#### University of Windsor Scholarship at UWindsor

**Electronic Theses and Dissertations** 

2010

## Biometric Applications Based on Multiresolution Analysis Tools

Abdul Mohammed University of Windsor

Follow this and additional works at: http://scholar.uwindsor.ca/etd

#### **Recommended** Citation

Mohammed, Abdul, "Biometric Applications Based on Multiresolution Analysis Tools" (2010). *Electronic Theses and Dissertations*. Paper 436.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

# BIOMETRIC APPLICATIONS BASED ON MULTIRESOLUTION ANALYSIS TOOLS

by

Abdul Adeel Mohammed

A Dissertation Submitted to the Faculty of Graduate Studies through the Department of Electrical and Computer Engineering in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy at the University of Windsor

Windsor, Ontario, Canada

2010

© 2010 Abdul Adeel Mohammed

Biometric Applications Based on Multiresolution Analysis Tools

by

Abdul Adeel Mohammed

APPROVED BY:

Dr. Ling Guan Ryerson University, Toronto, Canada

> Dr. Robin Gras School of Computer Science

Dr. Kemal Tepe Department of Electrical and Computer Engineering

Dr. Maher A. Sid-Ahmed Department of Electrical and Computer Engineering

Dr. Q. M. Jonathan Wu Department of Electrical and Computer Engineering

> Dr. James W. Gauld, Chair of Defense Faculty of Graduate Studies

> > 15th September, 2010

## **Co-Authorship Declaration**

I hereby declare that this dissertation incorporates the material that is the result of a joint research, as follows:

This dissertation incorporates the outcome of a joint research undertaken in collaboration with Dr. Rashid Minhas under the supervision of Dr. Jonathan Wu. The collaboration is covered in Chapters 3-4 of the dissertation. In all cases, the key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contributions of co-authors were primarily through the provision of proof reading, software design and reviewing the research papers regarding the technical content.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my dissertation, and have obtained a written permission from each of the co-authors to include the above materials in my dissertation.

I certify that, with the above qualification, this dissertation, and the research to which it refers, is the product of my own work.

### **Declaration of Previous Publications**

This dissertation includes following papers that have been published in peer reviewed journals and conferences.

I certify that I have obtained a written permission from the copyright owner(s) to include the published material(s) in my dissertation. I certify that the material describes the work completed during my registration as a graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my dissertation neither infringes upon anyone's copyright nor violates any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my dissertation, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my dissertation.

I declare that this is a true copy of my dissertation, including any final revisions, as approved by my dissertation committee and the Graduate Studies office, and that this dissertation has not been submitted for a higher degree to any other University or Institution.

Chapter Publication title/full citation		Status
Chapter 3 A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Sid-		Published
	Ahmed, "A Novel Technique for Human Face Recogni-	
	tion using Nonlinear Curvelet Feature Subspace", Lec-	
	ture Notes in Computer Science, vol. 5627 (2009), pp.	
	512-521	
Chapter 4 A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Pu		Published
	Sid-Ahmed, "Fingerprint Image Compression Standard	
	Based on Wave Atoms Decomposition and Self Orga-	
	nizing Feature Map", Proceedings of Electro/Iformation	
	Technology (2009), pp. 367-372	
Chapter 4	A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A.	Published
	Sid-Ahmed, "A Generic Fingerprint Image Compression	
	Technique based on Wave Atoms Decomposition"; Pro-	
	ceedings of International Conference on Image Processing	
	(2009), pp. 2837-2840	
Chapter 4 A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Sid-		Published
	Ahmed, "An Efficient Fingerprint Image Compression	
	Standard Technique Based on Wave Atoms Decompo-	
	sition and Multistage Vector Quantization"; Journal of	
	Integrated Compute Aided Engineering, vol. 17, issue 1	
	(2010), pp. 29-40	
Chapter 5	A. A. Mohammed, Q. M. J. Wu, M. A. Sid-Ahmed, "Ap-	Published
	plication of Wave Atoms Decomposition and Extreme	
	Learning Machine for Fingerprint Classification", Lec-	
	ture Notes in Computer Science vol. 6112 (2010), pp.	
	246-255	

### Abstract

This dissertation is dedicated to the development of new algorithms for biometric applications based on multiresolution analysis tools. Biometric is a unique, measurable characteristic of a human being that can be used to automatically recognize an individual or verify an individual's identity. Biometrics can measure physiological, behavioral, physical and chemical characteristics of an individual. Physiological characteristics are based on measurements derived from direct measurement of a part of human body, such as, face, fingerprint, iris, retina etc. We focussed our investigations to fingerprint and face recognition since these two biometric *modalities* are used in conjunction to obtain reliable identification by various border security and law enforcement agencies. We developed an efficient and robust human face recognition algorithm for potential law enforcement applications. A generic fingerprint compression algorithm based on state of the art multiresolution analysis tool to speed up data archiving and recognition was also proposed. Finally, we put forth a new fingerprint matching algorithm by generating an efficient set of fingerprint features to minimize false matches and improve identification accuracy.

Face recognition algorithms were proposed based on curvelet transform using kernel based principal component analysis and bidirectional two-dimensional principal component analysis and numerous experiments were performed using popular human face databases. Significant improvements in recognition accuracy were achieved and the proposed methods drastically outperformed conventional face recognition systems that employed linear one-dimensional principal component analysis. Compression schemes based on wave atoms decomposition were proposed and major improvements in peak signal to noise ratio were obtained in comparison to Federal Bureau of Investigation's wavelet scalar quantization scheme. Improved performance was more pronounced and distinct at higher compression ratios. Finally, a fingerprint matching algorithm based on wave atoms decomposition, bidirectional two dimensional principal component analysis and extreme learning machine was proposed and noteworthy improvements in accuracy were realized.

to my

#### MOTHER and FATHER

with love

## Acknowledgements

Firstly, I would like to thank the almighty Allah for his numerous blessings and bounties bestowed upon me. I am sincerely grateful to everyone who helped me during the course of my stay at University of Windsor. I would like to express my deep-felt appreciation to my advisor, Dr. Jonathan Wu for his advice, encouragement, enduring patience and support. He has always been open to suggestions, and offered me plenty of time and space to think about the ideas and research. His patience during initial years of PhD made me realize the fact that being an advisor demands a lot of endurance and tolerance.

I also wish to thank the other members of my committee, Dr. Maher Sid-Ahmed, Dr. Kemal Tepe, Dr. Robin Gras and Dr. Ling Guan. Their suggestions, comments and additional guidance were invaluable to the completion of this work. In difficult times, the support and important discussions with Rashid Minhas and Muhammad Iqbal rejuvenated my spirits.

Additionally, I want to thank staff and my group members at department of ECE for all their hard work and dedication, providing the means to achieve this important academic goal. Finally, I must appreciate my brother, wife and kids for putting up with me during the development of this work with continuous love and support.

## Table of Contents

				Ρ	age
C	o-Au	${\rm thorsh}$	ip Declaration	•	iii
D	eclar	ation o	of Previous Publications	•	iv
A	bstra	ct.		•	vi
D	edica	tion .		•	viii
A	cknov	wledge	ements	•	ix
Li	st of	Table	s	•	xiii
Li	st of	Figur	es	•	XV
Li	st of	Abbre	eviations	•	xviii
List of Symbols			XX		
1	Ove	erview	of Biometrics	•	1
	1.1	Outlin	e of a Biometric System		2
	1.2	Verific	ation vs. Identification		5
	1.3	Resear	ch Objectives		6
	1.4	Organ	ization of Dissertation		9
<b>2</b>	Mu	ltireso	lution Analysis Tools	•	12
	2.1	Wavel	et Transform		13
		2.1.1	Concept of Multiresolution Analysis		14
		2.1.2	Discrete Wavelet Transform		15
		2.1.3	Wavelet Filter		19
		2.1.4	Biorthogonal Spline Filter		21

		2.1.5 Comparison of filter properties
	2.2	Limitations of Wavelet Pyramid
	2.3	Curvelet Transform
		2.3.1 Continuous Time Curvelet Transform
		2.3.2 Fast Discrete Curvelet Transform
	2.4	Wave Atoms Decomposition
		2.4.1 1D Discrete Wave Atoms Decomposition
		2.4.2 2D Discrete Wave Atoms Decomposition
3	Hu	man Face Recognition 39
	3.1	Introduction
	3.2	Challenges in Face Recognition
		3.2.1 Databases
	3.3	Literature Review
	3.4	Kernel Principal Component Analysis
	3.5	Face Recognition using Curvelets and KPCA
		3.5.1 Experimental Results
	3.6	Bidirectional Two-Dimensional Principal Component Analysis 53
	3.7	Extreme Learning Machine
	3.8	Face Recognition using Multi-Dimensional PCA and ELM 58
		3.8.1 Results and Discussion
4	Fin	gerprint Compression
	4.1	Need for Image Compression
	4.2	Compression techniques
		4.2.1 Transform based compression
	4.3	Literature Review
	4.4	Vector Quantization
	4.5	Fingerprint Compression using Wave Atoms and Vector Quantization 78
		4.5.1 Results and Discussion

	4.6	Mathematical Morphology Operators	82
	4.7	Multistage Vector Quantization	83
	4.8	Fingerprint Compression using Mathematical Morphology and Multi-	
		stage Vector Quantization	86
		4.8.1 Comparative Results	89
<b>5</b>	Fin	gerprint Matching	94
	5.1	Introduction	94
	5.2	Typical Challenges	96
		5.2.1 FVC Datasets	97
	5.3	Fingerprint Feature Representation	98
	5.4	Literature Review	102
	5.5	Proposed Fingerprint Matching Algorithm	105
	5.6	Experimental Results	107
6	Co	nclusions and Future Research	111
	6.1	Summary of Contributions	115
R	References		
$\mathbf{V}$	ita A	Auctoris	133

## List of Tables

Property comparison of different wavelet filters	24
Kernels and their associated mathematical functions	49
Overview of KPCA Based Face Recognition	51
Average recognition rates (%) using Curvelet+PCA [59] and our KPCA	
based recognition scheme	52
Mean standard deviation of curvelet subbands in various databases .	60
Comparative Accuracy for YALE and ORL face database	62
Average recognition rates (%) for Sheffield and FERET database $~$	63
Average recognition rates (%) for ORL and GTech database $\ . \ . \ .$	64
Average recognition rates (%) and time complexity for Faces94 database	64
Average recognition rates (%) for JAFFE database at varying number	
of neurons	65
Average recognition rates (%) for FERET database using 2DPCA and	
B2DPCA	67
Average recognition rates (%) for JAFFE database using ELM and	
traditional BP based neural network	68
Filter coefficients for 9-7 wavelet filter	74
Bit rate vs. PSNR for FBI's WSQ and proposed method	81
Bit rate vs. PSNR for FBI's WSQ and the proposed methods	91
	Property comparison of different wavelet filters

5.1	Image details of each of the four FVC2000 database	98
5.2	Image details of each of the four FVC2002 database	99
5.3	Image details of each of the four FVC2004 database	100
5.4	Outline of the proposed fingerprint matching algorithm	106
5.5	Comparative results for various methods	108

# List of Figures

1.1	Biometric modalities that can be used for person identification $[2]$ .	3
1.2	Enrollment, verification and identification stages of a fingerprint sys-	
	tem [2] $\ldots$	7
1.3	Biometric revenues by technology, 2009 (AFIS revenues are not in-	
	cluded, source: International Biometric Group)	8
1.4	A fingerprint card containing rolls (1-10), slaps of left and right hand	
	(11,14), flat left and right thumb impression (12,13) [5] $\ldots \ldots \ldots$	10
2.1	Sinewave	16
2.2	Wavelet	16
2.3	1D DWT implementation using subband coding	17
2.4	Multilevel decomposition of 1D signal	18
2.5	2D DWT implementation using subband coding	18
2.6	Wavelet packet transform: Generalization of wavelet transform	19
2.7	Three scale wavelet decomposition	20
2.8	Filter structure and associating wavelets	23
2.9	Wavelets vs. Curvelets [17]	26
2.10	Overview of curvelet transform [9]	27
2.11	Spatial decomposition of a single subband [9]	29
2.12	Space frequency tiling in curvelet domain [20]	32
2.13	Wrapping a segment around the origin [22]	33

2.14	Comparison of different wave packets architectures with respect to mul-	
	tiscale nature and directional selectivity [24]	35
2.15	Wave atoms tiling in space and frequency [24]	35
3.1	Sample images of a subject from FERET database	41
3.2	Sample images from Faces94 database	42
3.3	Sample images of a subject from JAFFE database	42
3.4	Sample images of a subject from Georgia Tech database	43
3.5	Sample images from Sheffield face database	43
3.6	Sample images from ORL face database	44
3.7	Sample images of a subject from Yale face database	44
3.8	Sample images of a subject from AR face database	45
3.9	Block diagram of proposed B2DPCA algorithm	55
3.10	Architecture of an extreme learning machine classifier	58
3.11	Schematic diagram of proposed face recognition algorithm $\ldots \ldots$	59
3.12	(a) Original FERET image, (b) Curvelet subband at scale=1, (c-j)	
	Curvelet subbands at scale=2 and 8 angular orientations $\ldots \ldots \ldots$	60
3.13	Average recognition rate $(y$ -axis) vs. number of principal components	
	$(x-axis)$ for AR face database $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	65
3.14	Average recognition rate $(y$ -axis) vs. number of principal components	
	(x-axis) for ORL database at varying prototypes	66
3.15	Average recognition rate $(\%)$ for FERET database using kNN and ELM	67
3.16	Mean square error vs. number of epochs for JAFFE face database	69
4.1	FBI's 64-subband structure with a 5-level wavelet decomposition [87]	75
4.2	5-level decomposition subband structure	76
4.3	Block diagram of proposed fingerprint compression algorithm (WAVQ)	80
4.4	(a) Original fingerprint image (b) Compressed image at 0.25 bpp	81
4.5	Square structuring element	83

4.6	Dilation of binary map using a square structuring element $\ldots$ .	83
4.7	Erosion of binary map using a square structuring element	83
4.8	Block diagram of an MSVQ encoder	85
4.9	Block diagram of an MSVQ decoder	86
4.10	Overview of the proposed MSVQ based compression algorithm	87
4.11	Performance comparison of various methods at CR = 12.9:1. (a) Orig-	
	inal fingerprint image (b) Compressed image using FBI's WSQ (c)	
	Compressed image using wave atoms decomposition and VQ (d) Com-	
	pressed image using the proposed MSVQ based method	90
4.12	Original and reconstructed images using our proposed method (a-b)	
	tented arch type image (c-d) whorl type image	91
4.13	PSNR analysis for tented arch type fingerprint image using haar wavelets,	
	haar based contourlets and wave atoms	92
4.14	PSNR analysis for tented arch type fingerprint image using different	
	contourlets and wave atoms	92
5.1	Fingerprint image with the most common ridge characteristics $\ldots$ .	95
5.2	(a-b) Impressions from the same finger look significantly different (large	
	intra-class variation). (c-d) Impressions from different fingers look sim-	
	ilar to an untrained eye (small interclass variation) [119] $\ldots$ .	97
5.3	Sample fingerprint images from FVC 2000 dataset	99
5.4	Sample fingerprint images from FVC 2002 dataset	100
5.5	Sample fingerprint images from FVC 2002 dataset	101
5.6	Fingerprint classes with marked core (red) and delta (blue) points [121]	102
5.7	Recognition accuracy for DB1 database	109
5.8	Recognition accuracy for DB2 database	109
5.9	Recognition accuracy for DB3 database	110
5.10	Recognition accuracy for DB4 database	110

## List of Abbreviations

AFIS	Automated Fingerprint Identification System
$\mathbf{STFT}$	Short Time Fourier Transform
$\mathbf{FFT}$	Fast Fourier Transform
DCT	Discrete Cosine Transform
JPEG	Joint Picture Expert Group
JPEG-LS	Lossless JPEG
DWT	Discrete Wavelet Transform
FIR	Finite Impulse Response
$\mathbf{QMF}$	Quadrature Mirror Filter
DQMF	Dual Quadrature Mirror Filter
FDCT	Fast Discrete Curvelet Transform
PCA	Principal Component Analysis
KPCA	Kernel Principal Component Analysis
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
2DPCA	Two Dimensional Principal Component Analysis
B2DPCA	Bidirectional Two Dimensional Principal Component Analysis
$\mathbf{VQ}$	Vector Quantization
MSVQ	Multi Stage Vector Quantization
$\mathbf{SVM}$	Support Vector Machines

Evolutionary Pursuit
Elastic Bunch Group Matching
Kernel Associative Memory
k Nearest Neighbor
Average Recognition Rate
Dots Per Inch
Bits Per Pixel
Portable Network Graphics
Peak Signal to Noise Ratio
Integrated Services Digital Network
Wavelet Scalar Quantization
Federal Bureau of Investigation
Region of Interest
Fingerprint Verification Competition
Difference of Gaussian
Directional Filter Bank
Generalized Hough Transform
Feedforward Neural Network
Back Propagation
Extreme Learning Machine
Radial Basis Function Neural Network
Ellipsoidal Basis Function Neural Network

# List of Symbols

$\Psi$	Mother wavelet
$\psi$	Wavelet basis function
h	Low-pass decomposition filter (scaling filter)
g	High-pass decomposition filter (wavelet filter)
$ ilde{h}$	Low-pass reconstruction filter
${ ilde g}$	High-pass reconstruction filter
α	Multiscale nature of the transform
$\beta$	Wave packet's directional selectivity
$\dot{x}_i$	Input sample $i$
$\dot{t}_i$	Labeled output $i$
$ ilde{\xi}(.)$	Nonlinear activation function
$\check{w}_i$	Weight vector between input and hidden layer
$\check{\gamma}_i$	Weight vector between hidden and output layer
$ ilde{\delta}$	Hidden layer output matrix
$\mathbb{L}$	Number of hidden neurons
IN	Number of output neurons
$\check{b}_i$	Hidden layer bias term
$\mathbb{E}$	Error function to be minimized
$\mathbf{sign}()$	Signum function

$\bar{C}$	PCA covariance matrix
$\bar{\mathbb{C}}$	KPCA covariance matrix
$ar{\phi}(.)$	KPCA mapping function
$\mathbb{K}$	KPCA Gram matrix
V	KPCA eigenvector
$tr(\mathbb{S}_{\check{x}})$	Trace of the covariance matrix $\mathbb{S}_{\check{x}}$
$J(\mathbb{X})$	Total scatter criterion

### Chapter 1

## **Overview of Biometrics**

Biometric refers to the process of establishing an individual's identity based on physical, physiological, chemical or behavioral attributes of a person. The relevance of biometrics in the 21<sup>st</sup> century is dictated by the need for large-scale identity management systems that predominantly rely on accurate determination of an individual's identity. Examples of these applications include sharing networked computer resources, granting access to nuclear facilities, performing remote financial transactions. The advent of web-based services (e.g., online banking, online shopping) and the deployment of decentralized customer service centers (e.g., credit cards) have further emphasized the need for consistent identity management systems that can accommodate a large number of individuals.

The most vital task in an identity management system is the verification of an individual's claimed identity. An individual's authentication is extremely necessary for a variety of reasons but the primary intention, is to prevent unauthorized users from accessing protected/classified information. Traditional person identification methods include knowledge-based (e.g., passwords) and token-based (e.g., ID cards) mechanisms, but these surrogate representations can be easily lost, shared, stolen and manipulated thereby jeopardizing the intended security. Biometrics offer a natural and reliable solution to several aspects of identity management through a complete and/or partial automated system to recognize individuals based on their biological characteristics [1]. Biometrics enable us to establish an identity based on *who you* are, rather than by *what you possess*, such as an ID card, or *what you remember*, such as a password or a security question. In some applications, biometrics may be used to supplement ID cards and passwords thereby imparting an additional level of security.

Biometric systems use a variety of physical or behavioral characteristics that include fingerprint, face, hand/finger geometry, iris, retina, signature, palmprint, voice pattern, gait, ear, hand vein, odor or the DNA information to establish an individual's identity [3], [4]. These characteristics are commonly referred as *modalities, traits* and *indicators* in literature. Biometric systems offer multiple benefits over traditional security methods as they cannot be easily lost, stolen or shared. Besides bolstering security, biometric systems also add to user's convenience by alleviating the need to remember passwords.

#### 1.1 Outline of a Biometric System

A biometric system is typically a pattern recognition system that acquires biometric data from an individual, extracts a set of salient features, compares the feature set against the feature sets stored in the database, and executes an action based on the results of the comparison. Therefore, a generic biometric system encapsulates four main modules: a sensor module; a quality assessment and feature extraction module; a matching module; and a database module. Each of these modules is briefly described below.

1. Sensor Module. A suitable biometric reader or scanner is required to acquire the raw biometric data of an individual. To obtain face images a digital still camera or a video camera can be used where as for fingerprint images, an optical fingerprint sensor may be used to image the friction ridge structure of



Figure 1.1: Biometric modalities that can be used for person identification [2]

the fingertip. The sensor module determines the human machine interface and is, therefore, pivotal to the performance of any biometric system. A poorly designed or noisy interface can result in a low matching accuracy. Since most biometric *modalities* are acquired as images (except voice and odor), the quality of the raw data is to a great extent impacted by the characteristics of the camera technology used.

- 2. Quality Assessment and Feature Extraction. The quality of biometric data sensed by the sensor is first assessed in order to determine whether it meets the minimum standards required for the desired application. Typically, the acquired data is enhanced in order to improve its quality. However, in some cases, the quality of the data is poor and the user is asked to present the biometric data again. In some situations it is indispensable that data is acquired at varying displacements, orientations, atmospheric conditions, noise levels, image acquisition sessions etc. in order to generate a dataset that takes into account different challenges encountered in real world applications. The biometric data is then processed to extract a set of salient discriminatory features.
- 3. Matching and Decision Making. The extracted features are compared against the stored templates in order to generate match scores. The match score is moderated by the quality of the presented biometric data. The matcher module also encapsulates a decision making module that either validates a claimed identity or provides a ranking of the enrolled identities in order to identify an individual. The decision making module is dictated by the biometric system architecture and the desired accuracy rates.
- 4. System Database Module. During the enrollment process, the extracted features are stored in the database along with some biographic information (such as name, address, birth date etc.) that distinguishes each individual. The data captured during the enrollment process may or may not be supervised by a

human depending on the application. The user template is either extracted from a single biometric sample or generated by processing multiple samples. Thus, in some fingerprint recognition systems the minutiae template is extracted after mosaicing multiple samples of the same finger. Some systems, such as face recognition systems store multiple templates in order to account for large intraclass variations associated with each user. Depending on the application, the template is either stored in the central biometric system database or recorded on a token (e.g., smart card) issued to the individual.

### 1.2 Verification vs. Identification

In literature the terms verification and identification have been used interchangeably, but there is some tangible distinction that sets them apart. Depending on the specific application, a biometric system may operate as a verification module or an identification module. These two modules are briefly explained below in the following section. The basic processes involved in a fingerprint enrollment and recognition (verification and identification) module are depicted in Figure 1.2.

In the verification mode, the system establishes an individual's identity by comparing the captured biometric *trait* with his/her own biometric template(s) stored in the system database. In such a system, an individual who claims an identity is recognized via a PIN, a user name or a smart card, and the system conducts a one-to-one comparison to determine whether the claim is true or not. Verification is typically used for positive recognition and the aim is to prevent multiple people from using the same identity.

In the identification mode, the system recognizes an individual by searching the templates of all the users enrolled in the database. Therefore, the system conducts a one-to-many search to establish an individual's identity. The system, thus, fails if the subject was not already enrolled in the database. Identification is essentially employed for negative recognition where the system establishes whether the person is who he/she denies to be. The purpose of negative recognition is to prevent a single person from using multiple identities. Identification may also be used in positive recognition in some biometric applications.

### **1.3** Research Objectives

Establishing the identity of an individual with high confidence is becoming critical in a number of applications in our vastly interconnected society. Questions like "Is he/she really who they claim to be?", "Is this person authorized to use this facility?" or "Is he/she a fugitive wanted by the government?" are often posed in a variety of scenarios ranging from issuing a driver's licence to gaining entry into a country. The need for reliable user authentication techniques has increased in the wake of increased security concerns, and rapid advancements in the field of networking, communication and mobility. Thus, biometrics are being increasingly incorporated in several different applications, categorized into three main groups.

- Commercial: Computer network login, electronic data security, e-commerce, Internet access, ATM or credit card use, physical access control, mobile phone, PDA, medical records management, distance learning, etc.
- *Government*: Identity cards, managing inmates in a correctional facility, driver's license, welfare-disbursement, border control, passport control, etc.
- *Forensics*: Corpse identification, criminal investigation, parenthood determination, etc.

In the context of this work we will focus our attention to two biometric *traits*, namely, faces and fingerprints. These are the two most important biometric *modalities* employed in various commercial, government and law enforcement applications.



Figure 1.2: Enrollment, verification and identification stages of a fingerprint system [2]

These biometrics are used in conjunction by various border security and law enforcement agencies. Researchers are using these two *traits* in combination with other biometrics to achieve improved recognition. For example at United Stages land, water and air crossings all visitors are imaged and their fingerprints acquired, these biometric *indicators* are stored and matched against a criminal database to ensure that criminals are barred from entering into the country and all fugitives arrested. Figure 1.3 shows the percentage of generated revenues by different biometric technologies.



Figure 1.3: Biometric revenues by technology, 2009 (AFIS revenues are not included, source: International Biometric Group)

To acquire face images, cheap cameras and image capture devices offer sufficient image resolution. Images generated using such image acquisition devices are often stored using JPEG or similar compression format, therefore, image size is usually small and no additional compression is required. Face recognition is not as accurate as other biometrics, thus, fingerprints are used in conjunction with faces to enhance accuracy. Fingerprints are often sensed using sophisticated and expensive image acquisition devices that generate high resolution images. A single fingerprint card contains 14 different images: 10 rolled impression of each finger, duplicate (flat) impression of thumb and simultaneous impression (slap) of all fingers together (see Figure 1.4). Rolled fingerprints are captured by rolling the finger (tip to the first joint) from "nail-to-nail" on the sensing surface. It provides the largest fingerprint area and contains about 80 minutiae, on average, thus facilitating accurate classification and matching. Flat fingerprints are acquired by pressing the thumb against the sensing surface, whereas a slap fingerprint is achieved by pressing the four fingers (index, middle, ring and little) simultaneously against the sensing surface. Fingerprint images are digitized at a resolution of 500 dpi with 256 gray levels which entails approximately 10 MB of storage. Therefore, an efficient compression algorithm that retains detailed fingerprint minutiae, ridges and other fingerprint features is needed to minimize storage requirements, speed up data archiving and diminish transmission bandwidth.

#### 1.4 Organization of Dissertation

The remainder of this thesis consists of 5 chapters which are organized as follows:

- Chapter 2. Multiresolution Analysis Tools: Covers the mathematical properties of wavelets, curvelets and wave atoms decomposition. It also discusses several wavelets, including Haar, Daubechies and biorthogonal spline wavelets. Two curvelet transform architectures are detailed and an implementation of discrete digital wave atoms decomposition is also reviewed.
- Chapter 3. *Human Face Recognition:* This chapter includes a brief introduction of face recognition, survey of existing approaches, and the individual modules used in our face recognition system. A new technique for multidimensional principal component analysis is proposed and explained. The proposed



Figure 1.4: A fingerprint card containing rolls (1-10), slaps of left and right hand (11,14), flat left and right thumb impression (12,13) [5]

face recognition algorithms are detailed along with comparative results on wide ranging sets of face databases.

• Chapter 4. *Fingerprint Compression:* The chapter reviews some of the established techniques for fingerprint compression, the proposed fingerprint compression schemes are also described and experimental results compared with other state-of-the-art techniques.

- Chapter 5. *Fingerprint Matching:* Basic concepts of fingerprints, their associated challenges, feature representation and fingerprint verification datasets are reviewed. Our wave atoms based proposed fingerprint matching algorithm is discussed and experimental findings compared with wavelet based fingerprint matching methods.
- Chapter 6. *Conclusions and Future Research:* Some conclusions and considerations on how to enhance the work in the future are included.

### Chapter 2

### Multiresolution Analysis Tools

Fourier transform has been used as a principle tool for signal analysis since early 19th century. It was developed by French mathematician, J. Fourier, who showed that any periodic function can be expressed as a sum of periodic complex exponential function. Later, this idea was generalized to non-periodic functions, periodic and non-periodic discrete time signals. In frequency domain, Fourier transform constructs a sinusoidal basis to describe energy distribution of a signal. However, Fourier transform is not well suited to describe local changes in frequency since the frequency component has infinite time support i.e. time (spatial) information is lost and it is impossible to specify when a particular phenomenon took place. Most of the practical signals and images contain non-stationary signal components and capturing them is a crucial step in classification.

To alleviate the limitations of Fourier transform, the windowed Fourier transform (Short Time Fourier Transform - STFT) was proposed. STFT works by dividing the signal into small segments where each segment is assumed to be stationary. However, STFT has several limitations i.e. if we use an infinite length window; we obtain Fourier transform with perfect frequency resolution but no time information. On the other hand, to acquire a stationary sample we have to use a small enough window in which the signal is stationary. The narrower the window, the better is the time

resolution and the assumption of stationarity, but poorer the frequency resolution [6]. Therefore to strike a balance between the time resolution and frequency resolution we turn our focus to wavelet transform, which is based on multiresolution analysis.

### 2.1 Wavelet Transform

Wavelet transform is a widely accepted solution to overcome the shortcomings of the Fourier transform and STFT. In wavelet analysis the fully scalable wavelet solves the problem of time and frequency resolution. The flexible window is moved along the signal and the spectrum is calculated for every position. The process is repeated several times with a variable window size and the collection of time-frequency representation of the signal is obtained. In this manner, the big wavelets give an approximate image of the signal, while the smaller wavelets zoom in on the details. Therefore, wavelets adapt automatically to both the high-frequency and low-frequency components of a signal by varying the window size. The wavelet transform is well suited for nonstationary signals, brief signals and signals with interesting components at different scales [7]. Wavelets are dilated and translated versions of a single function  $\Psi$ , which is called mother wavelet.

$$\Psi_{a,b}x = |a|^{\frac{-1}{2}}\Psi\frac{(x-b)}{a}$$
(2.1)

where  $\Psi$  satisfies the condition

$$\int_{-\infty}^{+\infty} \Psi(t)dt = 0 \tag{2.2}$$

The basic idea of the wavelet transform is to represent any arbitrary function f as a decomposition of wavelet basis or write f as an integral over a and b. Where a is the scale parameter and b is the position parameter.

When dealing with sampled data that is discrete in time we need to have a discrete representation of time and frequency, which is called discrete wavelet transform. We will briefly discuss the concept of multiresolution analysis before we discuss about the discrete wavelet transform.

#### 2.1.1 Concept of Multiresolution Analysis

A signal/image can be viewed as combination of a smooth background and fluctuations(fine details). The distinction between the smooth part and the detail part of a signal is determined by the resolution. Image detail at one resolution will act as a smooth background at higher resolution. At a given resolution, a signal is approximated by ignoring all fluctuations below that scale. We can progressively increase the resolution; at each stage of the increase in resolution finer details are added to the coarser description, thus providing a successively better approximation of the signal.

A function f(t) at a resolution level j is denoted by  $f_j(t)$  and the details are denoted by  $d_j(t)$ . At the next higher resolution level j + 1; the new approximation to  $f_j(t)$  is

$$f_{j+1}(t) = f_j(t) + d_j(t)$$
(2.3)

The original function is recovered as the resolution approaches to infinity.

$$f(t) = f_j(t) + \sum_{k=j}^{\infty} d_k(t)$$
 (2.4)

Multiresolution analysis involves decomposition of the function space into a sequence of subspaces  $V_j$ . The subspace  $V_j$  is contained in all the higher subspaces. If the approximation of f(t) at a level j is denoted by  $f_j(t)$  then  $f_j(t) \in V_j$ . Since information at resolution level j is a part of information at a higher resolution level j + 1, mathematically  $V_j \in V_{j+1}(t)$  for all j. We can therefore decompose our subspaces accordingly as

$$V_{j+1} = V_j \oplus W_j \tag{2.5}$$
Where  $W_j$  is the detail space at a resolution level j and  $V_j$  is the approximation at resolution level j. The space V is decomposed in order to obtain

$$V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1} = \dots = W_j \oplus W_{j-1} \oplus W_{j-2} \oplus \dots \oplus W_0 \oplus V_0 \quad (2.6)$$

### 2.1.2 Discrete Wavelet Transform

Wavelet analysis is also based on a decomposition of a signal using an orthonormal family of basis functions. A wavelet has its energy concentrated in time and is well suited for the analysis of transient, time-varying signals. A wavelet expansion is defined by a two-parameter family of functions

$$f(t) = \sum_{j} \sum_{k} a_{j,k} \psi_{j,k}(t)$$
 (2.7)

Where j and k are integers and the function  $\psi_{j,k}(t)$  is the wavelet expansion function which form an orthogonal basis. The two parameter coefficients  $a_{j,k}(t)$  are the Discrete Wavelet Transform (DWT) coefficients. The DWT coefficients  $a_{j,k}(t)$  are obtained using the following formula

$$a_{j,k} = \int f(t)\psi_{j,k}(t)dt \qquad (2.8)$$

The wavelet basis functions are a two-parameter family of functions that are related to the function  $\Psi(t)$ , the mother wavelet by

$$\psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k) \tag{2.9}$$

k represents translation and j is the dilation parameter. Therefore wavelet basis functions are obtained from a single wavelet by dilating and translating the single mother wavelet  $\Psi(t)$ . The concept of dilation and translation allows the wavelet transform to be localized in both time and frequency (scale) domain. Wavelet reveals aspects of data that other transform techniques overlook i.e. trends, breakdown points and discontinuities. By analyzing the sine wave and wavelets depicted in Figure 2.1 and Figure 2.2 respectively, we can clearly state that signals with sharp changes and peaks will be better analyzed with an irregularly shaped wavelet rather than with a smooth sinusoid.

Figure 2.1: Sinewave



Figure 2.2: Wavelet

DWT is implemented using the Mallat algorithm [8] known as two-channel subband coder to obtain the discrete wavelet transform coefficients. A pair of Finite Impulse Response (FIR) quadrature mirror filters known as scaling filter and wavelet filter are used. The scaling filter h is a low-pass filter and the wavelet filter g is a highpass filter.  $\tilde{h}$  is the low-pass reconstruction filter and  $\tilde{g}$  is the high-pass reconstruction filter. Both g and h are related by the following equation.

$$g_n = (-1)^n h_{N-1-n}, \longrightarrow n = 0, 1, 2, \dots N - 1$$
 (2.10)

Filter implementation of DWT using a two channel subband coder is shown in Figure 2.3. Images are decomposed using 2D DWT which is applied by means of separability approach along its rows and columns alternatively resulting into four smaller subsets. The subband coder of Figure 2.3 is generalized for 2D signals as shown in Figure 2.5.



Figure 2.3: 1D DWT implementation using subband coding

In wavelet analysis, a 1-Dimensional signal is split into approximation and detail components. The approximation component is recursively decomposed into second level approximation and detail coefficients and this process is repeated for the number of required decomposition levels. A level 3 wavelet decomposition of a signal is illustrated in Figure 2.4, therefore for n level of decomposition there exists n + 1 possible ways to decompose or encode a signal.



Figure 2.4: Multilevel decomposition of 1D signal



Figure 2.5: 2D DWT implementation using subband coding

Wavelet packet analysis is an extension of wavelet transform as both the approximation and detail coefficients are recursively decomposed at each level of decomposition. This results in an increased range of possibilities for signal analysis. Wavelet packet decomposition tree is as shown in Figure 2.6.



Figure 2.6: Wavelet packet transform: Generalization of wavelet transform

Figure 2.7 illustrates three levels of wavelet decomposition wherein HL1, LH1 and HH1 represent the finest/detail coefficient of the original image, HL2, LH2 and HH2 represent the finest/detail coefficient of sub-band LL1 and similarly HL3, LH3 and HH3 represent the finest/detail coefficient of sub-band LL2. LL3 is the lowest frequency term that represents all the coarser levels.

### 2.1.3 Wavelet Filter

Wavelet bases are constructed with certain desired properties and quite a bit of freedom is exercised in choosing the wavelet function to generate a particular wavelet basis. Specific choice and method of construction of wavelet basis entirely depends on the requirements and motivation for its construction. There are two important classes

LL3 LH3	HL3 HH3	HL2	τττ 1
LH 2		HH 2	HLI
LH1			HH1

Figure 2.7: Three scale wavelet decomposition

of compactly supported wavelet bases, namely the compactly supported orthogonal and the biorthogonal wavelet bases. These wavelet bases give rise to FIR linear phase and FIR subband filtering schemes. Common examples of compactly supported orthogonal basis wavelets are the Haar wavelet basis and the Daubechies wavelet basis. In this section, we will briefly discuss about the Biorthogonal spline wavelets and finally compare the wavelet properties of Haar, Daubechies and Biorthogonal spline wavelets.

### 2.1.4 Biorthogonal Spline Filter

Most of the images are smooth and when dealing with images it is required that a wavelet filter should not deteriorate the smoothness of the image. Biorthogonal spline wavelets are a class of wavelet filters that use a smooth mother wavelet for image analysis. In addition to a smooth mother wavelet it is also required that mother wavelet is symmetric so that the corresponding wavelet transform could be implemented using mirror boundary conditions that reduce boundary artifacts. Except for the trivial case of Haar wavelets none of the wavelet filters are both symmetric and orthogonal.

Therefore to achieve symmetric property we relax the orthogonality constraint and construct a biorthogonal basis. Decomposition of an image is obtained using the following equation.

$$c_{m,n}(f) = \sum_{k} g_{2n-k} a_{m-1,k}(f)$$
(2.11)

$$a_{m,n}(f) = \sum_{k} h_{2n-k} a_{m-1,k}(f)$$
(2.12)

Where  $g_l = (-1)^l h_{-l+1}$  and  $h_n = 2^{1/2} \int \psi(x-n)\psi(2)dx$ . The image is reconstructed using the equation below:

$$a_{m-1,f}(f) = \sum_{n} [\tilde{h}_{2n-l}a_{m,n}(f) + \tilde{g}_{2n-l}c_{m,n}(f)]$$
(2.13)

 $\tilde{h}, \, \tilde{g}$  are different from h and g and the relationship between them is given by the following equation

$$\tilde{g}_n = (-1)^n h_{-n+1} \tag{2.14}$$

$$g_n = (-1)^n \tilde{h}_{-n+1} \tag{2.15}$$

$$\sum_{n} = h_n \tilde{h}_{n+2k} = \delta_{k,0} \tag{2.16}$$

Define

$$\phi(x) = \sum_{n} h_n \phi(x - 2n) \tag{2.17}$$

$$\tilde{\phi}(x) = \sum_{k} \tilde{h}_n \tilde{\phi}(x - 2n)$$
(2.18)

$$\psi(x) = \sum_{n} g_n \psi(x - 2n) \tag{2.19}$$

$$\tilde{\psi}(x) = \sum_{k} \tilde{g}_n \tilde{\phi}(x - 2n)$$
(2.20)

Therefore we can rewrite  $a_{m,n}(f)$  and  $c_{m,n}(f)$  as:

$$a_{m,n}(f) = 2^{-m/2} \int \phi_{m,n}(x) f(x) dx$$
(2.21)

$$c_{m,n}(f) = 2^{-m/2} \int \psi_{m,n}(x) f(x) dx$$
(2.22)

and the reconstruction equation thus becomes

$$f = \sum_{m} n < \psi_{m,n}, f > \tilde{\psi}_{m,n}$$
(2.23)

Figure 2.8 gives a relationship between filter structure and wavelet functions. For symmetric filters, the condition of exact reconstruction on h and  $\tilde{h}$  can be written as:

$$H(\xi) + \tilde{H}(\xi) + H(\xi + \pi) + \tilde{H}(\xi + \pi) = 1$$
(2.24)

Where

$$\tilde{H}(\xi) = 2^{-1/2} \sum_{n} \tilde{h}_{n} e^{-jn\xi}$$
(2.25)



Figure 2.8: Filter structure and associating wavelets

and

$$H(\xi) = 2^{-1/2} \sum_{n} h_n e^{-jn\xi}$$
(2.26)

### 2.1.5 Comparison of filter properties

A comparative study of wavelet filter properties of Haar wavelet, Daubechies wavelet and Biorthogonal spline wavelet is included in Table 2.1. Amongst the three wavelets only Haar and Daubechies wavelets possess orthogonality, which offers the following advantages:

- Scaling and Wavelet functions are same for both forward and inverse transform.
- Correlation in the signal between different subspaces is removed.

Haar wavelet is the simplest and the most fastest wavelet to implement but the major disadvantage of haar wavelet is its discontinuity, which makes it difficult to simulate a continuous signal. Daubechies invented the first continuous orthogonal compact support wavelet but this wavelet family is non-symmetric. The advantage of the wavelet possessing symmetric property is that the wavelet transform can be implemented using mirror boundary conditions that reduce boundary artifacts. Therefore Biorthogonal spline wavelet filters are the best available wavelets for image compres-

Property	Haar	Daubechies	Biorthogonal Spline
Explicit Function	Yes	No	Yes
Orthogonal	Yes	Yes	No
Symmetric	Yes	No	Yes
Continuous	No	Yes	Yes
Compact support	Yes	Yes	Yes
Maximum regularity(order L)	No	No	Yes
Shortest scaling function(order L)	Yes	No	Yes

Table 2.1: Property comparison of different wavelet filters

sion applications. The B-spline wavelets are smooth and since splines are piecewise polynomial they are easy to manipulate.

# 2.2 Limitations of Wavelet Pyramid

Wavelets have had a wide impact, both in theory and in practice, and especially in areas of data compression and signal restoration. The shrinking of wavelet coefficients proved to be a very powerful tool for statistical estimation, from both a theoretical and a practical standpoint. Therefore, wavelet based-coders have found wide applications in various data compression applications and have been included in JPEG-2000. Meyer portrayed wavelet theory as a unifying mathematical language for describing a set of connected ideas that arose in different areas. Researchers have proved that a lot of claims regarding the applicability of wavelets to image processing problems have been perhaps overstated and some of their limitations include:

• *Inefficient Representations*: From a theoretical viewpoint, wavelet series is not optimal for representing objects with discontinuities along curves.

- *Directional Limitations*: Wavelets have only a fixed number of directional elements, independent of scale.
- *Scale Limitations*: Wavelet pyramids are not anisotropic and contain elements at fixed scales/locations.

Wavelet transform is a well known multiresolution analysis tool capable of conveying accurate temporal and spatial information. Wavelet transform has been profusely used to address problems in; signal and speech processing, data compression, pattern recognition, image reconstruction, and biomedical engineering applications. Wavelets better represent objects with point singularities in 1D and 2D space but fail to deal with singularities along curves in 2D. Discontinuities in 2D are spatially distributed which leads to extensive interaction between discontinuities and wavelet expansion coefficients. Therefore wavelet representation does not offer sufficient sparseness for image analysis. Following wavelets, research community has witnessed intense efforts for development of better directional and decomposition tools, namely, curvelets [9], ridgelets [15] and contourlets [16].

# 2.3 Curvelet Transform

Curvelet transform [9] belongs to the family of multiresolution analysis tool that is designed and targeted to represent smooth objects with discontinuity along a general curve. Curvelet transform overcomes shortcomings of existing multiresolution analysis schemes and offers improved directional capacity to represent edges and other singularities along curves. Curvelets outperform wavelets in situations that require optimal sparse representation of objects with edges, representation of wave propagators, image reconstruction with missing data etc. Let us assume that a function fhas a discontinuity across a curve, and is otherwise smooth as shown in Figure 2.9. Approximating f from the best m-terms in the Fourier expansion at a specified error



Figure 2.9: Wavelets vs. Curvelets [17]

rate in wavelet domain would require  $O(m^{-1})$  terms, whereas a curvelet expansion demands only  $O(m^{-2})$  terms. Curvelet transform is a multiscale non-standard pyramid with numerous directions and positions at each length and scale. Curvelets offer anisotropic and a locally adaptive scaling unlike other pyramid schemes. Original implementation of curvelet transform is based on the combination of Ridgelets, Multiscale ridgelets and Bandpass filtering.

- *Ridgelets* analyze objects with discontinuities across straight lines.
- *Multiscale ridgelets* renormalize and transport data to a wide range of scales and locations.
- *Bandpass filtering* separates an object into a series of disjoint scales.

Analysis of curvelet transform based on its original implementation includes subband decomposition, smooth partitioning, renormalization and ridgelet analysis. An organizational outline of curvelet transform is also depicted in Figure 2.10.



Figure 2.10: Overview of curvelet transform [9]

1. Subband Decomposition. A bank of subband filters  $P_0$ ,  $(\Delta_s, s \ge 0)$  are defined and the object f is filtered into distinct subbands:

$$f \mapsto (P_0 f, \Delta_1 f, \Delta_2 f, \Delta_3 f, ...)$$

$$(2.27)$$

The different subbands contains details about  $2^{-2s}$  wide.

2. Smooth Partitioning. A collection of smooth windows  $w_Q(x_1, x_2)$  localized around dyadic squares are defined

$$Q = [k_1/2^s, (k_1+1)/2^s] \times [k_2/2^s, (k_2+1)/2^s]$$
(2.28)

Multiplication of a function by the corresponding window function generates a localized result, therefore different configuration of  $k_1$  and  $k_2$  with a fixed s produces a smooth dissection of the function into 'squares'. Smooth windowing dissection is applied to each of the filtered subbands.

$$\Delta_s f \mapsto (w_Q, \Delta_s f)_{Q \in \mathcal{Q}_s} \tag{2.29}$$

3. **Renormalization.** Renormalize each square generated in the previous stage to unit scale.

$$g_Q = (T_Q)^{-1}(w_Q \Delta_s f), \qquad Q \in \mathcal{Q}_s$$
(2.30)

where  $T_Q$  denotes the operator that transports and renormalizes f and is defined as:

$$(T_Q f)(x_1, x_2) = 2^s f(2^s x_1 - k_1, 2^s x_2 - k_2)$$
(2.31)

4. Ridgelet Analysis. Each 'square' is analyzed in the orthonormal ridgelet system with basis elements  $\rho_{\lambda}$ , thus making an orthobasis for  $L^2(\mathbf{R}^2)$ 

$$\alpha_{\mu} = \langle g_Q, \rho_{\lambda} \rangle, \qquad \mu = (Q, \lambda) \tag{2.32}$$

To better understand the spatial decomposition in curvelet transform; let us assume that an object f that exhibits an edge as shown in Figure 2.11. Each fine-scale subband filtered output  $\Delta_s f$  contains a map of the edge in f, widened to a width  $2^{-2s}$ as per the scale of subband filter operator. Each subband appears as a collection of smooth ridges, and when each subband is partitioned into 'squares', an empty square or a ridge fragment is observed. The ridge fragments are nearly straight at fine scales and thus form the desired input for ridgelet analysis.

The image is synthesized using a reverse process order that includes; *ridgelet syn*thesis, renormalization, smooth integration and subband recomposition, as described below.

1. **Ridgelet Synthesis.** Each 'square' is reconstructed using the orthonormal ridgelet system as:

$$g_Q = \sum_{\lambda} \alpha_{\lambda,Q} \rho_{\lambda} \tag{2.33}$$

2. **Renormalization.** Every individual 'square' reconstructed in the previous stage is renormalzied to its own proper square.

$$h_Q = (T_Q)g_Q, \qquad Q \in \mathcal{Q}_s \tag{2.34}$$



Figure 2.11: Spatial decomposition of a single subband [9]

3. **Smooth Integration.** Windowing dissection is reversed to each of the reconstructed windows.

$$\Delta_s f = \sum_{Q \in \mathcal{Q}_s} w_Q . h_Q \tag{2.35}$$

4. **Subband Recomposition.** The effect of the bank of subband filters in neutralized using the following recomposition formula.

$$f = P_0(P_0 f) + \sum_{s>0} \Delta_s(\Delta_s f) \tag{2.36}$$

### 2.3.1 Continuous Time Curvelet Transform

Since the introduction of curvelet transform researchers have developed numerous algorithmic strategies [18], [19], [20], [21] for its implementation based on its original architecture. Let us consider a 2D space, i.e.  $\Re^2$ , with a spatial variable x and a frequency-domain variable  $\omega$ , and let r and  $\theta$  represent polar coordinates in frequencydomain. W(r) and V(t) are radial and angular windows respectively. Both windows are smooth, nonnegative, real valued and supported by arguments  $r \in [1/2, 1]$  and  $t \in [-1, 1]$ . For  $j \geq j_0$ , frequency window  $U_j$  in Fourier domain is defined as,

$$U_j(r,\theta) = 2^{-3j/4} W(2^{-j}r) V(\frac{2^{\lfloor j/2 \rfloor}\theta}{2\pi}), \qquad (2.37)$$

where  $\lfloor j/2 \rfloor$  is the integral part of j/2. Thus, the support of  $U_j$  is a polar wedge defined by the support of W and V, applied with scale-dependent window widths in each direction. Windows W and V always obey the admissibility conditions as follows:

$$\sum_{j=-\infty}^{+\infty} W^2(2^{-j}r) = 1, r \in (3/4).$$
(2.38)

$$\sum_{l=-\infty}^{+\infty} V^2(t-l) = 1, t \in (-1/2, 1/2).$$
(2.39)

Curvelets are defined (as function of  $x = (x_1, x_2)$ ) at scale  $2^{-j}$ , orientation  $\theta_l$ , and position  $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 2^{-j}, k_2 2^{-j/2})$  by  $\varphi_{j,k,l}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)}))$ , where  $R_{\theta}$  is an orthogonal rotation matrix. A curvelet coefficient is simply computed by computing the inner product of an element  $f \in L^2(R^2)$  and a curvelet  $\varphi_{j,k,l}$ 

$$c_{j,k,l} = \langle f, \varphi_{j,k,l} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,k,l}} dx.$$
(2.40)

Curvelet transform also contains coarse scale elements similar to wavelet theory. For  $k_1, k_2 \in Z$ , we define a coarse level curvelet as:

$$\varphi_{j_0,k}(x) = \varphi_{j_0,k}(x - 2^{-j_0})k, \hat{\varphi}_{j_0(\omega)} = 2^{-j_0}W_0(2^{-j_0}|\omega|).$$
(2.41)

Curvelet transform is composed of fine-level directional elements  $(\varphi_{j,k,l})_{j\geq j_0,k,l}$  and coarse-scale isotropic father wavelet  $(\phi_{j_0,k})_k$ . Key components of the construction are summarized in Figure 2.12, left hand side represents the induced tiling of the Fourier frequency plane and the image on the right shows the associated spatial Cartesian grid at a given scale and orientation. The shaded region in Figure 2.12 represents a parabolic wedge in the Fourier frequency plane. The wedges are the consequence of the Fourier plane partitioning in radial (concentric circles) and angular divisions. Concentric circles are responsible for decomposition of the image in multiple scales (used for bandpassing the image) and angular divisions correspond to different angles or orientation. Therefore, to address a particular wedge we need to define both its scale and angle. Plancherel's theorem is applied in equation (2.42) to express  $c_{j,k,l}$  as an integral over the entire frequency plane.

$$c_{j,k,l} = \frac{1}{(2\pi)^2} \int \hat{f}(\omega) \overline{\hat{\varphi}_{j,k,l}(\omega)} d\omega = \frac{1}{(2\pi)^2} \int \hat{f}(\omega) U_j(R_{\theta_i}\omega) e^{i\langle x_k^{(j,l)},\omega\rangle} d\omega.$$
(2.42)



Figure 2.12: Space frequency tiling in curvelet domain [20]

### 2.3.2 Fast Discrete Curvelet Transform

New implementations of Fast Discrete Curvelet Transform (FDCT) are ideal for deployment in large-scale scientific applications due to their numerical isometry and an utmost 10 folds computational complexity as compared to Fast Fourier Transform (FFT) operating on a similar sized data. In our research work we used FDCT via wrapping, proposed by authors in [20] for image analysis. Interested readers are requested to refer to [20] for additional mathematical details.

- Compute 2D FFT coefficients and obtain Fourier samples  $\hat{f}[n_1, n_2]$  where  $-n/2 < n_1$  and  $n_2 < n/2$ .
- For each scale j and angle l, form the product  $\tilde{U}_{j,l}[n_1,n_2]\hat{f}[n_1,n_2]$
- Wrap this product around the origin and obtain  $\tilde{f}_{j,l}[n_1, n_2] = W(\tilde{U}_{j,l}\hat{f})$  $[n_1, n_2]$ , where the range of  $n_1$ ,  $n_2$  and  $\theta$  respectively are  $0 < n_1 < L_{1,j}$ ,  $0 < n_1 < L_{2,j}$  and  $(-\pi/4, \pi/4)$ .
- Apply inverse 2D FFT to each  $\tilde{f}_{j,l}$  and save discrete curvelet coefficients.



Figure 2.13: Wrapping a segment around the origin [22]

In the first two stages, Fourier frequency plane of the image is divided into radial and angular wedges owing to the parabolic relationship between a curvelets length and width, as demonstrated in Figure 2.12. Each wedge corresponds to curvelet coefficient at a particular scale and angle. Step 3 is essentially required to re-index the data around the origin as shown in Figure 2.13. Finally, inverse FFT is applied to collect discrete curvelet coefficients in the spatial domain.

# 2.4 Wave Atoms Decomposition

Wave atoms are a recent addition to the collection of mathematical transforms for harmonic computational analysis. They are a variant of 2D wavelet packets that retain an isotropic aspect ratio, tender a sharp frequency localization that cannot be achieved using a filter bank based on wavelet packets and offer a significantly sparser expansion for oscillatory functions than wavelets, curvelets and Gabor atoms. Curvelets capture coherence only along oscillations whereas wave atoms capture coherence of pattern both along and across oscillations. Wave atoms precisely interpolate between Gabor atoms [6] (constant support) and directional wavelets [23] (wavelength is directly proportional to diameter) in the sense that the period of oscillations of each wave packet (wavelength) is related to the size of essential support by parabolic scaling i.e. wavelength is directly proportional to diameter<sup>2</sup>.

Two distinct parameters  $\alpha$ ,  $\beta$  represent decomposition and directional ability and are sufficient for indexing all known forms of wave packet architectures namely wavelets, Gabor, ridgelets, curvelets and wave atoms. The triangle formed by wavelets, curvelets and wave atoms, as shown in the Figure 2.14, indicates the wave packet families for which sparsity is preserved under transformation. Wave atoms are defined for  $\alpha = \beta = 1/2$ , where  $\alpha$  indexes the multiscale nature of the transform, from  $\alpha =$ 0 (uniform) to  $\alpha = 1$  (dyadic).  $\beta$  measures the wave packet's directional selectivity (0 and 1 indicate best and poor selectivity respectively). Wave atoms represent a class of wavelet packets where directionality is sacrificed at the expense of preserving sparsity of oscillatory patterns under smooth diffeomorphisms. Essential support of wave packet in space (left) and in frequency (right) is shown in Figure 2.15.

A function is considered to be an *oscillatory pattern* if it is the image under a smooth diffeomorphism of a function that oscillates only in one direction, say along the coordinate  $x_1$ . For  $x = (x_1, x_2)$ , we formulate our mathematical model as:

$$f = \sin(Ng(x))h(x) \tag{2.43}$$

Lets suppose that a function f exists with an arbitrary N, such that g and hbelong to the class  $C^{\infty}$ , and h is compactly supported inside the interval  $[0,1]^2$ . Assume g has no critical points, then f can be represented to accuracy  $\epsilon$  in  $L^2$  by the largest  $\mathbf{C}_{\epsilon}\mathbf{N}$  wave atom coefficients in absolute value, where for all M > 0, there exists  $C_M > 0$  such that  $C_{\epsilon} \leq C_M \epsilon^{-1/M}$ . In other words, O(N) wave atom coefficients suffice to represent f to some given accuracy. In contrast, we would need  $O(N^{3/2})$ 



Figure 2.14: Comparison of different wave packets architectures with respect to multiscale nature and directional selectivity [24]



Figure 2.15: Wave atoms tiling in space and frequency [24]

curvelets coefficients; and  $O(N^2)$  wavelet coefficients to represent f with the same accuracy.

#### 2.4.1 1D Discrete Wave Atoms Decomposition

Wave atoms are constructed from tensor products of adequately chosen 1D wave packets. Let  $\psi_{m,n}^{j}(x)$  represent a one-dimensional family of wave packets, where  $j, m \geq 0$ and  $n \in Z$ , centered in frequency around  $\pm \omega_{j,m} = \pm \pi 2^{j}m$  with  $C_{1}2^{j} < m < C_{2}2^{j}$ and centered in space around  $x_{j,n} = 2^{-j}n$ . One-dimensional version of the parabolic scaling states that the support of each bump of  $\hat{\psi}_{m,n}^{j}(\omega)$  is of length O(2j) while  $\omega_{j,m} = O(2^{2j})$ . Dyadic scaled and translated versions of  $\hat{\psi}_{m}^{0}$  in frequency domain are combined and the basis function is written as:

$$\psi_{m,n}^{j}(x) = \psi_{m}^{j}(x - 2^{-j}n) = 2^{j/2}\psi_{m}^{0}(2^{j}x - n)$$
(2.44)

The coefficients  $c_{j,m,n}$ , for each wave number  $\omega_{j,m,n}$ , are obtained as a decimated convolution at scale  $2^{-j}$ .

$$c_{j,m,n} = \int \psi_m^j (x - 2^{-j}n) u(x) dx$$
 (2.45)

By Plancheler's theorem

$$c_{j,m,n} = \int e^{i2^{-j}n\omega} \overline{\psi}^j_m(\omega) \hat{u}(\omega) d\omega \qquad (2.46)$$

If the function u is discretized at xk = kh, h = 1/N, k = 1, 2, 3, ..., N, then with a small truncation error equation (2.46) is modified as:

$$c_{j,m,n}^{D} = \sum_{k=2\pi(-N/2+1:N/2)} e^{i2^{-j}nk} \overline{\hat{\psi}_{m}^{j}(k)} \hat{u}(k)$$
(2.47)

The data is supported inside two disjoint intervals of size  $2^{j+1}\pi$  which are symmetric about origin (2j + 1 points). Instead of an interval of length  $2^j x 2\pi$ , sum in

equation (2.47) is computed using a reduced inverse FFT inside an interval of size  $2^{j+1}\pi$  centered about origin as:

$$c_{j,m,n}^{D} = \sum k = 2\pi (-2^{j}/2 + 1 : 2^{j}/2) e^{i2^{-j}nk} \sum_{p \in 2\pi Z} \overline{\psi}_{m}^{j}(k+2^{j}p) \hat{u}(k+2^{j}p) \quad (2.48)$$

A simple wrapping technique similar to the one described for discrete curvelet transform is used for the implementation of discrete wavelet packets. The main steps involved are:

- Perform an FFT of size N on the samples of u(k)
- For each pair (j,m), wrap the product  $\overline{\psi}_m^j \hat{u}$  by periodicity inside the interval  $[-2^j, 2^j]$  and perform an inverse FFT of size  $2^j$  to obtain  $c_{j,m,n}^D$ .
- Repeat step 2 for all pairs (j, m).

The overall complexity of the algorithm is O(NLogN) and the wavelet packets are decomposed into positive and negative frequency components, represented by

$$\hat{\psi}_{m,n}^{j}(\omega) = \hat{\psi}_{m,n,+}^{j}(\omega) + \hat{\psi}_{m,n,-}^{j}(\omega)$$
(2.49)

Hilbert transform  $H\hat{\psi}_{m,n}^{j}(\omega)$  of equation (2.49) represents an orthonormal basis  $L^{2}(\mathbf{R})$  and is obtained through a linear combination of negative and positive frequency bumps weighted by **i** and **-i** respectively.

$$H\hat{\psi}^{j}_{m,n}(\omega) = -i\hat{\psi}^{j}_{m,n,+}(\omega) + i\hat{\psi}^{j}_{m,n,-}(\omega)$$
(2.50)

#### 2.4.2 2D Discrete Wave Atoms Decomposition

A two-dimensional orthonormal basis function with 4 bumps in frequency plane is formed by individually taking products of 1D wave packets. Mathematical formulation and implementations for 1D case was presented in the previous section. 2D wave atoms are indexed by  $\mu = (j, m, n)$ , where  $m = (m_1, m_2)$  and  $n = (n_1, n_2)$ . Construction is not a simple tensor product since there is only one scale subscript j. This is similar to the non-standard multi-resolution analysis wavelet bases where the emphasis is to enforce same scale in both directions in order to retain an isotropic aspect ratio. Equation (2.44) is modified in 2D as:

$$\varphi_{\mu}^{+}(x_{1}, x_{2}) = \psi_{m_{1}}^{j}(x_{1} - 2^{-j}n_{1})\psi_{m_{2}}^{j}(x_{2} - 2^{-j}n_{2})$$
(2.51)

The Fourier transform of equation (2.51) is separable and its dual orthonormal basis is defined by Hilbert transformed wavelet packets in equation (2.53).

$$\varphi_{\mu}^{+}(\omega_{1},\omega_{2}) = \hat{\psi}_{m_{1}}^{j}(\omega_{1})e^{-i2^{-j}n_{1}\omega_{1}}\hat{\psi}_{m_{2}}^{j}(\omega_{2})e^{-i2^{-j}n_{2}\omega_{2}}$$
(2.52)

$$\varphi_{\mu}^{-}(x_1, x_2) = H\psi_{m_1}^j(x_1 - 2^{-j}n_1)H\psi_{m_2}^j(x_2 - 2^{-j}n_2)$$
(2.53)

Combination of equation (2.51) and equation (2.53) provides basis functions with two bumps in the frequency plane, symmetric with respect to the origin and thus directional wave packets oscillating in a single direction are generated.

$$\varphi_{\mu}^{(1)} = \frac{\varphi_{\mu}^{(+)} + \varphi_{\mu}^{(-)}}{2}, \ \varphi_{\mu}^{(2)} = \frac{\varphi_{\mu}^{(+)} - \varphi_{\mu}^{(-)}}{2}$$
 (2.54)

 $\varphi_{\mu}^{(1)}$  and  $\varphi_{\mu}^{(2)}$  together form the wave atoms frame and are jointly denoted by  $\varphi_{\mu}$ . Wave atoms algorithm is based on the apparent generalization of the 1D wrapping strategy to two dimensions and its complexity is  $O(N^2 Log N)$ .

# Chapter 3

# Human Face Recognition

# 3.1 Introduction

Face recognition has attracted research community during the last few decades as they are the most common visual patterns in our environment. Significant development in this area has facilitated emergence of a wide range of commercial and law enforcement face recognition and classification systems. Typical applications include driver's license, passports, voter registration card, human-computer interaction, database security, law enforcement, virtual reality to name a few. Face recognition is non-intrusive i.e. images can be captured, identified or verified even without the knowledge and physical interaction of the subject. Moreover, an expert is not required to analyze and interpret the results and data can be easily collected with simple devices like camera. Development of a reliable face recognition and classification system is an intricate task since faces are complex, and belong to a class of natural objects that does not lend themselves to simple geometric interpretations. In spite of these challenges the human visual cortex does an excellent job in efficiently discriminating and recognizing these images. A fully automated face recognition system must reliably perform three subtasks: face detection, feature extraction and recognition/identification. However, each of these subtasks itself represents a separate area of research and isolating the subtasks simplifies our job and also enhances the assessment and advancement of the component techniques. Therefore, standard face databases have been used for the experiments; and our principal focus has been on the development of new and efficient feature extraction methods.

# 3.2 Challenges in Face Recognition

Faces belong to a class of natural objects that look similar but subtle features make them different. Humans recognize faces with natural ease but automated face recognition is very challenging. The advantage of computer-aided face recognition is its ability to handle large number of faces, whereas a human brain has limited memory. Aging, changes in facial hair, illumination, viewpoint variations, and cluttered background are some of the problems that need to be tackled by an automatic face recognition system. Despite these massive challenges the human visual system efficiently discriminates and recognizes faces.

- *Head pose*: Rotation or tilt of the head significantly affects the performance of the recognition system.
- *Aging*: Images acquired in different sessions separated by long intervals may seriously degrade the recognition rate.
- *Facial expression*: Facial expressions such as smiling, shouting, crying, frowning etc., critically affect recognition accuracy.
- *Occlusion*: Partial occlusion of facial features with sunglasses, scarfs and other objects further complication detection and identification.
- *Hair style*: Changes in hairstyle may affect performance.

- *Illumination*: Amount and direction of light illuminating the subjects greatly impacts the recognition accuracy. Inconsistent lighting causes shadowing effects that significantly degrades the system's performance.
- *Frontal vs. Profile*: Profile images can be difficult to recognize if the system has only been trained using frontal faces.

Different challenges are associated with various face databases; some of the databases used to rigorously test our proposed methods and to compare with existing approaches are described in the following section.

### 3.2.1 Databases

FERET [28] was sponsored by the Department of Defense in order to develop a system with automatic face recognition capability to be employed for assistance in security, intelligence and law enforcement. The final corpus consists of 14051 8-bit grayscale images of human faces with views ranging from frontal to left and right profiles.



Figure 3.1: Sample images of a subject from FERET database

Faces94 database [29] was generated at the University of Essex and contains a series of 20 images per individual. Faces94 database is wide-ranging and encompasses images of 152 distinctive individuals. The database contains images of people of various racial origins, mainly first year undergraduate students, so the majority of individuals are between 18-20 years old but some older staff member and students are also included in the database. Some individuals are wearing glasses and/or beards.



Figure 3.3: Sample images of a subject from JAFFE database

The JAFFE database [30] includes 220 images of 10 Japanese female models captured in front of a semi reflective mirror. Each subject was recorded 3-4 times with six basic emotions and a neutral face. The camera trigger was controlled by the subjects. The resulting images have been rated by 60 Japanese women on a 5-point scale for each of the six adjectives. Figure 3.3 shows example image of one such subject from JAFFE database with different facial expressions.

Georgia Tech database [31] contains images 15 color images of each of the 50

people with different emotions. Majority of the images are captured in two sessions to take into account the variations in illumination conditions, facial expression, and appearance. Additionally, images are acquired at varying scales and orientations. Figure 3.4 shows sample images from Georgia Tech database exhibiting various facial expressions, scale variations and image acquisition at different sessions.



Figure 3.4: Sample images of a subject from Georgia Tech database



Figure 3.5: Sample images from Sheffield face database

Sheffield [32] consists of 564 images of 20 individuals. The database consist of images of individuals with mixed race, gender and appearance. Each individual is imaged in a range of poses from left/right profile to frontal view with small angular rotations between successive images. The database has been pre-cropped so that the image size is uniformly reduced to 112x92 pixels, thus, the background information is eliminated from the images and only the central characteristics of the face are retained.

ORL database [33] contains 10 different images for each of the 40 distinctive subjects. Images of some subjects are taken under varying lighting conditions, facial expressions and details. All images are captured against a dark homogeneous background with the subjects in an upright, frontal position with a small tolerance for side movement.



Figure 3.6: Sample images from ORL face database

Yale face database [34] contains 165 grayscale images of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: centerlight, with glasses, happy, left-light, without glasses, normal, right-light, sad, sleepy, surprised, and winking.



Figure 3.7: Sample images of a subject from Yale face database

AR face database [35] contains over 4000 color images corresponding to 126 individuals, i.e., 70 men and 56 women. Images feature frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were captured under controlled conditions without any restrictions on clothes, glasses, make-up, hair style, etc. Every person was imaged in two sessions separated by a two week interval.



Figure 3.8: Sample images of a subject from AR face database

# 3.3 Literature Review

Automatic face recognition systems are classified into two broad categories, namely, geometric local feature based and holistic appearance based face recognition [36]. Local feature based face recognition extracts geometric relationship between human facial features such as eyes, nose, mouth and facial boundary [40]. This approach significantly relies on the accuracy of facial feature detection. Reliable extraction of facial features is an extremely complicated task since human faces have similar features with subtle changes in size and geometry that make them different from one another. Due to aforementioned complications researchers prefer the use of holistic appearance based face recognition systems [42] wherein a human face is treated as a two-dimensional intensity variation pattern and recognition is achieved through matching of statistical properties. In this work we focus on appearance based recognition systems and briefly discuss some of the well established techniques in literature.

In order to generate discriminative features, improve the speed and accuracy of a face recognition system researchers have proposed the use of various dimensionality reduction techniques over the last two decades. Kirby *et al.* [43] represented human faces as a linear combination of weighted eigenvectors using Principal Component Analysis (PCA). Turk and Pentland [44] used PCA to represent the intensity pattern of human faces in a lower dimensional feature space. PCA based face recognition systems suffer from poor discriminatory power and high computational load, and therefore to eliminate the shortcomings of standard PCA based systems, Bartlett *et al.* [45] proposed the use of Independent Component Analysis (ICA). In [46] authors utilized Linear

Discriminant Analysis (LDA) to maximize the ratio of between-class scatter matrix and the within-class scatter matrix for improved face recognition. An eigenspace based adaptive approach that uses a specific kind of genetic algorithm called Evolutionary Pursuit (EP) [47] and Elastic Bunch Graph Matching (EBGM) [48] have been proposed to generate the best set of projection axes. Bach *et al.* [49] have proposed the use of kernel Hilbert space for ICA to adaptively generate nonlinear functions and to devise a robust algorithm with regards to variations in source density, degree of non-Gaussianity, and presence of outliers. A kernel machine-based discriminant analysis method [50] that deals with the nonlinearity of face pattern was proposed for improved representation of nonlinear and complex distribution of faces under varying viewpoint, illumination and facial expression. To enhance the discriminative power of extracted features and to achieve superior face recognition researchers have also proposed the use of bayesian [51], [52] and Support Vector Machine (SVM) [53], [54] frameworks.

Multiresolution analysis based approaches have been proposed to improve the performance of a face recognition system, deal with high image dimensionality, variations in viewpoint, illumination and facial expression. Face images are transformed into a new domain and later PCA and/or other dimensionality reduction techniques are employed. Development of enhanced multiresolution analysis tools have encouraged researchers to apply them for pattern recognition applications to achieve a high level of accuracy and class separability. Some of the well known wavelet based face recognition architectures include wavelet based PCA [55], wavelet based LDA [56], wavelet based Kernel Association Memory (kAM) [57] and wavelet based modular weighted PCA [58]. Emergence of a new multiresolution analysis tool, namely, curvelets with enhanced directional and edge representation has prompted researchers to apply them to several areas of image processing. Recent works in literature that are based curvelet transform include curvelet based PCA [59], curvelet based LDA [60] and curvelet based PCA+LDA [60]. Some of the limitations of existing face recognition algorithms include large sensitivity to viewpoint variations and number of prototypes as well as slow classification speed. This chapter presents two curvelet based face recognition algorithms;

- Recognition with Kernel Principal Component Analysis [61]: A face recognition system based on curvelet transform and Kernel Principal Component Analysis (KPCA) is developed. In earlier works features extracted from curvelet subbands were dimensionally reduced using traditional PCA. KPCA transforms data into a non-linear space using an integral kernel operator function and generates features that are more meaningful than the ones extracted using a linear PCA.
- Recognition based on Multidimensional PCA and Extreme Learning Machine: In this work a new human face recognition algorithm based on Bidirectional Two Dimensional Principal Component Analysis (B2DPCA) and Extreme Learning Machine (ELM) is introduced. The proposed method is based on curvelet image decomposition of human faces and utilizes its selected subband for classification. Subband exhibiting a maximum standard deviation is dimensionally reduced using an improved dimensionality reduction technique. These feature sets are used for classification using an ELM classifier. Other notable contributions of the proposed work include significant improvements in classification rate, up to hundred folds reduction in classification time and negligible dependence on the number of prototypes.

The following sections briefly describe the individual components used in our recognition framework followed by a detailed overview of our proposed methods, experimental results and comparisons against existing techniques.

# **3.4** Kernel Principal Component Analysis

Karhunen-Loeve expansion, also known as principal component analysis, is a powerful technique for extracting structural information from higher dimension data. PCA is an orthogonal transformation of the coordinate system and is evaluated by diagonalizing the covariance matrix. Given a set of feature vectors  $\bar{x}_i \in \mathbb{R}^{\mathbb{N}}, i = 1, 2, ..., \bar{m}$ which are centered with zero mean, their covariance matrix is evaluated as:

$$\bar{C} = \frac{1}{\bar{m}} \sum_{j=1}^{\bar{m}} \bar{x}_i \bar{x}_j^T \tag{3.1}$$

Eigenvalue equation,  $\lambda v = \bar{C}v$  is solved where v is the eigenvector matrix. To generate data with  $\mathbb{Q}$  dimensions, eigenvectors corresponding to  $\mathbb{Q}$  largest eigenvalues are selected as basis vectors of the lower dimension subspace. KPCA [65] is a generalization of PCA to compute the principal components of a feature space that is nonlinearly related to the input space. Feature space variables are obtained by higher order correlations between input variables. KPCA operates as a nonlinear feature extractor by mapping input space to a higher dimension feature space through a nonlinear mapping function where the data is linearly separable. Cover's theorem [63] justifies the conversion of data into a higher dimensional space and formalizes the intuition that the number of separation classes increase with dimensionality, thus more views of the class and non class data become evident. Mapping achieved using a kernel based technique solves the problem of nonlinear distribution of low level image features and also acts as a dimensionality reduction step. Data is transformed from a lower dimension space to a higher dimension using the mapping function  $\bar{\phi} : \mathbb{R}^{\mathbb{N}} \to \mathbb{F}$ , and linear PCA is performed on  $\mathbb{F}$ . The covariance matrix in the new domain is calculated using equation (3.2).

$$\bar{\mathbb{C}} = \frac{1}{\bar{m}} \sum_{j=1}^{\bar{m}} \bar{\phi}(\bar{x}_i) \bar{\phi}(\bar{x}_j)^T$$
(3.2)

The modified eigenvalue equation,  $\lambda v = \mathbb{C}v$  is solved to reduce the data dimension. The nonlinear map  $\bar{\phi}$  is not computed explicitly and is evaluated using the kernel function  $\mathbb{K}(\bar{x}_i, \bar{x}_j) = (\bar{\phi}(\bar{x}_i)\bar{\phi}(\bar{x}_j))$ . The kernel function implicitly computes the dot product of vectors  $\bar{x}_i$  and  $\bar{x}_j$  in the higher dimension space. Kernels are considered as functions measuring similarity between instances. The kernel value is high if the two samples are similar and zero if they are distant. Some of the commonly used kernel functions are listed in Table 3.1.

Pairwise similarity amongst input samples is captured in a Gram matrix  $\mathbb{K}$  and each entry of the matrix  $\mathbb{K}_{ij}$  is calculated using the predefined kernel function  $\mathbb{K}(\bar{x}_i, \bar{x}_j)$ . Eigenvalue equation in terms of Gram matrix is written as  $\bar{m}\lambda\bar{\beta} = \mathbb{K}\bar{\beta}$ .

K represents a positive semi definite symmetric matrix and contains a set of Eigenvectors which span the entire space.  $\bar{\beta}$  denotes the column vector with entries  $\bar{\beta}_1, \bar{\beta}_2, \dots, \bar{\beta}_{\bar{m}}$ . Since the Eigenvalue equation is solved for  $\bar{\beta}$  instead of eigenvector  $\mathbb{V}_i$ of KPCA, the entries of  $\bar{\beta}$  are normalized in order to ensure that the eigen values have unit norm in the feature space. After normalization the matrix consisting of eigenvector is computed as  $\mathbb{V} = \mathbb{D}\bar{\beta}$ , where  $\mathbb{D} = [\bar{\phi}(\bar{x}_1), \bar{\phi}(\bar{x}_2), \dots, \bar{\phi}(\bar{x}_{\bar{m}})]$  is the data matrix in the feature space.

Kernel Type	Mathematical Identity		
Gaussian	$\mathbb{K}(\bar{x}_i, \bar{x}_j) = \exp \frac{(-\ \bar{x}_i, \bar{x}_j\ ^2)}{2\bar{\alpha}^2}$		
Polynomial	$\mathbb{K}(\bar{x}_i, \bar{x}_j) = (\bar{x}_i.\bar{x}_j + \bar{\alpha})^{\bar{d},  \bar{d}=1,2,3}$		
Sigmoid	$\tanh(\mathbb{K}(\bar{x}_i, \bar{x}_j) + \bar{\alpha})$		

Table 3.1: Kernels and their associated mathematical functions

# 3.5 Face Recognition using Curvelets and KPCA

This method deals with recognition of face images using kNN based classification, utilizing reduced dimension feature vectors obtained from curvelet space. Images from each dataset are converted into gray level image with 256 gray levels. Conversion from RGB to gray level format along with a two fold reduction in the image size are the only pre-processing operations performed on the images. Our proposed method is based on extraction of global image features. Though resizing will result in loss of some of the localized image information it is assumed that the overall global information extracted from the resized image will not degrade radically. We randomly divided the image database into two non-overlapping sets, i.e., training and testing set. All images within each datasets are of the same dimension, i.e.  $R \times C$ . Similar image sizes support the assembly of equal sized curvelet coefficients and feature vector extraction with identical level of global content. Curvelet transform of every image is computed and only coarse level coefficients are extracted. Vectorization is the next step to convert our curvelet coefficients into  $U \times V$  dimension vector, called as curvelet vector, whereas  $U \times V \ll R \times C$ .

Applying kNN on curvelet vectors could be computationally expensive due to higher dimensionality of data originating from large images. Inclusion of outliers and irrelevant image points into classification can also affect the algorithm's performance, hence, KPCA is applied to reduce the data dimension. Scholkopf *et al.* [65] proposed KPCA, wherein principal components are computed in a higher dimensional feature space that is non-linearly related to the input space. Thus, KPCA reliably extracts non-linear principal components while maintaining global content of the input space.

A Polynomial function based KPCA is used in our proposed method for dimensionality reduction of curvelet features. KPCA feature vectors retain the global structure of input space that facilitates accurate classification with lower computational complexity, diminished outliers and irrelevant information. Next, kNN algorithm is
#### Table 3.2: Overview of KPCA Based Face Recognition

**INPUT:** Randomly divide image database into two subsets  $TR_{\iota}$  and  $TE_{\kappa}$  where  $\iota = 1, 2, ..., \mu$  and  $\kappa = 1, 2, ..., \nu$  representing training and test image sets respectively.

**OUTPUT:** Classifier - f(x)

1. Resize images from all database to  $R\times C$ 

2. Compute curvelet transform of every train and test image

3. Vectorize coarse level features into  $U \times V$  dimension vector

4. Compute the kernel matrix  $\mathbb{K}_{TR_{\iota}}$  and  $\mathbb{K}_{TE_{\kappa}}$  where each entry of the matrix is computed using a polynomial kernel function (see Table 3.1)

5. Solve eigenvalue equations:

$$\bar{m}\lambda\bar{\beta}_{TR} = \mathbb{K}_{TR}\bar{\beta}_{TR}$$
$$\bar{m}\lambda\bar{\beta}_{TE} = \mathbb{K}_{TE}\bar{\beta}_{TR}$$

 $'\lambda'$  and  $'\bar\beta'_{\mathfrak x}$  represent eigenvalue and eigenvector matrices respectively

6. Obtain kernel PCA based feature vectors by computing principal

component projections of each image into non-linear subspace using  $\bar{\beta}_{\mathfrak{x}}$ 

7. Classify KPCA based feature vectors using kNN

trained using labeled KPCA feature vectors. We selected kNN based classification due to its attractive properties and better performance in image-to-class scenario compared with other parametric classification schemes as argued by Boiman *et al.* [66]. Finally, test feature vectors are classified using kNN scheme utilizing Euclidean distance to compute dissimilarity between input images. Table 3.2 presents step-by-step procedure of our proposed techniques.

Number of	FERET		0	RL	GTech		
Components	Curvelet	Proposed	Curvelet	Proposed	Curvelet	Proposed	
	+ PCA	Method	+ PCA	Method	+PCA	Method	
5	47.9	66.7	45.5	81	76.6	96	
10	56.5	82.3	69	91	79.7	96	
15	57	86	72.5	93	80.3	97	
20	54.3	88.7	71.5	95.5	81.3	97.6	
25	52.2	88.7	74.5	96	81.7	97.6	

Table 3.3: Average recognition rates (%) using Curvelet+PCA [59] and our KPCA based recognition scheme

#### 3.5.1 Experimental Results

As described earlier, images are converted from RGB to gray level with a 2 fold reduction in size. In the FERET database 50% of images from each subject are used as prototypes and the remaining 50% for testing. Five images of each subject from ORL database are randomly selected as prototype and the remaining 5 are used for testing. Similarly 8 images of each subject of the Georgia Tech dataset are randomly selected for training. Both the testing and training sets of images are decomposed using curvelet transform at 5 scales and 8 different angles. Amongst the curvelet coefficients only approximate coefficients are selected as feature vectors since they closely represent an approximation of the input image. The selected feature vectors are dimensionally reduced with KPCA using a  $3^{rd}$  degree polynomial kernel. kNN is applied on dimensionally reduced feature vectors using a neighborhood size of 5. The above process was repeated thrice for all databases and average results are tabulated.

The recognition accuracy for FERET, ORL and GTech databases using our proposed method is listed in Table 3.5.1. Varying number of principal components are used to emphasize the recognition accuracy achieved using PCA and KPCA prior to saturation. It is clearly evident that the proposed method outperforms curvelet+PCA based technique. In this thesis the term "number of components" refers to the dimensionality of the feature vectors.

# 3.6 Bidirectional Two-Dimensional Principal Component Analysis

PCA is a data representation technique widely used in pattern recognition and compression schemes. Pioneering work by Kirby and Sirovich [43] used PCA for enhanced representation of face images, however PCA fails to capture even a minor variance unless it is explicitly accounted in the training data. Wiskott *et al.* [48] proposed a bunch graph matching technique to overcome limitations and weakness of linear PCA. In [67] Yang *et al.* proposed two dimensional PCA (2DPCA) for image representation. As opposed to PCA, 2DPCA is based on 2D matrices rather than 1D vectors. Therefore, image matrix does not need to be vectorized prior to feature extraction. Instead an image covariance matrix is directly computed using the original image matrices.

Let X denote a  $\check{q}$  dimensional unitary column vector. To project a  $\check{p}x\check{q}$  image matrix A to X; linear transformation  $\mathbb{Y} = AX$  is used which results in a  $\check{p}$  dimensional projected vector Y. The total scatter of the projected samples is determined to measure the discriminatory power of the projection vector X. The total scatter is characterized by the trace of  $S_{\check{x}}$  i.e. covariance matrix of the projected feature vectors,  $J(X) = tr(S_{\check{x}})$ , where tr() represents the trace of  $S_{\check{x}}$ .

$$\mathbb{S}_{\check{x}} = E[\mathbb{Y} - E(\mathbb{Y})][\mathbb{Y} - E(\mathbb{Y})]^T = E[(\mathbb{A} - E\mathbb{A})\mathbb{X}][(\mathbb{A} - E\mathbb{A})\mathbb{X}]^T \qquad (3.3)$$

$$tr(\mathbb{S}_{\check{x}}) = \mathbb{X}^T [E(\mathbb{A} - E\mathbb{A})^T (\mathbb{A} - E\mathbb{A})] \mathbb{X}$$
(3.4)

 $\mathbb{G}_{\check{t}} = E[(\mathbb{A} - E\mathbb{A})^T(\mathbb{A} - E\mathbb{A})]$  is a nonnegative  $\check{q}x\check{q}$  image covariance matrix. If there are  $\mathbb{M}$  training samples, the  $\check{\alpha}^{th}$  image sample is denoted by  $\check{p}x\check{q}$  matrix  $\mathbb{A}_{\check{\alpha}}$ .

$$\mathbb{G}_{\check{t}} = \frac{1}{\mathbb{M}} \sum_{\check{\alpha}=1}^{\mathbb{M}} (\mathbb{A}_{\check{\alpha}} - \check{\mathbb{A}})^T (\mathbb{A}_{\check{\alpha}} - \check{\mathbb{A}})$$
(3.5)

$$\mathbf{J}(\mathbf{X}) = \mathbf{X}^T \mathbf{G}_{\check{t}} \mathbf{X} \tag{3.6}$$

Where  $\mathbb{A}$  represents an average image of all the training samples. The unitary vector  $\mathbb{X}_{opt}$  that maximizes the generalized total scatter criterion  $\mathbb{J}(\mathbb{X})$  is called the optimal projection axes.  $\mathbb{X}_{opt}$  represents a collection of  $\check{d}$  orthonormal eigen vectors  $\mathbb{X}_1, \mathbb{X}_2, \dots, \mathbb{X}_{\check{d}}$  of  $\mathbb{G}_{\check{t}}$  corresponding to  $\check{d}$  largest eigen values. Hence dimensionality of every image  $A_{\check{\alpha}}$  is reduced by post multiplying and pre-multiplying it by the optimal projection axes as  $\mathbb{X}_{opt}^T \mathbb{A}_{\check{\alpha}} \mathbb{X}_{opt}$ .

A limitation of 2DPCA based recognition is that it operates along row directions only. Zhang and Zhou [68] proposed  $(2D)^2$  PCA based on the assumption that training sample images are zero mean, and image covariance matrix can be computed from the outer product of row/column image vectors. In [68] two image covariance matrices  $G_{\tilde{t}Row}$  and  $G_{\tilde{t}Col}$  are calculated by representing equation (3.5) initially in terms of row vectors of  $A_{\tilde{\alpha}}$  and  $\tilde{A}$  and later as column vectors of  $A_{\tilde{\alpha}}$  and  $\tilde{A}$ . The optimal projection axes of  $G_{\tilde{t}Row}$  and  $G_{\tilde{t}Col}$  are evaluated and labeled as  $X_{1opt}$  and  $Z_{1opt}$ . Since both  $G_{\tilde{t}}$  and  $G_{\tilde{t}Row}$  are evaluated along rows, their projection axes are similar and hence dimensionally reduced image of  $A_{\tilde{\alpha}}$  is evaluated as  $Z_{1opt}^T A_{\tilde{\alpha}} X_{1opt}$ .

Our dimensionality reduction algorithm works along the row and column directions independently of one another in order to better preserve the neighborhood relationship and to generate distinctive feature sets. Our proposed technique closely follows the work of [67] and generates an image covariance matrix  $\mathbb{G}_{t\tilde{\alpha}}$  and optimizes it. Once optimal projection axes  $\mathbb{X}_{opt\tilde{\alpha}}$  is calculated, dimensionality of every image  $\mathbb{A}_{\tilde{\alpha}}$  is reduced along its columns to generate new image sets  $\mathbb{A}_{\tilde{\beta}}$  using equation (3.7). The process is repeated to further reduce row dimension of the newly generated image sets by analyzing image covariance matrix  $\mathbb{G}_{\tilde{t}\tilde{\beta}}$  and optimal projection axes  $\mathbb{X}_{opt\tilde{\beta}}$  and finally pre-multiplying every new image  $\mathbb{A}_{\tilde{\beta}}$  with  $\mathbb{X}_{opt\tilde{\beta}}^{T}$  using equation (3.8). Block schematic diagram of our proposed B2DPCA algorithm is shown in Figure 3.9.

$$\mathbb{A}_{\check{\beta}} = \mathbb{A}_{\check{\alpha}} \mathbb{X}_{opt\check{\alpha}} \tag{3.7}$$

$$\mathbb{A}_{\check{\Theta}} = \mathbb{X}_{opt\check{\beta}}^T \mathbb{A}_{\check{\beta}} \tag{3.8}$$



Figure 3.9: Block diagram of proposed B2DPCA algorithm

### 3.7 Extreme Learning Machine

Feedforward neural networks are ideal classifiers for non-linear mappings that utilize gradient descent approach for weights and bias optimization. Important factors that influence the performance of traditional feedforward neural learning algorithm like Back-Propagation (BP) include:

- A small value of learning parameter ρ̃ causes the learning algorithm to converge slowly whereas a higher value leads to instability and divergence to local minima.
- Neural network may be **over-trained** using BP and obtain **inferior generalization performance**.
- Gradient descent based learning is an extremely **time consuming** process for most applications.

To overcome innate shortcomings of traditional learning techniques Huang *et al.* [69] proposed ELM to train a Single-hidden Layer Feedforward neural Network (SLFNN) as shown in Figure 3.10. A random selection of input weights and the hidden layer biases transforms the SLFNN into a linear system. Consequently, the output weights (linking the hidden layer and output layer) can be analytically determined through a simple generalized inverse operation of the hidden layer output matrices. In an ELM an infinitely differentiable hidden layer activation function facilitates random assignment of input weights and hidden layer biases. Consider a collection of N distinct samples  $(\dot{x}_i, \dot{t}_i)$  where  $\dot{x}_i = [\dot{x}_{i1}, \dot{x}_{i2}, ... \dot{x}_{in}]^T \in \Re^n$  and  $\dot{t}_i = [\dot{t}_{i1}, \dot{t}_{i2}, ... \dot{t}_{im}]^T \in \Re^m$ , an ELM with L hidden nodes and an activation function  $\tilde{\xi}(\dot{x})$  is modeled as:

$$\sum_{i=1}^{\mathbb{L}} \check{\gamma}_i \tilde{\xi}_i(\dot{x}_n) = \sum_{i=1}^{\mathbb{L}} \check{\gamma}_i \tilde{\xi}_i(\check{w}_i \dot{x}_n + \check{b}_i) = \dot{o}_n, n = 1, 2, \dots \mathbb{N},$$
(3.9)

where  $\check{w}_i = [\check{w}_{i1}, \check{w}_{i2}, ... \check{w}_{in}]^T$  and  $\check{\gamma}_i = [\check{\gamma}_{i1}, \check{\gamma}_{i2}, ... \check{\gamma}_{i\mathbb{L}}]^T$  represent input and hidden layer weight vectors respectively. ELM reliably approximates  $\mathbb{N}$  samples with minimum error.

$$\sum_{i=1}^{\mathbb{L}} \check{\gamma}_i \tilde{\xi}_i (\check{w}_i \dot{x}_n + \check{b}_i) = \dot{t}_n, n = 1, 2, \dots \mathbb{N}$$
(3.10)

Equation (3.10) is modified as  $\tilde{\delta}\tilde{\gamma} = \check{\tau}$ ,  $\tilde{\delta} = (\check{w}_1, ..., \check{w}_{\mathbb{L}}, \check{b}_1, ..., \check{b}_{\mathbb{L}}, \dot{x}_1, ..., \dot{x}_{\mathbb{N}})$ , such that  $i^{th}$  column of  $\tilde{\delta}$  is the output of  $i^{th}$  hidden node with respect to inputs  $\dot{x}_1, \dot{x}_2, ... \dot{x}_{\mathbb{N}}$ . If the activation function  $\tilde{\xi}(\dot{x})$  is infinitely differentiable, it is proved that the number of hidden nodes satisfy  $\mathbb{L} \ll \mathbb{N}$ . Training of ELM requires minimization of an error function  $\mathbb{E}$ .

$$\mathbb{E} = \sum_{n=1}^{\mathbb{N}} (\sum_{i=1}^{\mathbb{L}} \check{\gamma}_i \tilde{\xi}_i (\check{w}_i \dot{x}_n + \check{b}_i) - \dot{t}_n)^2 \Rightarrow \mathbb{E} = \| \tilde{\delta} \check{\gamma} - \check{\tau} \|.$$
(3.11)

In classical neural networks  $\tilde{\delta}$  is determined using gradient descent optimization wherein the input weights  $\tilde{w}_i$ , hidden layer weights  $\tilde{\gamma}_i$  and bias parameters  $\tilde{b}_i$  are iteratively tuned with a learning rate  $\check{\rho}$ . A small value of  $\check{\rho}$  causes the learning algorithm to converge slowly whereas a higher value leads to instability and divergence to local minima. To avoid instability and divergence to local minima, ELM incorporates a minimum norm least-square solution. Hence, the problem is transformed to a new domain and an optimal solution in the simplified domain is evaluated using the leastsquare solution. Instead of tuning the entire network parameters, input weights and bias parameters are randomly allocated and the problem is curtailed to the leastsquare solution of  $\tilde{\delta}\check{\gamma} = \check{\tau}$ . The hidden layer output matrix  $\tilde{\delta}$  is a non-square matrix and the norm least-square solution reduces to  $\check{\gamma} = \tilde{\delta}^*\check{\tau}$ , where  $\tilde{\delta}^*$  is the moore-penrose generalized inverse of  $\tilde{\delta}$ . An infinitely small training error is achieved using the above model since it represents a least-square solution of the linear system.

$$\|\tilde{\delta}\bar{\gamma} - \check{\tau}\| = \|\tilde{\delta}\tilde{\delta}^*\check{\tau} - \check{\tau}\| \equiv \min_{\check{\gamma}}\|\tilde{\delta}\check{\gamma} - \check{\tau}\|.$$
(3.12)



Figure 3.10: Architecture of an extreme learning machine classifier

## 3.8 Face Recognition using Multi-Dimensional PCA and ELM

The proposed method is based on image decomposition of curvelet transform and uses dimensionally reduced coefficients for recognition and classification. Distinctive feature sets generated using B2DPCA are used to establish recognition accuracy. Block schematic diagram of our proposed algorithm with important steps is shown in Figure 3.11.

Images from each database are converted into gray level image with a two fold reduction in image size to extract global image features. Each database is randomly divided into non-overlapping training and testing set so that 40-45% of images of each subject are used as prototypes and the remaining images are used for testing. Curvelet transform is applied to generate initial feature vectors since it offers superior performance in presence of singularities in higher dimension, and enhances localization of higher frequency components with minimized aliasing effects. Input images are



Figure 3.11: Schematic diagram of proposed face recognition algorithm

resized to RxC, since analogous image sizes support generation of curvelet feature vectors with identical level of global information. Furthermore, curvelet transform of every image is computed at 3 scales and 8 angular orientations, thus, generating 25 distinct subbands.

Standard deviation of every subband is calculated and the subband that exhibits the highest standard deviation is selected as initial feature vector, UxV, where UxV<< RxC. In contrast to the most recent work in literature [60] that uses two subbands, we select only one subband since the difference of standard deviation among successive subbands is quite significant. This noteworthy disparity in standard deviation is consistent for all the tested databases as shown in Table 3.4, where l represents the orientations. The proposed approach based on selecting a subband with the utmost standard deviation leads to momentous savings in computational cost during dimensionality reduction and classification stages. It is noted that the approximate curvelet subband holds the maximum standard deviation from amongst the 25 curvelet subbands. Figure 3.12 justifies our approach of selecting only one subband, i.e., curvelet subband at scale=1 and is in agreement with the results presented in Table 3.4.

Dimensionality reduction techniques have been frequently applied for real-time,

Database	Scale=1		Scale=2						
		$l_1$	$l_2$	$l_3$	$l_4$	$l_5$	$l_6$	$l_7$	$l_8$
FERET	63.3	3.7	3.7	3.1	3.7	3.5	3.6	3.1	3.7
Faces94	74.6	4.9	4.3	3.6	3.9	5.0	4.5	3.4	3.4
JAFFE	87.2	6.7	6.9	3.7	4.4	6.7	6.9	3.7	4.3
GTech	68.9	3.0	2.8	4.9	4.6	3.3	3.3	5.5	5.4
ORL	56.1	5.1	3.9	6.1	5.9	3.9	3.5	5.4	5.3
Sheffield	51.0	5.2	3.4	4.0	5.1	7.8	4.4	3.6	4.8

Table 3.4: Mean standard deviation of curvelet subbands in various databases



Figure 3.12: (a) Original FERET image, (b) Curvelet subband at scale=1, (c-j) Curvelet subbands at scale=2 and 8 angular orientations

accurate and efficient processing. B2DPCA is used to achieve superior and unique feature sets and minimize computational complexity of our framework. 2DPCA was proposed in the seminal work of Yang *et al.* [67] wherein an image covariance matrix is computed directly using the original image matrices. In our proposed work features are extracted by initially reducing dimension of initial feature matrix, i.e., selected curvelet subband along its columns, called as intermediate features. Later dimensionality of intermediate features is reduced along its rows so as to generate final feature sets, each of size U'xV', where U'xV' << UxV. Our modified approach preserves information between adjacent pixels and generates distinctive feature vectors. An ELM classifier is trained and tested using labeled B2DPCA feature vectors to ascertain accuracy.

#### 3.8.1 Results and Discussion

All Images are resized with a 2 fold dimension reduction and converted from RGB to gray level image. In all databases 40-45% of images of each subject are used as prototypes and the remaining images for test purposes. Both the testing and training image sets are decomposed using curvelet transform at 3 scales and 8 different angles. Approximate curvelet coefficients are dimensionally reduced using B2DPCA, vectorized, trained and tested using ELM. Fast learning and testing speed offered by ELM enabled us to repeat the experiments several times; 100 experiments are conducted for each database and average results are calculated. We have compared our ELM based recognition scheme (50 hidden neurons) against methods utilizing kNN of neighborhood size 5.

A comparative study of recognition performance is compared using ORL and Yale face database. Results are obtained using various techniques with 60 principal components. It is evident from the results presented in Table 3.5 that our proposed method outperforms existing wavelet and/or curvelet based face recognition architectures. In the following sections for simplicity we have only compared our results with a curvelet

Method	YALE	ORL
Standard Eigenface	76	92.2
Waveletface	83.3	92.5
Curveletface	82.6	94.5
Waveletface $+$ PCA	84	94.5
Waveletface $+$ LDA	84.6	94.7
Waveletface + Weighted Modular PCA	83.6	95
Curvelet face + LDA	83.5	95.6
Waveletface $+$ KAM	84	96.6
Curveletface + PCA	83.9	96.6
Curvelet face + PCA + LDA	92	97.7
Curvelet face + B2DPCA + ELM	99.7	99.9

Table 3.5: Comparative Accuracy for YALE and ORL face database

based PCA+LDA [60] approach.

The recognition accuracy achieved for Sheffield and FERET database for varying number of principal components is compared with curvelet based PCA+LDA approach [60] in Table 3.6, whereas results obtained for ORL and GTech database are listed in Table 3.7. Results obtained using our proposed method consistently outperform PCA+LDA based approach for Sheffield, FERET, ORL and GTech datasets. Comparative results obtained using the AR face database are plotted in Figure 3.13. Improvements in recognition accuracy using the AR and the FERET database imply that our proposed method is suitable for dealing with challenging face databases. It is worth mentioning that increasing the number of principal components does not necessarily increase accuracy and the use of localized information for face recognition may be exploited to generate improved results.

Number of	Sheffiel	ld	FERE	Г
Components	PCA+LDA [60]	Proposed	PCA+LDA [60]	Proposed
5	93.89	93.99	77.42	92.27
10	96.11	99.31	80.65	93.03
15	97.78	99.80	77.41	93.08
20	99.44	99.91	87.09	90.46
25	99.44	100	90.32	97.83
30	98.88	100	75.8	99.09
35	98.46	100	88.17	96.09
40	97.12	100	80.64	99.70
45	97.77	100	67.20	98.74
50	97.22	100	66.67	99.63

Table 3.6: Average recognition rates (%) for Sheffield and FERET database

In addition to improved accuracy our proposed method is also independent of the number of prototypes in comparison to other face recognition algorithms. Recognition rates obtained for ORL database at 30%, 40%, 60% and 70% prototypes are plotted in Figure 3.14 (y-axis denotes the accuracy and x-axis denotes the number of principal components). In order to avoid within-scatter matrix singular cases, authors in [60] extracted curvelet coefficients at 4 scales. Our proposed method is robust and free of the singularity issues, i.e., independent of the scales of curvelet decomposition that radically degrade precision of the PCA+LDA based method.

Table 3.8 compares Average Recognition Rate (AVR) and time complexity for Faces94 database. Results clearly validate our claim that the proposed method achieves superior recognition at hundred folds faster speed than state-of-the-art technique [60]. Our method is suitable for real-time applications. In addition to im-

Number of	ORL		GTech	1
Components	PCA+LDA [60]	Proposed	PCA+LDA [60]	Proposed
5	79.12	94.05	88.32	89.14
10	89.16	99.56	71	93.53
15	94.21	98.19	90.33	97.43
20	98.33	99.73	95.65	97.09
25	97.5	99.56	96.34	96.81
30	97.5	99.94	94.67	97
35	98.42	99.96	96	97.42
40	96.67	99.99	96	97.71
45	97.45	100	94	97.6
50	97.52	100	93	97.87

Table 3.7: Average recognition rates (%) for ORL and GTech database

provements in classification time, our system also achieves significant computational savings during dimensionality reduction stage since only one subband is utilized.

Number of	PCA	+LDA	Pro	posed
Components	AVR(%) Time(sec)		AVR(%)	Time(sec)
10	92.17	30.74	94.92	0.1329
20	97.28	30.67	98.91	0.1340
30	99.29	31.55	99.54	0.1343
40	99.29	33.17	99.55	0.1321
50	99.29	33.36	99.87	0.1348

Table 3.8: Average recognition rates (%) and time complexity for Faces94 database



Figure 3.13: Average recognition rate (y-axis) vs. number of principal components (x-axis) for AR face database

Table 3.9:	Average recognition rates	(%)	for	JAFFE	database	at	varying	number
	of neurons							

Neurons	35	40	45	50	55	60	STD
Components							
5	92.56	92.92	92.97	92.95	92.62	92.55	0.2047
10	99.80	99.78	99.93	99.77	99.81	99.75	0.0641
15	99.01	99.07	99.01	98.98	99.04	98.94	0.0454
20	99.97	99.95	99.96	99.97	99.89	99.95	0.0299
25	100	100	100	100	100	100	0

Experiments are also carried out by varying number of hidden neurons from 35 to 60 in intervals of 5, however negligible variation in accuracy is observed, as indicated



Figure 3.14: Average recognition rate (y-axis) vs. number of principal components (x-axis) for ORL database at varying prototypes

by the recognition rates and standard deviation (STD) in Table 3.9. Results in Table 3.9 represent a significant departure from traditional classification schemes where correctness is greatly attributed to the classifier parameters, for example, neighborhood size for a kNN classifier. To further investigate the advantages associated with the use of an ELM classifier, we classified B2DPCA reduced feature vectors using a kNN and ELM classifier with 5 neighbors and 50 neurons respectively. Improved recognition accuracy is achieved using ELM in comparison with kNN at varying number of principal components, as presented in Figure 3.15.

In order to emphasize the benefits of our proposed dimensionality reduction technique, i.e., B2DPCA, we compared the accuracy achieving using Yang's 2DPCA [67]



Figure 3.15: Average recognition rate (%) for FERET database using kNN and ELM

Table 3.10:	Average recognition	rates $(\%)$ for	or FERET	database	using 2DPCA	and
	B2DPCA	. ,			-	

Number of	2DPCA+ELM	B2DPCA+ELM
Components	AVR(%)	AVR(%)
4	49.7	51.06
9	83.12	77.93
16	78.15	93.91
25	94.13	97.83
36	99.25	99.74
49	99.78	99.63

with our approach. In both situations we used an ELM classifier to train our system and to ascertain recognition rate. Table 3.10 compares the average recognition rates obtained for FERET database at varying principal components. It is worth mentioning that the number of principal components are represented in the form of square of integers because of operational behavior of 2DPCA that simultaneously reduces dimensions along rows and columns using a single set of optimized eigen vectors (please refer to Chapter 3.6 for details). Improvements in accuracy using our proposed dimensionality reduction technique are apparent from the results.

Number of	ELM		<b>BP</b> Network at Varying Iterations						
Components		50	100	150	200	250	300	350	400
5	92.95	42.54	55.64	57.27	58	65.09	65.45	66.72	69.45
10	99.77	58.73	80.36	77.64	74.90	74.91	75.64	72.91	73.82
15	98.98	80	85.63	85.64	82.73	82.36	81.45	79.82	81.09
20	99.97	78.18	80.91	78	79.45	77.45	78.90	78	79.46
25	100	73.63	82.18	76.91	79.82	78.55	77.09	77.45	78

Table 3.11: Average recognition rates (%) for JAFFE database using ELM and traditional BP based neural network

We also conducted tests to investigate the improvements in accuracy in comparison to BP based neural network architecture with a single hidden layer. We classified B2DPCA reduced feature vectors using ELM and the BP based neural network, in both situations the hidden layer consisted of 50 neurons. To analyze the effect of number of iterations on the BP based neural network, we varied it from 50 to 400 in steps of 50 and the averaged results after 5 independent experiments are tabulated in Table 3.11. As pointed out earlier, from the results it is clear that increasing the number of principal components do not always guarantee an improved accuracy; moreover, accuracy is directly related to the discriminative information contained within the feature vectors. In a back propagation algorithm an increase in the number



Figure 3.16: Mean square error vs. number of epochs for JAFFE face database of epochs shrinks the mean square error between the actual and the target output (demonstrated in Figure 3.16) but it does not ensure better classification accuracy.

## Chapter 4

## **Fingerprint Compression**

## 4.1 Need for Image Compression

The advent of high speed computing devices and rapid developments in the field of communication has created a tremendous opportunity for various image processing applications. The amount of data required to store a digital image is continuously increasing and overwhelming the storage devices. In addition to high storage requirements, transmission of images through communication channels necessitates the application of sophisticated compression techniques. An acutely designed data compression system is required to alleviate these problems. Data compression is a key to the rapid progress made in the field of information technology. It is highly impractical to put in uncompressed images, audio and video on websites.

The fundamental goal of image compression is to represent an image in digital form with as few bits as possible while maintaining an acceptable level of image quality [71]. A typical 4"  $\times$  4" image scanned at 300 dpi with 24 bpp (true color) corresponds to more than 4 mega bytes. Thus, each image typically requires a high storage and a transmission time of more than a minute through a typical ISDN channel. There are two ways to solve this problem in a distributed environment, either to increase the channel bandwidth or to compress the image. Towering costs associated with high bandwidth channels makes them less attractive when compared to image compression.

### 4.2 Compression techniques

Over the last two decades numerous image compression techniques have been proposed. A compression algorithm incorporates a corresponding decompression algorithm in order to reconstruct the original image. A compression algorithm takes an input image **U** to generate an output image  $\mathbf{U}_c$  with fewer bits, and a reconstruction algorithm operates on the compressed image  $\mathbf{U}_c$  to reconstruct **V**. Based on the reconstruction constraints either lossless compression or lossy compression is employed.

- Lossless compression belongs to a category of algorithms where the reconstructed image V is identical to input image U. Such compression techniques are employed in situations where any appreciable image degradation is highly detrimental (example medical images). Therefore, lossless compression achieves very limited compression rates. Widespread lossless compression techniques include run-length coding, huffman coding, Lempel-Ziv-Welch (LZW) algorithm and arithmetic coding.
- Lossy compression refers to techniques that sustain loss of information and the reconstructed image is dissimilar to the original one. Several applications exist where it is acceptable for a reconstructed image to be different from the original as long as the differences do not result in visually annoying artifacts. Most compression standards fall into one of the three broad categories: vector quantization, predictive coding and transform coding [72], [73]. Vector quantization and predictive coding achieve inferior compression quality and are not as competitive as transform coding techniques [74].

#### 4.2.1 Transform based compression

Transform based coding techniques statistically decor-relate the information contained in the image so that redundant data can be discarded [75]. Hence, a *dense* signal is converted to a *sparse* signal and most of the information is concentrated to a few significant coefficients. Transform based compression techniques allow efficient storage, display and transmission of images that would otherwise be impractical.

Discrete Cosine Transform (DCT) [77] is a popular transform coding method used in JPEG standard for lossy compression of images. In JPEG, an image is divided into series of blocks, converted from spatial domain to frequency domain using a 2D DCT, quantized and sent to a lossless entropy encoder. Due to the blocked nature of input, correlation across the block boundaries is not eliminated and thus noticeable and annoying "blocking artifacts" are encountered at low bit rates.

More recently, wavelet transform has emerged as a cutting edge technology, within the field of image compression. Wavelet-based coding [78] provides substantial improvement in picture quality at higher compression ratio. Over the last decade, a variety of powerful and sophisticated wavelet-based schemes for image compression have been developed and implemented. Wavelet based techniques do not divide the image into blocks, but analyzes the whole image at once. This prevents any blocking artifact within the reconstructed image and improves efficiency with respect to compression ratio and Peak Signal to Noise Ratio (PSNR).

#### 4.3 Literature Review

Law enforcement, border security and forensic applications are some crucial fields where fingerprint image compression plays an important role. Emergence of protocols and commercially available products has prompted law enforcement agencies to use Automated Fingerprint Identification Systems (AFIS) during criminal investigations. The US Federal Bureau of Investigation (FBI) deals with a massive collection of fingerprint database, comprising of more than 200 million cards, growing at the rate of 30,000-50,000 new cards daily [80]. The archive consists of inked impressions on paper cards. A single card contains 14 different images: 10 rolled impression of each finger, duplicate (flat) impression of thumb and simultaneous impression of all fingers together. Fingerprint images are digitized at a resolution of 500 dpi with 256 gray levels which results into a fingerprint card requiring approximately 10 MB of storage. The gray level images tender a refined natural appearance to humans than black and white images and allow a higher level of subjective discrimination for fingerprint examiners. This call for an efficient compression standard that can significantly reduce the image size while retaining distinctive information, in conjunction with the size of FBI fingerprint database.

Fingerprint images exhibit characteristic high energy in certain high frequency bands resulting from the ridge-valley pattern and other structures. To account for this property, the FBI Wavelet Scalar Quantization (WSQ) [81] standard for lossy fingerprint compression uses a specific wavelet packet subband structure that emphasizes the important high frequency bands. This standard has been shown to be clearly superior to JPEG compression in terms of psychovisual and PSNR quality [82]. Wavelet-based compression schemes have been subsequently investigated in great number for their usefulness in fingerprint compression. The choice of filters in classical pyramidal coding schemes specifically tuned for fingerprint compression is an active area of research. Sherlock *et al.* [83], [84] have identified that biorthogonal wavelet filters are superior to orthogonal wavelet filters.

Fingerprint compression standard developed for FBI [85], also known as Wavelet Scalar Quantization (WSQ) has incorporated biorthogonal 9-7 filter pair for highly reliable fingerprint compression and reconstruction. Discrete Wavelet Transform (DWT) is widely used for image processing applications due to its improved spacefrequency decomposition [86], energy compaction of low frequency subbands, space localization of high frequency subbands and flexibility in time frequency tiling. Wavelet packets facilitate a flexible representation by allowing decompositions at every node of the tree resulting in an explicit structure for specific applications. Image analysis using DWT is performed using a pair of Quadrature Mirror Filter (QMF) and a Dual Quadrature Mirror Filter (DQMF). These sets of filters are further decomposed into four subsets of floating point coefficients:  $h(Lo_D)$ ,  $g(Hi_D)$ ,  $\tilde{h}(Lo_R)$  and  $\tilde{g}(Hi_D)$ that define the wavelet and scaling functions for forward and inverse DWT respectively. Fingerprint images are decomposed using a 2D DWT by means of separability approach resulting into four smaller subsets. Subsets are selected based on energy content, variance, their effect on reconstructed image, and are iteratively decomposed until desired number of subbands are obtained. In FBI's WSQ the fingerprint image is recursively decomposed with a five level wavelet decomposition resulting in 64 distinct subbands as shown in Figure 4.1. These subbands are quantized and represented using different coding techniques.

Filter Type	Filter Coefficients
$h(Lo_D)$	0.03783, -0.02385, -0.11062, 0.37740, 0.85270
	0.37740, -0.11062, -0.02385, 0.03783
$g(Hi_D)$	0.06454, -0.04069, -0.41809, 0.78849, -0.41809
	-0.04069,  0.06454
$\tilde{h}(Lo_R)$	-0.06454, -0.04069, 0.41809, 0.78849, 0.41809
	-0.04069, -0.06454
$\tilde{g}(Hi_R)$	0.03783, 0.02385, -0.11062, -0.37740, 0.85270
	-0.37740, -0.11062, 0.02385, 0.03783

Table 4.1: Filter coefficients for 9-7 wavelet filter

Inspired by the WSQ algorithm, a few wavelet packet based fingerprint compression schemes have been developed. An improved WSQ entropy coding stage using lossless zerotree coding is proposed in [88]. A 73-Subband subband structure with

						π	/2		<sup>to</sup> row
4	7	8	19	20	23	24			
6	9	10	21	22	25	26			
12	15	16	27	28	31	32			
14	17	18	29	30	33	34	52	53	
36	39	40							
38	41	42							
44	47	48							
46	49	50	51				54	55	
			57				60	61	
			59				62	63	
	4 6 12 14 <u>36</u> 38 44 46	4       7         6       9         12       15         14       17         36       39         38       41         44       47         46       49	4       7       8         6       9       10         12       15       16         14       17       18         36       39       40         38       41       42         44       47       48         46       49       50	4       7       8       19         6       9       10       21         12       15       16       27         14       17       18       29         36       39       40       40         38       41       42       44         46       49       50       51         57	4       7       8       19       20         6       9       10       21       22         12       15       16       27       28         14       17       18       29       30         36       39       40       -       -         38       41       42       -       -         44       47       48       -       -         46       49       50       51       -         51	4       7       8       19       20       23         6       9       10       21       22       25         12       15       16       27       28       31         14       17       18       29       30       33         36       39       40       44       47       48         46       49       50       51       51	4       7       8       19       20       23       24         6       9       10       21       22       25       26         12       15       16       27       28       31       32         14       17       18       29       30       33       34         36       39       40       33       34         44       47       48       46       49       50       51	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

Figure 4.1: FBI's 64-subband structure with a 5-level wavelet decomposition [87]

subsequent lattice vector quantization [89] has been proposed by Kasaei *et al.*. As shown in Figure 4.2, further decomposition is not applied to the diagonal subimage of the first wavelet level due to its low energy content, low variance, and low effect of this subimage on the reconstructed image, whereas the second wavelet level is further decomposed due to its important innate characteristics. Hence, the 73-Subband is computationally less expensive than the 64-Subband and offers superior performance in comparison to JPEG, WSQ, EZW [90], and SPIHT [91].

The concept of adaptive subband structures to improve on the fixed structure of WSQ was proposed in [92]. Khuwaja [93] adapts the wavelet packet filters and the decomposition level in addition to the selection of subbands to better represent



(a) Kasaei's 73-Subband structure

(b) FBI's 64-Subband structure

Figure 4.2: 5-level decomposition subband structure

the actual frequency characteristics of the image. Some compression schemes exploit the strong directional features in fingerprint images caused by ridges and valleys. A scanning procedure following dominant ridge direction has shown to improve lossless coding results as compared to JPEG-LS and PNG [94]. A wavelet footprint representation characterizing efficiently singular structures (corresponding to ridges) [95] that delivers better results as compared to the SPIHT algorithm was proposed by Sudhakar *et al.*. Contourlets [96], [97], [98] and contourlet packets [99] are used to exploit directional information which also results in PSNR improvements as compared to classical algorithms. Recently, fingerprint image compression schemes that use genetic algorithm [100], [101], [102] to generate wavelet and scaling coefficients for each level of decomposition have also been proposed. Multiresolution analysis tools have been successfully applied to fingerprint image compression in the last two decades; we propose two new fingerprint image compression techniques based on wave atoms decomposition:

• Fingerprint Compression with Vector Quantization [103]: The pro-

posed compression scheme is based upon linear Vector Quantization (VQ) of decomposed wave atoms representation of fingerprint images. Later quantized information is encoded with arithmetic entropy scheme.

• Compression using Mathematical Morphology and Multistage Vector Quantization [104]: In this work, the compression scheme is based upon MultiStage Vector Quantization (MSVQ) of processed wave atoms representation of fingerprint images. Wave atoms expansion is processed using mathematical morphological operators to emphasize and retain significant coefficients for transmission. MSVQ quantized significance map and scalar quantized coefficients are encoded, and transmitted using arithmetic entropy scheme.

An approach similar to Chapter 3 is followed; individual components used in our compression scheme are briefly described followed by a detailed outline of our proposed method, experimental results and comparisons against existing techniques.

### 4.4 Vector Quantization

Quantization is a process that maps a signal  $\mathbf{p}(\mathbf{m})$  into a finite series of  $\mathbf{K}$  discrete messages. For every  $\mathbf{K}^{th}$  message, there exists a pair of thresholds  $\mathbf{t}_{\mathbf{k}}$  and  $\mathbf{t}_{\mathbf{k}+1}$ and output value  $\mathbf{q}_{\mathbf{k}}$  such that  $\mathbf{t}_{\mathbf{k}} < \mathbf{q}_{\mathbf{k}} < \mathbf{t}_{\mathbf{k}+1}$ . Concept of scalar or one-dimensional quantization is extended to vector data of any arbitrary dimension. Instead of output levels, vector quantization employs a set of representation vectors and matrices for one-dimensional and two-dimensional data respectively. The set of representation vector is often referred to as a codebook and the entries within the codebook are known as codewords. The thresholds are replaced by a decision surface defined by a distance metric such as euclidean distance. In vector quantization high degree of co-relation between neighboring pixels is exploited and the coding of vector can theoretically improve performance. During coding the image is divided into blocks of fixed size  $\mathbf{m} \times \mathbf{m}$  pixels. For each block of input the codeword that results in a minimum euclidean distance is found and transmitted. On reconstruction, the same codebook is used and a simple look-up operation is performed and the image is reconstructed.

The classical method for codebook construction is by use of Linde, Buzo and Gray (LBG) algorithm [105]. According to this method  $\mathbf{K}$  codebook entries are initially set to random values and on each iteration, each input space is classified based on euclidean distance. Each codebook is replaced by the mean of its resulting class and the iterations are continued until a minimum acceptable error is achieved.

# 4.5 Fingerprint Compression using Wave Atoms and Vector Quantization

Wave atoms decomposition is used for sparse representation of fingerprint images since they belong to a category of images that oscillate smoothly in varying directions. Schematic block diagram of the proposed method is shown in Figure 4.3. Discrete 2D wave atoms decomposition is applied on the original image in order to efficiently capture coherence of the fingerprint images along and across the oscillations. An orthonormal basis  $\varphi_{\mu} (\varphi_{\mu}^{(1)} + \varphi_{\mu}^{(2)})$  is used instead of a tight frame since each basis function oscillates in two distinct directions instead of one. This orthobasis variant property is significantly important in applications where redundancy is undesired.

Magnitudes of wave atoms decomposed coefficients, carrying low information content, are either zero or very close to zero hence these can be discarded without a substantial degradation in image quality. An appropriate global threshold is used to achieve desired transmission bit rate. After thresholding the wave atoms coefficients, a significance map matrix and a significant coefficient vector is generated. Significance map is a matrix of binary values that indicates the presence or absence of significant coefficient at a specific location. The significance map is divided into non-overlapping blocks of 4x4. These non-overlapping blocks of significance map are vectorized and quantized using a K-means vector quantization scheme with 64 code words. Small blocks of data are used in order to minimize the error during vector quantization. The significant coefficients are quantized using a uniform scalar quantizer with 512 distinct levels.

Quantized significance map and significant coefficients are encoded using an arithmetic encoder. Arithmetic coding is a variable length entropy scheme that attempts to minimize the number of required bits. It converts a string into another representation using more bits for infrequent characters and vice versa. As opposed to other entropy encoding techniques that convert the input message into component symbols and replace each symbol with a code word; arithmetic coding represents the entire message into a single number thereby achieving optimal entropy encoding.

#### 4.5.1 Results and Discussion

Various fingerprint images used in FBI's WSQ standard are compressed using the proposed method and substantial improvement in compression is achieved. The quality of various image compression techniques depends upon how close is the reconstructed image to the original one. Different metrics are proposed for investigating the quality of compression algorithms. Some methods investigate similarity while others explore the level of dissimilarity between reconstructed and the reference image. Mean Square Error (MSE) and PSNR are two celebrated metrics used to examine the qualitative performance. MSE is a distortion metric that provides a measure of dissimilarity between two images. MSE and PSNR are calculated using the equations below:

$$MSE = \frac{1}{\mathbf{RC}} \sum_{i=1}^{\mathbf{R}} \sum_{j=1}^{\mathbf{C}} | \mathbf{O}_{\mathbf{i},\mathbf{j}} - \bar{\mathbf{O}}_{\mathbf{i},\mathbf{j}} |^2$$
(4.1)



Figure 4.3: Block diagram of proposed fingerprint compression algorithm (WAVQ)

$$PSNR = 10\log_{10}\frac{255^2}{MSE}$$
(4.2)

where  $\mathbf{R}$  indicates the number of image rows and  $\mathbf{C}$  refers to the number of columns,  $\mathbf{O}$  represents the original image and  $\mathbf{\bar{O}}$  refers to the reconstructed image.

Figure 4.4 demonstrates a sample fingerprint and the reconstructed image at 0.25 bpp using the proposed compression method. From the figure it is evident that the

proposed method using wave atoms decomposition does an excellent job in preserving the fine details in a fingerprint image i.e. the minutiae (ridges ending and bifurcations) at lower bit rates. Table 4.2 compares the PSNR obtained using our proposed method with the FBI's WSQ compression standard at varying bitrates. As shown in Table 4.2 fingerprint compression based on wave atoms decomposition produces a significant improvement in PSNR at high compression ratios (low bit rates) in comparison to FBI's WSQ fingerprint compression standard.



Figure 4.4: (a) Original fingerprint image (b) Compressed image at 0.25 bpp

Bit Rate	FBI's WSQ [85]	Proposed WAVQ
(bpp)	PSNR (dB)	PSNR (dB)
0.1	23.72	28.62
0.2	26.25	31.21
0.3	27.96	31.68
0.4	29.36	32.42
0.5	30.37	32.65

Table 4.2: Bit rate vs. PSNR for FBI's WSQ and proposed method

### 4.6 Mathematical Morphology Operators

Mathematical morphology is the analysis of signals/images in terms of their shape. It is used in image processing applications so as to preserve edge information and create clusters of significant coefficients [106]. The basic building blocks of mathematical morphology are dilations and erosions.

The basic effect of dilation on a binary image is to gradually enlarge the boundaries of regions of foreground pixels. Thus areas of foreground pixels grow in size while holes within those regions become smaller. Dilation of a binary input image is computed by superimposing the structuring element on top of the input image so that the origin of the structuring element coincides with the input pixel position. If at least one pixel in the structuring element coincides with a foreground pixel (white pixel i.e. 1) in the image underneath, then the input pixel is set to the foreground value. If all the corresponding pixels in the image are background (black pixel i.e. 0), the input pixel is left at the background value. Erosion is the dual of dilation i.e. eroding foreground pixels is equivalent to dilating the background pixels. Dilation and erosion of an image **S** using a structuring element **A** are denoted by  $\delta_{\mathbf{A}}$  and  $\epsilon_{\mathbf{A}}$  respectively.

$$\delta_{\mathbf{A}}(\mathbf{s}) = \mathbf{S} \oplus \mathbf{A} \tag{4.3}$$

$$\epsilon_{\mathbf{A}}(\mathbf{s}) = \mathbf{S} \odot \mathbf{A} \tag{4.4}$$

Effect of dilation and erosion on a binary image using a  $3 \times 3$  square structuring element is shown in Figure 4.6 and Figure 4.7.

In order to create clusters of significant coefficients using mathematical morphology, in this work, the significance map was dilated twice and then eroded once using a  $3 \times 3$  square structuring element. In this sequence of operation [107], the first dilation is used to generate clusters of significant coefficient whereas the closing operation (a combination of dilation and erosion) is used to fill in small holes.

1	1	1
1	1	1
1	1	1

Set of coordinate points =

 $\{ (-1, -1), (0, -1), (1, -1), \\ (-1, 0), (0, 0), (1, 0), \\ (-1, 1), (0, 1), (1, 1), \}$ 





Figure 4.6: Dilation of binary map using a square structuring element



Figure 4.7: Erosion of binary map using a square structuring element

## 4.7 Multistage Vector Quantization

VQ [105] is a quantization technique applied to an ordered set of symbols. The superiority of vector quantization lies in its ability to partition vector space, its capability 83

to exploit intra-vector correlations and most importantly the block coding gain it achieves. MSVQ [108] divides the encoding stage into several smaller modules and reduces encoding complexity and memory requirements for vector quantization, especially at high compression ratios. In the first stage, a low rate vector quantizer is used to generate a relatively crude encoding of the input vector using a small codebook. The coarse approximation in the form of output labels of the vector quantizer is transmitted to the receiver. The error between the original input and the coarse representation of the first stage is quantized by the second stage quantizer and the label of the output point is transmitted to the receiver. Similarly the input to the  $\mathbf{N}^{th}$  stage vector quantizer is the difference between the original input and the reconstruction obtained from the preceding  $\mathbf{N} - \mathbf{1}$  stages.

In vector quantization, an input vector is quantized by selecting the best matching representation from amongst a codebook of 2lr stored code vectors each of dimension **1**. Vector quantization is an optimal coding technique since other coding methods for a specified number  $\mathbf{b} = \mathbf{lr}$  of bits are equivalent to special cases of VQ with suboptimal codebooks. However, optimal VQ assumes single and possibly very large codebook with no imposed constraints in its structure. The resulting encoding and storage complexity, of the order of 2lr, may be prohibitive for many applications. Multistage vector quantization is a structured VQ scheme that can achieve very low encoding and storage complexity. In MSVQ, the lr bits are divided between N stages with  $b_i$ bits for stage i. The storage complexity of MSVQ  $\sum_{i=1}^{N} 2^{b_i}$  vector is significantly less than the complexity of an unstructured VQ that requires  $\prod_{i=1}^{N} 2^{\mathbf{b}_i} = 2^{\mathbf{lr}}$ . A sequential quantization operation is performed in MSVQ where each stage quantizes the residual of the previous stage. The structure of MSVQ encoder [108] consists of a cascade of VQ stages as shown in Figure 4.8. For an N-stage MSVQ, each  $\mathbf{n^{th}}$  stage quantizer  $\mathbf{Q_n},\,\mathbf{n}=1,2,\ldots,N$  is associated with a stage codebook  $\mathbf{C_n}$  that contains  $\mathbf{L_n}$  stage code vectors. The set of stage quantizers  $Q_1, Q_2, \ldots, Q_N$  are equivalent to a single quantizer **Q**, referred as the direct-sum vector quantizer.



Figure 4.8: Block diagram of an MSVQ encoder

In the MSVQ encoder shown in Figure 4.8, the input vector  $\mathbf{X}$ , is quantized with the first stage codebook producing the first stage code vector  $\mathbf{Q_1X}$ , a residual vector is formed by subtracting  $\mathbf{Q_1X}$  from  $\mathbf{X}$ . Later the residual vector is quantized using the second stage codebook, with exactly the same procedure as in the first stage. Therefore, in each stage except the last stage, a residual vector is generated and passed to the next stage to be quantized independently of other stages. The quantized error vector provides a refinement to the previous vector quantizer output and the level of correlation decreases as the process continues. For a 3-stage MSVQ quantizer the encoder and decoder equations are shown below.

$$\begin{aligned} {\bf Y_1} &= {\bf Q_1}({\bf X}) \\ {\bf Y_2} &= {\bf Q_2}({\bf X} - {\bf Q_1}({\bf X})) \\ {\bf Y_3} &= {\bf Q_3}({\bf X} - {\bf Q_1}({\bf X}) - ({\bf Q_2}({\bf X} - {\bf Q_1}{\bf X}))) \end{aligned} \tag{4.5}$$

The block diagram of an MSVQ decoder is shown in Figure 4.9, the decoder receives for each stage an output label identifying the stage code vector selected and the reconstruction vector  $\hat{\mathbf{X}}$  is generated as:

$$\hat{\mathbf{X}} = \mathbf{Y}_1 + \mathbf{Y}_2 + \mathbf{Y}_3 \tag{4.6}$$

The overall quantization error is equal to the quantization residual from the last stage. Sequential searching of the stage codebooks reduces the encoding complexity to the storage complexity, i.e.,  $\sum_{i=1}^{N} 2^{\mathbf{b}_{i}}$ .



Figure 4.9: Block diagram of an MSVQ decoder

# 4.8 Fingerprint Compression using Mathematical Morphology and Multistage Vector Quantization

Our proposed algorithm is an extension of our earlier work [109], wherein we used mathematical morphology and SOFM to obtain improved PSNR values. Important steps involved in our proposed scheme are outlined in Figure 4.10; we do not assume any a priori information and acquisition constraints for the input data. Fingerprint images are digitized using 256 gray levels, therefore, a color space transformation is
not required. To retain fine details in an image no further pre-processing, contrast enhancement and filtering etc., is performed. Similar to the VQ based algorithm an orthonormal basis  $\varphi_{\mu} \left(\varphi_{\mu}^{(1)} + \varphi_{\mu}^{(2)}\right)$  is used instead of a tight frame to ensure that each function oscillates in two distinct directions instead of one.



Figure 4.10: Overview of the proposed MSVQ based compression algorithm

To achieve efficient compression, image data with low information content such as smooth areas can be represented with fewer numbers of coefficients. The magnitudes of wave atoms decomposed coefficients, carrying low information content, are either zero or very close to zero hence they can be discarded without a substantial degradation in reconstructed image quality. An appropriate global threshold is used to define each coefficient as either significant or not and this threshold is defined according to the user's need, and represents a tradeoff between quality and compression; a small threshold results in a better quality reconstructed image with a small compression ratio whereas a larger threshold attains reconstructed images with poor quality at high compression ratios. The thresholding operation generates a significance binary map whose values are selected as '1' if the magnitude of wave atoms decomposed coefficient is greater than predefined threshold at that location and '0' otherwise. In significance map, values of '1' and '0' indicate the presence or absence of significant coefficient at a specific location. Please note that the size of significance map is similar to the dimension of the coefficients obtained after wave atoms decomposition. To corroborate improved image reconstruction, we apply morphological operators on significance map to be used at later stage for extraction and quantization of actual coefficients. The mathematical morphology procedure comprises of dilations and erosions to minimize the loss of potentially important coefficients because of thresholding thereby ensuring a high peak signal to noise ratio of the reconstructed image. The AND operation among significance map and original wave atoms decomposed coefficients is used to extract coefficients carrying important and discriminative information of oscillatory patterns in an input image. Such extracted coefficients, termed as significant coefficients, are scalar quantized using a uniform scalar quantizer with 512 distinct levels.

Blocks of  $\mathbf{d} \times \mathbf{d}$  elements of the significance map are non-uniformly vector quantized using MSVQ. The quantization process tries to eliminate psycho-visual redundancy which causes degradation in image quality. Additional loss of coefficient information is attributed to thresholding. In our proposed method, the significance map is divided into non-overlapping blocks of  $4 \times 4$  elements and vectorized into small length vectors of dimension  $16 \times 1$ ; small blocks of data are used in order to minimize error during vector quantization. These vectors are quantized using MSVQ scheme, resulting into a codebook representation termed as MSVQ quantized significance map.

The scalar quantized significant coefficients and MSVQ quantized significance map are transformed using an arithmetic encoder. Arithmetic coding is a variable length entropy scheme that attempts to minimize the number of bits by converting a string into another representation using more bits for infrequent characters and vice versa. As opposed to other encoding techniques which convert the input message into the component symbols and replace each symbol with a code word; arithmetic coding represents the entire message into a single number thereby achieving optimal entropy encoding. For a communication setup, the computation of encoded information using entropy scheme is a last step at the senders end before information is being transferred whereas decoding, un-quantization and inverse 2D wave atoms decomposition are applied to reconstruct the transformed image.

#### 4.8.1 Comparative Results

A sample fingerprint image used in evaluating the performance of the proposed compression algorithm is shown in Figure 4.11(a). Compressed images at a Compression Ratio (CR) of 12.9 using FBI's Wavelet Scalar Quantization (WSQ), VQ on wave atoms decomposition and the proposed MSVQ based wave atoms scheme are shown in Figure 4.11(b), Figure 4.11(c) and Figure 4.11(d) respectively. It is evident from Figure 4.11(d) that the proposed method using wave atoms decomposition performs better and preserves fine fingerprint details and contains less blur in comparison with FBI's WSQ standard and VQ based wave atoms decomposition.

Table 4.3 compares the PSNR values obtained using our proposed method, for different bit rates, against various methods. As shown in Table 4.3, fingerprint compression based on wave atoms decomposition and multistage vector quantization produces a significant improvement in PSNR at low bit rates (high compression ratios). We compare the performance of our proposed MSVQ based compression standard with WSQ [85] and our VQ based compression algorithm [103]. In [103], we developed a compression scheme based on linear vector quantization of decomposed wave atoms representation of fingerprint images; wave atoms decomposed coefficients were thresholded and a significance map matrix and a significant coefficient vector generated. The significance map was divided into non-overlapping blocks, vectorized and quantized using a K-means vector quantization scheme with 64 code words. Results are compared at higher compression since most standards perform reasonably well at lower compression rates but their performance drastically deteriorates at higher values. At a compression ratio of 19 : 1, genetic algorithm based multiresolution analysis algorithm [102] achieves a maximum MSE improvement of 16.71% and a PSNR gain





Figure 4.11: Performance comparison of various methods at CR = 12.9:1. (a) Original fingerprint image (b) Compressed image using FBI's WSQ (c) Compressed image using wave atoms decomposition and VQ (d) Compressed image using the proposed MSVQ based method

of 0.794 dB over FBI's WSQ. However, our proposed algorithm offers a maximum PSNR gain of 8.07 dB at a compression ratio of 16 : 1 compared to WSQ.

Additional experiments are performed to establish the improved performance of our proposed method against state-of-the-art schemes. Two publicly available fingerprint images, whorl and tented arch type [110], are used and compared against contourlets [98] for varying combinations of filter banks. The subjective evaluations

Bit Rate	FBI's WSQ [85]	WaveAtoms+VQ	WaveAtoms+MSVQ
(bpp)	PSNR (dB)	PSNR (dB)	PSNR (dB)
0.1	23.72	28.62	29.98
0.15	25.15	29.06	30.53
0.2	26.25	31.21	33.23
0.25	27.14	31.52	34.84
0.3	27.96	31.68	35.78
0.35	28.72	32.36	36.07
0.4	29.36	32.42	36.59
0.45	29.89	32.47	37.38
0.5	30.37	32.65	38.44

Table 4.3: Bit rate vs. PSNR for FBI's WSQ and the proposed methods



Figure 4.12: Original and reconstructed images using our proposed method (a-b) tented arch type image (c-d) whorl type image

for achieved compression using wave atoms decomposition for tented arch and whorl type images are presented in Figure 4.12. It is clearly evident that the reconstructed images at a rate of 1 bpp have close resemblance with the source images along with all the minute details being precisely preserved.



Figure 4.13: PSNR analysis for tented arch type fingerprint image using haar wavelets, haar based contourlets and wave atoms



Figure 4.14: PSNR analysis for tented arch type fingerprint image using different contourlets and wave atoms

For fair analysis we compared PSNR values obtained for various methods as presented in Figure 4.13 and Figure 4.14. The methods used for performance analysis against our proposed scheme include Haar wavelets, contourlets based on varying combinations of Pyramidal and Directional filter banks (5-3, 9-7, Haar and pkva). It is evidently clear that the PSNR values for compression based on wave atoms decomposition is monotonically increasing and consistently higher against other methods for rising compression rates. From Figure 4.13, it is noticeable that the performance of Haar wavelets and contourlets is analogous however none of the above methods visibly outperform the other.

## Chapter 5

# **Fingerprint Matching**

#### 5.1 Introduction

Fingerprints are the appearance of graphical ridge and valley patterns on the tips of human fingers. Owing to their uniqueness and stability [111], [112], the use of fingerprints is considered to be one of the most popular biometric modality for personal verification. Automated recognition of an individual based on fingerprints is preferred since they are less vulnerable to be copied, stolen and lost [113]. The growing needs of law enforcement agencies and opportunities arising for civilian applications mean that automated fingerprint matching systems are becoming increasingly prevalent. Taking into account the nature of most civilian and criminal applications, the reliability of a biometric system is particularity important since the consequences of false matching can often lead to irreparable damage while false rejection may be highly deplorable. Fingerprints are routinely used in forensic laboratories and identification units around the world [114] and are accepted as witness in the courts of law for nearly a century [115]. Despite significant development in automated recognition, reliable automatic fingerprint verification is still a challenging problem [116].

A fingerprint is the pattern of friction ridges on a human finger that provides increased friction for gripping. Scientists believe that the friction ridges are constructed



Figure 5.1: Fingerprint image with the most common ridge characteristics

from small *ridge units* whose size, shape, density and alignment are remarkably unique to individuals. During friction ridge formation, *ridge units* are merged into various ridge characteristics, the most representative of which are ridge bifurcations and endings. Fingerprints are extremely unique that no two persons, not even twins share exactly the same location, shape and inter-relationship of these ridge characteristics [117]. However, friction ridge formation is genetically controlled as statistically significant familial correlation and high heritability estimates have been observed for some of the ridge characteristics such as forks, endings and the total number of ridge characteristics [118].

Various fingerprint acquisition techniques have been developed. Based on the acquisition process, a fingerprint is either acquired as a *latent* or a *tenprint*. A *latent* acquisition is one in which an accidental impression left behind by an unknwon individual whereas a *tenprint* consists of a set of fingerprint impressions collected with an individual's consent. A complete *tenprint* is a collection of 14 fingerprint images collected as rolls, plains and slaps from all the ten fingers.

### 5.2 Typical Challenges

A fingerprint matching algorithm compares two given fingerprints and returns a degree of similarity between the template and the test image. Few matching algorithms operate directly on the grayscale values and most of them require that an intermediate fingerprint representation be derived through feature extraction. Fingerprint matching is an extremely difficult task, mainly due to large intra-class variations in different impressions of the same finger. The main factors responsible for intra-class variations are:

- *Displacement*: The same finger may be placed at different sensor locations at different time instances resulting in a global translation of fingerprint area.
- Rotation: Fingers may be rotated at different angles with respect to the sensor surface during different acquisition sessions. Inspite of guides mounted in certain commercial scanners, involuntary finger rotations of up to  $\pm 20$  degrees have been observed.
- *Pressure and skin conditions*: Ridge structure is accurately captured if uniform contact is established between all the fingers and the sensor surface. However due to variations in finger pressure, skin dryness, skin disease, sweat, humidity, dirt and grease a non-uniform contact is established, which causes noisy acquisition of fingerprint images.
- *Partial overlap*: Fingerprint displacement and rotation often cause part of the fingerprint area to fall outside the sensor's field of view thus resulting in a small Region of Interest (ROI) overlap between the template and the input fingerprints.
- *Non-linear distortion*: Sensing 3D shape of a fingerprint on a 2D sensor surface results in non-linear distortion.

- *Noise*: This is mainly introduced into the system due to the presence of left over residues on the glass platen as a result of previous fingerprint capture.
- *Feature extraction errors*: Feature extraction algorithms are imperfect and often introduce measurement errors. Errors include estimation of orientation and frequency images, detection of the number, type and position of the singularities, and segmentation of the fingerprint area from the background.

Figure 5.2 shows one pair of fingerprint image with high variability (large *intra-class* variation) and another pair of fingerprint impressions with small *intra-class* variation. To test the effectiveness of our proposed fingerprint matching algorithm we used benchmark fingerprint verification competition (FVC) datasets, namely, FVC 2000, FVC 2002 and FVC 2004.



Figure 5.2: (a-b) Impressions from the same finger look significantly different (large intra-class variation). (c-d) Impressions from different fingers look similar to an untrained eye (small interclass variation) [119]

#### 5.2.1 FVC Datasets

FVC datasets are used to evaluate and compare emerging fingerprint matching algorithms, and are not intended for performance evaluation in a real application. It attempts to track the state-of-the-art in fingerprint recognition and provide updated benchmarks and a testing protocol for fair and unambiguous evaluation of fingerprint verification algorithms. Each FVC collection of dataset contains four distinct fingerprint databases (DB1,DB2,DB3,DB4) generated using different image acquisition techniques. Each database contains 8 fingerprints of each of the 100 distinctive subjects. Detailed description of the image acquisition techniques used for generating a particular database, image size and resolution are given in Tables 5.1, 5.2 and 5.3 respectively . Sample images from each of the FVC2000, FVC2002 and FVC2004 dataset are also shown below.

	Sensor	Sensor Model	Image Size	Resolution
	Technology	(Manufacturer)		
DB1	Low-cost	Secure desktop scanner	$300 \times 300$	500 dpi
	Optical sensor	(KeyTronic)		
DB2	Low-cost	TouchChip	$256 \times 364$	500 dpi
	Capacitive sensor	(ST Microelectronics)		
DB3	Optical sensor	DF-90	$448 \times 478$	500 dpi
		(Identicator Technology)		
DB4	Synthetic fingerprint generator		$240 \times 320$	500 dpi

Table 5.1: Image details of each of the four FVC2000 database

### 5.3 Fingerprint Feature Representation

Fingerprint images are not directly compared in fingerprint matching. Instead, a set of salient and discriminatory features that represent the underlying fingerprint characteristics are extracted from the images before matching. Feature extraction is extremely critical in reducing dimensionality of the data since a raw image is high



Figure 5.3: Sample fingerprint images from FVC 2000 dataset

	Sensor	Sensor Model	Image Size	Resolution
	Technology	(Manufacturer)		
DB1	Optical sensor	TouchView II	$388 \times 374$	500 dpi
		(Identix)		
DB2	Optical sensor	FX2000	$296 \times 560$	569 dpi
		(Biometrika)		
DB3	Optical sensor	100SC	$300 \times 300$	500 dpi
		(Precise Biometrics)		
DB4	Synthetic fingerprint generator v2.51		$288 \times 384$	500 dpi

Table 5.2: Image details of each of the four FVC2002 database

dimensional and contains redundant information. Typically features are more robust to noise and distortion than raw gray level pixel values. As a result, the performance of a fingerprint matching algorithm greatly depends on the selection and extraction of fingerprint features.

In manual feature extraction, a large variety of features have been established,



Figure 5.4: Sample fingerprint images from FVC 2002 dataset

Table 5.3: Image details of each of the four FVC2004 database

	Sensor	Sensor Model	Image Size	Resolution
	Technology	$({ m Manufacturer})$		
DB1	Optical sensor	V300	$640 \times 480$	500 dpi
		(CrossMatch)		
DB2	Optical sensor	U.are.U4000	$328 \times 364$	500 dpi
		(Digital Persona)		
DB3	Thermal sweeping	FingerChip FCD4B14CB	$300 \times 480$	$512 \mathrm{~dpi}$
	sensor	(Atmel)		
DB4	Synthetic fingerprint generator v3.0		$288 \times 384$	500 dpi

encoded and recorded based on their observed evidential value during decades of forensic practice. These characteristic features are generally categorized into three levels [120]. Level 1 features, or patterns, are the macro details of the fingerprint such



Figure 5.5: Sample fingerprint images from FVC 2002 dataset

as ridge flow and pattern type (loop, arch, whorl etc.) [117] as shown in Figure 5.6. Level 2 features occur on individual ridge paths, including the turns that each ridge takes and the places where ridges terminate or split, such as ridge bifurcations and endings [117]. Unlike Level 1 patterns, Level 2 features have individualization power and contribute significantly to reliable fingerprint matching. On an average, a fingerprint generally contains 75-175 minutiae. However, at times only a small number of minutiae are available in the captured fingerprint image and extraction of additional Level 3 features may be necessary. Level 3 features include all dimensional attributes of the ridge such as ridge width, shape, pores, incipient ridges, breaks, creases, scars, and other permanent details [117]. Therefore, a higher feature level contains finer details.

In automatic feature extraction, pre-compiled algorithms are used to determine the strength of ridge-valley signals in order to determine the area where extractable features are present. Later, a set of pre-defined features are extracted and encoded for matching. Unlike manual matching, the features extracted in automatic matching do not encompass a particular physical counterpart (e.g., local orientation map or filter responses) as long as they lead to high matching accuracy.



Figure 5.6: Fingerprint classes with marked core (red) and delta (blue) points [121]

#### 5.4 Literature Review

Analogous to other pattern recognition problems, there are different approaches to fingerprint matching. Correlation-based techniques utilize gray level information of an image and take into account all dimensional attributes of a fingerprint, thereby providing enough image resolution. Fingerprint images are superimposed and correlation between corresponding pixels is computed for varying rotations and translations. These techniques have been successfully applied for fingerprint matching [122] but they suffer from extremely high computational cost.

Ridge bifurcations and endings are the most common minutiae that are found in plenty in every individual's fingerprint. The occurrence of other minutiae types, such as islands, dots, enclosures, bridges, double bifurcations, trifurcations, are relatively rare. Minutiae-based techniques that use minutiae points such as ridge endings or bifurcations as features for matching are the most popular approaches in literature [123], [124], [125], [126]. Minutiae based techniques [127] extract minutiae from two fingerprints and store them as sets of points in a 2D plane. A match is established by searching alignments between the template and the input minutiae set that results in maximum pairings. Each minutiae is described by its location in the fingerprint, orientation and type of minutiae either ridge ending or ridge bifurcation. In minutiae based methods singularities i.e. core and delta points are used to align the images. Despite their relative simplicity and storage efficiency, minutiaebased fingerprint matching has its own limitations. Firstly, it is not always easy to extract minutiae points accurately, especially for low-quality images and thus ridge patterns [120] are reliably extracted for classification. Secondly, minutiae points do not necessarily embody the most significant component of the rich discriminatory information available in the fingerprints. In addition, there are difficulties related to aligning the minutiae point patterns from the query and template fingerprints due to the lack of knowledge about the correspondence between two point sets. Generally, finding the best alignment between two minutiae sets is an extremely difficult problem [123].

Texture features such as orientation fields [128] and ridge patterns (ridge shape and ridge density) [129], [130] have also been used for fingerprint matching. In comparison to minutiae, texture features are more robust with respect to the extraction process and less sensitive to noise. However, these features do not adequate discriminative information as fingerprints from different fingers may share similar orientation fields and/or ridge patterns. Typically, matching two fingerprints using such features requires a proper alignment of the query and template fingerprints. However, similar to the alignment of minutiae point patterns, establishing an accurate alignment between two orientation fields or ridge patterns is a complex task. In [129], [130], the ridge patterns are aligned based on a single reference point, however, such schemes are not robust with respect to errors in the location of the reference point. Liu et al. [128] aligned orientation fields using a steepest descent algorithm that is sensitive to initial alignment configurations and susceptible to local optima.

Different features available in fingerprint images have also been combined to improve the performance of a fingerprint matching system [131], [132], [133]. The which features are usually incorporated at the verification stage in the processing chain. Ross *et al.* [132] merged minutiae with ridge flow information to verify fingerprints by initially aligning images using minutiae information by the elastic string matching technique [123]. The verification is then carried out using the minutiae set along with ridge feature map matching. In [131], [133] fingerprint images are first aligned based on minutiae points using Generalized Hough Transform (GHT) [134] and later verification is performed by combining minutiae with text feature matching. These schemes are essentially minutiae-based and their performance greatly depends on the quality of minutiae information. Erroneous detections during minutiae extraction are propagated to the alignment stage and may eventually lead to meaningless match scores.

Researchers have also used fast Fourier transforms (FFT) and multi-resolution analysis tools to extract global features from fingerprint images for classification. Fitz and Green [135] used a hexagonal FFT to transform fingerprint images into frequency domain and employed a *wedge-ring detector* to extract features. A fingerprint classifier based on wavelet transform and probabilistic neural network was proposed in [136]. Wilson *et al.* [137] developed a FBI fingerprint classification standard that incorporates a massively parallel neural network structure. In [129], a Gabor filter bank was used for fingerprint matching and Park *et al.* [130] proposed the use of Directional Filter Bank (DFB) for efficient fingerprint feature extraction and matching. Parsons *et al.* [138] proposed a pore extraction method for 1000 dpi fingerprint images using a Difference of Gaussian (DOG) filtering that approximates the Mexican-hat wavelet [139]. Neural network schemes based on self organizing feature map, fuzzy neural networks, Radial Basis Function Neural Network (RBFNN) and Ellipsoidal Basis Function Neural Network (EBFNN) have also been proposed [140].

#### 5.5 Proposed Fingerprint Matching Algorithm

In this chapter, a fast and accurate fingerprint matching algorithm that extracts sparse fingerprint representation using wave atoms decomposition is presented. The generated coefficients are dimensionally reduced using bidirectional two-dimensional principal component analysis. An ELM classifier is trained and tested using dimensionally reduced extracted features. The proposed recognition algorithm requires minimum human interventions and performs learning at multiple folds faster speed than conventional neural networks. ELM determines network parameters analytically, avoids trivial human intervention and makes it efficient for real-time applications.

The proposed scheme is independent of fingerprint patterns and is based on individual features and the number of trained fingerprint classes. Table 5.4 consists of detailed steps that demonstrate our proposed technique. Our system classifies fingerprint images into one of the trained classes; therefore, only one verification process is required per image. Our proposed scheme deals with recognition of fingerprint images using ELM design and utilizes dimensionally reduced feature vectors generated using wave atoms decomposition. Wave atoms decomposition is used for sparse representation of fingerprint images since they belong to a category of images that oscillate smoothly in varying directions. Discrete 2D wave atoms decomposition is applied on the original fingerprint image to efficiently capture coherence patterns along and across the oscillations. Fingerprint images are digitized using 256 gray levels therefore a transformation in color space is not required. Dimension of fingerprint images is reduced to  $64 \times 64$  prior to wave atoms decomposition. Image resizing is the only pre-processing performed on all datasets to minimize computational cost and to guarantee uniformity with other methods used for comparison. An orthonormal basis is used instead of a tight frame since each basis function oscillates in two distinct directions instead of one. This orthobasis variant property is important in applications where redundancy is undesired.

Table 5.4: Outline of the proposed fingerprint matching algorithm

**INPUT:** Randomly divide image database into two subsets  $TR_v$  and  $TE_{\varsigma}$  where  $v = 1, 2, ..., \mu$  and  $\varsigma = 1, 2, ..., \nu$  representing training and test image sets respectively.

**OUTPUT:** Classifier - f(x)

1. Resize images from all database to  $\mathcal{R} \times \mathcal{C}$ 

2. Compute the wave atoms decomposition of each training and test images and extract feature sets. Each feature set is of dimension  $\mathcal{R}x\mathcal{C}$  (Refer to Chapter 2.4 for details of wave atoms decomposition)

3. Calculate image covariance matrix of test and train images to obtain intermediate feature matrix

$$\begin{aligned} \mathbb{G}_{TR_{\upsilon}} &= \frac{1}{\mu} \sum_{\upsilon=1}^{\mu} (\mathbb{A}_{\upsilon} - \check{\mathbb{A}})^{T} (\mathbb{A}_{\upsilon} - \check{\mathbb{A}}) \\ \mathbb{G}_{TE_{\varsigma}} &= \frac{1}{\nu} \sum_{\varsigma=1}^{\nu} (\mathbb{A}_{\varsigma} - \check{\mathbb{A}})^{T} (\mathbb{A}_{\varsigma} - \check{\mathbb{A}}) \end{aligned}$$

4. Evaluate the maximizing criteria using the image covariance matrix according to the method described in Chapter 3.6 and generate B2DPCA based feature vectors, each of size  $\mathcal{U} \times \mathcal{V}$ 

5. Vectorize the B2DPCA based feature vectors obtained in the previous stage and train the ELM classifier

6. Classify images with test feature vectors using trained ELM

We randomly divide image database into two, namely, training set and testing set. All images within each database have the same dimension, i.e.  $\mathcal{R} \times \mathcal{C}$ . Similar image sizes support the assembly of equal sized wave atoms coefficients and feature vector extraction with identical level of global content. 2D wave atoms decomposition of every image is computed and coefficients are saved as initial feature matrix. Wave atoms decomposition is a relatively new technique for multiresolution analysis that offers significantly sparser expansion, for oscillatory functions, than other fixed standard representations like wavelets, curvelets and Gabor atoms.

Application of ELM based classification on original wave atoms coefficients is computationally expensive due to higher dimensionality of data originating from large image datasets. Outliers and irrelevant image points being included into classification task can also degrade the performance of our algorithm; hence B2DPCA is employed to reduce dimensionality of initial feature vectors. The optimal projection axes is calculated and dimensionality of every image is reduced along its columns to generate new image sets. An image covariance matrix of the new image sets is generated, its optimal projection axes is evaluated and the row dimension of the newly generated image is reduced to produce final feature matrix. Application of B2DPCA using the modified approach retains better structure and correlation information amongst neighboring pixel coefficients. Dimensionally reduced wave atoms coefficients are vectorized into a  $\mathcal{U} \times \mathcal{V}$  dimension vector, where  $\mathcal{U} \times \mathcal{V} \ll \mathcal{R} \times \mathcal{C}$ . ELM is trained using labeled B2DPCA feature vectors and recognition accuracy achieved using the trained network.

#### 5.6 Experimental Results

Extensive experiments are performed using standard and distinctive collections of fingerprint datasets, i.e., FVC2000, FVC2002 and FVC2004 [111] to test the practicality of our proposed method. All images are resized to  $64 \times 64$  in our experiments and 5 images from each database are used as prototypes and the remaining 3 for testing to ensure consistency with other methods used for comparison. Experiments were also performed on original fingerprint images and consistently better results were obtained since detailed fingerprint information was incorporated at the expense of large feature vectors. Both the testing and training sets of images are decomposed using 2D wave atoms transform using an orthonormal basis function and dimensionally reduced through application of B2DPCA. Dimensionally reduced features are vectorized and classification is performed by using ELM. The above process was repeated 10 times for all the databases and averaged results of few experiments are documented in the paper. The recognition accuracy for Db1 database from FVC2000, FVC2002 and FVC2004 is compared with Wavelet Transform (WT) based RBFNN and EBFNN fingerprint recognition algorithms. Results, obtained with the proposed method (only 6 principal components are used for consistency with other methods), are compared with the accuracy reported in [140] using WT-2DPCA-RBFNN and WT-2DPCA-EBFNN.

Database	WT-2DPCA-	WT-2DPCA-	Proposed
	RBFNN	EBFNN	Method
FVC2000	91	91	93.25
FVC2002	87	87	92.63
FVC2004	86.5	87	89.62

Table 5.5: Comparative results for various methods

We conclude from the results in Table 5.5 that our proposed fingerprint classification algorithm performs significantly better than the wavelet based RBFNN and EBFNN fingerprint classification algorithms. In addition to the improved classification accuracy, our proposed ELM based scheme performs training and testing thousands folds faster than conventional neural network based classification algorithms[69].

Classification accuracy for various databases from each of the three FVC datasets are plotted in Figures 5.7, 5.8, 5.9, 5.10 for varying number of principal components. It is evident from the Figures that several factors influence classification accuracy, namely, fingerprint acquisition techniques, climatic and environmental conditions and most notably the number of principal components. Dataset Db4 from each of the databases is generated using a synthetic fingerprint generator; consequently the effects of environment and other irrepressible conditions are trifling and are substantiated



Figure 5.7: Recognition accuracy for DB1 database



Figure 5.8: Recognition accuracy for DB2 database

by improved classification accuracy at low principal components. It is also worth noting that increasing the number of principal component does not always lead to an improved classification.



Figure 5.9: Recognition accuracy for DB3 database



Figure 5.10: Recognition accuracy for DB4 database

### Chapter 6

## **Conclusions and Future Research**

Security applications like user authentication for access to physical and virtual spaces provide and ensure higher security. Now days, Robust face recognition systems are in great demand to deal with crime and terrorism. However, identifying a person by taking an input face image and matching with the images present in the database is still a very challenging problem. This is due to the variability of human faces under different operational scenarios such as: illumination, rotation, expression, occlusion, camera view point, aging, makeup, and eyeglasses. These conditions immensely affect the performance of a face recognition system especially, when a match is to be established against large scale databases. This under performance of automated face recognition system prevents their deployment in real applications. Therefore, additional biometrics i.e., fingerprints are incorporated to improve the system performance and to facilitate its application to real world problems. Fingerprints are the manifestation of interleaved ridge and valley patterns on the tips of human fingers. Fingertip ridges have evolved over the years to allow humans to grasp and grip objects. Similar to other body parts, fingerprint ridges are also formed through a combination of genetic and environmental factors. In fact, fingerprint formation is similar to the growth of capillaries and blood vessels in angiogenesis. The genetic code in DNA gives general instructions on the way skin should form in a developing foetus. This is the most important reason why even fingerprints of identical twins are different. Consequently, human faces are used in conjunction with fingerprints by various law enforcement agencies and border security forces to track suspects and criminals. Automatic human identification systems have played a major role in commercial, governmental and forensics applications. However, these systems have not yet completely eliminated the need for manual examination, especially in matching latent fingerprints. Automatic systems utilize a limited feature set compared to human experts and are not able to easily adapt to handle variations in image quality and resolution. In this thesis, we proposed an enhanced and robust human face recognition algorithm for potential law enforcement applications. We also developed a generic fingerprint compression algorithm based on state of the art multiresolution analysis tools, i.e., wave atoms to speed up data archiving and recognition. Finally, we proposed an improved fingerprint matching algorithm to overcome some of the challenges associated with traditional fingerprint recognition algorithms. A summary of the thesis is given below.

**Chapter 1** provided a brief introduction of biometrics as a tool for security and identification. We reviewed some of the important biometric modalities being used for various commercial, governmental and forensic applications and briefly described the architecture of a typical biometric system. As mentioned earlier, we focused our discussion on two important biometric *traits*; namely, face and fingerprint. We commented on the differences between a verification module and the identification module. Lastly, we laid down the research objectives and included a brief outline of the thesis.

**Chapter 2** included a summary of the history of mathematical analysis tools, their limitations and multiresolution concept. We discussed in length about the implementation and mathematical details associated with DWT. We also included a thorough discussion of curvelet transform and provided a brief discussion highlighting the key implementation issues of FDCT using the wrapping technique. Finally, we

compared various multiresolution tools and reviewed the wave atoms decomposition including its application to both 1D and 2D signals.

**Chapter 3** proposed two algorithms for automatic face recognition using curvelets. Firstly, we developed a novel human face recognition system using curvelet transform and KPCA. In the past features extracted from curvelet subbands were dimensionally reduced using PCA for obtaining an enhanced representative feature set. In this work we used KPCA to generate a comprehensive feature set. kNN based classification scheme was employed for ascertaining accuracy. Experiments were performed using popular human face databases and significant improvement in recognition accuracy was achieved. The proposed method considerably outperformed conventional face recognition systems using standard PCA.

Secondly, we proposed an efficient human face recognition technique based on B2DPCA and ELM. Curvelets are used for image decomposition and subband exhibiting a highest standard deviation was selected. A B2DPCA algorithm was proposed to achieve superior and unique feature sets and minimize computational complexity of our framework. The feature sets were trained and tested using a fast and accurate ELM classifier. Extensive experiments were performed using 7 challenging databases and results substantiated our claim that the proposed method achieves improved recognition rate with a considerably smaller time complexity. In addition, our proposed method was also independent of the number prototypes used for training, scales of curvelet decomposition and the number of hidden neurons. Border security, video surveillance and database security are some areas where face categorization plays a very critical role and these applications can potentially benefit from our proposed recognition scheme.

**Chapter 4** thoroughly discussed two fingerprint compression algorithms based on orthobasis variant of wave atoms decomposition. They have been specifically designed for enhanced representation of oscillatory patterns and to convey temporal and spatial information. Wave atoms efficiently captured coherence of the fingerprint images along and across the oscillations. Initially, we proposed a compression scheme based upon linear vector quantization of decomposed wave atoms representation of fingerprint images. Later quantized information was encoded with arithmetic entropy scheme. The proposed image compression standard significantly outperformed the FBI fingerprint image compression standard.

We also proposed a compression scheme based upon multistage vector quantization of processed wave atoms representation of fingerprint images. Wave atoms expansion was processed using mathematical morphological operators to emphasize and retain significant coefficients for transmission. The scalar quantized significant coefficients and MSVQ quantized significance map are transformed using an arithmetic encoder. The proposed image compression standard outperformed other well established methods and achieved a PSNR gain of 8.07 dB in comparison to FBI's wavelet scalar quantization. The results obtained drastically outperformed genetic algorithm based fingerprint compression methods that generated a maximum MSE improvement of 16.71% and a PSNR gain of 0.794 dB over the FBI fingerprint compression standard. We also compared our proposed algorithm to a contourlet based fingerprint compression algorithm and improved results were achieved. Law enforcement and forensic applications can potentially benefit from our compression scheme.

**Chapter 5** discussed a fast and accurate technique for fingerprint matching based on wave atoms decomposition. B2DPCA was used to obtain improved feature sets and the system was classified using ELM. The foremost contribution of our method is the application of 2D wave atoms decomposition on original fingerprint images to obtain sparse and efficient coefficients. Secondly, distinctive feature sets were extracted using our dimensionality reduction technique. ELM eliminated limitations of classical training paradigm and trained our network at a considerably fast speed. Our algorithm combined optimization of B2DPCA and the speed of ELM to put together a superior and efficient algorithm for fingerprint classification. Experiments were performed on twelve distinct fingerprint databases and results compared against wavelet based fingerprint matching techniques.

#### 6.1 Summary of Contributions

In conclusion, the main contributions of this thesis are outlined below:

- 1. Two curvelet based algorithms were proposed for improved human face recognition. The proposed methods were tested using challenging face databases and significantly improved performance was achieved.
- 2. A new dimensionality reduction technique, i.e., B2DPCA was developed to generate distinctive feature vectors. The proposed dimensionality reduction technique is mathematically plausible since it first reduces dimensionality along the column dimension. Later, the transformed image sets are dimensionally reduced along the row direction to eventually achieve the desired data dimension.
- 3. Wave atoms based fingerprint compression schemes were put forth. To the best of my knowledge there is no other work in literature that has dealt with fingerprint compression using wave atoms decomposition, therefore, this represents the most significant contribution of my research.
- 4. Finally, we proposed a fingerprint matching algorithm using wave atoms decomposition and compared results with existing wavelet based approaches.

In future, I would like to exploit dependency amongst wave atoms coefficients at various scales and orientations to facilitate automated threshold computation for improved compression and to further improve the PSNR gains for fingerprint images. With regards to fingerprint matching, specific information regarding local features is completely lost at the expense of global features. Researchers are actively working on matching fingerprints in situations where the degree of overlap is quite low and it is only feasible if local features are stored and matched. Research has also shown that using key point alone for detection sacrifices shape information available in smooth portions of object contours thus approaches based on extracting edge points are not universally applicable. Therefore, I would like to develop a correlation based fingerprint matching algorithm using local brightness and texture gradients to generate oriented edge channels. It is perceived that the use of localized fingerprint information would boost accuracy in situations where partial overlap exists among fingerprint images.

# References

- A. K. Jain, A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition", IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image and Video-Based Biometrics, vol. 14, issue 1 (2004), pp. 4-20.
- [2] A. K. Jain, P. Flynn, A. Ross, "Handbook of Biometrics", Boston, Massachusetts: Springer Series (2008).
- [3] A. K. Jain, R. Bolle, S. Pankanti, "Biometrics: Personal Identification in Networked Society", Kluwer Academic Publishers, (1999).
- [4] J. L.Wayman, A. K. Jain, D. Maltoni, D. Maio, "Biometric Systems: Technology, Design and Performance Evaluation", Springer, (2005).
- [5] Pennsylvania Department of Aging. http://www.aging.state.pa.us/psonlinetraining/cwp.
- [6] S. G. Mallat, "A Wavelet Tour of Signal Processing", San Diego, California: Academic Press (1999).
- [7] B. B. Hubbard, "The World According to Wavelets", Wellesley, Massachusetts: A.K. Peters (1998).
- [8] S. G. Mallat, " A Theory of Multiresolution Signal Decomposition: the wavelet representation", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 11 (1989), pp. 674-693.

- [9] E. J. Cands, D. L. Donoho, "Curvelets A Surprisingly Effective Nonadaptive Representation for Objects with Edges", In: A. Cohen, C. Rabut and L. Schumaker, Editors, Curves and Surface Fitting: Saint-Malo 1999, Vanderbilt University Press, Nashville (2000), pp. 105-120.
- [10] E. J. Candes, "Ridgelets: Theory and Applications", Ph.D. Thesis, Department of Statistics, Stanford University (1998).
- [11] E. J. Candes, "Monoscale Ridgelets for the Representation of images with Edges", Technical Report, Statistics, Stanford University (1999).
- [12] E. J. Candes, "On the Representation of Mutilated Sobolev Functions", Technical Report, Statistics, Stanford University (1999).
- [13] E. J. Candes, D. L. Donoho, "Ridgelets: The Key to High-Dimensional Intermittency?", Phil. Transactions Royal Society of London, A. 357 (1999), pp. 2495-2509.
- [14] D. L. Donoho, M. R. Duncan, "Digital Curvelet Transform: Strategy, Implementation and Experiments" Proceedings of SPIE - The International Society for Optical Engineering, vol. 4056 (2000), pp. 12-30.
- [15] M. N. Do, M. Vetterli, "The Finite Ridgelet Transform for Image Representation", IEEE Transactions on Image Processing, vol. 12, no. 1 (2003), pp. 16-28.
- [16] M. N. Do, M. Vetterli, "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation", IEEE Transactions on Image Processing, vol. 14, no. 12 (2005), pp. 2091-2106.
- [17] J. L. Starck, Image Processing by the Curvelet Transform, powerpoint presentation.

- [18] J. L. Starck, N. Aghanim, O. Forni, "Detecting Cosmological Non-Gaussian Signatures by Multiscale Methods", Astronomy and Astrophysics, vol. 416, issue 1 (2004), pp. 9-17.
- [19] J. L. Starck, M. Elad, D. L. Donoho, "Redundant Multiscale Transforms and their Application for Morphological Component Analysis", Advances in Imaging and Electron Physics, vol. 132 (2004), pp. 287-342.
- [20] E. J. Candes, L. Demanet, D. L. Donoho, L. Ying, "Fast Discrete Curvelet Transforms" Multiscale Model Simulation, vol. 5, issue 3 (2006), pp. 861-899.
- [21] F. J. Herrmann, U. Boniger, D. J. Verschuur, "Nonlinear Primary-Multiple Separation with Directional Curvelet Frames", Geophysical International Journal, vol. 170, no. 2 (2007), pp. 781-799.
- [22] B. Eriksson, www.homepages.cae.wisc.edu/ ece734/project/s06/eriksson.ppt.
- [23] J. P. Antoine, R. Murenzi, "Two-Dimensional Directional Wavelets and the Scale-Angle Representation", Elsevier Signal Processing, vol. 52 (1996), pp. 259-281.
- [24] L. Demanet, L. Ying, "Wave Atoms and Sparsity of Oscillatory Patterns", Applied and Computational Harmonic Analysis, vol. 23 (2007), pp. 368-387.
- [25] L. Villemoes, "Wavelet Packets with Uniform Time-Frequency Localization", In Comptes Rendus Mathematique, vol. 335 (2002), pp. 793-796.
- [26] L. Demanet, L. Ying, "Scattering in Flatland: Efficient Representations via Wave Atoms", In Foundation of Computational Mathematics (2008).
- [27] L. Demanet, L. Ying, "Curvelet and Wave Atoms for Mirror Extended Images", Proceedings of SPIE, vol. 6701 (2007), pp. 6701-6719.

- [28] P. J. Phillips, H. Wechsler, J. Huang, P. J. Rauss, "The FERET Database and Evaluation Procedure for Face-Recognition Algorithms", Elsevier Image and Vision Computing, vol. 16, no. 5 (1998), pp. 295-306.
- [29] L. Spacek, The Essex Faces94 Database, http://cswww.essex.ac.uk/ mv/allfaces/.
- [30] M. J. Lyons, J. Budynek, S. Akamatsu, "Automatic Classification of Single Facial Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 12 (1999), pp. 1357-1362.
- [31] A. V. Nefian, M. Khosravi, M. H. Hayes, "Real-Time Human Face Detection from Uncontrolled Environments", SPIE Visual Communications on Image Processing (1997), pp. 211-219.
- [32] D. B. Graham, N. M. Allinson, "Characterizing Virtual Eigensignatures for General Purpose Face Recognition", Face Recognition: From Theory to Applications, NATO ASI Series F, Computer and Systems Sciences, vol. 163 (1998), pp. 446-456.
- [33] F. Samaria, A. Harter, "Parameterisation of a Stochastic Model for Human Face Identification", 2<sup>nd</sup> IEEE Workshop on Applications of Computer Vision, (1994), pp. 138-142.
- [34] M. S. Bartlett, J. R. Movellan, T. J. Sejnowski, "Face Recognition by Independent Component Analysis", IEEE Transactions on Neural Networks, vol. 13 (2002), pp. 1450-1464.
- [35] A. M. Martinez, R. Benavente, "The AR face database", CVC Technical Report, No.24 (1998).
- [36] B. S. Manjunath, R. Chellappa, C. V. Malsburg, "A Feature Based Approach

to Face Recognition", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (1992), pp. 373-378.

- [37] G. Chow, X. Li, "Towards a System for Automatic Facial Feature Detection", Elsevier Pattern Recognition, vol. 26, issue 12 (1993), pp. 1739-1755.
- [38] A. Pentland, B. Moghaddam, T. Starner, "View-Based and Modular Eigenspaces for Face Recognition", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (1994), pp. 84-91.
- [39] D.Valentin, H. Abdi, A. J. O'Toole, G. W. Cottrell, "Connectionist Models of Face Processing: A Survey", Elsevier Pattern Recognition, vol.27 (1994), pp. 1209-1230.
- [40] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma, N. Otsu, "Face Recognition System using Local Autocorrelations and Multiscale Integration", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no.10 (1996), pp. 1024-1028.
- [41] D. L. Swets, D. L. Weng, "Using Discriminant Eigenfeatures for Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, issue 8 (1996), pp. 831-836.
- [42] W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, "Face Recognition: A Literature Survey", ACM Computing Surveys, vol. 35, issue 4 (2003), pp. 99-458.
- [43] M. Kirby, L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no.1 (1990), pp. 103-108.
- [44] M. A.Turk, A. P. Pentland, "Face Recognition using Eigenfaces", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (1991), pp. 586-591.

- [45] M. S. Bartlett, J. R. Movellan, T. J. Sejnowski, "Face Recognition by Independent Component Analysis", IEEE Transactions on Neural Networks, vol. 13, no. 6 (2002), pp. 1450-1464.
- [46] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, "Face Recognition using LDA-Based Algorithms", IEEE Transactions on Neural Networks, vol. 14, no. 1 (2003), pp. 195-200.
- [47] C. Liu, H. Wechsler, "Evolutionary Pursuit and its Application to Face Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 6 (2000), pp. 570-582.
- [48] L. Wiskott, J. M. Fellus, N. Kruger, C. VonDerMalsburg, "Face Recognition by Elastic Bunch Graph Matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no.7 (1997), pp. 775-779.
- [49] F. R. Bach, M. I. Jordan, "Kernel Independent Component Analysis", Journal of Machine Learning Research, vol. 3 (2002), pp. 1-48.
- [50] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, "Face Recognition using Kernel Direct Discriminant Analysis Algorithms", IEEE Transactions on Neural Networks, vol. 14, no. 1 (2003), pp. 117-126.
- [51] C. Liu, H. Wechsler, "A Unified Bayesian Framework for Face Recognition", Proceedings of International Conference on Image Processing (1998), pp. 151-155.
- [52] B. Moghaddam, T. Jebara, A. Pentland, "Bayesian Face Recognition", Elsevier Pattern Recognition, vol. 33, issue 11 (2000), pp. 1771-1782.
- [53] K. Jonsson, J. Matas, J. Kittler, Y. P. Li, "Learning Support Vectors for Face Verification and Recognition", Proceedings of IEEE Conference on Automatic Face and Gesture Recognition (2000), pp. 208-213.
- [54] B. Heisele, P. Ho, T. Poggio, "Face Recognition with Support Vector Machines: Global Versus Component-Based Approach", Proceedings of International Conference on Computer Vision, vol. 2 (2001), pp. 688-694.
- [55] G. C. Feng, P. C. Yuen, D. Q. Dai, "Human Face Recognition using PCA on Wavelet Subband", Journal of Electronic Imaging, vol. 9, issue 2 (2000), pp. 226-233.
- [56] J. T. Chien, C. C. Wu, "Discriminant Wavelet-Faces and Nearest Feature Classifiers for Face Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, issue 2 (2002), pp. 1644-1649.
- [57] B. L. Zhang, H. Zhang, S. Sam Ge, "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory", IEEE Transactions on Neural Networks, vol. 15, issue 1 (2004), pp. 166-177.
- [58] M. Zhao, P. Li, Z. Liu, "Face Recognition Based on Wavelet Transform Weighted Modular PCA", Proceedings of the Congress in Image and Signal Processing (2008), pp. 589-593.
- [59] T. Mandal, Q. M. Jonathan Wu, "Face Recognition using Curvelet Based PCA", Proceedings of International Conference on Pattern Recognition, (2008), pp. 1-4.
- [60] T. Mandal, Q. M. Jonathan Wu, Y. Yuan, "Curvelet Based Face Recognition via Dimension Reduction", Elsevier Signal Processing, vol. 89, issue 3 (2009), pp. 2345-2353.
- [61] A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Sid-Ahmed, "A Novel Technique for Human Face Recognition using Nonlinear Curvelet Feature Subspace", Lecture Notes in Computer Science, vol. 5627 (2009), pp. 512-521.
- [62] A. A. Mohammed, Q. M. J. Wu, M. A. Sid-Ahmed, "Application of Bidirectional Two-dimensional Principal Component Analysis to Curvelet Feature Based Face

Recognition", Proceedings of IEEE International Conference on Systems, Mann and Cybernetics (2009), pp. 4124-4130.

- [63] T. M. Cover, "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition" IEEE Transactions on Electronic Computers, vol. 14, issue 3 (1965), pp. 326-334.
- [64] T. M. Cover, J. A. Thomas, "Elements of Information Theory", John Wiley, Second Edition (2006).
- [65] B. Scholkopf, A. Smola, K. R. Muller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", Neural Computation, vol. 12, issue 5 (1998), pp. 1299-1319.
- [66] O. Boiman, E. Shechtman, M. Irani, "In Defense of Nearest-Neighbor Based Image Classification", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (2008), pp. 1-8.
- [67] J. Yang, D. Zhang, A. F. Frangi, J. Yang, "Two-Dimensional PCA: A New Approach to Appearance Based Face Representation and Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no.1 (2004), pp. 131-137.
- [68] D. Zhang, Z. H. Zhou, "(2D)2 PCA: Two-Directional Two-Dimensional PCA for Efficient Face Representation and Recognition", Elsevier Neurocomputing, vol. 69, issue 1 (2005), pp. 224-231.
- [69] G. Huang, Q. Zhu, C. Siew, "Extreme Learning Machine: Theory and Applications", Elsevier Neurocomputing, vol. 70, issue 1-3 (2006), pp. 489-501.
- [70] J. D. Eggerton, M. D. Srinath, "A Visually Weighted Quantization Scheme for Image Bandwidth Compression at Low Data Rates", IEEE Transactions on Communication, vol. 34 (1986), pp. 840-847.

- [71] J. D. Gibson, T. Berger, T. Lookabaugh, D. Linghbergh, R. L. Baker, "Digital Compression for Multimedia", San Francisco, California: Morgan Kaufmann Publishers (1998).
- [72] A. Gersho, R. M. Gray, "Vector Quantization and Signal Compression", Norwell, Massachusetts: Kluwer (1992).
- [73] N. Jayant, J. Johnston, R. Safranek, "Signal Compression Based on Models of Human Perception", Proceedings of the IEEE, vol. 81, issue 10 (1993), pp. 1385-1422.
- [74] K. Sayood, "Introduction to Data Compression", 2<sup>nd</sup> Edition, San Diego, California: Academic Press (2000).
- [75] R. C. Gonzalez, R. E. Woods, "Digital Image Processing", 2<sup>nd</sup> Edition, Upper Saddle River, New Jersey: Prentice Hall (2002).
- [76] N. Ahmed, T. Natarajan, K. R. Rao, "Discrete Cosine Transform", IEEE Transactions on Computer, vol. C-23, no. 1 (1974), pp. 90-93.
- [77] G. Strang, "The Discrete Cosine Transform", Society for Industrial and Applied Mathematics, vol. 41 (1999), pp. 135-147.
- [78] M. Vetterli, J. Kovavcevic, "Wavelets and Subband Coding", 1<sup>st</sup> Edition, Englewood Cliffs, New Jersey: Prentice Hall (1995).
- [79] R. R. Coifman, M. V. Wickerhauser, "Entropy Based Algorithm for Best Basis Selection", IEEE Transactions on Information Theory, vol. 38 (1992), pp. 713-718.
- [80] C. N. Brislawn, J. N. Bradley, R. Onyshczak, T. Hopper, "The FBI Compression Standard for Digitized Fingerprint Images", Proceedings of SPIE, vol. 2847 (1996), pp. 344-355.

- [81] J. N. Bradley, C. M. Brislawn, T. Hopper, The FBI Wavelet/Scalar Quantization Standard for Gray-Scale Fingerprint Image Compression, Proceedings of SPIE, Visual Information Processing II, vol. 1961 (1993), pp. 293-304.
- [82] R. Kidd, "Comparison of Wavelet Scalar Quantization and JPEG for Fingerprint Image Compression", Journal of Electronic Imaging, vol. 4, issue 1 (1995), pp. 31-39.
- [83] B. G. Sherlock, D. M. Monro, "Optimized Wavelets for Fingerprint Compression", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3 (1996), pp. 1447-1450.
- [84] B. G. Sherlock, D. M. Monro, "Psychovisually Tuned Wavelet Fingerprint Compression", Proceedings of International Conference on Image Processing, vol. 2 (1996), pp. 585-588.
- [85] J. N. Bradley, C. N. Brislawn, "Proposed First Generation WSQ Bit Allocation Procedure", Proceedings of Criminal Justice Information Servies Technology (1993), pp. C11-C17.
- [86] R. A. DeVore, B. Jawerth, B. Lucier, "Image Compression through Wavelet Transforms Coding", IEEE Transactions on Information Theory, vol. 38 (1992), pp. 719-746.
- [87] "Criminal Justice Information Services (CJIS) WSQ Gray-scale Fingerprint Image Compression Specification", Federal Bureau of Investigation, Document No. IAFIS-IC-0110(V3) (1997).
- [88] J. Scharinger, H. Gaderbauer, G. Wilflingseder, "Redesigning the FBI Fingerprint Compression Standard", Pattern Recognition and Medical Computer Vision, vol. 106 (1998), pp. 215-223.

- [89] S. Kasaei, M. Deriche, B. Boashash, "Fingerprint Compression Using a Modified Wavelet Transform and Pyramid Lattice Vector Quantization", Proceedings of TENCON, vol. 2 (1996), pp. 798-803.
- [90] J. M. Shapiro, "Embedded Image Coding using Zerotrees of Wavelet Coefficients", IEEE Transactions on Signal Processing, vol. 41 (1993), pp. 3445-3462.
- [91] A. Said, W. A. Pearlman, "A New, Fast and Efficient Image Codec based on Set Partitioning in Hierarchical Trees", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6 (1996), pp. 243-249.
- [92] P. Saeedian, B. Shirazi, "A Novel Fingerprint Image Compression Technique using Adaptive Subband Image Coding", Proceedings of Picture Coding Symposium (2004), pp. 139-143.
- [93] G. Khuwaja, Best Parameter Based Compression of Fingerprints with Wavelet Packets", International Journal of Computer Applications in Technology, vol. 19, issue 1 (2004), pp. 51-62.
- [94] J. Thärna, K. Nilsson, J. Bigun, "Orientation Scanning to Improve Lossless Compression of Fingerprint Images", Lecture Notes in Computer Science, vol. 2688 (2003), pp. 343-350.
- [95] R. Sudhakar, N. Vigesh,S. Jayaraman, "Application of Wavelet Footprints for Fingerprint Compression", International Journal on Graphics, Vision, and Image Processing, vol. 5, issue 6 (2005), pp. 39-45.
- [96] R. Sudhakar, R. Karthiga, S. Jayaraman, "Fingerprint Compression using Contourlet Transform with Modified SPIHT Algorithm", Iranian Journal of Electrical and Computer Engineering, vol. 5, issue 1 (2005), pp. 3-10.
- [97] R. Eslami, H. Radha, "Wavelet-Based Contourlet Transform and its Application

to Image Coding", Proceedings of International Conference on Image Processing, vol. 5 (2004), pp. 3189-3192.

- [98] S. Esakkirajan, T. Veerakumar, V. Murugan, R. Sudhakar, "Fingerprint Compression Using Contourlet Transform and Multistage Vector Quantization", International Journal of Biological and Medical Sciences, vol. 1, no. 2 (2006), pp. 140-147.
- [99] R. Eslami, H. Radha, "Wavelet-Based Contourlet Packet Image Coding", Conference on Information Sciences and Systems, John Hopkins University (2005).
- [100] U. Grasemann, R. Mikkulainen, "Evolving Wavelets using a Coevolutionary Genetic Algorithm and Lifting", Proceedings of Sixth Genetic and Evolution Computation Conference (2004), pp. 969-980.
- [101] F. Y. Moore, "Evolved Multi-resolution Analysis Transforms for Optimized Image Compression and Reconstruction under Quantization", WSEAS Transactions on Computers, vol. 1 (2006), pp. 97-104.
- [102] B. Babb, F. Y. Moore, "The Best Fingerprint Compression Standard Yet", Proceedings of IEEE International Conference on Systems, Mann and Cybernetics (2007), pp. 2911-2916.
- [103] A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Sid-Ahmed, "A Generic Fingerprint Image Compression Technique Based on Wave Atoms Decomposition", Proceedings of International Conference on Image Processing (2009), pp. 2837-2840.
- [104] A. A. Mohammed, R. Minhas, Q. M. J. Wu, M. A. Sid-Ahmed, "An Efficient Fingerprint Image Compression Standard Technique Based on Wave Atoms Decomposition and Multistage Vector Quantization", Journal of Integrated Compute Aided Engineering, vol. 17, issue 1 (2010), pp. 29-40.

- [105] Y. Linde, A. Buzo, R. M. Gray, "An Algorithm for Vector Quantizer Design", IEEE Transactions on Communication, vol. 28 (1980), pp. 84-95.
- [106] J. Serra, "Image Analysis and Mathematical Morphology", Academic Press, New York (1988).
- [107] E. Morales, F. Shih, "Wavelet Coefficients Clustering using Morphological Operations and Pruned Quadtree", Elsevier Pattern Recognition, vol. 10, issue 10 (2000), pp. 1611-1620.
- [108] H. Khalil, K. Rose, "Multistage Vector Quantizer Optimization for Packet Networks", IEEE Transactions on Signal Processing, vol. 51 (2003), pp. 1870-1879.
- [109] A. A. Mohammed, R. Minhas, Q. M. J. Wu, M.A. Sid-Ahmed, "Fingerprint Image Compression Standard Based on Wave Atoms Decomposition and Self Organizing Feature Map", Proceedings of Electro/Iformation Technology (2009), pp. 367-372.
- [110] http://biometrics.cse.msu.edu/fingerprint.html
- [111] D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar, "Handbook of Fingerprint Recognition", New York: Springer (2003).
- [112] N. Yager, A. Amin, "Fingerprint Classification: A Review", Pattern Analysis and Applications vol. 7, issue 1 (2004), pp. 77-93.
- [113] A. K. Jain, A. Ross, S. Pankanti, "Biometrics: A Tool for Information Security", IEEE Transactions on Information Forensics Security, vol. 1, issue 2 (2006), pp. 125-143.
- [114] H. C. Lee, R. E. Gaensslen, "Advances in Fingerprint Technology", Second Edition, New York: CRC Press (2001).
- [115] Latent Print Examination, http://onin.com/fp/ (2002).

- [116] B. Bhanu, X. Tan, "Computational Algorithms for Fingerprint Recognition", Kluwer Academic Publishers, Dordrecht (2003).
- [117] D. R. Ashbaugh, "Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology", CRC Press (1999).
- [118] S. Chandra, "Genetics of Epidermal Ridge Minutiae", International Journal of Anthropology, vol. 20, issue 1 (2005), pp. 51-62.
- [119] S. Pankanti, S. Prabhakar, A. K. Jain, "On the Individuality of Fingerprints", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, issue 8 (2002), pp. 1010-1025.
- [120] A. Jain, Y. Chen, M. Demitrius, "Pores and Ridges: Fingerprint Matching using Level 3 Features", Proceedings of International Conference on Pattern Recognition, vol. 4 (2006), pp. 477-480.
- [121] Y. Chen, "Extended Feature Set and Touchless Imaging For Fingerprint Matching, Ph.D. Thesis, Computer Science Department, Michigan State University (2009).
- [122] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, A. K. Jain, "FVC 2004: Third Fingerprint Verification Competition", Lecture Notes in Computer Science, vol. 3072 (2004), pp. 1-7.
- [123] A. K. Jain, L. Hong, R. Bolle, "On-line Fingerprint Verification, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19 (1997), pp. 302-314.
- [124] T. Y. Jea, V. Govindaraju, "A Minutiae Based Partial Fingerprint Recognition System", Elsevier Pattern Recognition, vol. 38 (2004), pp. 1672-1684.
- [125] D. Maio, D. Maltoni, "Direct Gray-Scale Minutiae Detection in Fingerprints", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19 (1997), pp. 27-40.

- [126] E. Zhu, J. P. Yin, G. M. Zhang, "Fingerprint Matching Based on Global Alignment of Multiple Reference Minutiae", Elsevier Pattern Recognition, vol. 38, issue 10 (2005), pp. 1685-1694.
- [127] B. Bhanu, X. Tan, "Fingerprint Indexing Based on Novel Features of Minutiae Triplets", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, issue 5 (2003), pp. 616-22.
- [128] L. Liu, T. Jiang, J. Yang, C. Zhu, "Fingerprint Registration by Maximization of Mutual Information", IEEE Transactions on Image Processing, vol. 15, issue 5 (2006), pp. 1100-1110.
- [129] A. K. Jain, S. Prabhakar, L. Hong, S. Pankanti, "Filterbank-Based Fingerprint Matching", IEEE Transactions on Image Processing, vol. 9, issue 5 (2000), pp. 846-859.
- [130] C. Park, J. Lee, M. Smith, S. Park, K. Park, "Directional Filter Bank-Based Fingerprint Feature Extraction and Matching, IEEE Transactions on Circuits System and Video Technology, vol. 14, issue 1 (2004), pp. 74-85.
- [131] J. Gu, J. Zhou, C. Yang, "Fingerprint Recognition by Combining global Structure and Local Cues", IEEE Transactions on Image Processing, vol. 15, issue 7 (2006), pp. 1952-1964.
- [132] A. Ross, A. K. Jain, J. Reismanb, "A Hybrid Fingerprint Matcher", Elsevier Pattern Recognition, vol. 36 (2003), pp. 16611673.
- [133] D. Wan, J. Zhou, "Fingerprint Recognition using Model-Based Density Map", IEEE Transactions on Image Processing, vol. 15, issue 6(2006), pp. 1690-1696.
- [134] N. K. Ratha, K. Karu, S. Chen, A. K. Jain, "Areal-Time Matching System for large Fingerprint Databases", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, issue 8 (1996), pp. 799-813.

- [135] A. Fitz, R. Green, "Fingerprint Classification using a Hexagonal Fast Fourier Transform", Elsevier Pattern Recognition, vol. 29, issue 10 (1996), pp. 1587-1597.
- [136] L. Seokwon, N. Boohee, "Fingerprint Recognition using Wavelet Transform and Probabilistic Neural Network", Proceedings of International Joint Conference on Neural Networks, vol. 5 (1999), pp. 3276-3279.
- [137] C. Wilson, G. Candela, P. Grother, C. Watson, R. Wilkinson, "Massively Parallel Neural Network Fingerprint Classification System", National Institute of Standards and Technology, NISTIR vol. 4880 (1992).
- [138] N. R. Parsons, J. Q. Smith, E. Thonnes, L. Wang, R. G. Wilson, "Rotationally Invariant Statistics for Examining the Evidence from the Pores in Fingerprints", Technical report, Center for Research in Statistical Methodology (CRiSM) (2007).
- [139] A. K. Jain, Y. Chen, M. Demirkus, "Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, issue 1 (2007), pp. 15-27.
- [140] J. Luo, S. Lin, M. Lei, J. Ni, "Application of Dimensionality Reduction Analysis to Fingerprint Recognition", ISCID, vol. 2 (2008), pp. 102-105.
- [141] A. Bazen, S. Gerez, "Systematic Methods for the Computation of the Direction Fields and Singular Points of Fingerprints", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, issue 7 (2002), pp. 905-919.
- [142] A. A. Mohammed, Q. M. J. Wu, M. A. Sid-Ahmed, "Application of Wave Atoms Decomposition and Extreme Learning Machine for Fingerprint Classification", Lecture Notes in Computer Science, vol. 6112 (2010), pp. 246-255.

## Vita Auctoris

Abdul Adeel Mohammed was born in Abu Dhabi, United Arab Emirates. He obtained his Bachelor of Engineering (Electronics and Communication) degree from Osmania University Hyderabad, India in 2001. He moved to Canada and started Master's of Applied Science program at Ryerson University in September 2002. He worked in the area of Image processing and Computer Vision and eventually graduated from Ryerson University in Winter 2005. During his stay at Ryerson University he won the prestigious Ontario graduate scholarship and two graduate awards in the Department of Electrical and Computer Engineering.

In 2006, he started PhD program in the Department of Electrical and Computer Engineering at University of Windsor, Canada. Adeel won the doctoral tuition scholarship and the NSERC Industrial Post Graduate Scholarship in collaboration with Vista Solutions Inc. At Vista Solutions Inc. he worked on single and multi-camera robot guidance industrial projects with specific applications to automotive industry. He is a member of IEEE and a licensed Professional Engineer in the province of Ontario.