

NOVEL FILTERING METHODS FOR IMAGE AND VIDEO PROCESSING APPLICATIONS

by

Muhammad Talal Ibrahim

M.Sc., Computer Engineering, UET, Taxila, Pakistan, 2006

B.Sc., Software Engineering, NUML, Islamabad, Pakistan, 2004

A dissertation

presented to Ryerson University

in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

in the Program of

Electrical and Computer Engineering

Toronto, Ontario, Canada, 2012

©Muhammad Talal Ibrahim 2012

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

Novel Filtering Methods for Image and Video Processing Applications

Doctor of Philosophy 2012

Muhammad Talal Ibrahim

Electrical and Computer Engineering

Ryerson University

Abstract

During the last few years, digital filtering methods for image/video processing applications have reached a satisfactory level. However, their performance degrades in the presence of noise, trend, motion, shape deformation, intensity inhomogeneity, shadows, or low image quality, to name a few. To cope with these challenges, this dissertation presents novel filtering methods for image/video processing applications that outperform the existing and state-of-the-art methods.

The dissertation starts by introducing a novel trend filtering method that transforms the inter-frame registration problem into low complexity trend filtering problem. In the proposed method, Laplacian eigenmaps in conjunction with the modified empirical mode decomposition has been used to suppress the noise artifacts and the trend term.

In multi-dimensional signals, the trend term is often referred to as non-uniform illumination or global intensity inhomogeneity. This dissertation presents a new filtering method for estimating the global intensity inhomogeneity in two dimensional and volume images. Global intensity inhomogeneity often arises due to the imperfections of data acquisition

device, direction of source light, and properties of the subject under study. The proposed method generates a high-pass filter based on the grey-weighted distance transform of the frequency content of an image/volume. It provides an accurate estimation of global intensity inhomogeneity without any parameter tweaking, which makes it applicable to many imaging modalities.

The dissertation also presents a filtering methodology to cope with local intensity inhomogeneity that gives rise to shadow artifacts. These artifacts appear as sharp discontinuities and are often corrected at different scales and orientations. The proposed method makes use of decimation-free directional filter bank to suppress the local intensity inhomogeneity and shadow artifacts irrespective of scale and orientation.

In addition to intensity inhomogeneity correction, the dissertation also presents a filtering method that utilizes the Gabor filter bank to generate rotation invariant feature codes. The effectiveness of the proposed method has been evaluated in both identification and verification modes for fingerprint recognition.

The uniqueness of the presented filtering methods lies in the fact that they are essentially parameter free and can easily be scaled to higher dimensions. This makes them applicable to many different image/video processing applications with least of effort from the end user, i.e., eliminating the user biases.

Acknowledgement

I would like to express my sincere gratitude to Prof. Ling Guan, my dissertation supervisor, for his invaluable guidance, excellent suggestions, constant encouragement, and kind understanding during the dissertation research period. I would like to thank Prof. Anastasios N. Venetsanopoulos and Prof. Liang Zhao for taking the time and effort to review this dissertation and provide me with their insightful comments. I would also like to express my sincere gratitude to Prof. Jiying Zhao and Prof. Bruce Elder for their invaluable time to serve as external examiners. I would also like to thank Prof. Soosan Beheshti for her guidance during dissertation writing.

I would also like to take this opportunity to thank the faculty members in the Electrical and Computer Engineering Department for their constant support during my period of study at Ryerson. I would like to acknowledge the Department of Electrical and Computer Engineering, Ryerson University, providing me financial support throughout my research work. I would like to express my great thanks to Prof. Ingela Nyström from Centre for Image Analysis, Uppsala University, Sweden for the technical and moral support during the last two years. I would also like to express my special thanks to Dr. M. Khalid Khan Niazi for the technical and moral support during my studies.

I am grateful to the members of the Ryerson Multimedia Processing Lab in Toronto. In particular, Matthew Kyan, Yongjin Wang, Rui Zhang, Ning Zhang, Yun Tie, Xiaoming Nan, Naimul Mefraz Khan, Yifeng He, Chunhao Wang, Dong Nan, Adrian Bulzacki, Naimul Mefraz Khan, Harmanjot Singh Sandhu, Zhibing Xie, and Thambu Kuganeswaran. I thank

you all for your friendship, unconditional support and valued assistance.

For all others who, more or less, helped me in this work, please be assured that I am forever grateful to you all.

Last but not the least, I would like to express my heartiest gratitude, and love to my family for their great love, continuous support and encouragement during my graduate studies.

Dedication

To my father Muhammad Ibrahim (late) for his unconditional love, sacrifice, hard work and strength.

To my mother Rehana Bibi for her unconditional love, support and patience.

To my brother Muhammad Bilal and his family for supporting me throughout my stay in Canada.

To my sister Razia Sultana who is a true blessing of Allah.

Without all of you, this work would have not been possible. Thank you all for sharing this dream.

Contents

1	Introduction	1
1.1	Trend Filtering	2
1.2	Global Intensity Inhomogeneity Correction	3
1.3	Local Intensity Inhomogeneity Correction	5
1.4	Image filtering based feature extraction	7
1.5	Contributions	9
1.6	Organization	11
2	Materials and Methods	12
2.1	Digital Signal/Image Filtering	12
2.1.1	Two Dimensional Sinusoids	13
2.1.2	Two Dimensional Fourier Transform	15
2.2	Directional Analysis	16
2.2.1	Gabor Filter	17
2.2.2	Directional Filter Bank (DFB)	17
2.3	Empirical Mode Decomposition (EMD)	23
2.4	Clustering	26
2.4.1	K-Medoids Clustering	26
2.4.2	Fuzzy C-Means Clustering	27
2.5	Dimensionality Reduction	28
3	Novel Trend Filtering Method	31
3.1	Introduction	31
3.2	Proposed Method	33
3.2.1	Heart Signal Generation	33
3.2.2	Heart Signal Analysis	35

3.3	Results and Discussion	38
3.4	Summary	39
4	Global Intensity Inhomogeneity Correction	42
4.1	Introduction	42
4.2	Proposed Method for Global Intensity Inhomogeneity Correction	43
4.3	Experimental Setups	48
4.3.1	Experimental Setup 1	48
4.3.2	Experimental Setup 2	50
4.3.3	Experimental Setup 3	51
4.3.4	Experimental Setup 4	53
4.3.5	Experimental Setup 5	53
4.4	Results and Discussion	55
4.4.1	Results of Setup 1	55
4.4.2	Results of Setup 2	57
4.4.3	Results of Setup 3	60
4.4.4	Results of Setup 4-5	68
4.5	Global Intensity Inhomogeneity Correction in Volume Images	75
4.5.1	Results and Discussion	77
4.6	Summary	78
5	Local Intensity Inhomogeneity Correction	80
5.1	Introduction	80
5.2	Proposed Method	81
5.2.1	Decimation-Free Directional Filter Banks (DDFB)	82
5.3	Results and Discussion	85
5.3.1	Database Preparation	85
5.3.2	Matching	87
5.4	Summary	88
6	Filter-Bank based Rotation Invariant Feature Codes	91
6.1	Introduction	91
6.2	Proposed System for Fingerprint Recognition	93
6.2.1	Fingerprint Image Enhancement	94
6.2.2	Core Point Detection	96

6.2.3	ROI and Feature Extraction	98
6.2.4	Dimensionality Reduction	102
6.3	Experimental Results	103
6.3.1	Identification Results	103
6.3.2	Verification Results	104
6.4	Discussion	107
6.5	Summary	107
7	Conclusions and Future Work	109
7.1	Conclusions	109
7.2	Future Work	112
	Bibliography	114

List of Tables

4.1	Segmentation Results for DRIVE Database.	63
4.2	Vessel extraction results on the DRIVE database.	66
4.3	Vessel extraction results on the STARE database.	66
4.4	Segmentation Results for STARE Database.	67
4.5	White matter and grey matter restoration accuracy of various inhomogeneity levels (IL) in terms of cross-correlation coefficient	73
4.6	2D - White and grey matter restoration accuracy in terms of CC and CJV .	74
4.7	3D - White and grey matter restoration accuracy in terms of CC and CJV .	77
5.1	Recognition rates (%) when using images of subset 1 as training set	90
6.1	Details of FVC2002 Database Set A	103
6.2	Identification results (CRR in %).	104
6.3	Verification results of user-independent threshold (EER in %).	105
6.4	Verification results of user-specific thresholding schemes (EER in %).	106
6.5	EER (%) comparison of existing methods with proposed system on FVC2002 set_a	107

List of Figures

2.1	a) 2D sinusoid. b) Cross sectional view of 2.1(a).	14
2.2	Examples of 2D sinusoids of varying frequencies of size 255×255 with $A = 10$ and $\phi = 0$	14
2.3	a) Original TEM image of a c-elegans. b) Magnitude of the CSFT of 2.3(a).	16
2.4	Power spectrum of the 2D Gabor filter.	18
2.5	Passbands are represented by the greyish area. a) Frequency partition map of DFB [85]. b) Frequency partition map of Gabor filter bank.	18
2.6	3-stage DFB structure proposed in [85]. a) DFB modulated structure. b) DFB non-modulated structure.	19
2.7	Example of Quincunx downsampled image. a) Cameraman. b) Downsampled by Q	20
2.8	Types of ideal passbands for DFB structure proposed in [85]. a) For modulated structure shown in Fig. 2.6(a). b) For non-modulated structure shown in Fig. 2.6(b).	21
2.9	Sang DFB structure proposed in [85]. a) First stage. b) Second stage. c) Third stage.	22
2.10	3-stage DDFB structure proposed in [87].	23
2.11	DDFB structure proposed in [87]. a) First stage. b) Second stage. c) Third stage.	24
3.1	a) Number of heart beats counted by different methods on 151 videos. b) Difference between G_t and the number of beats counted by the method proposed in [110]. c) Difference between G_t and the number of beats counted by the methods proposed in [104, 111]. d) Standard deviation of the inter-beat time computed individually for every video.	40

4.1	a) Original TEM image of a c-elegans. b) Magnitude of the CSFT of 4.1(a). c) Result after applying GWDt on 4.1(b).	46
4.2	Percentage of the image power for a ball of radius \mathbf{r} . Here, the threshold value ($T = 4$) is highlighted in red.	47
4.3	a) Ideal low-pass filter. b) Zoomed version of the filter in 4.3(a). c) Smoothed version of the filter in 4.3(b). d) High-pass filter.	49
4.4	a) Binary retinal image $B(x, y)$ from the DRIVE database [120]. b) Global intensity inhomogeneity pattern $P(x, y)$. c) Global intensity inhomogeneity image $I_n(x, y)$ produced by adding the global intensity inhomogeneity pattern of 4.4(b) to the binary retinal image.	49
4.5	Synthetic image $S_g(x, y)$, created by using Gaussians of varying sigmas. . . .	51
4.6	a) Global intensity inhomogeneity pattern created by using Eq. 4.11. b) Global intensity inhomogeneity pattern created by a second order polynomial. c) Global intensity inhomogeneity pattern created by using Eq. 4.13. d) Global intensity inhomogeneity image created by adding 4.6(a) with Fig. 4.5. e) Global intensity inhomogeneity image created by adding 4.6(b) with Fig. 4.5. f) Global intensity inhomogeneity image created by adding 4.6(c) with Fig. 4.5.	52
4.7	a), b) Images from the DRIVE database [120]. c), d) Images from the STARE database [121]. e), f) Images from the VICAVR database [122].	54
4.8	TEM image of c-elegans suffering from global intensity inhomogeneity.	55
4.9	Transaxial view of a slice from the brain MR volume. a) Original image from the brainweb database [123]. b) Result after applying the method proposed in [41]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.	56

4.10	Percentage histograms. Here, the horizontal axis represents the percentage of correctly classified white pixels, while the vertical axis represents the number of test images. a) Histogram after application of the proposed method. The histogram shows that the proposed method was able to recover 98.5% of the pixels. b) Results of applying direct Otsu's thresholding on images with intensity inhomogeneity pattern created by standard normal distribution. The histogram shows that one can correctly recover 12% of the pixels. c) Result of the proposed method when a sinusoidal wave was added as an inhomogeneous pattern. Here, our method was able to recover 100% of the pixels. d) Results of applying direct Otsu's thresholding on images with intensity inhomogeneity pattern created by using the sinusoidal wave. On average, one can only recover 18% of the pixels in an image.	58
4.11	a) Original image. b) Global intensity inhomogeneity pattern created by a third degree polynomial. c) Global intensity inhomogeneity image. d) Adjusted image after 5 th iteration of the proposed method.	59
4.12	Histograms, where the horizontal axis represents the cross-correlation value between two images, while the vertical axis represent the number of test images. a) Cross-correlation of images produced by our method and the ground truth. Here, the images were corrupted by third degree polynomials. The histogram is centered around 0.84. This clearly shows that our results are quite similar to the ground truth. b) Cross-correlation between the ground truth and the intensity inhomogeneity images produced using third degree polynomials. The histogram is centered around 0.06.	60
4.13	a) Global intensity inhomogeneity pattern created by a sinusoidal wave. b) Global intensity inhomogeneity image created by adding 4.13(a) and Fig. 4.5. c) Corrected image at 5 th iteration of the proposed method.	61
4.14	Histogram images. a) Cross-correlation of images produced by our method and the ground truth. Here, the images were corrupted by using sinusoids. The histogram is centered around 0.96 which clearly shows the ability of our method to restore the original image. b) Cross-correlation between the ground truth and intensity inhomogeneity images produced by using sinusoids. Here, the mean value is centered around 0.15.	62

4.15	a), e), i), m) Original images from DRIVE database. b), f), j), n) Results after applying [35]. c), g), k), o) Results after applying [36]. d), h), l), p) Results after applying proposed method.	64
4.16	a), e), i), m) Original images from STARE database. b), f), j), n) Results after applying [35]. c), g), k), o) Results after applying [36]. d), h), l), p) Results after applying proposed method.	65
4.17	a) Original image from the VICAVR database [122]. b) Result after applying the method proposed in [35]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.	68
4.18	a) Original image from the VICAVR database [122]. b) Result after applying the method proposed in [35]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.	69
4.19	Average subjective evaluation of the retinal images after applying the proposed method.	70
4.20	a) Average subjective evaluation of TEM images. b) Average subjective evaluation of the brain MR images.	71
4.21	a) Original image suffering from global intensity inhomogeneity. b) Result after applying HUM [11]. c) Result after applying the method proposed in [35]. d) Result after applying the method mentioned in [36]. e) Result produced by the proposed method.	72
4.22	a) A frontal view of a slice from a 3D MRI volume corrupted with 90% global intensity inhomogeneity. b) The same slice of the volume after applying the proposed method. Significant improvement is visible in the lower part of slice.	78
5.1	Original image.	81
5.2	a), b) Original images of the same subject. c), d) Images having shadow artifacts after gamma transformation with $\gamma^* = 5$	83
5.3	a) Frequency response of 2D low-pass filter. b) Frequency response of 2D filter after applying Quincunx downsampling matrix. c) Frequency response of $v(x,y)$. d) Frequency response of $h(x,y)$	84
5.4	a), e) Original images of the same subject. b), f) Images having shadow effects after gamma transformation with $\gamma^* = 5$. c), g) Horizontal features after applying DDFB on the normalized images. d), h) Vertical features after applying DDFB on the normalized images.	86

5.5	a) Some images of subset 5 of a randomly selected subject from the Yale Face Database B. b) Results of the proposed system on 5.5(a).	87
5.6	a) Average recognition rate for horizontal features. b) Average recognition rate for vertical features.	89
6.1	Block diagram of biometric recognition systems.	92
6.2	Block diagram of the proposed system.	94
6.3	Block diagram of the enhancement method proposed in [150].	95
6.4	Images from FVC2002 <i>set_a</i> : a) Original image 1_1 from Db_1a. b) Original image 1_2 from Db_1a. c) Result of enhanced image 1_1. d) Result of enhanced image 1_2.	97
6.5	Images from FVC2002 <i>set_a</i> : a) Enhanced image 1_1 from Db_1a. b) Enhanced image 1_2 from Db_1a. c) Detected core point in image 1_1. d) Detected core point in image 1_2.	99
6.6	a) Detected core point represented by black circle. b) Region of Interest. c) to f) Results after applying four out of eight oriented Gabor filters on Fig. 6.6(b).101	
6.7	AAD features corresponding to the images c) to f) shown in Fig. 6.6, respectively.102	

List of Abbreviations

Abbreviation	Description
1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
AAD	Absolute Average Deviation
CC	Coefficient of Contrast
CJV	Coefficient of Joint Variation
CRR	Correct Recognition Rate
CSFT	Center Shifted Fourier Transform
DDFB	Decimation-Free Directional Filter Bank
DFB	Directional Filter Bank
DFT	Discrete Fourier Transform
EER	Equal Error Rate
EMD	Empirical Mode Decomposition
FAR	False Acceptance Rate
FCM	Fuzzy C-Means
FFT	Fast Fourier Transform
FOV	Field of View
FPR	False Positive Rate
FRR	False Rejection Rate
GD	Gestation Day
GM	Grey Matter
GWDT	Grey-weighted Distance Transform
HUM	Homomorphic Unsharp Masking
LDA	Linear Discriminant Analysis

LEM	Laplacian Eigenmaps
MRI	Magnetic Resonance Imaging
PCA	Principal Component Analysis
ROI	Region of Interest
TEM	Transmission Electron Microscopy
TPR	True Positive Rate
WM	White Matter

List of Publications

Journal Publications

1. M. K. K. Niazi, **M. T. Ibrahim**, I. Nyström, and L. Guan, “An Iterative Method for Intensity Inhomogeneity Correction Based on the Grey-weighted Distance Transform of the Magnitude Spectrum,” submitted Tier-1 Journal in Image Processing, 2012.
(M. K. K. Niazi and M. T. Ibrahim have contributed equally in this paper.)
2. **M. T. Ibrahim**, T. M. Khan, S. A. Khan, M. A. Khan, and L. Guan, “Iris Localization Using Local Histogram and Other Image Statistics,” Journal of Optics and Lasers in Engineering, vol. 50, pp. 645-654, Feb. 2012.
3. **M. T. Ibrahim**, Y. Wang, L. Guan, and A. N. Venetsanopoulos, “A Filter Bank Based Approach for Rotation Invariant Fingerprint Recognition,” Journal of Signal Processing Systems, DOI: 10.1007/s11265-011-0630-x, 2011.
4. **M. T. Ibrahim**, M. A. Khan, K. S. Alimgeer, M. K. Khan, I. A. Taj, and L. Guan, “Velocity and pressure-based partitions of horizontal and vertical trajectories for on-line signature verification,” Pattern Recognition, vol. 43(8), pp. 2817-2832, 2010.
5. L. Guan, P. Muneesawang, Y. Wang, R. Zhang, T. Yun, A. Bulzacki, and **M. T. Ibrahim**, “Multimodal information fusion for selected multimedia applications,” International Journal of Multimedia Intelligence and Security, vol. 1(1), pp. 5-32, 2010.

Conference Publications

6. **M. T. Ibrahim**, M. K. K. Niazi, and L. Guan, “Horizontal Features based Illumination Normalization Method for Face Recognition,” in *IEEE International Symposium on Image and Signal Processing and Analysis (ISPA 2011)*, pp. 684-689, 2011.
7. M. K. K. Niazi, **M. T. Ibrahim**, M. F. Nilsson, A. Sköld, L. Guan, and I. Nystörm, “Robust Signal Generation and Analysis of Rat Embryonic Heart Rate in Vitro Using Laplacian Eigenmaps and Empirical Mode Decomposition,” in *14th International Conference on Computer Analysis of Images and Pattern (CAIP 2011)*, vol. 6855 of Lecture Notes in Computer Science, pp. 523-530, Springer-Verlag Berlin/Heidelberg, 2011.
8. **M. T. Ibrahim**, Y. Wang, L. Guan, and A. N. Venetsanopoulos, “Fingerprint Verification Using Rotation Invariant Feature Codes,” in *International Conference on Image Analysis and Recognition (ICIAR 2011)*, vol. 6754 of Lecture Notes in Computer Science, pp. 111-119, Springer-Verlag Berlin/Heidelberg, 2011.
9. **M. T. Ibrahim**, T. M. Khan, M. A. Khan, and L. Guan, “A Novel and Efficient Feedback Method for Pupil and Iris Localization,” in *International Conference on Image Analysis and Recognition (ICIAR 2011)*, vol. 6754 of Lecture Notes in Computer Science, pp. 79-88, Springer-Verlag Berlin/Heidelberg, 2011.
10. M. K. K. Niazi, **M. T. Ibrahim**, L. Guan, and I. Nyström, “Bias field correction using grey-weighted distance transform applied on MR volumes,” in *IEEE International Symposium on Biomedical Imaging (ISBI 2011)*, pp. 357-360, 2011.
11. **M. T. Ibrahim**, M. J. Kyan, M. A. Khan, and Ling Guan, “On-Line Signature Verification Using 1-D Velocity-Based Directional Analysis,” in *IEEE International Conference on Pattern Recognition (ICPR 2010)*, pp. 3830-3833. 2010.
12. **M. T. Ibrahim**, T. M. Khan, M. A. Khan, and L. Guan, “Automatic Segmentation of Pupil Using Local Histogram and Standard Deviation,” in *Proceedings of Visual Communications and Image Processing (VCIP 2010)*, vol. 7744, pp. 77442S-77442S-8,

- 2010.
13. **M. T. Ibrahim**, R. Zhang, and L. Guan, "An Efficient Feedback Method for Iris Localization," in *Proceedings of Visual Communications and Image Processing (VCIP 2010)*, 2010. [Live Demo]
 14. R. Zhang, **M. T. Ibrahim**, and L. Guan, "A Collaborative Bayesian Image Retrieval Framework," in *Proceedings of Visual Communications and Image Processing (VCIP 2010)*, 2010.[Prototype System Demo]
 15. L. Guan, P. Muneesawang, Y. Wang, R. Zhang, T. Yun, A. Bulzacki, and **M. T. Ibrahim**, "Multimedia multimodal methodologies," in *IEEE International Conference on Multimedia and Expo (ICME 2009)*, pp. 1600-1603, 2009.
 16. **M. T. Ibrahim**, M. J. Kyan, M. A. Khan, K. K. Alimgeer, and Ling Guan, "On-Line Signature Verification: Directional Analysis of a Signature Using Weighted Relative Angle Partitions for Exploitation of Inter-Feature Dependencies," in *IEEE International Conference on Document Analysis and Recognition (ICDAR 2009)*, pp. 41-45, 2009.
 17. **M. T. Ibrahim**, M. J. Kyan, and Ling Guan, "On-line Signature Verification Using Global Features," in *Proceedings of 22nd Canadian Conference on Electrical and Computer Engineering (CCECE 2009)*, pp. 682-685, 2009.
 18. **M. T. Ibrahim** and Ling Guan, "On-line signature verification by using most discriminating points," in *IEEE International Conference on Pattern Recognition (ICPR 2008)*, pp: 1-4, 2008.
 19. **M. T. Ibrahim**, M. J. Kyan, and Ling Guan, "On-line Signature Verification Using Most Discriminating Features and Fisher Linear Discriminant Analysis (FLD)," in *IEEE International Symposium on Multimedia (ISM 2008)*, pp. 172-177, 2008.
 20. **M. T. Ibrahim**, T. Bashir, and L. Guan, "Robust Fingerprint Image Enhancement: An Improvement to Directional Analysis of Fingerprint Image Using Directional Gaussian Filtering and Non-subsampled Contourlet Transform," in *IEEE International Symposium on Multimedia (ISM 2008)*, pp. 280-285, 2008.

Chapter 1

Introduction

Vision is often considered as the most powerful human sense. The human visual system in conjunction with the brain produce an organized perception in terms of motion, size, shape, distance, relative position, and texture [1–3]. Recently, many applications are being developed to mimic the human vision in areas related to multimedia processing, autonomous navigation, motion analysis, surveillance, biometrics, optical character reading and medical imaging [4–6]. Generally, there are two principal approaches taken towards solving the problem of visual processing - reproducing the human visual system and simulating a similar system [7]. Since, it is still largely unknown how our brain processes the visual information; the simulation approach is the most common one. It has brought diversity to the field of digital image processing where researchers from different areas are trying to solve the problem in their own way, making image processing a multidisciplinary field of study.

Digital image processing generally refers to the use of computer algorithms that operate on or analyse digital images [8]. It enables the computer to extract useful information from the increasingly complex scenes, volumes, and time sequences [9]. Due to the rapid development of data acquisition devices, digital image processing has gained a lot of attention in fields such as biometrics and biomedical [6, 10]. One of the earliest applications of digital image processing was in the newspaper industry in the early 1920s [11]. Since then, the field of digital image processing has progressed immensely, especially with the advent of cheaper and smaller computers, digital cameras, television channels, and the internet.

Digital filtering is an integral part of many digital image/video processing applications.. It is not only used to improve the quality of an image but can also serve as a preprocessing step for subsequent processing, for instance, trend removal, noise suppression, motion analysis, intensity inhomogeneity correction, restoration, segmentation and feature extraction, to name a few [11–16]

In this dissertation, we focus on the research and development of advanced filtering methods for image and video processing applications. Specifically, our focus is on the filtering methods for trend removal/intensity inhomogeneity correction and feature extraction.

1.1 Trend Filtering

Trend filtering is an important task in signal analysis. It has a wide range of applications in areas like macroeconomics [17, 18], social sciences [19], revenue management [20], and biological and medical sciences [21–23]. The presence of the trend term in a signal often gives rise to false extrema which in return degrades the performance of data analysis methods [24]. For instance, in a heart signal, the localization and the counting of the heart beat can be affected by the presence of these false extrema [25]. Therefore, it is highly desirable to suppress the trend term before performing any kind of data analysis.

Generally, the trend of a signal is defined as a slowly varying component [16]. It is often approximated with the help of regression methods [16]. These methods transform the trend filtering problem into a curve fitting problem where the parameters of the curve (lower order polynomial) are approximated with the help of weighted least squares. However, estimation of the trend term with the help of data regression is a highly complex problem. Moreover, the performance of such methods is highly sensitive to the choice of polynomial, i.e., the order and the number of terms in the polynomial.

Trend filtering plays a vital role in signal analysis, as it may hide or give rise to false extrema [24]. The Fourier transform is a widely used tool for signal analysis. It has sinusoidal basis functions which are global in nature [11]. Due to this global nature, the Fourier transform is considered as a weak tool to perform non-stationary signal analysis. In case of

non-stationary signal analysis, STFT is often preferred over the Fourier transform due to its local nature [26]. In STFT, a non-stationary signal is assumed to be stationary within a small window. However, determination of the stationary portion and its size is a dilemma in itself.

Wavelets are often considered to be one of the best tools to perform the non-stationary signal analysis [11,12]. Wavelets perform the signal analysis depending on the selection of the mother wavelet and the number of scales. Ideally, these parameters should be derived from the underlying signal. But, there is no such methodology to provide any a priori knowledge about these parameters.

On the other hand, empirical mode decomposition (EMD) is a fully data driven method [24, 27]. It does not use any predetermined filter or fixed basis functions [24]. It provides a parameter free methodology to decouple the trend term from a 1D non-stationary signal.

In case of video analysis, image registration is often considered to be a prerequisite to a meaningful data/signal analysis [28]. Especially, in motion analysis, it is often preferred to perform the non-rigid inter-frame registration before extracting the motion information from the videos [28–31]. Moreover, the presence of the trend term further complicates the motion analysis process [25]. To cater this problem, this dissertation presents a robust method that transforms the high complexity non-rigid inter-frame registration problem into a 1D trend filtering problem. Furthermore, the proposed method extracts the trend term without defining any regression model.

1.2 Global Intensity Inhomogeneity Correction

In multi-dimensional signals, the trend term is often referred to as shading, non-uniform illumination, intensity non-uniformity, global intensity inhomogeneity, or bias field [32]. Image filtering has been widely used for the correction of global intensity inhomogeneity [32–34], where inhomogeneity can be described as smooth variations in intensities. Global intensity inhomogeneity is very smooth across the whole image, i.e., it varies slowly and is often characterized by low frequency components [34]. In the Fourier transform, it is often char-

acterised by low frequencies located in the proximity of the DC component or the zeroth frequency component [33].

During the last decade, a number of intensity inhomogeneity correction methods have been proposed in the literature [32,34]. Generally, these methods can be classified into surface fitting, filtering, and histogram thresholding [32,34]. In [35], a surface fitting method based on bivariate polynomial of degree N was used to approximate the intensity inhomogeneity in an image. Furthermore, Gaussian filter was used to suppress the additive noise followed by the logarithmic transformation. This transformation helps to convert the multiplicative image model to an additive one for subsequent processing. The parameters of the surface fitting model were calculated from the gradient of the smoothed image by using a weighted least squares method. An exponential weighting function was used to give small weights to the pixels with large gradient magnitudes and vice versa. The method is computationally expensive, as one needs to solve a system of $2P(N+1)(N+2)/2 - 1$ linear equations to find the $(N+1)(N+2)/2$ parameters, where P is the total number of pixels and N is the degree of the polynomial. The performance of the method is dependent on the degree of the polynomial used to estimate the intensity inhomogeneity. Furthermore, the method is sensitive to the choice of the parameter in the weighting function.

In [36], maximum a posteriori estimation with the sparseness prior of the image gradient was used to estimate the global intensity inhomogeneity in natural images. It is based on the assumption that adjacent pixels have the same intensities unless separated by edges [36]. The method works well as long as the image follows the sparseness property of the gradient probability distribution function. This method is also computationally expensive and heavily depends on the parameters of the weighting function. Here, it is worth mentioning that the gradient magnitude must in general be sampled at twice the original sampling rate to avoid aliasing [37], which was not followed in [36]. Doubling the sampling rate will at least double the computational cost of the method proposed in [36].

A well-known iterative method (N3) for global intensity inhomogeneity correction in magnetic resonance (MR) images was proposed in [38]. It corrects global intensity inhomogeneity

without requiring a priori knowledge of the tissue classes. It iteratively estimates the global intensity inhomogeneity by maximizing the high frequency content of the distribution of the tissue classes.

Homomorphic unsharp masking (HUM) is a frequency domain filtering method to decouple the global intensity inhomogeneity from an image [11, 39]. HUM scales down the magnitude of the low frequency components by multiplying them with a smaller value (γ_L) and amplifies the magnitude of the high frequency components by multiplying them with a larger value (γ_H). For better performance, HUM requires fine tuning of three parameters, γ_L , γ_H and c , where c controls the transition of the filter. One of the main challenges in using HUM is to determine the order and the cut-off of the filter. In a qualitative analysis [40], it was reported that HUM-based mean filtering outperforms HUM-based median filtering. It was also concluded that a window size of 65×65 or larger (to apply this method to other images, the window size should be scaled inversely with voxel size to maintain the window size in millimeters) is appropriate. If HUM is applied to an image composed by a grey foreground with a dark background, it produces halo artifacts on the boundaries [41]. In HUM, these artifacts become more prominent when a large window is used. The Guillemaud filter [42] was successfully used in [41, 43, 44] to avoid these artifacts. In [45], an automatic method was proposed to select the cut-off frequency for HUM. Later on, this method was generalized to 3D [41].

1.3 Local Intensity Inhomogeneity Correction

One of the issues closely related to global intensity inhomogeneity is shadow artifacts and illumination variations. These shadow artifacts appear as sharp discontinuities. Unlike the global intensity inhomogeneity, these sharp changes are only visible at shadow boundaries, making intensity inhomogeneity local in nature. The literature shows that the presence of local intensity inhomogeneity affects the performance of image matching [46, 47]. Especially, the suppression of local intensity inhomogeneity in facial images has been one of the major challenges in most of the current face recognition systems [48]. Presence of the local intensity

inhomogeneity can alter the face appearance and can produce strong shadow artifacts [49]. These artifacts can degrade the performance of a face recognition system. It is revealed in the face recognition vendor tests [50,51] that the performance of the face recognition system is highly sensitive to shadow artifacts and the variations produced in the face appearance by the illumination conditions can be much larger than the changes caused by the personal identity.

In literature, a large number of methods have been proposed for the correction of local intensity inhomogeneity in facial images [52]. In [53, 54], self quotient image (SQI) based methods were used to estimate the local intensity inhomogeneity in facial images. These methods model the reflectance as a ratio between the image and the estimated illumination. However, these methods produce halo effects at large illumination discontinuities. In order to reduce the halo effects, a multi scale retinex (MSR) model was proposed in [55] which attempts to estimate the local intensity inhomogeneity at different scales. This method is computationally expensive and requires the selection of parameters, i.e., number of scales, weights associated with each scale and the standard deviation of the smoothing function at each scale, which itself is an optimization problem.

A number of edge-preserving filtering methods have also been used to estimate the local intensity inhomogeneity in facial images, such as anisotropic diffusion, bilateral filtering and mean shift filtering [48]. A logarithmic total variation (LTV) method was proposed in [56], which exploits the edge-preserving and multi-scale decomposition capability of the $TV - L^1$ [57] model. This method also helps to extract the smaller intrinsic facial structures.

In [58], a wavelet based method has been proposed, where the histogram equalization was applied to the low-low sub-band image and the high frequency components were enhanced in the other 3 sub-band images. Discrete cosine transform (DCT) transform has also been used for the illumination normalization in the logarithmic domain [59] where some of the DCT coefficients corresponding to the low frequency contents were discarded to compensate for illumination variations.

1.4 Image filtering based feature extraction

In literature, a number of filtering methods have been proposed to extract the most compact and informative set of features from images/videos [15]. This compact representation is usually preferred to improve the efficiency of classification problems. The importance of feature extraction is constantly increasing with the fast growing diversity of signals/images. However, the extraction of stable and discriminatory features from a signal/image is a dilemma in itself.

Feature extraction is an essential pre-processing step to pattern recognition and machine learning problems. It has been widely used in extracting discriminatory information from the biometric samples which can be further used for person identification. Among various biometrics, fingerprint has gained popularity due to its high accuracy, stability, public acceptability, and low cost [5, 60]. Many different methods have been proposed in the literature for fingerprint recognition, which can be categorized as minutiae-based and image-based methods [5]. In minutiae-based methods, a set of un-ordered minutiae points are identified after some pre-processing steps and compared with the minutiae points of the template [5, 61, 62]. The dissimilarity measure is based on the set difference. A post-processing step is usually adopted to remove the false minutiae points [62]. In general, minutiae-based methods are computationally complex due to the requirement of time-consuming preprocessing, and the performance of such methods are highly dependent on the quality of the fingerprint images.

In contrast to minutiae-based methods, image-based methods extract the features directly from the fingerprint image. It requires less pre-processing and hence is computationally more efficient. Moreover, image-based methods generally work well in case of low quality images in comparison to minutiae-based methods. A number of image-based methods have been proposed in the past, majority of which employ texture analysis using filter banks [5]. In general, local texture analysis has shown to be more effective than the global counterparts [63]. In [64], a local texture analysis based approach was proposed, where a set of Gabor filters with four and eight different orientations were applied to the region of interest (ROI) of 96×120 pixels, cropped with the core point as its center. The detection of the core point

was performed manually and the ROI was divided into non-overlapping blocks of size 8×8 . The magnitudes of each non-overlapping block corresponding to different orientations of the Gabor filters were used as the features. In [65, 66], the core point was detected manually and the ROI of 64×64 pixels was cropped centered at the core point. The ROI was further divided into four non-overlapping blocks of size 32×32 , and a set of different wavelets up to the fourth level were applied to each block. On each level, the standard deviation of the wavelet coefficients was calculated and a feature vector of length 48 was extracted for the representation of each fingerprint. Finally, Euclidean distance was used to measure the similarity between the feature vectors of a template and a query image. A similar approach was presented in [67]. The ROI of size 64×64 , centered at the manually detected core point, was first identified. The ROI was divided into four non-overlapping blocks of size 32×32 , and DCT was applied on each sub-image independently. The standard deviations of six pre-defined blocks in each sub-image were used to generate a feature vector. In all of the above mentioned methods, core point was detected manually and the features were extracted from the ROI of the original input image without any fingerprint enhancement.

In [68], a filter-bank based approach was proposed, where a ROI centered at the core point was divided into B concentric bands and each sub-band was further divided into k sectors. A set of Gabor filters with eight orientations were applied to the image, and the absolute average deviation (AAD) of each sector was used as a feature. A cyclic shift method was also proposed to make the images invariant to the rotation of 11.25° . Euclidean distance was used as the similarity metric. However, this method is computationally complex and requires the storage of a large number of cyclicly rotated images as templates. Furthermore, it is only invariant to the rotation of 11.25° , therefore provides limited rotation invariance.

In [69], Fourier Mellin transform has been used in conjunction with wavelet. The ROI of 128×128 pixels centered at the core point was cropped and a two-level wavelet was applied. At each level of wavelet, Fourier Mellin transform was used to make the ROI translation, rotation and scale invariant. Finally, Euclidean distance was used to find the matching score between the two fingerprints.

Recently, the Hu’s invariant moments [70] based methods have been utilized in [71, 72]. The ROI of size 96×96 centered at the core point was divided into 16×16 non-overlapping blocks. The invariant moments were calculated for each non-overlapping block and a feature vector of length 252 was used to represent each fingerprint.

1.5 Contributions

The main contributions of this dissertation are enumerated as follows:

- One of the main contributions of this dissertation is the development of novel trend filtering method that transforms the inter-frame registration problem into a low complexity trend filtering problem. The proposed method provides a mechanism to transform the motion information from a video sequence to a 1D signal. The proposed method has been used to analyse the embryonic heart rate *in vitro* from the video recordings of rat. In the proposed method, LEM in conjunction with correlation coefficient is used to extract the motion information from the video sequence whereas the modified EMD is used for the signal analysis. Moreover, the modified EMD provides a data dependent decomposition which helps in the suppression of noise and trend term, i.e., motion artifacts. Furthermore, the proposed method finds the trend term without defining any regression model which makes it attractive for the biologists to analyse the heart motion without any need of parameter tweaking and expensive imaging equipment.
- Another major contribution is the accurate estimation of global intensity inhomogeneity in two dimensional (2D) and volume images. In the proposed method, global intensity inhomogeneity has been estimated by using grey-weighted distance transform (GWDT) in conjunction with Fourier transform. The proposed method is based on the observation that if global intensity inhomogeneity is visible in an image then the magnitude of affected frequency components should have a higher magnitude than their neighbouring components. In the proposed method, these affected frequency components are localised in the frequency domain by GWDT. The proposed method

generates a regular/irregular shaped high-pass filter based on the frequency content of an image affected by the global intensity inhomogeneity. The proposed method iteratively suppresses the magnitude of the affected frequency components without any prior knowledge about the imaging modality. Moreover, all of the filter parameters are automatically determined from the affected low frequency components. The effectiveness of the proposed methods has been evaluated on the images acquired from several imaging modalities, which makes it different and unique from most of the existing methods in this field of study. Experimental results demonstrate that the proposed method performs better, in some cases, strikingly better than some of the existing and state-of-the-art methods.

The same idea has been extended to 3D for the volume images, especially for brain MR volumes. The proposed method constructs a 3D high-pass filter with the help of 3D GWDT in the frequency domain. The comparison of the experimental results with some of the recent and state-of-the-art methods demonstrates the versatility and robustness of the proposed 3D method.

- Another main contribution of this dissertation is the development of filtering method for local intensity inhomogeneity correction and shadow artifacts suppression. The proposed method is based on the gamma transformation and the decimation-free directional filter bank (DDFB). In the proposed method, gamma transformation is used to normalize the input image whereas DDFB is used to suppress the shadow artifacts and the local intensity inhomogeneity irrespective of scale and orientation. The proposed method has been deployed to suppress these artifacts in facial images. Moreover, it has been empirically shown that most of the discriminatory power lies within the horizontal facial features in presence of shadows and local intensity inhomogeneity.
- A novel filtering method has been proposed to generate rotation invariant feature codes from the 2D images. However, in this dissertation the proposed method has been used to extract the rotation invariant feature codes from the fingerprint images. The

proposed method helps to improve the recognition accuracy of a fingerprint recognition system. The proposed method generates the rotation invariant feature codes based on the output of the complex filters and the Gabor filter bank. The effectiveness of the proposed approach has been demonstrated on the FVC2002 set_a public database, in both identification and verification modes.

1.6 Organization

The organization of this dissertation is as follows:

Chapter 2 provides a description and the background of the methods used in this dissertation.

Chapter 3 introduces the proposed method for the signal generation and trend removal from the video recordings. The effectiveness of the proposed method is demonstrated through extensive experimentation on the video recordings of rat embryos.

Chapter 4 presents the proposed filtering method for global intensity inhomogeneity correction in 2D and volume images. The proposed method has been tested on several imaging modalities. Theoretical foundations of the proposed method along with experimental results are also reported.

Chapter 5 presents the proposed filtering method for local intensity inhomogeneity correction in facial images. The proposed method has been evaluated on two public databases.

Chapter 6 presents the proposed filtering method for extracting the rotation invariant feature codes from the fingerprint images. The effectiveness of the proposed method in both identification and verification scenarios is also reported.

Chapter 7 summarizes the work presented in this dissertation and outlines the directions for future research.

Chapter 2

Materials and Methods

In an effort to make the dissertation self contained, this chapter provides a brief description of the material and methods employed in this dissertation.

2.1 Digital Signal/Image Filtering

Digital filtering is often considered as a key step in digital signal and image processing. It is mainly used for amplification or attenuation of some frequencies depending on the nature of application. In general, filtering can be performed in either time/spatial or frequency domain [12]. In spatial domain, filtering is generally performed with the help of spatial masks (*also known as spatial filters, kernels, or templates*) that are convolved over the entire image. Mathematically, it can be written as:

$$g(x, y) = h(x, y) * f(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(i, j) f(x - i, y - j) \quad (2.1)$$

where $*$ represents the convolution between the input image f and the spatial mask h .

The values or the co-efficients of the mask determine the nature and the properties of filtering technique [9]. Generally, the spatial mask whose co-efficients sums to one is known as a low-pass filter and it is used for smoothing purpose [11]. On the other hand, the co-efficients of a high-pass filter sums to zero and it is used to detect sharp changes and edges. In spatial filtering, the output value at the location (x, y) not only depends on the input

value at (x, y) , but also on the specified number of neighbouring pixel values around (x, y) determined by the size of mask.

In the frequency domain, filtering can be performed by the point-wise multiplication:

$$G(u, v) = F(u, v)H(u, v) \quad (2.2)$$

where $F(u, v)$ represents the Fourier transform of $f(x, y)$ and $H(u, v)$ represents the Fourier transform of $h(x, y)$. Once, the filtering has been performed in the frequency domain, inverse Fourier transform is applied to $G(u, v)$ to bring it back to the spatial domain. For better understanding of the frequency domain, a brief description of the 2D sinusoids and the Fourier transform is presented in the sections to follow.

2.1.1 Two Dimensional Sinusoids

A 2D sinusoid is characterized by phase shift, frequency and the direction of oscillation [12]. Mathematically, a 2D sinusoid can be written as:

$$S(x, y) = A \sin(2\pi(Ux + Vy) + \phi), \quad (2.3)$$

where A represents the amplitude, ϕ represents the phase shift, U and V (with units of cycles/pixel) represents the frequency of oscillation along the horizontal and the vertical spatial image dimensions.

In case of 2D sinusoids defined on a finite grid of size $M \times N$, it is often preferred to use scaled frequencies (u, v) that have the visually intuitive units of cycles/image [11, 12]. So, a 2D sinusoid defined on a finite grid of size $M \times N$ can be written as:

$$S(x, y) = A \sin(2\pi(\frac{ux}{M} + \frac{vy}{N}) + \phi), \quad (2.4)$$

Fig. 2.1 represents a 2D sinusoid with its cross sectional view.

Generally, a 2D sinusoid oscillates along every direction except for the direction orthogonal to the direction of fastest oscillation [12]. The direction of the fastest oscillation is given

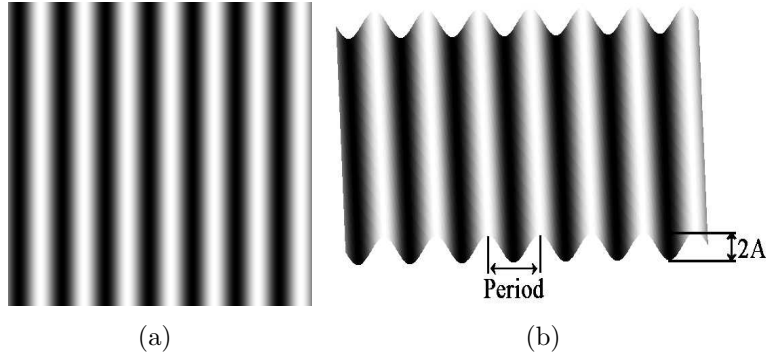


Figure 2.1: a) 2D sinusoid. b) Cross sectional view of 2.1(a).

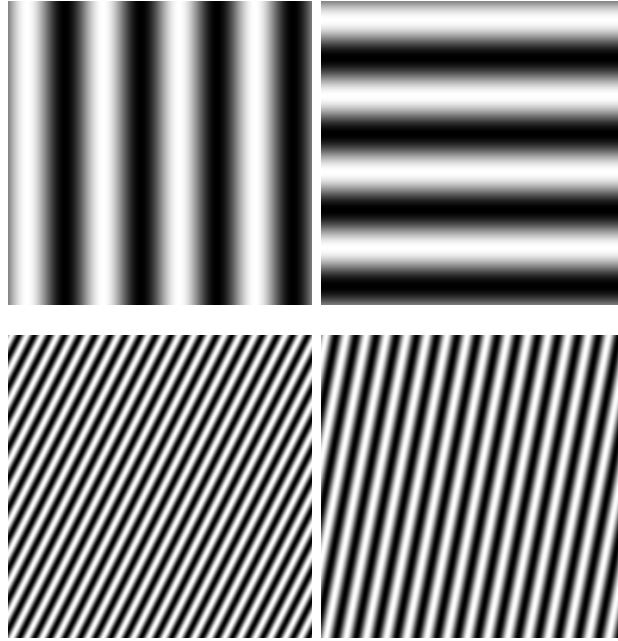


Figure 2.2: Examples of 2D sinusoids of varying frequencies of size 255×255 with $A = 10$ and $\phi = 0$.

by:

$$\theta = \tan^{-1}\left(\frac{v}{u}\right), \quad (2.5)$$

whereas its frequency can be expressed as:

$$\omega = \sqrt{u^2 + v^2}. \quad (2.6)$$

Fig. 2.2 shows multiple sinusoids of varying frequencies and orientations.

2.1.2 Two Dimensional Fourier Transform

Fourier transform is a signal/image analysis tool that decomposes a signal/image into its sine and cosine components [12]. It transforms a spatial domain image into its frequency domain representation. In the frequency domain, each point represents a particular frequency contained in the spatial domain image [11]. As, we are dealing with the digital images only, so we will restrict our discussion to the Discrete Fourier Transform (DFT) which is a sampled version of the Fourier transform.

According to the DFT, an image can be resolved into sinusoids of varying frequencies [73, 74]. Mathematically, it can be written as:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi j(\frac{ux}{M} + \frac{vy}{N})}, \quad (2.7)$$

where $0 \leq u \leq M - 1$, $0 \leq v \leq N - 1$, and $F(u, v)$ represents a complex-valued function of the sinusoid passing through the whole image $f(x, y)$ of size $M \times N$. Here, frequency (u, v) is defined in terms of cycles/image. The basis functions of the DFT are complex exponentials with increasing frequencies that exist through out the signal/image at each time/spatial location [11]. Due to this global nature, Fourier transform provides the information only about the frequency components that are present in a signal/image irrespective of where these components appear in the signal/image.

Generally, a 2D DFT $F(u, v)$, is characterised by the frequency (ω), magnitude ($|F(u, v)|$), phase ($\angle F(u, v)$) and direction of the fastest oscillation (θ). Mathematically, these four characteristics can be expressed as:

$$\begin{aligned} \omega &= \sqrt{u^2 + v^2}, \\ \theta &= \tan^{-1}\left(\frac{v}{u}\right), \\ |F(u, v)| &= \sqrt{\Re(F(u, v))^2 + \Im(F(u, v))^2}, \\ \phi &= \tan^{-1}\left(\frac{\Im(F(u, v))}{\Re(F(u, v))}\right). \end{aligned} \quad (2.8)$$

In the polar form, $F(u, v)$ can be written as:

$$F(u, v) = \Re(F(u, v)) + j\Im(F(u, v)) = |F(u, v)|e^{j\phi} \quad (2.9)$$

where $|F(u, v)|$ is the magnitude and ϕ represents the phase. The magnitude of the Fourier transform decreases rapidly as one moves away from origin [11]. Fig. 2.3(b) represents the magnitude of the center shifted Fourier transform (CSFT).

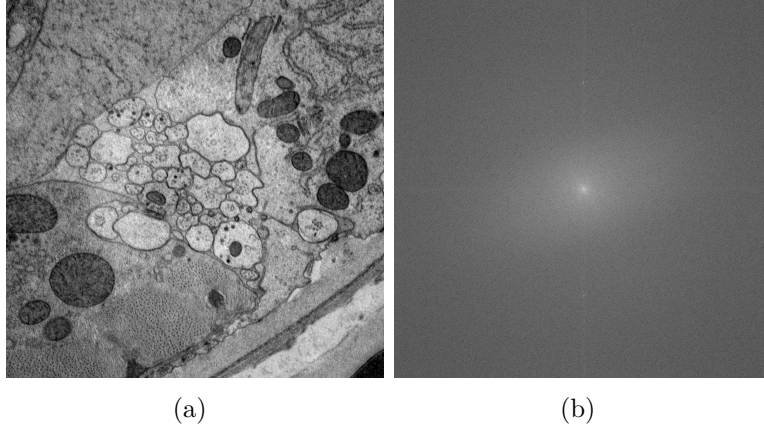


Figure 2.3: a) Original TEM image of a c-elegans. b) Magnitude of the CSFT of 2.3(a).

The 2D DFT can be implemented efficiently by computing the 1D DFT row-wise followed by column-wise computation of 1D DFT or vice versa. This separability property of the DFT helps to reduce the number of computations. Even with these computational savings, an N -point 1D DFT has the computational complexity of $O(N^2)$. However, this computational complexity can be further reduced to $O(N \log_2 N)$ by using the Fast Fourier Transform (FFT) to compute the 1D DFT [11].

2.2 Directional Analysis

Directional analysis is a process of decomposing an image into a set of several directional components, each one containing a set of frequency subbands. These subbands can be represented in the frequency domain by partitioning the Fourier spectrum of an image [75].

Directional analysis has been widely used for texture analysis, feature extraction, image enhancement, segmentation, edge detection, and target detection [64, 68, 76, 77].

2.2.1 Gabor Filter

The most commonly used directional analysis tool is Gabor filter which is a frequency and orientation selective filter [78]. It is a modified version of Gaussian filter which can be constructed by modulating a Gaussian in some particular frequency and orientation [79]. Mathematically, the 2D Gabor filter can be written as:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \cos(2\pi u_o x) \quad (2.10)$$

where u_o represents the center frequency of a cosine wave along the x-axis and (σ_x, σ_y) defines the standard deviation of the 2D Gaussian along the x - and y -axes, respectively [79, 80].

The 2D Gabor filter $g(x, y)$ has the following Fourier transform:

$$G(u, v) = A \left[\exp \left\{ -\frac{1}{2} \left(\frac{(u - u_o)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right) \right\} + \exp \left\{ -\frac{1}{2} \left(\frac{(u + u_o)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right) \right\} \right] \quad (2.11)$$

where $A = 2\pi\sigma_x\sigma_y$, $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$ [81].

The 2D Gabor filter of arbitrary orientation can be obtained by rotating $x - y$ coordinate system in Eq. 2.10. Fig. 2.4 shows the power spectrum of the 2D Gabor filter, which is a sum of a pair of Gaussians oriented in a particular direction with some radial center frequency u_o .

Gabor filters have been widely used for local texture analysis [12]. However, in most of the natural images, it is hard to find a single filter resolution at which one can localize a spatial structure in an image [82].

2.2.2 Directional Filter Bank (DFB)

Another directional analysis tool is Directional Filter Bank (DFB) which is capable of extracting the spatial directional information that is global in nature [83]. DFB was proposed

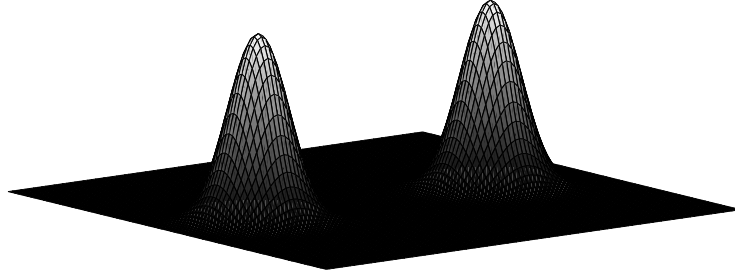


Figure 2.4: Power spectrum of the 2D Gabor filter.

by Bamberger [84] and later refined by Sang [85]. It decomposes the spectrum of an image into wedge-shaped like passbands as shown in Fig. 2.5. These passbands corresponds to global features in a specific direction in spatial domain [84]. The outputs of DFB are known

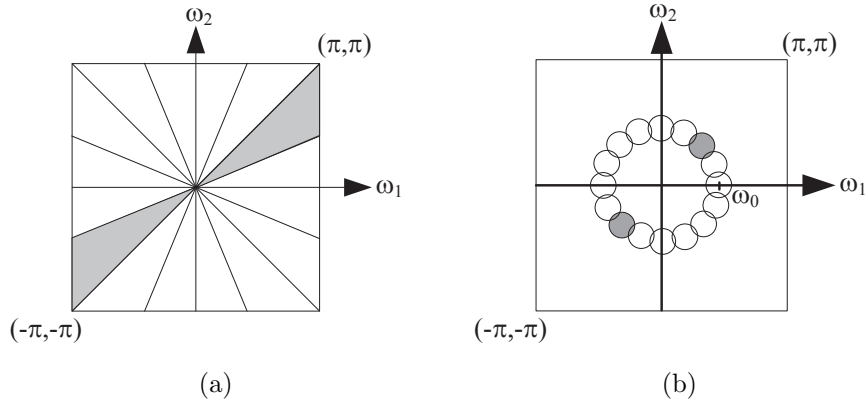


Figure 2.5: Passbands are represented by the greyish area. a) Frequency partition map of DFB [85]. b) Frequency partition map of Gabor filter bank.

as directional subbands which are maximally decimated and are orthogonal to each other. Each subband is of smaller size as compared to the original image because of the presence of downsamplers. The original structure of DFB suffers from frequency scrambling [84], i.e., after the second stage of DFB, subbands become visually distorted. The solution to this frequency scrambling was proposed in [85], where all the sampling matrices were moved to

the end of analysis section of DFB. Fig. 2.6 shows the structure of latest version of 3-stage DFB proposed in [85], which eliminates frequency scrambling.

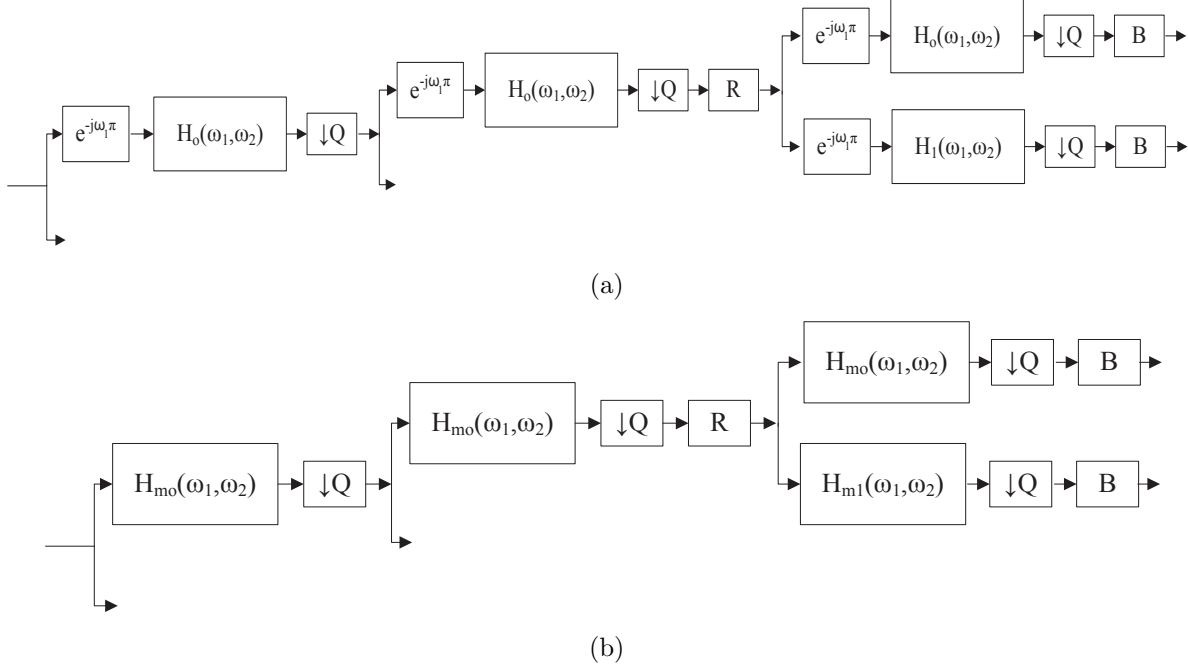


Figure 2.6: 3-stage DFB structure proposed in [85]. a) DFB modulated structure. b) DFB non-modulated structure.

The basic structure of the DFB shown in Fig. 2.6(a) consists of diamond shaped low-pass filter $H_o(\omega_1, \omega_2)$, diamond shaped high-pass filter $H_1(\omega_1, \omega_2)$, Quincunx down sampler Q , diamond conversion matrices R , modulators $e^{-j\omega_1\pi}$, and postsampling matrix B . The modulator varies the spectrum of the image so that the modulated image can be divided into two directional subbands using $H_o(\omega_1, \omega_2)$ and $H_1(\omega_1, \omega_2)$. The Quincunx downsampling matrix used in DFB structure is given below:

$$Q = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

The matrix Q not only downsamples the image diagonally but also rotates the image by 45 degrees around its origin [85]. Figure 2.7 shows the effect of Quincunx downsampling on the cameraman image. The four diamond conversion matrices R applied in the DFB structure are given below:

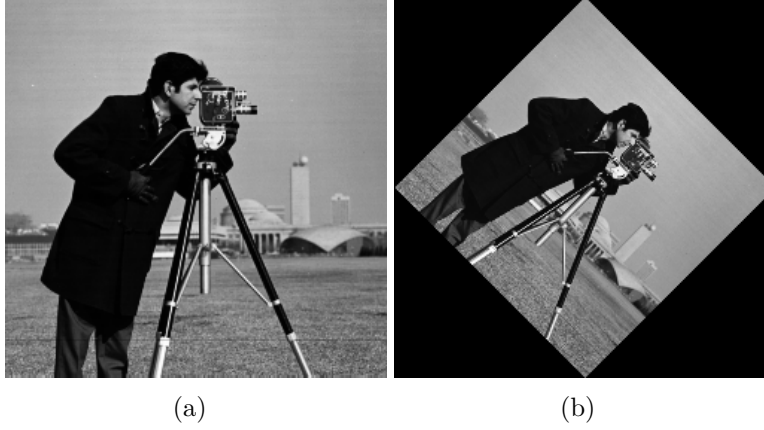


Figure 2.7: Example of Quincunx downsampled image. a) Cameraman. b) Downsampled by Q .

$$R_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$R_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad R_4 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

These matrices help to map the parallelogram shaped passband to a diamond shaped passband [85]. The matrix B in the DFB structure is the post sampling matrix which is used to correct the frequency scrambling associated with each subband [75, 85].

In Fig. 2.6(b), the modulators $e^{-j\omega_1\pi}$ has been moved inside the diamond shaped filters to give rise to hour-glass shaped filters $H_{m0}(\omega_1, \omega_2)$ and $H_{m1}(\omega_1, \omega_2)$, as shown in Fig. 2.8(b). However, the outputs of the two structures shown in Fig. 2.6 are equivalent to each other [85]. Fig. 2.9 illustrates the filtering steps associated with a 3-stage DFB decomposition. The two subbands corresponding to the first stage of DFB are shown in Fig. 2.9(a). In the second stage, the outputs of the first stage are further divided into four subbands using a procedure shown in Fig. 2.9(b). In the third stage, eight directional subband outputs are generated. Fig. 2.9(c) shows two of the eight directional subbands created by the third stage of DFB.

In [75, 86], DFB has been used for the image enhancement, where directional energies of subbands were calculated for the enhancement purpose. Interpolation was used before calculating directional energies to establish a one-to-one correspondence between all the

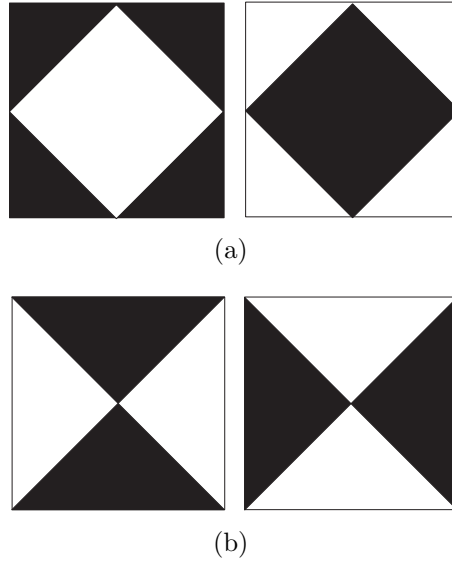
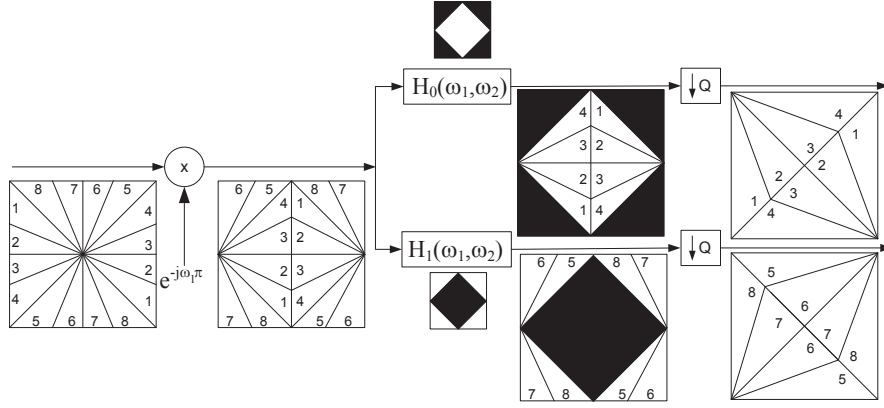


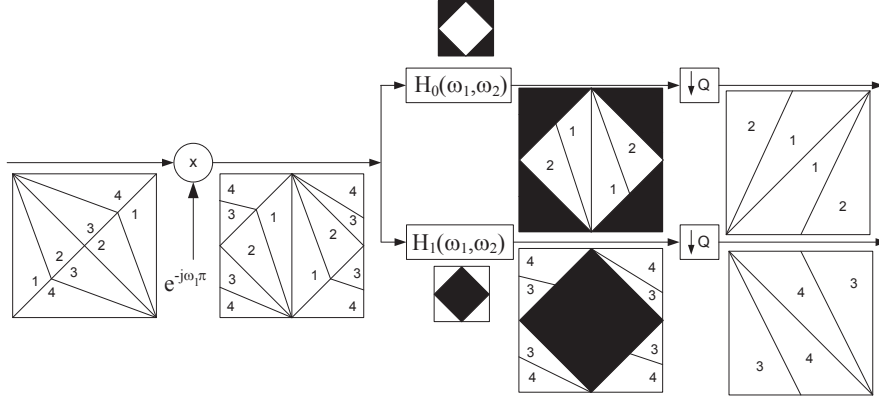
Figure 2.8: Types of ideal passbands for DFB structure proposed in [85]. a) For modulated structure shown in Fig. 2.6(a). b) For non-modulated structure shown in Fig. 2.6(b).

directional subbands of DFB. To overcome the extra step of interpolation, Decimation-Free Directional Filter Bank (DDFB) was proposed in [87], where the decimators and resamplers were shifted to the right of the filters by using the multi-rate noble identities [88]. The outputs of the DDFB are known as directional images and they are of same size as the input image. In DDFB, there is a one-to-one correspondence between pixels of all directional images, i.e., two pixels located at a spatial position (x, y) in two different directional images corresponds to the same position (x, y) in the original image. Fig. 2.10 shows the structure of 3-stage DDFB proposed in [87]. Here, Q and R corresponds to the same Quincunx downsampling matrix and the diamond conversion matrices used in DFB [87].

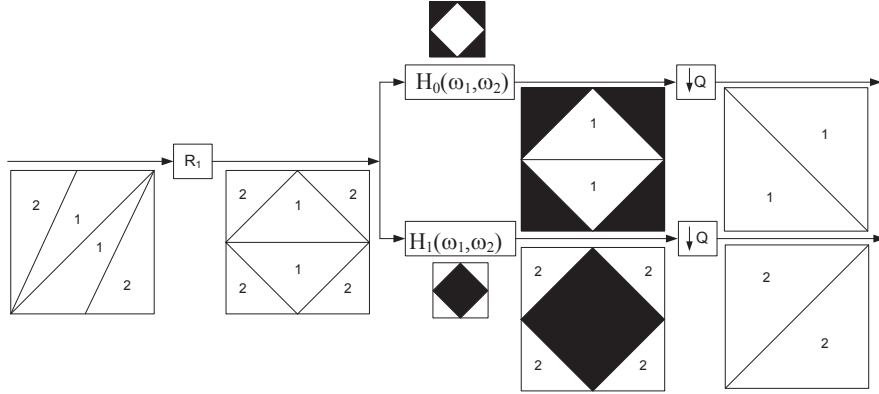
Fig. 2.11 shows a 3-stage DDFB decomposition where the outputs corresponding to the first stage of DDFB are shown in Fig 2.11(a). Fig. 2.11(b) and Fig. 2.11(c) shows some of the filters corresponding to the second and the third stage of DDFB .



(a)



(b)



(c)

Figure 2.9: Sang DFB structure proposed in [85]. a) First stage. b) Second stage. c) Third stage.

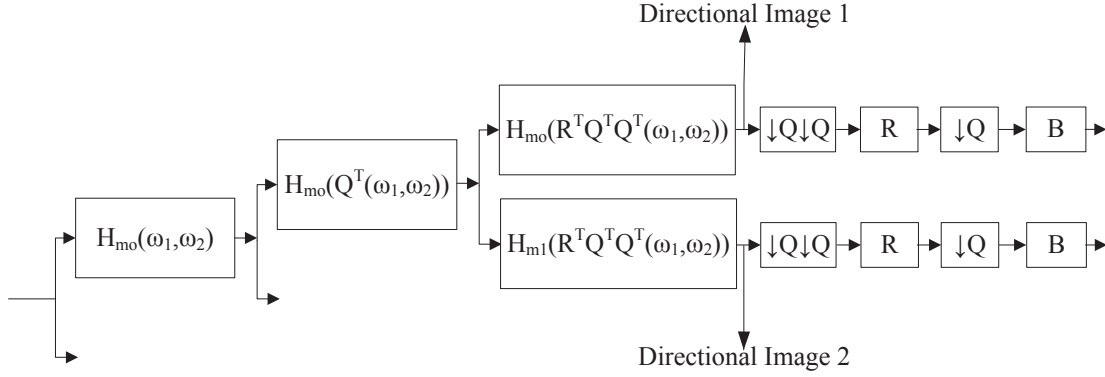


Figure 2.10: 3-stage DDFB structure proposed in [87].

2.3 Empirical Mode Decomposition (EMD)

EMD is an adaptive multi-scale representation that decomposes a non-stationary signal into a sum of oscillatory functions known as *intrinsic mode functions* (IMFs) [24,89], that: 1) have the same number of extrema and zero-crossings or differ by at most one; 2) are symmetric with respect to local zero mean.

To be successfully decomposed into IMFs, a signal must have at least two extrema, i.e., one maximum and one minimum. Each IMF is computed through an iterative process known as *sifting process* [89], which is summarized in Algorithm 1.

At the end of this process, the original signal $z(t)$ can be expressed as:

$$z(t) = \sum_{i=0}^{n-1} h_i(t) + r(t), \quad (2.12)$$

where n is the total number of IMFs, $r(t)$ represents the final residue also known as the trend term of the signal $z(t)$, and $h_i(t)$ is an IMF.

It is important to mention that sifting is an iterative process that normally takes more than one iteration to generate a single IMF. Generally, the number of iterations are controlled by a stopping criterion (SD). In literature, several different types of stopping criteria have been proposed [27]. For instance, in conventional EMD [24], the sifting process continues until the normalized squared difference between two successive $p(t)$ is smaller than a pre-

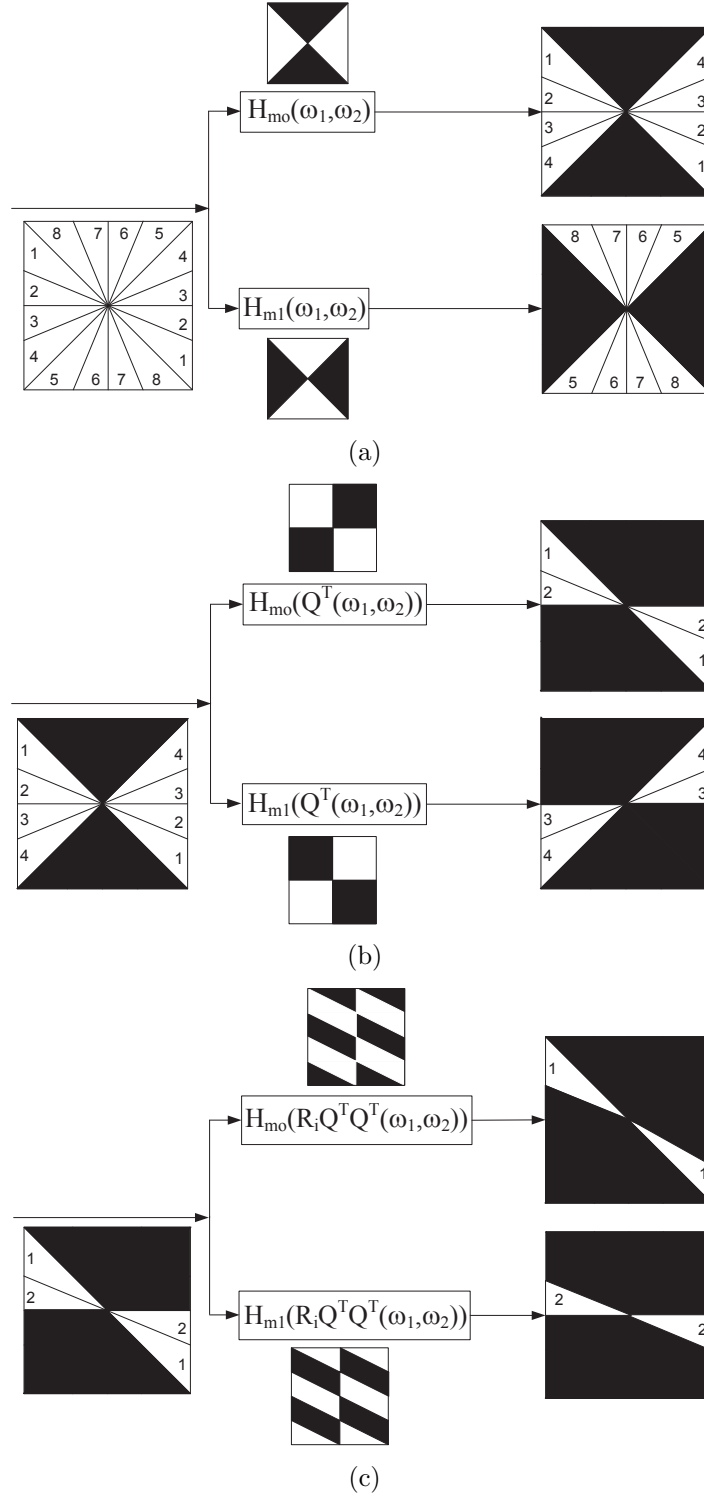


Figure 2.11: DDFB structure proposed in [87]. a) First stage. b) Second stage. c) Third stage.

Algorithm 1 Empirical Mode Decomposition

```
1:  $i \leftarrow -1$ .
2:  $j \leftarrow 0$ .
3: repeat
4:   Find all the extrema of a given signal  $z(t)$ .
5:   Create an upper and lower envelope,  $z_{up}(t)$  and  $z_{low}(t)$ , by fitting a cubic spline through
      maxima and minima, respectively.
6:   Find the mean envelope  $m(t)$  by:

$$m(t) = \frac{z_{up}(t) + z_{low}(t)}{2}$$

7:   if  $z(t) - m(t)$  is an IMF then
8:      $i \leftarrow i + 1$ 
9:     Find the IMF,  $h_i(t)$ :

$$h_i(t) \leftarrow z(t) - m(t)$$

10:    Find the residual,  $r(t)$ :

$$r(t) \leftarrow z(t) - h_i(t)$$


$$z(t) \leftarrow r(t)$$

11:     $j \leftarrow 0$ .
12:  else
13:     $p_j(t) \leftarrow z(t)$ ,
14:     $z(t) \leftarrow z(t) - m(t)$ ,
15:     $j \leftarrow j + 1$ .
16:  end if
17: until  $z(t)$  is monotonic or one extremum is left
```

determined value:

$$\frac{\sum_{t=0}^T |p_{j-1}(t) - p_j(t)|^2}{\sum_{t=0}^T p_{j-1}^2(t)} \leq SD. \quad (2.13)$$

Typically, the value of SD is chosen between 0.2 and 0.3 [24]. In [90], a second type of criterion, termed as the S stoppage criterion was proposed. According to S stoppage criterion, the sifting process stops only if: 1) the number of zero crossings and extrema are equal or differ by at most one and 2) remain same for S successive $p(t)$. The optimal value of S is found to be between 3 and 8 [27].

In contrast to the traditional signal analysis methods like Fourier transform, STFT and

wavelets, EMD is a fully data driven method [27]. It does not use any predetermined filter or fixed basis functions [24]. Due to these reasons, it has been widely used in the areas like oceanography, feature extraction, noise removal, trend removal, biometrics, biomedical signal and image analysis [91–93].

2.4 Clustering

Clustering is a supervised/unsupervised classification of patterns (objects, observations, data items, or feature vectors) into groups or clusters [94]. The main goal of clustering is to minimize the within cluster dissimilarity and to maximize the between cluster dissimilarity [95]. It can also be considered as finding groups of patterns such that the patterns within a group are similar to each other and different from the patterns in other groups [96].

2.4.1 K-Medoids Clustering

K-medoids is a classical partitioning method where each cluster is represented by one of the data points in the cluster as its center, i.e., medoid [97]. A medoid can be defined as the data point of a cluster, whose average dissimilarity to all the data points in the cluster is minimal [98].

Given an unlabelled data set $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ of n data points where $x_i \in \mathbb{R}^d$, K-medoids algorithm partitions \mathbf{X} into K clusters by minimizing the total squared error between \mathbf{X} and all the medoids \mathbf{V} , where $\mathbf{V} = \{v_1, v_2, \dots, v_K\} \subset \mathbf{X}$ with cardinality K [99]. Mathematically, it can be written as:

$$\min_{\mathbf{V}} J(\mathbf{V}) = \sum_{j=1}^K \sum_{i=1}^n \|x_i - v_j\|^2, \quad (2.14)$$

K-medoids algorithm works as follows:

1. Fix the number of clusters K , $K \in [2, n)$
2. Randomly choose the initial configuration of medoids \mathbf{V} .

3. Associate each data point to the closest medoid using a distance measure and calculate the cost.
4. For each medoid v_j
 - (a) Swap the non-medoid data point with v_j .
 - (b) Compute the cost of the configuration, i.e., total squared error.
 - (c) Select the configuration with the lowest cost.
5. Repeat steps 3 to 4 until there is no change in the medoids.

2.4.2 Fuzzy C-Means Clustering

Fuzzy C-Means (FCM) is a clustering method [100], which allows a pattern x_k to belong to more than one cluster with a certain degree of membership $\mu_{ik} \in [0, 1]$. The degree of membership μ_{ik} represents the fuzzy membership of k -th pattern to the i -th cluster.

Given an unlabelled data set $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$, where n is the total number of data points in \mathbf{X} and $x_i \in \mathbb{R}^d$. FCM partitions \mathbf{X} into c clusters where $c \in [2, n]$. It iteratively updates the membership μ_{ik} and the cluster centers v_i by minimizing the following objective function:

$$\min_{\mathbf{U}, \mathbf{V}} J(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^m \|x_k - v_i\|^2, \quad 1 \leq m < \infty \quad (2.15)$$

where $\mathbf{V} = \{v_1, v_2, \dots, v_c\} \subset \mathbb{R}^d$, m represents the weighting exponent and $\mathbf{U} = [\mu_{ik}]_{c \times n}$ is a fuzzy c-partition matrix of \mathbf{X} under the conditions:

$$\mu_{ik} \in [0, 1] \mid \sum_{i=1}^c \mu_{ik} = 1 \forall k \wedge 0 < \sum_{k=1}^n \mu_{ik} < n \forall i.$$

The standard FCM algorithm can be summarized as follows:

1. Select $c \in [2, n]$, $m \in [1, \infty)$, $\epsilon > 0$, $h = 0$ and initialize the partition matrix $\mathbf{U} = \mathbf{U}^{(0)}$.

- At iteration h , compute the cluster centers $v_i^{(h)}$ using:

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^m x_k}{\sum_{k=1}^n \mu_{ik}^m}, i = 1, 2, \dots, c, \quad (2.16)$$

where $\mu_{ik} \in \mathbf{U}^{(h)}$.

- Compute the coefficients in $\mathbf{U}^{(h+1)}$ using:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|x_k - c_i\|}{\|x_k - c_j\|} \right)^{\frac{2}{m-1}}}, k = 1, 2, \dots, n, \quad (2.17)$$

where $\|\cdot\|$ denotes the Euclidean norm. Typically, the value of m is set to 2, which is equivalent to normalizing the coefficient linearly to make their sum 1.

- If $\|\mathbf{U}^{(h+1)} - \mathbf{U}^{(h)}\| < \epsilon$, then stop; otherwise return to step 2 and increase h by 1, i.e., $h = h + 1$.

2.5 Dimensionality Reduction

Dimensionality reduction is the mapping of the higher dimensional data to a lower dimensional space such that a subspace in which the data lives is detected [101]. For instance, let $\mathcal{Z} = \{\mathcal{Z}_i\}_{i=1}^C$ denote a set of training images (or feature vectors), containing C subjects with each subject $\mathcal{Z}_i = \{\mathbf{z}_{ij}\}_{j=1}^{C_i}$, consisting of C_i images \mathbf{z}_{ij} . Each image is represented by a column vector of length J , i.e., $\mathbf{z}_{ij} \in \mathbb{R}^J$. In general, the objective of dimensionality reduction is, based on optimization of some criterion, to find a map $\varphi(\cdot)$ to produce a lower dimensional feature representation $\mathbf{y}_{ij} = \varphi(\mathbf{z}_{ij})$, $\mathbf{y}_{ij} \in \mathbb{R}^M$, $M \ll J$.

Here, we will address the two most widely used subspace techniques for linear dimensionality reduction, i.e., Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

PCA [102] is one of the the most influential subspace techniques for linear dimensionality

reduction. It is an unsupervised learning technique which provides an optimal, in the mean-squared-error sense, representation of the input data in a lower dimensional space. Let \mathbf{S}_t be the sample covariance matrix defined as follows,

$$\mathbf{S}_t = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} (\mathbf{z}_{ij} - \bar{\mathbf{z}})(\mathbf{z}_{ij} - \bar{\mathbf{z}})^T \quad (2.18)$$

where $\bar{\mathbf{z}} = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} \mathbf{z}_{ij}$ is the mean of all gallery samples and N represents the total number of images available in the training set. The PCA feature space is thus spanned by the first M ($M < N$) eigenvectors of \mathbf{S}_t corresponding to the largest eigenvalues. The feature representation of a feature vector \mathbf{z} in the PCA space can be obtained by a linear mapping : $\mathbf{y} = \mathbf{A}_{PCA}^T (\mathbf{z} - \bar{\mathbf{z}})$, where $\mathbf{A}_{PCA} = [\mathbf{v}_1, \dots, \mathbf{v}_M]$ is the transformation matrix consisting of the first M eigenvectors.

LDA is another commonly used method for linear dimensionality reduction [103]. In contrast to PCA, LDA is a supervised learning technique which produces a class-specific feature space based on the maximization of the so-called Fisher's criterion defined as the ratio of between-class scatter to within-class scatter, i.e.,

$$\mathbf{A}_{LDA} = \arg \max_{\mathbf{A}} \frac{|\mathbf{A}^T \mathbf{S}_b \mathbf{A}|}{|\mathbf{A}^T \mathbf{S}_w \mathbf{A}|}, \quad \mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_M], \quad \mathbf{a}_k \in \mathbb{R}^J, \quad (2.19)$$

$$s.t. \quad \mathbf{a}_i^T \mathbf{a}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

where \mathbf{S}_b and \mathbf{S}_w denote the between- and within-class scatter matrices defined as follows,

$$\mathbf{S}_b = \frac{1}{N} \sum_{i=1}^C C_i (\bar{\mathbf{z}}_i - \bar{\mathbf{z}})(\bar{\mathbf{z}}_i - \bar{\mathbf{z}})^T \quad (2.20)$$

$$\mathbf{S}_w = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} (\mathbf{z}_{ij} - \bar{\mathbf{z}}_i)(\mathbf{z}_{ij} - \bar{\mathbf{z}}_i)^T \quad (2.21)$$

where $\bar{\mathbf{z}}_i = \frac{1}{C_i} \sum_{j=1}^{C_i} \mathbf{z}_{ij}$ is the mean of the i -th class and $\bar{\mathbf{z}} = \frac{1}{N} \sum_{i=1}^C \sum_{j=1}^{C_i} \mathbf{z}_{ij}$ is the mean of

all classes. Thus, LDA-based feature representation of an input vector \mathbf{z} can be obtained by a linear projection: $\mathbf{y} = \mathbf{A}_{LDA}^T \mathbf{z}$. The maximization of the Fisher's criterion given in Eq. 2.19 is equivalent to solving an eigenvalue problem of $\mathbf{S}_w^{-1} \mathbf{S}_b$. If \mathbf{S}_w is non-singular, the transformation matrix \mathbf{A}_{LDA} is comprised of the first M most significant generalized eigenvectors of \mathbf{S}_w and \mathbf{S}_b corresponding to M largest eigenvalues λ_k , i.e., $\mathbf{A}_{LDA} = [\mathbf{v}_1, \dots, \mathbf{v}_M]$, $\lambda_1 > \lambda_2 > \dots, \lambda_M$, where $\mathbf{S}_b \mathbf{v}_k = \lambda_k \mathbf{S}_w \mathbf{v}_k$.

Chapter 3

Novel Trend Filtering Method

3.1 Introduction

Trend filtering plays a vital role in data analysis. The presence of the trend term in a signal may hide or give rise to false extrema which severely affects the detection of true extrema [24]. For instance, localization and the counting of heart beat in an embryonic heart signal becomes a challenging problem in presence of trend term [25].

Estimation of the motion signal from a video recording of an embryonic heart has been widely used for the drug screening process [104, 105]. However, the extracted motion signal is often degraded by the trend term which leads to poor data analysis [105]. To solve this problem, we present a novel filtering method that transforms the motion information from the videos to a 1D signal. Moreover, the proposed method estimates the trend term in the generated motion signal without defining any regression model and also helps in screening of pharmaceutical drugs.

Prior to dosing in human subjects, these pharmaceutical drugs are often tested on rat embryos. For this purpose, rat whole embryo culture on gestation day (GD) 13 is often used to investigate if drugs may have an effect on the embryonic heart *in vitro*. If a drug cause reduced heart rate and/or irregular heart rate, it may have the potential to damage the embryo by causing periods of reduced oxygen supply. There is a clear relationship between reduced oxygen supply during mammalian development and an increased risk of birth defects

and embryonic death [106]. For instance, in pregnant women a $\sim 20\%$ reduction in fetal heart rate during the first trimester is associated with a markedly increased risk of spontaneous abortion [107].

The *in vitro* method of investigating drug effects on the rat embryonic heart has improved over the years. However, it is just recently that feasible image analysis tools have been introduced to assess the effect on embryonic heart rate *in vitro* [104, 108]. In comparison, a lot more work has been done to set up image analysis tools in the zebrafish *in vivo* culture system. For instance, in [109], the changes in average light intensity of each frame was combined with pixel differential computed across the frames to construct the heart signal. In [110], the heart signal was visualised as a waveform of dynamic pixels produced by the oscillatory movement of blood cells. Later on, short-time Fourier transform (STFT) was used to analyse the non-stationary heart signal. In STFT, it is assumed that some portion of a non-stationary signal is stationary. However, the determination of the stationary portion and its size is a dilemma in itself. In [111], fast differential interference contrast imaging carried out at 250 frames/sec was combined with autocorrelation to measure the heart activity. To have a true representation of the heart activity, the reference image should represent either of the extreme states in the heart. However, the selection of the reference image, manual or automatic, is un-conclusive in [111].

In the current study, rat embryos are cultured with an open yolk sac. This, together with the vigorous heart beats at GD 13, results in a non-rigid deformation in the shape of the embryo. It can also result in translation of the embryo in the culture medium. Furthermore, in case of an irregular heart rate, the motion of the heart can become non-stationary in nature.

To cope with the above mentioned issues, we carried out the assessment of the heart activity in two steps, *heart signal generation* and *heart signal analysis*. Signal generation is achieved by LEM in conjunction with correlation coefficient, while the signal analysis has been performed by the modified EMD.

3.2 Proposed Method

Existing methods to assess embryonic heart rate requires a skilful operator to perform the experiment by means of expensive imaging equipment [109]. The aim of this study is to present a simple and low cost solution to measure the rat embryonic heart rate *in vitro*. The proposed solution requires an ordinary light microscope with a video camera capable of capturing a video at 30 frames per second along with a desktop computer. The section to follow will present a robust method to generate a signal representing the heart motion from video recording of the cultured rat embryos.

3.2.1 Heart Signal Generation

LEM is a non-linear dimensionality reduction method which aims to find the lower dimensional manifold embedded in the higher dimensional space while preserving the spatial relationship [112]. To accomplish this task, LEM constructs a graph G in which every data point d_i is connected to its k -nearest neighbours. All edges between the connected data points in a graph G have a cost equal to one. It is followed by the construction of an adjacency matrix C_{ij} which have an entry $\mathbf{1}$ at location (i, j) , if the data point d_i is among the k -nearest neighbours of d_j . The rest of the locations in C_{ij} are set to zero. LEM requires the construction of Laplacian matrix L , which can be computed as:

$$L_{ij} = \begin{cases} D_{ij} & \text{if } i = j \\ -C_{ij} & \text{if } d_i \text{ and } d_j \text{ are adjacent data points} \\ 0 & \text{otherwise.} \end{cases} \quad (3.1)$$

Here, D is a diagonal matrix computed as $D_{ii} = \sum_j C_{ij}$. The final step in the LEM method is to find the generalized eigenvector solution to:

$$L\mathbf{f} = \lambda D\mathbf{f}. \quad (3.2)$$

As L is a symmetric positive semi-definite matrix with $\lambda_0 = 0$ as a trivial eigenvalue, it implies that all eigenvalues can be ordered as: $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \cdots \leq \lambda_{N-1}$. Now, the

mapping of d_i in a low-dimensional space m is computed by leaving the eigenvector \mathbf{f}_0 and using the next m eigenvectors as [112]:

$$d_i \longrightarrow (\mathbf{f}_1(i), \mathbf{f}_2(i), \dots, \mathbf{f}_m(i)). \quad (3.3)$$

Here, the mapping \mathbf{F} is the solution to the following minimization problem [112]:

$$\mathbf{F} = \arg \min_{\mathbf{f}_1 \dots \mathbf{f}_m} \sum_{(i,j)} C_{ij} \|\mathbf{f}^{(i)} - \mathbf{f}^{(j)}\|^2 \text{ with: } \mathbf{F}^T \mathbf{F} = \mathbf{I} \quad (3.4)$$

where $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m]$ is a $n \times m$ matrix, n represents the total number of data points, and $\mathbf{f}^{(i)}$ represents the i -th row of matrix \mathbf{F} , i.e., corresponds to the m -dimensional Euclidean coordinates of the i -th data point d_i .

As, we are interested in measuring the heart rate in the atrium and the ventricle, respectively, the method requires manual selection of these regions. In our case, the area representing either chamber would be treated as d_i . It is intuitive that LEM on such data will provide quite logical representation by mapping the filled, semi-filled, emptied atrium frames close to filled, semi-filled, emptied atrium frames, respectively in the lower dimensional space. The same would be the case with the ventricle. K -medoids clustering of this low-dimensional mapping into three clusters yields intuitive classification, i.e., filled, semi-filled, emptied chamber classes. This mapping eases the selection of the template frame for the atrium and the ventricle. It is worth mentioning that each low-dimensional point has a one-to-one correspondence with the higher dimensional point. The class having the least average value (computed among the corresponding higher dimensional points) is considered as the state when the specific heart chamber is fully filled. It is mainly because the least light would pass through blood in the case when chamber is fully filled. The point in the higher dimensional space is considered as a template if the corresponding lower dimensional point represents the centre of this chamber class. The template selected using this method is relatively insensitive to outliers and noise [112]. To construct the adjacency matrix, we have used 5-nearest neighbours and for the low-dimensional embedding, we have set $m = 2$.

The choice of 5-nearest neighbours is motivated by the fact that it always resulted in one-connected component in the graph for our test videos.

Later on, cross-correlation of the template frame (representing the atrium filled class) with the area representing the atrium will yield the heart signal representing motion of the atrium (A_s). The ventricle activity (V_s) is achieved similarly by using the template frame (representing the ventricle filled class) with the area representing the ventricle. For the sake of simplicity, we will use the term *heart signal* to refer to both, A_s and V_s .

3.2.2 Heart Signal Analysis

Depending on the heart conditions, the associated heart motion is intrinsically non-stationary in nature. It is also observed that the spatial movement of the embryo due to placement in a culture medium (liquid in nature) results in a trend term in the associated heart signal. In case of arrhythmia, the heart motion will change over time which demands for time-frequency analysis tool. Generally, wavelets and filter banks provide a pre-defined multi-scale representation of a non-stationary signal. The application of wavelets require the definition of a mother wavelet as well as the total number of scales that are both user defined parameters [24, 113]. However, it will be more logical to select these parameters based on the frequency content of the underlying signal but they are a priori unknown. The recently developed EMD seems suitable, as it overcomes the above mentioned shortcomings.

EMD is an adaptive multi-scale representation that decomposes a non-stationary signal into symmetrical oscillating functions known as intrinsic mode functions (IMF) and a less oscillating local mean [24, 89]. Each IMF has the same number of extrema and zero-crossings or can differ at most by one. IMF should also be symmetric with respect to local zero mean. To extract each IMF, EMD uses an iterative procedure known as sifting process [89]. In [114], the sifting process was replaced with partial differential equations to improve the performance of EMD. Another approach to improve the performance of EMD is to better approximate the mean envelope as the conventional mean envelope can result in undershoot, overshoot, instabilities to noise, and erroneous detection of extrema [115].

To overcome the above mentioned problems, we have followed the same lines as discussed in [116] and [89] along with a new relaxation in the stopping criterion, which helps in faster convergence. The proposed method provides the steps to remove the trend term and noise from the heart signal as summarized in Algorithm 2.

Algorithm 2 Modified EMD for Signal Analysis and Trend Removal

- 1: $f_r \leftarrow \text{frame rate of the camera}$
- 2: $f_h \leftarrow 6 \text{ Hz}$ (*Maximum heart beat rate*)
- 3: $i \leftarrow -1, \varepsilon_1 \leftarrow 0.05, \varepsilon_2 \leftarrow 0.10, \alpha \leftarrow 2$
- 4: **repeat**
- 5: Find all extrema locations (t_k) in $z(t)$.
- 6: Compute the centroid $(\bar{z}_k(\bar{t}_k))$ between two consecutive extremum as:

$$\bar{z}_k \leftarrow \frac{1}{t_{k+1} - t_k} \int_{t_k}^{t_{k+1}} z(t) dt$$

$$\bar{t}_k \leftarrow \frac{\int_{t_k}^{t_{k+1}} t |z(t) - \bar{z}_k|^2 dt}{\int_{t_k}^{t_{k+1}} |z(t) - \bar{z}_k|^2 dt}$$

- 7: Find a mean envelope $m(t)$ by fitting a cubic spline through all centroids $(\bar{z}_k(\bar{t}_k))$.
- 8: **if** $(z(t) - m(t))$ is an *imf* **then**
- 9: $i \leftarrow i + 1$
- 10: Find the IMF, $h_i(t)$:

$$h_i(t) \leftarrow z(t) - m(t)$$

- 11: Find the residual, $r(t)$:

$$r(t) \leftarrow z(t) - h_i(t)$$

$$z(t) \leftarrow r(t)$$

- 12: Find all extrema locations t_k in $z(t)$.
- 13: Find the symmetry ratio, s_r :

$$\lambda(k) \leftarrow \frac{z(t_{k+1}) - z(t_{k-1})}{t_{k+1} - t_{k-1}}(t_k - t_{k-1}) + z(t_{k-1})$$

$$s_r \leftarrow \frac{|z(t_k) + \lambda(k)|}{|z(t_k) - \lambda(k)|}$$

14: **else**

15:

$$z(t) \leftarrow z(t) - m(t)$$

16: Find all extrema locations t_k in $z(t)$.

17: Find the symmetry ratio, s_r :

$$\lambda(k) \leftarrow \frac{z(t_{k+1}) - z(t_{k-1})}{t_{k+1} - t_{k-1}}(t_k - t_{k-1}) + z(t_{k-1})$$

$$s_r \leftarrow \frac{|z(t_k) + \lambda(k)|}{|z(t_k) - \lambda(k)|}$$

18: **end if**

19: Find the number of extrema n_e in $z(t)$.

20: **until** $((s_r \geq \varepsilon_1) \wedge (n_e > \alpha))$

21: Find the cut-off frequency cut_{off} to remove noise and higher order harmonics:

$$cut_{off} \leftarrow \frac{2f_h}{f_r} + \varepsilon_2$$

22: Convolve h_0 with a FIR low-pass filter fil , with a cut-off frequency of cut_{off} .

23: Drop the trend term to compensate for the spatial heart motion in the culture medium:

$$z_{comp}(t) \leftarrow \sum_{i=0}^{n-1} h_i(t)$$

where n is the total number of IMFs.

The proposed method helps to find the trend term without defining the regression model. Moreover, the non-rigid shape deformation produced due to the vigorous heart motion, the noise induced during the data acquisition, and the higher order harmonics (produced due to unequal stay of blood in each of the heart chambers) are also catered during step 21 to 23 of Algorithm 2. Taking benefit from the zero mean property of the IMF, we have defined a beat

in a different way. We count it as a beat if there is a negative minimum between two positive maxima. Here, maxima is defined as the highest positive value between every two negative minima. The positive minimum separating the two positive maxima is considered as false minima and the negative maxima separating the two minima is considered as false maxima. This methodology avoids false maxima/minima detection and also avoids the computation of differential for maxima/minima detection. Such a definition assures that every beat has substantial amplitude and also justifies the removal of the trend term.

3.3 Results and Discussion

For evaluation of the proposed method, we have used 151 videos to investigate the effects of pharmaceutical drugs on the heart activity of rat embryos. To generate the videos, we used a culture system (with some modification) previously published in [105]. In this study, each embryo was cultured in 25 ml bottles containing 4 ml of Dulbeccos Modified Eagles Medium (DMEM, Ref. No. D1145, Sigma Chemical Co., St. Louis, MO, USA). Intermittent gassing (95% O₂, 5% CO₂) for 2 minutes instead of continuous gassing was used, and the bottles were gassed after addition of the drugs to the culture media, and after every recording of the heart rate.

After 1 hour of incubation, each embryo still in its bottle, was examined under a light microscope (Olympus SZ-40) equipped with a camera (uEye UI-2210-M/C). A 30-second video of the embryos was recorded to be used for later analysis. Damaged and dead embryos or embryos with a heart rate less than 160 beats per minute were discarded. After recording the embryonic heart rate, the test compound (or vehicle serving as control) was added to the culture medium and the bottles were gassed for 2 minutes with the same gas mixture as above. The embryos were then incubated for 1 hour before being re-examined again as described above. In total, we have used 151 videos for the evaluation of our proposed method.

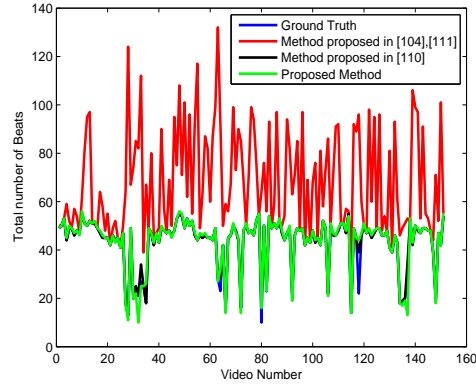
We opted to compare our heart signal analysis method with the analysis methods mentioned in [104,110,111]. We skipped the comparison between the proposed signal generation

method and the signal generation methods mentioned in [110,111], because they have used different imaging technique to acquire images. However, it seems fair to compare the signal analysis methods as there is no difference in the generated signals. For validation of our method, we have generated a ground truth (G_t) where videos were stepped through frame by frame and the exact heart rate was determined visually. Fig. 3.1(a) shows the number of heart beats counted by each of these methods in 151 different videos. It is clear from Fig. 3.1(a) that the proposed method is quite accurate in counting the heart beats in comparison to the others.

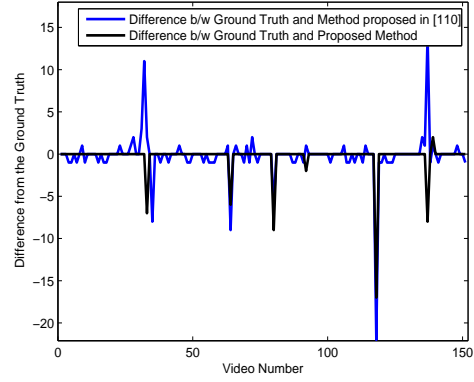
Fig. 3.1(b) shows the difference between G_t and the number of beats counted by the proposed method for all 151 videos. It also shows the difference between G_t and the beats counted by [110]. It is clear from Fig. 3.1(b) that the difference between G_t and the number of beats counted by the proposed method is quite small as compared to [110]. Fig. 3.1(c) shows the difference between G_t and the number of heart beats counted by [104,111], which is quite higher as compared to the proposed method. It is interesting to mention that the method proposed in [110], provides quite accurate count of the heart beats by finding the fundamental frequency in the frequency domain but fails to localize the location of the peaks in the time domain. Fig. 3.1(d) shows the standard deviation of the inter-beat time, which is often used by biologists to judge the heart condition, i.e., to differentiate between regular and irregular heart activity. Here, the higher values are often associated with the irregularity of the heart motion. It is evident from the results that the proposed method outperforms the other methods but still there are 5 instances when the proposed method fails to correctly count the heart beats.

3.4 Summary

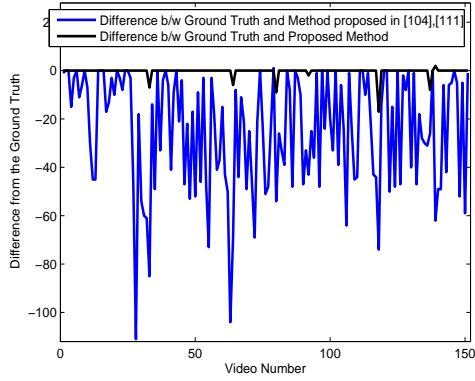
To develop an accurate and suitable method for measuring the embryonic heart rate *in vitro*, a system combining LEM and EMD has been proposed. The proposed method assess the heart activity in two steps; *heart signal generation* and *heart signal analysis*. Signal generation is achieved by LEM in conjunction with correlation coefficient, while the signal



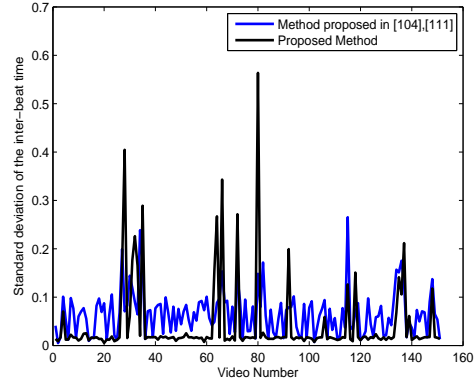
(a)



(b)



(c)



(d)

Figure 3.1: a) Number of heart beats counted by different methods on 151 videos. b) Difference between G_t and the number of beats counted by the method proposed in [110]. c) Difference between G_t and the number of beats counted by the methods proposed in [104,111]. d) Standard deviation of the inter-beat time computed individually for every video.

analysis of the heart motion has been performed by the modified EMD. LEM helps to find the template for the atrium and the ventricle respectively, whereas EMD helps to find the trend term without defining any regression model.

The proposed method also removes the motion artifacts produced due to the non-rigid deformation in the shape of embryo, the noise induced during the data acquisition, and the higher order harmonics. To check the authenticity of the proposed method, 151 videos have been investigated. The results clearly show the accuracy of the proposed method in counting and localizing different heart states. The efficiency and the robustness of the proposed method makes it more attractive for the biologists to analyse the heart motion without the need of expensive imaging equipment and parameter tweaking.

Chapter 4

Global Intensity Inhomogeneity Correction

4.1 Introduction

In multi-dimensional signals, the trend term is often referred to as shading, non-uniform illumination, intensity non-uniformity, global intensity inhomogeneity, or bias field [32]. Global intensity inhomogeneity is a multiplicative field that induces smooth variations across the whole image [33]. Due to these smooth variations, it is often characterised by the low frequency components [35]. It is mainly caused due to the imperfections of data acquisition device, direction of light source, and subject topology, etc [32]. Global intensity inhomogeneity correction is often considered as a prerequisite to a successful image segmentation [33]. The performance of image matching, retrieval and tracking algorithms are also affected by the presence of global intensity inhomogeneity [46, 47]. This wide range of applications has lead to the development of a plethora of methods. A thorough review of the development during the last decade is reported in [32, 34].

Generally, an image is modelled as a product of the true signal $t(x, y)$ and intensity inhomogeneity $i(x, y)$ with some noise $\eta(x, y)$ added by the acquisition device. Mathematically, it can be written as:

$$f(x, y) = t(x, y)i(x, y) + \eta(x, y). \quad (4.1)$$

In the discussion to follow, we will assume an additive image model as the noise $\eta(x, y)$ in Eq. 4.1 can be suppressed by low-pass filtering. This allows us to transform the multiplicative image model into an additive model by using a logarithmic transformation.

The main task of the global intensity inhomogeneity correction is to adjust the magnitudes of those low frequencies that are affected by the addition of global intensity inhomogeneity. Now, a simple question arises: how and where to find those frequencies in the Fourier transform? Theoretically, the question is complex to answer. Usually, it is solved empirically with user interaction where the user selects the appropriate parameters of the filter which suppresses the affected frequency components [11]. Normally, circular filters are used due to the lack of information about the location of the frequencies that are affected by global intensity inhomogeneity. Another reason for using circular filters is the ease in their design [41]. The circular filter not only suppresses the affected frequencies but also removes additional frequency components of an image. This often results in a high-pass look of the resulting image and can also result in inducing further inhomogeneity in an image. So, it is desirable to only suppress the frequencies that are affected by the global intensity inhomogeneity and to have very low effect on rest of the frequency components of an image. During a literature review [32], we have found that most of the Fourier transform based global intensity inhomogeneity correction methods require user interaction where the user fine tunes the parameters based on his/her visual inspection [32, 34]. This user interaction makes the use of those methods impractical, especially in a large dataset.

4.2 Proposed Method for Global Intensity Inhomogeneity Correction

The proposed method constructs a low-pass filter based on the assumption that the addition of the global intensity inhomogeneity will amplify the magnitude of the affected low frequencies. So, we can expect a relatively higher magnitude for those low frequencies as compared to their neighbours in the Fourier transform. As the proposed method is based

on the magnitude of the affected low frequencies, it can produce a filter whose shape (regular or irregular) depends on the distribution of frequencies affected by the global intensity inhomogeneity. To the best of our knowledge, this prior information has never been used in filter design nor has been considered by the user while choosing the cut-off of the filter.

Another crucial issue is to determine the magnitude of the affected low frequencies that needs to be suppressed. Should it be suppressed completely? In homomorphic unsharp masking (HUM), this is controlled by a parameter γ_L which is used to scale down the magnitude of the low frequency components [11]. Generally, the value of γ_L is selected by the user on the basis of visual inspection which tends to make the results highly subjective and dependent on user's experience. To overcome these issues, the proposed method iteratively selects the amount of magnitude of the affected low frequencies that needs to be suppressed. This is extremely important as complete suppression of these low frequencies can produce further inhomogeneity in an image. In other words, retaining a certain amount of magnitude of these low frequencies is necessary to recover the true image. The proposed method is summarized in Algorithm 3.

Algorithm 3 Global Intensity Inhomogeneity Correction based on GWDT in 2D Images

```

1:  $f = \text{input\_image}$ 
2:  $F = \text{fftshift}(\text{fft2}(f))$ ; // Compute the centered shifted Fourier transform (CSFT).
3:  $j = 0$ ; // initialization
4:  $\text{similarity}[j] = 1$ ; // initialization.
5: repeat
6:    $j = j + 1$ ;
7:    $F_{\text{mag}} = \text{abs}(F)$ ; // Magnitude calculation.
8:    $\text{Dist}_{\text{map}} = \text{GWDT}(F_{\text{mag}})$ ; // Compute the distance map using GWDT.
9:   for  $r = 1, \dots, n$  do
10:     $\text{Dist} = (\text{Dist}_{\text{map}} \leq r) \times F_{\text{mag}}$ ;
11:     $P_R(r) = \frac{\text{sum}(\text{abs}(\text{Dist})^2)}{\text{sum}(\text{abs}(F)^2)} \times 100$ ; // Calculating the percentage of the image power.
12:   end for
```

```

13:  Select a threshold value ( $T$ );
14:   $Mask = Dist_{map} \leq T$ ;
15:   $Low\_Pass = Mask * Gaussian\_filter$ ; // Here,  $*$  represents the convolution operator.
16:   $High\_Pass = \max(Low\_Pass) - Low\_Pass$ ;
17:   $ADJUSTED = High\_Pass \times F$ ; // Frequency domain filtering.
18:   $F = ADJUSTED$ ;
19:   $adjusted = \text{real}(\text{ifft2}(\text{ifftshift}(ADJUSTED)))$ ;
20:   $similarity[j] = \text{correlation}(adjusted, f)$ ;
21: until  $0 \leq (similarity[j - 1] - similarity[j]) < \epsilon$ 

```

The proposed method starts by calculating the magnitude of the Fourier transform of the given input image as shown in Fig. 4.1(b). As, we are interested in suppressing the magnitude of the affected low frequency components, i.e., relatively higher in magnitude and spatially close to the DC component, GWDT is one of the methods that can be used to fulfil both constraints. GWDT provides a framework to combine spatial information with magnitude information [117]. For this reason, the magnitude image F_{mag} , becomes an input to the GWDT. For instance, let \mathcal{P}_{pq} represents the set of all possible paths between pixels p and q . Let $\pi \in \mathcal{P}_{pq}$ such that $\pi = \{p = p_0, p_1, \dots, p_{n-1} = q\}$, where p_i and p_{i+1} are adjacent pixels. Then, the length \mathcal{L} of the path π is defined as:

$$\mathcal{L}(\pi) = \sum_{i=0}^{n-1} |F_{mag}(p_i) - F_{mag}(p_{i+1})| \cdot \|p_i - p_{i+1}\|, \quad (4.2)$$

where $F_{mag}(p_i)$ and $F_{mag}(p_{i+1})$ represents the magnitude value of Fourier transform at p_i and p_{i+1} . Here, $|\cdot|$ represents the absolute and $\|\cdot\|$ denotes the Euclidean norm. Now, the grey-weighted distance between p and q can be defined as:

$$d(p, q) = \begin{cases} \min_{\pi \in \mathcal{P}_{pq}} \mathcal{L}(\pi) & \text{if } p \neq q \\ 0 & \text{if } p = q. \end{cases} \quad (4.3)$$

Readers interested in deep understanding of GWDT and its efficient implementation are

referred to [117,118]. In the proposed method, the DC component of the magnitude of CSFT serves as the only seed point for the GWDT. Fig. 4.1(c) shows the distance map ($Dist_{map}$) computed using the GWDT on the magnitude (F_{mag}) of the CSFT. Here, the range of the values from low to high shows the grey-weighted distance from the seed.

The next step of the proposed method is to compute the percentage of the image power

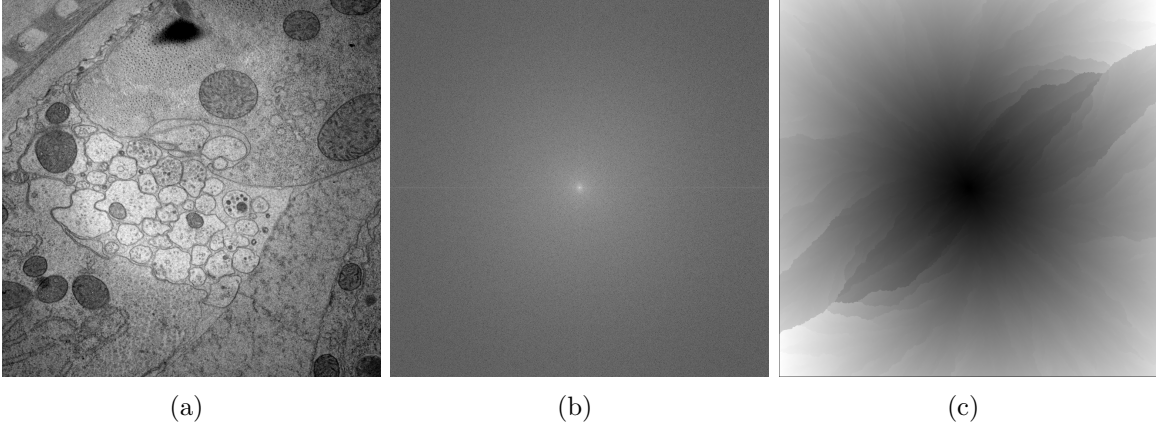


Figure 4.1: a) Original TEM image of a c-elegans. b) Magnitude of the CSFT of 4.1(a). c) Result after applying GWDT on 4.1(b).

depending on the grey-weighted distance from the DC component. This distance shows the change in the magnitude of the frequencies from the DC component in each direction independently, which helps to find the irregular distribution of frequencies affected by the global intensity inhomogeneity. It is followed by finding the locations of these affected frequencies. To accomplish this, let $M : Dist_{map} \rightarrow F_{mag}$, and $s_0 \in Dist_{map}$ is the location of the seed. Let $B(s_0, r)$ be a ball with the centre in s_0 and radius r , i.e.,

$$B(s_0, r) = \{q \mid d(s_0, q) \leq r\}, \quad (4.4)$$

where q defines the location of the pixels contained within a radius r . Then, the set of the frequencies within radius r is defined as:

$$K(s_0, r) = \{M(q) \mid q \in B(s_0, r)\}. \quad (4.5)$$

It is followed by computing the percentage of the image power ($P_R(r)$) as shown in Fig. 4.2. Mathematically, it can be written as:

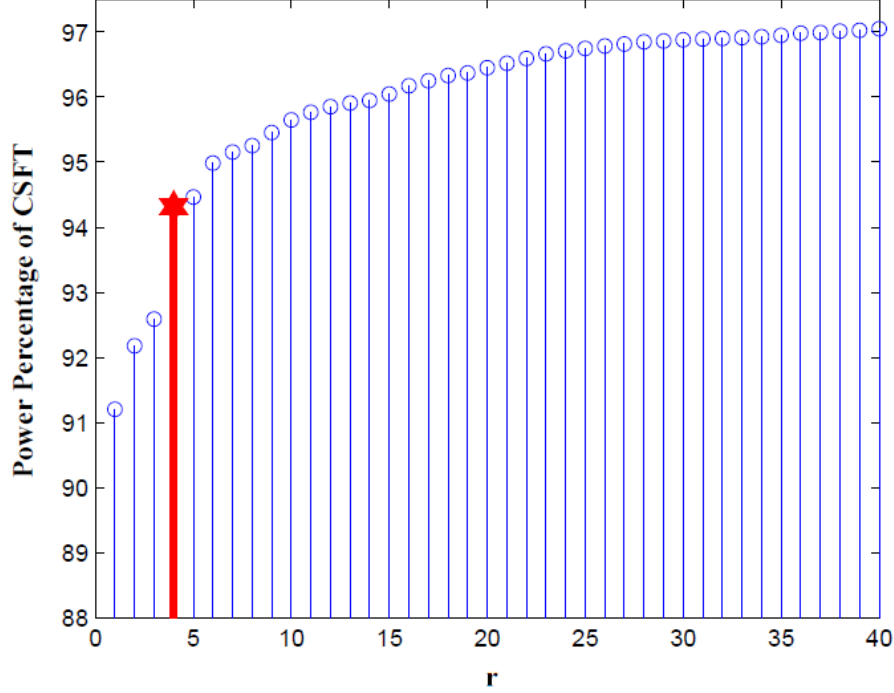


Figure 4.2: Percentage of the image power for a ball of radius \mathbf{r} . Here, the threshold value ($T = 4$) is highlighted in red.

$$P_R(r) = \frac{\Omega(K(s_0, r))}{\Omega(F)} \times 100, \quad (4.6)$$

where the total image power is computed by:

$$\Omega(F) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F(u, v)|^2. \quad (4.7)$$

where the M and N represents the total number of rows and columns of an image.

The value of r at which there is a sharp change in $P_R(r)$ for the ball $B(s_0, r)$ will serve as the threshold (T) as shown in Fig. 4.2. This threshold T , corresponds to the ball outside which the magnitude F_{mag} , drops abruptly. Essentially, this means that all the low frequencies contained within this ball have relatively much higher magnitude than the frequencies

outside the ball. This ball can take on any shape depending on the magnitude of the affected low frequencies relative to the DC component. Fig. 4.3(a) shows an example of a filter resulting from Fig. 4.1(c). Fig. 4.3(b) shows the zoomed version of Fig. 4.3(a). According to our observation this threshold is usually between 3 to 5. This threshold value will help us to find an ideal low-pass filter (F_{low}), which will correspond to the frequencies affected by the global intensity inhomogeneity. F_{low} can be expressed as:

$$F_{low} = \begin{cases} 1 & \text{if } F_{mag} \in K(s_0, T) \\ 0 & \text{if } F_{mag} \in \overline{K(s_0, T)}. \end{cases} \quad (4.8)$$

To avoid the ringing artifacts produced by an ideal filter, we have smoothed our filter with a Gaussian of $\sigma = 1$ as shown in Fig. 4.3(c). To suppress these frequencies, we have used a *High_Pass* filter as shown in Fig. 4.3(d), which is produced in step 16 of Algorithm 3. The next step of the proposed method is to filter the given image by using the *High_Pass* filter and take the resulting image (*adjusted*) back to the spatial domain. The whole process is repeated until the difference between the similarity (as computed in step 20 of Algorithm 3) is less than ϵ . We have set ϵ to 0.05, however any value less than 0.1 works fine for the applications that we have tested. In each iteration, the filter will suppress the magnitude of the affected low frequencies, gradually from one iteration to the next. Empirically, we have noticed that the proposed method takes between 3 to 5 iterations to converge.

4.3 Experimental Setups

For the experimental evaluation of the proposed method, we have used five different setups.

4.3.1 Experimental Setup 1

In experimental setup 1, we have used 40 binary images from the digital retinal images for vessel extraction (DRIVE) database [120]. Here, a binary retinal image $B(x, y)$ represents the pure reflectance which essentially means that illumination is constant throughout the

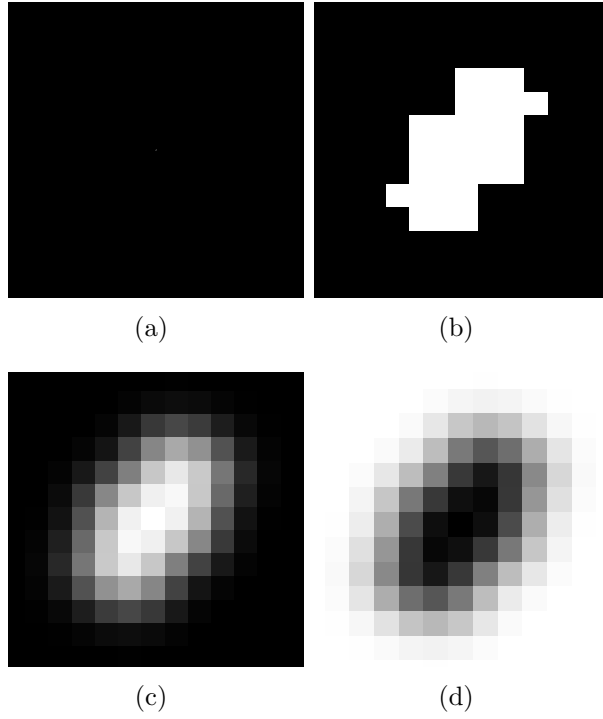


Figure 4.3: a) Ideal low-pass filter. b) Zoomed version of the filter in 4.3(a). c) Smoothed version of the filter in 4.3(b). d) High-pass filter.

image as shown in Fig. 4.4(a). To synthesize the image $I_n(x, y)$ having global intensity

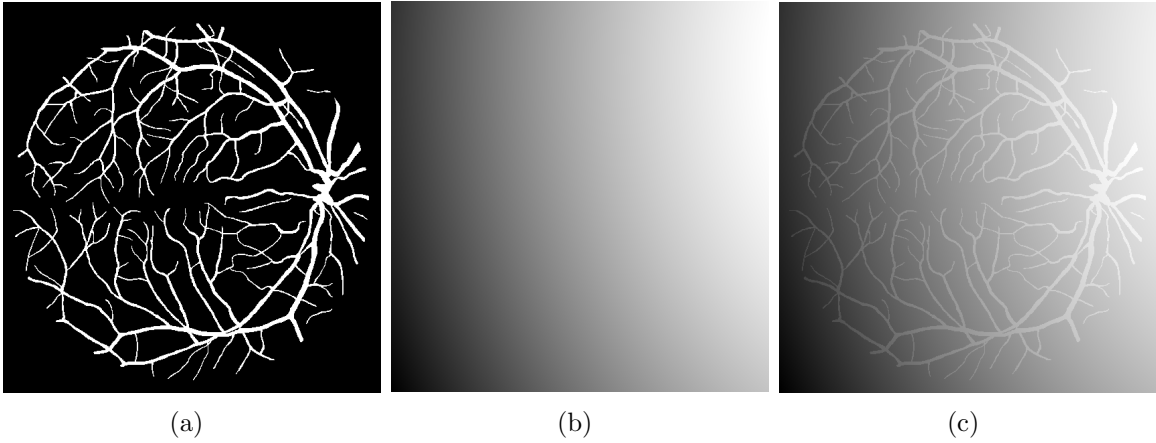


Figure 4.4: a) Binary retinal image $B(x, y)$ from the DRIVE database [120]. b) Global intensity inhomogeneity pattern $P(x, y)$. c) Global intensity inhomogeneity image $I_n(x, y)$ produced by adding the global intensity inhomogeneity pattern of 4.4(b) to the binary retinal image.

inhomogeneity, the first step is to generate an image $Q(x, y)$ having values randomly drawn from a standard normal distribution $\mathcal{N}(0, 1)$. It is convolved with a Gaussian low-pass filter G by using the following equation:

$$P(x, y) = Q(x, y) * G_{\mu, \sigma}, \quad (4.9)$$

where $P(x, y)$ is the synthetic global intensity inhomogeneity pattern and $G \sim \mathcal{N}(0, 75)$. Finally, an inhomogeneous synthetic image $I_n(x, y)$ is created by:

$$I_n(x, y) = B(x, y) + P(x, y). \quad (4.10)$$

One such example is shown in Fig. 4.4(c), where the synthetic global intensity inhomogeneity pattern $P(x, y)$ is added to the binary retinal image $B(x, y)$. For simplicity, we are considering the additive image model in Eq. 4.10, as one can easily transform a multiplicative model into additive by using the logarithmic transform. We have added a random pattern to each of the 40 binary retinal images. This process was repeated 10 times to produce a test set of 400 images.

4.3.2 Experimental Setup 2

In experimental setup 2, we have created 400 synthetic images $S_g(x, y)$ by using Gaussians of varying sigmas. All of these Gaussians were placed randomly in a digital grid with their centers apart from each other by $6\sigma_{max}$, where σ_{max} is the maximum standard deviation among all the Gaussians used to create $S_g(x, y)$. Their centers were kept $6\sigma_{max}$ apart to avoid the overlap between the neighbouring Gaussians. For each image, four Gaussians of different sigmas were used and they were placed randomly to form an image. One such synthetic image is shown in Fig. 4.5. Now, to induce global intensity inhomogeneity into the synthetic image $S_g(x, y)$, we created different kinds of sinusoidal intensity inhomogeneity patterns using:

$$P(x, y) = A \sin(2\pi(\frac{ux}{M} + \frac{vy}{N})) + A, \quad (4.11)$$

where A is the amplitude, u is the horizontal frequency, v is the vertical frequency, M is the total number of rows of an image and N is the total number of columns of an image. Different values of u and v were picked randomly between $(0, 1]$ for creating sine waves of different frequencies as it resembles the global intensity inhomogeneity pattern produced by different imaging modalities. Finally, the inhomogeneous image $I_n(x, y)$ is created by:

$$I_n(x, y) = S_g(x, y) + P(x, y). \quad (4.12)$$

We have also used global intensity inhomogeneity pattern generated by polynomial of degree two and three. A third degree polynomial has the following form:

$$P(x, y; a) = a_1x^3 + a_2y^3 + a_3, \quad (4.13)$$

where $a = a_1, a_2, a_3$ is the parameter vector that defines the surface. Finally, inhomogeneous image is created by using Eq. 4.12. Fig. 4.6 illustrates few examples of our synthetic images.

4.3.3 Experimental Setup 3

A confocal scanning laser ophthalmoscope is widely used to assess the health of the retina. Retinal images often suffer from global intensity inhomogeneity [119], which in return can affect the feature extraction, registration, and segmentation process. Moreover, the performance of the biometric systems based on the retinal vessel tree structure are also affected

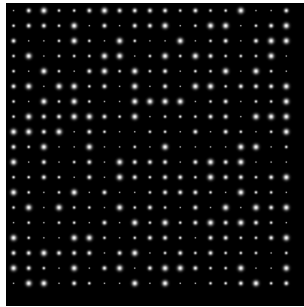


Figure 4.5: Synthetic image $S_g(x, y)$, created by using Gaussians of varying sigmas.

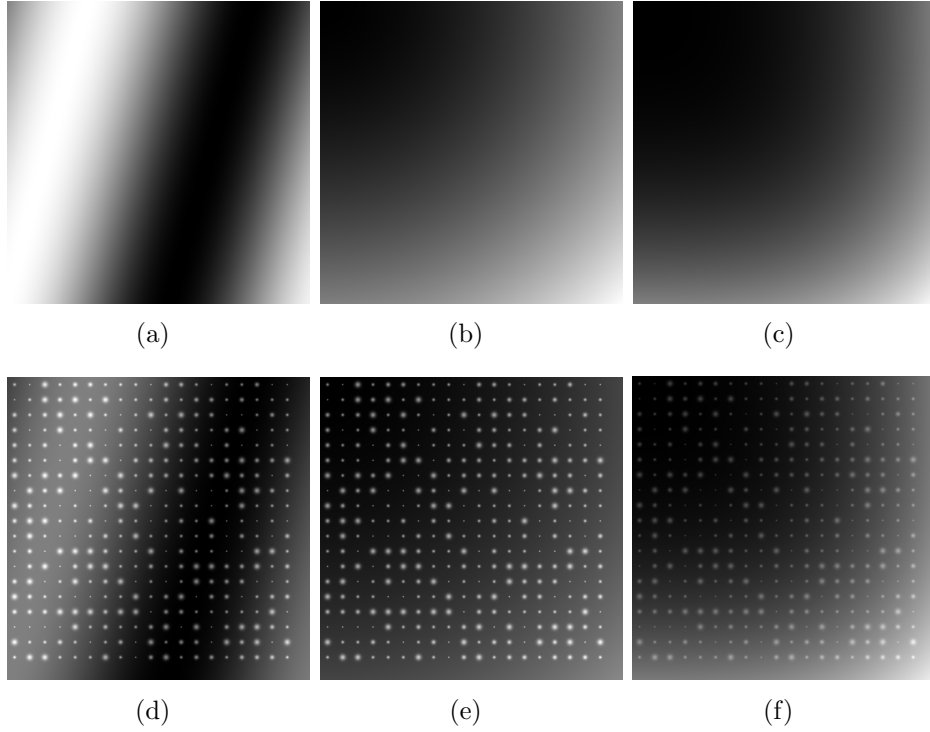


Figure 4.6: a) Global intensity inhomogeneity pattern created by using Eq. 4.11. b) Global intensity inhomogeneity pattern created by a second order polynomial. c) Global intensity inhomogeneity pattern created by using Eq. 4.13. d) Global intensity inhomogeneity image created by adding 4.6(a) with Fig. 4.5. e) Global intensity inhomogeneity image created by adding 4.6(b) with Fig. 4.5. f) Global intensity inhomogeneity image created by adding 4.6(c) with Fig. 4.5.

by the presence of global intensity inhomogeneity. This makes the retinal image an ideal candidate to test the proposed method.

To evaluate the performance of the proposed method, we have used three public databases: DRIVE [120], STARE [121], and VICAVR [122]. The DRIVE database consists of 40 eye-fundus color images taken with a Canon CR5 nonmydriatic 3CCD camera with a 45° field-of-view (FOV). Each of the image is captured at a resolution of 768×584 pixels with 8 bits per color channel and stored in TIFF format. The database is divided in two sets: a test set and a training set. Each of the sets contains 20 images. The database also provides the FOV masks corresponding to each image and the manually labelled ground truth images provided by the two experts. Two examples from the DRIVE database are shown in Fig. 4.7(a) and

Fig. 4.7(b). The STARE database contains 20 images captured by a TopCon TRV-50 fundus camera at 35° FOV. All the images are captured at a resolution of 605×700 pixels with 8 bits per colour channel. All the images are manually labelled by the two experts. Fig. 4.7(c) and Fig. 4.7(d) shows some of randomly selected images from the STARE database.

The VICAVR database contains 58 images. All the images are taken with TopCon non-mydratic camera NW-100 with a resolution of 768×584 pixels. Examples from the VICAVR database are shown in Fig. 4.7(e) and Fig. 4.7(f).

4.3.4 Experimental Setup 4

In this experimental setup, we have used the transmission electron microscopy (TEM) images of c-elegans as used in [35]. TEM images often suffer from global intensity inhomogeneity due to electron imaging defects (non-uniform support films) and specimen staining [35]. One such image is shown in Fig. 4.8, where the global intensity inhomogeneity is evident in the center of the image. Presence of such inhomogeneities implies a need for image restoration prior to further image analysis. Hence, they are suitable candidates for the proposed method.

4.3.5 Experimental Setup 5

The literature shows that many of the filtering methods for global intensity inhomogeneity correction are tested on MR images [32]. In MR images, the global intensity inhomogeneity can be induced by the choice of the radio-frequency coil, the acquisition pulse sequence and by the nature and geometry of the sample itself [34]. To have a fair comparison, we have opted to evaluate our method on 3D MR volumes obtained from the BrainWeb Simulated Brain (BSB) Database [123]. This database provides both, the images suffering with global intensity inhomogeneity as well as the ground truth. Here, the inhomogeneous images are created using inhomogeneity estimated from real MR scans [123]. The induced global intensity inhomogeneity is non-linear and complex in shape. In total, we have used 48 MR volumes of dimensions $181 \times 217 \times 181$. The simulated data sets are obtained with the following

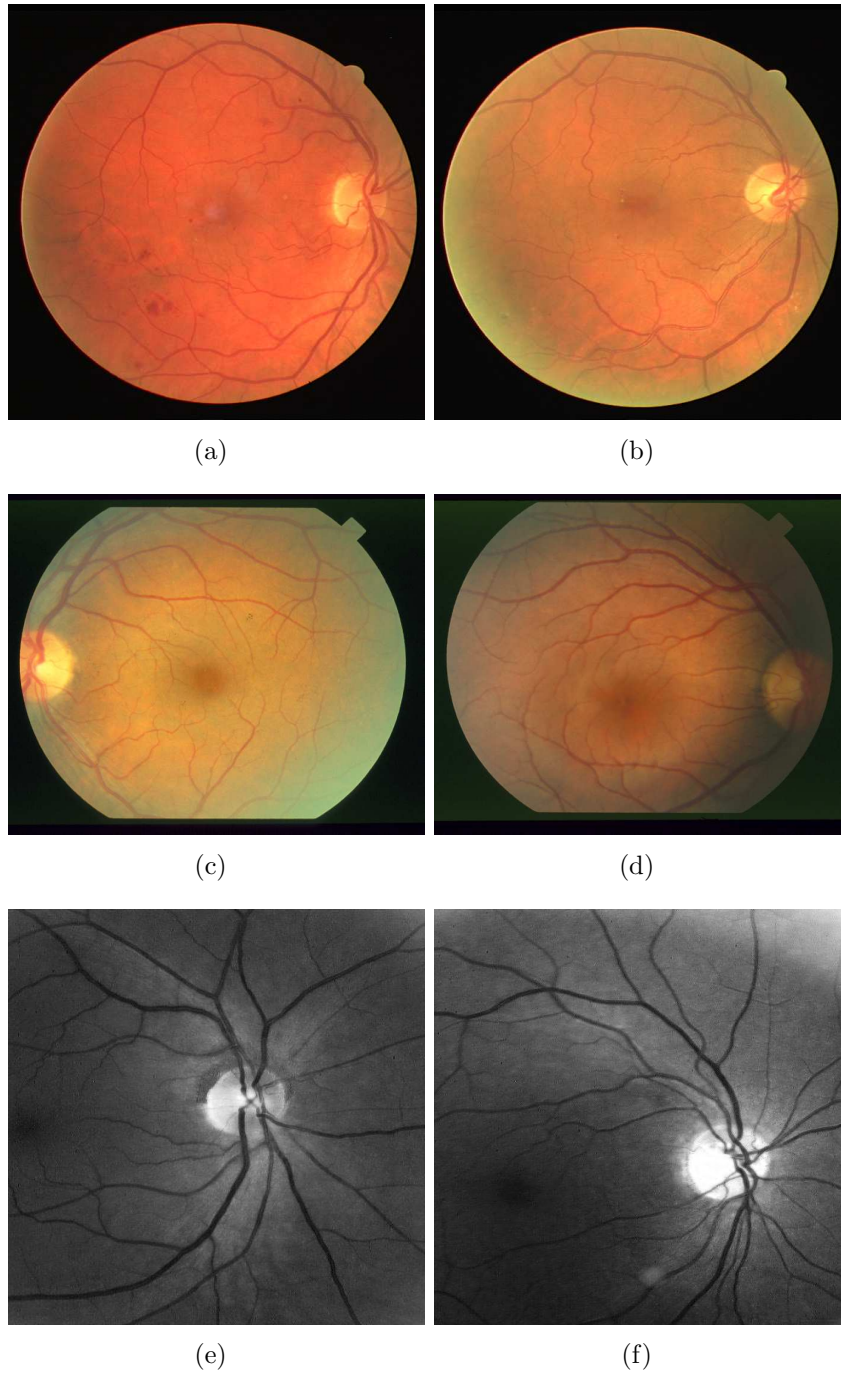


Figure 4.7: a), b) Images from the DRIVE database [120]. c), d) Images from the STARE database [121]. e), f) Images from the VICAVR database [122].

settings: T1, T2 relaxation times, slice thickness of 1, 3, and 5 mm, all global intensity inhomogeneity shapes, and 0%, 20%, 40%, 50%, 60%, 70%, 80%, and 90% inhomogeneity levels. Here, 90% inhomogeneity level means that the global multiplicative inhomogeneity has a range of values between 0.55 and 1.45 over the brain area. Linear scaling is used to produce other inhomogeneity levels. We created six different (with various slice thickness, inhomogeneity shapes, and relaxation times) MR volumes for each intensity inhomogeneity level. Fig. 4.9(a) shows the transaxial view of a slice from the test volume.

4.4 Results and Discussion

4.4.1 Results of Setup 1

In experimental setup 1, a synthetic inhomogeneous pattern $P(x, y)$ of size $M \times N$ was added to a binary retinal image $B(x, y)$ to produce a global intensity inhomogeneous image $I_n(x, y)$. One way to evaluate the proposed method in such a situation would be to recover $B(x, y)$ from $I_n(x, y)$. To accomplish this task, we run the proposed method on $I_n(x, y)$, which resulted in an image $R(x, y)$. Then Otsu's thresholding was applied afterwards on $R(x, y)$ to convert it into a binary image $B_R(x, y)$ [124]. To quantify the accuracy of the

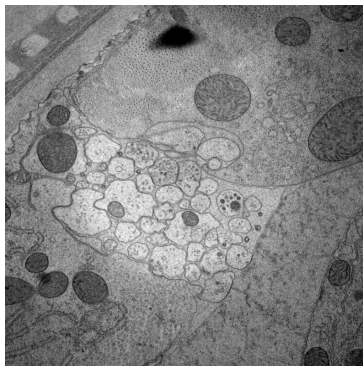


Figure 4.8: TEM image of c-elegans suffering from global intensity inhomogeneity.

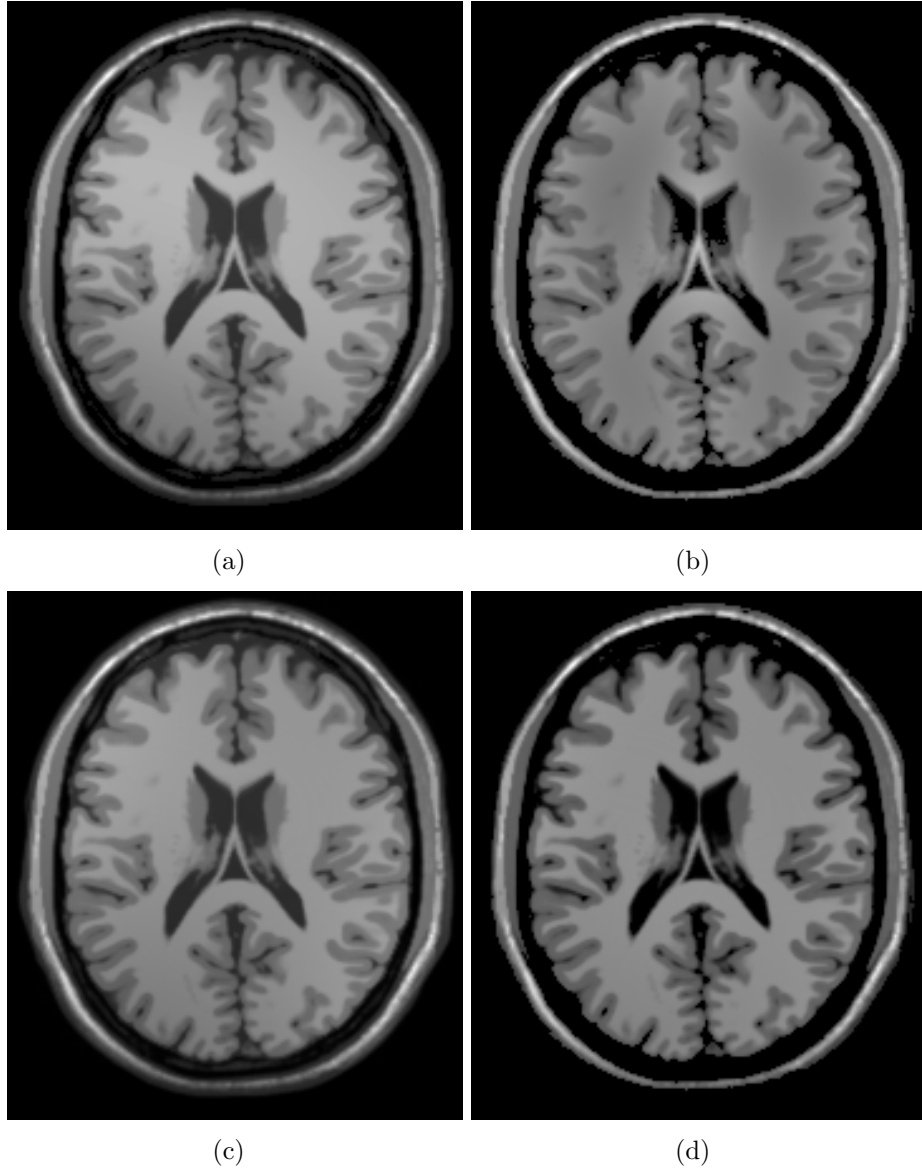


Figure 4.9: Transaxial view of a slice from the brain MR volume. a) Original image from the brainweb database [123]. b) Result after applying the method proposed in [41]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.

proposed method, we computed the percentage of correctly classified pixels as:

$$percent_c = 1 - \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (|B(x, y) - B_R(x, y)|)}{MN}, \quad (4.14)$$

where M and N represents the total number of rows and columns of an image.

Fig. 4.10(a) shows the percentage of correctly classified pixels on 400 test images. Application of Otsu's thresholding on images preprocessed by the proposed method recovered more than 98.5% of the pixels in all 400 test cases. Fig. 4.10(b) shows that the direct application of Otsu's method on average recovers 12% of the pixels only.

Moreover, the sinusoidal pattern $P(x, y)$ (as explained in experimental setup 2) was added to the binary retinal image $B(x, y)$ to produce a global intensity inhomogeneity image $I_n(x, y)$. The test set was increased from 40 to 400 in the same manner as mentioned in experimental setup 1. The evaluation was performed along the same lines as mentioned earlier in the current section. The result of applying Otsu's thresholding after preprocessing by the proposed method is shown in Fig. 4.10(c). The result shows that the proposed method helps to recover 100% pixels in all the 400 test images, while direct application of Otsu's method recovers less than 18% pixels as shown in Fig. 4.10(d). Although, it is trivial to find the edges in our intensity inhomogeneous images, but to find the internal areas of the vessels is a bit tricky. However, the proposed method is able to find the whole structure very efficiently. It is mainly due to the reason that the proposed method is able to exactly locate the low frequency components which are affected by the global intensity inhomogeneity in the image.

4.4.2 Results of Setup 2

In experimental setup 2, a polynomial pattern $P(x, y)$ was added to a synthetic image $S_g(x, y)$ to produce a global intensity inhomogeneity image $I_n(x, y)$. It is worth mentioning that the maximum brightness of $P(x, y)$ was increased to 1500 to make things more challenging and to prove that the proposed method does not produce any artifacts while removing or reducing the global intensity inhomogeneity. One such example of pattern $P(x, y)$ is shown in Fig. 4.11(b). To evaluate the proposed method, we approximated/recovered $S_g(x, y)$ from $I_n(x, y)$. To approximate $S_g(x, y)$, we ran the proposed method on $I_n(x, y)$, which resulted in an image $R(x, y)$. Fig. 4.11(c) shows an example of a global intensity inhomogeneity image

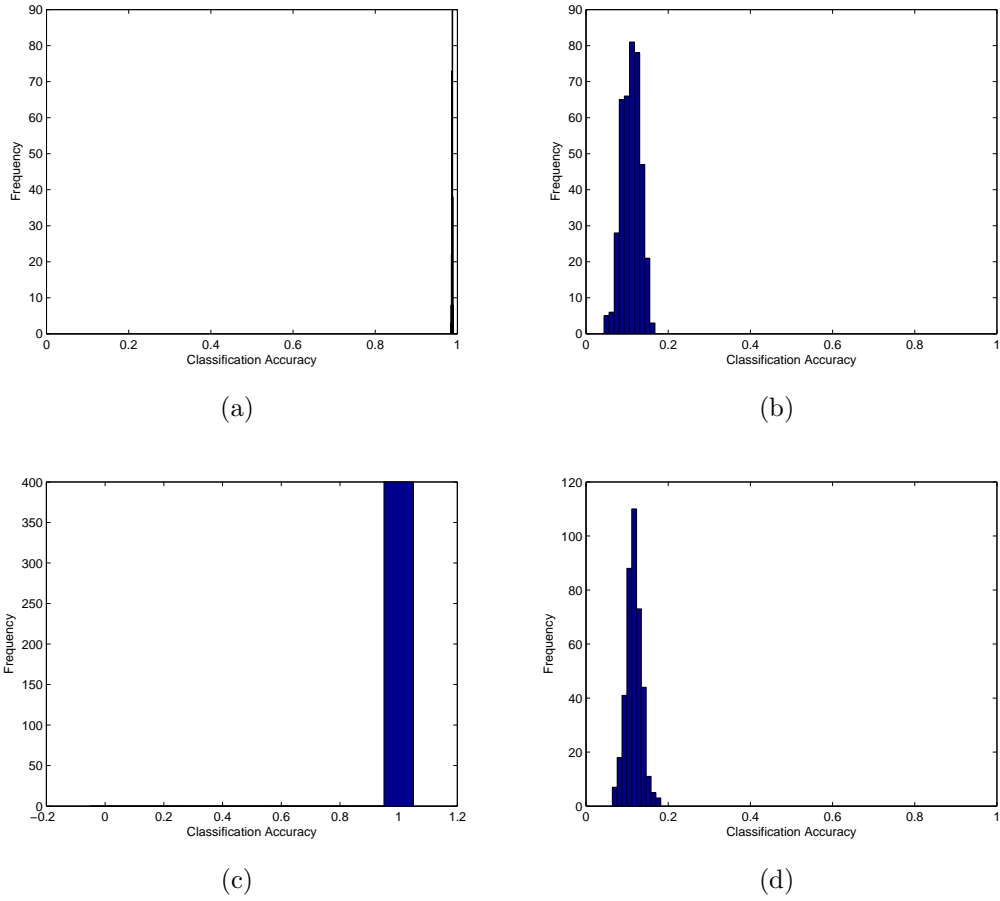


Figure 4.10: Percentage histograms. Here, the horizontal axis represents the percentage of correctly classified white pixels, while the vertical axis represents the number of test images. a) Histogram after application of the proposed method. The histogram shows that the proposed method was able to recover 98.5% of the pixels. b) Results of applying direct Otsu's thresholding on images with intensity inhomogeneity pattern created by standard normal distribution. The histogram shows that one can correctly recover 12% of the pixels. c) Result of the proposed method when a sinusoidal wave was added as an inhomogeneous pattern. Here, our method was able to recover 100% of the pixels. d) Results of applying direct Otsu's thresholding on images with intensity inhomogeneity pattern created by using the sinusoidal wave. On average, one can only recover 18% of the pixels in an image.

$I_n(x, y)$ produced by adding Fig. 4.11(b) with Fig. 4.11(a), while Fig. 4.11(d) shows the approximated image $R(x, y)$ using the proposed method. To quantify the similarity between the approximated image $R(x, y)$ and the ground truth $S_g(x, y)$, we chose cross-correlation as a similarity measure. Fig. 4.12(a) shows the cross-correlation values between $S_g(x, y)$ and

the approximated image $R(x, y)$ for 400 test images. Fig. 4.12(b) shows the result of taking direct cross-correlation between $S_g(x, y)$ and $I_n(x, y)$. The severity of the inhomogeneity is obvious from the horizontal axis of Fig. 4.12(b), where the maximum cross-correlation value is less than 0.1. The proposed method brought significant improvement as the average cross-correlation value improved from 0.06 to 0.84.

Moreover, the sinusoidal pattern $P(x, y)$ was added to the synthetic image $S_g(x, y)$

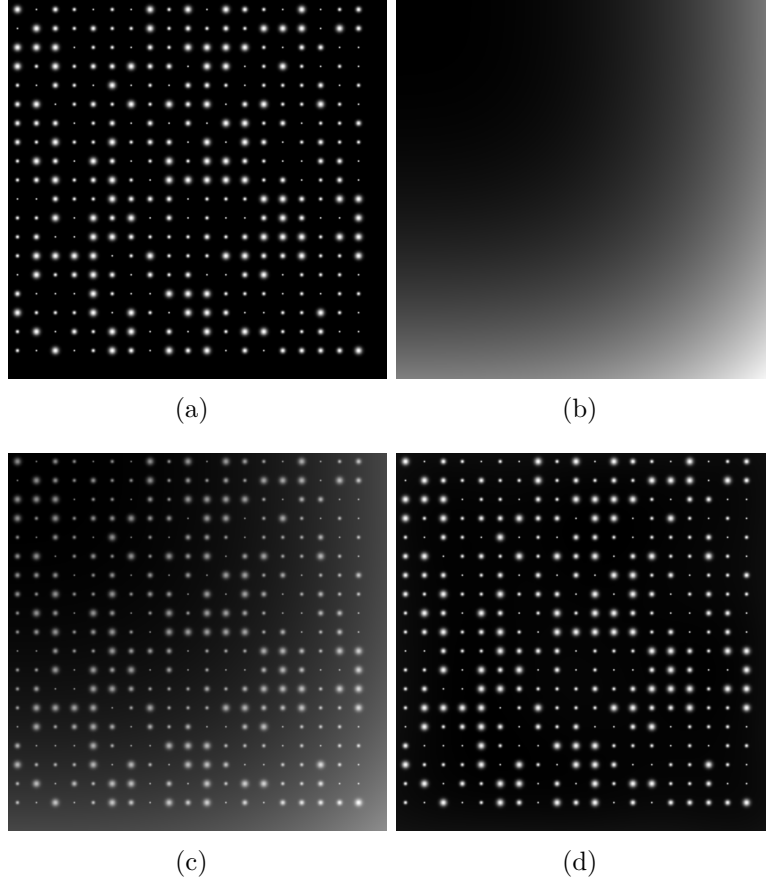


Figure 4.11: a) Original image. b) Global intensity inhomogeneity pattern created by a third degree polynomial. c) Global intensity inhomogeneity image. d) Adjusted image after 5th iteration of the proposed method.

to produce a global intensity inhomogeneity image $I_n(x, y)$. Once again, the maximum brightness of $P(x, y)$ was increased to 800 for the same reasons as mentioned earlier in this section. Fig. 4.13(a) shows an example of a sinusoidal pattern $P(x, y)$, while Fig. 4.13(c)

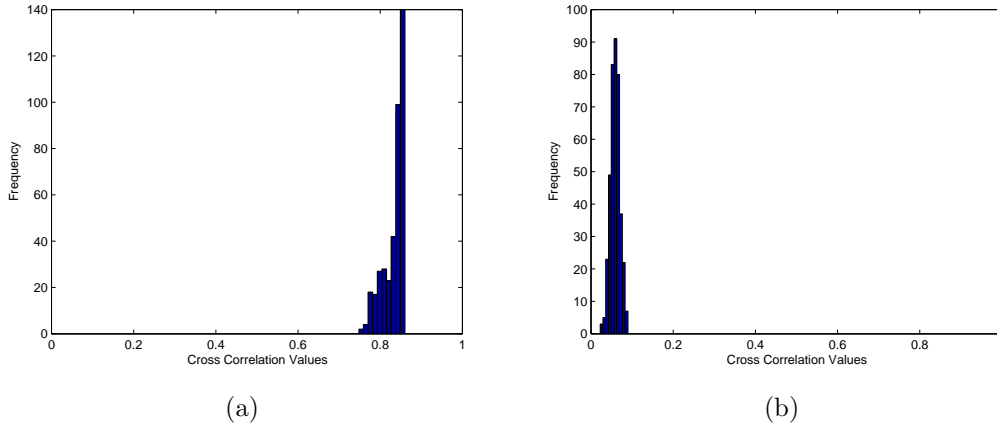


Figure 4.12: Histograms, where the horizontal axis represents the cross-correlation value between two images, while the vertical axis represent the number of test images. a) Cross-correlation of images produced by our method and the ground truth. Here, the images were corrupted by third degree polynomials. The histogram is centered around 0.84. This clearly shows that our results are quite similar to the ground truth. b) Cross-correlation between the ground truth and the intensity inhomogeneity images produced using third degree polynomials. The histogram is centered around 0.06.

shows the approximated image $R(x, y)$ using the proposed method. The evaluation was performed along the same lines as mentioned earlier in the current section. Fig. 4.14(a) shows the cross-correlation values between $S_g(x, y)$ and $R(x, y)$ for 400 test images, while Fig. 4.14(b) shows the result of taking cross-correlation between $S_g(x, y)$ and $I_n(x, y)$. It is evident from the results that the proposed method has improved the average approximation accuracy from 0.15 to 0.96.

4.4.3 Results of Setup 3

In this experimental setup, we have applied the proposed method on the green channel of the original RGB retinal images. In RGB retinal images, the green channel provides the best vessel-background contrast whereas the red channel is the brightest color channel and has low contrast, and the blue channel has poor dynamic range [125]. For these reasons, the green channel is preferred for the vessel extraction in the RGB retinal images [126]. Here, we have compared the performance of our method with the surface fitting method [35]

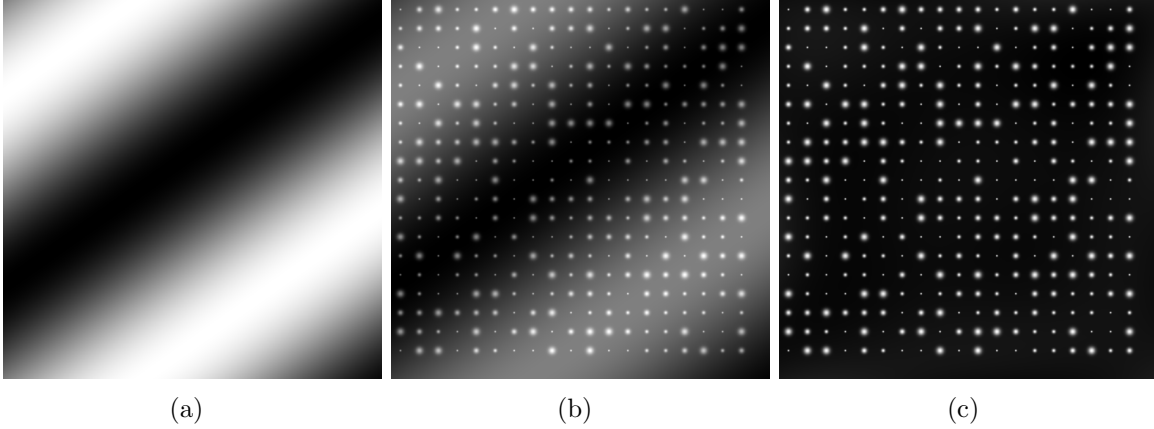


Figure 4.13: a) Global intensity inhomogeneity pattern created by a sinusoidal wave. b) Global intensity inhomogeneity image created by adding 4.13(a) and Fig. 4.5. c) Corrected image at 5th iteration of the proposed method.

and the method based on the sparseness property of the gradient probability distribution function [36]. Fig. 4.15 and Fig. 4.16 shows the result of applying the proposed method on some of the randomly selected images from the DRIVE and STARE database. Here, it is important to mention that we have used normalized convolution to avoid the halo artifacts [42]. It is clear from the results that the proposed method has suppressed the global intensity inhomogeneity with much less artifacts introduced in comparison to the results produced by [35] and [36].

For the quantitative evaluation of the proposed method, we have used (1) detection accuracy, (2) the corresponding true positive rate (TPR), and (3) the false positive rate (FPR) at that accuracy as the performance measures [127]. The detection accuracy is defined as the ratio of the total number of correctly classified pixels to the number of pixels inside the FOV. The TPR is defined as the ratio of the number of correctly classified vessel pixels to the number of total vessel pixels in the ground truth. The FPR is defined as the ratio of the number of non-vessel pixels inside the FOV but classified as vessel pixels, to the number of non-vessel pixels inside FOV in the ground truth. Here, we have used the vessel extraction method proposed in [127] to extract the vessel tree structure from the homogenized images.

In case of DRIVE database, we have applied the proposed method on the 20 test images

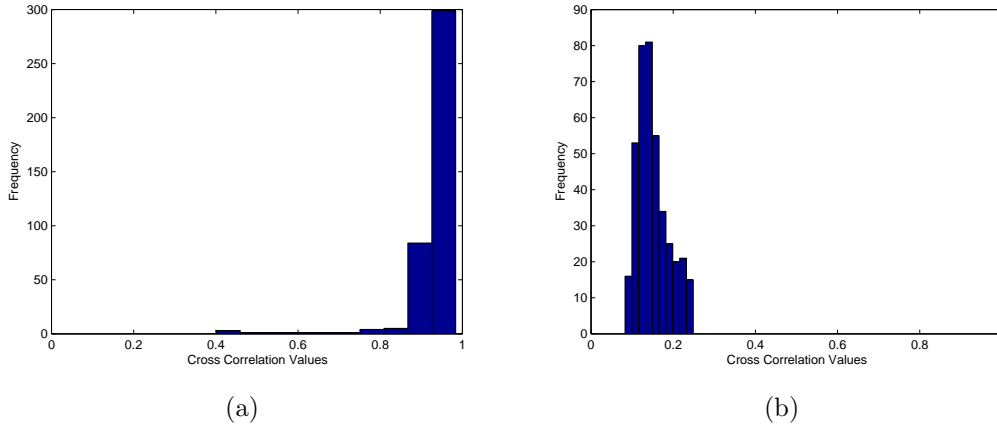

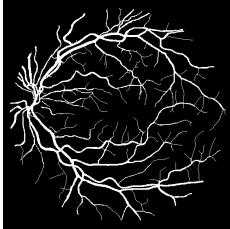
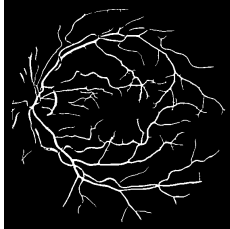
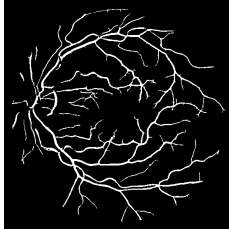
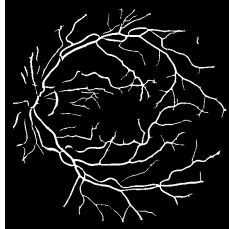

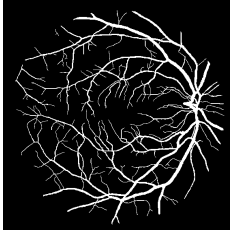
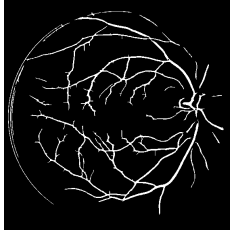
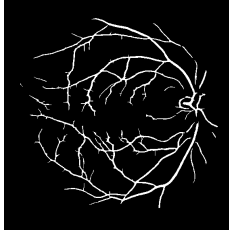
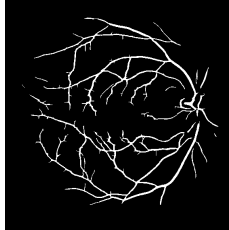

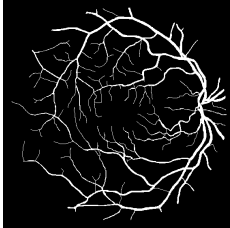
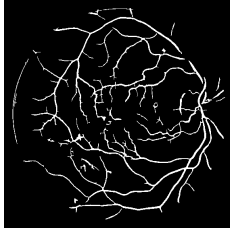
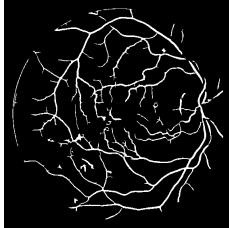
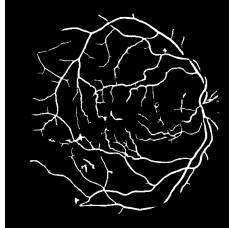

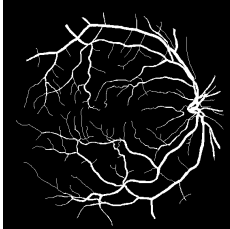
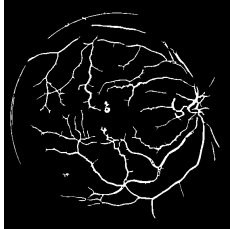
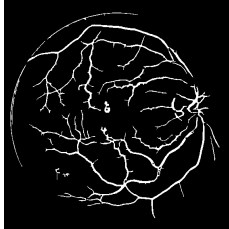
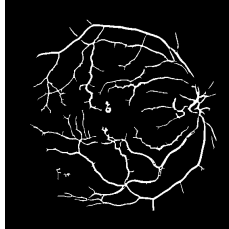

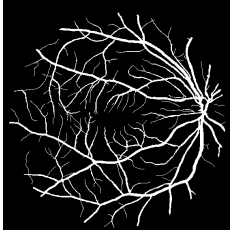
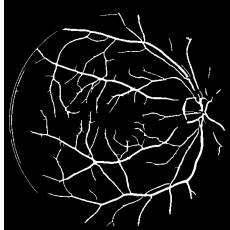
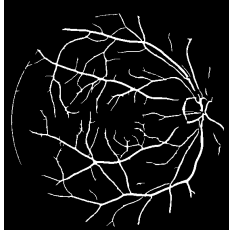
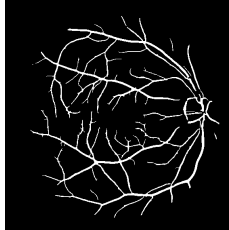


Figure 4.14: Histogram images. a) Cross-correlation of images produced by our method and the ground truth. Here, the images were corrupted by using sinusoids. The histogram is centered around 0.96 which clearly shows the ability of our method to restore the original image. b) Cross-correlation between the ground truth and intensity inhomogeneity images produced by using sinusoids. Here, the mean value is centered around 0.15.

and have used the hand-labelled images by the second expert as the ground truth. Fig. 4.15 shows the result of applying the proposed method on some of the randomly selected images from the DRIVE database in comparison with the results produced by [35] and [36]. Once the global intensity inhomogeneity has been suppressed, we have extracted the retinal vessel tree structure for the quantitative evaluation of the proposed method. Table 4.1 shows the results after applying the vessel extraction method on the homogenised images. Table 4.2 shows the vessel extraction accuracy of the proposed method in comparison with some of the existing methods for global intensity inhomogeneity correction. It is clear from Table 4.2 that the proposed method has outperformed the global intensity inhomogeneity correction methods proposed in [35] and [36]. Furthermore, the proposed method has achieved the highest accuracy rate and TPR with the lowest FPR.

For STARE database, we have used the hand-labelled images by the first expert as the ground truth. Fig. 4.16 shows the result of applying the proposed method on some of the randomly selected images from the STARE database. It is clear from Fig. 4.16 that the proposed method has been able to correct the global intensity inhomogeneity. Table 4.3 shows the vessel extraction accuracy of the proposed method. It is clear from the results

Table 4.1: Segmentation Results for DRIVE Database.

Original	Expert 2	T. Tasdizen [35]	Y. Zheng [36]	Proposed
				
				
				
				
				

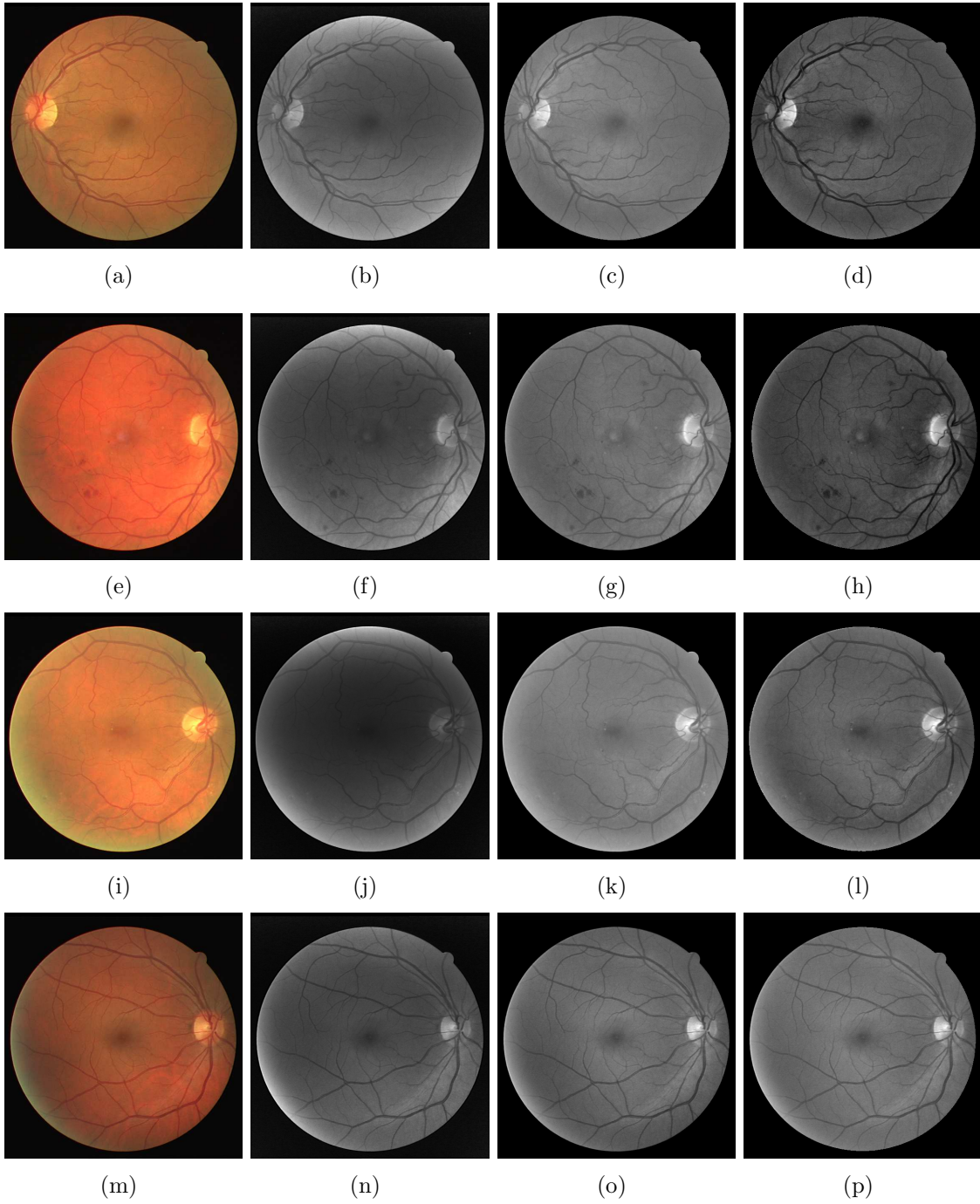


Figure 4.15: a), e), i), m) Original images from DRIVE database. b), f), j), n) Results after applying [35]. c), g), k), o) Results after applying [36]. d), h), l), p) Results after applying proposed method.

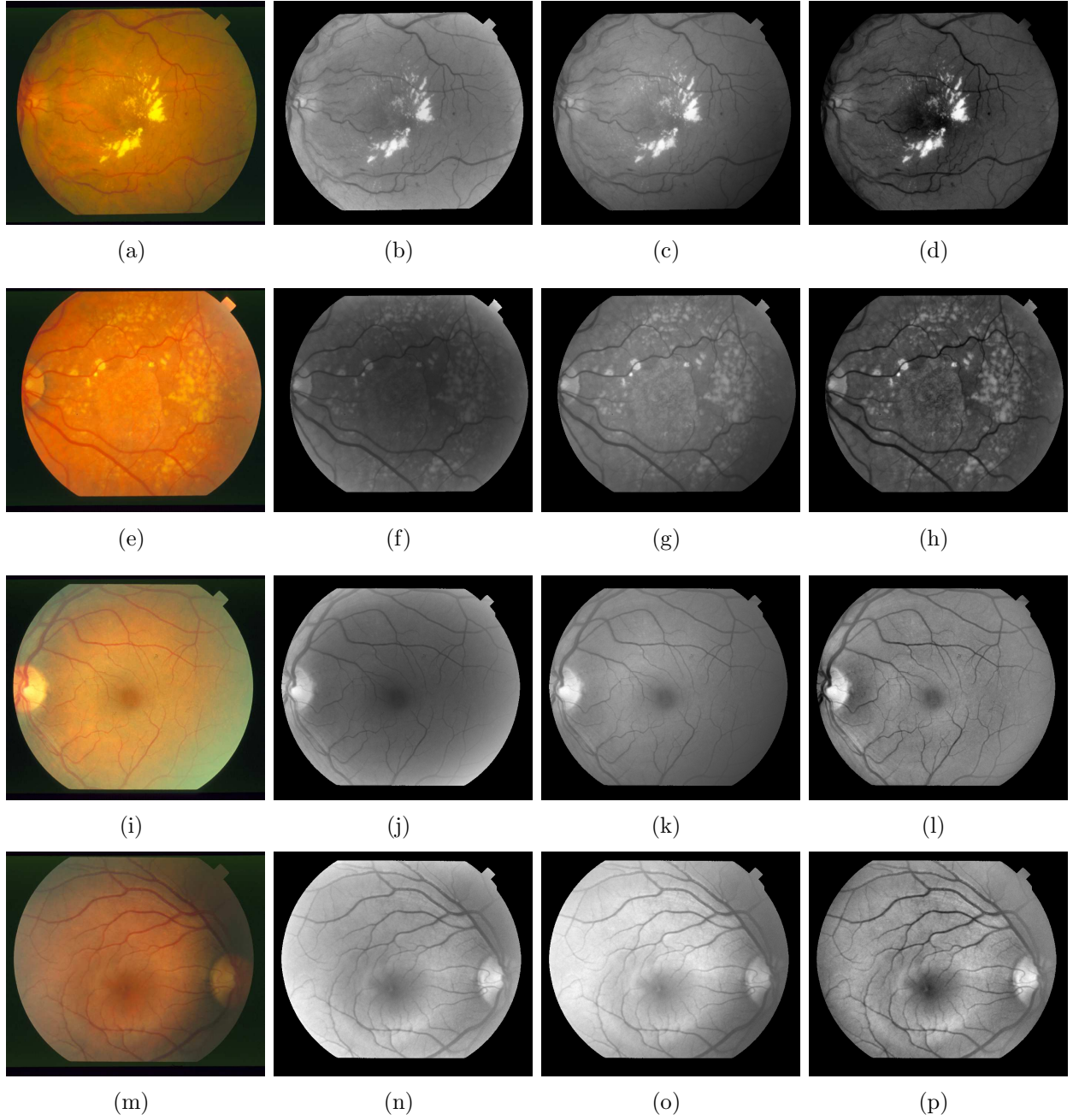


Figure 4.16: a), e), i), m) Original images from STARE database. b), f), j), n) Results after applying [35]. c), g), k), o) Results after applying [36]. d), h), l), p) Results after applying proposed method.

shown in Table 4.3 and Table 4.4 that the proposed methods helps in the suppression of global intensity inhomogeneity while preserving the true vessel structures.

Table 4.2: Vessel extraction results on the DRIVE database.

Methods	TPR	FPR	Accuracy
T. Tasdizen [35]	0.5825	0.0175	0.9337
Y. Zheng [36]	0.5785	0.0169	0.9323
Proposed	0.5922	0.0135	0.9359


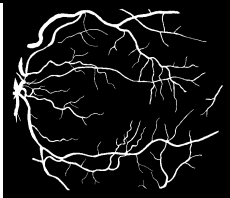
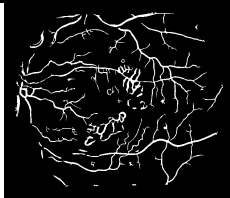
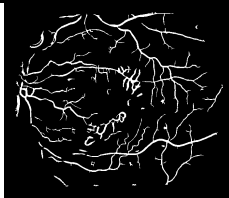
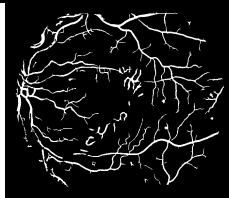

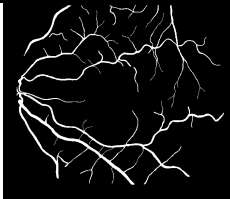

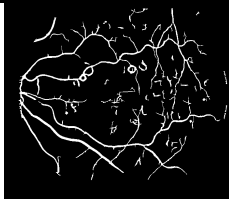
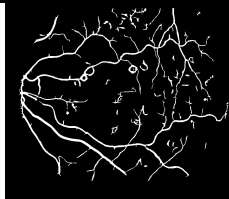

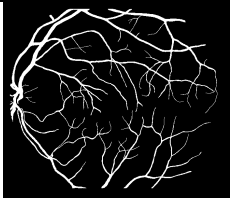
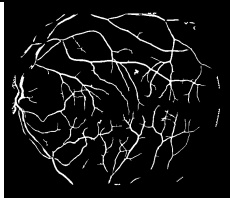
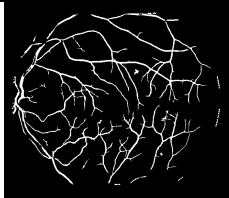
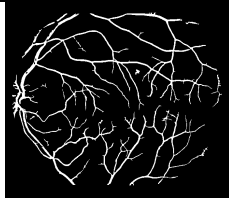

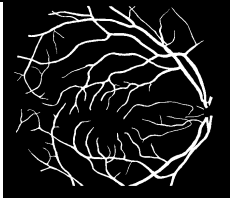
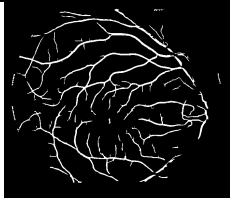
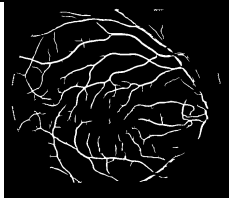
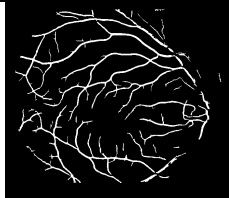
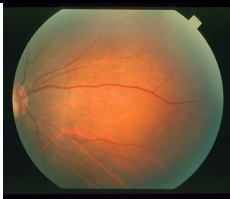
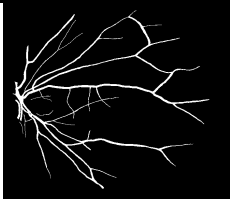


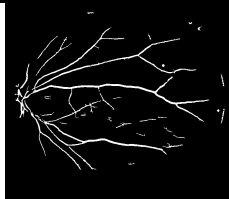
Table 4.3: Vessel extraction results on the STARE database.

Methods	TPR	FPR	Accuracy
T. Tasdizen [35]	0.6939	0.0257	0.9449
Y. Zheng [36]	0.6967	0.0254	0.9454
Proposed	0.7376	0.0265	0.9485

Due to the absence of the ground truth in case of VICAVR database [122], we have followed the subjective evaluation by humans that is considered as the most appropriate criterion for the assessment of the image quality [11]. The subjective evaluations can be done by using absolute rating scale or by means of side-by-side comparisons of the original image and the resulting image. In our subjective evaluation, we have followed the side-by-side scale that ranges from -3 to 3 (-3 =Much Worse, -2 =Worse, -1 =Slightly Worse, 0 =Same, 1 =Slightly Better, 2 =Better, 3 =Much Better). All of the results were displayed on the monitor screen and the students were asked to do the evaluation. In total 18 students participated in this evaluation process who marked the images according to their visual judgement. To make the evaluation process fair, the results were shown to the students without mentioning the results of the proposed method and the results obtained by applying other methods.

For the evaluation of VICAVR database, we compared our results with the methods proposed in [35] and [36]. Fig. 4.17 and Fig. 4.18 shows the result of applying the proposed method on some of the randomly selected images from VICAVR database [122]. The evaluation in Fig. 4.19 shows that the proposed method works better than the methods developed in [35] and [36].

Table 4.4: Segmentation Results for STARE Database.

Original	Expert 1	T. Tasdizen [35]	Y. Zheng [36]	Proposed
				
				
				
				
				

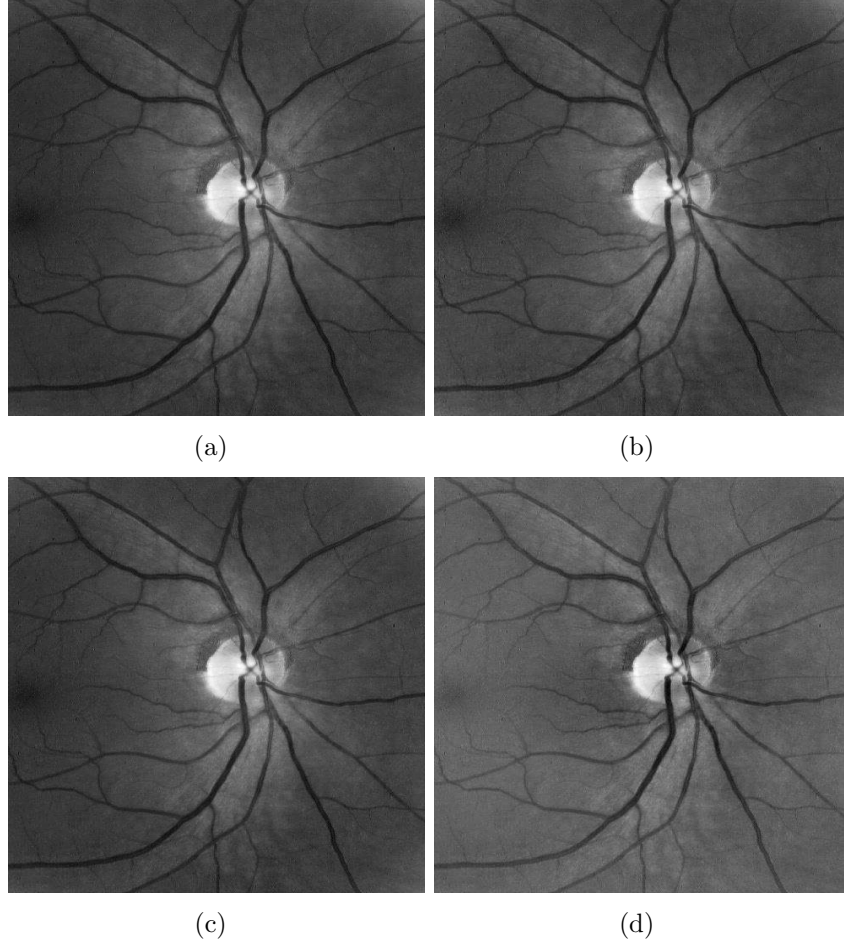


Figure 4.17: a) Original image from the VICAVR database [122]. b) Result after applying the method proposed in [35]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.

4.4.4 Results of Setup 4-5

For the evaluation of the experimental setups 4 and 5, we have followed the same subjective evaluation by humans as used in experimental setup 3. In this subjective evaluation, 18 students participated to mark the images according to their visual judgement. On average, each student evaluated 80 images in four different sessions. Fig. 4.20 shows the histograms of the evaluation for the databases used in this work for experimental setup 4 and 5.

For TEM images, we compared our results with the methods developed in [11, 35, 36], as shown in Fig. 4.21. We opted to compare our results with these recent methods as they

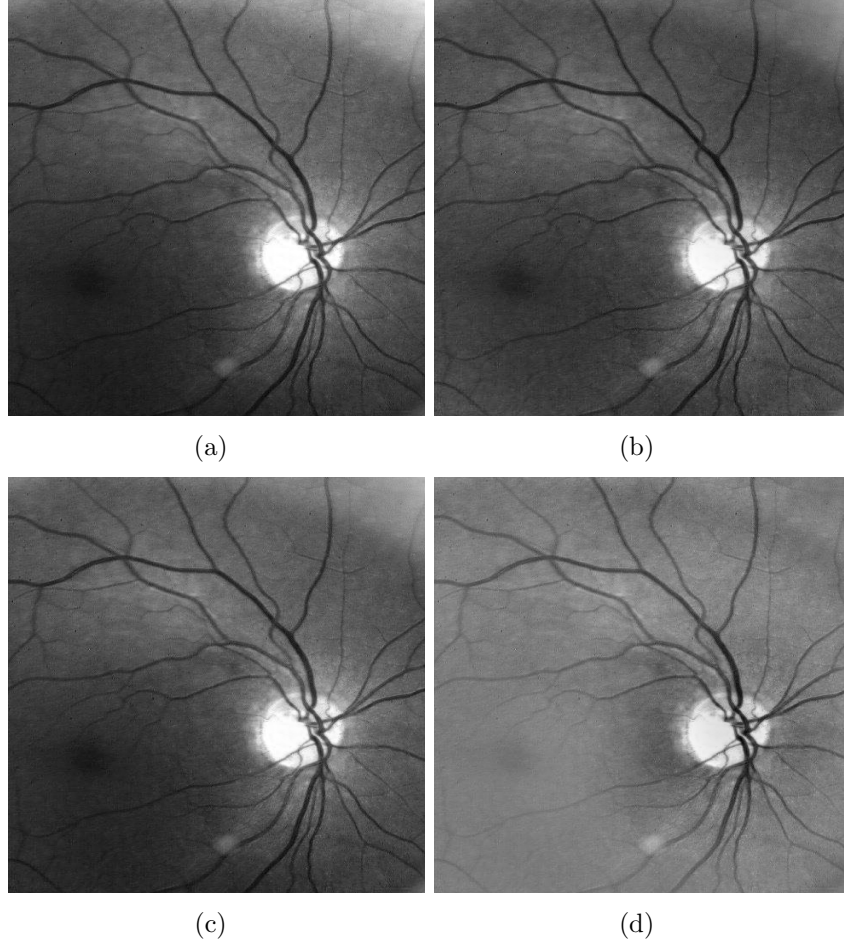


Figure 4.18: a) Original image from the VICAVR database [122]. b) Result after applying the method proposed in [35]. c) Result after applying the method proposed in [36]. d) Result after applying the proposed method.

are reported [35, 36] to outperform the traditional methods. The evaluation in Fig. 4.20(a) shows that the proposed method works better than the methods developed in [11, 35, 36].

We followed the same lines to evaluate experimental setup 5, as we did in experimental setup 4. Here, we have compared our results with methods developed in [36, 41]. It is also important to mention that we have used normalized convolution as proposed in [42], to avoid the halo artifacts. Fig. 4.9 presents a comparison of the proposed method with methods developed in [36, 41]. The evaluation in Fig. 4.20(b) also depicts the effectiveness of the proposed method.

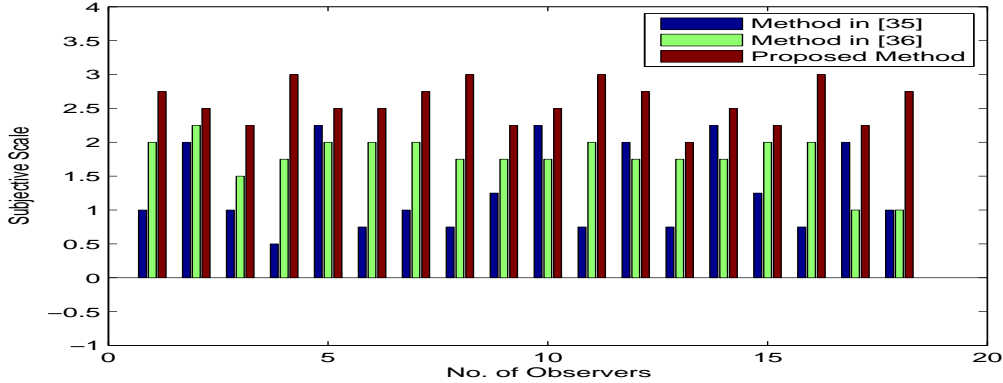
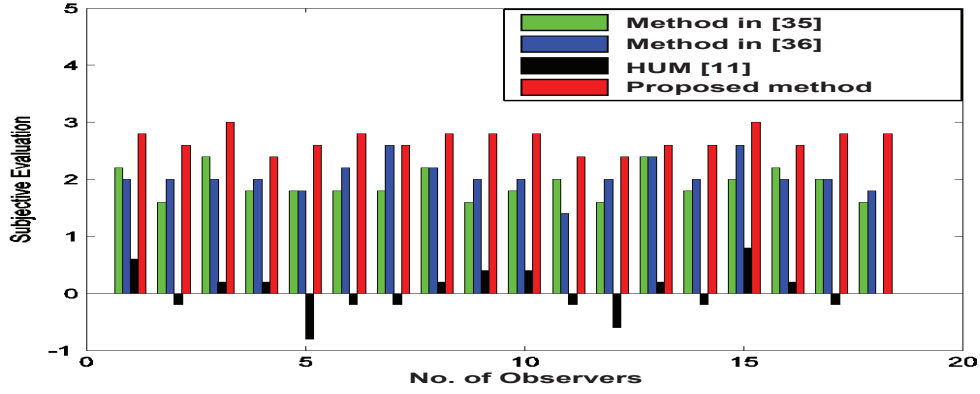


Figure 4.19: Average subjective evaluation of the retinal images after applying the proposed method.

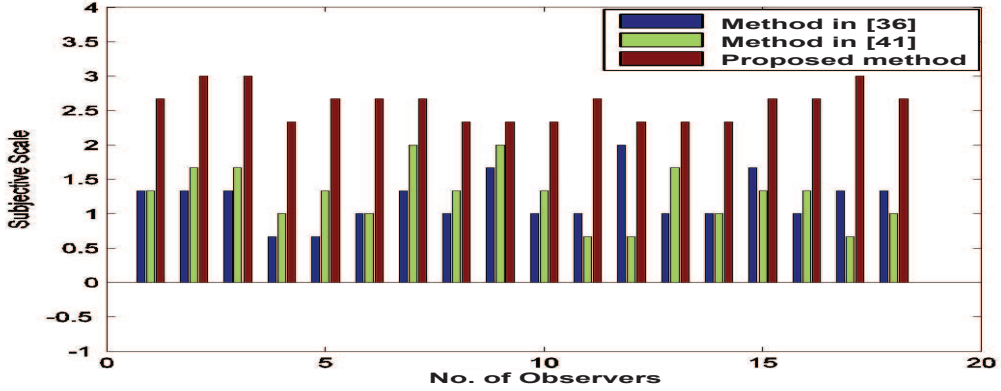
Along with the subjective evaluations of experimental setup 5, we also report the quantitative evaluations. As mentioned in experimental setup 5, the BSB database provides the ground truth which can be used for evaluation purposes. It provides two types of ground truth, either in the form of labelled volume, where each label represents a tissue class (0=Background, 1=CSF, 2=Grey Matter, 3=White Matter, 4=Fat, 5=Muscle, 6=Muscle/Skin, 7=Skull, 8=Vessels, 9=Around Fat, 10=Dura Matter, and 11=Bone Marrow) or in the form of fuzzy volumes, where each fuzzy volume represents the contribution of each tissue class. To keep the evaluation process simple, we opted to measure the quantity of the white matter and the grey matter tissues among the 12 different tissue types. For this purpose, we used the fuzzy volumes to construct a binary volume (V^g) by assigning all the voxels to 1, where the grey matter membership value is higher than all other tissue types. On the same lines, we constructed another binary volume (V^w) with value 1 at locations where the white matter membership value is greater than all other tissue types.

To quantify the accuracy of the proposed method, we segmented the output of our method using fuzzy c-means (FCM) clustering [128] into three clusters which resulted in three fuzzy volumes, i.e., the grey matter (V^{gf}), the white matter (V^{wf}) and the background (V^{bf}). It is worth mentioning that the output of the proposed method was processed by equation:

$$adjusted^p = adjusted \times (V^g + V^w), \quad (4.15)$$



(a)

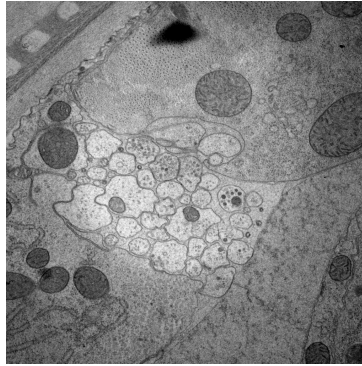


(b)

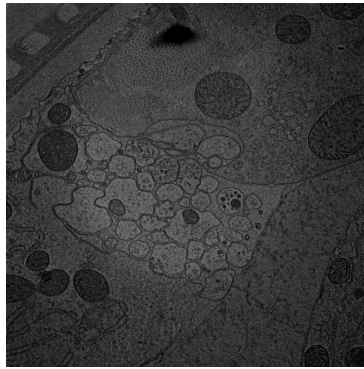
Figure 4.20: a) Average subjective evaluation of TEM images. b) Average subjective evaluation of the brain MR images.

where *adjusted* represents the volume produced by applying the proposed method slice by slice. This processing ensures that we are only processing the grey matter and the white matter tissues. So, essentially our goal is to segment the grey matter and the white matter tissues as accurately as possible. The final segmentation can be achieved by using the following equations:

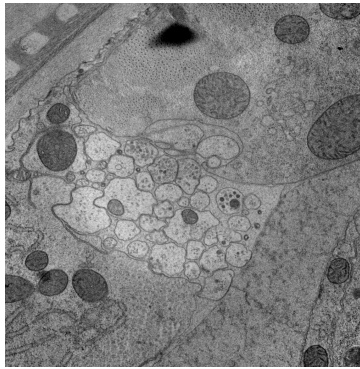
$$adjusted^{pg} = \begin{cases} 1 & \text{if } V^{gf} > \max(V^{wf}, V^{bf}) \\ 0 & \text{otherwise,} \end{cases} \quad (4.16)$$



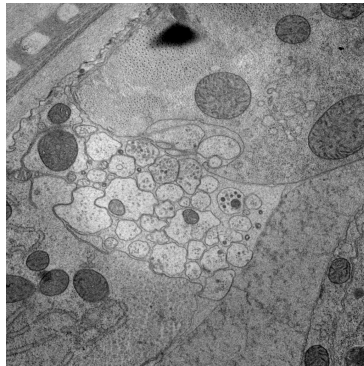
(a)



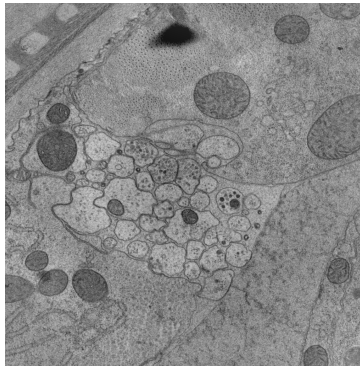
(b)



(c)



(d)



(e)

Figure 4.21: a) Original image suffering from global intensity inhomogeneity. b) Result after applying HUM [11]. c) Result after applying the method proposed in [35]. d) Result after applying the method mentioned in [36]. e) Result produced by the proposed method.

$$adjusted^{pw} = \begin{cases} 1 & \text{if } V^{wf} > \max(V^{gf}, V^{bf}) \\ 0 & \text{otherwise.} \end{cases} \quad (4.17)$$

To quantify the recovery of grey matter tissue, we computed the cross-correlation slice by slice between $adjusted^{pg}$ and V^g volumes as:

$$corr_s^g = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (adjusted_s^{pg}(x, y) - \overline{adjusted_s^{pg}})(V_s^g(x, y) - \overline{V_s^g})}{\sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (adjusted_s^{pg}(x, y) - \overline{adjusted_s^{pg}})^2 \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (V_s^g(x, y) - \overline{V_s^g})^2}}, \quad (4.18)$$

where $\overline{adjusted_s^{pg}}$ and $\overline{V_s^g}$ represents the mean of the s^{th} slice in $adjusted^{pg}$ and V^g , respectively. The white matter tissue accuracy ($corr_s^w$) is computed by taking the cross-correlation between $adjusted^{pw}$ and V^w , slice by slice. Table 4.5 shows the mean and the standard deviation of $corr_s^g$ and $corr_s^w$ for 36 test volumes. It also reports the results of applying the same evaluation procedure on the images processed by some of the recent methods [36, 41]. As, it is evident from Table 4.5 that the proposed method outperforms the other methods for all the test volumes. It is also evident from the results that [36] and [41] segments the white matter more accurately as compared to the grey matter tissue. However, the proposed method segments the white matter tissue as accurately as the grey matter tissue.

Furthermore, we have compared the results in terms of coefficient of contrast (CC) [41]

Table 4.5: White matter and grey matter restoration accuracy of various inhomogeneity levels (IL) in terms of cross-correlation coefficient

Tissue Type		White Matter			Grey Matter		
Methods		Y. Zheng [36]	Ardizzone [41]	Proposed	Y. Zheng [36]	Ardizzone [41]	Proposed
# of Vols.	IL[%]	Mean(Std.)	Mean(Std.)	Mean(Std.)	Mean(Std.)	Mean(Std.)	Mean(Std.)
6	40	0.8599(0.0024)	0.8896(0.0033)	0.9317(0.0173)	0.6718(0.0042)	0.7800(0.0006)	0.9433(0.0073)
6	50	0.8504(0.0025)	0.8823(0.0023)	0.9322(0.0171)	0.6531(0.0045)	0.7757(0.0008)	0.9437(0.0071)
6	60	0.8420(0.0026)	0.8720(0.0022)	0.9323(0.0168)	0.6381(0.0054)	0.7735(0.0026)	0.9436(0.0068)
6	70	0.8328(0.0029)	0.8640(0.0028)	0.9325(0.0167)	0.6199(0.0092)	0.7694(0.0028)	0.9434(0.0066)
6	80	0.8208(0.0045)	0.8553(0.0047)	0.9327(0.0164)	0.5935(0.0111)	0.7568(0.0055)	0.9433(0.0064)
6	90	0.7512(0.0924)	0.8058(0.0263)	0.9327(0.0164)	0.5152(0.0687)	0.7122(0.0298)	0.9433(0.0064)

and coefficient of joint variation (CJV) [129]. Here, CC and CJV are defined as:

$$CC = \frac{\mu(WM)}{\mu(GM)}, \quad (4.19)$$

and

$$CJV = \frac{\sigma(WM) + \sigma(GM)}{|\mu(WM) - \mu(GM)|}, \quad (4.20)$$

respectively. Here, μ and σ represents the mean and the standard deviation. WM stands for the white matter tissue and GM represents the grey matter tissue. CC measures the contrast between WM and GM tissues. So, higher the ratio, the better it is. On the other hand, CJV measures the overlap between WM and GM tissue distributions. So, smaller the value, the less overlap it has between the distributions. It is worth mentioning that all the methods were applied slice by slice to construct the corrected volumes. Table 4.6 shows average CC and CJV values for each inhomogeneity level. It is evident from Table 4.6, that the proposed method outperforms the methods developed in [36, 38, 41], both in terms of CC and CJV . It is interesting to mention that the value of CJV decreases with increasing intensity inhomogeneity for the proposed method. This is due to the fact that the increase in inhomogeneity will increase the magnitude of the affected frequencies substantially, and as a consequence, it will become trivial for GWDT to localize such frequencies.

Table 4.6: 2D - White and grey matter restoration accuracy in terms of CC and CJV

# of Vols.	IL[%]	Y. Zheng [36]		Ardizzone [41]		N3 [38]		Proposed	
		CC	CJV	CC	CJV	CC	CJV	CC	CJV
6	0	1.305276	0.483392	1.162004	0.013417	1.3269	0.008374	1.642670	0.006085
6	20	1.313817	0.449433	1.186173	0.012347	1.3263	0.008511	1.618956	0.005996
6	40	1.323322	0.452676	1.237673	0.011331	1.3269	0.008566	1.596365	0.005942
6	50	1.328033	0.460662	1.267643	0.010623	1.3275	0.008596	1.585846	0.005912
6	60	1.332611	0.467540	1.285935	0.010130	1.3282	0.008611	1.575819	0.005892
6	70	1.337113	0.471941	1.302322	0.009738	1.3292	0.008617	1.565964	0.005875
6	80	1.341433	0.475713	1.310956	0.009445	1.3302	0.008612	1.557453	0.005862
6	90	1.345732	0.479162	1.318917	0.009206	1.3315	0.008592	1.548814	0.005851

4.5 Global Intensity Inhomogeneity Correction in Volume Images

In volume images, slice by slice global intensity inhomogeneity correction often results in inter-slice variations, which may be problematic for subsequent image analysis [41]. To overcome this problem, the method for global intensity inhomogeneity correction proposed in Section 4.2 is extended to 3D for the volume images, especially for the brain MR volumes.

The proposed method is based on the 3D GWDT of the magnitude spectrum of MR volume. The 3D GWDT of the magnitude spectrum helps to find the frequencies that are affected by the global intensity inhomogeneity. These affected frequencies are localized by defining a ball whose radius changes according to the GWDT. It is important to mention that filtering is performed by normalized convolution [130] to avoid the halo artifacts [42]. The proposed method for the 3D global intensity inhomogeneity correction is summarized below.

Algorithm 4 Global Intensity Inhomogeneity Correction in Volume Images

1: $f(x, y, z) \leftarrow \text{input image}$

2:

$$F(u, v, w) \leftarrow \frac{1}{MNO} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \sum_{z=0}^{O-1} f(x, y, z) e^{-2\pi j(\frac{ux}{M} + \frac{vy}{N} + \frac{wz}{O})},$$

3: $F(u, v, w) \leftarrow \text{fftshift}(F(u, v, w)),$

4: $j = 0;$ // initialization

5: $\text{similarity}[j] = 1;$ // initialization.

6: **repeat**

7: $j = j + 1;$

8: $F_{\text{mag}}(u, v, w) \leftarrow |F(u, v, w)|,$

9:

$$d(p, q) = \begin{cases} \min_{\pi \in \mathcal{P}_{pq}} \mathcal{L}(\pi) & \text{if } p \neq q \\ 0 & \text{if } p = q. \end{cases}$$

where $\mathcal{L}(\pi)$ is computed as:

$$\mathcal{L}(\pi) = \sum_{i=0}^{n-1} | F_{mag}(p_i) - F_{mag}(p_{i+1}) | \cdot \| p_i - p_{i+1} \|,$$

where $F_{mag}(p_i)$ is a grey-level value at voxel p_i . where p_i and p_{i+1} are adjacent voxels.

10: **for** $r = 1, \dots, n$ **do**

11:

$$B(s_0, r) = \{q \mid d(s_0, q) \leq r\},$$

Then the set of low frequencies within radius r is defined as:

$$K(s_0, r) = \{M(q) \mid q \in B(s_0, r)\}.$$

where $M : Dist_{map} \rightarrow F_{mag}$, and $s_0 \in Dist_{map}$ is the location of the seed.

12: It is followed by computing the percentage of the power spectrum ($P_R(r)$):

$$P_R(r) = \frac{\Omega(K(s_0, r))}{\Omega(F)} \times 100,$$

where the total image power is computed by:

$$\Omega(F) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \sum_{w=0}^{O-1} | F(u, v, w) |^2.$$

13: **end for**

14: Select a threshold value (T);

15:

$$F_{low} = \begin{cases} 1 & \text{if } F_{mag} \in K(s_0, T) \\ 0 & \text{if } F_{mag} \in \overline{K(s_0, T)}. \end{cases}$$

16: Smooth F_{low} with a Gaussian of $\sigma = 1$.

17: $High_Pass \leftarrow \max(F_{low}) - F_{low}$;

18: $ADJUSTED \leftarrow High_Pass \times F(u, v, w)$;

```

19:  adjusted  $\leftarrow$  real(ifftn(ifftshift(ADJUSTED)));
20:  similarity[j] = correlation(adjusted,f);
21: until  $0 \leq (\textit{similarity}[j - 1] - \textit{similarity}[j]) < \epsilon$ 

```

4.5.1 Results and Discussion

We have opted to evaluate the proposed method on 3D MR volumes obtained from the BrainWeb Simulated Brain Database [123]. We have used the same database as used in Section 4.3.5. Here, we opted to compare our results with some of the recent methods proposed in [36, 38, 41]. Fig. 4.22 shows the frontal view of a slice of the brain MR volume after applying the proposed method. Table 4.7 shows the average *CC* and *CJV* values computed over the database used in experimental setup 5. In Table 4.7, results are grouped according to the inhomogeneity levels, as we are interested in observing the behaviour of each method under different inhomogeneity levels. It is evident from the results that the proposed 3D global intensity inhomogeneity correction method outperforms the 3D version of the methods proposed in [36, 38, 41], both in terms of *CC* and *CJV*. The performance of all methods goes down with increasing inhomogeneity. This is mainly because the spectrum of global inhomogeneity and the spectrum of the volume starts to overlap.

Table 4.7: 3D - White and grey matter restoration accuracy in terms of CC and CJV

# of Vols.	IL[%]	Y. Zheng [36]		Ardizzone [41]		N3 [38]		Proposed	
		CC	CJV	CC	CJV	CC	CJV	CC	CJV
6	0	1.2865	0.195845	1.170462	0.012497	1.268995	0.006750	1.586980	0.0000283
6	20	1.3054	0.259103	1.179238	0.012237	1.267653	0.006692	1.611466	0.0000316
6	40	1.3243	0.370014	1.151662	0.013959	1.271201	0.006851	1.636120	0.0000367
6	50	1.3336	0.424884	1.255659	0.010895	1.271684	0.007098	1.648297	0.0000396
6	60	1.3427	0.460994	1.191073	0.013999	1.270182	0.006828	1.660282	0.0000424
6	70	1.3517	0.487349	1.217886	0.031443	1.272087	0.007061	1.672088	0.0000450
6	80	1.3604	0.516353	1.218399	0.047353	1.274926	0.007294	1.683705	0.0000475
6	90	1.3702	0.516594	1.173487	0.067265	1.274868	0.007845	1.696117	0.0000486

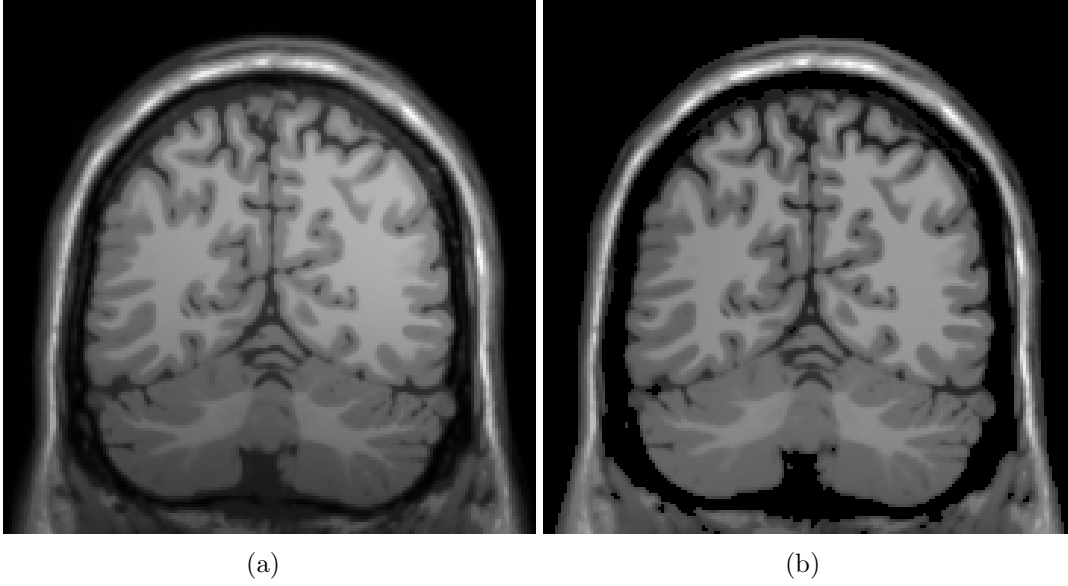


Figure 4.22: a) A frontal view of a slice from a 3D MRI volume corrupted with 90% global intensity inhomogeneity. b) The same slice of the volume after applying the proposed method. Significant improvement is visible in the lower part of slice.

4.6 Summary

The success of the method is based on the observation that if global intensity inhomogeneity is visible in an image, then the frequency components affected by such inhomogeneity should have a higher magnitude than its neighbouring components. A filter is automatically generated based on the frequency content of the degraded image. The results reported here clearly shows that the natural look of the image has remained intact to much of the extent after the global intensity inhomogeneity correction, which is often lost in classical filtering methods. The method is general in nature as we have tested it on five different types of images. It does not seem to require any prior knowledge of the imaging modality. It only assumes that the input image is suffering from global intensity inhomogeneity. Moreover, the proposed method is not user dependent as all of the filter parameters are automatically generated based on the GWDT of the input image. In case of homogeneous image, the method will not induce any inhomogeneity or any artifacts, however, it will decrease the overall energy of the image. It works fine as long as the intensity inhomogeneity is global in

nature otherwise it fails. The proposed method can also help to improve the segmentation accuracy if used in conjunction with state-of-the-art segmentation methods. In future, we are planning to extend this method to other image modalities and use a more sophisticated method to convert our ideal filter to a smoothed one.

Chapter 5

Local Intensity Inhomogeneity Correction

5.1 Introduction

Global intensity inhomogeneity is considered as a slowly varying component. However, this assumption does not hold in case of images having shadow artifacts. These artifacts appear as sharp discontinuities and are only visible at shadow boundaries. The presence of the shadow artifacts shifts the paradigm from global to local intensity inhomogeneity correction. The literature shows that the presence of local intensity inhomogeneity affects the performance of image matching [46,47]. Especially, in face recognition, the suppression of local intensity inhomogeneity is considered to be one of the major challenges [48]. It is revealed in the face recognition vendor tests [50,51] that the performance of the face recognition system is highly sensitive to local intensity inhomogeneity and shadow artifacts. These artifacts are often corrected at different scales and orientations [52]. However, the selection of scale and orientation is a problem in itself.

Generally, it is believed that the facial features invariant to expressions and local intensity inhomogeneity are most discriminatory and helpful in face recognition [52]. Facial features include eyes, irises, lips, nasal cavities and eyebrows [131]. In frontal pose images, most of the facial features do not change their direction during different facial expressions. For

instance, the orientation of an eye remains horizontal ($-\frac{\pi}{4}$ to $\frac{\pi}{4}$) during different facial expressions. However, cheek wrinkles (due to facial expressions) are vertical ($\frac{\pi}{4}$ to $\frac{3\pi}{4}$) in nature as shown in Fig. 5.1. So, one can interpret that the vertical features are relatively less invariant to facial expressions as compared to horizontal. Moreover, the topographic structure of facial features often gives rise to vertical shadows. This suggests that the



Figure 5.1: Original image.

horizontal features are least effected by the local intensity inhomogeneity. Fig. 5.1 shows some of the vertical and horizontal features, where a dotted line is superimposed on some of the vertical features and an ellipse is superimposed on the horizontal features. In this study, we present a comparison between horizontal and vertical facial features in the presence of local intensity inhomogeneity. We have empirically shown that in presence of local intensity inhomogeneity, most of the discriminatory power lies within the horizontal features. The details of the proposed method are given in the section to follow.

5.2 Proposed Method

The proposed method is based on the systematic approach where gamma transformation is used for image normalization and DDFB is used to suppress the local intensity inhomogeneity. Mathematically, the gamma transformation can be written as:

$$\begin{aligned} I'_{xy} &= G(I_{xy}; \gamma^*) \\ G(I_{xy}; \gamma) &= I_{xy}^{\frac{1}{\gamma}} \end{aligned} \tag{5.1}$$

where I'_{xy} is the grey-level of an image I_{xy} at location (x, y) . Here, the optimal value of gamma is represented by γ^* and $\gamma \in \mathbb{Z}^+$. In practice, gamma transformation is preferred over log transformation, as log transformation often turns out to be too strong and tends to over-amplify the noise in dark regions of an image [132]. A gamma transformation with $\gamma \geq 2$ is often considered as a good compromise [133]. Here, we have used $\gamma^* = 5$ as suggested in [133]. Fig. 5.2 shows the result after gamma transformation with $\gamma^* = 5$.

5.2.1 Decimation-Free Directional Filter Banks (DDFB)

Generally, global intensity inhomogeneity is modelled by low frequency components. However, this assumption does not hold in case of local intensity inhomogeneity as it gives rise to shadows artifacts. These shadow artifacts are often considered as high frequency components and they appear as edges after any global normalization method. The gamma transformation in itself is not an exception as it also results in edges at shadow boundaries as shown in Fig. 5.2. Depending on the severity of the local intensity inhomogeneity, these edges at the shadow boundaries can appear at different scales. Gabor filters provide a mechanism to deal with edges occurring at different scales and orientation [134]. But to choose a specific scale and orientation is an open question up till now. Concluding from the discussion in Section 5.1, we can construct a partial solution to the orientation problem by using the prior that the shadow boundaries are mainly vertical in nature. The ambiguity in scale selection leaves us with no other option except to use all scales at the same time.

To suppress the edges produced at the shadow boundaries, DDFB appears to be an appropriate candidate as it provides a mechanism to design orientation selective filters irrespective of the scale. These filters split the Fourier transform of an image into wedge shaped passbands. In the spatial domain, these passbands correspond to all features in a specific orientation irrespective of the scale. Readers interested in further details about the topic are advised to consult [87, 135]. As, we are interested in suppressing the vertical shadows, we have designed a filter pair using DDFB, which decomposes a face features into its horizontal and vertical components.

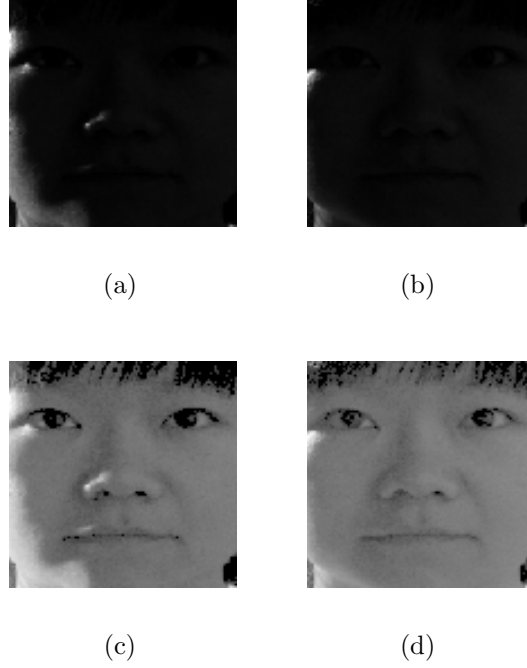


Figure 5.2: a), b) Original images of the same subject. c), d) Images having shadow artifacts after gamma transformation with $\gamma^* = 5$.

To make the discussion self-contained, we have enlisted the steps which are necessary to construct the desired filter pair.

1. Construct a 1D linear phase FIR low-pass filter \mathbf{u} , with a cut-off frequency of 0.47π to avoid aliasing.
2. Construct a 2D filter by taking the tensor product of the 1D filter \mathbf{u} , with itself. For instance, a 1-dimensional filter $\mathbf{u} = (u_1, u_2, \dots, u_m)$ having m coefficients, after the tensor product can be expressed as:

$$X = \mathbf{u} \otimes \mathbf{u} = \begin{bmatrix} u_1u_1 & u_1u_2 & \dots & u_1u_m \\ u_2u_1 & u_2u_2 & \dots & u_2u_m \\ \vdots & \vdots & \ddots & \vdots \\ u_mu_1 & u_mu_2 & \dots & u_mu_m \end{bmatrix}$$

where the \otimes symbol defines the tensor product. Fig. 5.3(a) shows the Fourier transform

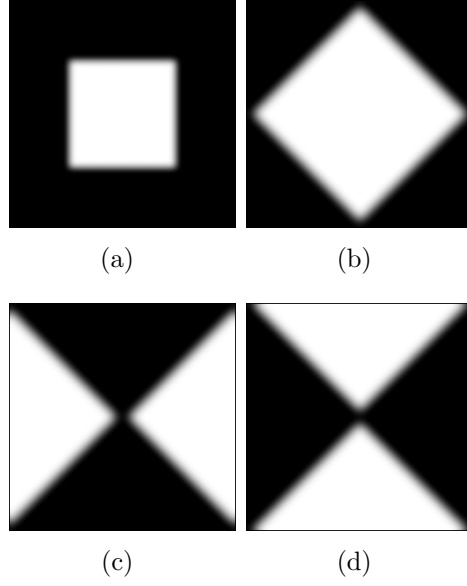


Figure 5.3: a) Frequency response of 2D low-pass filter. b) Frequency response of 2D filter after applying Quincunx downsampling matrix. c) Frequency response of $v(x,y)$. d) Frequency response of $h(x,y)$.

of the 2D filter X constructed by using the tensor product.

3. Now, perform the directional downsampling on the 2D filter X by using the Quincunx downsampling matrix M , where M is:

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Mathematically, the whole operation in the frequency domain can be represented as:

$$X_w(\omega) = \frac{1}{|\det(M)|} \sum_{k \in N(M^T)} X(M^{-T}(\omega - 2\pi k)) \quad (5.2)$$

where $N(M^T)$ is the set of integers of the form $M^T x$, with $x \in [0, 1)^D$. The result of this transformation on Fig. 5.3(a) is shown in Fig. 5.3(b). Here, it is trivial that the directional downsampling maps the square-shaped passband in Fig. 5.3(a) to diamond-shaped passband as shown in Fig. 5.3(b). It is important to mention that the directional downsampling will not decrease the total number of filter coefficients in the spatial domain, however, it will introduce zeros in the filter.

4. As a last step to construct the desired filter pair, we modulate the 2D filter having a diamond shaped passband with π in each frequency dimension. Mathematically, it can be stated as:

$$\begin{aligned} v(x, y) &= x_w(x, y)e^{j2\pi(\frac{\pi x}{R})} \\ h(x, y) &= x_w(x, y)e^{j2\pi(\frac{\pi y}{C})} \end{aligned} \quad (5.3)$$

where $x_w(x, y)$ is the spatial domain representation of the diamond shaped passband filter, $v(x, y)$ will capture the vertical components, $h(x, y)$ will capture the horizontal components of the image, R and C defines the size of the filter.

5. As, it is obvious from Eq. 5.3, this transformation will map the diamond shaped passband to an hour glass shaped passband. Fig. 5.3(c) and Fig. 5.3(d) shows the Fourier transform after such a transformation. It is also worth mentioning that application of these filters will decompose the input image into its directional components irrespective of their scale. We will use the filters shown in Fig. 5.3(c) and Fig. 5.3(d) to decompose the face image into its directional components. The filter in Fig. 5.3(c) will capture the vertical components while the filter in Fig. 5.3(d) will capture the horizontal components of the image. This is mainly due to the fact that the horizontal components (having gradient in the vertical direction) are localized vertically in the Fourier transform and vice versa. Results after applying DDFB on the normalized images are shown in Fig. 5.4.

5.3 Results and Discussion

To evaluate the performance of the proposed method, we have used two very well known public databases: Yale Face Database B [136] and the Extended Yale Face Database B [137].

5.3.1 Database Preparation

Yale Face Database B contains 5,760 single light source images of 10 different subjects. Each subject has 9 poses with 64 different illumination conditions, inducing local intensity inhomogeneity. The Extended Yale Face Database B contains 16,128 images of 28 subjects

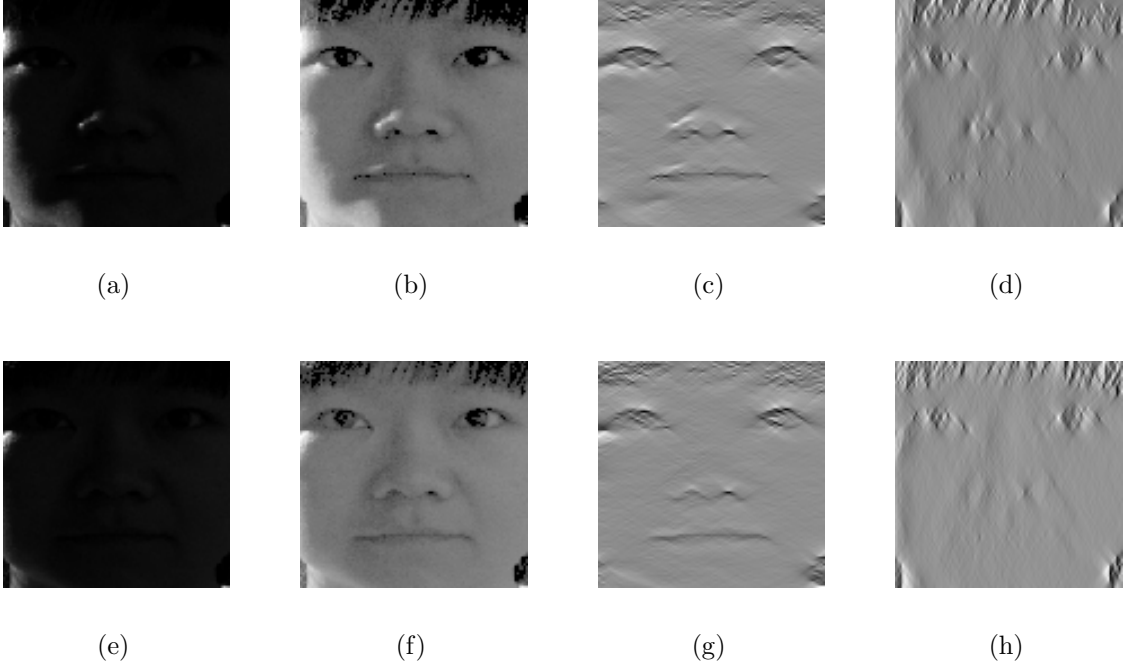


Figure 5.4: a), e) Original images of the same subject. b), f) Images having shadow effects after gamma transformation with $\gamma^* = 5$. c), g) Horizontal features after applying DDFB on the normalized images. d), h) Vertical features after applying DDFB on the normalized images.

under 9 poses and 64 illumination conditions. In our experiments, these two databases were combined to form a single large database. Since, the proposed method deals with the local intensity inhomogeneity correction, only the frontal pose images for each subject with different 64 illumination conditions were selected. In total, the combined database contains 2,432 images from 38 subjects. Out of 2,432 images, 25 corrupted images were discarded as reported in [137, 138]. The whole database was divided into five subsets according to the angle between the light source direction and the camera axis [137].

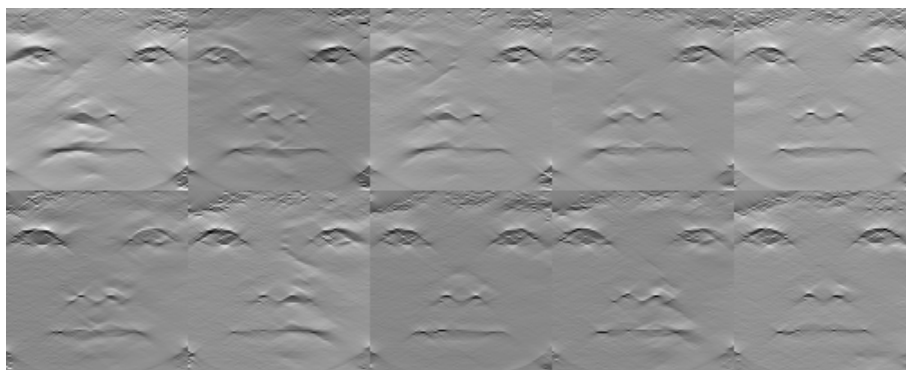
All face images were rotated by finding the slope between the two eye coordinates and then rotating the images in such a way that the slope between the two eye coordinates become zero. Once all the images were aligned, they were cropped and resized to 100×100 .

These images were given as an input to the proposed method. Fig. 5.5 shows the results of the proposed method on some images of subset 5 of a randomly selected subject from the

Yale Face Database B. It is clear from the Fig. 5.5 that the proposed method not only helps in the face normalization but also helps in the suppression of local intensity inhomogeneity.



(a)



(b)

Figure 5.5: a) Some images of subset 5 of a randomly selected subject from the Yale Face Database B. b) Results of the proposed system on 5.5(a).

5.3.2 Matching

In this study, we have used the correlation coefficient to measure the similarity between test and training images. For the quantitative evaluation of the proposed method, images in subset 1 taken under minor illumination variations were used for the training purpose while the other images from subsets 2 to 5 were used as test images. It is often impractical to have a large number of images per subject in the training set. So, it is desirable to have a face recognition system that could achieve a higher recognition rate with fewer images per subject.

To check the robustness of the proposed method given fewer images per subject, we have performed sampling without replacement on the training images. However, such sampling with replacement will introduce biases that can result in poor statistics. To overcome the sampling biases, we have used all possible k -combinations per subject within the subset 1. The total number of combinations for each value of k is computed using:

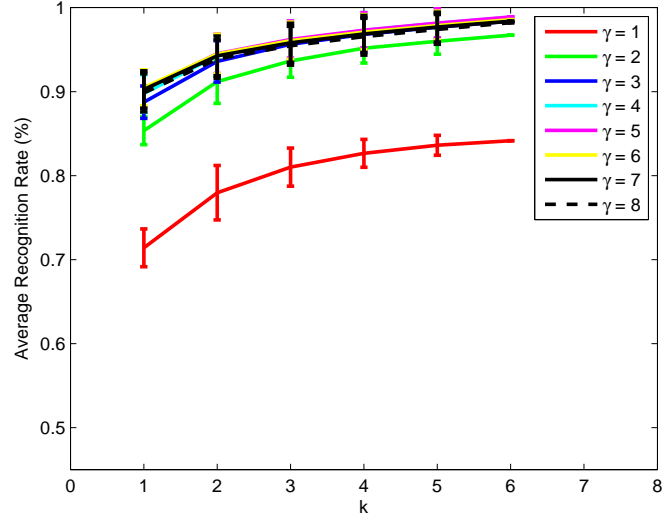
$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (5.4)$$

where n is the total number of images per subject in subset 1. The presence of the three corrupted images in subset 1 as reported in [137, 138] has limited us to use $n = 6$. Fig. 5.6 shows the average recognition rate for the horizontal and the vertical features for different values of n . For instance, $k = 1$ implies that training was accomplished with one image per subject. It is obvious from Fig. 5.6(a) that $\gamma^* = 5$ used by the proposed method is the optimal value which produces the best average recognition rate. The average recognition rate achieved by the horizontal and the vertical features are shown in Fig. 5.6. The results support our observation that horizontal features have more discriminating power in comparison to vertical features.

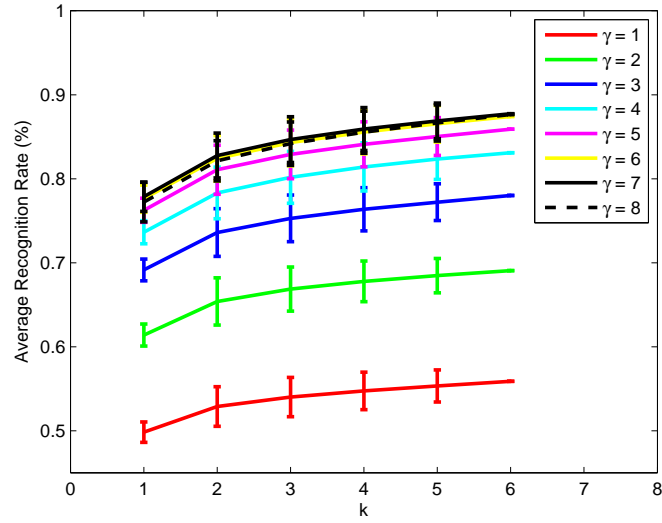
Table 5.1 shows the comparison of the proposed method with some of the existing methods. On average, the recognition rate achieved by the proposed method is far better than the other methods.

5.4 Summary

We have presented a novel filtering method for the local intensity inhomogeneity correction in facial images. Gamma transformation has been used to normalize the input image whereas DDFB is used to suppress the shadow artifacts and the local intensity inhomogeneity in the normalized image. We have used correlation coefficient as a similarity measure for face recognition. Empirically, it is shown that most of the discriminatory features in a human face are horizontal in nature. The efficiency of the proposed method is evaluated on two public



(a)



(b)

Figure 5.6: a) Average recognition rate for horizontal features. b) Average recognition rate for vertical features.

Table 5.1: Recognition rates (%) when using images of subset 1 as training set

Methods	Subset 2	Subset 3	Subset 4	Subset 5	Average
Original	100.00	74.17	45.00	18.95	54.01
SQI [53]	99.56	98.68	88.74	91.8	94.13
DCT [59]	97.81	96.92	90.08	64.78	84.71
TVQI [139]	100.00	96.48	91.41	93.21	94.87
LTV [56]	100.00	99.78	82.25	94.06	93.62
DWT(Hist.) [58]	99.56	77.36	20.23	20.93	49.21
RLS (LOG-DCT) [140]	100.00	87.10	87.60	84.80	89.15
DWT+LTV [138]	100.00	92.53	85.11	86.00	90.12
DWT+TVQI [138]	100.00	99.78	96.18	97.17	98.10
WIIP [141]	94.17	90.00	82.86	55.26	77.51
INPS [142]	100.00	99.17	94.29	95.79	97.04
Proposed Method	100.00	100.00	96.24	99.44	98.89

databases: Yale Face Database B, and the Extended Yale Face Database B. Experimental results demonstrate that the proposed method is superior to other methods in performance and robustness.

Chapter 6

Filter-Bank based Rotation Invariant Feature Codes

6.1 Introduction

A biometric system determines or validates an individual's identity based on his/her physiological and/or behavioral traits. Physiological traits are related to the physical aspects of a human such as fingerprint, iris, face and DNA, which are generally stable over a relatively long period of time. On the other hand, behavioral traits depict the behavior of a person. They tend to change with the time. Examples of behavioral traits include signature, keystrokes, hand writing, and gait [60]. Biometrics are superior to traditional password or token based methods in both security and convenience.

Depending on the application context, a biometric system can operate in either an identification mode or a verification mode [143]. Fig. 6.1 shows the block diagram of a biometric system in both modes. Identification is a one-to-many comparison to find an individual's identity. Verification is a one-to-one match that determines whether the claimed identity is true or false.

During enrolment, a feature vector $\mathbf{x}_i, i = 1, 2, \dots, C$, where C is the total number of participants, is extracted from the biometric data of each individual and stored in the system database as a template. In identification mode, given an input feature vector \mathbf{y} , if the identity

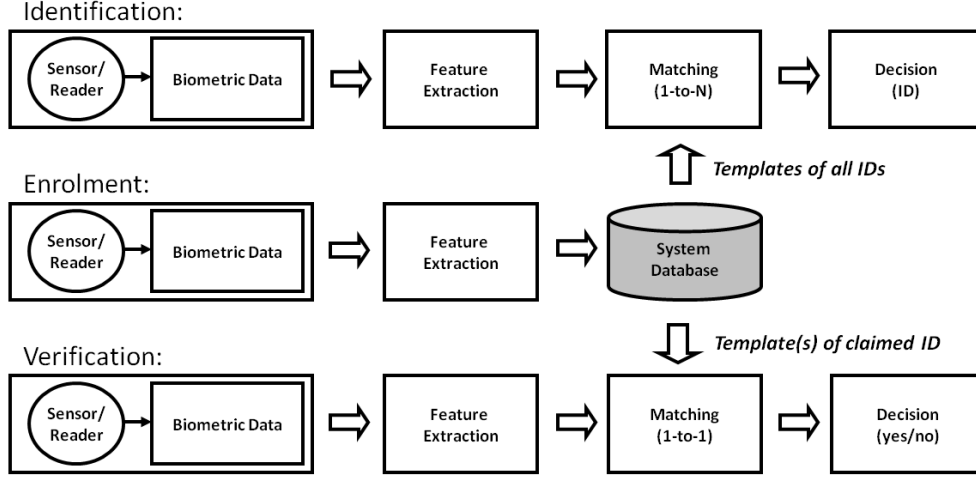


Figure 6.1: Block diagram of biometric recognition systems.

of \mathbf{y} , \mathcal{ID}_y , is known to be in the system database, i.e., $\mathcal{ID}_y \in \{\mathcal{ID}_1, \mathcal{ID}_2, \dots, \mathcal{ID}_C\}$, then \mathcal{ID}_y can be determined by $\mathcal{ID}_y = \mathcal{ID}_k = \arg \min_k \{S(\mathbf{y}, \mathbf{x}_k)\}$, $k = 1, 2, \dots, C$, where S denote the dissimilarity measure and a smaller value represents closer in distance. The performance of a biometric identification system can be evaluated by the percentage of correctly recognized attempts, i.e., correct recognition rate (CRR).

In the verification mode, a feature vector \mathbf{y} is extracted from the biometric signal of the authenticating individual \mathcal{ID}_y , and compared with the stored template \mathbf{x}_j of the claimed identity \mathcal{ID}_j , through a similarity function S . The evaluation of a verification system can be performed in terms of hypothesis testing [144]: \mathbf{H}_0 : $\mathcal{ID}_y = \mathcal{ID}_j$, the claimed identity is correct; \mathbf{H}_1 : $\mathcal{ID}_y \neq \mathcal{ID}_j$, the claimed identity is not correct. The decision is made based on the system threshold τ : \mathbf{H}_0 is decided if $S(\mathbf{y}, \mathbf{x}_j) \leq \tau$ and \mathbf{H}_1 is decided if $S(\mathbf{y}, \mathbf{x}_j) > \tau$.

The performance of a biometric verification system is usually evaluated in terms of *false accept rate* (FAR), *false reject rate* (FRR), and *equal error rate* (EER). FAR is the probability of deciding \mathbf{H}_0 when \mathbf{H}_1 is true, i.e., $P(\mathbf{H}_0|\mathbf{H}_1)$. FRR is the probability of deciding \mathbf{H}_1 when \mathbf{H}_0 is true, i.e., $P(\mathbf{H}_1|\mathbf{H}_0)$. In experimental evaluation, the FAR can be computed as $FAR = \frac{1}{C} \sum_{j=1}^C e_j^{FAR}$, where C is total number of human individuals in the evaluation, and e_j^{FAR} is the FAR for individual \mathcal{ID}_j . Let N_j^s denote the number of successful imposter

attempts against individual \mathcal{ID}_j , and N_j^a denote the total number of imposter attempts against individual \mathcal{ID}_j , then $e_j^{FAR} = N_j^s/N_j^a$. Similarly, the FRR can be computed as $FRR = \frac{1}{C} \sum_{j=1}^C e_j^{FRR}$, $e_j^{FRR} = M_j^r/M_j^a$, where M_j^r denote the number of rejected genuine attempts for individual \mathcal{ID}_j , and M_j^a denote the total number of genuine attempts for individual \mathcal{ID}_j . The FAR and FRR are closely related functions of the system decision threshold τ . EER is defined as the operating point where FAR and FRR are equal. The lower the EER, the better the system performance.

Among various biometrics, fingerprint has been one of the most extensively studied biometric traits due to its high accuracy, stability, and low cost [5,60]. It has been used as a form of identification since 7000 BC [145]. One of the first commercial applications of modern biometric systems which was deployed in early 1970s was also based on fingerprint to provide time keeping and monitoring applications. The location of the minutiae points (ridge endings and bifurcations), and the arrangement of the ridges/valleys make the fingerprints unique. Due to this uniqueness, it has been widely used in many commercial applications. Most government and forensic applications operate in identification mode, such as border control, social security, criminal investigation, and terrorist identification [146]. Examples of verification application include computer and network login, physical access control, electronic data security, and ATM.

A large number of methods have been proposed in the literature for fingerprint recognition, which can be categorized as minutiae-based and image-based methods [5]. This chapter presents an image-based fingerprint system, whose details are given in the section to follow.

6.2 Proposed System for Fingerprint Recognition

This section presents the detailed methodology of the proposed fingerprint recognition system. Fig. 6.2 depicts a block diagram of the proposed system. An input fingerprint image is first enhanced by using a contextual filtering based method [150]. After fingerprint enhancement, the core point is automatically detected by using complex filters [151], and a ROI of size 223×223 centered around the core point is extracted. The ROI is made rotation

invariant by rotating it around the core point in such a way that its core point is at an angle of $\pi/2$. Then a set of Gabor filters oriented at eight different directions are applied to the ROI. The filtered ROI is divided into 5 concentric bands, and each subband is further divided into 16 sectors. For each sector, the AAD is calculated and finally a feature vector of length $8 \times 5 \times 16 = 640$ is generated for each fingerprint image. This feature vector becomes an input to the dimensionality reduction module, and the resulting lower-dimensional feature vector is stored in the system database as a biometric template. At the recognition stage, the same procedure is applied on a probing image, and the resulting lower-dimensional feature vector is matched against those stored in the system database.

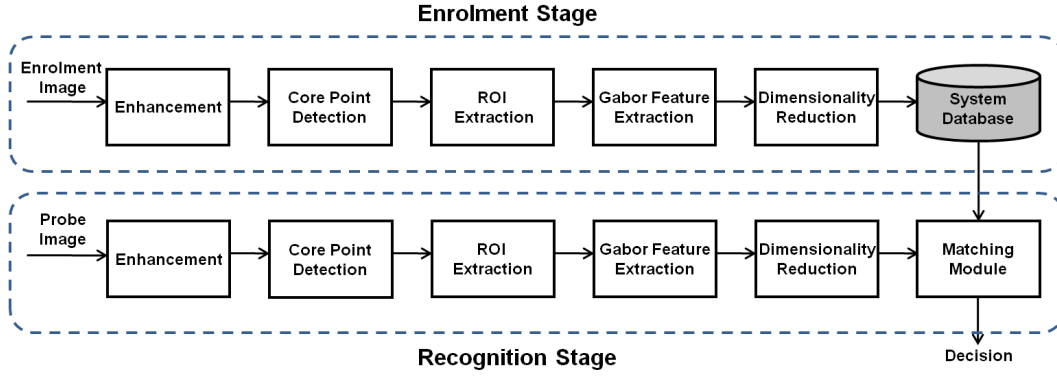


Figure 6.2: Block diagram of the proposed system.

6.2.1 Fingerprint Image Enhancement

The performance of a fingerprint recognition system highly depends on the quality of fingerprint image. To enhance the image quality and obtain accurate representation of the fingerprint characteristics, it is important to perform image enhancement. For this purpose, a contextual filtering based method [150] has been employed. Fig. 6.3 depicts the block diagram of the enhancement method. The enhancement method consists of two steps. It starts by taking STFT to calculate the local ridge orientation, local ridge frequency, and the region mask. In the second step, enhancement is performed in the frequency domain. The details of each step are given below:

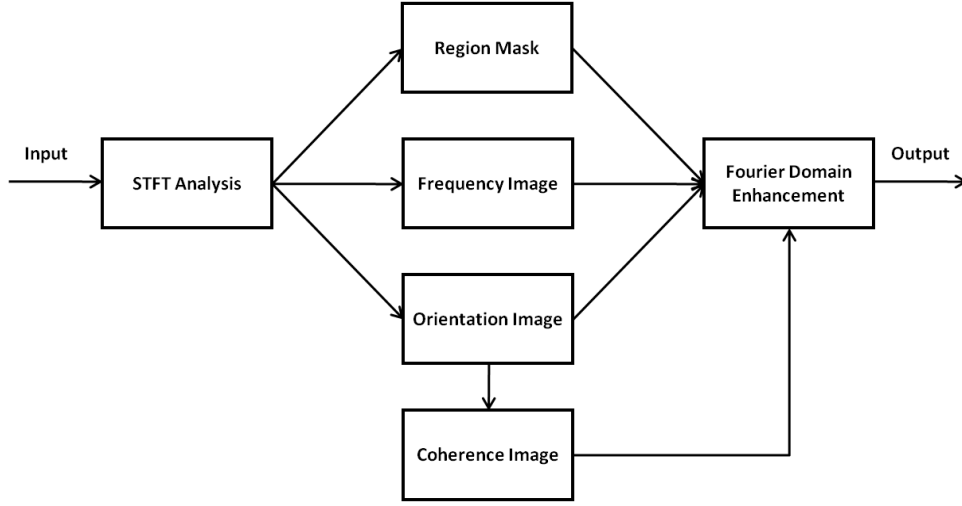


Figure 6.3: Block diagram of the enhancement method proposed in [150].

- Step 1: STFT Analysis

1. For each overlapping block $B(x, y)$ in the image:

- a. Calculate a zero mean block by using:

$$B(x, y) = B(x, y) - \text{mean}(B(x, y)). \quad (6.1)$$

- b. Multiply the zero mean block $B(x, y)$ with the raised cosine spectral window, which tapers smoothly near the border and is unity at the center of the window.

- c. Take the fast Fourier transform (FFT) of the block $F_B = FFT(B)$ and perform the root filtering on F_B :

$$R(x, y) = FFT^{-1}(F_B(u, v)|F_B(u, v)|^k). \quad (6.2)$$

In the above equation, the elements of the power spectrum ($|F_B(u, v)|$) are raised to a power of k (0.4 is used) and then multiplied by $F_B(u, v)$. Finally, the inverse Fourier transform is calculated to generate the filtered block $R(x, y)$. The multiplication of the FFT elements by a power of the power

spectrum has the effect of amplifying the dominant frequencies in the block. Presumably, the dominant frequencies of the block are those corresponding to the ridges, thereby increasing the ratio of ridge information to non-ridge and adapting to variations in ridge frequency from one block to the next.

- d. Perform the STFT analysis on the root filtered block R to obtain ridge orientation image $O(x, y)$, energy image $E(x, y)$ and ridge frequency image $F(x, y)$.
2. Generate a region mask $M(x, y)$ by thresholding the energy image $E(x, y)$ and calculate the coherence image $C(x, y)$ from the smoothed orientation image $O'(x, y)$, which is obtained by performing vector averaging on the orientation image $O(x, y)$.
- Step 2: Enhancement
 - 3) For each overlapping block $B(x, y)$ in the image:
 - a. Compute angular filter F_A centered around $O(x, y)$ with a bandwidth inversely proportional to $C(x, y)$.
 - b. Compute radial filter F_R centered around the frequency $F(x, y)$.
 - c. Filter the block in the FFT domain $F_0 = F \times F_A \times F_R$, and generate the enhanced block by taking the inverse Fourier transform of F_0 .
 - 4) Reconstruct the enhance image by composing the enhanced blocks.

The above enhancement steps are performed on overlapping blocks of size 32×32 pixels. Fig. 6.4 shows examples of enhanced fingerprint images, which demonstrate significantly improved image quality.

6.2.2 Core Point Detection

The accurate detection of the core point is a challenging problem. Many different methods have been proposed in the past. One of the most common and widely used methods is the Poincare index method. It was first developed in [152], used in [153] [154], and was enhanced in [155]. Nilsson et. al [151] proposed an alternative method for the extraction of core and

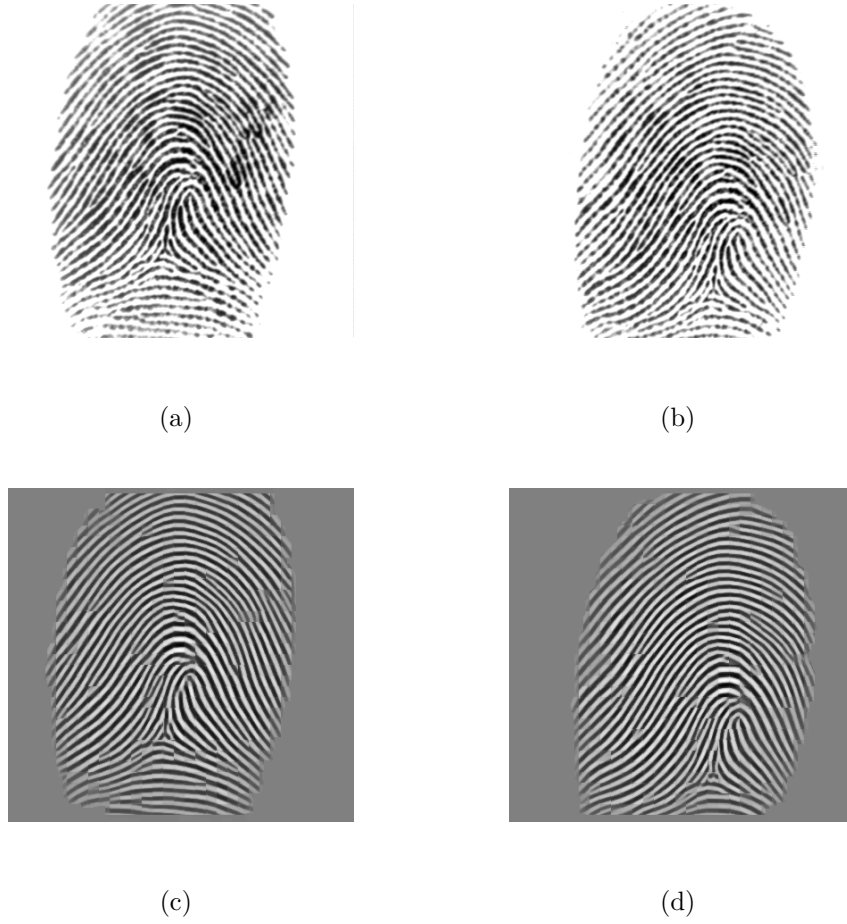


Figure 6.4: Images from FVC2002 *set_a*: a) Original image 1_1 from Db_1a. b) Original image 1_2 from Db_1a. c) Result of enhanced image 1_1. d) Result of enhanced image 1_2.

deltas based on complex filters. By using this method, an input fingerprint image was first converted into a complex image, and a sub-sampled Gaussian pyramid was then applied up to four levels, where level 3 has the lowest resolution, and level 0 has the highest. Only the angle of the complex orientation field was used while the magnitude was set to one in multiscale filtering. Filters of first order symmetry or parabolic symmetry were applied to each resolution for the detection of singular points. Finally, singular points were tracked from the lowest to the highest level. We have adopted this approach for core point detection due to its accuracy and robustness. The detailed algorithm works as follows.

An input fingerprint image is first converted to complex-valued orientation tensor field

image:

$$z(x, y) = (f_x + if_y)^2, \quad (6.3)$$

where f_x is the first derivative of the input image in the horizontal direction and f_y is the first derivative in the vertical direction. They are computed by taking the first derivative of a 1D Gaussian with standard deviation $\sigma = 0.8$, and then convolving it in horizontal and vertical directions, respectively. The small value of σ is selected to capture the small variations in the fingerprint. A sub-sampled Gaussian pyramid is then applied to the resulting complex image $z(x, y)$. At each level of the pyramid, a Gaussian with $\sigma = 0.8$ is applied before downsampling by 2. Finally, the first order symmetry or parabolic symmetry filter is applied at each level of the pyramid to detect the singular points. The first order symmetry or parabolic symmetry filters are computed as:

$$\begin{aligned} h1 &= (x + iy)g(x, y), \\ h2 &= (x - iy)g(x, y), \end{aligned} \quad (6.4)$$

where $h1$ is used to detect the core point, $h2$ is used to detect the delta, and $g(x, y)$ is a Gaussian low-pass filter with $\sigma = 1.5$ (as suggested in [151]). In this work, since we are interested in the detection of the core point, only $h1$ is used. Fig. 6.5 shows some examples of core point detection, in which the effectiveness of the employed method can be easily observed.

6.2.3 ROI and Feature Extraction

Once the core point has been detected, the next step is to identify the ROI and perform feature extraction. One of the most well-known feature extraction methods in fingerprint recognition is the filter-bank based approach [156], which utilizes a bank of Gabor filters to capture both local and global characteristics of a fingerprint. The filter-bank based approach works as follows:

1. Detect the core point and extract the ROI centered around it.

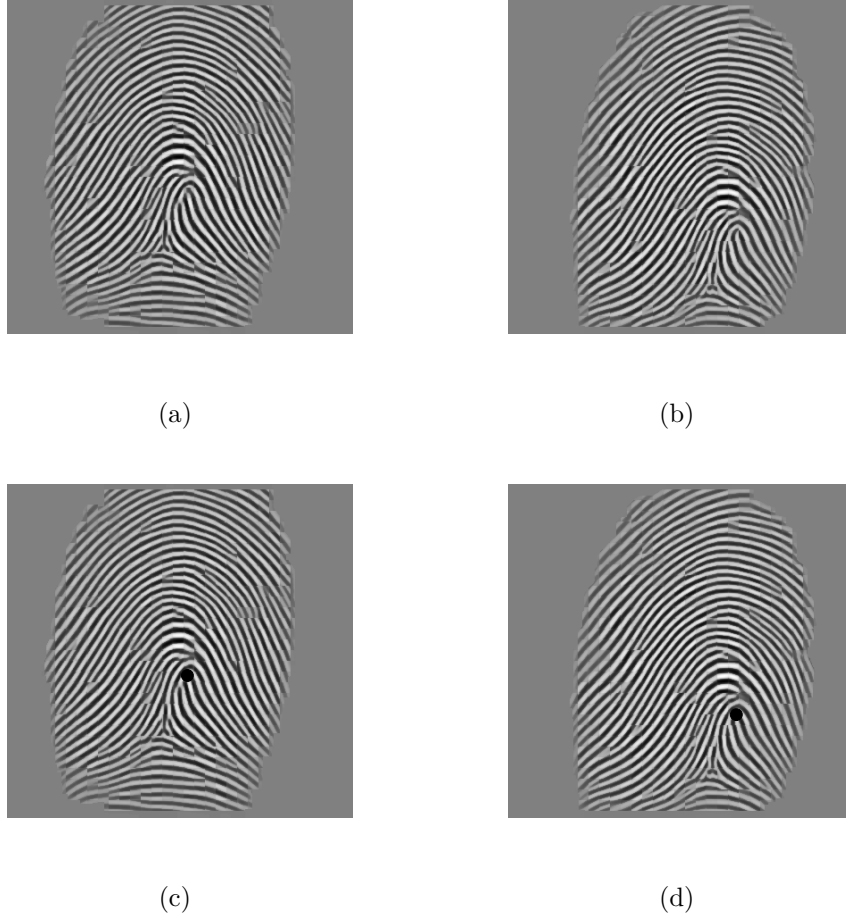


Figure 6.5: Images from FVC2002 set_a : a) Enhanced image 1_1 from Db_1a . b) Enhanced image 1_2 from Db_1a . c) Detected core point in image 1_1. d) Detected core point in image 1_2.

2. The ROI is then divided into B concentric bands and each subband further is divided into k sectors. For a fingerprint scanned at 500 dpi, the value of B is set to 5, the value of k is set to 16, and the width of each sector is 20 pixels as suggested in [156].
3. Normalize each sector such that the grey levels in each sector has a specified mean and variance.
4. Apply a Gabor filter bank on the ROI. Gabor filters will enhance the ridges and valleys that are oriented at the same angle as the filter, and suppresses the ridges and valleys oriented at different angles as shown in Fig. 6.6.

5. Compute AAD of the Gabor responses for each sector:

$$V_{i\theta^\circ} = \frac{1}{n_i} \sum_{n_i} |F_{i\theta^\circ}(x, y) - m_{i\theta^\circ}|, \quad (6.5)$$

where $F_{i\theta^\circ}(x, y)$ is the filtered image at direction θ° , $m_{i\theta^\circ}$ is the mean value of $F_{i\theta^\circ}(x, y)$ in sector S_i , and n_i is the number of pixels in S_i . In the above equation, for all $i \in \{0, 1, \dots, 79\}$ and $\theta \in \{0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ\}$, the AAD feature value ($V_{i\theta^\circ}$) is calculated. To this end, each filtered image will be represented by a feature vector of size $16 \times 5 = 80$, as shown in Fig. 6.7, and a fingerprint image will be encoded by a feature vector of length $8 \times 80 = 640$.

6. Rotate the features in the feature vector cyclically to generate more templates corresponding to different rotations of 22.5° . This cyclic rotation of the feature vector is equivalent to the feature vector generated by the rotation of the original image corresponding to different rotations of 22.5° . For each fingerprint, five different templates corresponding to different rotations of 22.5° will be saved in the database.
7. Rotate the original image by an angle of 11.25° , generate its feature vector, and perform cyclic rotation. Overall, the feature vector was made invariant to the rotation of 11.25° . For each fingerprint, ten templates corresponding to different rotations of the original image were stored in the database.
8. Match the feature vector of an input fingerprint with each of the ten templates generated in the above steps. The final results is the minimum of the matching scores.

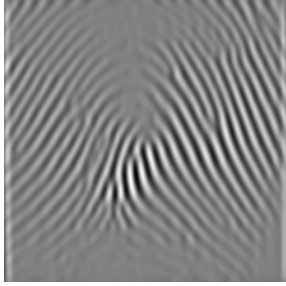
The main drawback of this approach is that it is computationally expensive and it is only invariant to the rotation of 11.25° . For each fingerprint, one needs to generate ten templates. It requires more enrolment time, more recognition time, and larger template storage. In this work, we address the limitation of this approach by enabling the images rotation invariant before feature extraction. Once the core point has been detected, we identify an ROI of size 223×223 with the core point as its center. To produce rotation invariance, the detected



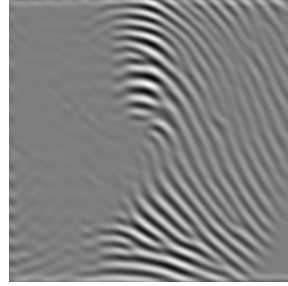
(a)



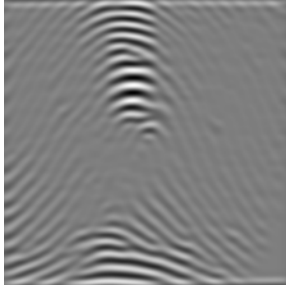
(b)



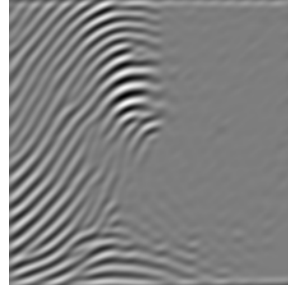
(c)



(d)



(e)



(f)

Figure 6.6: a) Detected core point represented by black circle. b) Region of Interest. c) to f) Results after applying four out of eight oriented Gabor filters on Fig. 6.6(b).

ROI is rotated by an angle of $\pi/2 - \theta_{core}$, where θ_{core} is the angle of the core point, such that the core point is aligned at an angle of $\pi/2$. As such, a sector to sector correspondence

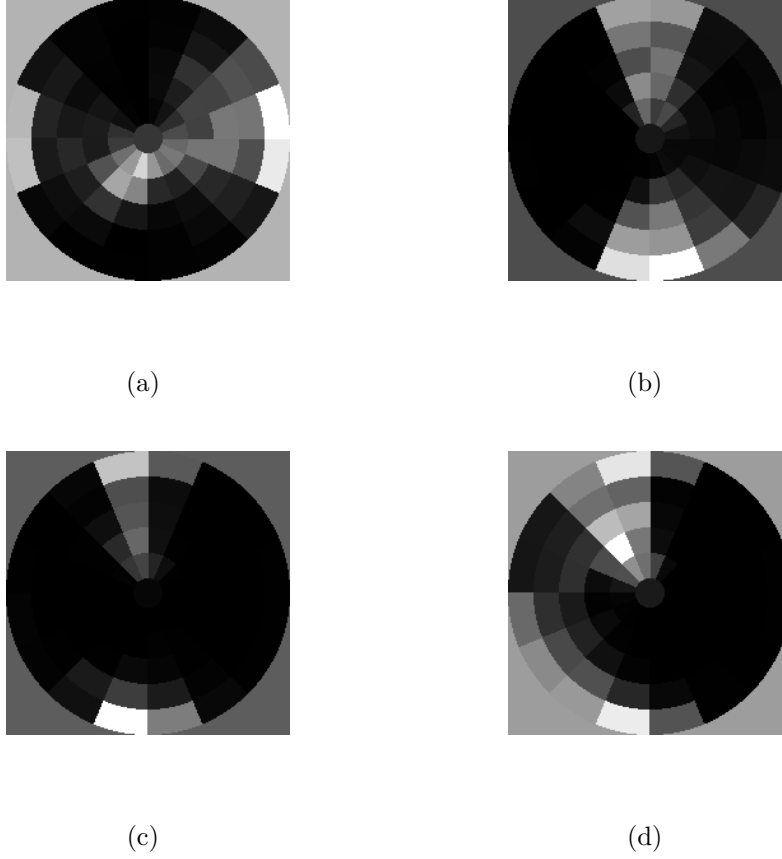


Figure 6.7: AAD features corresponding to the images c) to f) shown in Fig. 6.6, respectively.

between different feature vectors can be established. By using this method, it relaxes the requirement of generating and storing multiple templates corresponding to different rotations of the original image, and solves the problem of limited rotation invariance.

6.2.4 Dimensionality Reduction

The Gabor filter bank method captures the characteristics of a fingerprint image with respect to different orientations, and outputs a feature vector of length 640. Although, it is possible to perform matching using these feature vectors directly but it is generally computationally complex, and provides limited discriminatory power. To address this problem, we apply subspace techniques for dimensionality reduction. Specifically, we compare the performance

of two representative techniques, PCA and LDA.

6.3 Experimental Results

The effectiveness of the proposed method is evaluated through experiments on the public database FVC2002 set_a, which contains four subsets: DB1_a, DB2_a, DB3_a, and DB4_a. The details of each data set are shown in Table 6.1. Each of the subsets contains 800 images from 100 subjects with 8 images per subject. In this work, we use six out of eight fingerprints from each person to train our proposed system and the remaining two for testing. This process is repeated four times by selecting different fingerprints for training and testing, and the average of the experimental results is reported. For PCA method, the reduced dimensionality M is selected automatically based on $\sum_{i=1}^M \lambda_i / \sum_{i=1}^Q \lambda_i \geq 99\%$, where λ_i is the i -th largest eigenvalue, and Q is the total number of computed eigenvalues. For LDA, the reduced dimensionality M is selected as $M = C - 1 = 100 - 1 = 99$, where C is the number of human subjects for training in each data set.

Name	Sensor Type	Image Size	Resolution
DB1_a	Optical Sensor	388×374 pixels	500 dpi
DB2_a	Optical Sensor	296×560 pixels	569 dpi
DB3_a	Capacitive Sensor	300×300 pixels	500 dpi
DB4_a	SFinGe v2.51	288×384 pixels	500 dpi

Table 6.1: Details of FVC2002 Database Set A

6.3.1 Identification Results

For identification, CRR is used to evaluate the system performance, which is defined as the ratio of the number of correctly identified attempts over the total number of identification attempts. Since there are 800 samples in each data set, the experiments are composed of 800 identification attempts for each subset, with 3200 attempts in total. Table 6.2 shows the experimental results of each data set respectively, as well as the average CRR for the whole

database, using original Gabor features, PCA and LDA reduced lower-dimensional features, with Euclidean and Cosine distance as similarity measure respectively. It can be seen that the original Gabor features and the PCA dimensionality reduced features produce similar performance. This is due to the fact that the PCA method only provides a more compact representation of the original features, with no class information being taken into account. On the other hand, it can be seen that the supervised LDA method, which provides a class specific solution, is capable of improving the identification accuracy in all the data sets.

		DB1_a	DB2_a	DB3_a	DB4_a	Average
Gabor	Euc.	97.5	97.5	82.25	84.3750	90.4063
	Cos.	97.75	97.8750	82.00	85.1250	90.6875
PCA	Euc.	97.5	97.5	81.75	84.1250	90.2188
	Cos.	97.25	97.25	81.8750	83.8750	90.0625
LDA	Euc.	98.25	99.00	88.1250	94.1250	94.875
	Cos.	99.3750	99.00	90.6250	95.75	96.1875

Table 6.2: Identification results (CRR in %).

6.3.2 Verification Results

In the verification mode, EER is used to evaluate the performance of the proposed system. It is defined as the operating point where FAR and FRR are equal. A lower EER indicates a better verification performance. As illustrated before, the FAR is computed as the ratio of the number of false accepted impostor claims over the total number of impostor attempts, while the FRR corresponds to the ratio of the number of false rejected genuine claims over the total number of genuine attempts. To compute the FAR and the FRR, the assessment for the genuine match was conducted by comparing the fingerprint for each person with other fingerprints of the same person, while impostor match was carried out by comparing fingerprints of each person with fingerprints of other subjects. For each data set, there are 200 testing samples, therefore the number of genuine and impostor matches are $1 \times 200 = 200$ and $99 \times 200 = 19800$, respectively.

1. User-independent Threshold:

For user-independent threshold, the same system threshold is applied to all the users. Table 6.3 shows the experimental results. Overall, similar as the identification mode, PCA and the original Gabor features obtain similar EER. LDA based method provides better verification performance, and the best result is achieved when cosine distance is employed as the similarity measure.

		DB1_a	DB2_a	DB3_a	DB4_a	Average
Gabor	Euc.	2.0208	3.4583	6.1136	7.4261	4.7547
	Cos.	2.9602	3.8068	7.2134	9.4628	5.8608
PCA	Euc.	1.8775	3.1831	5.9899	7.1010	4.5379
	Cos.	2.0903	3.1351	6.1111	6.9110	4.5619
LDA	Euc.	1.8794	2.0366	7.5657	5.8403	4.3305
	Cos.	0.3194	0.4886	3.4785	2.9173	1.8009

Table 6.3: Verification results of user-independent threshold (EER in %).

2. User-specific Threshold:

For verification, it is possible to utilize user-specific threshold for the final decision based on the characteristics of a specific user. To provide some insight into the design of a user-specific thresholding scheme, we examine three schemes where the intra-class, inter-class, and both intra- and inter-class distributions are taken into consideration.

Let μ_{intra}^k and σ_{intra}^k denote the mean and standard deviation of intra-class (or within-class) distance distribution obtained from the training samples of a specific user \mathcal{ID}_k , with respect to certain distance metrics, such as the Euclidean or cosine distance in this work. Let μ_{inter}^k and σ_{inter}^k denote the mean and standard deviation of the inter-class (or between-class) distance distribution obtained for a user, which is estimated by labelling all the samples from other users as one class, and computing their distance to the samples of the user \mathcal{ID}_k . Then for the intra-class scheme, the user-specific threshold for \mathcal{ID}_k is computed as: $\tau_k = \mu_{intra}^k + \xi \sigma_{intra}^k$, where ξ is a controlling parameter such that different FAR and FRR can be obtained depending the requirement of the

application. For the inter-class scheme, the user-specific threshold for \mathcal{ID}_k is computed as: $\tau_k = \mu_{inter}^k - \xi \sigma_{inter}^k$. For the intra- and inter-class scheme, we assume the intra-class and inter-class distributions are both Gaussian, and compute a point τ_0 at which $P_{intra}(x > \tau_0) = P_{inter}(x < \tau_0)$, where P denote probability. The user-specific threshold for subject \mathcal{ID}_k is then defined as $\tau = \xi \tau_0$.

User-specific threshold			DB1_a	DB2_a	DB3_a	DB4_a	Average
Intra-class	PCA	Euc.	2.6730	4.6850	8.9135	7.2696	5.8853
		Cos.	2.6319	4.4091	9.3864	6.7702	5.7994
	LDA	Euc.	4.5802	4.6275	12.2330	8.4255	7.4666
		Cos.	5.7771	4.0795	16.4792	7.9596	8.5739
Intra- and Inter-class	PCA	Euc.	1.8396	2.9981	6.4514	6.9912	4.5701
		Cos.	2.0676	2.9198	6.3144	5.4722	4.1935
	LDA	Euc.	2.0884	2.5284	8.3870	5.5271	4.6327
		Cos.	1.1957	1.2866	5.3864	3.8687	2.9344
Inter-class	PCA	Euc.	1.7241	2.9905	5.9268	6.5657	4.3018
		Cos.	1.7538	2.5492	6.6963	5.3125	4.0779
	LDA	Euc.	0.8340	1.6117	7.1231	4.8062	3.5937
		Cos.	0.2740	0.5032	3.1812	2.0044	1.4907

Table 6.4: Verification results of user-specific thresholding schemes (EER in %).

Table 6.4 compares the experimental results obtained using different user-specific thresholding schemes. It can be seen that compared with the user-independent scheme presented in Table 6.3, the intra-class scheme degrades the performance significantly. This is due to the fact that there are only a small number of training samples for each subject, and hence the estimation of the intra-class distribution is not accurate. By using both intra- and inter-class distributions, it can be improved to a level of approaching the performance of user-independent thresholding scheme. Again, due to the inaccurate estimation of the intra-class distribution, the verification accuracy still degrade slightly. On the other hand, the inter-class distribution is computed based on a relatively large number of training samples, and therefore provides a better estimation of the characteristics of a certain user with respect to other users. Consequently, it can be seen from Table 6.4 that the inter-class thresholding scheme provides a perfor-

mance improvement over the user-independent scheme in Table 6.3. This observation offers a guideline for the design of a user-specific scheme, in which the investigation of inter-class distribution may possibly provide better recognition performance in a small sample size problem.

6.4 Discussion

The proposed inter-class distribution based user-specific thresholding scheme achieves the best recognition accuracy, when applied on LDA reduced low-dimensional Gabor features with cosine distance as the similarity metric. Furthermore, the effectiveness of the proposed solution is illustrated through a comparison with representative existing works in image-based fingerprint verification, including the score-fusion based hybrid verification system [146], and those based on the Fourier-Mellin transform [69], DCT [67], and Hu’s invariant moments [71]. Table 6.5 compares the performance of the proposed system with the above mentioned methods in terms of EER. It can be seen that the proposed method has outperformed some of the existing methods with an average EER of 1.49%.

Method	DB1_a	DB2_a	DB3_a	DB4_a	Average
Ross [146]	1.87	3.98	4.64	6.21	4.17
Jin [69]	2.43	4.41	5.18	6.62	4.66
Amornraksa [67]	2.96	5.42	6.79	7.53	5.68
Park [71]	1.63	3.78	4.20	4.68	3.57
Proposed	0.27	0.50	3.18	2.00	1.49

Table 6.5: EER (%) comparison of existing methods with proposed system on FVC2002 set_a

6.5 Summary

We have presented a new method for rotation invariant fingerprint recognition. The proposed method employs contextual filtering based technique to enhance the quality of the fingerprint

images, followed by a complex filtering based method for automatic core point detection. To achieve rotation invariant features, an ROI of predefined size is identified and rotated according to the core point angle. This invariant ROI becomes an input to a Gabor filter bank of eight orientations and the AADs are computed as the features. To further reduce the dimensionality of the feature space, whilst obtaining discriminatory representations, two dimensionality reduction tools are examined and compared. The effectiveness of the proposed solution is demonstrated on the well known public database of FVC2002 set_a, in both identification and verification scenarios. To further explore the possibility of utilizing user-specific thresholds for verification, we have studied three different thresholding schemes. Experimental results demonstrate that the proposed method has improved the recognition accuracy in both the identification and the verification scenarios.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

The dissertation presents novel filtering methods for image and video processing applications. The overall objective has been to develop filtering methods for trend removal, intensity inhomogeneity correction, and feature extraction. One of the contributions of the dissertation is the development of filtering method for trend removal and data analysis in videos. The proposed filtering method transforms the non-rigid inter-frame registration problem into low complexity trend filtering problem, i.e, transforming the motion information from the video to a 1D signal. In this dissertation, the proposed method has been employed to perform the heart motion analysis of an embryonic rat from its video recording. In the proposed method, LEM in conjunction with the correlation coefficient is used to extract the motion information from the videos whereas the modified EMD is used to suppress the noise and motion artifacts. Furthermore, the proposed method also helps in the trend removal without defining any regression model. The effectiveness of the proposed method has been evaluated on 151 videos. The results show an accuracy and robustness of the proposed method in counting and localizing different heart states that makes it attractive for the biologists to analyse the heart motion without any need of parameter tweaking and expensive imaging equipment.

In multi-dimensional signals, the trend term is often referred to as shading, non-uniform

illumination, intensity non-uniformity, global intensity inhomogeneity, or bias field [32]. Global intensity inhomogeneity often arises due to the imperfections of data acquisition device, direction of source light, and properties of the subject under study. It affects the performance of image segmentation, matching, retrieval and tracking algorithms [32, 46, 47], to name a few. In this dissertation, we have presented a filtering method for the global intensity inhomogeneity correction. The proposed method is based on the observation that if an image is degraded by the global intensity inhomogeneity, then the magnitude of the frequency components affected by the global intensity inhomogeneity is much higher than its counterparts. In the proposed method, GWDT has been used to automatically localise the affected frequencies in the frequency domain. Moreover, the filter and its parameters are automatically determined from the affected frequency components of an image.

To check the robustness of the proposed method, we have presented a comparison with the recent and state-of-the-art methods. The performance has been evaluated both quantitatively and qualitatively. For the qualitative analysis, we have followed the subjective evaluation by humans that is reported as the most appropriate criterion for the assessment of the image quality. Whereas, the quantitative evaluation was performed by measuring segmentation accuracy, true positive rate, false positive rate, similarity with ground truth measured in terms of cross correlation, coefficient of contrast, and coefficient of joint variation. Experimental results demonstrate the superiority of the proposed method in comparison with some of the existing and state-of-the-art methods.

The same idea has been extended to 3D for the volume images, especially for the brain MR volumes. The proposed method is based on the GWDT of the magnitude spectrum of the volume image. Here, a filter is automatically generated based on the content of the MR volume. To the best of our knowledge, the proposed filtering method is the first of its kind where the shape of the filter solely depends on the global intensity inhomogeneity present in the image.

One of the problems closely related to global intensity inhomogeneity is local intensity inhomogeneity. Global intensity inhomogeneity is considered as a slowly varying compo-

nent but this assumption does not hold in case of images having shadow artifacts. These shadow artifacts appear as sharp changes at shadow boundaries [55], making intensity inhomogeneity local in nature. To cope with this problem, we have presented a local intensity inhomogeneity correction method based on gamma transformation and DDFB. In the proposed method, gamma transformation is used to normalize the input images whereas DDFB is used to suppress the local intensity inhomogeneity and shadow artifacts irrespective of scale and orientation. In this dissertation, the proposed method has been deployed to suppress these artifacts in facial images. Empirically, it is shown that most of the discriminatory power lies within the horizontal facial features under shadow artifacts. The efficiency of the proposed method has been evaluated on two public databases: Yale Face Database B, and the Extended Yale Face Database B. It is clear from the experimental results that proposed filtering method is superior to the existing methods in performance and robustness, as it does not require any parameter tweaking.

One of the key applications of image filtering is feature extraction. Feature extraction has been widely used in extracting discriminatory information from the biometric samples which can be further used for person identification. In this dissertation, we have presented a filter-bank based approach to extract the rotation invariant feature codes from the fingerprint images. The proposed method generates the rotation invariant feature codes based on the output of the complex filters and the Gabor filter bank. After generating the rotation invariant feature codes, each fingerprint is represented by a feature vector of length 640. To further reduce the dimensionality of the feature space, we have applied subspace techniques for dimensionality reduction. Specifically, we have compared the performance of two representative techniques, namely PCA and LDA. User-specific thresholding schemes are also investigated to improve the performance of the system. We have evaluated the performance of the proposed method on FVC2002 set_a public database, in both identification and verification scenarios.

This dissertation demonstrates the importance of filtering methods that are independent of parameter tweaking and user interaction for image and video processing applications.

Such filtering methods will help to eliminate user biases.

7.2 Future Work

As an extension of this dissertation, we propose the following possible directions for future research.

- The methodology proposed for the key frame selection can be used in surveillance applications. Especially, in motion detection, selection of a key frame plays a vital role.
- Development of an iterative method for global intensity inhomogeneity correction in color images and videos. Such method can have the wide range of applications in image and video processing.
- The proposed method for global intensity inhomogeneity correction can also help to improve the performance of different descriptors like the scale invariant feature transform (SIFT) [157]. The classical framework of SIFT is not illumination invariant and its performance degrades in the presence of non-uniform illumination.
- Extension of the proposed global intensity inhomogeneity correction method to suppress atmospheric haze in color images. Generally, atmospheric haze is created by small water particles, dirt, and/or soot, and also by the air molecules.
- The proposed global intensity inhomogeneity correction method can also help to improve the segmentation accuracy if used in conjunction with state-of-the-art segmentation methods.
- GWDT in the frequency domain can be used to detect dominant frequency regions in directional images.
- To concentrate more on the filter design aspect for 2D and 3D global intensity inhomogeneity correction methods. State-of-the-art methods can be used for the conversion

of the ideal filter to non-ideal.

- Development of a method that can keep the facial features intact while removing the local intensity inhomogeneity in images and videos.
- Integrating the local histogram with the proposed rotation invariant feature codes to form an image descriptor. Such kind of descriptor can have a wide range of applications in image classification and retrieval.

Bibliography

- [1] W. S. Cleveland, *The Elements of Graphing Data*. AT&T Bell Laboratories, 1994.
- [2] D. M. Regan, *Human Perception of Objects: Early Visual Processing of Spatial Form Defined by Luminance, Color, Texture, Motion, and Binocular Disparity*. Sinauer Associates Inc., 1st ed., March 2000.
- [3] K. K. Evans and A. Treisman, “Perception of objects in natural scenes: Is it really attention free?,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 31, no. 6, pp. 1476–1492, 2005.
- [4] F. Camastra, “Image processing: Principles and applications [book review],” *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 610, 2007.
- [5] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. London: Springer, 2nd ed., 2009.
- [6] E. C. Kyriacou, C. S. Pattichis, M. S. Pattichis, C. P. Loizou, C. S. Christodoulou, S. K. Kakkos, and A. Nicolaides, “A review of noninvasive ultrasound image processing methods in the analysis of carotid plaque morphology for the assessment of stroke risk,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 1027–1038, 2010.
- [7] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, Aug. 2010.
- [8] A. K. Jain, *Fundamentals of Digital Image Processing*. Prentice Hall, 1st ed., Oct. 1988.

- [9] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Thomson-Engineering, 3rd ed., March 2007.
- [10] Z. Jing and B. Nath, “Image processing techniques of landmines: A review,” in *Proceedings of International Conference on Intelligent Sensing and Information Processing*, pp. 143–148, Aug. 2004.
- [11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Singapore: Pearson Education, Inc., 3rd ed., 2007.
- [12] A. C. Bovik, *Handbook of Image and Video Processing (Communications, Networking and Multimedia)*. Orlando, FL, USA: Academic Press, Inc., 1st ed., 2000.
- [13] H. Greenspan, C. H. Anderson, and S. Akber, “Image enhancement by nonlinear extrapolation in frequency space,” *IEEE Transactions on Image Processing*, vol. 9, pp. 1035–1048, Aug. 2000.
- [14] F. Russo, “An image enhancement technique combining sharpening and noise reduction,” *IEEE Transactions on Instrumentation and Measurement*, vol. 51, pp. 824–828, Dec. 2002.
- [15] I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh, eds., *Feature Extraction, Foundations and Applications*. Series Studies in Fuzziness and Soft Computing, Physica-Verlag, Springer, 2006.
- [16] T. Alexandrov, S. Bianconcini, E. B. Dagum, P. Maass, and T. McElroy, “A review of some modern approaches to the problem of trend extraction,” Research Report Series, Statistics 2008-3, Statistical Research Division, U.S. Census Bureau, Washington, March 2008.
- [17] R. Hodrick and E. Prescott, “Postwar U.S. business cycles: An empirical investigation,” *Journal of Money, Credit, and Banking*, vol. 29, pp. 1–16, 1997.

- [18] K. Singleton, “Econometric issues in the analysis of equilibrium business cycle models,” *Journal of Monetary Economy*, vol. 21, pp. 361–368, 1988.
- [19] S. Levitt, “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not,” *Journal of Economic Perspectives*, vol. 18, pp. 163–190, 2004.
- [20] R. Talluri and G. van Ryzin, *The Theory and Practice of Revenue Management*. Kluwer Academic, Boston, MA, 2004.
- [21] S. Greenland and M. Longnecker, “Methods for trend estimation from summarized doseresponse data, with applications to meta-analysis,” *American Journal of Epidemiology*, vol. 135, pp. 1301–1309, 1992.
- [22] W. Link and J. Sauer, “Estimating equations estimates of trend,” *Bird Populations*, vol. 2, pp. 23–32, 1994.
- [23] A. Schlogl, J. Fortin, W. Habenbacher, and M. Akay, “Adaptive mean and trend removal of heart rate variability using Kalman filtering,” in *IEEE 23rd Annual International Conference of Engineering in Medicine and Biology Society*, vol. 1, pp. 571–573, Oct. 2001.
- [24] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, March 1998.
- [25] M. K. K. Niazi, M. T. Ibrahim, M. F. Nilsson, A. Sköld, L. Guan, and I. Nystörm, “Robust Signal Generation and Analysis of Rat Embryonic Heart Rate in Vitro Using Laplacian Eigenmaps and Empirical Mode Decomposition,” in *14th International Conference on Computer Analysis of Images and Pattern (CAIP 2011)*, vol. 6855 of Lecture Notes in Computer Science, pp. 523–530, Springer-Verlag Berlin/Heidelberg, 2011.

- [26] E. Jacobsen and R. Lyons, “The sliding DFT,” *Signal Processing Magazine*, vol. 20, no. 4, pp. 74–80, March 2003.
- [27] N. E. Huang and W. Zhaohua, “A review on Hilbert-Huang transform: Method and its applications to geophysical studies,” *Reviews of Geophysics*, vol. 46, pp. 1–23, June 2008.
- [28] M. Shah and R. Kumar, *Video Registration (The International Series in Video Computing)*. Kluwer Academic Publishers, 2003.
- [29] S. S. Beuchemin and J. L. Barron, “The computation of optical flow,” *ACM Computing Surveys*, vol. 27, pp. 433–467, 1995.
- [30] B. Zitov and J. Flusser, “Non-rigid image registration: Theory and practice,” *The British Journal of Radiology*, vol. 77, pp. 140–153, 2004.
- [31] P. R. Andersen and M. Nielsen, “Non-rigid registration by geometry constrained diffusion,” *Medical Image Analysis*, vol. 5, pp. 81–88, 2001.
- [32] U. Vovk, F. Pernus, and B. Lika, “A review of methods for correction of intensity inhomogeneity in MRI,” *IEEE Transactions on Medical Imaging*, vol. 26, pp. 405–421, March 2007.
- [33] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, “Parametric estimate of intensity inhomogeneities applied to MRI,” *IEEE Transactions on Medical Imaging*, vol. 19, pp. 153–165, March 2000.
- [34] B. Belaroussi, J. Milles, S. Carme, Y. M. Zhu, and H. Benoit-Cattin, “Intensity non-uniformity correction in MRI: Existing methods and their validation,” *Medical Image Analysis*, vol. 10, no. 2, pp. 234–246, 2006.
- [35] T. Tasdizen, R. Whitaker, R. Marc, and B. Jones, “Non-uniform illumination correction in transmission electron microscopy,” in *MICCAI Workshop on Microscopic Image Analysis with Applications in Biology*, Sept. 2008.

- [36] Y. Zheng, M. Grossman, S. Awate, and J. Gee, “Automatic correction of intensity nonuniformity from sparseness of gradient distribution in medical images,” in *Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part II*, vol. 5762, pp. 852–859, Springer-Verlag Berlin, Heidelberg, 2009.
- [37] U. Köthe, “Edge and junction detection with an improved structure tensor,” in *Pattern Recognition, 25th DAGM Symposium, Magdeburg, Germany* (B. Michaelis and G. Krell, eds.), vol. 2781 of *Lecture Notes in Computer Science*, pp. 25–32, Springer, September 2003.
- [38] J. G. Sled, A. F. Zijdenbos, and A. C. Evans, “A nonparametric method for automatic correction of intensity nonuniformity in MRI data,” *IEEE Transactions on Medical Imaging*, vol. 17, pp. 87–97, Feb. 1998.
- [39] L. Axel, J. Costantini, and J. Listerud, “Intensity correction in surface-coil MR imaging,” *American Journal of Roentgenology*, vol. 148(2), pp. 418–420, 1987.
- [40] B. H. Brinkmann, A. Manduca, and R. A. Robb, “Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction,” *IEEE Transactions on Medical Imaging*, vol. 17(2), pp. 161–171, 1998.
- [41] E. Ardizzone, R. Pirrone, and O. Gambino, “Bias artifact suppression on MR volumes,” *Computer Methods and Programs in Biomedicine*, vol. 92, pp. 35–53, 2008.
- [42] R. Guillemaud, “Uniformity correction with homomorphic filtering on region of interest,” in *IEEE International Conference on Image Processing*, vol. 2, pp. 872–875, 1998.
- [43] E. Ardizzone, R. Pirrone, and O. Gambino, “Pervasive access to MRI bias artifact suppression service on a grid,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13(1), pp. 87–93, 2009.
- [44] E. Ardizzone, R. Pirrone, and O. Gambino, “Fuzzy c-means segmentation on brain MR

- slices corrupted by RF-Inhomogeneity,” in *Proceedings of the 7th International Workshop on Fuzzy Logic and Applications: Applications of Fuzzy Sets Theory*, pp. 378–384, 2007.
- [45] E. Ardizzone, R. Pirrone, and O. Gambino, “Frequency determined homomorphic unsharp masking algorithm on knee MR images,” in *13th International Conference on Image Analysis and Processing*, pp. 922–929, 2005.
- [46] L. Juan and O. Gwun, “A comparison of SIFT, PCA-SIFT and SURF,” *International Journal of Image Processing (IJIP)*, vol. 3, no. 4, pp. 143–152, 2010.
- [47] B. Georgescu and P. Meer, “Point matching under large image deformations and illumination changes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 674–688, 2004.
- [48] X. Zou, J. V. Kittler, and K. Messer, “Illumination invariant face recognition: A survey,” in *IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pp. 1–8, Sept. 2007.
- [49] R. Gross and V. Brajovic, “An image preprocessing algorithm for illumination invariant face recognition,” in *4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pp. 10–18, Springer, June 2003.
- [50] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone, “FRVT 2002: Evaluation report,” Tech. rep., March 2002.
- [51] P. Phillips, W. Scruggs, A. O’Toole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, “FRVT 2006 and ICE large-scale experimental results,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 831–846, May 2010.
- [52] H. Lee and D. Kim, “Expression-invariant face recognition by facial expression transformations,” *Pattern Recognition Letters*, vol. 29, pp. 1797–1805, Oct. 2008.

- [53] H. Wang, S. Z. Li, and Y. Wang, “Face recognition under varying lighting conditions using self quotient image,” in *IEEE International Conference on Automatic Face and Gesture Recognition*, AFGR’04, pp. 819–824, 2004.
- [54] G. An, J. Wu, and Q. Ruan, “An illumination normalization model for face recognition under varied lighting conditions,” *Pattern Recognition Letters*, vol. 31, pp. 1056–1067, July 2010.
- [55] D. J. Jobson, Z. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE Transactions on Image Processing*, vol. 6, pp. 965–976, 1997.
- [56] T. Chen, W. Yin, X. Zhou, D. Comaniciu, and T. Huang, “Total variation models for variable illumination face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1519–1524, 2006.
- [57] T. F. Chan and S. Esedoglu, “Aspects of total variation regularized L1 function approximation,” Tech. Rep. CAM Report 04-07, University of California, Feb. 2004.
- [58] S. Du and R. Ward, “Wavelet-based illumination normalization for face recognition,” in *IEEE International Conference on Image Processing*, pp. 954–957, 2005.
- [59] W. Chen, M. J. Er, and S. Wu, “Illumination compensation and normalization for robust face recognition using Discrete Cosine Transform in logarithm domain,” *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 36, no. 2, pp. 458–466, 2006.
- [60] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 3, pp. 4–20, 2004.
- [61] A. K. Jain, L. Hong, and R. Bolle, “Real-time matching system for large fingerprint databases,” *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol. 19, no. 4, pp. 302–314, 1997.

- [62] D. D. Hung, "Enhancement and feature purification of fingerprint images," *Pattern Recognition*, vol. 26, no. 11, pp. 1661–1671, 1993.
- [63] M. S. Khalil, D. Mohamad, M. K. Khan, and Q. Al-Nuzaili, "Fingerprint pattern classification," *Digital Signal Processing*, vol. 20, pp. 1264–1273, July 2010.
- [64] C. L. S. Wang, "Fingerprint feature extraction using Gabor filters," *Electronic Letters*, vol. 35, no. 4, pp. 288–290, 1999.
- [65] M. Tico, E. Immomen, P. Ramo, P. Kuosmanen, and J. Saarinen, "Fingerprint recognition using wavelet features," in *Proc. ISCAS*, vol. 2, (Australia), pp. 21–24, May 2001.
- [66] M. Tico, P. Kuosmanen, and J. Saarinen, "Wavelet domain features for fingerprint recognition," *Electronic Letters*, vol. 37, no. 1, pp. 21–22, 2001.
- [67] T. Amornraksa and S. Tachaphetpi boon, "Fingerprint recognition using DCT features," *Electronics Letters*, vol. 42, no. 9, pp. 522–523, 2006.
- [68] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "Filterbank-based fingerprint matching," *IEEE Transactions on Image Processing*, vol. 9, pp. 846–859, May 2000.
- [69] A. T. B. Jin, D. N. C. Ling, and O. T. Song, "An efficient fingerprint verification system using integrated wavelet and fourier-mellin invariant transform," *Image and Vision Computing*, vol. 22, pp. 503–513, June 2004.
- [70] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, pp. 179–187, 1962.
- [71] J. C. Yang and D. S. Park, "A fingerprint verification algorithm using tessellated invariant moment features," *Neurocomputing*, vol. 71, no. 10-12, pp. 1939–1946, 2008.
- [72] J. C. Yang and D. S. Park, "Fingerprint verification based on invariant moment features and nonlinear bpnn," *International Journal of Control, Automation, and Systems*, vol. 6, no. 6, pp. 800–808, 2008.

- [73] J. Pawley, ed., *Handbook of Biological Confocal Microscopy, 3rd Edition*. New York: Springer, 2006.
- [74] D. J. Fleet, “Disparity from local weighted phase-correlation,” in *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 48–56, Oct. 1994.
- [75] S. Park, *New Directional Filter Banks and Their Applications in Image Processing*. PhD thesis, Georgia Institute of Technology, Nov. 1999.
- [76] C. H. Park, J. J. Lee, M. J. T. Smith, S. Park, and K. J. Park, “Directional filter bank-based fingerprint feature extraction and matching,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 74–85, Jan. 2004.
- [77] J. Yang, L. Liu, T. Jiang, and Y. Fan, “A modified Gabor filter design method for fingerprint image enhancement,” *Pattern Recognition Letters*, vol. 24, pp. 1805–1817, Aug. 1995.
- [78] D. Dunn and W. E. Higgins, “Optimal Gabor filters for texture segmentation,” *IEEE Transactions on Image Processing*, vol. 4, pp. 947–964, July 1995.
- [79] J. Daugman, “Two-dimensional spectral analysis of cortical receptive field profiles,” *Vision Research*, vol. 20, no. 10, pp. 847–856, 1980.
- [80] J. Daugman, “Complete discrete 2D Gabor transforms by neural networks for image analysis and compression,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1169–1179, 1988.
- [81] A. K. Jain and F. Farrokhnia, “Unsupervised texture segmentation using Gabor filters,” *Pattern Recognition*, vol. 24, pp. 1167–1186, Dec. 1991.
- [82] V. S. Vyas and P. Rege, “Automated texture analysis with Gabor filter,” *International Journal on Graphics, Vision and Image Processing*, vol. 6, p. 3541, July 2006.

- [83] R. H. Bamberger, *The Directional Filter Bank: A Multirate Filterbank for the Directional Decomposition of Images*. PhD thesis, Georgia Institute of Technology, Nov. 1990.
- [84] R. H. Bamberger and M. J. T. Smith, “A filter bank for the directional decomposition of images: Theory and Design,” *IEEE Transactions on Signal Processing*, vol. 40, pp. 882–893, April 1992.
- [85] S. Park, M. J. T. Smith, and R. M. Mersereau, “A new directional filter bank for image analysis and classification,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1417–1420, 1999.
- [86] S. Park, M. J. T. Smith, and J. J. Lee, “Fingerprint enhancement based on the directional filter bank,” in *IEEE International Conference on Image Processing*, pp. 793–796, Sept. 2000.
- [87] M. A. U. Khan, M. K. Khan, and M. A. Khan, “Coronary angiogram image enhancement using decimation-free directional filter banks,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. V – 441–4, May 2004.
- [88] B. L. Evans, R. H. Bamberger, and J. H. McCellan, “Rules for multidimensional multirate structures,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 762–771, April 1994.
- [89] T. Oberlin, S. Meignen, and Valérie Perrier, “An Alternative Formulation for the Empirical Mode Decomposition,” Tech. Rep. hal-00553107, HAL, Jan. 2011.
- [90] N. E. Huang, M. L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, “A confidence limit for the empirical mode decomposition and hilbert spectral analysis,” *Proceedings of the Royal Society of London. Series A*, vol. 459, pp. 2317–2345, Sept. 2003.
- [91] W. Huang, Z. Shen, N. E. Huang, and Y. C. Fung, “Engineering analysis of biological variables: an example of blood pressure over 1 day,” in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 4816–4821, April 1998.

- [92] L. Hualou, L. Qiu-Hua, and J. D. Z. Chen, “Application of the empirical mode decomposition to the analysis of esophageal manometric data in gastroesophageal reflux disease,” *IEEE Transactions on Biomedical Engineering*, vol. 52, pp. 1692–1701, Oct. 2005.
- [93] H. Liang, Z. Lin, and R. McCallum, “Artifact reduction in electrogastrogram based on empirical mode decomposition method,” *Medical and Biological Engineering and Computing*, vol. 38, pp. 35–41, 2000.
- [94] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review,” *ACM Computing Surveys*, vol. 31, pp. 264–323, Sept. 1999.
- [95] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988.
- [96] U. von Luxburg, “A tutorial on spectral clustering,” *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.
- [97] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data, An Introduction to Cluster Analysis*. JohnWiley & Sons, Brussels, Belgium, 1990.
- [98] A. P. Reynolds, G. Richards, and V. J. Rayward-smith, “The application of K-medoids and PAM to the clustering of rules,” in *Intelligent Data Engineering and Automated Learning*, vol. 3177 of *Lecture Notes in Computer Science*, pp. 173–178, Springer Berlin, 2004.
- [99] L. Kaufman and P. J. Rousseeuw, “Clustering by means of medoids,” *Statistical Data Analysis Based on the L1 Norm*, Y. Dodge, Ed., pp. 405–416, 1987.
- [100] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters,” *Journal of Cybernetics and Systems*, vol. 3, no. 3, pp. 32–57, 1973.
- [101] C. Burges, “Dimension reduction a guided tour,” Tech. Rep. MSR-TR-2009-2013, Microsoft Research, 2009.
- [102] L. T. Jolliffe, *Principle Component Analysis*. New York: Springer-Verlag, 1986.

- [103] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.
- [104] M. Khan, M. Nilsson, B. Danielsson, and E. Bengtsson, "Fully automatic heart beat rate determination in digital video recordings of rat embryos," in *Advances in Mass Data Analysis of Images and Signals in Medicine, Biotechnology, Chemistry and Food Industry* (P. Perner and O. Salvetti, eds.), vol. 5108 of *Lecture Notes in Computer Science*, pp. 27–37, Springer Berlin / Heidelberg, 2008.
- [105] D. Abela, H. Ritchie, D. Ababneh, C. Gavin, M. F. Nilsson, M. K. Khan, K. Carlsson, and W. S. Webster, "The effects of drugs with ion channel-blocking activity on the early embryonic rat heart," *Birth Defects Research Part B: Developmental and Reproductive Toxicology*, vol. 89, pp. 429–440, Oct. 2010.
- [106] W. S. Webster and D. Abela, "The effect of hypoxia in development," *Birth Defects Res Part C*, vol. 81, pp. 215–228, 2007.
- [107] P. M. Doubilet and C. B. Benson, "Outcome of first-trimester pregnancies with slow embryonic heart rate at 6-7 weeks gestation and normal heart rate by 8 weeks at us," *Radiology*, vol. 236, pp. 643–636, Aug. 2005.
- [108] A. C. Sköld, C. Danielsson, B. Linder, and B. R. Danielsson, "Teratogenicity of the IKr-Blocker Cisapride: Relation to Embryonic Cardiac Arrhythmia," *Reproductive Toxicology*, vol. 16, no. 4, pp. 333–342, 2002.
- [109] M. Fink, C. Callol-Massot, A. Chu, P. Ruiz-Lozano, J. C. I. Belmonte, W. Giles, R. Bodmer, and K. Ocorr, "A new method for detection and quantification of heartbeat parameters in drosophila, zebrafish, and embryonic mouse hearts," *Journal of BioTechniques*, vol. 46, pp. 101–113, Feb. 2009.
- [110] P. K. Chan, C. C. Lin, and S. H. Cheng, "Noninvasive technique for measurement of

- heartbeat regularity in zebrafish (*Danio rerio*) embryos,” *Journal of BMC Biotechnology*, vol. 9, pp. 1–10, Feb. 2009.
- [111] J. T. Zhu, J. He, J. Y. Chen, D. R. Lu, and L. W. Zhou, “Fast differential interference contrast imaging combined with autocorrelation treatments to measure the heart rate of embryonic fish,” *Journal of Biomedical Optics*, vol. 13, pp. 020503–1–020503–3, May 2008.
- [112] M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data representation,” *Neural Computation*, vol. 15, pp. 1373–1396, June 2003.
- [113] P. Yiou, D. Sornette, and M. Ghil, “Data-adaptive wavelets and multi-scale singular-spectrum analysis,” *Physica D: Nonlinear Phenomena*, vol. 142, pp. 254–290, August 2000.
- [114] S. E. Hadji, R. Alexandre, and A. O. Boudraa, “Analysis of Intrinsic Mode Functions: A PDE Approach,” *IEEE Signal Processing Letters*, vol. 17, pp. 398–401, April 2010.
- [115] H. Hong, X. Wang, Z. Tao, and S. Du, “Centroid-based sifting for empirical mode decomposition,” *Journal of Zhejiang University - Science C*, vol. 12, pp. 88–95, 2011.
- [116] H. Hong, X. Wang, and Z. Tao, “Local integral mean-based sifting for empirical mode decomposition,” *IEEE Signal Processing Letters*, vol. 16, pp. 841–844, Oct. 2009.
- [117] K. C. Strasters, A. W. M. Smeulders, and H. T. M. V. Voort, “3-d texture characterized by accessibility measurements, based on the grey weighted distance transform,” *BioImaging*, vol. 2, no. 1, pp. 1–21, 1994.
- [118] P. K. Saha, F. W. Wehrli, and B. R. Gomberg, “Fuzzy distance transform: Theory, algorithms, and applications,” *Computer Vision and Image Understanding*, vol. 86(3), pp. 171–190, 2002.
- [119] L. Kubecka, J. Jan, and R. Kolar, “Retrospective illumination correction of retinal images,” *International Journal of Biomedical Imaging*, vol. 2010, pp. 1–10, 2010.

- [120] J. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. V. Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 501–509, April 2004.
- [121] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piece wise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19(3), pp. 203–210, 2000.
- [122] "VICAVER, VARPA images for the computation of the Arterio/Venular Ratio database," <http://www.varpa.org/RetinalResearch.html> 2009.
- [123] R.-S. Kwan, A. Evans, and G. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 1085–1097, Nov. 1999.
- [124] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Jan. 1979.
- [125] D. Marín and A. Aquino and M. E. Gegúndez-Arias and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on Medical Imaging*, vol. 30(1), pp. 146–158, Jan. 2011.
- [126] T. Walter, P. Massin, A. Erginay, R. Ordonez, C. Jeulin, and J. C. Klein, "Automatic detection of microaneurysms in color fundus images," *Medical Image Analysis*, vol. 11(6), no. 2, pp. 555–566, Dec. 2007.
- [127] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," *Computers in Biology and Medicine*, vol. 40), pp. 438–445, 2010.
- [128] R. L. Cannon, J. V. Dave, and J. C. Bezdek, "Efficient implementation of the fuzzy

- c-means clustering algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 2, pp. 248–255, 1986.
- [129] B. Likar, M. Viergever, and F. Pernus, “Retrospective correction of MR intensity inhomogeneity by information minimization,” *IEEE Transactions on Medical Imaging*, vol. 20, pp. 1398–1410, Dec. 2001.
- [130] H. Knutsson and C. F. Westin, “Normalized and differential convolution: Methods for interpolation and filtering of incomplete and uncertain data,” in *Proceedings of Computer Vision and Pattern Recognition*, pp. 515–523, 1993.
- [131] M. Sto, H. Murakami, and M. Kasugaa, “Automatic facial feature extraction,” in *Computer Vision and Graphics* (M. Viergever, K. Wojciechowski, B. Smolka, H. Palus, R. S. Kozera, W. Skarbek, and L. Noakes, eds.), vol. 32 of *Computational Imaging and Vision*, pp. 400–405, Springer Netherlands, 2006.
- [132] X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” in *Analysis and Modeling of Faces and Gestures (AMFG’07)*, vol. 4778 of *Lecture Notes in Computer Science*, pp. 168–182, 2007.
- [133] X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” *IEEE Transactions on Image Processing*, vol. 19, pp. 1635–1650, June 2010.
- [134] H. Knutsson and G. H. Granlund, “Texture analysis using two-dimensional quadrature filters,” in *In Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, pp. 206–213, 1983.
- [135] A. L. da Cunha, Z. Jianping, and M. N. Do, “The nonsubsampling contourlet transform: Theory, design, and applications,” *IEEE Transactions on Image Processing*, vol. 15, pp. 3089–3101, Oct. 2006.

- [136] A. S. Georgiades, P. L. Belhumeur, and D. J. Kriegman, "From few to many: Generative models for recognition under variable pose and illumination," in *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 277–284, 2000.
- [137] A. S. Georgiades, P. L. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 643–660, March 2001.
- [138] A. Petpon and S. Srisuk, "Illumination normalization for robust face recognition using discrete wavelet transform," in *Proceedings of the 6th international conference on Advances in visual computing - Volume Part III*, vol. 6455 of *ISVC'10*, (Berlin, Heidelberg), pp. 69–78, Springer-Verlag, 2010.
- [139] T. Chen, W. Yin, X.-S. Zhou, D. Comaniciu, and T. Huang, "Illumination normalization for face recognition and uneven background correction using total variation based image models," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 532–539, 2005.
- [140] X. Xie, W. Zheng, J. Lai, and P. Yuen, "Face illumination normalization on large and small scale features," in *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR08, pp. 1–8, 2008.
- [141] Y. Z. Goh, A. B. Teoh, and M. K. O. Goh, "Wavelet based illumination invariant preprocessing in face recognition," in *Proceedings of the 2008 Congress on Image and Signal Processing, Vol. 3 - Volume 03*, CISP '08, (Washington, DC, USA), pp. 421–425, IEEE Computer Society, 2008.
- [142] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult illumination conditions," in *3rd international conference on Analysis and modelling of faces and gestures (2007)*, vol. 4778, pp. 168–182, Springer, 2007.
- [143] A. K. Jain, A. Ross, and S. Prabhakar, "Biometrics: A tool for information security,"

- IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 125–143, 2006.
- [144] R. M. Bolle, J. H. Connel, and N. K. Ratha, “Biometric perils and patches,” *Pattern Recognition*, vol. 35, pp. 2727–2738, 2002.
 - [145] A. K. Jain, R. M. Bolle, and S. Pankanti, *Biometrics: The Personal Identification in Networked Society*. Kluwer, Jan. 1999.
 - [146] A. Ross, A. K. Jain, and J. Reisman, “A hybrid fingerprint matcher,” *Pattern Recognition*, vol. 36, no. 7, pp. 1661–1673, 2003.
 - [147] N. Ratha, S. Chen, and A. K. Jain, “Adaptive flow orientation-based feature-extraction in fingerprint images,” *Pattern Recognition*, vol. 28, no. 11, pp. 1657–1672, 1995.
 - [148] N. Ratha, K. Karu, S. Chen, and A. K. Jain, “Real-time matching system for large fingerprint databases,” *IEEE Trans on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 799–813, 1996.
 - [149] A. K. Jain, L. Hong, S. Pankanti, and R. Bolle, “An identity-authentication system using fingerprints,” *Proceeding of the IEEE*, vol. 85, no. 9, pp. 1365–1388, 1997.
 - [150] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, “Fingerprint enhancement using STFT analysis,” *Pattern Recognition*, vol. 40, no. 1, pp. 198–211, 2007.
 - [151] K. Nilsson and J. Bigun, “Complex filters applied to fingerprint images detecting prominent symmetry points used for alignment,” in *Biometric Authentication*, pp. 39–47, 2002.
 - [152] M. Kawagoe and A. Tojo, “Fingerprint pattern classification,” *Pattern Recognition*, vol. 17, no. 3, pp. 295–303, 1984.
 - [153] K. Karu and A. K. Jain, “Fingerprint classification,” *Pattern Recognition*, vol. 29, no. 3, pp. 389–403, 1996.

- [154] A. K. Jain, S. Prabhakar, and L. Hong, “A multichannel approach to fingerprint classification,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 348–358, 1999.
- [155] A. M. Bazen and S. H. Gerez, “Systematic methods for the computation of the directional fields and singular points of fingerprints,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 905–919, 2002.
- [156] S. Prabhakar, *Fingerprint Classification and Matching Using a Filterbank*. PhD thesis, Michigan State University, 2001.
- [157] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.