, I.B. Ayed, Effective level set image segmentation with a kernel induced data term, *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 220-232, 2010.
- [104] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.
- [105] A. Tsai, A. Yezzi, A.S. Willsky and et al., Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification, *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169-1186, 2001.
- [106] M. Heiler, C. Schnorr, Natural image statistics for natural image segmentation, *International Journal on Computer Vision*, vol. 63, no. 1, pp. 5-19, 2005.

-
- [107] S. Gao, T.D. Bui, Image segmentation and selective smoothing by using Mumford-Shah model, *International Journal on Computer Vision*, vol. 14, no. 10, pp. 1537-1549, 2005.
- [108] J. Mukherjee, MRF clustering for segmentation of color images, *Pattern Recognition Letters*, vol. 23, no. 8, pp. 917-929, 2002.
- [109] H. Deng, D.A. Clausi, Unsupervised image segmentation using a simple MRF model with a new implementation scheme, *Pattern Recognition*, vol. 37, no. 12, pp. 2323-2335, 2004.
- [110] Y. Xia, D. Feng, R. Zhao, Adaptive segmentation of textured images by using the coupled Markov random field model, *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3559-3566, 2006.
- [111] Z. Kato, T.C. Pong, A Markov random field image segmentation model for color textured images, *Image and Vision Computing*, vol. 24, no. 10, pp. 1103-1114, 2006.
- [112] C. Nikou, A.C. Likas, N.P. Galatsanos, A Bayesian framework for image segmentation with spatially varying mixtures, *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2278-2289, 2010.
- [113] M. Mignotte, A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation, *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1610-1624, 2010.
- [114] S. Chen, Image segmentation by MAP-ML estimations, *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2254-2264, 2010.
- [115] L. Zhang, Q. Ji, Image segmentation with a unified graphical model, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1406-1425, 2010.
- [116] S.H. Lee, H.I. Koo, N.I. Cho, Image segmentation algorithms based on the machine learning of features, *Pattern Recognition Letters*, vol. 31, no. 14, pp. 2325-2336, 2010.

- [117] C. Carson, Blobworld: image segmentation using expectationmaximization and its application to image querying, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026-1038, 2002.
- [118] N. Greggio, Fast estimation of Gaussian mixture models for image segmentation, *Machine Vision and Applications, Special Issue: Microscopy Image Analysis for Biomedical Applications*, vol. 23, no. 4, pp. 773-789, 2012.
- [119] Z. Liu, Color image segmentation using nonparametric mixture models with multivariate orthogonal polynomials, *Neural Computing and Applications*, vol. 21, no. 4, pp. 801-811, 2012.
- [120] Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: theory and its application to image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1101-1113, 1993.
- [121] J. Shi, J. Malik, Normalized cuts and image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, 2000.
- [122] W. Tao, H. Jin, Y. Zhang, Color image segmentation based on mean shift and normalized cuts, *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 37, no. 5, pp. 1382-1389, 2007.
- [123] S. Sarkar, P. Soundararajan, Supervised learning of large perceptual organization: graph spectral partitioning and learning automata, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 5, pp. 504-525, 2000.
- [124] S. Wang, J.M. Siskind, Image segmentation with minimum mean cut, *Proceeding of 2001 8th IEEE International Conference on Computer Vision*, vol. 1, pp. 517-524, 2001.
- [125] S. Wang, J.M. Siskind, Image segmentation with ratio cut, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 675-690, 2003.
- [126] J.S. Kim, K.S. Hong, Color texture segmentation using unsupervised graph cuts, *Pattern Recognition*, vol. 42, no. 5, pp. 735-750, 2009.
- [127] Y.Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut / max-flow algorithms for energy minimization in vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124-1137, 2004.

- [128] S. Han, Image segmentation based on grabcut framework integrating multiscale nonlinear structure tensor, *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2289-2302, 2009.
- [129] Q. Yang, Progressive cut: an image cutout algorithm that models user intentions, *IEEE Multimedia*, vol. 14, no. 3, pp. 56-66, 2007.
- [130] P.F. Felzenszwalb, D.P. Huttenlocher, Efficient graph-based image segmentation, *International Journal on Computer Vision*, vol. 59, no. 2, pp. 167-181, 2004.
- [131] L. Grady, E.L. Schwartz, Isoperimetric graph partitioning for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 469-475, 2006.
- [132] P.R. Hill, C.N. Canagarajah, D.R. Bull, Image segmentation using a texture gradient based watershed transform, *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1618-1633, 2003.
- [133] H.T. Nguyen, M. Worring, R. Boomgaard, Watersnakes: energydriven watershed segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 3, pp. 330-342, 2003.
- [134] I. Vanhamel, I. Pratikakis, H. Sahli, Multiscale gradient watersheds of color images, *IEEE Transactions on Image Processing*, vol. 12, no. 6, pp. 617-626, 2003.
- [135] J.B. Kim, H.J. Kim, Multiresolution-based watersheds for efficient image segmentation, *Pattern Recognition Letters*, vol. 24, no. 1-3, pp. 473-488, 2003.
- [136] P. Lang, K. Yeow, A. Nichols and et al., Cellular imaging in drug discovery, *Nature Reviews Drug Discovery*, vol. 5, pp. 343-356, 2006.
- [137] A. Bullen, Microscopic imaging techniques for drug discovery, *Nature Reviews Drug Discovery*, vol. 7, pp. 54-67, 2008.
- [138] V. Starkuviene, R. Pepperkok, The potential of high-content high-throughput microscopy in drug discovery, *British Journal of Pharmacology*, vol. 152, no. 1, pp. 62-71, 2007.
- [139] K.A. Baraa, J.M. Philip, M.N. Nicole and et al., An automated cell-counting algorithm for fluorescently-stained cells in migration assays, *Biological Procedures Online*, vol. 13, no. 9, pp. 1-5, 2011.

- [140] K. Bredies, H. Wolinski, An active-contour based algorithm for the automated segmentation of dense yeast populations on transmission microscopy images, *Computing and Visualization in Science*, vol. 14, no. 7, pp. 341-352, 2011.
- [141] V.T. Ta, O. Lezoray, A. Elmoataz and et al., Graph-based tools for microscopic cellular image segmentation, *Pattern Recognition*, vol. 42, no. 6, pp. 1113-1125, 2009.
- [142] P.M. Pilarski, C.J. Backhouse, Towards robust cellular image classification: theoretical foundations for wide-angle scattering pattern analysis, *Biomedical Optics Express*, vol. 1, no. 4, pp. 1225-1233, 2010.
- [143] L. Shamir, J.D. Delaney, N. Orlov and et al., Pattern recognition software and techniques for biological image analysis, *PLOS Computational Biology*, vol. 6, no. 11, pp. e1000974:1-10, 2010.
- [144] G. Dougherty, Digital image processing for medical applications, *Cambridge University Press, Cambridge*, 2009.
- [145] P. Sarder, A. Nehorai, Deconvolution methods for 3D fluorescence microscopy images, *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 32-45, 2006.
- [146] A. Papini, A new algorithm to reduce noise in microscopy images implemented with a simple program in python, *Microscopy Research and Technique*, vol. 75, no. 3, pp. 334-342, 2011.
- [147] Y.L. Wang, Computational restoration of fluorescence images: noise reduction, deconvolution, and pattern recognition, *Methods in Cell Biology*, vol. 72, pp. 337-348, 2003.
- [148] C.W. Chang, M.A. Mycek, Total variation versus wavelet-absed methods for image denoising in fluorescence lifetime imaging microscopy, *Journal of Biophotonics*, vol. 5, pp. 449-457, 2012.
- [149] C.W. Chang, M.A. Mycek, Precise fluorophore lifetime mapping in live-cell, multiphoton excitation microscopy, *Optical Express*, vol. 18, no. 8, pp. 8688-8696, 2010.
- [150] S.T. Li, L.Y. Fang, H.T. Yin, An efficient dictionary learning algorithm and its application to 3D medica image denoising, *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 417-427, 2012.

- [151] J.L. Starck, F. Murtagh, A. Bijioui, Image processing and data analysis: the multiscale approach, *Cambridge University Press*, 1988.
- [152] S.Y. Chen, M.Z. Zhao, G. Wu and et al., Recent Advances in Morphological Cell Image Analysis, *Computational and Mathematical Methods in Medicine*, vol. 2012, pp. 1-10, 2012.
- [153] Z. Pincus, J.A. Theriot, Comparison of quantitative methods for cell-shape analysis, *Journal of Microscopy*, vol. 227, no. 2, pp. 140-156, 2007.
- [154] E. Meijering, Cell Segmentation: 50 Years Down the Road, *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 140-145, 2012.
- [155] W. Jiang, M.L. Baker, Q. Wu and et al., Applications of a bilateral denoising filter in biological electron microscopy, *Journal of Structural Biology*, vol. 144, no. 1-2, pp. 114-122, 2003.
- [156] R.S. Pantelic, R. Rothnagel, C.Y. Huang and et al., The discriminative bilateral filter: an enhanced denoising filter for electron microscopy data, *Journal of Structural Biology*, vol. 155, no. 3, pp. 395-408, 2006.
- [157] J. Rajan, K. Kannan, M.R. Kaimal, An improved hybrid model for molecular image denoising, *Journal of Mathematical Imaging and Vision*, vol. 31, no. 1, pp. 73-79, 2008.
- [158] S. Schulte, B. Huysmans, A. Pizurica et al., A new fuzzy-based wavlet shrinkage image denoising technique, *Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science*, vol. 4179, pp. 12-23, 2006.
- [159] S. Schulte, V.D. Witte, E.E. Kerre, A fuzzy noise reuction method for color images, *IEEE Transactions on Image Processing*, vol. 16, no. 5, pp. 1425-1436, 2007.
- [160] M. Wilscy, M.S. Nair, A new fuzzy-based additive noise removal filter, *Advances in Electrical Engineering and Computational Science, Lecture Notes in Electrical Engineering*, vol. 39, pp. 87-97, 2009.
- [161] M.S. Nair, G. Raju, Additive noise removal using a novel fuzzy-based filter, *Computers and Electrical Engineering*, vol. 37, no. 5, pp. 644-655, 2011.

- [162] S. Paris, S.W. Hasinoff, J. Kautz, Local laplacian filters: edge-aware image processing with a laplacian pyramid, *ACM Transactions on Graphics*, vol. 30, no. 4, pp. 1-11, 2011.
- [163] I. Clark, Practical geostatistics, *Applied Science Publishers, London*, 1979.
- [164] J.P. Chiles, P. Delfiner, Geostatistics: modeling spatial uncertainty, *John Wiley and Sons, New York*, 1999.
- [165] M.B. Wilk, R. Gnanadesikan, Probability plotting methods for the analysis of data, *Biometrika*, vol. 55, no. 1, pp. 1-17, 1968.
- [166] B. Efron, R. Tibshirani, An Introduction to the Bootstrap, *Chapman and Hall, Boca Raton*, 1993.
- [167] H.J. Ader, G.J. Mellenbergh, D.J. Hand, Advising on research methods: a consultant's companion, *Johannes van Kessel Publishing, The Netherlands*, 2008.
- [168] J.S. Lim, Two-dimensional signal and image processing, *Prentice Hall, NJ*, 1990.
- [169] M. Riffle, T.N. Davis, The yeast resource center public image repository: a large database of fluorescence microscopy images, *BMC Bioinformatics*, vol. 11, no. 263, pp. 1-9, 2010.
- [170] Z. Wang, A.C. Bovik, H.R. Sheikh and et al., Image quality assessment: From error measurement to structural similarity, *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [171] M. Clarke, U. Engel, J. Giorgione and et al., Curvature recognition and force generation in phagocytosis, *BMC Bioinformatics*, vol. 8, no. 154, pp. 1-22, 2010.
- [172] Marker-controlled watershed segmentation, <http://www.mathworks.com.au/help/images/examples/marker-controlled-watershed-segmentation.html>
- [173] O. Dzyubachyk, W.A.V. Cappellen, J. Essers and et al., Advanced level-set-based cell tracking in time-lapse fluorescence microscopy, *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 852-867, 2010.
- [174] D.J. Stephens, V.J. Allan, Light microscopy techniques for live cell imaging, *Science*, vol. 300, no. 5616, pp. 82-86, 2003.

- [175] M.D. Acunto, O. Salvetti, Image analysis aided by segmentation algorithms for techniques ranging from scanning probe microscopy to optical microscopy, *Microscopy: Science, Technology, Applications and Education*, vol. 2, pp. 1455-1466, 2010.
- [176] A. Danckaert, E.G. Couto and L. Bollondi and et al., Automated recognition of intracellular organelles in confocal microscope images, *Traffic*, vol. 3, no. 1, pp. 66-73, 2002.
- [177] R.F. Murphy, M. Velliste and G. Porreca, Robust numerical features for description and classification of subcellular location patterns in fluorescence microscope images, *Journal of VLSI Signal Processing*, vol. 35, no. 3, pp. 311-321, 2003.
- [178] A.M. Base, Pattern recognition for medical imaging, *Elsevier Academic Press*, 2004.
- [179] G. Yang, X. Yu, X. Zhuang, The current status and development of pattern recognition diagnostic methods based on medical imaging, *IEEE International Conference on Networking, Sensing and Control*, pp. 567-572, 2008.
- [180] A.P. Dhawan, Medical image analysis, *Wiley-IEEE Press*, 2011.
- [181] P. Corr, Pattern recognition in diagnostic imaging, *World Health Organization*, 2001.
- [182] D.Y. Tsai, Y.B. Lee, S. Masaru and et al., Medical image classification using genetic-algorithm based fuzzy-logic approach, *Journal of Electronic Imaging*, vol. 13, no. 4, pp. 780-788, 2004.
- [183] T. Le, D. Tran, W. Ma and et al., A new support vector machine method for medical image classification, *Proceeding of 2010 2nd European Workshop on Visual Information Processing*, pp. 165-170, 2010.
- [184] H. Wu, H. Zhang, C. Li, Medical image classification with multiple kernel learning, *Proceeding of 2010 2nd International Conference on Internet Multimedia Computing and Service*, pp. 189-192, 2010.
- [185] P. Hossein, G. Hassan, Content-based medical image classification using a new hierarchical merging scheme, *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 651-661, 2008.

- [186] Y.G. Zhang, B.L. Zhang, F. Coenen and et al., One-class kernel subspace ensemble for medical image classification, *Eurasip Journal on Advances in Signal Processing*, vol. 2014, no. 17, pp. 1-13, 2014.
- [187] G. Castellano, L. Bonilha, L.M. Li and et al., Texture analysis of medical images, *Clinical Radiology*, vol. 59, no. 12, pp. 1061-1069, 2004.
- [188] K.C. Jayashree, H. William, Effectiveness of global features for automatic medical image classification and retrieval - the experiences of OHSU at ImageCLEFmed, *Pattern Recognition Letters*, vol. 29, no. 15, pp. 2032-2038, 2008.
- [189] J.D. Deng, Improving feature extraction for automatic medical image categorization, *Proceeding of 2009 24th International Conference on Image and Vision Computing*, pp. 379-384, 2009.
- [190] T.D. Pham, D.T.P. Le, J.W. Xu and et al., Personalized identification of abdominal wall hernia meshes on computed tomography, *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 153-161, 2014.
- [191] R. Parekh, Using texture analysis for medical diagnosis, *IEEE Transaction on Multimedia*, vol. 19, no. 2, pp. 28-37, 2012.
- [192] T. Tuytelaars, K. Mikolajczyk, Local invariant feature detectors: a survey, *Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177-280, 2007.
- [193] J. Zhang, H. Zhao, J.M. Liang, Continuous rotation invariant local descriptors for texton dictionary-based texture classification, *Computer Vision and Image Understanding*, vol. 117, no. 1, pp. 56-75, 2013.
- [194] Z.L. Zhu, C.X. Zhao, Y.K. Hou and et al., Rotation invariant texture image retrieval based on combined feature sets, *International Journal of Digital Content Technology and its Applications*, vol. 5, no. 3, pp. 287-292, 2011.
- [195] G.Y. Zhao, T. Ahonen, J. Matas and et al., Rotation invariant image and video description with local binary pattern features, *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1465-1477, 2012.

- [196] M. Varma, A. Zisserman, A statistical approach to material classification using image patches, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 2032-2047, 2009.
- [197] M. Varma, A. Zisserman, A statistical approach to texture classification from single images, *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61-81, 2005.
- [198] C. Schmid, Constructing models for content-based image retrieval, *Proceedings of 2001 IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 39-45, 2001.
- [199] T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29-44, 2001.
- [200] R. Krishnamoorthi, S. Sathiya, A multiresolution approach for rotation invariant texture image retrieval with orthogonal polynomials model, *Journal of Visual Communication and Image Representation*, vol. 23, no. 1, pp. 18-30, 2012.
- [201] C. Liu and R. Szeliski, S.B. Kang and et al., Automatic estimation and removal of noise from a single image, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 299-314, 2008.
- [202] Z. Li, G.Z. Liu, Y. Yang and et al., Scale and rotation-invariant local binary pattern using scale adaptive texton and subuniform-based circular shift, *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2130-2140, 2012.
- [203] X.J. Li, J.W. Zhang, Y.D. Ma, Texture classification using combined rotation and scale invariant Gabor features, *Journal of Computational Information Systems*, vol. 7, no. 12, pp. 4241-4253, 2011.
- [204] X.D. Xie, Q.H. Dai, K.M. Lam and et al., Efficient rotation and scale invariant texture classification method based on Gabor wavelets, *Journal of Electronic Imaging*, vol. 17, no. 4, pp. 1-7, pp. 2008.
- [205] J. Han, K.K. Ma, Rotation invariant and scale invariant Gabor features for texture image retrieval, *Image and Vision Computing*, vol. 25, no. 9, pp. 1474-1481, 2007.

- [206] J.L. Crowley, A.C. Sanderson, Multiple resolution representation and probabilistic matching of 2D gray scale shape, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 113-121, 1987.
- [207] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-985, 2002.
- [208] Z.H. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1657-1663, 2010.
- [209] P. Janney, G. Geers, Texture classification using invariant features of local textures, *IET Image Processing*, vol. 4, no. 3, pp. 158-171, 2010.
- [210] J. Chen, S.G. Shan, C. He and et al., WLD: a robust local image descriptor, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705-1720, 2010.
- [211] F.R. Siqueira, W.R. Schwartz, H. Pedrini and et al., Multi-scale gray level co-occurrence matrices for texture description, *Neurocomputing*, vol. 120, pp. 336-345, 2013.
- [212] T. Lindeberg, Scale-space theory: a basic tool for analysing structures at different scales, *Journal of Applied Statistics*, vol. 21, no. 2, pp. 224-270, 1994.
- [213] T. Lindeberg, Feature detection with automatic scale selection, *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79-116, 1998.
- [214] T. Lindeberg, Scale selection, *Computer Vision: A Reference guide*, Springer, 2014.
- [215] H. Bay, A. Ess, T. Tuytelaars and et al., SURF: speeded up robust features, *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.
- [216] C. Harris, M. Stephens, A combined corner and edge detector, *Proceedings of 1988 4th Alvey Vision Conference*, pp. 147-151, 1988.
- [217] J. Shi, C. Tomasi, Good features to track, *Proceedings of 1994 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593-600, 1994.

- [218] E. Rosten, T. Drummond, Fusing points and lines for high performance tracking, *Proceedings of 2005 IEEE International Conference on Computer Vision*, vol. 2, pp. 1508-1511, 2005.
- [219] S. Leutenegger, M. Chli, R. Siegwart, BRISK: binary robust invariant scalable keypoints, *Proceedings of 2011 IEEE International Conference on Computer Vision*, pp. 2548-2555, 2011.
- [220] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [221] D. Unay, A. Ekin, R.S. Jasinschi, Local structure-based region of interest retrieval in brain MR images, *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 897-903, 2010.
- [222] J.M. Aman, J.H. Yao, R.M. Summers, Content-based image retrieval on CT colonography using rotation and scale invariant features and bag-of-words model, *Proceedings of 2010 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1357-1360, 2010.
- [223] D. Hlindzich, R. Maenner, Medical feature matching and model extraction from MRI / CT based on the invariant generalized hough / radon transform, *IFMBE Proceedings of 4th European Conference of the International Federation for Medical and Biological Engineering*, vol. 22, pp. 608-612, 2008.
- [224] B. Anja, R. Rainer, F. Thomas and et al., Wavelet based noise reduction in CT-images using correlation analysis, *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1685-1703, 2008.
- [225] A.F. Santiago, A.L. Carlos, C.F. Westin, Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach, *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1383-1398, 2008.
- [226] P.L. Shui, W.C. Zhang, Noise-robust edge detector combining isotropic and anisotropic Gaussian kernels, *Pattern Recognition*, vol. 45, no. 2, pp. 806-820, 2012.

- [227] T. Lindeberg, J. Garding, Shape-adapted smoothing in estimation of 3D depth cues from affine distortions of local 2D structure, *Image and Vision Computing*, vol. 15, no. 6, pp. 415-434, 1997.
- [228] A. Almansa, T. Lindeberg, Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection, *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2027-2042, 2000.
- [229] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, *Proceedings of 2005 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 886-893, 2005.
- [230] J.R. Carr, F.P. de Miranda, The semivariogram in comparison to the co-occurrence matrix for classification of image texture, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 6, pp. 1945-1952, 1998.
- [231] A. Yue, C. Zhang, J. Yang and et al., Texture extraction for object-oriented classification of high spatial resolution remotely sensed images using a semivariogram, *International Journal of Remote Sensing*, vol. 34, no. 11, pp. 3736-3759, 2013.
- [232] X. Wang, T.X. Han, S. Yan, A HOG-LBP human detector with partial occlusion handling, *Proceedings of 2009 IEEE International Conference on Computer Vision*, pp. 32-39, 2009.
- [233] W. Zuo, L. Zhang, C. Song and el at., Texture enhanced image denoising via gradient histogram preservation, *Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1203-1210, 2013.
- [234] H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction, *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349-366, 2007.
- [235] P. Chatterjee, P. Milanfar, Clustering-based denoising with locally learned dictionaries *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1438-1451, 2009.
- [236] P.P. Kastendeuch, Pressure gradient force, atmospheric circulation and climate in western Europe (1899-2002), *International Journal of Climatology*, vol. 27, no. 15, pp. 2055-2067, 2007.

- [237] C.E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, 1948.
- [238] O. Christensen, An introduction to frames and riesz bases, *Birkhauser*, 2003.
- [239] R. Balan, P. Casazza, C. Heil and et al., Density, overcompleteness, and localization of frames, *The Journal of Fourier Analysis and Applications*, vol. 12, no. 2, pp. 105-143, 2006.
- [240] B. Leon, Science and information theory, *Courier Dover Publications*, 2004.
- [241] Q. Shan, Z.R. Li, J.Y. Jia and et al., Fast image / video upsampling, *ACM Transactions on Graphics*, vol. 27, no. 5, pp. 153:1-8, 2008.
- [242] Y.B. Zhang, D.B. Zhao, J. Zhang and et al., Interpolation dependent image downsampling, *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3291-3296, 2011.
- [243] K.M. Horton, R.A. Abrams, E.K. Fishman, Spiral CT of colon cancer: imaging features and role in management, *The Journal of Continuing Medical Education in Radiology*, vol. 20, no. 2, pp. 419-430, 2000.
- [244] R.N. Okaili, J. Krejza, and S.M. Wang et al., Advanced MR imaging techniques in the diagnosis of intraaxial brain tumor in adults, *The Journal of Continuing Medical Education in Radiology*, vol. 26, no. s1, pp. s173-s189, 2006.
- [245] H. Faulds, On the skin-furrows of the hand, *Nature*, XXII: 605, 1880.
- [246] D. Maltoni, D. Maio, A.K. Jain and et al., *Hand book of fingerprint recognition*, Springer Verlag, 2009.
- [247] P. Komarinski, Automated fingerprint identification systems (AFIS), *Elsevier Academic Press*, 2005.
- [248] Y. Wang, J.K. Hu, Global ridge orientation modelling for partial fingerprint identification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 72-87, 2011.
- [249] A.K. Jain, J.J. Feng, K. Nandakumar, On matching latent fingerprints, *Proceedings of IEEE Computer Vision and Pattern Recognition Workshop Biometrics*, pp. 1-8, 2008.

- [250] A.K. Jain, J.J. Feng, Latent fingerprint matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 88-100, 2011.
- [251] A.A. Paulino, J.J. Feng, A.K. Jain, Latent fingerprint matching using descriptor-based hough transform, *Proceedings of IEEE International Joint Conference on Biometrics Compendium*, pp. 1-7, 2011.
- [252] A.A. Paulino, J.J. Feng, A.K. Jain, Latent fingerprint matching using descriptor-based hough transform, *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 31-45, 2013.
- [253] K. Cao, E.Y. Liu, A.K. Jain, Segmentation and enhancement of latent fingerprints: a coarse to fine ridge structure dictionary, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 9, pp. 1847-1859, 2014.
- [254] J.Y. Zhang, R.J. Lai, C.J. Kuo, Adaptive directional total-variation model for latent fingerprint segmentation, *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 8, pp. 1261-1273, 2013.
- [255] J.Y. Zhang, R.J. Lai, C.J. Kuo, Latent fingerprint detection and segmentation with a directional total variation model, *Proceedings of IEEE International Conference on Image Processing*, pp. 1145-1148, 2012.
- [256] S.K. Ashtiani, C.J. Kuo, A robust technique for latent fingerprint image segmentation and enhancement, *Proceedings of IEEE International Conference on Image Processing*, pp. 1492-1495, 2008.
- [257] H. Choi, M. Boaventura, I.A.G. Boaventura and et al., Automatic segmentation of latent fingerprints, *Proceedings of IEEE International Conference on Biometrics Compendium*, pp. 303-310, 2012.
- [258] N.J. Short, M.S. Hsiao, A.L. Abbott and et al., Latent fingerprint segmentation using ridge template correlation, *Proceedings of International Conference on Imaging for Crime Detection and Prevention*, pp. 1-6, 2011.
- [259] X.J. Tan, B. Bhanu, Fingerprint matching by genetic algorithms, *Pattern Recognition*, vol. 39, no. 3, pp. 465-477, 2006.
- [260] K. Engan, S.O. Aase, J.H. Husoy, Multi-frame compression: theory and design, *Signal Processing*, vol. 80, no. 10, pp. 2121-2140, 2000.

-
- [261] I. Todic, P. Frossard, Dictionary learning, *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27-38, 2011.
- [262] J. Mairal, F. Bach, J. Ponce and et al., Online dictionary learning for sparse coding, *Proceedings of International Conference on Machine Learning*, pp. 689-696, 2009.
- [263] J. Tropp, Greed is good: Algorithmic results for sparse approximation, *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231-2242, 2004.
- [264] E. Agoston, S. James, Introduction to evolutionary computing, *Springer*, 2003.
- [265] B.H. Li, Q.G. Meng, H. Holstein, Point pattern matching and applications - a review, *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, pp. 729-736, 2003.
- [266] L.H. Zhang, W.L. Xu, C. Chang, Genetic algorithm for affine point pattern matching, *Pattern Recognition Letters*, vol. 24, no. 1-3, pp. 9-19, 2003.
- [267] J.W. Xu, J.K. Hu, X.P. Jia, Genetic algorithm for distorted point set matching, *Proceedings of International Congress on Image and Signal Processing*, vol. 3, pp. 1724-1729, 2013.
- [268] 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.