2017

High Performance Techniques for Face Recognition

Ahmed Aldhahab
University of Central Florida

HIGH PERFORMANCE TECHNIQUES FOR FACE RECOGNITION

by

AHMED QASIM JUMAAH ALDHAHAB
B.Sc. University of Babylon, 2004
M.Sc. University of Technology, 2006

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical and Computer Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2017

Major Professor: Wasfy B. Mikhael

ABSTRACT

The identification of individuals using face recognition techniques is a challenging task. This is due to the variations resulting from facial expressions, makeup, rotations, illuminations, gestures, etc. Also, facial images contain a great deal of redundant information, which negatively affects the performance of the recognition system. The dimensionality and the redundancy of the facial features have a direct effect on the face recognition accuracy. Not all the features in the feature vector space are useful. For example, non-discriminating features in the feature vector space not only degrade the recognition accuracy but also increase the computational complexity.

In the field of computer vision, pattern recognition, and image processing, face recognition has become a popular research topic. This is due to its wide spread applications in security and control, which allow the identified individual to access secure areas, personal information, etc. The performance of any recognition system depends on three factors: 1) the storage requirements, 2) the computational complexity, and 3) the recognition rates.

Two different recognition system families are presented and developed in this dissertation. Each family consists of several face recognition systems. Each system contains three main steps, namely, preprocessing, feature extraction, and classification. Several preprocessing steps, such as cropping, facial detection, dividing the facial image into sub-images, etc. are applied to the facial images. This reduces the effect of the irrelevant information (background) and improves the system performance. In this dissertation, either a Neural Network (NN) based classifier or Euclidean distance is used for classification purposes. Five widely used databases, namely, ORL, YALE, FERET, FEI, and LFW, each containing different facial variations, such as light condition, rotations, facial expressions, facial details, etc., are used to evaluate the proposed systems. The experimental results of the proposed systems are analyzed using K-folds Cross Validation (CV).

In the family-1, Several systems are proposed for face recognition. Each system employs different integrated tools in the feature extraction step. These tools, Two Dimensional Discrete Multiwavelet Transform (2D DMWT), 2D Radon Transform (2D RT), 2D or 3D DWT, and Fast Independent Component Analysis (FastICA), are applied to the processed facial images to reduce the dimensionality and to obtain discriminating features. Each proposed system produces a unique representation, and achieves less storage requirements and better performance than the existing methods.

For further facial compression, there are three face recognition systems in the second family. Each system uses different integrated tools to obtain better facial representation. The integrated tools, Vector Quantization (VQ), Discrete cosine Transform (DCT), and 2D DWT, are applied to the facial images for further facial compression and better facial representation. In the systems using the tools VQ/2D DCT and VQ/ 2D DWT, each pose in the databases is represented by one centroid with $4 \times 4 \times 16$ dimensions. In the third system, VQ/ Facial Part Detection (FPD), each person in the databases is represented by four centroids with $4 \times Centroids$ $(4 \times 4 \times 16)$ dimensions.

The systems in the family-2 are proposed to further reduce the dimensions of the data compared to the systems in the family-1 while attaining comparable results. For example, in family-1, the integrated tools, FastICA/ 2D DMWT, applied to different combinations of sub-images in the FERET database with K-fold=5 (9 different poses used in the training mode), reduce the dimensions of the database by 97.22% and achieve 99% accuracy. In contrast, the integrated tools, VQ/ FPD, in the family-2 reduce the dimensions of the data by 99.31% and achieve 97.98% accuracy. In this example, the integrated tools, VQ/ FPD, accomplished further data compression and less accuracy compared to those reported by FastICA/ 2D DMWT tools.

Various experiments and simulations using MATLAB are applied. The experimental results of both families confirm the improvements in the storage requirements, as well as the recognition rates as compared to some recently reported methods.

To my father, mother, brothers, and sister. To my beautiful wife Rasha and my sweet kids Danaih, Durar, and Rose.

ACKNOWLEDGMENTS

First, and foremost, all praise and thanks are to Allah for his inspiration, persistent generosity and blessings. Then, I would like to convey my deep gratitude and appreciation to my supervisor, Prof. Wasfy B. Mikhael, for his peerless encouragement, guidance, and support. His help and guidance throughout my study at UCF are invaluable. His advice in academia and otherwise have undoubtedly left a mark on me. Also, I would like to express my thanks and appreciation to my Co-advisor Dr. George Atia for his support and guidance.

I would like to express my thanks and gratitude to my committee members Prof. W. Linwood Jones, Dr. Lei Wei, and Prof. Ahmad Elshennawy for providing insightful comments and stimulating thoughtful discussion.

Also, I would like to express my deepest thanks and appreciation to Diana Camerino for her help, support, and guidance.

I am also very grateful to my Lab-mates. I would like to express a special thanks to my dearest friend Ahmed Hemida for his brotherhood and for sharing my memories, my sorrows and critical situations. The help all of you offered, the thoughts we exchanged, and the laughter we shared made these Ph.D. years fun and unforgettable.

I would like to express my deepest gratitude and appreciation to my parents who have offered everything to me to reach this point in my life. I would like to express my sincere love and gratitude to my brothers and sister. Their encouragement and support are unbounded.

I am truly lucky to have my Wife Rasha and My kids Daniah, Durar, and Rose, who always sacrifice their happiness for me. My wife always does her best in order to make me happy and successful. I am really lucky that I will spend my life with such a loving and respectful person. I love you.

TABLE OF CONTENTS

LIST OF FIGURES xviii

LIST OF TABLES xxiv

LIST OF ABBREVIATIONS xxviii

CHAPTER 1: INTRODUCTION 1

1.1 Introduction 1

1.1.1 Types of Biometrics 2

1.2 Face Recognition 2

1.2.1 Face Recognition Categorization 4

1.2.2 Advantages and Challenges for Face Recognition 5

1.2.2.1 Advantages of Using Face Recognition 5

1.2.2.2 Challenges of Face Recognition 5

1.3 Databases 6

1.3.1 Olivetti Research Lab (ORL) Database 6

1.3.2 YALE Database 7

1.3.3 The Facial Recognition Technology (FERET) Database 8

1.3.4 FEI Database 9

1.3.5 LFW Database 10

1.4 Organization of this Dissertation 11

CHAPTER 2: LITERATURE REVIEW 14

2.1 Introduction 14

2.1.1 Subspace Methods 14

2.1.2 Transform Domain 17

2.1.3 Compression Methods 20

2.1.4 Artificial Neural Network and Deep Learning (ANN & DL) 21

2.1.5 Methods for Improving the Recognition Rates 22

CHAPTER 3: METHODOLOGY 24

3.1 Introduction 24

3.2 Discrete Cosine Transform (DCT) 24

3.3 Discrete Wavelet Transform (DWT) 25

3.4 Multiwavelet Transform (MWT) 31

3.4.1 Preprocessing of Multiwavelet 35

3.4.1.1 A Critically-Sampled Scheme (Approximation Based Preprocessing) . . . 37

3.5 Radon Transform . . . 38

3.6 Vector Quantization (VQ) . . . 39

3.7 Independent Component Analysis (ICA) . . . 41

3.7.1 Fast Independent Component Analysis (FastICA) . . . 44

3.7.2 ICA-Architectures . . . 47

CHAPTER 4: SUPERVISED FACIAL RECOGNITION BASED ON MULTI-RESOLUTION ANALYSIS AND FEATURE ALIGNMENT . . . 49

4.1 Proposed System . . . 49

4.1.1 Preprocessing . . . 49

4.1.2 Feature Extraction . . . 50

4.1.3 Classification . . . 53

4.2 Experimental Results . . . 53

4.2.1 Experimental Results for the ORL Database . . . 55

4.2.2 Experimental Results for the YALE Database . . . 56

4.2.3 Experimental Results for the FERET Database . . . 56

4.3 Discussion and Neural Network Performance . . . 57

4.4 Conclusion 59

CHAPTER 5: SUPERVISED FACIAL RECOGNITION BASED ON MULTIRESOLUTION ANALYSIS WITH RADON TRANSFORM 60

5.1 Proposed Approach 60

5.1.1 Preprocessing 60

5.1.2 Feature Extraction 61

5.1.3 Recognition 66

5.2 Experimental Results 66

5.2.1 Experimental Results for the ORL database 66

5.2.2 Experimental Results for the YALE database 67

5.2.3 Experimental Results for the FERET database 67

5.3 Discussion and Neural Network Performance 69

5.4 Conclusion 70

CHAPTER 6: THREE SUPERVISED FACIAL RECOGNITION TECHNIQUES BASED ON FASTICA/DMWT 71

6.1 Proposed Techniques 72

6.1.1 Preprocessing 72

6.1.2 Feature Extraction (Selection) 73

6.1.3 Recognition 78

6.2 Experimental Results 79

6.2.1 Experimental Results for the ORL Database 80

6.2.2 Experimental Results for the YALE Database 81

6.2.3 Experimental Results for the FERET Database 82

6.2.4 Experimental Results for the FEI Database 83

6.2.5 Experimental Results for the LFW Database 84

6.3 Discussion 85

6.4 Neural Network Performance 87

6.5 Conclusion 89

CHAPTER 7: SUPERVISED FACIAL RECOGNITION BASED ON EIGENANALYSIS OF MULTIRESOLUTION AND INDEPENDENT FEATURES 91

7.1 Proposed Techniques 91

7.1.1 Preprocessing 91

7.1.2 Feature Extraction 92

7.1.3 Recognition 96

7.2 Experimental Results 97

7.2.1 Experimental Results for the ORL Database 97

7.2.2 Experimental Results for the YALE Database 98

7.2.3 Experimental Results for the FERET Database 98

7.3 Discussion and Neural Network Configuration 99

7.4 Conclusion 100

CHAPTER 8: HIGH PERFORMANCE AND EFFICIENT FACIAL RECOGNITION USING NORM OF ICA/MULTIWAVELET FEATURES 101

8.1 Proposed Approach 101

8.1.1 Preprocessing 101

8.1.2 Feature Extraction 102

8.1.3 Recognition 107

8.2 Experimental Results and Discussion 108

8.2.1 Experimental Results for the ORL Database 108

8.2.2 Experimental Results for the YALE Database 109

8.2.3 Experimental Results for the FERET Database 110

8.3 Discussion 110

8.4 Conclusion 111

CHAPTER 9: A FACIAL RECOGNITION METHOD BASED ON DMW TRANSFORMED PARTITIONED IMAGES 112

9.1 Proposed Techniques 112

9.1.1 Preprocessing 112

9.1.2 Feature Extraction 114

9.1.3 Recognition 117

9.2 Experimental Results 118

9.2.1 Experimental Results for the ORL Database 118

9.2.2 Experimental Results for the YALE Database 119

9.2.3 Experimental Results for the FERET Database 120

9.2.4 Experimental Results for the FEI Database 120

9.3 Discussion 121

9.4 Neural Network Performance 122

9.5 Conclusion 124

CHAPTER 10: FACE RECOGNITION EMPLOYING DMWT FOLLOWED BY FASTICA 125

10.1 Proposed System 126

10.1.1 Preprocessing 127

10.1.2 Feature Extraction 128

10.1.3 Recognition 134

10.1.4 Decision 135

10.2 Experimental Results 136

10.2.1 Experimental Results for the ORL Database 136

10.2.2 Experimental Results for the YALE Database 137

10.2.3 Experimental Results for the FERET Database 138

10.2.4 Experimental Results for the FET Database 139

10.2.5 Experimental Results for the LFW Database 140

10.3 Discussion 141

10.4 Performance of Neural Network 144

10.5 Conclusion 146

CHAPTER 11: EMPLOYING VECTOR QUANTIZATION ALGORITHM IN A TRANSFORM DOMAIN FOR FACIAL RECOGNITION 148

11.1 Proposed System 148

11.1.1 Preprocessing 149

11.1.2 Feature Extraction 150

11.1.3 Classification 153

11.2 Experimental Results 154

11.2.1 Experimental Results for the ORL Database 154

11.2.2 Experimental Results for the YALE Database 155

11.2.3 Experimental Results for the FERET Database 155

11.2.4 Experimental Results for the FEI Database 155

11.3 Discussion 156

11.4 Conclusion 157

CHAPTER 12: EFFICIENT FACIAL RECOGNITION USING VECTOR QUANTIZATION OF 2D DWT FEATURES 158

12.1 Proposed System 158

12.1.1 Preprocessing 159

12.1.2 Feature Extraction 160

12.1.3 Classification 163

12.2 Experimental Results 164

12.2.1 Experimental Results for the ORL Database 164

12.2.2 Experimental Results for the YALE Database 165

12.2.3 Experimental Results for the FERET Database 165

12.2.4 Experimental Results for the FEI Database 166

12.3 Discussion 166

12.4 Conclusion 167

CHAPTER 13: EMPLOYING VECTOR QUANTIZATION ON DETECTED FACIAL PARTS FOR FACE RECOGNITION 168

13.1 Proposed System 168

13.1.1 Preprocessing 169

13.1.2 Feature Extraction 171

13.1.3 Classification 174

13.2 Experimental Results 175

13.2.1 Experimental Results for the ORL Database 175

13.2.2 Experimental Results for the YALE Database 175

13.2.3 Experimental Results for the FERET Database 176

13.2.4 Experimental Results for the FEI Database 177

13.3 Discussion 178

13.4 Conclusion 178

CHAPTER 14: CONCLUSION AND FUTURE WORKS 179

14.1 Future Works 182

LIST OF REFERENCES 183

LIST OF FIGURES

1.1 Figure 1.1-a shows the scenario of using biometric MRTD system for Passport. Figure 1.1-b presents the various biometrics features based on MRTD compatibility. 3

1.2 Sample images of ORL Database 7

1.3 Sample images of YALE Database 7

1.4 Sample images of FERET Database 8

1.5 Sample images of FEI Database 9

1.6 Sample images of LFW Database 10

3.1 DWT structures. Figure 3.1-a shows the input image.Figure 3.1-b demonstrates the first level of decomposition, which divids the signal into four frequency subbands. Figure 3.1-c shows the second level of decompositions of DWT, where L and H correspond to the low pass and high pass filters, respectively. 30

3.2 Shows the 1-level of DWT and DMWT decomposition. Figure 3.2-a is the 1-level of 2D DWT decomposition. Figure 3.2-b is the 1-level of 2D DMWT decomposition. 31

3.3 Multi-scaling and Multi-wavelet functions for GHM filter 32

3.4 Shows different preprocessing algorithms. 36

3.5 Radon Transform 38

3.6 Block diagram for Independent Component Analysis (ICA). 42

3.7 ICA model for Architecture I. 47

3.8 Statistically independent image basis employed with the coefficients a_i to reconstruct the input facial images. 48

4.1 Application of 2D DMWT on different databases. 52

4.2 Application of 2D RT on the resultant of 2D DMWT. 54

4.3 MSE performance of the NNT for our proposed system compared with other methods. 58

5.1 Application of 2D DMWT on different databases. 63

5.2 An example of 3D arrangement applied to figure 5.1-c. 64

5.3 Applying the 3D DWT to the result of figure (5.2). 65

5.4 MSE performance of the NNT for the YALE database during the training phase. For both approaches, we used one hidden layer with 512 neurons. The activation function used is the hyperbolic tangent sigmoid, Back Propagation is used for training, and the goal is 10^{-7}. 69

6.1 The Three Proposed Techniques 72

6.2 Figure 6.2-A shows 5 original samples of one person for all different databases. Figure 6.2-B shows samples after cropping. 74

6.3 An example of applying 2-D DMWT to the different databases. 75

6.4 Applying 2D FastICA into extracted subimages. S_x represents Subimages of LL sub-band. IS_x represents Independent Subimages of LL sub-band where $x = 1, 2, 3, and\ 4$. 77

6.5 Hyperbolic tangent sigmoid transfer function. The activation function tansig(x) is expressed as $tansig(x) = \frac{2}{1+e^{-2x}} - 1$ and it can be considered as $tansh(x)$. Its output range between (-1,1). 78

6.6 The NN performance for the three proposed techniques compared with the other approaches. 88

7.1 The Proposed Techniques. 92

7.2 Application of 2D DMWT on different databases. 94

7.3 Applying 2D FastICA into extracted subimages. Sx represents Subimages of LL subband. ISx represents Independent Subimages of LL subband where $x = 1, 2, 3, and\ 4$. 95

8.1 Application of 2D DMWT on different databases. 104

8.2 Figure 8.2-A represents the subimages of the LL subband of 2D DMWT for P training poses from each person. Figure 8.2-B is rearranging Figure 8.2-A in four combinations. 105

8.3 An example of applying 2D FastICA to the first combination. IF_p represents the resultant Independent features where $p \in \{1, 2, 3, \ldots, P\}$. 105

8.4 Proposed Technique 1. 106

8.5 Proposed Technique 2. 106

8.6 Proposed Technique 3. 106

8.7 Proposed Technique 4. 107

9.1 The Proposed Techniques. 113

9.2 Dividing the image of ORL database into 6 parts. 113

9.3 Applying 2-D DMWT to all partitions of one sample of each different databases. 115

9.4 Shows the extracted features for one pose after converted into one dimensional form. 116

9.5 Shows how the parts are participating with each other 122

9.6 The performance of Neural Network for ORL, FERET, and FEI databases. . 123

10.1 The Proposed System 126

10.2 Dividing the first pose of person one of ORL database into 6 parts. 127

10.3 An example of applying 2-D DMWT to ORL and FEI databases. 129

10.4 Two representations. $S_1^* \ldots S_6^*$ are denoted as sub-images with maximum ℓ_2-Norm. 130

10.5 Applying FastICA to retained LL sub-band. S_l represents Sub-images of LL sub-band. IF_l represents Independent features. Where $l = 1, 2, 3, and\, 4$. 132

10.6 Hyperbolic tangent sigmoid transfer function. The activation function is expressed as $tansig(x) = \frac{2}{1+e^{-2x}} - 1$ and considered as $tansh(x)$. Its output range is between (-1,1). 135

10.7 Shows input image division and partitioning construction. 142

10.8 NN performance for the proposed techniques compared with the other approaches based first representation. 145

10.9 NN performance for the proposed techniques compared with the other approaches based second representation. 146

11.1 Proposed system. Figure 11.1-a shows the Preprocessing and the Feature extraction phases. Figure 11.1-b illustrates the classification phase. 149

11.2 Figure 11.2-A shows 5 original samples of one person for all different databases. Figure 11.2-B shows samples after cropping. 151

11.3 Shows the original pose and its DCT version 152

11.4 Calculating the first mean of the retained matrix of each pose. 153

12.1 Proposed system. Figure 12.1-a shows the Preprocessing and the Feature extraction phases. Figure 12.1-b illustrates the classification phase. 159

12.2 Application of 2D DMWT on different databases. 162

13.1 Proposed system. Figure 13.1-a shows the Preprocessing and the Feature extraction phases. Figure 13.1-b illustrates the classification phase. 169

13.2 Figure 13.2-(a & e) show the original poses. Figure 13.2-(b & f) illustrate the detected faces and facial parts. Figure 13.2-(c & g) represent the detected faces of the two poses. Detected facial parts of the two poses are displayed in figure 13.2-(d & h). 170

13.3 Constructing four groups using six training poses of the first person of YALE database. 172

LIST OF TABLES

4.1 The Dimensions of the Databases 50

4.2 The Recognition Rate of the ORL Database 55

4.3 The Recognition Rate of the YALE Database 56

4.4 The Recognition Rate of the FERET Database 57

5.1 The Dimensions of the Databases 61

5.2 The Recognition Rate of the ORL Database 67

5.3 The Recognition Rate of the YALE Database 68

5.4 The Recognition Rate of the FERET Database 68

6.1 The Dimensions of The Databases 73

6.2 Experimental Results for the ORL Database 80

6.3 Experimental Results for the YALE Database 81

6.4 Experimental Results for the FERET Database 82

6.5 Experimental Results for the FEI Database 83

6.6 Experimental Results for the LFW Database 84

7.1 The Dimensions of the Databases 92

7.2 Experimental Results for the ORL Database 97

7.3 Experimental Results for the YALE Database 98

7.4 Experimental results for the FERET Database 99

8.1 The Dimensions of the Databases 102

8.2 Experimental Results for the ORL Database 109

8.3 Experimental Results for the YALE Database 109

8.4 Experimental results for the FERET Database 110

9.1 The Dimensions of the Databases 114

9.2 Experimental Results for the ORL Database 119

9.3 Experimental Results for the YALE Database 119

9.4 Experimental Results for the FERET Database 120

9.5 Experimental Results for the FEI Database 121

10.1 The Dimensions of The Databases 128

10.2 Experimental Results for the ORL Database 137

10.3 Experimental Results for the YALE Database 138

10.4 Experimental Results for the FERET Database 139

10.5 Experimental Results for the FEI Database 140

10.6 Experimental Results for the LFW Database 141

11.1 The Dimensions of the Databases 149

11.2 Experimental Results for the ORL Database 154

11.3 Experimental Results for the YALE Database 155

11.4 Experimental Results for the FERET Database 156

11.5 Experimental Results for the FEI Database 156

12.1 The Dimensions of the Databases 160

12.2 Experimental Results for the ORL Database 164

12.3 Experimental Results for the YALE Database 165

12.4 Experimental Results for the FERET Database 165

12.5 Experimental Results for the FEI Database 166

13.1 The Dimensions of the Detected Faces/Facial Parts 171

13.2 Proposed Dimensions of the Detected Faces/Facial Parts 171

13.3 Experimental Results for the ORL Database 176

13.4 Experimental Results for the YALE Database 176

13.5 Experimental Results for the FERET Database 177

13.6 Experimental Results for the FEI Database 177

LIST OF ABBREVIATIONS

1D . . . . . One-Dimensional

2D DMWT . . . . . Two Dimensional Discrete Multiwavelet Transform

2D DWT . . . . . Two Dimensional Discrete Wavelet Transform

2D FastICA. . . . . Two Dimensional Fast Independent Component Analysis

2D RT . . . . . Two Dimensional Radon Transform

3D DWT . . . . . Three Dimensional Discrete Wavelet Transform

ANN . . . . . Artificial Neural Network

BMP . . . . . Bitmap

BPNN . . . . . Back Propagation Neural Network

BSS . . . . . Blind Source Separation

CVA . . . . . Common Vector Approach

CWT . . . . . Continuous Wavelet Transform

DARPA . . . . . Defense Advanced Research Products Agency

DCT . . . . . Discrete Cosine Transform

DFT . . . . . Discrete Fourier Transform

DL . . . . . Deep Learning

DNA . . . . . Deoxyribonucleic Acid

DoD . . . . . Department of Defense

FERET . . . . . The Facial Recognition Technology

FI . . . . . Feature Information

FPD . . . . . Facial Parts Detection

FSVM . . . . . Fuzzy Supported Vector Machine

GHM . . . . . Geronimo, Hardian, and Massopust

GIF . . . . . Graphics Interchange Format

GWT . . . . . Gabor Wavelet Transform

HBM . . . . . histogram-based method

ICA . . . . . Independent Component Analysis

ISVM . . . . . Improved Supported Vector Machine

JPEG . . . . . Joint Photographic Experts Group

KFCG . . . . . Kekre Fast Codebook Generation

KMCG . . . . . Kekres Median Codebook Generation Algorithm

KPE . . . . . Kekres Proportionate Error Algorithm

LBG . . . . . Linde, Buzo, and Gray

LBP . . . . . Local Binary Patterns

LBPP . . . . . Local Binary Probabilistic Pattern

LDA . . . . . Linear Discriminant Analysis

LFW . . . . . Labeled Faces in the Wild

LL, LH, HL, and HH. . . . . Low-Low, Low-High, High-Low, and High-High

LPP . . . . . Locality Preserving Projection

MLP . . . . . Muti-Layer Perceptron

MPMP . . . . . Multi-partitioning Max Pooling

MRA . . . . . Multi-Resolution Analysis

MRTD . . . . . Machine Readable Travel Document

MSE . . . . . Mean Squared Error

MSF . . . . . Markov Stationary Features

NIST . . . . . National Institute of Standards and Technology

NMF . . . . . Non-negative Matrix Factorization

ORL . . . . . Olivetti Research Lab

PCA . . . . . Principal Component Analysis

PDF . . . . . Probability Density Function

PIN . . . . . An Identification Number of an individual

PSO . . . . . Particle Swam Optimization

RBF . . . . . Radial Basis Function

ROK-VISIT. . . . . Republic of Korea Visitor and Immigrant Status Indicator Technology

SI . . . . . Structural Information

SID . . . . . Stacked Image Descriptor

SLBFLE. . . . . Simultaneous Local Binary Feature Learning and Encoding

SLF . . . . . Statistical Local Features

SVM . . . . . Supported Vector Machine

TIFF . . . . . Tag Image File Format

UIDAI . . . . . The Unique Identification Authority of India

US-VISIT . . . . . United States Visitor and Immigrant Status Indicator Technology

VQ . . . . . Vector Quantization

WFT . . . . . Window Fourier Transform

WT . . . . . Wavelet Transform

CHAPTER 1: INTRODUCTION

1.1 Introduction

Biometrics refers to metrics as they relate to human characteristics. The term biometrics is derived from two ancient Greek words: *bio* means life and *metrics* mean measure. Biometrics can be classified as either identification (i.e. recognition) or verification (i.e. authentication). Biometric identifiers are the distinctive, measurable characteristics used to label and describe individuals. The identification system seeks to answer the question "Who are you?" Biometric verification is a way to verify the individual, since the individual is uniquely distinguishable through one or more biological traits. Verification systems seek to answer the question "Is this person who they say they are?" Distinctive verifiers and identifiers include fingerprints, hand geometry, earlobe geometry, retina and iris patterns, voice waves, DNA, signatures, palm print, face, etc. The biometrics field is multidisciplinary and includes sensor design, computer vision, image processing, pattern recognition, machine learning, signal processing, and information fusion. Biometrics are unique to each individual and are more reliable in verifying an identity than passwords, PIN, smart cards, tokens, keys, plastic cards, etc [1].

Biometric-Based Techniques have shown to be the most committed option to recognize the individual in recent years. Instead of authenticating people through physical and virtual domains, which are based on passwords, PIN, smart cards, tokens, keys, plastic cards, and so forth, biometrics recognition seeks to recognize the persons themselves. Since all physical and virtual domain techniques are subject to be forgotten, stolen, guessed, misplaced, etc., biometrics recognition is a promising way to recognize the individual [2, 3].

1.1.1 Types of Biometrics

Biometrics is the measurement and statistical analysis of people's physical and behavior characteristics. It refers to a study in Biological Science that can measure the biological statistics. In most studies, biometrics relates to authentication, which is a process to determine whether something/someone is what or who it is declared to be. Biometrics can be categorized as [4]:

1. Chemical biometrics: which involves measuring chemical cues, such as DNA and Odor.
2. Visual biometrics (Physical): which involves measurements and characterization of face, for example Face Recognition, Fingerprint Recognition, Finger Geometry Recognition, Eyes - Iris Recognition, Eyes - Retina Recognition, Hand Geometry Recognition, Signature Recognition, Typing Recognition, etc.
3. Behavioral biometrics: which measures the performance of people during their daily activities, such as Speech, Signature Recognition, Gait Analysis, Keystrokes Dynamics, etc.

1.2 Face Recognition

Identification by face recognition is one of the biometric techniques that receives a considerable attention in recent years. In the field of Computer Vision, Pattern Recognition, and Image Processing, face recognition has become a popular research topic due to its wide potential applications in access control, personal identification, security in which several security biometric systems are contacted based on biometric face recognition, etc [5, 6, 7, 8]. The efficiency and the quality of any recognition system depend on several factors: 1) the storage requirements, 2) the computational complexity, and 3) the recognition rates. Efficient and compact features are considered as an essential part of every reliable recognition system.

(a) (b)

Figure 1.1: Figure 1.1-a shows the scenario of using biometric MRTD system for Passport. Figure 1.1-b presents the various biometrics features based on MRTD compatibility.

Face recognition is a routine activity that humans and most of animals do in their daily lives. Due to low cost and improvement in integrated circuit and computer technology, the interest in automated image recognition and video analysis/ recognition that mimics human's skills is increased. This is due to the fact that these techniques (face recognition and video analysis/ recognition) are inherent in several applications, such as biometric authentication, security, control, surveillance, human-computer interaction, and multimedia management. Based on the six biometric traits studies by Hjelms and Kee in 2001 [9], facial features have the highest capability in Machine Readable Travel Document (MRTD) system [10] due to several evaluation factors, such as enrollment, renewal, machine requirements, and public perception as shown in figure 1.1 [11].

In 1973, the first automated face recognition system was presented by Kanade [12] in his Ph.D. dissertation. Then, the automated face recognition experienced a latent period until the work presented by Sirovich and Kirby in 1987 and 1990 [13, 14] on a low dimensional face representation derived using Karhunen-Loeve transform or Principal Component Analysis (PCA). Thereafter, Turk and Pentland in 1991 [15, 16] expanded the results presented in [13, 14] and applied the

eigenfaces method as a face recognition approach. Several commercial face recognition systems have been implemented such as FaceIt, MorphoTrust Technology [17], Cognitec [18], Facekey [19], LATHEM [20], FingerTec [21], etc. Also, the face recognition system is used as a security application in several countries, such as United States Visitor and Immigrant Status Indicator Technology (US-VISIT), The Unique Identification Authority of India (UIDAI), Vision-Box (Singapore), ROK-VISIT (Republic of Korea Visitor and Immigrant Status Indicator Technology), Germany introduced use of automatic face recognition to tackle rising terrorism threats, the Malaysian government is installing 300 facial recognition devices to scan travelers entering and exiting the country at key entry points by the end of 2017, Exit & Entry Permit (Republic of China), etc.

The stored face images in the databases are called the *training images*.The procedure using certain algorithms to prepare for the training images is called the *training phase procedure*. The unknown face image required to be identified is called the *testing face image*. The procedure used to prepare for the test image is called the *testing phase procedure*.

1.2.1 Face Recognition Categorization

Face recognition systems are expected to identify faces present in images or video automatically. Face recognition system can be classified into two modes: Face Identification and Face Verification.

Face Identification system (or recognition) is a one-to-n matching system, where n is the total number of biometrics stored. It compares the queried face image to all facial images in the databases to identify the claimed face image. This system answers the question “Who is this person?”

Face verification system (or Authentication) is a one-to-one matching system that compares the queried face image to a sample face image whose identity is being claimed. This system is trying

to answer the question “Am I who I say I am?” [11, 4].

1.2.2 Advantages and Challenges for Face Recognition

1.2.2.1 Advantages of Using Face Recognition

There are numerous advantages of using face recognition technology for identification instead of using identification based on finger print, geometry, hand print, iris, retina, ear, voice, etc. All of these technologies require voluntary actions by users, for example a person being required to put his/her hand on a hand rest for hand or finger prints detection. Also, he/she has to stand in a fixed position in front of the camera for iris or retina identification. However, the face identification is done passively without any necessary action on the part of the user since the face image can be taken from a distance by using a camera. Also, some of these techniques require very expensive equipment and are very sensitive to body motions. Even voice recognition is subject to noise contamination in public places or noise in phone lines or tape recordings. However, facial images can be obtained using inexpensive fixed cameras. Efficient facial recognition systems are not sensitive to the noise, variation in orientations, scales, and illuminations [2, 3].

1.2.2.2 Challenges of Face Recognition

Despite these advantages, face recognition is still a challenging task [22, 23, 24]. This is due to several problems. The first problem is illumination variations, which change the semblance of the subject significantly and hence result in dissimilarity between the testing and training poses. Pose change is also a major problem in face recognition. The available images (poses), for example, in the training phase were taken from frontal angle (view) of the face while the testing poses might be taken form another view that leads to fault-identification. Another problem is facial

expressions. Different facial expressions for one person might be similar to another person. Taking pictures of people of different ages is also one of the related problems and leads to incorrect identification, since the appearance of the face is changed. There are other problems, such as occlusion, unavailability of the whole face, and low resolution of pictures, which result in less information about the persons and thus reduced recognition rates [22, 25, 26, 27].

1.3 Databases

The databases used in this dissertation are described below:

1.3.1 Olivetti Research Lab (ORL) Database

The ORL database, shown in figure 1.2, was collected between April 1992 to April 1994. The database was used in the context of a face recognition project carried out in collaboration with the Speech, Vision and Robotics Group of the Cambridge University Engineering Department (AT&T Laboratories Cambridge). The ORL database consists of 10 different poses of 40 different persons with the resolution of 112×92 pixels in BMP format. The total number of poses in this database is 400 poses. The poses of all subjects have different facial variations, such as facial expressions (open / close eyes, smiling / not smiling) and facial details (glasses/ no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position [28].

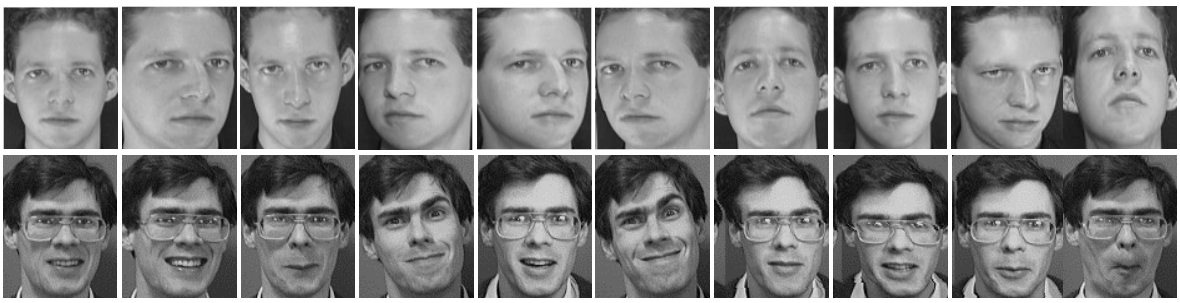


Figure 1.2: Sample images of ORL Database

1.3.2 YALE Database

There are 15 persons in YALE database, each with 11 different poses with resolution of 320×243 pixels in GIF format. Therefore, the total number of poses in the YALE database is 165 poses. All of them have lighting variations including left light, center light, and right light. Images of faces in YALE database have different configurations including glasses/ without glasses. Also, faces have different facial variations including happy, normal, sad, sleepy, surprised, and wink [29]. Samples of database are shown in figure 1.3.

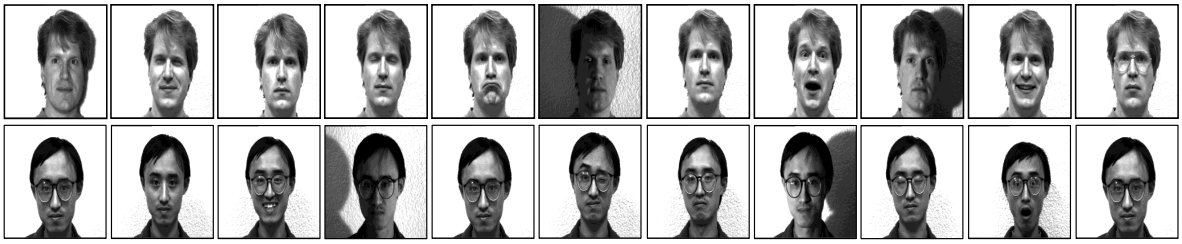


Figure 1.3: Sample images of YALE Database

1.3.3 The Facial Recognition Technology (FERET) Database

The FERET Database is a standard database that was proposed to help build an automatic face recognition system that could be applied to assist security, intelligent, and law enforcement personnel in the performance of their duties. The FERET program is sponsored by the US Department of Defense (DoD) Counter-drug Technology Development Program. The FERET program is managed through the Defense Advanced Research Products Agency (DARPA) and the National Institute of Standards and Technology (NIST). The target of FERET program's sponsored was to develop face recognition algorithms. The Face Recognition Technology (FERET) Database was collected during the period between 1993-1997 [30, 31, 11]. Samples from the database are shown in figure 1.4.

The subset *fc* of the FERET database, which is used in this dissertation, contains 200 individual, each with 11 different poses with resolution 256×384 pixels in TIFF format. Therefore, the total number of poses in the subset *fc* of the FERET database is 2200. All poses in the database have different facial variations, such as light conditions, rotations, background, facial expressions, and glasses/no glasses [30, 31].

Figure 1.4: Sample images of FERET Database

1.3.4 FEI Database

The FEI database is a Brazilian face database, which contains face images collected from June 2005 to March 2006, recorded at the Artificial Intelligence Laboratory of FEI in São Bernardo do Campo, São Paulo, Brazil. The FEI consists of 200 individuals, each with 14 different poses, for a total of 2800 images. Each pose has a resolution of 640×480 pixels in the JPG format. All of the facial images are in color and taken against a white homogeneous background in an upright frontal position with rotation up to about 180 degrees. Face images in the FEI database are mainly represented by student and staff between 19 and 40 years old with distinct appearance, hairstyle, and adorns. The number of males and females in this database is equal to 100 [32]. The gray scale version of the database is used in this dissertation. Samples from the database are shown in figure 1.5.

Figure 1.5: Sample images of FEI Database

1.3.5 LFW Database

Labeled Faces in the Wild (LFW) is a large database of photographs designed to study the problem of unconstrained face recognition, which contains 13233 images of 5749 the persons. Each pose is labeled with the name of person. Each person is given a unique name, such as "Angelina_Jolie, Leonardo_DiCaprio, etc". These facial images were collected from the web; therefore, the face images in this database have different facial variations in pose, lighting, focus, resolution, facial expression, age, gender, race, accessories, make-up, occlusions, background, and photographic quality. Most of the face images are in color, and there are some other faces in gray-scale. The images are available in the size of 250×250 pixels in the .JPG format [33]. The gray-scale version of this database and the cropped version of this database are used in this dissertation. Samples of database are shown in figure 1.6.

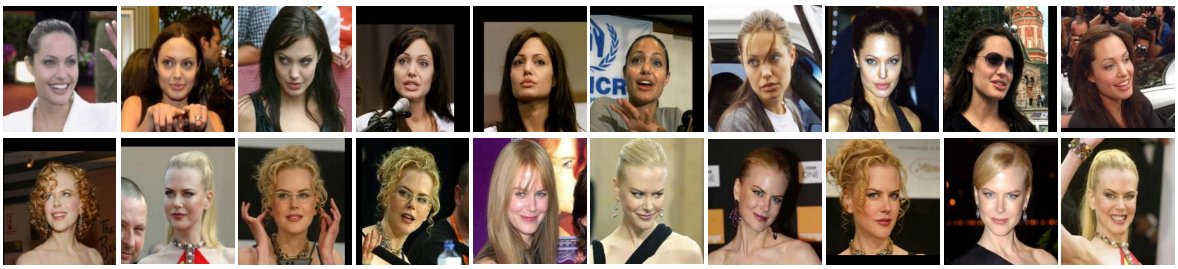


Figure 1.6: Sample images of LFW Database

All the experimental results that are presented in this dissertation are based on the face identification scenario.

1.4 Organization of this Dissertation

The dissertation is organized as follow:

1. **Chapter Two**: In this chapter, a literature review about the algorithms and techniques that have been used in face recognition systems is presented.

2. **Chapter Three**: In this chapter, different algorithms that were used in our research and applied to the facial images to achieve the best facial image representations are discussed.

3. **Chapter Four**: A new supervised algorithm fo face recognition based on the integration of Two-Dimensional Discrete Multiwavelet Transform (2D DMWT), 2D Radon Transform (2D RT), and 2D Discrete Wavelet Transform (2D DWT) is proposed in this chapter [34]. It is shown that this approach can significantly improve the classification performance and the storage requirements of the overall recognition system.

4. **Chapter Five**: In this chapter, a supervised facial recognition approach based on the integration of 2D DMWT, 2D RT, and 3D DWT is discussed [35]. The system proposed shows a significant improvement in the recognition rates and achieves less storage requirements.

5. **Chapter Six**: Three supervised techniques are proposed in this chapter for facial recognition [36, 37]. In this chapter, The first two techniques apply FastICA/ 2D DMWT with ℓ_2-norm to the facial images, while *technique 3* employs only 2D DMWT with ℓ_2-norm to the facial images. The proposed *techniques 1* $\&$ *2* accomplish higher recognition rates compared to those rates achieved by technique 3 and the other methods.

6. **Chapter Seven**: A supervised facial recognition system using 2D DMWT, Two Dimensional Fast Independent Component Analysis (2D FastICA) is presented in this chapter [38]. In this chapter, the eigenvalues and eigenvectors decomposition technique was used to represent the

facial images. The proposed approach shows a significant improvement in the recognition rate, storage requirements, as well as computational complexity.

7. **Chapter Eight**: In this Chapter, 2D DMWT in conjunction with FastICA are proposed for facial recognition [39]. For feature extraction, the integrated tools 2D FastICA/ 2D DMWT are applied to different combinations of poses corresponding to the sub-images of the low-low frequency sub-band of the MWT, and the ℓ_2-norm of the resulting features are computed to obtain discriminating and independent features, while achieving significant dimensionality reduction. The proposed approach is shown to yield significant improvement in storage requirements, computational complexity, as well as recognition rates over existing approaches.

8. **Chapter Nine**: A new approach based on applying 2D DMWT to processed/partitioned facial images is proposed in this chapter for facial recognition [40]. In this chapter, the facial images are divided into six parts. Each pose is represented by six feature vectors. The proposed system achieves comparable results to those rates accomplished by existing approaches.

9. **Chapter Ten**: The same tools, used in chapter 6, are applied to the processed/ partitioned facial images in this chapter for face recognition [41]. Each pose in this chapter has two representations. Each pose in the first representation has six feature vectors, while it has three feature vectors in the second representations. The results achieved in this chapter are the highest rates accomplished compared to the other results reported in other chapters and the state-of-the-art approaches.

10. **Chapter Eleven**: A Vector Quantization (VQ) algorithm in the Discrete Cosine Transform (DCT) domain is proposed for facial recognition in this chapter [42]. VQ algorithm employing Kekre Fast Codebook Generation (KFCG) approach for codebook initialization is

applied to the transformed truncated features. The proposed approach is shown to improve the recognition rates as well as the storage requirements in comparison with some of the existing state-of-the art approaches.

11. **Chapter Twelve**: A new approach for facial recognition employing 2D DWT and VQ is proposed in this chapter [43]. The results show that the proposed approach improves the recognition rates and reduces the storage requirements compared with existing methods.

12. **Chapter Thirteen**: The Facial Parts Detection (FPD) approach in conjunction with the Vector Quantization (VQ) algorithm are proposed in this chapter for face recognition [44]. Detecting facial parts, which are nose, eyes, and mouth, and choosing appropriate dimensions for each part are done in this chapter. Experimental work is performed to evaluate the performance of the proposed technique and the state-of-the-arts approaches. The proposed system consistently improves the recognition rates as well as the storage requirements.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

There are a large number of algorithms that have been proposed for facial recognition. Designing an efficient facial recognition system is a challenging task due to the large dimensions of the data. Data compression/feature extraction is a necessary part of all algorithms.

2.1.1 Subspace Methods

The most broadly used method for face recognition is Principal Component Analysis (PCA). PCA, subspace method, is one of the signal representation approaches that can extract facial basis (eignfaces) from a Covariance matrix constructed from a set of training faces [45]. The idea of eignfaces is first presented in [13, 14] for low-dimensional facial characteristics and representations. Later, Truk and Pentland [15, 16] in 1991 expanded the results presented in [13, 14] and applied the eigenfaces method for the face recognition approach. The eigenfaces, which define a feature vector space that dramatically reduces the dimensions of the original image faces, correspond to the eigenvectors associated with the dominant eigenvalues of the known face Covariance matrix. Eigenfaces using the Karhunen-Love Transform have two limitations. First, increasing the data size leads to reduction in the feature discrimination (eigenfaces) . Second, both the computational complexity of finding the eigenvectors and the storage requirements are increased with large training databases [46]. PCA is widely used in the area of face recognition. The performance of PCA will suffer significant degradation when the training face images used have different facial configurations in pose, illuminations, expressions, occlusions or other localized disorder (hair style and facial hair). Also, PCA fails to give meaningful representations when images have significant

variations in geometry [47]. Research has been done to make PCA more robust to facial variations [48, 49, 50, 51, 52]. Since PCA is very sensitive to the outliers, one way to overcome some of these problems is using ℓ_1 eigenfaces. Fast ℓ_1 PCA reported in [53] integrated with the greedy search algorithm were introduced in [54] to overcome these problems and to improve recognition accuracy.

Another subspace approach called Linear Discriminant Analysis (LDA), fisherfaces, is a supervised method and uses the data class label information. In contrast to PCA that projects the data along the direction of maximum variance, LDA seeks the directions for maximum discriminations of classes. LDA constructs a linear subspace in which data is optimally discriminated by maximizing inter class scatter and minimizes intra class scatter in the projective linear subspace [48, 55]. LDA based techniques are widely used in several computer vision and pattern recognition applications, such as face recognition [56], facial expression classification [57], LDA based electromyography Pattern Recognition for anthropomorphic robotic hand [58], handwritten recognition [59], etc. LDA in the area of the face recognition outperformed the PCA approach [48, 60].

Locality Preserving Projection (LPP) is an alternative approach for PCA, which seeks to find the optimal linear approximation to the eigenfunctions of the Laplace Beltrami operator on the face manifold. LPP, a linear subspace method, tries to optimally preserve the local neighborhood information. LPP is used for dimensionality reduction and feature selection, hence it is employed for facial representation and classification [61, 62]. Local features were extracted using the LPP approach in [63] for facial recognition in which the effect of facial variations in pose, facial expressions, and illuminations is reduced.

Independent Component Analysis (ICA), an extended version of PCA, is considered as another subspace approach. In contrast to PCA that only considers the second-order statistics, it is well known that ICA employs the high-order statistics in the data [64, 65, 66, 67]. ICA is successfully

used for feature extraction in the facial representation and recognition aspect [68, 64]. Here, the basis images extracted from the training database are as independent as possible compared with the basis images extracted using PCA that are uncorrelated. Both ICA and Supported Vector Machine (SVM) were introduced by Zhang in [58] for face recognition. Gabor-FastICA in conjunction with Linear Kernal SVM were proposed for thermal face classification [69]. Both PCA and ICA were presented in [70] for face recognition. The integrated tools achieved a recognition rate of 97.33% based on YALE database.

Facial recognition based on 2D ICA and Fuzzy Supported Vector Machine (FSVM) is proposed in [71]. There, the extracted features are obtained by applying 2D ICA to the low-frequency band of different levels of 2D discrete wavelet transform (2D DWT) based on Daubechies 4. Then, the extracted features are fed to FSVM for classification. The algorithm was tested on the ORL and YALE databases. In [72], 2D DWT, 2D ICA, and Radial Basis Function (RBF) were used for facial recognition. The 2D DWT based on Daubechies 4 is used to extract useful information from the face images and then 2D ICA is used to get discriminating and independent features. Then the extracted features are fed to a RBF for the recognition task. The algorithm was tested on the ORL database and the recognition rate achieved was 93%. Also, DWT using various wavelet filters, ICA, and Euclidean distance were employed in [65] for face recognition.

Another subspace method is Non-negative Matrix Factorization (NMF) [73, 74, 75], which tries to find basis images from original data without using class information. In the face space, NMF can capture facial parts and Euclidean structure of the facial images. Also, it can decompose facial images into their constituent parts, such as eyes, nose, mouth, etc. In contrast to PCA, which preserves the global facial structure, and LDA, which tries to search discrimination by maximizing between classes and minimizing within classes, NMF is a part-based algorithm that can capture the local facial features [73, 76, 77]. NMF has been recently shown to be one of the most powerful decomposition techniques for multivariate data, which is inherent to several applications in signal

processing, computer vision, pattern recognition, image processing, and biomedical engineering. NMF requires that the projective vectors and associated weighted coefficients are non-negative. Both sparse facial representations and spatial facial information were considered in [78] for face recognition by using NMF based on spatial pyramid matching. In [79], a hybrid technique based on NMF and Gabor Wavelet Transform (GWT) was applied to the facial images to extract the local facial features.

2.1.2 Transform Domain

Discrete Wavelet Transform (DWT), based on Multi-Resolution Analysis (MRA), is broadly and successfully used in different fields. DWT is used for dimensionality reduction (compression), feature extraction, and noise mitigation. Hence, it is widely used in the face recognition area. In [80], DWT was proposed for facial classification in which both data compression/feature extraction were achieved. The authors in [80] tried different wavelet filters to obtain different results and the highest recognition rate accomplished was 89.42% for ORL database using the Haar filter. Based on favorable properties of wavelet filters, such as orthogonality, symmetry, etc., DWT integrated with the other approaches, such as PCA, ICA, etc. achieves superior performance. In [81], 2D PCA was applied to the approximation coefficients of the third level of DWT decompositions for feature selection/data compaction. The authors in [81] demonstrated that using the integrated tools PCA/DWT led better results than those reported using PCA or DWT alone. The authors in [82] combined ear features with frontal face features to improve the recognition accuracy, where both features were extracted using PCA/DWT. ICA and 2D DWT were proposed in [65, 83] for facial representations and recognition. Several levels of DWT decompositions were applied to the facial images for dimensionality reduction by retaining only the Low-Low (LL) frequency sub-band, and then ICA applied to obtain different discriminating extracted features. Their system achieved results higher than those obtaining using traditional DWT. PCA and Daubechies wavelet subbands

were presented in [84]. First, PCA was applied to the four levels of Daubechies wavelet subbands to select the features. Then, the City Block distance and Euclidean distance measures were used for accuracy measurements. The system was tested on the ORL database and a recognition rate of 96.87% was reported.

An extended version of DWT is a Discrete Multiwavelet Transform (DMWT). MWT, based on MRA, is used for signal representation and data compression. Filters used in MWT have more degrees of freedom than those filters used in WT. In the area of face representation and classification, MWT is used for dimensionality reduction (data compression), noise alleviation, and feature extraction. Tarik *et al.* [85] proposed 2D Stationary MWT, PCA, and a histogram-based method (HBM) for facial recognition. Their system was evaluated using the ORL database. The highest recognition rate achieved was 94.5% based on HBM. 2D DMWT is applied to facial images in [34] (explained in detail in Chapter 4) for feature extraction and dimensionality reduction. The images are divided into four frequency sub-bands as explained in Chapter 4. Then the features of each sub-band were aligned around the origin using 2D Radon Transform (RT). Furthermore, 2D DWT was applied to the aligned features to obtain less storage requirements and more compacted features. Also, the authors expanded their work presented in [34] to use a 3D DWT instead of 2D DWT for facial recognition as explained in detail in Chapter 5 [35]. In [35], after applying 2D RT to each sub-band of DMWT matrix, the four resultant sub-bands were arranged in 3D matrix. Then the 3D DWT was applied to the 3D matrix to achieve high dimensionality reduction and hence accomplished less storage requirements. 2D DMWT was proposed in [86] for infrared facial recognition. In [86], infrared facial images were first represented using DMWT. Then PCA was applied to different MWT sub-bands for further feature compaction. Since each sub-band has different information, different weight was assigned to each using Fisher discriminant criterion. 4-levels of DMWT decompositions were applied to facial images of the ORL database in [87] to achieve highly dimensionality reduction. Then Particle Swam Optimization (PSO) was used to

seek the optimal facial representations.

Also, eigenvalues and eigenvectors extracted from facial images can be considered as another means of facial representation. Therefore, several researchers, such as [15, 16, 88, 89], etc. used the eigenfaces approach in the face recognition system. Other researchers transform images into different domains for further dimensionality reduction and then they applied the eigenfaces approach for optimal facial representations, for example in [90, 84, 81], etc. Facial images were efficiently represented using eigenvalues and eigenvectors of FastICA based on MWT features for face recognition [38] (see Chapter 7 for more detail). In [38], the system was extensively evaluated using the ORL, YALE, and FERET databases, and the accuracies accomplished were higher than those reported in [90, 84, 81]. Furthermore, FastICA in conjunction with 2D DMWT was also used in [39] for face recognition. In [39], as explained in detail in Chapter 8, FastICA was applied to different combinations of poses corresponding to the subimages of the low-low frequency subband of the MWT, and the ℓ_2-norm of the resulting features are computed to obtain discriminating and independent features, while achieving significant dimensionality reduction. The results achieved were higher than the results reported in [65, 90].

Although Discrete Cosine Transform (DCT) is well-known in the image compression, DCT is also widely used in the area of face representation and classification. DCT not only transforms the image from the spatial to the frequency domain, but also preserves most of the information in the low frequency region. By retaining only the low frequency region, the data compression and feature extraction are accomplished. Tools are integrated to overcome the weakness found in a single transform and to obtain facial features less sensitive to the facial variations. Face recognition based on DCT-based system was proposed in [91]. Sisodia *et al.* [92] employed DCT in conjunction with Improved Supported Vector Machine (ISVM) for face recognition. The authors in [93] introduced Local Binary Probabilistic Pattern (LBPP) and DCT for face recognition. LBPP was used to simplify the facial image into its principal components, and then these components

were captured by DCT. A hybrid technique based on DCT and GWT was presented in [94] to obtain discriminant facial features. Facial images were efficiently represented using the integrated tools DCT/PCA for feature extraction and NN for classification [95].

2.1.3 Compression Methods

Vector Quantization (VQ) has been successfully used in image compression since Linde, Buzo, and Gray (LBG) published their distinguished work in 1980 [96]. In [97], face recognition using four different algorithms was studied for face representations and classification. The four algorithms, namely, LBG, Kekres Proportionate Error Algorithm (KPE), Kekres Median Codebook Generation Algorithm (KMCG), and Kekres Fast Codebook Generation Algorithm (KFCG) were applied to Georgia Tech and Indian face databases to generate a codebook with suitable sizes. In [97], the algorithms were tested against DCT-based system, and the KFCG algorithm accomplished highest results among all the algorithms used. Later, the VQ algorithm was integrated with the other tools to improve the recognition rates. In [98], the facial features were extracted using VQ histogram and Markov Stationary Features (MSF), and the proposed system achieved recognition rate of 96.16% for ORL database. Both VQ and DCT were integrated in [42] for face recognition, See Chapter 11. In [42], DCT with appropriate truncation dimensions was applied to the processed faces for dimensionality reduction. Thereafter, VQ algorithm, employing KFCG approach for codebook initialization, was applied to the transformed truncated features for further feature compression. The final feature matrices for four databases used in their system, namely ORL, YALE, FERET, and FEI, were of size $4 \times 4 \times 16$ for each pose of each person. The results presented in [42] were higher than the results reported in [91, 97, 90]. Later, Aldhahab *et al.* [43] used VQ in conjunction with DWT for face recognition, as explained in Chapter 12. In [43], the DWT was applied to each pose in the applicable databases and only the LL frequency sub-band was retained and all other sub-bands were eliminated. Then, VQ, using KFCG for codebook initialization,

was used for further dimensionality reduction and best facial representation. The dimensionality reduction achieved in [43], which is $4 \times 4 \times 16$, was the same as in [42], while the recognition rates accomplished were higher than the accuracies reported in [42] for all databases. Subsequently, Aldhahab *et al.* [44] employed new techniques based on Facial Part Detection (FPD) with VQ for face recognition, see Chapter 13. Nose, both eyes, and Mouth were detected using FPD for dimensionality reduction and removing unimportant information. In the feature extraction phase, four groups for each person, one group for each detected part, were constructed for dimensionality reduction and feature discrimination by considering all parts of all training poses. For further data compression, the VQ algorithm, employing the KFCG approach for codebook initialization, was applied to each of the four groups. In [44], the authors achieved high dimensionality reduction compared with [42, 43]. Each person in [44] has features of $4 \times 4 \times 4 \times 16$ ($4 \times Centroid$) dimensions regardless of the number of poses used in the training mode, while the final features of each pose of each person in [42, 43] has $4 \times 4 \times 16$ (Centroid) dimensions.

2.1.4 Artificial Neural Network and Deep Learning (ANN & DL)

Artificial Neural Network (ANN) and Deep Learning (DL) are used to improve the performance of the facial recognition system. DL based multimodal face representation was proposed to improve the performance of the recognition system in [99], where faces appearing in multimedia applications (such as social media) were evaluated. Lei *et al.* [100] proposed Stacked Image Descriptor (SID), which employs deep structure, for face recognition. More complex facial information was extracted using deep structure. Also, features' compactness and discriminant were improved. In [101], Muti-Layer Perceptron (MLP) NN was applied to the extracted features, based on GWT/LDA algorithms, to improve the recognition accuracy. Better results were accomplished when DL NN was applied to the extracted features [102]. In [103], PCA and deep learning were both used to obtain different facial representation from different projections for face recognition.

2.1.5 Methods for Improving the Recognition Rates

To overcome limitations of some algorithms, such as PCA, LDA, and LPP when applied to the databases that have different facial variations, local-based features can achieve promising results under different facial variations [104]. Extracting local features from facial images can be done automatically using suitable tools or by partitioning faces into several parts. The most common methods used for extracting local features are Local Binary Patterns (LBP), GWT, etc. The LBP was originally applied to the texture representation and classification [105], then proposed in [106] for facial recognition. Using LBP in the feature extraction step will make the system more robust to facial variations in illumination and rotation. GWT facial representations are also more robust to facial variations in illumination and expressions [107].

In 2013, N.K. Patil *et al.* [108] used decorrelating local features based on DWT for facial recognition. The shape and textures were captured by employing LBP on GWT preprocessing facial images. Then, the local facial representation were obtained by decorrelating the features extracted by employing 2-level of DWT decompositions of normalized LBP features. Discriminant local facial features were obtained using Statistical Local Features (SLF) based on Multipartitioning Max Pooling (MPMP) with kernal representation [109]. In [110], global and local facial features, extracted using the combination of GWT and PCA, were used for face recognition. X. Duan and Z. Tan in [111] segmented the facial images to n overlapping subregions. Local facial features were acquired by applying PCA onto each normalized subregion. Even though the computational complexity increased in [111], since there were n overlapping subregions, the highest recognition rate accomplished was 97.7% based on the FERET database. In contrast to LBP and GWT, which required prior knowledge and two stages to obtain discriminant features, Simultaneous Local Binary Feature Learning and Encoding (SLBFLE) method, unsupervised learning method, learned from raw pixels and required only one stage to obtain the discriminant local features [112]. Face

images in [113] were partitioned to thirteen non-overlapping parts. Common Vector Approach (CVA), which is a subspace method that finds common vectors for each class and then uses those vectors in the recognition of classes, was applied to each part to obtain different facial feature representations. In [114], both feature information (FI) and structural information (SI) were extracted by applying ICA to fifteen non-overlapping regions. Although the computational complexity increased, the 95.138% recognition rate was achieved based on the FRGC V2 database.

CHAPTER 3: METHODOLOGY

3.1 Introduction

Since the input facial images have large dimensions, which have a lot of unnecessary information, the recognition rates are decreased and the computational complexity is increased. Therefore, it is necessary to select or extract efficient features to better represent the faces and to reduce the dimensions, which leads to less storage requirements. To do this task, different algorithms are applied to the facial images.

In this chapter, different algorithms, which are used in our research, are applied to the facial images to achieve better facial image representations. The algorithms are: 1. Discrete Cosine Transform (DCT). 2. Discrete Wavelet Transform (DWT), which is used for dimensionality reduction and noise eliminations. 3. Discrete Multiwavelet transform (DMWT). It is the same as DWT, but filters used in DMWT have more degrees of freedoms than those filters used in DWT. 4. Radon Transform (RT), which is used to get directional features of images. 5. Fast Independent Analysis (fastICA), which is employed to achieve more discriminating features than the original facial image features. 6. Vector Quantization (VQ). It is used here to compress the facial images for better representation of the data.

3.2 Discrete Cosine Transform (DCT)

DCT is considered one of the most powerful transforms in Signal and Image Processing. A signal can be more efficiently represented in a transform domain than in a spatial domain. In the DCT domain, it is possible to represent the original signal with fewer coefficients since most of the

signal energy is concentrated in the low-frequency regions. DCT is used in the lossy compression algorithms, e.g. MP3 for audio and JPEG for images. DCT yields only real number coefficients and therefore it is easier to use. The 2D DCT for a 2D signal $g(u, v)$ and its inverse can be expressed as [115]:

$$G(m,n) = \frac{2}{\sqrt{M \times N}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} g(u,v) c_m \\ cos(\frac{m(2u+1)\pi}{2M}) c_n cos(\frac{n(2v+1)\pi}{2N}) \tag{3.1}$$

where $g(u, v)$ is the signal in the time domain, M and N are the row and the column size of $g(u, v)$, and $G(m, n)$ is the (m^{th} row, n^{th} column) DCT coefficient for $u = 0, 1, \ldots M - 1$ and $v = 0, 1, \ldots N - 1$.

$$g(u,v) = \frac{2}{\sqrt{M \times N}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} G(m,n) c_m \\ cos(\frac{m(2u+1)\pi}{2M}) c_n cos(\frac{n(2v+1)\pi}{2N}) \tag{3.2}$$

where c_m, and c_n are defined as [115]:

$$c_m = \begin{cases} \frac{1}{\sqrt{2}} & \text{for m=0} \\ 1 & \text{otherwise} \end{cases} \tag{3.3}$$

3.3 Discrete Wavelet Transform (DWT)

Window Fourier Transform (WFT) is considered one of the solutions for representing a non-periodic signal. WFT gives information about the signals in both the time and frequency domains with limitations. The WFT has two drawbacks: using a window with an infinite length leads to per-

fect frequency resolution, but no time resolution. Furthermore, a window with a finite length gives better time resolution, but poorer frequency resolution. Hence, the Discrete Wavelet Transform overcomes the drawbacks of the WFT. The DWT, considering both time and frequency resolutions at the same time, is considered as an extended version of the Discrete Fourier Transform (DFT) [116].

Information in different frequency bands plays important roles in the face recognition system. For facial images, the global characteristics of the facial images are located in the low frequency band, while all the facial details are localized in the high frequency band. Facial expressions are a facial detail that can affect the frequency components of the facial images. To alleviate the impact of the facial expressions on the face recognition system, the high frequency bands are removed and only the low frequencies are retained [117]. Hence, Wavelet Transform (WT), one of the signal processing transforms, represents the facial images in both time and frequency domains. In each level of decomposition, WT divides the signal into four sub-bands corresponding to the low and high frequency components. Therefore, the effect of different facial expressions can be reduced by removing the high frequency components.

After DWT is applied to the face image, four different frequency sub-bands, namely, LL, LH, HL, and HH result, where L and H are associated with the low pass and high pass filters respectively. For further decomposition, the DWT is applied again to the LL, low-low frequency sub-band. Global and detailed information about the face image are extracted by using DWT. The low frequency sub-band, which contains most of the face image energy, has lower resolution than the original face image. Based on DWT theory, the high frequency components are more sensitive to the facial variations than the low frequency components. Therefore, only the LL sub-band of DWT is retained for dimensionality reduction and for better facial representations [116, 118].

The DWT, the discrete form of Continuous Wavelet Transform (CWT), is a famous transformation

due to it is wide employment in several areas, such as Image Processing, Computer Vision, Pattern Recognition, and Signal Processing. The DWT, based on Multi-Resolution Analysis (MRA), uses two sets of functions called scaling function $\phi(t)$ and wavelet function $\psi(t)$. Both functions ($\phi(t)$ & $\psi(t)$) are associated with low pass and high pass filters, respectively. To decompose an image into multiple frequency bands, it goes through successive low pass and high pass filters [116]. In DWT, the signal $x[n]$ is passed through low pass and high pass filters and then half of the samples are eliminated using Nyquist's criterion. The first level of decomposition of the DWT can be expressed mathematically as:

$$y_{low} = \sum_{n} x[n] \cdot L[2k - n] \tag{3.4}$$

$$y_{high} = \sum_{n} x[n] \cdot H[2k - n] \tag{3.5}$$

where $L[\cdot]$ & $H[\cdot]$ are corresponding to low pass and high pass filters, respectively. $x[n]$ is the input signal. y_{low} and y_{high} are the output from the low pass and high pass filters, respectively [116, 118, 119].

The 2D DWT is an extended version of the 1D DWT. The 2D scaling function $\phi(x, y)$ is the extended version of 1D scaling function $\phi(x)$. Similarly the 2D wavelet function $\psi(x, y)$ that is the extended version of the 1D wavelet function $\psi(x)$ [119, 120].

In 2D DWT, a 2D scaling function $\phi(x, y)$ and three 2D wavelet functions $\psi^H(x, y)$, $\psi^V(x, y)$, and $\psi^D(x, y)$ are needed. Each function can be represented as the product of their respective one-dimensional (1D) functions. The separable scaling and wavelet functions are expressed as [119, 120]:

$$\phi(x,y) = \phi(x)\phi(y) \tag{3.6}$$

$$\psi^H(x,y) = \psi(x)\phi(y) \tag{3.7}$$

$$\psi^V(x,y) = \phi(x)\psi(y) \tag{3.8}$$

$$\psi^D(x,y) = \psi(x)\psi(y) \tag{3.9}$$

The scaling function $\phi(x,y)$ gives the approximation coefficients of an image. The wavelet functions measure the intensity variations of an image along different directions. $\psi^H(x,y)$, $\psi^V(x,y)$, and $\psi^D(x,y)$ measure variations along Horizontal edge, Vertical edge, and Diagonal, respectively [119, 120].

Given the two dimensional separable scaling and wavelet functions shown in Eq. 3.6 - 3.9, the 2D DWT is a straightforward. The scaled and translated basis functions, which are extended version of 1D basis functions, are defined as:

$$\phi_{j,m,n}(x,y) = 2^{j/2}\phi(2^j x - m, 2^j y - n) \tag{3.10}$$

$$\psi^i_{j,m,n}(x,y) = 2^{j/2}\psi^i(2^j x - m, 2^j y - n), \quad i = \{H,V,D\} \tag{3.11}$$

where the index i identifies the directional wavelets in Eq. 3.6 - 3.9. The 2D DWT for 2D image $f(x,y)$ of a size $M \times N$ is expressed as:

$$W_\phi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \phi_{j_0,m,n}(x, y) \tag{3.12}$$

$$W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \psi_{j,m,n}^i(x, y) \tag{3.13}$$

Where $i = \{H, V, D\}$, j_0 is an arbitrary start scale. $W_\phi(j_0, m, n)$ represents the approximation of the image $f(x, y)$ at scale j_0. The $W_\psi^i(j, m, n)$ denotes the Horizontal, Vertical, and diagonal coefficients at scale $j \geq j_0$. Usually, $j_0 = 0$ and $N = M = 2^J$ are selected such that $j = 0, 1, 2, \ldots, J-1$ and $m = n = 0, 1, 2, \ldots, 2^j - 1$ [119].

For a 2D image $I(x, y)$, akin to the 1D DWT explained in details in [118, 119], the 2D DWT uses Low pass filter L with the impulse response $l(n)$, high pass filter H with the impulse response $h(n)$, and down-sampling by a factor of 2. Convolving the rows of the input image $I(x, y)$ with both L filter and H filter produces two sub-images $I_L(x, y)$ and $I_H(x, y)$ that have horizontal resolution down-sampled by a factor of 2. In the same manner, the two sub-images are filtered column-wise and down-sampled by a factor of 2 to yield four sub-images $I_{LL}^d(x, y)$, $I_{LH}^d(x, y)$, $I_{HL}^d(x, y)$, and $I_{HH}^d(x, y)$, where the superscript d represents the decomposition level. $I_{LL}^d(x, y) = W_\phi$ is the approximation coefficients of the image $I(x, y)$. $I_{LH}^d(x, y) = W_\psi^V$, $I_{HL}^d(x, y) = W_\psi^H$, and $I_{HH}^d(x, y) = W_\psi^D$ are the detail-sub-images, which correspond to the Vertical, Horizontal, and Diagonal directions of the original image $I(x, y)$, respectively. The same procedures are performed on the $I_{LL}^1(x, y)$ sub-image for the second level of decomposition, i.e., $d = 2$ [120, 119, 118]. Figure 3.1 shows the structure of the first and second levels of decompositions of DWT for 2D signals. In this dissertation, one level of decomposition is applied to the input facial images and the well-known Haar filter is employed in the DWT [119, 118].

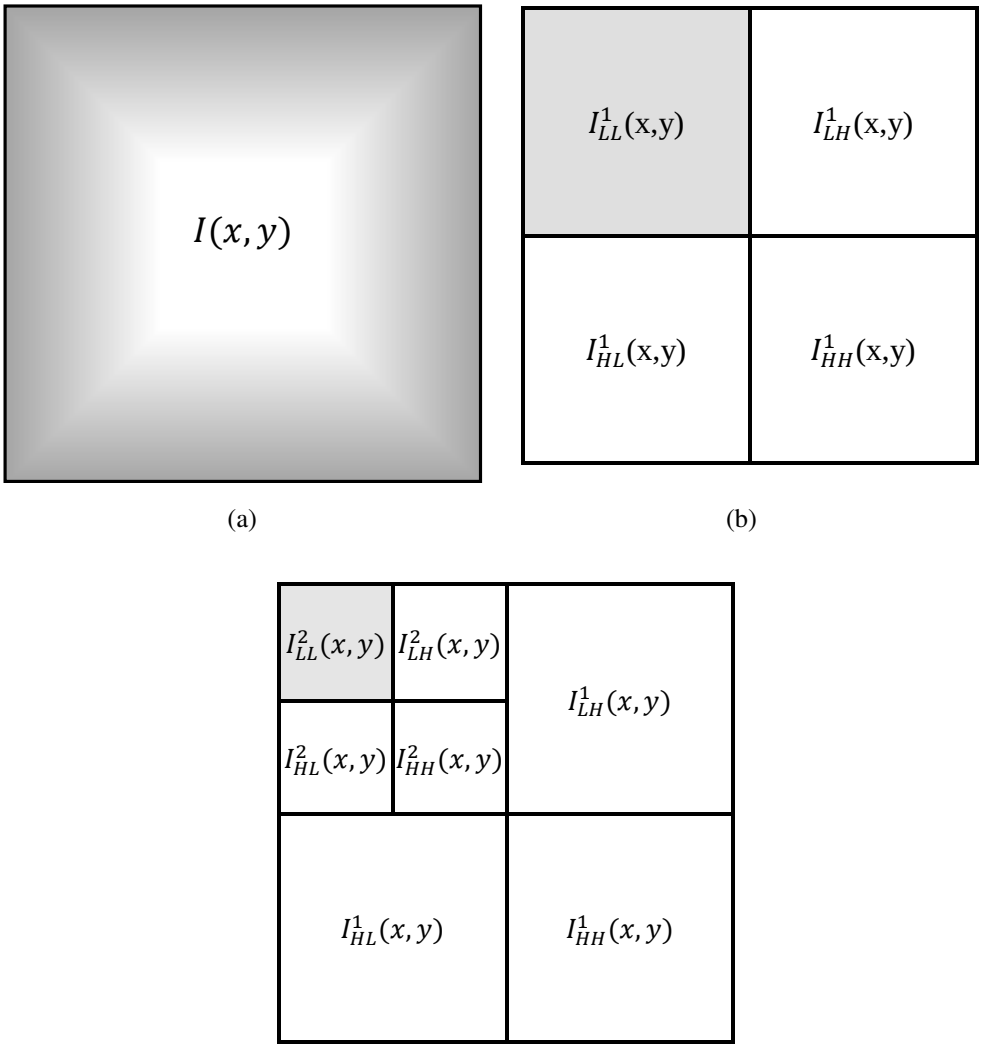
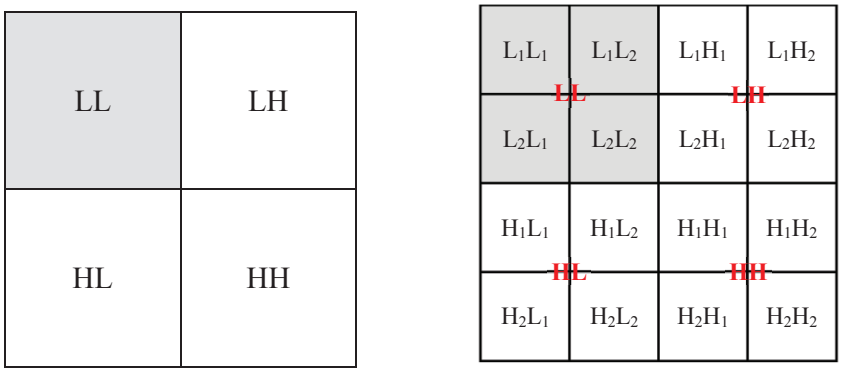


Figure 3.1: DWT structures. Figure 3.1-a shows the input image.Figure 3.1-b demonstrates the first level of decomposition, which divids the signal into four frequency subbands. Figure 3.1-c shows the second level of decompositions of DWT, where L and H correspond to the low pass and high pass filters, respectively.



(a) 1-Level of DWT

(b) 1-Level of DMWT

Figure 3.2: Shows the 1-level of DWT and DMWT decomposition. Figure 3.2-a is the 1-level of 2D DWT decomposition. Figure 3.2-b is the 1-level of 2D DMWT decomposition.

3.4 Multiwavelet Transform (MWT)

The MRA for multiwavelet case is the same as the standard wavelet transform (MRA) except that MWT uses Q scaling functions, $\phi_1(t-k),\ \phi_2(t-k), \ldots,\ \phi_Q(t-k)$. The Q scaling functions expressed in the vector notation as $\Phi(t) = [\phi_1(t), \ldots,\ \phi_Q(t)]^T \in L^2(R)^Q$ is called the multi-scaling function. Analogous to the scalar case, this can satisfy the matrix dilations as:

$$\Phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} H_k \cdot \Phi(2t-k) \tag{3.14}$$

where H_k is a $Q \times Q$ matrix dimensions for each integer k and it is associated with the low pass filter. The Q scaling functions are associated with the Q wavelet functions that can satisfy the vector notation $\psi(t) = [\psi_1(t), \ldots,\ \psi_Q(t)]^T \in L^2(R)^Q$, which is called multi-wavelet functions. The Q multiwavelet functions satisfy the wavelet matrix equation as:

$$\Psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} G_k \cdot \Phi(2t-k) \tag{3.15}$$

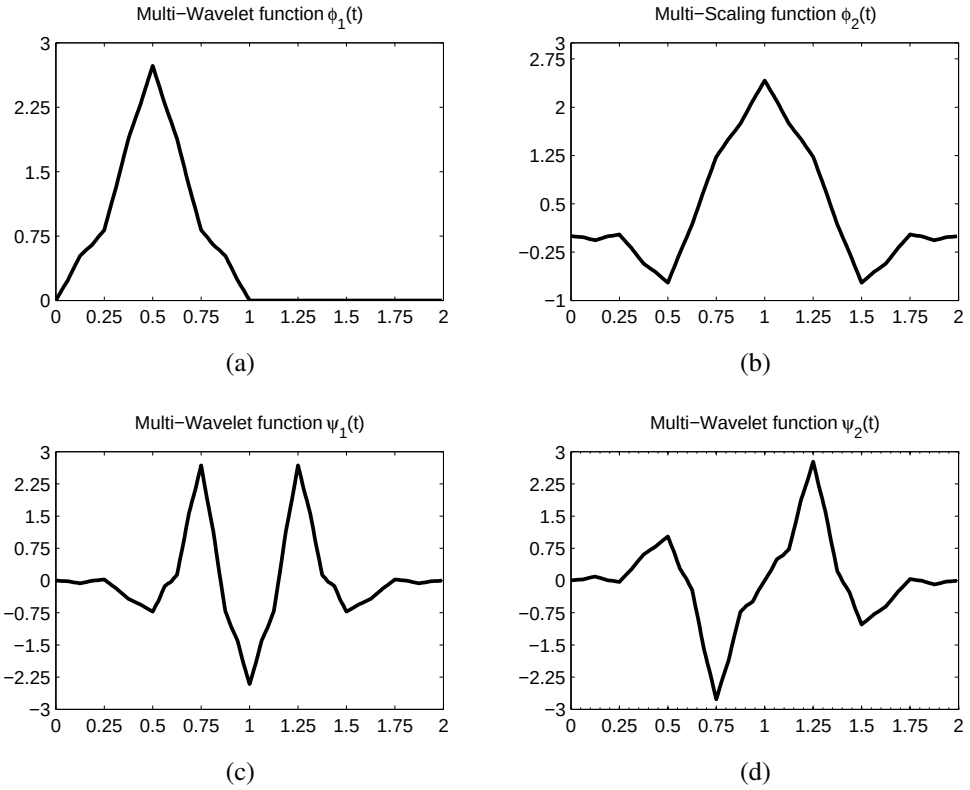


Figure 3.3: Multi-scaling and Multi-wavelet functions for GHM filter

where G_k is a $Q \times Q$ matrix dimensions for each integer k and G_k is associated with high pass filter [121, 122].

One of the famous multiwavelets system was constructed by Geronimo, Hardian, and Massopust (GHM) [123]. Their system uses multiplicity $Q = 2$. In such a case, there are two scaling functions $\Phi(t) = [\phi_1(t) \;\; \phi_2(t)]^T$ as shown in figure 3.3-(a & b), and two wavelet functions $\Psi(t) = [\psi_1(t) \;\; \psi_2(t)]^T$ as shown in figure 3.3-(c & d) [121]. The matrix elements in the GHM filter provide extra degrees of freedom in comparison with the scalar case. The GHM filter combines orthogonality, high order of approximation, and symmetry. Also, both the multi-scaling and multi-wavelet functions of GHM filter are quite smooth [121]. For the GHM case, both Eqs. (3.14)

& (3.15) are written as [121]:

$$\Phi = \begin{bmatrix} \phi_1(t) \\ \phi_2(t) \end{bmatrix} = \sqrt{2} \sum_k H_k \cdot \begin{bmatrix} \phi_1(2t-k) \\ \phi_2(2t-k) \end{bmatrix} \tag{3.16}$$

$$\Psi = \begin{bmatrix} \psi_1(t) \\ \psi_2(t) \end{bmatrix} = \sqrt{2} \sum_k G_k \cdot \begin{bmatrix} \phi_1(2t-k) \\ \phi_2(2t-k) \end{bmatrix} \tag{3.17}$$

Hence, the dilation Eq. (3.14) and wavelet Eq. (3.15) can be written as:

$$\Phi = \begin{bmatrix} \phi_1(t) \\ \phi_2(t) \end{bmatrix} = H_0\Phi(2t) + H_1\Phi(2t-1) + H_2\Phi(2t-2) + H_3\Phi(2t-3), \tag{3.18}$$

$$\Psi = \begin{bmatrix} \psi_1(t) \\ \psi_2(t) \end{bmatrix} = G_0\Phi(2t) + G_1\Phi(2t-1) + G_2\Phi(2t-2) + G_3\Phi(2t-3), \tag{3.19}$$

and H_k & G_k, where $k = 0, 1, 2, 3$, are expressed as [121, 122]:

$$H_0 = \begin{bmatrix} \frac{3}{5\sqrt{2}} & \frac{4}{5} \\ \frac{-1}{20} & \frac{-3}{10\sqrt{2}} \end{bmatrix} H_1 = \begin{bmatrix} \frac{3}{5\sqrt{2}} & 0 \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix}, H_2 = \begin{bmatrix} 0 & 0 \\ \frac{9}{20} & \frac{-3}{10\sqrt{2}} \end{bmatrix}, H_3 = \begin{bmatrix} 0 & 0 \\ \frac{-1}{20} & 0 \end{bmatrix}$$

$$G_0 = \begin{bmatrix} \frac{-1}{20} & \frac{-3}{10\sqrt{2}} \\ \frac{1}{10\sqrt{2}} & \frac{3}{10} \end{bmatrix}, G_1 = \begin{bmatrix} \frac{9}{20} & \frac{-1}{\sqrt{2}} \\ \frac{-9}{10\sqrt{2}} & 0 \end{bmatrix}, G_2 = \begin{bmatrix} \frac{9}{20} & \frac{-3}{10\sqrt{2}} \\ \frac{9}{10\sqrt{2}} & \frac{-3}{10} \end{bmatrix}, G_3 = \begin{bmatrix} \frac{-1}{20} & 0 \\ \frac{-1}{10\sqrt{2}} & 0 \end{bmatrix}$$

There are four favorable properties for GHM scaling functions [121, 122]:

1. The short support for both scaling functions of GHM filters are [0,1] and [0,2].

2. The approximation of the GHM system based on the second order.

3. Translated version of the scaling functions and the wavelet functions are orthogonal.

4. Symmetry is a property of both the scaling function and wavelet function.

The transformation matrix T, which is the GHM filter matrix, of DMWT is constructed using the above scaling and wavelet functions and can be written as [121, 122]:

$$T = \begin{bmatrix} H_0 & H_1 & H_2 & H_3 & 0 & 0 & \cdots \\ G_0 & G_1 & G_2 & G_3 & 0 & 0 & \cdots \\ 0 & 0 & H_0 & H_1 & H_2 & H_3 & \cdots \\ 0 & 0 & G_0 & G_1 & G_2 & G_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \tag{3.20}$$

As we mentioned earlier, both H_k and G_k are associated with low pass and high pass filters, respectively. Both of them are 2×2 matrices for each integer k [121, 122]. Hence the transformation matrix T is written as [121, 122, 124]:

$$T = \begin{bmatrix} H_{0_{0,0}} & H_{0_{0,1}} & H_{1_{0,0}} & H_{1_{0,1}} & H_{2_{0,0}} & H_{2_{0,1}} & H_{3_{0,0}} & H_{3_{0,1}} & 0 & 0 & 0 & 0 & \cdots \\ H_{0_{1,0}} & H_{0_{1,1}} & H_{1_{1,0}} & H_{1_{1,1}} & H_{2_{1,0}} & H_{2_{1,1}} & H_{3_{1,0}} & H_{3_{1,1}} & 0 & 0 & 0 & 0 & \cdots \\ G_{0_{0,0}} & G_{0_{0,1}} & G_{1_{0,0}} & G_{1_{0,1}} & G_{2_{0,0}} & G_{2_{0,1}} & G_{3_{0,0}} & G_{3_{0,1}} & 0 & 0 & 0 & 0 & \cdots \\ G_{0_{1,0}} & G_{0_{1,1}} & G_{1_{1,0}} & G_{1_{1,1}} & G_{2_{1,0}} & G_{2_{1,1}} & G_{3_{1,0}} & G_{3_{1,1}} & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & H_{0_{0,0}} & H_{0_{0,1}} & H_{1_{0,0}} & H_{1_{0,1}} & H_{2_{0,0}} & H_{2_{0,1}} & H_{3_{0,0}} & H_{3_{0,1}} & \cdots \\ 0 & 0 & 0 & 0 & H_{0_{1,0}} & H_{0_{1,1}} & H_{1_{1,0}} & H_{1_{1,1}} & H_{2_{1,0}} & H_{2_{1,1}} & H_{3_{1,0}} & H_{3_{1,1}} & \cdots \\ 0 & 0 & 0 & 0 & G_{0_{0,0}} & G_{0_{0,1}} & G_{1_{0,0}} & G_{1_{0,1}} & G_{2_{0,0}} & G_{2_{0,1}} & G_{3_{0,0}} & G_{3_{0,1}} & \cdots \\ 0 & 0 & 0 & 0 & G_{0_{1,0}} & G_{0_{1,1}} & G_{1_{1,0}} & G_{1_{1,1}} & G_{2_{1,0}} & G_{2_{1,1}} & G_{3_{1,0}} & G_{3_{1,1}} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \tag{3.21}$$

The results of the one level of decomposition of DMWT are shown in figure 3.2-B. It is obvious there that the input data is divided into four main sub-bands and each sub-band is further divided

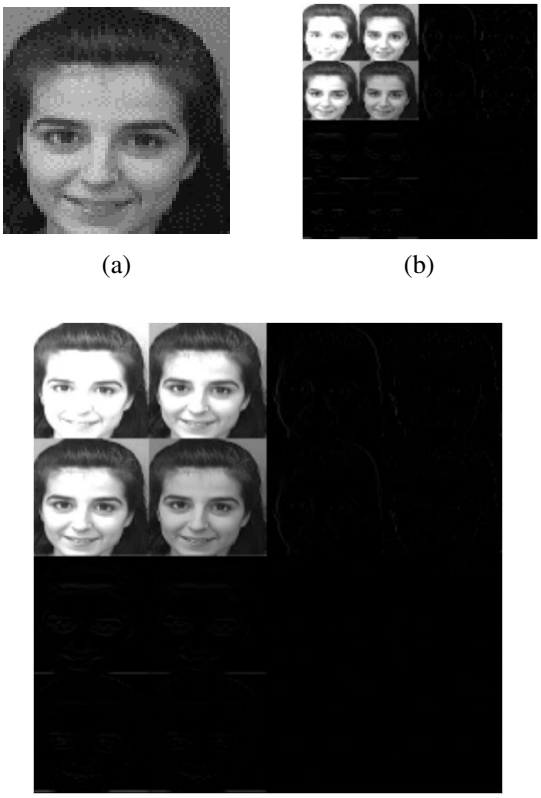
into four sub-sub-bands. For example, the sub-sub-band L_2H_1 corresponds to the data from the first channel (high pass filter) in the horizontal direction and second low pass filter in the vertical direction. The total number of sub-bands related to the number of decompositions can be expressed as $4 + 12 \times L^*$, where L^* is the number of multiwavelet decomposition levels.

In contrast to the DWT, the DMWT has several favorable features, such as achieving a good reconstruction (orthogonality), better performance (linear phase symmetry), high order of approximation (vanishing moments), and compact support. These desirable features cannot be achieved at the same time in the scalar wavelet, while in MWT provide more degree of freedom and give excellent performance in signal and image applications [121].

3.4.1 Preprocessing of Multiwavelet

The objective of using the preprocessing step in the Multiwavelet transform is to associate a scalar input signal of length Z to a sequence of length-2 vectors [122]. In the multiwavelets, both the low pass filter H_k and the high pass filter G_k, are $Q \times Q$ matrix filters. Thus, in contrast to DWT, DMWT requires a vector valued input signal instead of using a scalar input signal during the convolution step. In multiwavelet transform, the signal is passed through a so called preprocessing step. In the preprocessing step, a sequence of input vectors is extracted from the scalar input to achieve better performance [121, 122, 125].

Assume that the dimensions of the input image after resizing to power of two are $Z \times Z$. The preprocessing is called an oversampled-scheme (repeated row preprocessing) if it produces a Z-length 2 vector. The preprocessing is called a critically-sampled scheme (approximation based preprocessing) if the preprocessing produces a $Z/2$-length 2 vector. In our dissertation, we used critically-sampled scheme preprocessing [121, 122, 125]. Moreover, there are several other ways to do preprocessing, as described in [126, 127, 128, 129, 130].



(a) (b)

(c)

Figure 3.4: Shows different preprocessing algorithms.

Figure 3.4 shows different preprocessing techniques. Figure 3.4-a, shows the original image. Figure 3.4-b illustrates the result of applying DMWT to the original image, preprocessed using critically-sampled scheme preprocessing (approximation-based preprocessing). Figure 3.4-c applies DMWT to the original image using oversampled scheme preprocessing (repeated row preprocessing).

The transformation matrix T of DMWT using critically-sampled scheme preprocessing has the same dimensions as the input image ($Z \times Z$ dimensions, where Z is power of 2).

3.4.1.1 A Critically-Sampled Scheme (Approximation Based Preprocessing)

Herein, the continuous-time multiwavelet (based on approximation properties) is used to describe the preprocessing algorithm that yields the critical-sampled scheme signal representation. A 1-D signal $f(t)$ is used to obtain a vector input from the scalar input signal [121, 122, 125, 131]. Approximation based preprocessing is inherently connected to the GHM scaling functions. A basis of subspace V_0 is produced using the translated version of the GHM scaling functions. Therefore, $f(t)$ is a linear combination of these scaling functions:

$$f(t) = \sum_{n} V_{1,n}^{(0)} \phi_1(t-n) + V_{2,n}^{(0)} \phi_2(t-n) \tag{3.22}$$

Since the support of both $\phi_1(t)$ and $\phi_2(t)$ are [0,1] and [0,2], respectively as shown in figure 3.3-(a & b), the input sequence $f[n]$ has samples of $f(t)$ at integer and half-integer points. $\phi_1(t)$ is zero at all integer points while $\phi_2(t)$ is non-zero at integer 1. Accordingly:

$$f[2n] = f[n] \qquad f[2n+1] = f[n+\frac{1}{2}]$$

Hence,

$$\begin{aligned} f[2n] &= V_{2,n}^{(0)} \phi_2(1) \\ f[2n+1] &= V_{2,n-1}^{(0)} \phi_2(\frac{3}{2}) + V_{1,n}^{(0)} \phi(\frac{1}{2}) + V_{2,n}^{(0)} \phi_2(\frac{1}{2}) \end{aligned} \tag{3.23}$$

Taking into a consideration the symmetry property $\phi_2(\frac{3}{2}) = \phi_2(\frac{1}{2})$, the coefficients $V_{1,n}^{(0)}$ and $V_{2,n}^{(0)}$ in Eq.(3.23) can be written as:

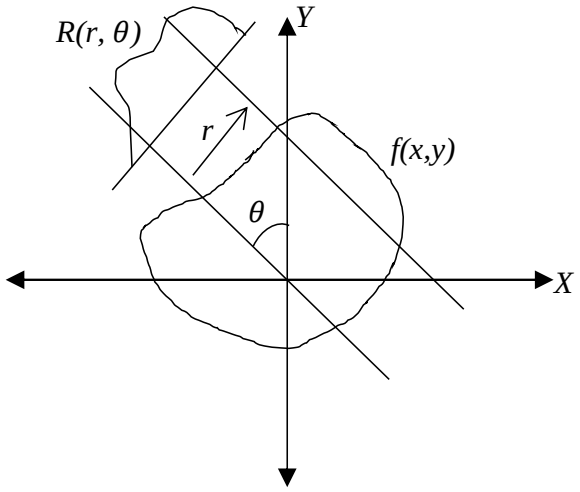


Figure 3.5: Radon Transform

$$V_{2,n}^{(0)} = \frac{f[2n+2]}{\phi_2(1)}$$
$$V_{1,n}^{(0)} = \frac{\phi_2(1)f[2n+1] - \phi_2(\frac{1}{2})(f[2n+2] + f[2n])}{\phi_2(\frac{1}{2})\phi_1(\frac{1}{2})} \tag{3.24}$$

As we can see, the two input rows $V_{1,n}^{(0)}$ and $V_{2,n}^{(0)}$ are obtained from the input signal $f[n]$. The 1-D approximation-based preprocessing for computing the discrete multiwavelet transform can be extended into 2-dimensional approximation-based preprocessing [121, 122, 125, 131].

3.5 Radon Transform

The Radon Transform $R(r, \theta)$ of an image $f(x, y)$ is defined as a slant or sloping line with angle θ from the y axis and with distance r from the origin [132], as shown in figure 3.5. It is a very

powerful tool to capture all the useful directional features of images. In image processing, the Radon Transform is the projection of an image $f(x, y)$ along the specified direction or angles. The result of the radon transform is represented by the sum of all intensities along these directions or angles. The Radon Transform for a two-dimensional image $f(x, y)$ is:

$$r(r, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(r - xcos\theta - ysin\theta)dxdy \tag{3.25}$$

where $\delta(.)$ is the Dirac delta function, $r \in (-\infty, \infty)$is the perpendicular distance from the origin, and $\theta \in (0, 2\pi)$ is an angle formed by the distance vector [133], as shown in the figure 3.5.

3.6 Vector Quantization (VQ)

Vector Quantization is a lossy compression technique in image processing in which each block (having the dimensions of $p \times q$) of an image is replaced by a codeword, or centroid, drawn from a closed set. The set is called a Codebook, and the set consists of C codewords. The famous Linde, Buzo, and Gray (LBG) algorithm was presented in [96], which was based on the Lloyd algorithm. In that algorithm, the substitution process is based on the similarity between each image block and each codeword. This is done by employing the correlation between adjacent pixels in the spatial domain. The codeword that yields the minimum distortion replaces an image block in the quantized representation. Five parameters should be considered when designing a Vector Quantizer VQ. The first two are the dimensions of the codeword, i.e., p and q. Inherently, image blocks and codewords should have the same dimensions. The third parameter is the number of codewords in a codebook, i.e., C, that is chosen to minimize the total error between the original image and its quantized version. Obviously, more codewords result in less distortion, but also more codewords require more storage space. The last two parameters chosen when designing a Vector Quantizer are

codebook initialization, and the distance measure criterion. Different codebook initialization techniques give different final codebooks when the same number of iterations is considered. The most well-known distance measure criteria are: Mean Squared Error (MSE), norm-based distance criteria, and Modified Itakura-Satio [134]. All design parameters should be optimized at the same time since different combinations of different parameters result in different codebooks. For instance, increasing the dimensions of the codewords decreases the number of blocks to be processed, but increases the total error given that the distance measure and the number of iterations are the same in both cases.

The algorithm starts by calculating a mean vector of all image blocks, and this mean vector is the first centroid. For the LBG algorithm, the next two centroids are found by adding and subtracting a small number ϵ to/from the first centroid. The image blocks are replaced by either one of these two centroids depending on which one is nearest. The updated two centroids are calculated by averaging all image blocks that were encoded using each of them, i.e., each new centroid is the average of all image blocks that were closer to it than the other centroid. Four new centroids are computed by adding and subtracting ϵ to/from these two centroids. All image blocks that are closer to each centroid are averaged and this average replaces that centroid. The splitting process continues until the number of required codewords, i.e., C, is reached. In image compression applications, for each image block the algorithm finds the nearest codeword and overwrites the data in it by that codeword. This procedure is repeated as many times as the number of image blocks expressed in Eq.13.1. The final outputs of the LBG algorithm are: codebook, and Index table. The Index table is a matrix having the same dimensions as the input image but containing numbers referring to the index of the codeword that replaced this image block. When VQ was used in data transmission, both transmitter and receiver had the same codebook and the transmitter sends only the index for each codeword. The receiver decodes that index and places the corresponding codeword in the position of that block. Obviously, a transmitting protocol had to be used to keep

the spatial order of the image blocks in the correct form. For face recognition applications, the outputs of the VQ might be used together, or only the codebook is used.

This LBG initialization has the disadvantage of being a time consuming process. Therefore, [135] presented a new alternative algorithm by which initial codebooks can be computed faster than LBG. Kekre Fast Codebook Generation (KFCG) is the name of that algorithm. The initialization process begins by calculating the mean of all image blocks. Based on the value of the first element in the mean vector, image blocks are divided into two groups. This division depends on whether or not the first value in each block exceeds the corresponding value in the mean vector. The means of the new groups are calculated. Again, the split process is repeated with the same comparison but with different values satisfactory. The comparison/averaging steps are repeated until the total number of codewords is achieved.

3.7 Independent Component Analysis (ICA)

Blind Source Separation (BSS) is a method of restoring the original source signal from a set of mixed signals without any prior knowledge of the source signals and mixing signals. BSS is a fundamental and challenging problem in signal processing. Several algorithms and approaches have been proposed to solve the BSS problems, such as PCA [136], Projection Pursuit [137], Blind Convolution [138], BP Neural Network [139], etc. ICA has been developed for a long time to solve the BSS problems, and it is a powerful tool in signal processing and data analysis [64, 68, 140]. Hence various approaches have been developed based on the ICA approach to solve the problems of BSS, such as [64, 66, 138, 140, 141, 142], etc.

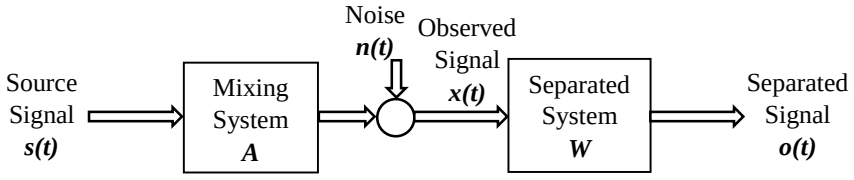


Figure 3.6: Block diagram for Independent Component Analysis (ICA).

The principle of ICA is to express a set of random variables as a linear combination of statistically independent components (source signals), where both source signal and mixing matrix are unknown. Figure 3.6 shows the block diagram of ICA. Let a vector $S = [s_1,\ s_2,\ \cdots,\ s_r]^T$ be denoted as r-dimensional unknown source signals with zero mean and $X = [x_1,\ x_2,\ \cdots,\ x_r]^T$ is denoted as r-dimensional observed signals. Assume a matrix A is a mixing matrix. Then the mixture module can be expressed as:

$$X = A.S = \sum_{i=1}^{r} a_i s_i, \qquad i = 1,\ 2,\ 3,\ \cdots,\ r \tag{3.26}$$

where $A = (a_1,\ a_2,\ \cdots,\ a_r) \in R^{r \times r}$ is a full rank matrix and a_i is the base vector of the mixing matrix. s_i is the i^{th} independent component [75]. Based on ICA assumptions explained in detail in [75], the noise $n(t)$ introduced in figure 3.6 is very small compared with the observed signal X, hence, it is neglected.

Each observed x_j is a linear combination of weighted independent components s_i. Based on several assumptions and constraints corresponding to ICA problems, which are explained in detail in [75], the ICA algorithm attempts to find the transformation matrix W of the mixing signals such that

$$O = W.X = W.A.S \tag{3.27}$$

is an estimation of the source signal and each o_i is as independent as possible [67, 75]. ICA, which is an extension and a generalization of PCA that decorrelates the second order statistic, is well known to decorrelate the high order statistics of the input data. Therefore, if the source signal has a Gaussian distribution, PCA can only be performed to achieve the separation. This is due to the fact that the uncorrelation property is equivalent to statistical independence for normal distribution random process. Hence, the signal can be described by only the second order statistic. Data representation using ICA is more meaningful than data representation based on PCA [64, 68, 75]. In the field of face recognition, important information may be localized in the high order relationships between pixels. Also, ICA can capture the local characteristics of the input face images. Due to these favorable properties of ICA, the ICA bases vectors are extremely robust to the natural face variations such as facial expressions, rotations, and illuminations [64, 67, 68, 143]. As discussed in [64, 75, 144], there is another way of representing the input facial images by using an unknown mixing matrix A as is used in Eq.3.27.

Both Centering and Whitening are preprocessing steps for ICA. Applying a linear transformation to the r-dimensional signal X such that the output $O = W.X$ is a white signal is called a whitening process. Before employing ICA, a transformation matrix W is applied to the observed vector X such that the covariance matrix of the output $O = W.X$ is an identity matrix I (O is a white signal). Let X be an r-dimensional random vector with zero mean and

$$C_X = Cov(X) = E[X.X^T] = EDE^T \tag{3.28}$$

is a covariance matrix of X, which is positive definite. D is the diagonal matrix of its eigenvalues and E is an orthogonal matrix of its eigenvectors such that $EE^T = I$. Therefore the linear whitening transform is $W = D^{-\frac{1}{2}}E^T$. Hence, the output signal $y = WX$ has a covariance:

$$
\begin{aligned}
Cov(y) &= E[yy^T] \\
&= E[WX.X^TW^T] \\
&= E[D^{-\frac{1}{2}}E^T.XX^T.ED^{-\frac{1}{2}}] \\
&= D^{-\frac{1}{2}}E^T.E[XX^T].ED^{-\frac{1}{2}} \\
&= D^{-\frac{1}{2}}E^T.EDE^T.ED^{-\frac{1}{2}} \\
&= I
\end{aligned}
$$

Applying the whitening process to the observed signal leads to the transformation matrix W being an orthogonal matrix, hence, the output vector O is white. ICA uses the whitening preprocessing to reduce the complexity of the problem [64, 66, 67, 68, 75].

3.7.1 Fast Independent Component Analysis (FastICA)

FastICA, proposed and developed by [140, 144], is a faster algorithm for ICA. FastICA, which tries to find the maximum of the non-Gaussianity of W^TX, uses a fixed point algorithm. There are several advantages of using FastICA compared with ICA algorithms, for example FastICA has faster convergence compared with traditional ICA convergence, FastICA does not required to select a step size to be selected compared with ICA based on the gradient based algorithm. FastICA finds the independent non-Gaussian signal using any arbitrary non-linear function $\theta(t)$, while other algorithms required the Probability Density Function (PDF) of the selected non-linear function to be evaluated [75, 144].

The method of FastICA tries to find a local extreme of the Kurtosis of a linear combination of

the observed mixing signals or variables. In the learning process, the method of FastICA finds the direction of the weight vector w_i of W. Moreover, the goal of the fixed point is to find the maximum of the non-Gaussianity of $o_i = w_i^T x$. FastICA uses negentropy as the objective function [75]. [145] introduced a simple way to approximate the negentropy function:

$$J(o_i) \approx [E\{G(o_i)\} - E\{G(\nu)\}]^2 \tag{3.29}$$

where o_i is the output random variable with zero mean and unit variance, G is a non-quadratic function, and ν is a Gaussian random variable with zero mean and unit variance. The approximation given in Eq. (3.29) leads to a new objective function for estimating ICA. Hence, maximizing the function J_G

$$J_G(w) = [E\{G(w^T x)\} - E\{G(\nu)\}]^2 \tag{3.30}$$

gives one independent component and w is the r-dimensional weighted vector constrained such that $E\{(w^T x)^2\} = 1$. The algorithm for FastICA based on the maximum negentropy objective function is explained in details in [75, 140, 144, 145, 146].

The one unit fixed point algorithm based on the sphere data is derived below. The maximum value of $J_G(w)$ is obtained when $E\{G(w^T x)\}$ reaches the optimum under the constraint of $E\{w^T x)\} = \|w\|^2 = 1$.

$$\begin{cases} \underset{w}{\text{Maximize}} & J_G(w) = E\{G(w^T x)\} \\ \text{s.t} & \|w\|^2 = 1 \end{cases} \tag{3.31}$$

Based on the Kuhn-Tucker condition, the optimum of $E\{G(w^T x)\}$ under the constraint of $E\{w^T x)\} = \|w\|^2 = 1$ is achieved where

$$E\{xg(w^T x)\} - \beta w = 0 \tag{3.32}$$

where β is a constant and given by $\beta = E\{w_0^T x g(w_o^T x)\}$, w_0 is the optimum value of w. The Newton's iterative method is used to solve the Eq.(3.32). By referring to the left hand side by F, we find the Jacobian matrix $JF(w)$

$$JF(w) = E\{xx^T g\prime(w^T x)\} - \beta I \tag{3.33}$$

Eq.3.33 can be simplified by approximating the first term. Since the data is spherical, then we can split the expectation as

$$E\{xx^T g\prime(w^T x)\} \approx E\{xx^T\}E\{g\prime(w^T x)\} = E\{g\prime(w^T x)\}$$

Since $E\{xx^T\} = I$. Therefore, the Jacobian matrix becomes diagonal and can be inverted. Using w instead of w_0, the approximation Newton iteration is expressed as

$$w_{k+1} = w_k - \frac{[E\{xg(w_k^T x)\} - \beta w_k]}{E\{g\prime(w_k^T x)\} - \beta} \tag{3.34}$$

To improve the stability, w_{k+1} is normalized by $(w_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|})$, and $\beta = E\{w^T x g(w^T x)\}$. The Eq.3.34 can be more simplified by multiplying both sides by $\beta - E\{g\prime(w_k^T x)\}$, thus Eq.3.34 is written as

$$\begin{aligned} w_{k+1} &= E\{xg(w_k^T x)\} - E\{g\prime(w_k^T x)\}w_k \\ w_{k+1} &= \frac{w_{k+1}}{\|w_{k+1}\|} \end{aligned} \tag{3.35}$$

Therefore, the algorithm for FastICA based on the 4maximum negentropy objective function is illustrated in the following steps [140, 75, 144, 145, 146]:

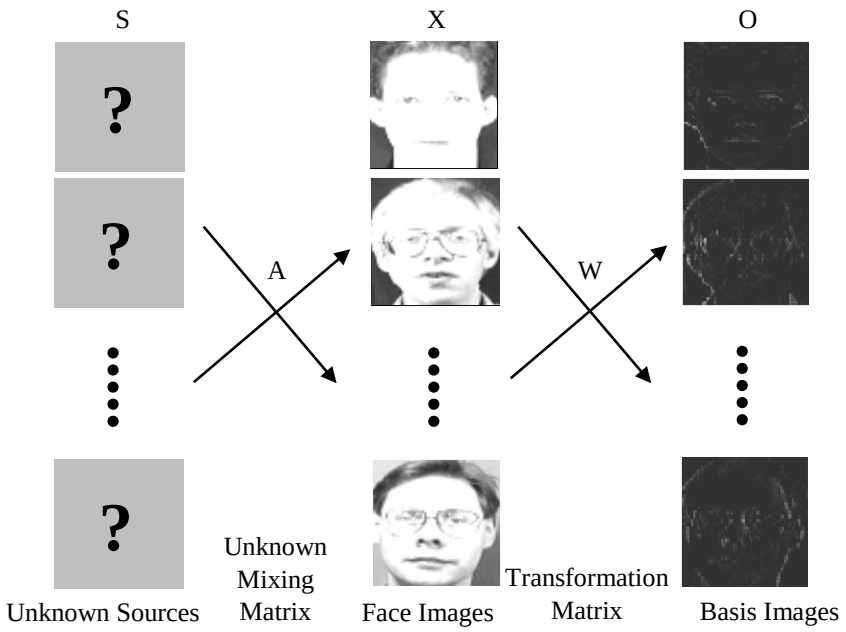


Figure 3.7: ICA model for Architecture I.

1. Initialize weight $w_k = 0$, where $k = 0$.

2. Update weight by $w_{k+1} = E\{xg(w_k^T x)\} - E\{g\prime(w_k^T x)\}w_k$.

3. Normalize the weighted vector as $w_{k+1} = \frac{w_{k+1}}{\|w_{k+1}\|}$.

4. If there is no convergence go to step 2.

5. If the convergence achieved, the independent component are obtained $y = s = W^T X$.

3.7.2 ICA-Architectures

There are two different ICA architectures employed for face recognition. In architecture I, which is used in this dissertation as shown in figure 3.7, the goal is to find a statistically independent basis images [64, 67].

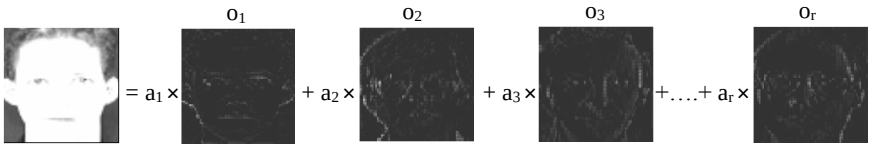


Figure 3.8: Statistically independent image basis employed with the coefficients a_i to reconstruct the input facial images.

In this architecture, the input facial images are organized in the rows and pixels are in the columns. The input facial images $(X = A.S)$ are considered a linear combination of unknown statistically independent components S by unknown mixing matrix A. Then the basis components (images) are found by projecting input facial images X onto transformation matrix W. As mention earlier, the goal of ICA is to find the transformation matrix W such that the rows of $O = W.X$ are as statistically independent as possible. Thereafter, the statistically independent source images estimated from the rows of O are considered as basis images that represent the input facial images [64, 67].

For input facial image reconstruction as shown in figure 3.8, the rows of O are used with the linear coefficients a_i. These coefficients are contained in the inverse of the transformation matrix $W^{-1} \triangleq A$ [64, 67].

In architecture II, the input facial images in X are arranged so that the images are in the columns and pixels are in the rows. Therefore, the ICA bases are arranged in the columns of $O = W.X$. The independent basis coefficients contained in the columns of O are used with the coefficients contained in W^{-1} to reconstruct the input facial images in X. As mentioned earlier the target of ICA is to make the facial basis images as closely independent as possible [64, 75, 67].

CHAPTER 4: SUPERVISED FACIAL RECOGNITION BASED ON MULTI-RESOLUTION ANALYSIS AND FEATURE ALIGNMENT

A new supervised algorithm based on the integration of Two-Dimensional Discrete Multiwavelet Transform (2-D DMWT), 2-D Radon Transform, and 2-D DWT is proposed in this chapter for face recognition [34]. In the feature extraction step, multiwavelet filter banks are used to extract useful information from the face images. The extracted information is then aligned using the Radon Transform, and localized into a single band using 2-D DWT for efficient sparse data representation. This information is fed into a neural network for training and testing. The proposed method is tested on three different databases, namely, ORL, YALE and subset fc of FERET, which comprise different poses and lighting conditions. It is shown that this approach can significantly improve the classification performance and the storage requirements of the overall recognition system.

4.1 Proposed System

The proposed algorithm for face recognition consists of 3 phases: preprocessing, feature extraction, and classification.

4.1.1 Preprocessing

The preprocessing phase consists of:

1. The first step aims to convert different image databases with different dimensions into a general common dimension (128x128). In this study, we consider three databases, with dimensions shown in Table 4.1.

Since the databases have different dimensions, and since the algorithm used in this chapter requires dimensions that are power of two, all the databases (images) are converted to a common dimension (128×128).

Table 4.1: The Dimensions of the Databases

Databases	ORL	YALE	FERET
Sizes	112×92	243×320	384×256

2. All images use a uint8 data type, which is not suitable for the transforms used in this algorithm. Therefore, the second step is to convert all images from ($X \times Y$ uint8) to (128×128 double).

4.1.2 Feature Extraction

To extract useful features from the images of the considered databases we apply 2-D DMWT, 2-D RT, and 2-D DWT. Mutliresolution analysis based on multiwavelets is used to extract the salient features, which are then aligned using the Radon Transform for efficient data representation. Since the filter banks in MWT are matrix-valued filters, the data needs to be preprocessed to obtain vector inputs for these filters [121]. Herein, we adopt a critically sampled scheme for preprocessing as in [121, 131, 147], suited for the 2-D image data as described next.

Algorithm for computing the 2-D DMWT coefficients using a critically sampled scheme for preprocessing:

1. Construction of the T matrix shown in Eq. 3.20 & Eq. 3.21.

2. Preprocessing the rows: we obtain a $N \times N$ matrix from a $N \times N$ input image. The odd rows of the resultant matrix correspond to the original matrix and the even rows are properly scaled versions of the input rows [121, 131, 147]. These serve as vector-valued inputs to the aforementioned GHM filter banks.

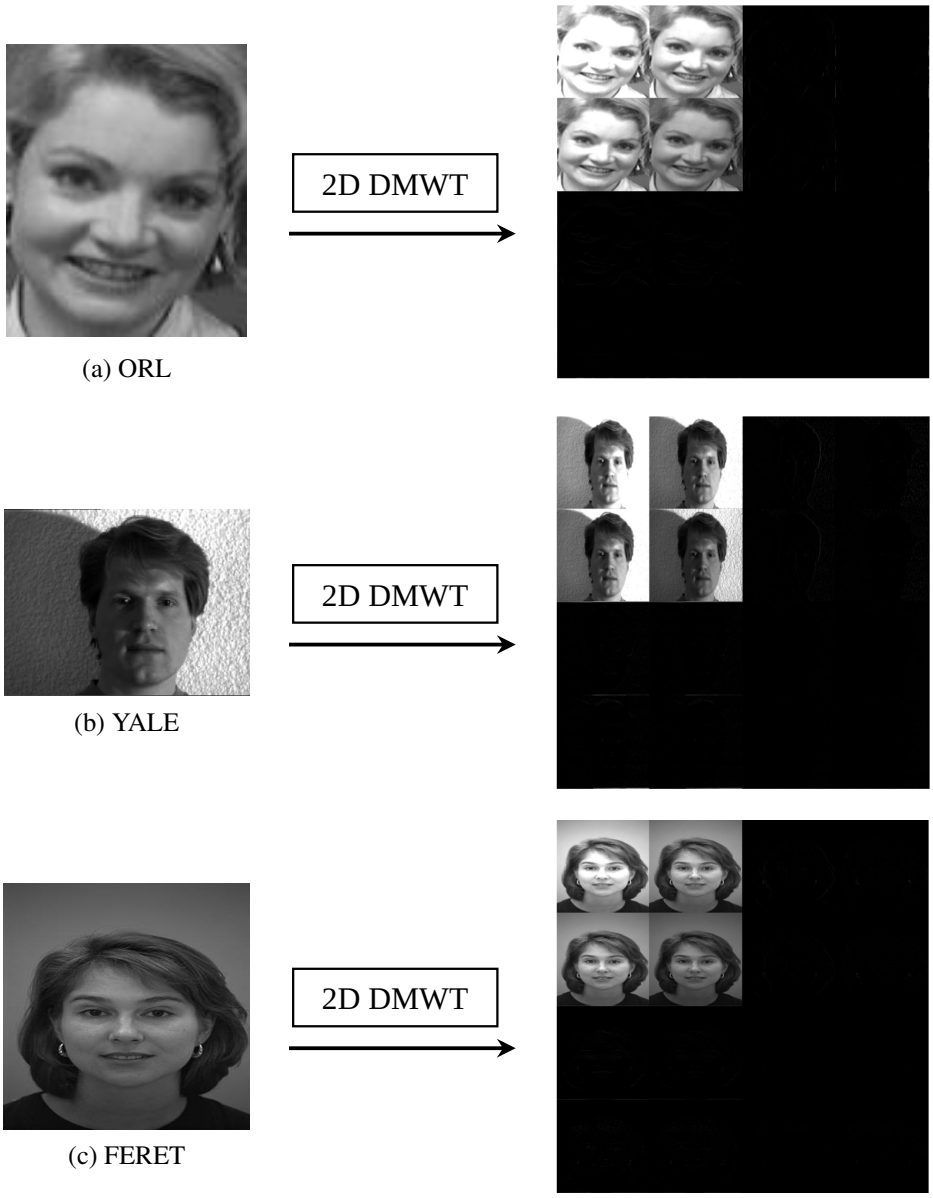
3. Transformation of the image rows: first, we apply the transformation T to the preprocessed matrix from step 2. Second, we permute the resulting matrix rows by rearranging the row pairs $1, 2$ & $5, 6, \ldots, N-3, N-2$ after each other in the upper half, then the row pairs $3, 4$ & $7, 8, \ldots, N-1, N$ below them in the lower half.

4. Preprocessing the Columns: the same preprocessing procedure is repeated here for the columns by transposing the matrix from step 3 and repeating step 2.

5. Transformation of the image Columns: we apply the transformation T to the resultant matrix from step 4, then permute the resulting matrix rows as in step 3.

6. The matrix from step 5 is transposed and we permute the coefficients for the resultant matrix. The obtained coefficients represent a one level decomposition of the 2-D DMWT.

An example of applying 2-D DMWT on the ORL, YALE, and FERET databases is shown in figure 4.1.



(a) ORL

(b) YALE

(c) FERET

Figure 4.1: Application of 2D DMWT on different databases.

In figure 4.1, the resultant image is sub divided into four sub bands and each sub band is further divided into four sub images as shown. We can see that the useful information exists only in the first subband, which contains all the energy from the original image. After extracting the features using 2-D DMWT, the Radon Transform is applied to the first subband. The final step of feature extraction is applying the 2-D DWT on each of the resultant sub-images to localize the information into a single band, which is the low frequency band as shown in figure 4.2.

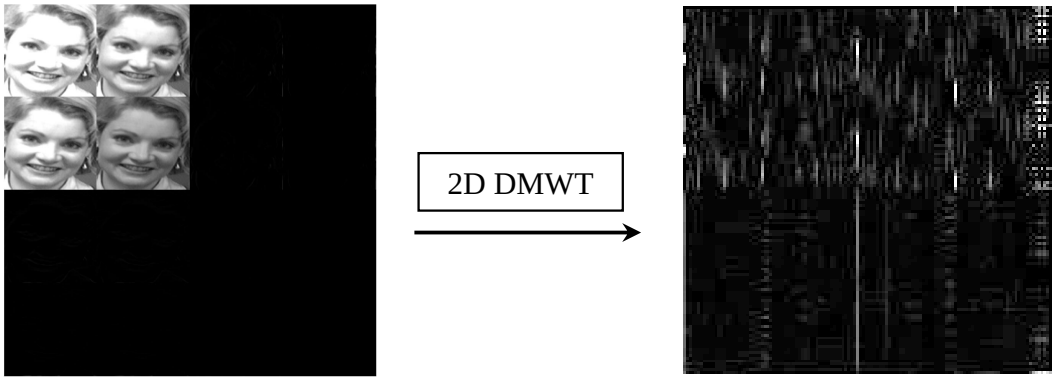
The final matrix is now fed into a Neural Network classifier for training and testing.

4.1.3 Classification

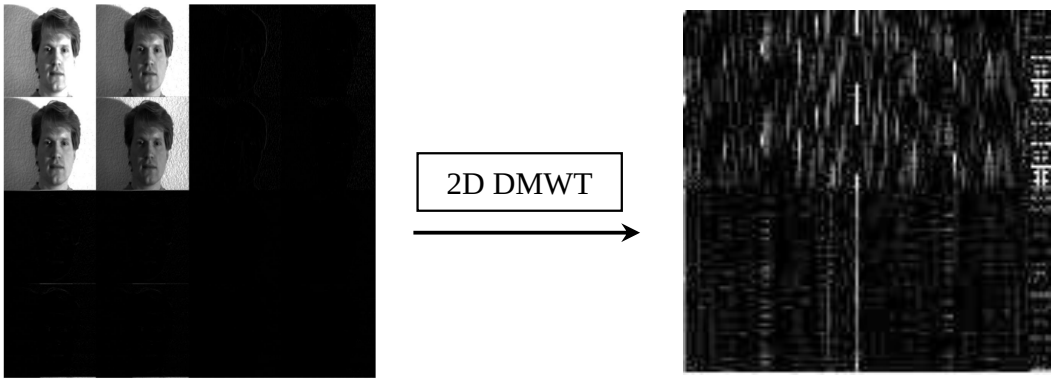
The algorithm we propose for training is based on the Back Propagation Training Algorithm for Neural Networks. The matrix of features is first converted to a 1-D form that is suitable for training the neural network. Since the proposed approach is supervised we need to choose a desired output vector for each database. There are 40 desired output vectors for ORL, 15 desired output vectors for YALE, and 200 desired output vectors for the subset fc of FERET database. The Neural Network consists of three layers, namely, an input layer, a hidden layer and an output layer.

4.2 Experimental Results

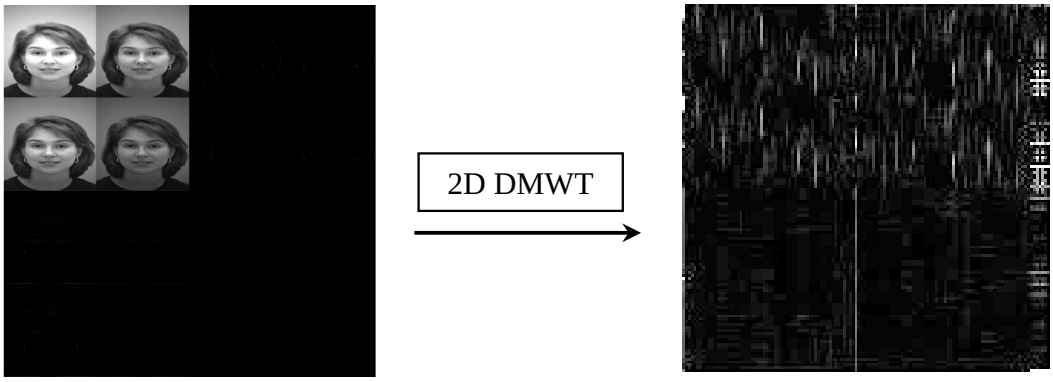
The results of the proposed algorithm are presented in this section. The combination of 2-D DMWT, 2-D RT, and 2-D DWT are presented and compared with the combination of 2-D DMWT and 2-D RT and the combination of 2-D DMWT and 2-D DWT. All the algorithms were tested using the ORL, YALE, and subset fc of FERET databases.



(a) ORL



(b) YALE



(c) FERET

Figure 4.2: Application of 2D RT on the resultant of 2D DMWT.

4.2.1 Experimental Results for the ORL Database

The ORL database consists of images for 40 persons, each with 10 different poses [28]. Let P be the number of poses that are used for training. Hence, 10-P poses are used for testing. In the training phase, we use P=1, P=3, and P=5 poses (of the 40 persons). Table 4.2 summarizes the performance of the three aforementioned algorithms. It is shown that the system that uses the 3 transforms significantly outperforms the other systems.

As shown, even when a small number of poses is used, the algorithm achieves good recognition accuracy. As we increase the size of the training data, the performance is improved. The proposed system also reduces the storage requirements. The size of the feature matrix is shown in the last row of Table 4.2.

Table 4.2: The Recognition Rate of the ORL Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode		
			2D DMWT and 2D RT	2D DMWT and 2D DWT	Proposed System
P=1	P=9	100%	67.78%	72.5%	90.56%
P=3	P=7	100%	75.35%	83.21%	96.07%
P=5	P=5	100%	83.5%	89.5%	99.5%
Size of the resultant matrix			64×64	32×32	32×32

4.2.2 Experimental Results for the YALE Database

The YALE database consists of images for 15 persons each with 11 different poses [29]. In the training phase, we use P=1, P=3, and P=5 poses. Table 4.3 summarizes the performance of the three aforementioned algorithms. Again, it is shown that the system that uses the 3 transforms significantly outperforms the other systems.

Table 4.3: The Recognition Rate of the YALE Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode		
			2D DMWT and 2D RT	2D DMWT and 2D DWT	Proposed System
P=1	P=10	100%	58.67%	60.67%	88.67%
P=3	P=8	100%	81.67%	85.83%	97.5%
P=5	P=6	100%	87.78%	92.22%	98.89%
Size of the resultant matrix			64×64	32×32	32×32

As before, the algorithm achieves good recognition accuracy even with a small number of poses and reduces the storage requirements. The size of the feature matrix is shown in the last row of Table 4.3.

4.2.3 Experimental Results for the FERET Database

The subset fc of FERET database consists of images for 200 persons each with 11 different poses [30, 31]. In the training phase, we use P=1, P=3, and P=5 poses. Table 4.4 summarizes the

performance of the three aforementioned algorithms. It is shown that the system that uses the three transforms significantly outperforms the other systems.

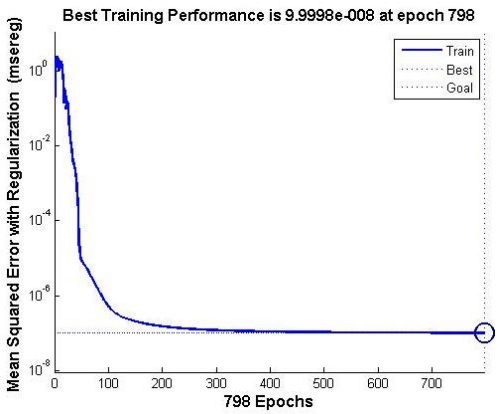
Table 4.4: The Recognition Rate of the FERET Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode		
			2D DMWT and 2D RT	2D DMWT and 2D DWT	Proposed System
P=1	P=10	100%	62.25%	64.4%	89.5%
P=3	P=8	100%	75.94%	80.56%	95.18%
P=5	P=6	100%	84.08%	90.75%	97.91%
Size of the resultant matrix			64×64	32×32	32×32

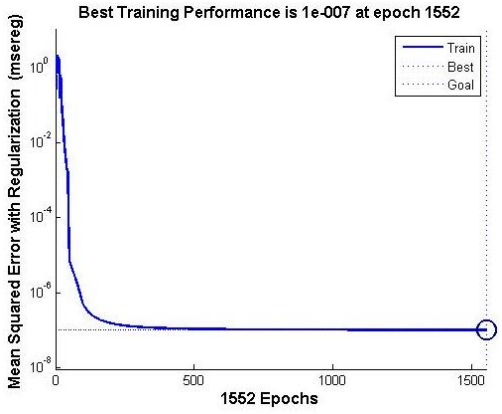
The results exhibit the same behavior as for the previous databases. Specifically, high recognition rates are achieved by the system that adopts the three transforms and the performance can be further improved with more training.

4.3 Discussion and Neural Network Performance

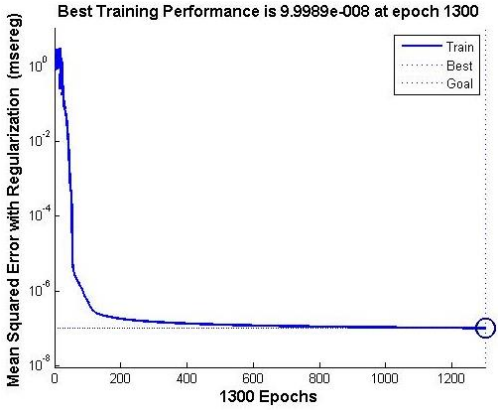
The performance of the neural network during the training phase can affect the overall performance of the proposed algorithm. Choosing the number of hidden layers, the number of neurons in the hidden layers, the type of the activation function, and the training function plays an important role on the overall performance of the system. In this study, we used one hidden layer with 1024 neurons for the ORL and YALE databases. For subset fc of FERET we used 2 hidden layers with 2096 and 1024 neurons for the first and the second layer, respectively.



(a) Proposed System



(b) 2D DMWT and 2D RT



(c) 2D DMWT and 2D DWT

Figure 4.3: MSE performance of the NNT for our proposed system compared with other methods.

The activation function used is the hyperbolic tangent sigmoid and Back Propagation is used for training. We plot the MSE versus the number of iterations for the three proposed algorithms during the training phase. An example of the performance of the Neural Network for the ORL database is shown in figure 4.3. It is shown that the proposed system that combines the three transforms achieves the fastest convergence and attains a MSE of 10^{-7}.

4.4 Conclusion

In this chapter, we proposed a new supervised classification algorithm that uses three consecutive transforms (2-D DMWT, 2-D RT, and 2-D DWT) and a neural network for facial recognition. Combining these transforms led to a significant improvement in the classification performance, as well as the storage requirement and the execution time of the recognition task. An extensive study of the performance of the proposed algorithm was conducted using three databases (ORL, YALE and subset fc of FERET), which have different poses and lighting conditions. The classification rates achieved were 99.5%, 98.667%, and 97.91%.

CHAPTER 5: SUPERVISED FACIAL RECOGNITION BASED ON MULTIRESOLUTION ANALYSIS WITH RADON TRANSFORM

In this chapter a supervised facial recognition approach based on the integration of Two Dimensional Discrete Multiwavelet Transform (2D DMWT), 2D Radon Transform (2D RT), and 3D DWT is proposed [35]. In the feature extraction step, 2D DMWT is used to extract the useful information from the image. The extracted features are then aligned using 2D RT and localized in one single band using 3D DWT. The resulting features are fed into a Neural Network for both training and testing. The proposed algorithm is tested on different databases, namely, ORL, YALE, and FERET. It is shown that the proposed approach can significantly improve the recognition rate and the storage requirements of the overall recognition system.

5.1 Proposed Approach

The proposed algorithm for facial recognition consists of three main steps, namely, Preprocessing, Feature Extraction, and Classification.

5.1.1 Preprocessing

The preprocessing step can be further divided into two steps:

1. The first step aims to convert different image databases with different dimensions (see Table 5.1) into one common dimension (128×128). This common dimension is chosen since the algorithm used in this chapter requires dimensions that are power of two.
2. All images use a uint8 datatype, which is not suitable for the transforms used in this algo-

rithm. Therefore, the second step is to convert all images from $(X \times Y$ uint8) to $(128 \times 128$ double).

Table 5.1: The Dimensions of the Databases

Databases	ORL	YALE	FERET
Sizes	112×92	243×320	384×256

5.1.2 Feature Extraction

The image of a human face has a lot of redundant information and is high-dimensional, hence we use the following to get an efficient compressive representation of the images and to extract reliable features:

1. The 2D DMWT based on Multiresolution Analysis (MRA) is used for:
   (a) Dimensionality reduction.
   (b) Localization of all the useful information in a single band.
   (c) Noise reduction.
2. The 2D RT is used to capture all the useful directional features of images and localize them around the origin.
3. The 3D DWT is also based on MRA and is used for:
   (a) Further dimensionality reduction.
   (b) Localization of directional features in a single band.

(c) Further noise reduction and enhanced smoothness.

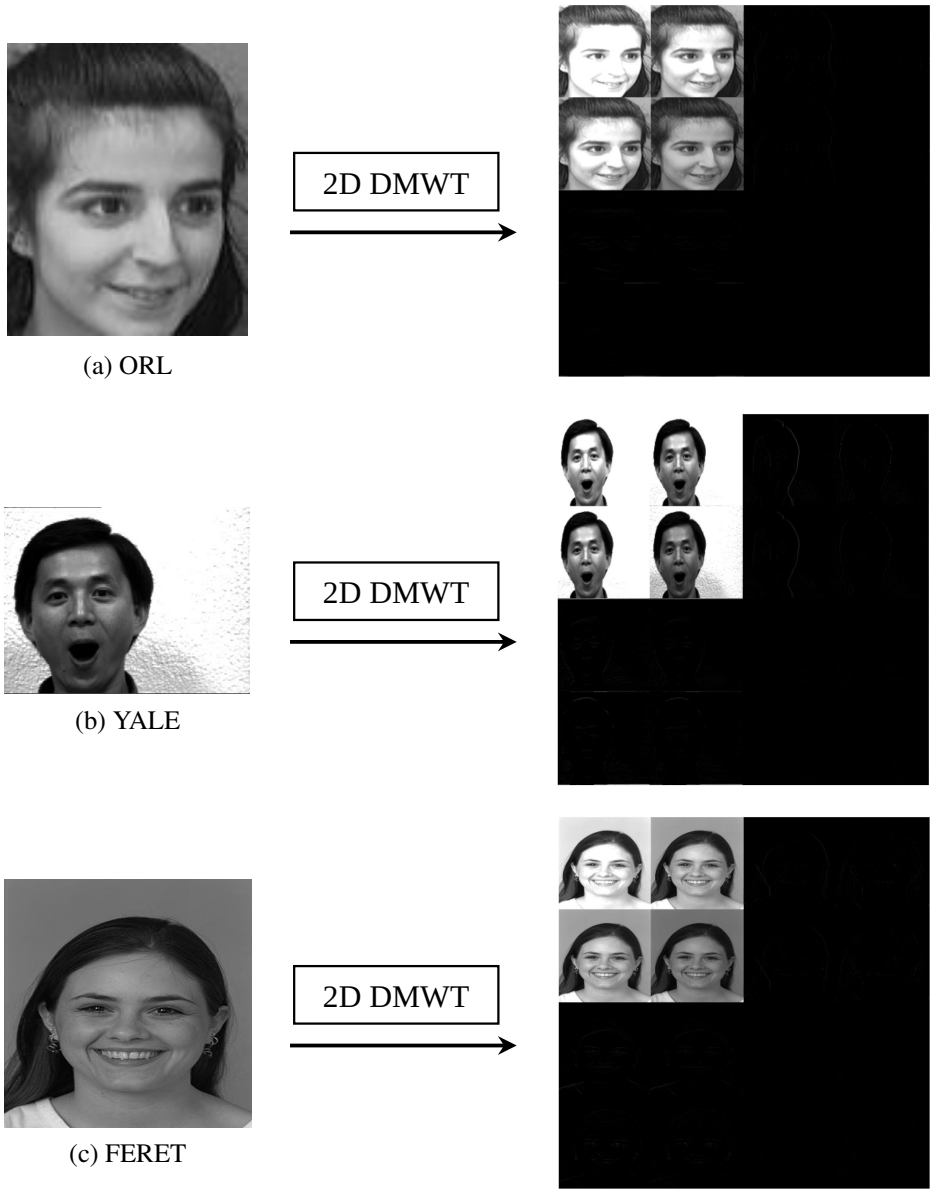
First, the 2D DMWT is applied to the images. As shown in figure 5.1, the image is subdivided into four subbands and each subband is further divided into four subimages.

From subfigures (5.1-a), (5.1-b), and (5.1-c), it is shown that all the useful information is localized in the upper left band, i.e., the LL band of the Multiwavelet transform. We retain the LL band and the other subbands are discarded. Hence, the dimension of the image matrix is reduced to 64×64. Each of the subimages in the LL subband is 32×32. The 2D RT is then applied to the first subband for feature alignment. In the final step, the 3D DWT is applied to the resultant matrix after rearranging the data in a 3D form as shown in figure 5.2.

Figure (5.2-a) shows the 3D arrangement of the LL subband of Figure 5.1-c *after applying* 2D RT on each of the subimages. Figure (5.2-b) represents the 3D arrangement of the subimages of the LL subband of figure 5.1-c. The 3D DWT is implemented on two steps:

1. We apply a 2D DWT on each 32×32 subimage of the LL subband of the 2D DMWT.
2. Then, we apply a 1D DWT on the depth of figure 5.2, which is of dimension 4×1.

The result of the feature extraction step is shown in figure 5.3. From the definition of the DWT, all the useful information will be localized in the LL frequency subband, and by eliminating the other subbands the resultant matrix is (16×16). Figure (5.3-A) is the result of applying 3D DWT to figure (5.2-a). Figure (5.3-B) is the result of applying 3D DWT to figure (5.2-b).



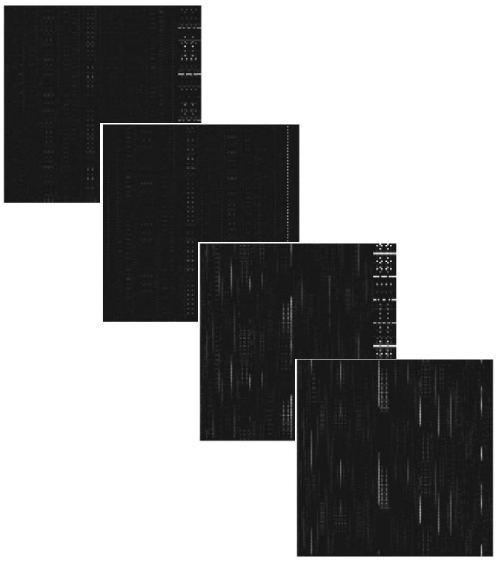
(a) ORL

(b) YALE

(c) FERET

Figure 5.1: Application of 2D DMWT on different databases.

(a)

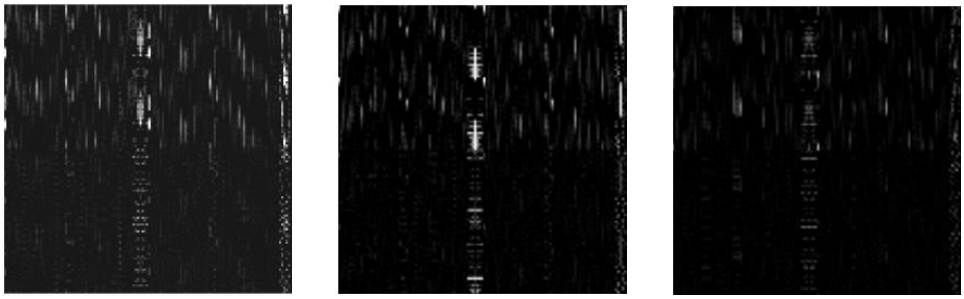


(b)

Figure 5.2: An example of 3D arrangement applied to figure 5.1-c.

ORL YALE FERET

(A)



ORL YALE FERET

(B)

Figure 5.3: Applying the 3D DWT to the result of figure (5.2).

5.1.3 Recognition

In the recognition step, we use a Neural Network (NNT) trained using the Back Propagation Training Algorithm (BPTA). The matrix of features is first converted to a 1D form that is suitable for training the NNT. BPTA is a supervised learning algorithm, hence we choose a desired output vector for each database. Since the used databases have different numbers of persons, there are 40, 15, and 200 different desired outputs for ORL, YALE and FERET, respectively. The NNT has three layers, namely, an input layer, a hidden layer, and an output layer.

5.2 Experimental Results

The results of the proposed algorithm are presented in this section. The proposed system will be compared with an algorithm that uses a combination of (2D DMWT and 3D DWT with NNT).

5.2.1 Experimental Results for the ORL database

ORL consists of 40 persons, each with 10 different poses [28]. Let P denote the number of poses used for training. Hence, $10 - P$ poses are used for testing. We use $P = 1$, $P = 3$, and $P = 5$ poses. Table 5.2 summarizes the results for the different algorithms.

As shown, even with a small number of poses for training, the proposed algorithm achieves a high recognition rate. As we increase the size of the training data, the performance is improved.

Table 5.2: The Recognition Rate of the ORL Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode	
			2D DMWT and 3D DWT	Proposed System
P=1	P=9	100%	76.11%	90.56%
P=3	P=7	100%	87.85%	95%
P=5	P=5	100%	94.5%	99.5%
Size of the resultant matrix			16×16	16×16

5.2.2 Experimental Results for the YALE database

The YALE database consists of 15 persons, each with 11 different poses [29]. Table 5.3 summarizes our results. Again, the proposed approach with the 3 transforms is shown to outperform the other recognition system.

5.2.3 Experimental Results for the FERET database

This database consists of 200 persons, each with 11 different poses [30, 31]. Table 5.4 summarizes the results for the aforementioned algorithms. The results exhibit the same behavior as in the previous databases.

Table 5.3: The Recognition Rate of the YALE Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode	
			2D DMWT and 3D DWT	Proposed System
P=1	P=10	100%	68.66%	87.33%
P=3	P=8	100%	85%	96.67%
P=5	P=6	100%	93.33%	98.89%
Size of the resultant matrix			16×16	16×16

Table 5.4: The Recognition Rate of the FERET Database

NO. of poses used in training mode	NO. of poses used in testing mode	Recognition rate for training mode	Recognition rate for testing mode	
			2D DMWT and 3D DWT	Proposed System
P=1	P=10	100%	69.45%	88.4%
P=3	P=8	100%	85.5%	95.25%
P=5	P=6	100%	93.25%	98.17%
Size of the resultant matrix			16×16	16×16

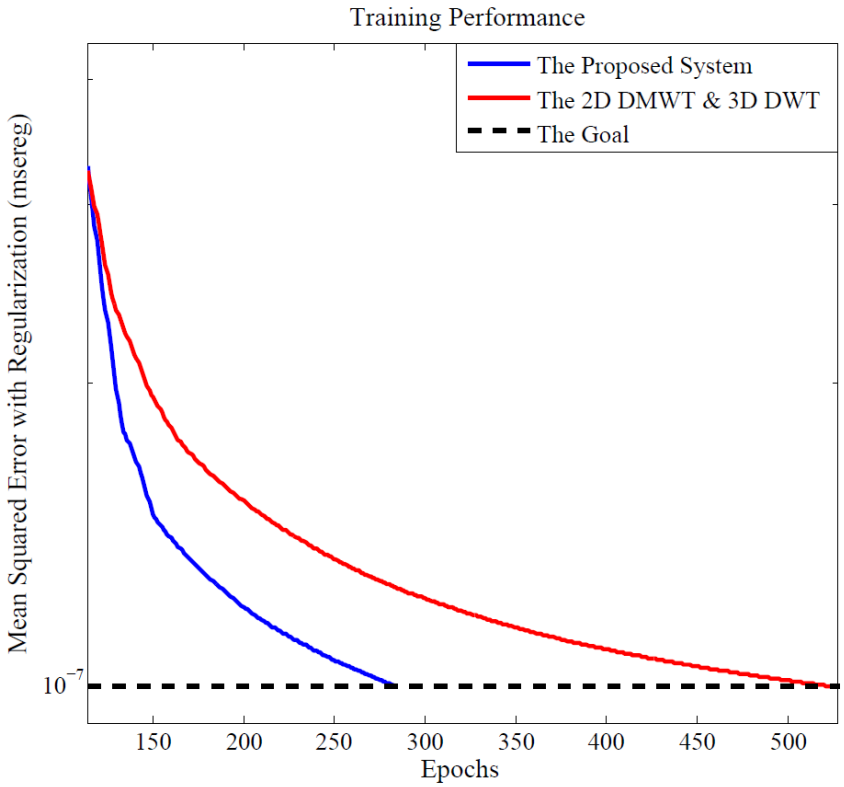


Figure 5.4: MSE performance of the NNT for the YALE database during the training phase. For both approaches, we used one hidden layer with 512 neurons. The activation function used is the hyperbolic tangent sigmoid, Back Propagation is used for training, and the goal is 10^{-7}.

5.3 Discussion and Neural Network Performance

The configuration of the NNT during the training phase can affect the overall performance of the proposed algorithm. Choosing the number of hidden layers, the number of neurons in the hidden layers, the types of the activation functions, the training function, the training algorithm, and the target performance can impact the overall performance of the system.

In this chapter, we used one hidden layer with 512 neurons for the ORL and YALE databases. For FERET we used 2 hidden layers with 1024 and 512 neurons for the first and the second hidden

layer, respectively. The activation function used is the hyperbolic tangent sigmoid and Back Propagation is used for training and testing. We plot the MSE versus the number of iterations for the two proposed algorithms during the training phase. Figure 5.4 illustrates an example of the performance of the NNT for the YALE database during the training phase. It is shown that the proposed system that combines three successive transforms achieves a faster convergence.

5.4 Conclusion

We proposed a supervised facial recognition algorithm consisting of three consecutive transforms (2D DMWT, 2D RT, and 3D DWT). The proposed approach was shown to improve the recognition rate, as well as the storage requirements and the computational complexity due to the extraction of efficient and compact features using successive compressive and aligning transforms. While the input images are 128×128, the feature matrices are of dimension 16×16. An extensive study of the performance of the proposed algorithm was conducted using three databases that have different lighting conditions, rotation angles and facial expressions. The recognition rates achieved are 99.5%, 98.89%, and 98.17% for ORL, YALE, and FERET databases, respectively.

CHAPTER 6: THREE SUPERVISED FACIAL RECOGNITION TECHNIQUES BASED ON FASTICA/DMWT

In this chapter, three supervised facial recognition techniques are proposed, based on the two-dimensional discrete multiwavelet transform (2D DMWT) and fast independent component analysis (2D FastICA), as described in [36, 37]. Each technique has three main steps, namely, preprocessing, feature extraction, and classification. Three preprocessing steps, namely, a cropping approach, choosing appropriate dimensions, and prefiltering of the images, are performed in each proposed technique. In the first step of the feature extraction, the 2D DMWT is applied to the facial images for the purpose of data compression (dimensionality reduction) and feature selection. Then, extracted features of 2D DMWT are processed using the 2D FastICA to obtain dominant features with enhanced discriminant and independent properties. Finally, the 2D DMWT features and the discriminant/independent features are further compacted by using the ℓ_2-norm. It is noted that the first and the second technique employ 2D DMWT/2D FastICA with ℓ_2-norm, while the third technique uses 2D DMWT with the ℓ_2-norm. Then the computed ℓ_2-norm features, techniques' features, are fed into a neural network (NN) classifier. The NN classifier employs a back propagation training algorithm (BPTA) for the recognition task. The three proposed techniques are evaluated using five different databases, namely, ORL, YALE, FERET, FEI, and LFW. These databases have different facial configurations, such as facial expressions, light conditions (illuminations), rotations, makeups, etc. The experimental results of the proposed system are analyzed using K-folds cross validation (CV). The results confirm the improvements in the recognition rates as well as the storage requirements as compared to some recently reported methods.

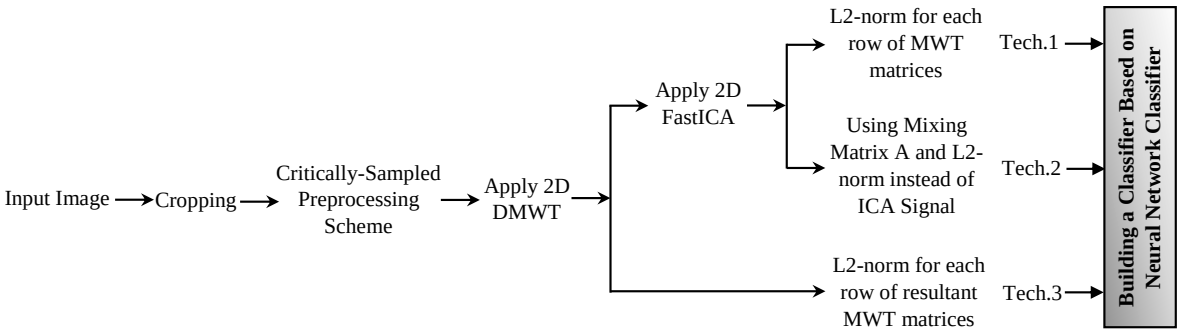


Figure 6.1: The Three Proposed Techniques

6.1 Proposed Techniques

Each of the three proposed face recognition techniques consists of three main steps, namely, preprocessing, feature selection, and classification, as shown in figure 6.1. Cropping approach, choosing appropriate dimensions, and prefiltering are done in the preprocessing step. The second step aims to find a better facial representation with less storage requirements. Finally, neural network classifier (NN) is used in the classification step. The proposed techniques are evaluated using five databases, namely, ORL, YALE, FERET, FEI, and LFW that have different facial variations, such as illuminations, rotations, facial expressions, makeups, etc. Samples of the databases are shown in figure 6.2-A.

6.1.1 Preprocessing

It consists of three steps:

1. The first step aims to find the dimensions of the face images. Then, a cropping approach, which is manually performed using MATLAB 2016a, is applied to all poses in the databases, as shown in figure 6.2-B. Table 6.1 shows the dimensions of all databases after cropping.

2. The second step aims to convert the dimensions of the cropped databases shown in Table 6.1 to appropriate ones. Power of two dimensions are chosen due to DMWT prefiltering requirements.

Table 6.1: The Dimensions of The Databases

Databases	Actual Size	Dimensions After Cropping	Proposed Dimensions
ORL	112×92	112×92	128×128
YALE	243×320	180×150	128×128
FERET	384×256	220×170	128×128
FEI	480×640	320×290	256×256
LFW	250×250	64×64	64×64

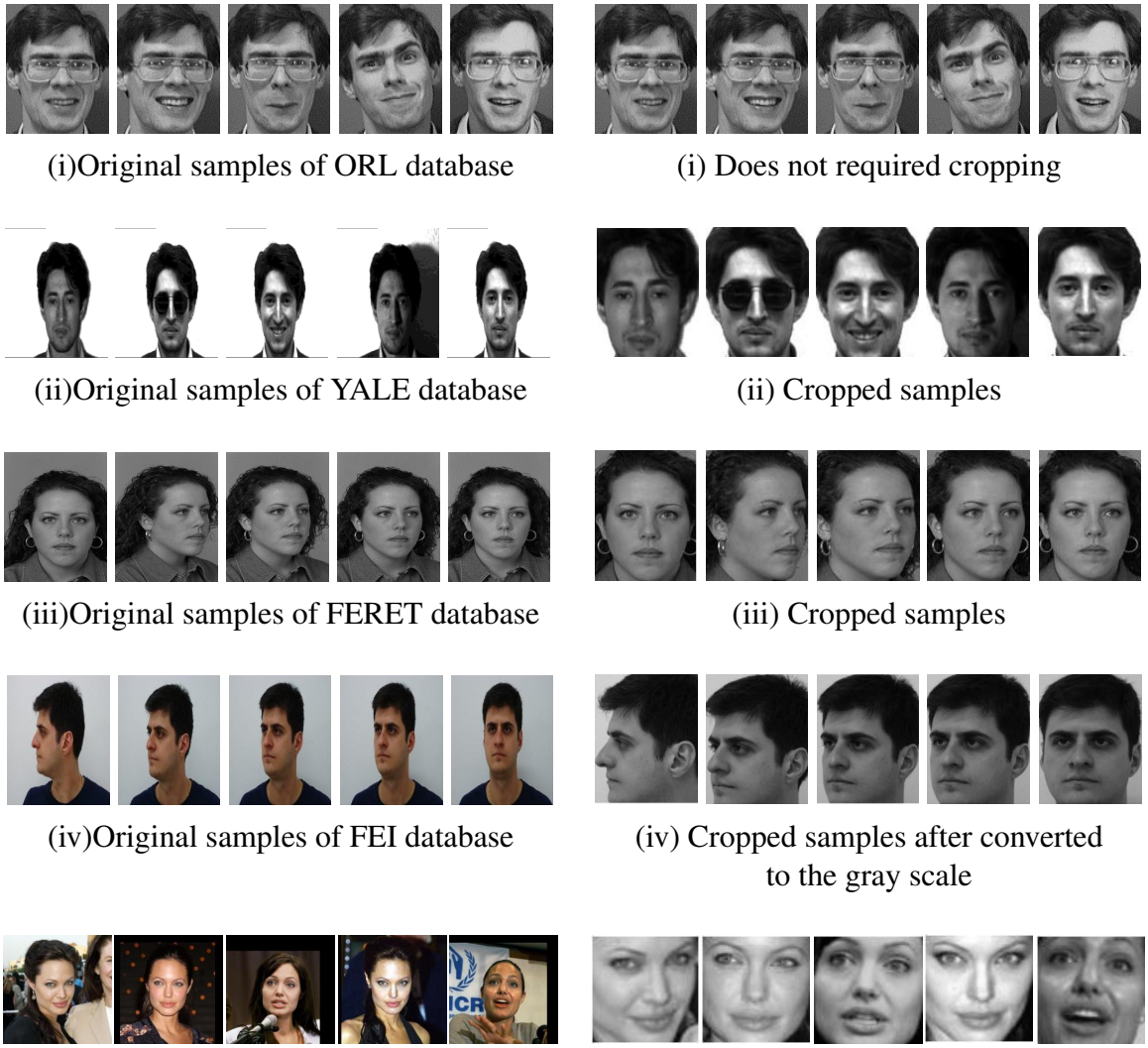
3. The critically-sampled preprocessing scheme, approximation based preprocessing, is applied to the processed facial images ([121, 122, 125]).

6.1.2 Feature Extraction (Selection)

Human faces have large dimensions that increase the computational complexity and hence affect the overall performance of the recognition system. The feature extraction step is the key part of the facial recognition system. A facial image has highly irrelevant information that is not necessary for the classification task. Therefore, the following efficient tools are applied to obtain features that are more discriminant and less dependent:

1. The 2D DMWT based on Multiresolution Analysis (MRA) is used for:

   (a) Dimensionality reduction.



(i)Original samples of ORL database

(i) Does not required cropping

(ii)Original samples of YALE database

(ii) Cropped samples

(iii)Original samples of FERET database

(iii) Cropped samples

(iv)Original samples of FEI database

(iv) Cropped samples after converted to the gray scale

(iv)Original samples of LFW database

(iv) Cropped samples after converted to the gray scale

(A) Original samples

(B) Cropped Samples

Figure 6.2: Figure 6.2-A shows 5 original samples of one person for all different databases. Figure 6.2-B shows samples after cropping.

(a) ORL (b) YALE (c) FERET (d) FEI (e) LFW

Figure 6.3: An example of applying 2-D DMWT to the different databases.

(b) Noise alleviation.

(c) Localization of most efficient information in one single sub-band based on low-low (LL) frequency band.

2. The FastICA is used for:

   (a) Decorrelating the high order statistics in addition to decorrelating the second order moments. Most of the efficient information about the local characteristics of facial images is contained in the high order statistics. Hence, ICA basis is used for best representing the input facial images ([65, 64, 144, 67, 68, 66, 148]).

   (b) Improving the convergence rate and reducing the computational complexity ([64, 144])

Then the resultant ICA features are:

a) Independent, leading to an efficient representation and hence better identification and classification rates.

b) Less sensitive to the facial variations arising from different facial expressions, illuminations, rotations, different poses, etc. ([75, 64, 144, 67, 68]).

As shown in figure 6.1, the 2D DMWT is applied to the cropped processed databases. Figure 6.3 shows an example of applying 2D DMWT into facial images of five different databases. It is obvious from figure 6.3 that the facial image is divided into four main sub-bands with dimensions of 64×64 for ORL, YALE, and FERET databases, 128×128 for FEI database, and 32×32 for LFW database. Each one of these sub-bands are further divided into four sub-images with dimensions of 32×32 for ORL, YALE, and FERET databases, 64×64 for FEI database, and 16×16 for LFW database. From sub-figures (6.3-a), (6.3-b), (6.3-c), (6.3-d), and (6.3-e), most of the facial information features center in one single sub-band, related to the low-low (LL) frequency sub-band of the MWT. The LL sub-band is maintained and all other sub-bands are eliminated. Hence, the resultant matrix has 64×64 dimensions for ORL, YALE, and FERET databases, 128×128 dimensions for FEI database, and 32×32 for LFW database. For each sub-image in the LL sub-band, the following procedures are executed to obtain an efficient feature matrix:

1) Convert each resultant sub-image of LL sub-band of DMWT into a vector.

2) Apply (1) to all sub-images of each pose. Therefore, the resultant feature matrix of each pose has 1024×4, 1024×4, 1024×4, 4096×4, 256×4 dimensions for ORL, YALE, FERET, FEI, and LFW databases, respectively.

Extracting new features from the original one further reduces the dimensions of the extracted features, gets efficient representation of the images, achieves good performance, and then obtains better results. These are accomplished by applying the following three techniques:

A. *Technique 1*

   i. Apply FastICA to the extracted matrix resulted from ORL, YALE, FERET, FEI, and LFW databases as shown in figure 6.4. The resulting features are more efficient, discriminant, and independent.

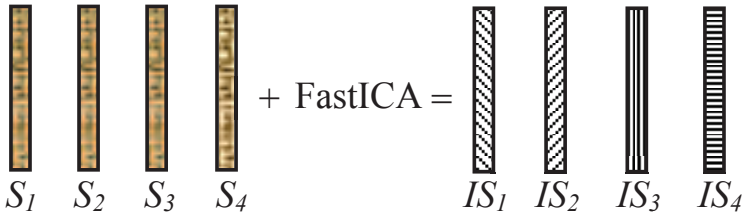


Figure 6.4: Applying 2D FastICA into extracted subimages. S_x represents Subimages of LL sub-band. IS_x represents Independent Subimages of LL sub-band where $x = 1, 2, 3, and\ 4$.

ii. Apply ℓ_2-norm for each row of the resulting matrix to concentrate all the energy in one single column. As a result, the resulting feature matrix for each pose has 1024×1 dimensions for ORL, YALE, and FERET databases, 4096×1 dimensions for FEI database, and 256×1 dimensions for LFW database.

B. *Technique 2*: It is the same as Tech.1. Herein, the mixing matrix A, which appeared in Eq. (3.26), is used instead of using the FastICA feature matrix. As mentioned, it is considered as an alternative way to represent the input facial images ([64, 144, 67, 68]).

C. *Technique 3*: It is the same as Tech.1 without applying Tech.$1 - i$.

3) Repeat steps [1-2] for each pose in the five databases. Each person has a resultant extracted features with $X^* \times Y^*$ dimensions. Where X^* is 1024 for ORL, YALE and FERET databases, 4096 for FEI database, and 64 for LFW. $Y*$ depends on number of poses used in the training mode of each person for each database.

Finally, the resultant features are fed to the classification step to build a neural network (NN) based classifier.

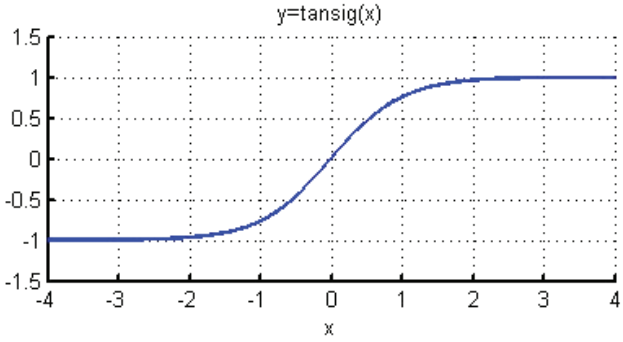


Figure 6.5: Hyperbolic tangent sigmoid transfer function. The activation function tansig(x) is expressed as $tansig(x) = \frac{2}{1+e^{-2x}} - 1$ and it can be considered as $tansh(x)$. Its output range between (-1,1).

6.1.3 Recognition

A neural network, based on a back propagation training algorithm (BPTA), is used in this chapter for the recognition task. BPTA is used in both training and testing modes. Since the BPTA requires supervised learning, it is necessary to choose the desired output for each person in the database. Five different databases, each with a different number of persons, are used to evaluate the three proposed techniques. Each person has a desired output different from the desired outputs of other persons. Since there are 15 persons in the YALE database, there are 15 neural network outputs, one for each person. All poses of each person have the same desired output. For example, the desired output of the first person of the YALE database is $[1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1]$, where number (1) means the output is active, and the test facial image is correctly matched. The number (-1) means the output is inactive, and the test facial image is not identified correctly. Note that the output might not be equal to 1 or -1 but may reach these values since the activation function used in the BPTA is hyperbolic tangent sigmoid transfer function that is shown in figure 6.5 ([149]).

The structure of the NN in the training mode has three layers, namely, an input layer, a hidden layer, and an output layer. The classifier in the training mode is configured using the resultant features from the feature extraction step. For computing accuracy, the recognition rates are measured as $\frac{M}{T} \times 100\ \%$, where M is the total number of poses correctly matched and T is the total number of poses in the database.

The same steps as in the training modes are followed in the testing mode. First, the cropping method, re-sizing, and prefiltering are applied. Then, the 2D DMWT is applied to the processed facial image. As before, the only LL sub-band is retained and all the other sub-bands are eliminated. Then, each sub-image in the LL sub-band is converted to 1D form. Thereafter, *techniques 1-3* are applied to the 1024×4 feature matrix, obtained by arranging the four 1D sub-images of LL sub-band of ORL, YALE, and FERET databases in a matrix form. The three proposed techniques are also applied to the 4096×4 and 256×4 feature matrix of FEI and LFW databases, respectively that resulted and arranged in the same manner. Then, the ℓ_2-norm is applied for further data compaction. The feature vector for each technique of each test pose of each person has dimensions of 1024×1 for ORL, YALE, and FERET databases, 4096×1 for FEI database, and 256×1 for LFW database. Finally, the tested feature vector is fed into BPTA to obtain the matching results.

6.2 Experimental Results

The results of the three proposed techniques are discussed and compared with the other approaches . Databases discussed in sub-subsections 1.3.1-1.3.5 are used to evaluate the techniques. To analyze the results, K-folds cross validation is used. Different values of K are chosen, $K = 2, K = 3, and\ K = 5$. The rates reported in Tables 6.2-6.6 are the average of the rates obtained across the K-folds CV. Simulations of the proposed techniques are presented and compared with the existing approaches. All techniques achieved 100% accuracies when tested with the training poses.

6.2.1 Experimental Results for the ORL Database

In the ORL database, there are 40 people, each with 10 different facial poses, as mentioned in sub-subsection 1.3.1. The results of the proposed techniques and the other approaches are summarized in the Table 6.2.

Table 6.2: Experimental Results for the ORL Database

Method	**Recognition rates**
Proposed Technique 1	**98.5%**
Proposed Technique 2	**98.25**%
Proposed Technique 3	95.5%
RV-LDA [55]	97%
OLPP [62]	93.5%
LPP [63]	95.83%
DWT-PCA [81]	96%
DWT-ICA [83]	96.75%
SNMFSPM [78]	98.15%
DCT [91]	94.25%
LBPP-DCT [93]	95.5%
GWT-DCT [94]	96.5%
DCT-PCA [95]	95.75%
SADL [102]	97.5%

6.2.2 Experimental Results for the YALE Database

There are 15 persons in the YALE database, each with 11 different poses, as mentioned in sub-subsection 1.3.2. Table 6.3 summarizes the results of the three proposed techniques in comparison with the state-of-the-art approaches.

Table 6.3: Experimental Results for the YALE Database

Method	**Recognition rates**
Proposed Technique 1	**98.33%**
Proposed Technique 2	**98%**
Proposed Technique 3	95.5%
OLPP ([62])	98.2%
LPP ([63])	97.14%
DWT-PCA ([81])	95.7%
DWT-ICA ([83])	96.25%
SNMFSPM ([78])	97.69%
GWT-DCT ([94])	96%
DCT-PCA ([95])	95.5%
SADL ([102])	94.67%
NFLS-II ([150])	82.42%

6.2.3 Experimental Results for the FERET Database

There are 200 different persons, each with 11 different poses in the FERET database, as mentioned in sub-subsection 1.3.3. The results of the proposed techniques and other approaches are summarized in Table 6.4

Table 6.4: Experimental Results for the FERET Database

Method	**Recognition rates**
Proposed Technique 1	**98.25%**
Proposed Technique 2	**97.89%**
Proposed Technique 3	95.51%
OLPP ([62])	92.1%
DWT-PCA ([81])	96%
DWT-ICA ([83])	96.67%
GWT-DCT ([94])	96.5%
DCT-PCA ([95])	95.73%
LDA-SID ([100])	96.1%
SADL ([102])	92.78%
SLF-RKR-ℓ_2 ([109])	89.2%
LFLM-SIFT ([111])	92.4%
SLBFLE ([112])	97.7%

6.2.4 Experimental Results for the FEI Database

There are 200 persons in the FEI database, each with 14 different poses, as mentioned in sub-subsection 1.3.4. Table 6.5 summarizes the results of the three proposed techniques.

Table 6.5: Experimental Results for the FEI Database

Method	**Recognition rates**
Proposed Technique 1	**98.4%**
Proposed Technique 2	**98%**
Proposed Technique 3	95.6%
OLPP ([62])	98.9%
DWT-PCA ([81])	96.4%
DWT-ICA ([83])	96.7%
GWT-DCT ([94])	96.51%
DCT-PCA ([95])	95.9%
LFLM-SIFT ([111])	85.3%
NFLS-II ([150])	93%
PCNC ([151])	94.17%

6.2.5 Experimental Results for the LFW Database

Labeled faces in the wild (LFW), which contains 13233 images of 5749 persons, is a large database of photographs designed for unconstrained face recognition, see sub-subsection 1.3.5. Table 6.6 summarizes the results of all techniques proposed compared with the other methods.

Table 6.6: Experimental Results for the LFW Database

Method	**Recognition rates**
Proposed Technique 1	**97.91%**
Proposed Technique 2	**97.23%**
Proposed Technique 3	93.67%
DWT-PCA ([81])	94.34%
DWT-ICA ([83])	95%
GWT-DCT ([94])	94.7%
DCT-PCA ([95])	93.9%
CNN-SAE ([99])	98.24%
LDA-SID ([100])	95.65%
SLF-RKR-ℓ_2 ([109])	81.9%
SLBFLE ([112])	84.2%

6.3 Discussion

The discussion of the results of the three proposed techniques is presented in this section. In contrast to the DWT, the DMWT has several favorable features, such as good reconstruction (orthogonality), better performance (linear phase symmetry), high order of approximation (vanishing moments), and compact support, as mentioned in Section 2. These desirable features cannot be achieved simultaneously in the scalar wavelet, while in MWT provide more degree of freedom and give perfect performance in signal and image applications as explained in [121]. Based on a number of decomposition levels required and the resulted sub-bands, which mathematically expressed as $3 \times L^* + 1$ for DWT and $4 + 12 \times L^*$ for DMWT, DMWT achieves higher dimensionality reduction than DWT. Where L^* is the number of decomposition levels required. Therefore, all techniques, in the manner proposed, accomplished less storage requirements compared with storage requirements reported in [81, 82, 83] that used DWT. Similarly, the techniques proposed in this chapter achieved less storage requirements compared to some of the recently reported approaches.

The objective of using FastICA is due to its favorable properties compared with the traditional ICA. FastICA has faster convergence than traditional ICA, since FastICA does not require a step size in contrast to ICA based gradient algorithm. Also, FastICA finds the independent non-Gaussian signal by using any arbitrary non-linear function $\theta(t)$, while other ICA algorithms require evaluation of the PDF of the selected non-linear function ([75, 144]). As a consequence, the FastICA efficiently estimates the statistical components. Therefore, the goal of using FastICA in this chapter is to find a basis of facial images that are statistically independent or as independent as possible for better facial representation. This is due to the fact that the basis images of FastICA maintain the structural information of the facial images, which are presented in the illumination, facial expressions, and rotation. According to [75, 64, 144, 67, 68], ICA is more robust to the variations in the facial expressions, rotations, and light conditions.

As shown in Tables 6.2-6.6 above, combining these two described tools (2D DMWT and 2D FastICA) achieved high recognition rates. *Techniques 1 & 2* led to a notable improvement in the recognition rates, storage requirements, as well as computational complexity in comparison with the traditional ICA implementations and almost all the other approaches. As shown in the experimental results above, the proposed techniques using both tools (2D DMWT and 2D FastICA) outperformed *technique 3*, which does not use the FastICA step. Furthermore, *techniques 1 & 2* show an improvement in the recognition rates in comparison with almost all the other approaches. For example, [83] applied DWT/ICA to the FEI facial images. In [83], 128×128 (16384×1) features represented each pose. The recognition rate achieved by [83] was 96.7%. The proposed *techniques 1 & 2* achieved 98.4% and 98%, respectively while using 64×64 features to represent each pose. [62] applied OLLP method to the FEI facial images. Although the recognition rate achieved by [62] was slightly higher by 0.5% than the rates achieved by our *techniques 1 & 2*, the storage requirements accomplished by *techniques 1 & 2* were less by 33.33% than the storage requirements achieved by [62]. Also, comparing the results reported for ORL, YALE, and FERET databases shown in Tables 6.2-6.4, *techniques 1 & 2* achieved higher recognition rates than those rates reported by [83] and [62]. Similarly, our *techniques 1 & 2* outperformed almost all other approaches in term of recognition rates, as shown in Tables 6.2-6.6.

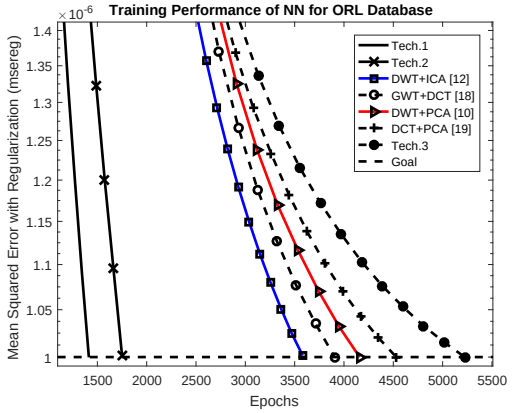
In almost all the simulations, *techniques 1 & 2* proposed in this chapter produced higher or comparable recognition accuracy, as shown in Tables 6.2-6.6. The OLPP, as shown in Table 6.5, and CNN-SAE, as shown in Table 6.6, have slightly higher recognition rates compared with the first two proposed techniques. It is to be noted that the test results for OLLP and CNN-SAE were obtained under different conditions. The OLPP results reported in [62] and given in Table 6.5 were obtained using 10 poses out of the 14 poses in the FEI database. Also, in [62], 9 poses out of the 10 were used for training, and one pose was used for testing. Our test results in Table 6.5 for *techniques 1 & 2* were obtained using all 14 poses in the FEI database. K-folds CV was used to

analyze the results, and three values of K (K=2, K=3, K=5) were chosen, which on average used fewer training images compared with the OLPP results. The recognition rates reported in the tables above were the average of the rates obtained across the K-folds CV. In addition, our techniques have lower computational complexity and storage requirements.

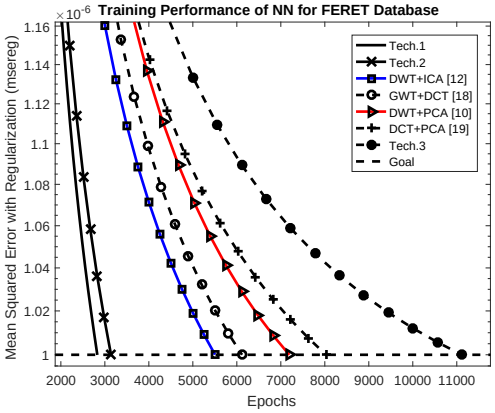
Similarly, the CNN-SAE ([99]) has higher computational complexity and storage requirements compared with *techniques 1 & 2*. In like manner, our proposed *techniques 1 & 2* outperformed the other reported methods in terms of the storage requirements, the computational complexity, and the recognition rates.

6.4 Neural Network Performance

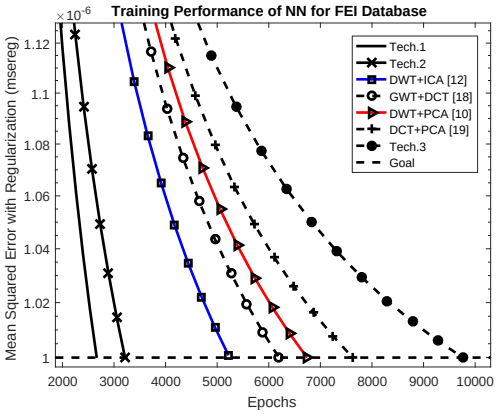
The NN configuration during the training mode has a direct effect on the performance of the proposed techniques and hence affects the recognition rates. There are three main layers in the structure of the NN, namely, input, hidden, and output layers. The number of neurons in the input and output layers is known and fixed based on the input data dimensions or the number of persons in each database, while the number of neurons in the hidden layers is somehow flexible. There are several factors that have a direct effect on the overall performance of the recognition system, such as number of hidden layers, number of neurons in the hidden layers, type of activation functions, training algorithms, and the target performance. In this chapter, one hidden layer with 512 neurons was chosen for ORL and YALE databases. For FERET and FEI databases, two hidden layers were selected with 1024 and 512 neurons for the first and second layers, respectively. For LFW database, two hidden layers were chosen with 512 and 256 neurons for the first and second hidden layers, respectively. The hyperbolic tangent sigmoid transfer function ($tansig(x)$) is used for all layers. The back propagation training algorithm (BPTA) is used for training and testing.



(a) ORL



(b) FERET



(c) FEI

Figure 6.6: The NN performance for the three proposed techniques compared with the other approaches.

Samples of the NN performance using the three proposed techniques in the training mode in comparison with the some of the state-of-the-arts methods are shown in figure 6.6. As shown in figure 6.6, the proposed *techniques 1 & 2* achieved the ***goal*** faster than *technique 3* and the other approaches reported in [81, 83, 94, 95] due to the efficient facial representations resulting from combining the DMWT and FastICA tools.

6.5 Conclusion

Three supervised facial recognition techniques based on the integrated tools of 2D DMWT and FastICA were presented in this chapter. Low storage requirements and high recognition rates were achieved in this contribution. *Techniques 1 & 2* applied 2D DMWT/FastICA with the ℓ_2-norm to the facial images, while *technique 3* employed only the 2D DMWT with the ℓ_2-norm to the facial images. In this chapter, DMWT/FastICA were employed to reduce the dimensionality of the data and produce more discriminant and independent facial features. The ℓ_2-norm was applied to the extracted features for further feature compression. In the classification step, a NN based on BPTA was used during the training and testing modes.

The three proposed techniques were extensively evaluated using five widely used databases, namely, ORL, YALE, FERET, FEI, and LFW. These databases have different facial variations, such as light condition, rotations, facial expressions, makeups, etc. The experimental results were analyzed using K-folds CV. Different values of K-folds, $K = 2, 3, \& \ 5$, were selected. The recognition rates reported in Tables 6.2-6.6 were the average of the rates obtained across the K-folds CV.

The proposed *techniques 1 & 2* outperformed *technique 3* and almost all the other methods, as shown in Tables 6.2-6.6. Also, the NN, employing *techniques 1 & 2*, converged faster than *technique 3* and the other approaches ([81, 83, 94, 95]), as shown in figure 6.6. *Technique 3* was presented to illustrate the performance improvement of *techniques 1 & 2* relative to *technique 3* due to incorporating the FastICA in *techniques 1 & 2*.

Then, the 2D FastICA was employed on the DMWT features to obtain an efficient, discriminating, and less dependent features and also to reduce the computational complexity in comparison with traditional ICA algorithms. Then, the ℓ_2-Norm was applied into each row of the extracted features to further reduce the storage requirements and to concentrate all the efficient features in

one single column, which led to the data compaction. The Neural Network (NN), based on using Back Propagation Training Algorithm, was contacted during both training and testing modes. As discussed in the experimental results, the proposed *Tech.1* and *Tech.2* outperformed *Tech.3* and the state-of-the-arts approaches [152, 153, 65, 91]. Also, the NN convergence of *Tech.1* and *Tech.2* was faster at achieving the goal as shown in figure 6.6 than *Tech.3*, [152], [65].

CHAPTER 7: SUPERVISED FACIAL RECOGNITION BASED ON EIGENANALYSIS OF MULTIRESOLUTION AND INDEPENDENT FEATURES

A supervised facial recognition system using Two Dimensional Multiwavelet Transform (2D DMWT), Two Dimensional Fast Independent Component Analysis (2D FastICA) is presented in this chapter [38]. In the feature extraction step, the 2D DMWT is used to extract useful information from the face images. The 2D DMWT is followed by a 2D FastICA and eigendecomposition to obtain discriminating and independent features. The resulting compressed features are fed into a Neural Network (NNT) based classifier for training and testing. All techniques are tested using ORL, YALE, and FERET databases. The proposed approach shows a significant improvement in the recognition rate, storage requirements, as well as computational complexity.

7.1 Proposed Techniques

The proposed facial recognition system consists of three phases: Preprocessing, Feature Extraction, and Recognition. The proposed techniques are shown in figure 7.1.

7.1.1 Preprocessing

The preprocessing phase consists of two steps:

1) The first step aims to convert the different image dimensions (see Table 12.1) into one common dimension (128×128) since the algorithm used in this chapter requires dimensions that are power of two.

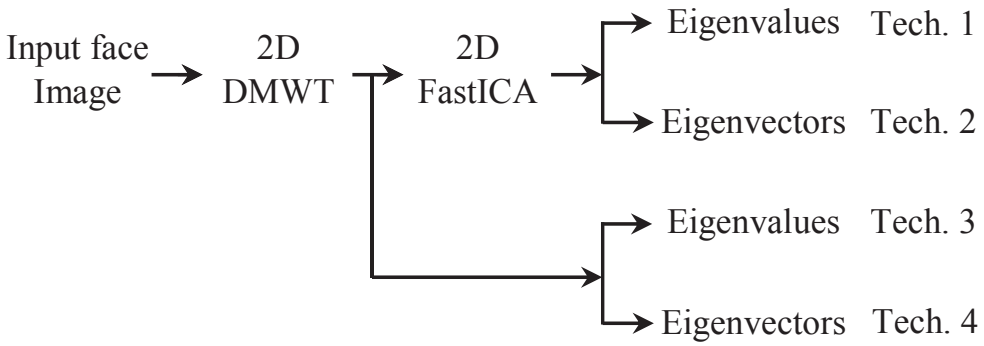


Figure 7.1: The Proposed Techniques.

Table 7.1: The Dimensions of the Databases

Databases	ORL	YALE	FERET	The Proposed
Size	112×92	243×320	384×256	128×128

2) The second step aims to convert all the images to $(128 \times 128\ double)$ extension instead of the $unit8$ datatype, which is not suitable for the transform used in this chapter.

7.1.2 Feature Extraction

A facial image has highly redundant information and large dimensions. Hence, the following effective tools are applied to get an efficient representation of the images by extracting discriminating and independent features. The four techniques used in this chapter for feature extraction are described later in the section and are illustrated in figure 7.1.

1) The 2D DMWT based on Multiresolution Analysis (MRA) is used for:

   i. Dimensionality reduction.

ii. Localizing all the useful information in one single band.

iii. Noise reduction.

2) The FastICA is used for:

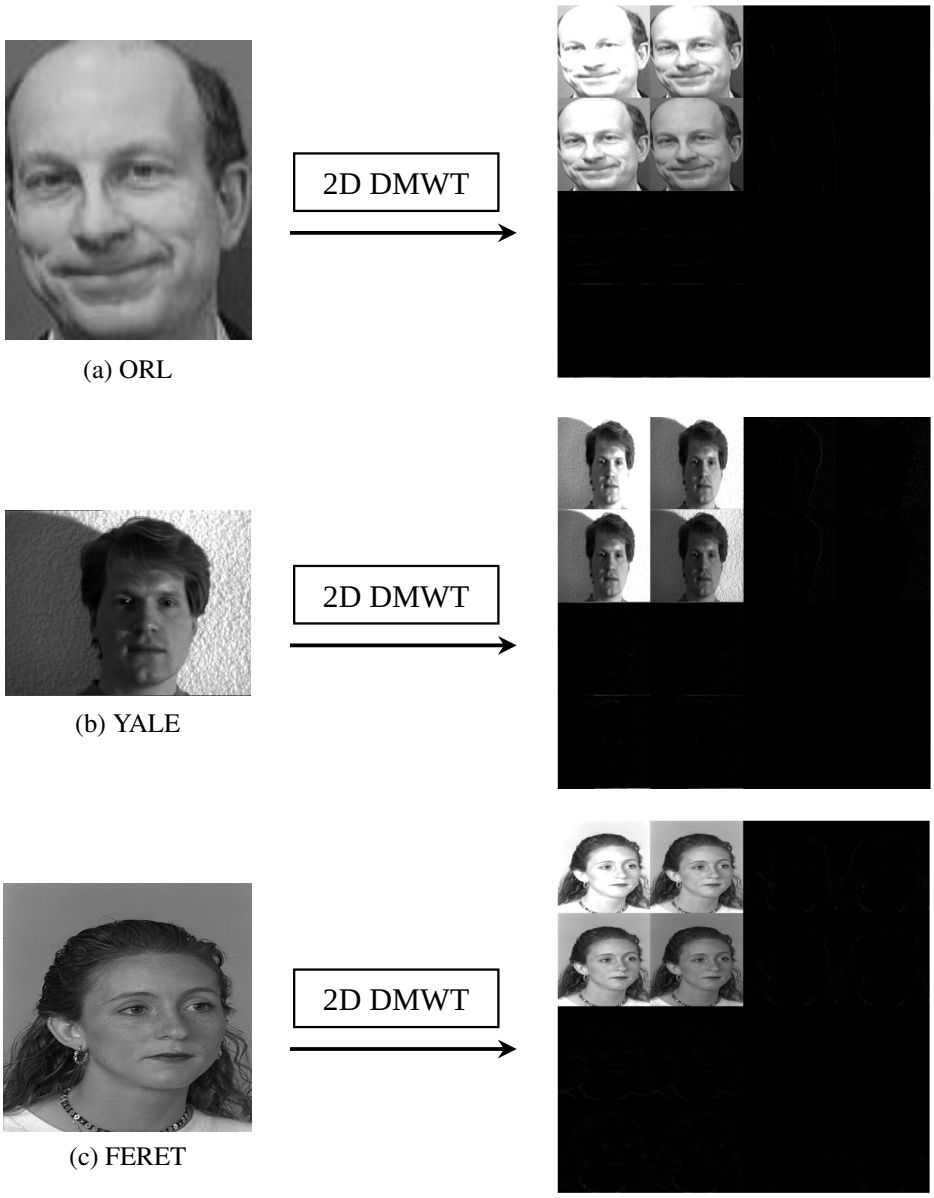
i. Decorrelating the high order statistics since most of the important information is contained in the high order statistics of the image [65].

ii. Reducing the computational complexity and improving the convergence rate.

iii. The ICA features are less sensitive to the facial variations arising from different facial expressions and different poses [71].

iv. The ICA features are independent leading to a better representation and hence better identification and recognition rates.

First, the 2D DMWT is applied to different databases. Figure 7.2 illustrates an example of applying 2D DMWT to the face images. As shown in figure 7.2, the images are divided into four main subbands with dimension 64×64 and each one is further divided into four 32×32 subimages. From subfigures (7.2-a), (7.2-b), and (7.2-c), all the useful information is localized in the upper left band, which corresponds to the low-low (LL) frequency band of the Multiwavelet transform. The LL subband is retained, while the remaining subbands are eliminated. Therefore, the resultant image matrix is 64×64. For the four subimages of the LL subband, the following procedures are performed to obtain the final feature matrix:



(a) ORL

(b) YALE

(c) FERET

Figure 7.2: Application of 2D DMWT on different databases.

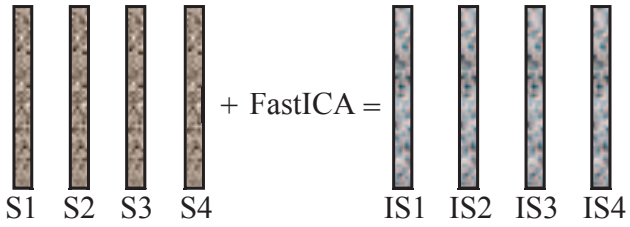


Figure 7.3: Applying 2D FastICA into extracted subimages. Sx represents Subimages of LL subband. ISx represents Independent Subimages of LL subband where $x = 1, 2, 3, and\ 4$.

1) Convert each 32×32 subimage into a 1024×1 vector.

2) Repeat (1) for all subimages of each pose. The resulting features for each pose have dimension 1024×4.

3) Apply FastICA to the resulting feature matrix of each pose as shown in figure 7.3.

4) Extract new features from the original features to reduce the dimensions of the feature space and achieve better performance. This is achieved by first converting each column of ISx (the independent subimages) into 2D form.

5) Find the eigenvalues and eigenvectors as follows:

   A. The extracted eigenvalues for each pose.

      i. Find the eigenvalues for each ISx, which has 32×1 dimension.

      ii. Repeat (i) for all ISx of each pose. The resultant features will have dimension 32×4.

      iii. Convert the 32×4 resultant features into 128×1 1D form, which correspond to each pose. (**Tech. 1**).

   B. The extracted eigenvectors of each pose.

      i. Find the eigenvectors for each ISx, which has 32×32 dimension.

ii. Convert the matrix of eigenvectors to 1D form with 1024×1 dimension.

iii. Repeat (ii) for all ISx of each pose. The resultant feature matrix has dimension 1024×4.

iv. Find the ℓ_2-norm for each row of the resultant matrix to reduce the dimensionality and constrain all the energy of each pose in a 1024×1 column. (**Tech. 2**).

C. In (**Tech. 3**) 2D DMWT is applied and step 5.A without applying FastICA.

D. In (**Tech. 4**) 2D DMWT is applied and step 5.B without applying FastICA.

6) Repeat steps [1-5] for all poses of three different databases. For all techniques, there are two features for each pose. One corresponds to the eigenvalues with 128×1 dimension and the other corresponds to the eigenvectors with 1024×1 dimension.

The feature matrices are fed to a NNT for recognition.

7.1.3 Recognition

In the recognition step, a Neural Network based on the Back Propagation Training Algorithm (BPTA) is used for training and testing. BPTA is a supervised learning algorithm. Therefore, It is necessary to choose a desired output for each database. There are 40, 15, and 200 different desired outputs for the ORL, YALE, and FERET databases, respectively, corresponding to the different number of persons in each database. Three layers are used in the NNT, namely, an Input, a Hidden, and an Output layer.

7.2 Experimental Results

The results of the proposed techniques are presented in this section. The techniques are tested using the ORL, YALE, and FERET databases, which have different light conditions, facial expressions, angle rotation, and a large number of poses and persons. The proposed system achieved 100% recognition rates when tested with the training poses.

7.2.1 Experimental Results for the ORL Database

There are 40 persons in the ORL database, each with 10 different poses [28]. Let P denote the number of poses used for training. Hence, $10 - P$ poses are used for testing. Table 7.2 summarizes the results for the different approaches. As shown, even with a small number of poses for training, the proposed techniques achieve a high recognition rate. As the number of poses increases, the performance is improved.

Table 7.2: Experimental Results for the ORL Database

P for Training	P for Testing	Recognition Rates for Testing Phase			
		Eigenvalues Based 2D DMWT	Eigenvectors Based 2D DMWT	Eigenvalues Based 2D DMWT and FastICA	Eigenvectors Based 2D DMWT and FastICA
1	9	78.05%	83.61%	86.94%	90.56%
3	7	89.64%	92.5%	93.57%	96.07%
5	5	96%	97.5%	98%	99%

7.2.2 Experimental Results for the YALE Database

This database consists of 15 persons, each with 11 different poses [29]. Table 7.3 summarizes the results of the proposed techniques.

Table 7.3: Experimental Results for the YALE Database

P for Training	Poses for Testing	Recognition Rates for Testing Phase			
		Eigenvalues Based 2D DMWT	Eigenvectors Based 2D DMWT	Eigenvalues Based 2D DMWT and FastICA	Eigenvectors Based 2D DMWT and FastICA
1	10	80.67%	87.33%	84%	89.33%
3	8	85.83%	91.67%	90.83%	94.17%
5	6	93.33%	95.56%	96.67%	97.78%

7.2.3 Experimental Results for the FERET Database

There are 200 persons in this database, each with 11 different poses [30, 31]. Table 7.4 summarizes the results of the proposed techniques. The results exhibit the same behavior as in the previous databases.

Table 7.4: Experimental results for the FERET Database

P for Training	Poses for Testing	Recognition Rates for Testing Phase			
		Eigenvalues Based 2D DMWT	Eigenvectors Based 2D DMWT	Eigenvalues Based 2D DMWT and FastICA	Eigenvectors Based 2D DMWT and FastICA
1	10	79.8%	82.65%	83.05%	89.9%
3	8	89.75%	92.56%	93.31%	94.81%
5	6	96.08%	97.25%	97.75%	98.58%

7.3 Discussion and Neural Network Configuration

Note that FastICA decorrelates the images and produces statistically independent sets of images. Then eigenanalysis of the resulting features generates an efficient image representation. The above results confirm that combining these tools with the described techniques yields significant improvement in the recognition rates. As shown in the experimental results, the proposed techniques (1 & 2) outperform techniques (3 & 4) which do not use the FastICA step. This is due to the fact that the eigenvalues and eigenvectors features of (1 & 2) are selected from the independent features (ISx). In other words, the independent features outperform the correlated features. The configuration of the neural network during the training phase can affect the overall performance of the proposed techniques. Choosing the number of hidden layers, the number of neurons in the hidden layers, the types of the activation functions, the training function, the training algorithm, and the target performance can impact the overall performance of the system.

7.4 Conclusion

A facial recognition system based on two consecutive transforms (2D DMWT and 2D FastICA) was proposed. Using eigenanalysis of the multiresolution and independent features of the integrated transforms (2D DMWT and 2D FastICA) led to a significant improvement in the computational complexity, the storage requirements, as well as the recognition rates. Each input image has dimension 128×128 (16384×1) while the eigenvalue features have dimension 128×1 and the eigenvector features have dimension 1024×1 for each input image. This is due to using the LL subband information of the 2D DMWT transform, thereby leading to a significant dimensionality reduction. The performance of the proposed techniques are evaluated using three different databases that have different light conditions, face rotations, and facial expressions. The highest recognition rates recorded are 99%, 97.78%, and 98.58% for the ORL, YALE, and FERET databases, respectively.

CHAPTER 8: HIGH PERFORMANCE AND EFFICIENT FACIAL RECOGNITION USING NORM OF ICA/MULTIWAVELET FEATURES

In this Chapter, a supervised facial recognition system is proposed [39]. For feature extraction, a Two-Dimensional Discrete Multiwavelet Transform (2D DMWT) is applied to the training databases to compress the data and extract useful information from the face images. Then, a Two-Dimensional Fast Independent Component Analysis (2D FastICA) is applied to different combinations of poses corresponding to the subimages of the low-low frequency subband of the MWT, and the ℓ_2-norm of the resulting features are computed to obtain discriminating and independent features, while achieving significant dimensionality reduction. The compact features are fed to a neural network (NN) based classifier to identify the unknown images. The proposed techniques are evaluated using three databases, namely, ORL, YALE, and FERET. The recognition rates are measured using K-fold Cross Validation. The proposed approach is shown to yield significant improvement in storage requirements, computational complexity, as well as recognition rates over the existing approaches.

8.1 Proposed Approach

In this section, the proposed system is presented. The supervised facial recognition system consists of three phases, namely, Preprocessing, Feature Extraction, and Recognition.

8.1.1 Preprocessing

This phase consists of two steps:

1) The first step aims to convert all dimensions of different image databases used in this chapter (see Table 12.1) to an appropriate one. This dimension is chosen since the algorithm used requires dimensions that are power of two.

Table 8.1: The Dimensions of the Databases

Databases	ORL	YALE	FERET	The Proposed
Size	112×92	243×320	384×256	128×128

2) This step aims to convert all the images into $(128 \times 128\ double)$ instead of using $uint8$ extension, which is not suitable for the transform used in this chapter.

8.1.2 Feature Extraction

Feature extraction is an essential component of facial recognition systems. A face image may contain information that is not essential for the recognition purpose. Hence, the following effective tools are applied to get an efficient representation of the images by extracting discriminating and independent features:

1) The 2D DMWT based on MRA is used for:

   i. Dimensionality reduction.

   ii. Noise reduction.

   iii. Localizing all the useful information in one single band.

2) The FastICA is used for:

i. Decorrelating the high order statistics since most of the significant information is contained in the high order statistics of the image [65, 90, 144, 143, 68, 66, 64].

ii. The ICA features are independent leading to a better representation and hence better identification and recognition rates.

iii. The ICA features are less sensitive to the facial variations arising from different facial expressions and different poses [65, 90, 144, 68, 64].

iv. Reducing the computational complexity and improving the convergence rate.

First, the 2D DMWT is applied to the different databases. Figure 8.1 shows an example of applying 2D DMWT to different face images. As shown in figure 8.1, the images are divided into four main subbands with 64×64 dimension and each one is further divided into four 32×32 subimages. From subfigures (8.1-a), (8.1-b), and (8.1-c), it is easily seen that most of the useful information is localized in the upper left band, which is related to the low-low (LL) frequency band of the DMWT. Thus the LL subband is retained, and all the remaining subbands are eliminated. Therefore, the resulting image matrix is 64×64. The following procedures are applied to the four subimages of the LL subband of DMWT in order to get an efficient representation for the input images.

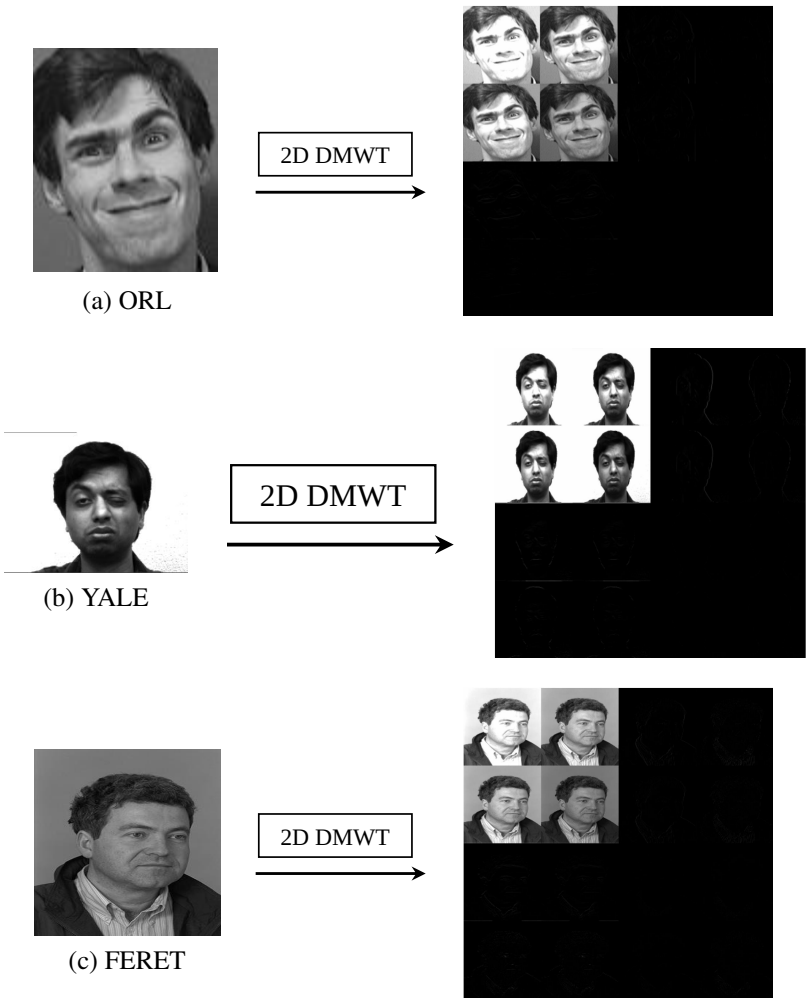
1) Convert each 32×32 subimage into a 1024×1 vector.

2) Repeat (1) for all subimages of each pose. The resulting features for each pose have dimension 1024×4.

3) For each of the four subimages of the LL subband, combine the 1024×1 subimage vectors of the different poses of each person as shown in figure 8.2-B.

Thus, the combinations shown in figure 8.2-B are defined as:

$$\text{Combination}_i = [S_{pi}] \tag{8.1}$$



(a) ORL

(b) YALE

(c) FERET

Figure 8.1: Application of 2D DMWT on different databases.

where $p \in \{1, 2, 3, \ldots, P\}$ and P is the number of poses used in the training mode, $i \in \{1, 2, 3, 4\}$ indexes the four subimages of the LL subband.

It is very useful to extract new features from the original features to reduce the dimensionality of the feature space. To this end, we apply the following four techniques to the *Combinations* above:

A. *Technique 1* (shown in figure 8.4)

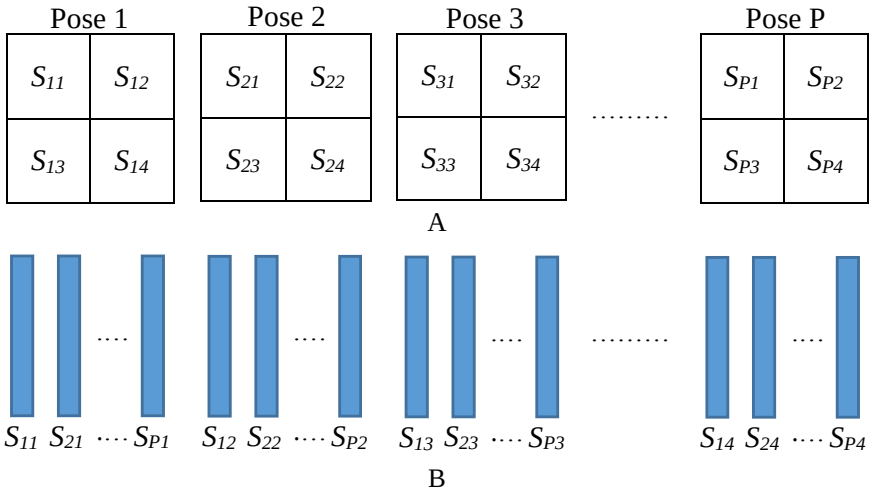


Figure 8.2: Figure 8.2-A represents the subimages of the LL subband of 2D DMWT for P training poses from each person. Figure 8.2-B is rearranging Figure 8.2-A in four combinations.

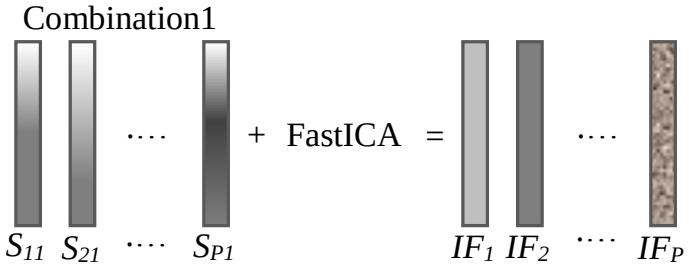


Figure 8.3: An example of applying 2D FastICA to the first combination. IF_p represents the resultant Independent features where $p \in \{1, 2, 3, \ldots, P\}$.

i. Apply 2D FastICA on each combination as shown in figure 8.3.

ii. Choose the IF_p column that has the maximum ℓ_2-norm to represent the features of each combination. In this case, the resultant features of each person have dimensions 1024×4 regardless of the number of poses used in both the training and the testing modes.

B. *Technique 2* (shown in figure 8.5)

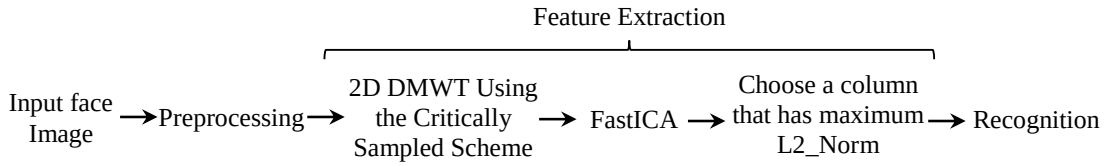


Figure 8.4: Proposed Technique 1.

i. Apply 2D FastICA on each combination.

ii. Form a 1024×1 column vector whose entries consist of the ℓ_2-Norm for each row of the resultant features to reduce the dimensionality. The resultant features of each person have dimension 1024×4 regardless of the number of poses used in training mode.

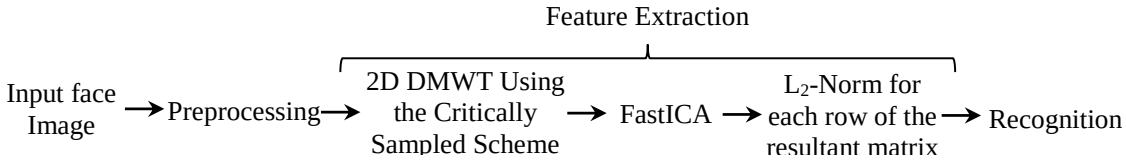


Figure 8.5: Proposed Technique 2.

C. *Technique 3* (shown in figure 8.6): This is similar to Tech.2, except that we use the mixing matrix A shown in (3.26) instead of using the resultant FastICA matrix. This is an alternative approach for representing the face images [144, 64].

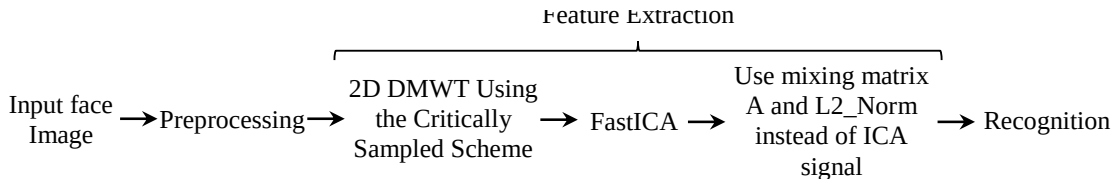


Figure 8.6: Proposed Technique 3.

D. *Technique 4* (shown in figure 8.7): Same as Tech.2 without applying step B-i.

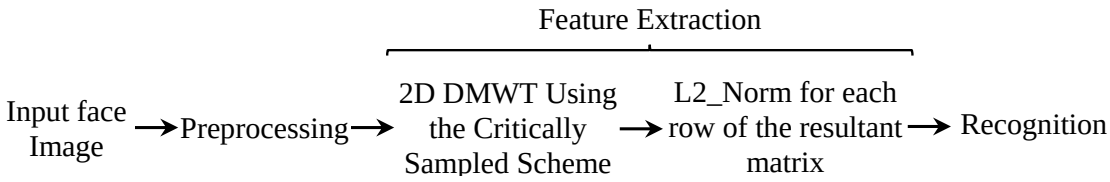


Figure 8.7: Proposed Technique 4.

4) Repeat [1-3] for all poses of the three different databases.

5) Repeat Tech.1-Tech.4 for all persons in the three different databases. For all techniques, there are 1024×4 features corresponding to each person regardless of the number of poses used in training mode.

8.1.3 Recognition

A Neural Network (NNT) based on the Back Propagation Training Algorithm (BPTA) is used during the training and the testing phases. Since BPTA is a supervised learning algorithm, it is necessary to choose a desired output for each database used. In this chapter, the system is evaluated using three different databases, each with a different number of persons. Therefore, there are 40, 15, 200 desired outputs for the ORL, YALE, and FERET databases, respectively, corresponding to the number of persons in each database. The configuration of the NNT has three layers, namely, an input layer, a hidden layer, and an output layer.

In the training mode, the classifier is configured using the described training features. For testing, the recognition rate is measured as $\frac{D}{L} \times 100\ \%$, where D is the total number of images correctly matched and L is the total number of images in the database.

In the testing mode, we follow the same steps of the training phase. First, the 2D DMWT is applied to the test image. Second, only the LL subband is retained. Third, each subimage is converted to 1D form. Then, Tech.1-Tech.4 are applied to the 1024×4 resultant features corresponding to the four subimages of the LL subband. For each technique, there is a 1024×1 feature vector for each test image. This vector is fed into the NNT for classification.

8.2 Experimental Results and Discussion

The experimental results of the proposed techniques are presented in this section. We also compare the results of the proposed approach to some of the state-of-the-art methods. The techniques are tested using three different databases, ORL, YALE, and FERET, which have different facial expressions, light conditions, and rotations. K-fold Cross Validation [154] is used to evaluate our proposed techniques. Different values of K are used in this chapter as show in Table 8.2, 8.3, and 8.4.

8.2.1 Experimental Results for the ORL Database

The ORL database consists of 40 persons, each with 10 different poses [28]. Therefore, P different poses are used in the training mode, and $10 - P$ poses are used in the testing mode. Table 8.2 summarizes the results for all different techniques. As shown, even when a small number of poses are used in the training mode, the algorithm achieves high recognition accuracy. As the number of poses increased, the performance is improved.

Table 8.2: Experimental Results for the ORL Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and 2D ICA [65]	2D DWT and 2D PCA [90]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	96.75%	**98.25%**	97.25%	96.25%	93.75%	92.5%
K=3	100%	97.5%	**98.75%**	97.75%	96.75%	95.25%	94.75%
K=5	100%	98.25%	**99.25%**	98.75%	97.75%	96.75%	96.5%

8.2.2 Experimental Results for the YALE Database

This database consists of 15 persons, each with 11 different poses [29]. Table 8.3 summarizes the results of the proposed techniques. As before, our proposed techniques show superior performance.

Table 8.3: Experimental Results for the YALE Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and 2D ICA [65]	2D DWT and 2D PCA [90]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	96.97%	**98.18%**	96.97%	95.15%	92.73%	92.12%
K=3	100%	97.58%	**98.79%**	97.58%	96.36%	94.55%	93.94%
K=5	100%	98.18%	**99.39%**	98.78%	97.58%	96.36%	95.15%

8.2.3 Experimental Results for the FERET Database

There are 200 persons in this database, each with 11 different poses [30, 31]. Table 8.4 summarizes the results for the proposed techniques. The results exhibit the same behavior as in the previous databases.

Table 8.4: Experimental results for the FERET Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DWT and 2D ICA [65]	2D DWT and 2D PCA [90]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	96.68%	**98.32%**	97.36%	95.86%	92.87%	92.5%
K=3	100%	97.6%	**98.6%**	98.1%	96.9%	95.14%	94.86%
K=5	100%	97.78%	**99%**	98.46%	96.96%	96.14%	95.68%

8.3 Discussion

The goal of using FastICA in this chapter is to find a set of statistically independent basis images that preserve the structural information of the face images present in the facial expressions, illumination, and rotations. ICA representations [64] are more robust to variations due to different light conditions, changes in hair location, make up, and facial expressions. This efficient representation leads to significant improvement in the recognition rates. As demonstrated in the previous results, Tech. 1, Tech. 2, Tech. 3 outperform Tech. 4, which does not use the FastICA step. Our proposed techniques are shown to outperform the existing method [65, 90] A in the recognition rates, storage

requirements, and computational complexity. The configuration of the Neural Network during the training mode can effect the overall accuracy of the system. Hence, choosing the number of hidden layers, number of neurons in each hidden layer, activation function, training algorithms, and the target performance play an important role in the overall system performance.

8.4 Conclusion

A facial recognition system based on the integration of Multiresolusion Analysis and FastICA was proposed. The proposed approach leads to notable dimensionality reduction and subsequently less storage requirements, as well as higher recognition rates. These benefits are due to using combinations of different poses for the subimages of the LL subband of the 2D DMWT and the ℓ_2 Norm for 2D FastICA features in the feature extraction step. Therefore, we exploit the redundancy present in these subimages to reduce the dimensions of the used features, which also account for all the variations in one short feature vector per combination. Each person is represented using features of dimension 1024×4 regardless of the number of poses in the training phase. Compared with the existing methods, the proposed approach achieves higher recognition rates with less storage and computation requirements. The proposed techniques were evaluated using three different databases with different light conditions, rotations, and facial expressions. The recognition rates are analyzed using K-fold CV. For example, for K=2, the highest recognition rates achieved were 98.25%, 98.18%, and 98.32% for the ORL, YALE, and FERET databases, respectively.

CHAPTER 9: A FACIAL RECOGNITION METHOD BASED ON DMW TRANSFORMED PARTITIONED IMAGES

In this chapter, Two Dimensional Discrete Multiwavelet Transform (2D DMWT) is applied to the processed/partitioned facial images for face recognition [40]. First, the input facial image is divided into six parts in the preprocessing step to reduce the effect of the non-efficient/discriminant features on the system performance. Then the 2D DMWT is applied to each part for feature extraction and dimensionality reduction. Finally, the ℓ_2-Norm is applied for further feature compression. These features are classified using a Neural Network (NN) that is trained using Back Propagation Training Algorithm (BPTA). The proposed techniques are evaluated using four databases, namely, ORL, YALE, FERET, and FEI. These databases cover a wide variety of facial variations, such as facial expressions, rotations, illuminations, etc. K-folds Cross Validation (CV) is used to analyze the experimental results. The proposed technique is shown to improve the recognition rates as well as the storage requirements in comparison with the other approaches.

9.1 Proposed Techniques

The proposed approach has three important steps, namely, preprocessing, feature extraction, and classification, as shown in figure 9.1. The proposed system is evaluated using four databases, namely, ORL, YALE, FERET, and FEI. Samples of the databases are shown in figure 11.2.

9.1.1 Preprocessing

The preprocessing phase has further steps:

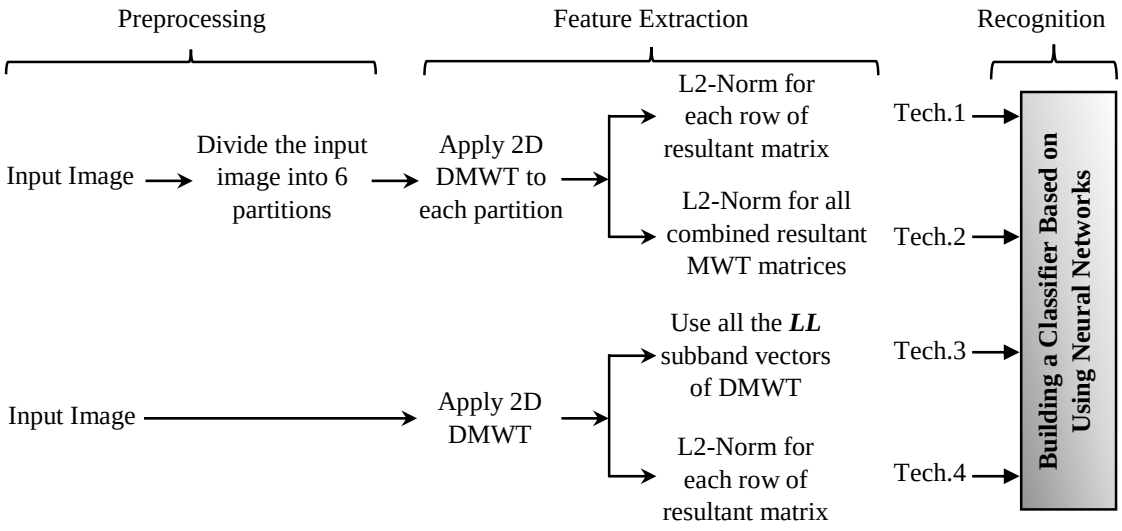


Figure 9.1: The Proposed Techniques.

1) Apply cropping technique to the databases to remove the redundant background information. The dimensions of the databases after cropping are shown in Table.9.1.

2) Divide all different poses of all persons into six parts as shown in figure 9.2. The motivation here is to reduce the effect of non-efficient/discriminating features on the system performance and to improve the NN convergence rates, which lead to improve the recognition rates.

Figure 9.2: Dividing the image of ORL database into 6 parts.

3) Convert all different dimensions of all parts of all persons of four databases to appropriate one (see Table 9.1). These dimensions are chosen because the algorithm used (DMWT) required dimensions that are power of two.

Table 9.1: The Dimensions of the Databases

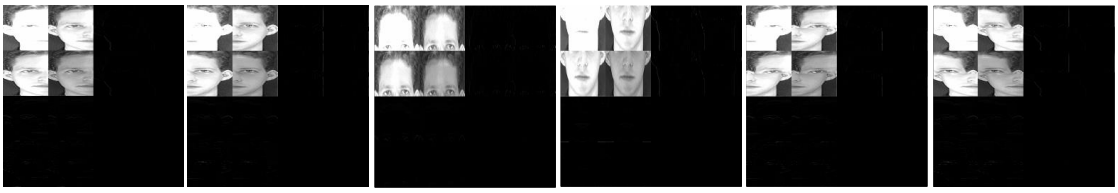
Databases	Actual Size	Dimension After Cropping	Proposed Dimension for each part
ORL	112×92	112×92	64×64
YALE	243×320	180×150	128×128
FERET	384×256	220×170	128×128
FEI	480×640	320×290	256×256

4) Finally Convert the **uint8** data-type extension of all parts, which is not suitable for the transform used in this chapter, into double data-type.

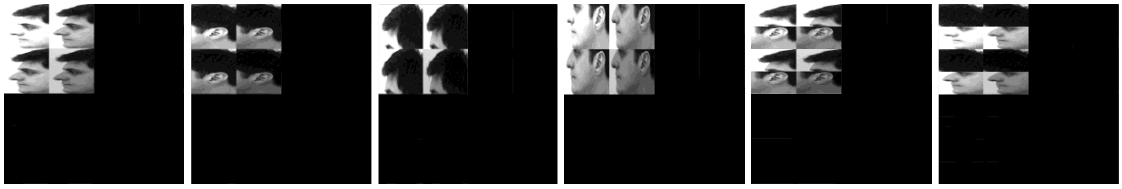
5) Apply a critically-sampled scheme Preprocessing (approximation based preprocessing) to input facial images [121].

9.1.2 Feature Extraction

The goal of this step is to find the best features in the database. Feature extraction step compromises of reducing number of elements required to describe the large databases. Hence, it leads to dimensionality reduction and reduces the computational complexity.



(a) ORL



(b) FEI

Figure 9.3: Applying 2-D DMWT to all partitions of one sample of each different databases.

Therefore, the feature extraction considered as a substantial part in the recognition system. Because the human faces are not rigid bodies, there is lot of irrelevant information that is effecting the convergence (recognition rates) and increasing the computational complexity. Therefore, the following efficient transform based on multiresolution analysis, 2D DMWT, is used for:

1) Dimensionality reduction.

2) Noise reduction.

3) Localizing all the useful information in one single band.

As mentioned that each pose in the databases is divided to six parts for minimizing the effect of non useful features and for faster convergence. In the feature extraction step, the 2D DMWT is applied to each part of each pose. Figure 9.3 shows the result of applying 2D DMWT to one sample of ORL database. In figure 9.3, each part is divided into four sub-bands with 32×32, 64×64, and 128×128 dimensions for ORL, YALE $\&$ FERET, and FEI database, respectively.

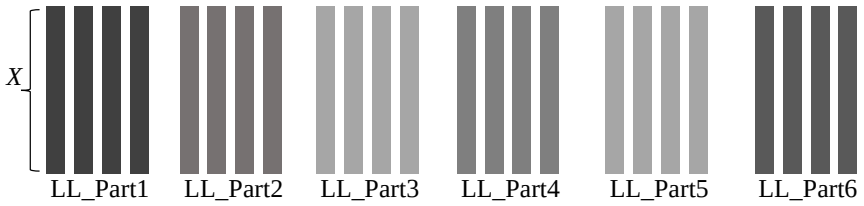


Figure 9.4: Shows the extracted features for one pose after converted into one dimensional form.

Additionally, each sub-band is further divided into four sub-images with dimensions of 16×16, 32×32, 64×64 for ORL database, YALE and FERET databases, and for FEI database, respectively. As shown from sub-figures (9.3-a) and (9.3-b) all the useful information is localized in the low-low (LL) frequency sub-band, upper left band, of DMWT. Hence, the LL sub-band is extracted and all the other sub-bands are excluded. Therefore, the resultant features of each pose has $16 \times 16 \times 24$, $32 \times 32 \times 24$, and $64 \times 64 \times 24$ dimensions for ORL, YALE $\&$ FERET, and FEI databases, respectively. The following steps are applied to the databases to get better facial representation:

1) Convert each sub-image in the LL sub-band of each part to a vector of 256×1, 1024×1, and 4096×1 dimensions for ORL, YALE and FERET, and FEI databases, respectively.

2) Repeat (1) to all sub-images in the LL sub-band. Hence, the extracted features for each pose has $X \times 24$ dimensions all databases, as shown in figure 9.4, where X is varied based on which database get involved.

3) ***Technique 1*** is achieved by find ℓ_2-Norm for each row of each part. Therefore, the resultant features extracted for each pose have dimensions of $X \times 6$.

4) ***Technique 2*** is accomplished by find ℓ_2-Norm for each row of each pose. Thus, the resultant features extracted for each pose have dimensions of $X \times 1$.

5) ***Technique 3*** is the same as ***Technique 1*** but without applying **partitioning step** and the

ℓ_2**-Norm**. This means that all the sub-images are used. The final feature matrix of each pose has $X \times 4$ dimensions.

6) ***Technique 4*** is the same as ***Technique 1*** but without applying **partitioning step**. In this case, the final feature matrix of each pose has $X \times 1$ dimensions.

7) Repeat [3-6] for all poses in the four databases.

8) The final feature matrix for each person has $X \times Y$ dimensions, where X is varied according to which database is used. Y is corresponding to the number of poses of each person used in the training mode, and to which technique is employed.

Then the final features are forwarded to the recognition step.

9.1.3 Recognition

There are two modes in the system, training and testing modes. A NN based on BPTA is used in both modes. Since the BPTA is a supervised learning algorithm, it is necessary to choose the desired output for each person. The desired output of each person should be different from the other persons. For example, the desired output of person 1 in the YALE database is $[1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1\ -1]^T$. The number 1 means that the output is active and the number -1 means that the output is inactive. There are 40, 15, 200, and 200 different desired outputs for ORL, YALE, FERET, and FEI databases, respectively. The structure of the NN consists of three layers, namely, an input layer, a hidden layer, and an output layer.

The classifier is built in the training mode using the ***Techniques'*** features. For testing, the recognition rate is measured as $\frac{D}{L} \times 100\ \%$, where D is the total number of poses correctly matched and L is the total number of poses in the database. For ***Technique 1***, if 4 partitions out of 6 partitions

are correctly matched , the input pose is correctly matched the exact person. In case of ***Technique 3***, if 3 sub-images out of 4 sub-images are correctly matched , the input pose is correctly matched the exact person.

The same steps are followed to generate the testing features. First, the input pose is divided into 6 parts. Then the Multiwavelet preprocessing step is applied. Then the 2D DMWT is applied to the preprocessed data and only the LL sub-band is retained. Thereafter, ***Techniques 1-4*** are computed. Finally, the final feature matrix is fed into NN for classification.

9.2 Experimental Results

The results of the proposed techniques are presented in this section. The proposed techniques are evaluated using four databases, namely, ORL, YALE, FERETt, and FEI that have different facial variations, such as light conditions, rotations, facial expressions, etc. We compared our results with the other methods. The results are analyzed using K-fold Cross Validation (CV) [154]. In this chapter, three values of K-fold are used as shown in Tables 9.2-9.5. The rates reported in Tables 9.2-9.5 are the average of the rates obtained across the K-folds CV.

9.2.1 Experimental Results for the ORL Database

There are 40 individuals in ORL database, each with 10 different poses. All poses in the database have lot of diversities, such as facial details (glasses/ no glasses) and facial expressions (open / closed eyes, smiling / not smiling) [28]. The results are summarized in the Table 9.2. As we can see when the value of K increased, the performance is improved.

Table 9.2: Experimental Results for the ORL Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DCT 2D PCA[95]	2D DWT and 2D PCA[81]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	96%	95.75%	94.5%	94.25	92%	95%
K=3	100%	96.8%	96.25%	95.3%	95.1	94.75%	95.5%
K=5	100%	97.5%	97.25%	96.25%	96.25	96%	96.75%

9.2.2 Experimental Results for the YALE Database

The YALE database consists of 15 persons, each with 11 different poses. All poses in the database have different variations, such as happy, sad, w/no glasses, left-right-center light, sleepy, surprised, normal, and wink [29]. The results are summarized in Table 9.3.

Table 9.3: Experimental Results for the YALE Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DCT 2D PCA[95]	2D DWT and 2D PCA[81]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	94.67%	94.56%	93.44%	93.33	93.22%	93.44%
K=3	100%	96.11%	95.97%	95.28%	95	94.44%	95.56%
K=5	100%	96.81%	96.4%	95.97%	95.7	96.34%	96.55%

9.2.3 Experimental Results for the FERET Database

There are 200 persons in the FERET database, each with 11 different poses. FERET database consists of different facial configurations, such as facial expressions, rotations, light conditions, w/no glasses [30, 31]. Table 9.4 summarizes the results and the proposed techniques exhibit the same behavior as previous results.

Table 9.4: Experimental Results for the FERET Database

K	Training	Recognition Rates for Testing Phase				2D DCT	2D DWT and
Fold	Rates	Tech.1	**Tech.2**	Tech.3	Tech.4	2D PCA[95]	2D PCA[81]
K=2	100%	95.75%	95.43%	94.73%	94.55	93.94%	94.83%
K=3	100%	96.68%	96.35%	95.41%	95.1	94.74%	95.71%
K=5	100%	97.24%	96.9%	96.27%	96.1	95.7%	96.4%

9.2.4 Experimental Results for the FEI Database

In the FEI database, there are 200 individuals, each with 14 different poses. All poses are colored with 180 degree rotations, hairstyle, distinct appearance, and adorns [32]. The results are summarized in Table 9.5. As shown in the preceding databases the proposed techniques outperformed the existing methods.

Table 9.5: Experimental Results for the FEI Database

K Fold	Training Rates	Recognition Rates for Testing Phase				2D DCT 2D PCA[95]	2D DWT and 2D PCA[81]
		Tech.1	**Tech.2**	Tech.3	Tech.4		
K=2	100%	95.75%	95.46%	94.68%	94.5	93.97%	94.8%
K=3	100%	96.64%	96.4%	95.3%	95.15	94.83%	95.64%
K=5	100%	97.44%	97.1%	96.63%	96.42	95.85%	96.25%

9.3 Discussion

The goal of dividing the input pose into six parts is to reduce the effect of unnecessary and non-discriminating features on the overall system performance. Including inefficient and redundant features in the feature vector space not only degrades the recognition accuracy but also increases the computational complexity.

The 2D DMWT is used to reduce the dimensions of the databases and extract the most relevant features from the input facial images by retaining only the LL sub-band, as shown in figure 9.3. Since each part is shared with the other two parts as seen in figure 9.5, the matching probability for each pose is increased. Since ***Technique 1*** is applying DMWT to each part that leads to six resultant features for each pose, There are six experimental results for each pose, one for each part. To obtain the final decision, we applied the majority vote. If four parts are correctly matched, the input test image is declared as correctly matched to the right person.Therefore, ***Technique 1*** outperformed the other ***Techniques*** due to this division preprocessing.

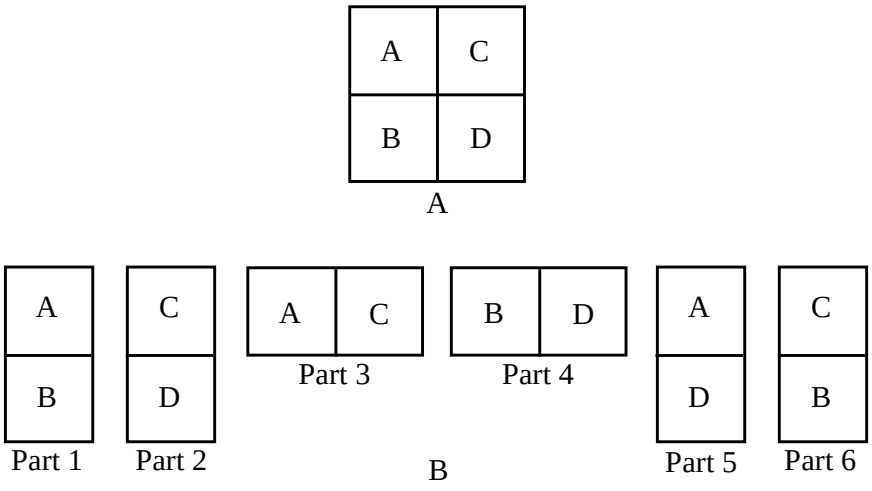
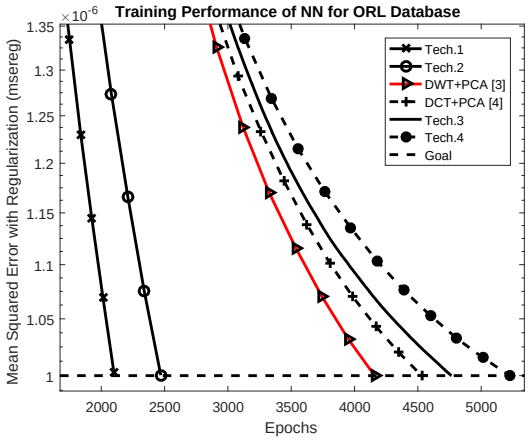


Figure 9.5: Shows how the parts are participating with each other

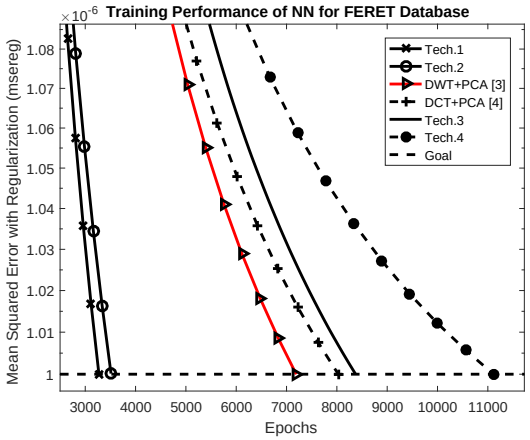
Figure 9.5 shows how parts are engaged with each other. Therefore, Using the division preprocessing step leads to a notable improvement in the recognition rats. If the first four parts are correctly matched the right person, the majority is achieved and the decision is made and the program is stopped. Hence, the computational complexity is improved. As shown in the Tables above, our proposed techniques also outperformed the existing approaches presented in [81, 95].

9.4 Neural Network Performance

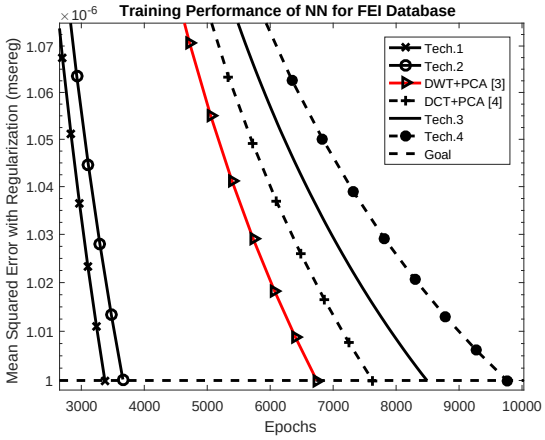
Even though efficient features are obtained, the recognition rates might not be improved. This is due to the fact that the arrangement of the NN in the training phase has a direct effect on the performance of the recognition system. Several parameters can have that effect, such as choosing number of hidden layers, number of neurons in each hidden layer, types of activation functions, training algorithms, desired outputs, learning rates, and the target. In this paper, we used one hidden layer with 256 neurons, 512 neurons, and 1024 neurons for ORL, YALE & FERET, and FEI databases, respectively.



(a) ORL



(b) FERET



(c) FEI

Figure 9.6: The performance of Neural Network for ORL, FERET, and FEI databases.

The hyperbolic tangent sigmoid transfer function ($tansig(x)$) is used for all layers. During the training and testing modes, the BPTA was employed. Figure 9.6 shows as example of NN performance in the training mode. As show in figure 9.6, our proposed techniques (1 & 2) achieved the **goal** faster than the other techniques and the other approaches [81, 95].

9.5 Conclusion

A new system based on applying 2D DMWT to the partitioned faces was proposed in this chapter for face recognition. The partitioning step employed in the preprocessing steps was used to reduce the effect of inefficient features on the system performance, improve the NN performance, and improve the recognition rates. The dimensionality reduction and efficient features extracted were achieved by retaining the LL sub-band of 2D DMWT. The features were further compacted into one single column by employing the ℓ_2-Norm to each row of the resultant DMWT matrix. The final features extracted using the proposed techniques that fed to NN for training/testing have dimensions of $X^* \times Y^*$, where X^* is 256, 1024, and 4096 dimensions for ORL, YALE $\&$ FERET, and FEI databases, respectively, and Y^* depends on which technique is involved. The proposed techniques were tested using four databases, namely, ORL, YALE, FERET, and FEI that have different facial configurations, such as light conditions, rotation, facial expressions, etc. K-fold CV was used to evaluate our experimental results and different values of K were selected as shown in the Tables above. From the Tables 9.2-9.5, our proposed techniques outperformed the methods presented in [81, 95]. The proposed Techniques (1 $\&$ 2) achieved at the goal faster than the existing approaches and Techniques (3 $\&$ 4) as shown in figure 9.6.

CHAPTER 10: FACE RECOGNITION EMPLOYING DMWT FOLLOWED BY FASTICA

Face recognition become a challenging topic in several fields since images of faces are varied by changing illuminations, facial rotations, facial expressions, etc. In this chapter, Two Dimensional Discrete Mulltiwavelet Transform (2D DMWT) and Fast Independent Component Analysis (FastICA) are proposed for face recognition [41]. These algorithms are integrated to achieve less storage requirements and high accuracies. Preprocessing, feature extraction, and classification are the main steps in the proposed system. In the preprocessing step, Each pose in the database is divided into six parts to reduce the effect of unnecessary facial features and highlight the local features in each part. For feature extraction, the 2D DMWT is applied to each part for dimensionality reduction and features extraction. There are two representations resulted from DMWT step. Then FastICA followed by ℓ_2-norm are applied to each representation, which produce different techniques. This results in features that are more discriminating, less dependent, and more compressed. In the recognition step, the resulted compressed features from two representations are fed to a Neural Network (NN) based classifier for training and testing. The proposed techniques are extensively evaluated using five databases, namely, ORL, YALE, FERET, FEI, and LFW, which have different facial variations, such as illuminations, rotations, facial expressions, etc. The results are analyzed using K-fold Cross Validation (CV). Sample results and comparison with a large number of recently proposed approaches are provided. The proposed approach is shown to yield significant improvement compared with the other approaches.

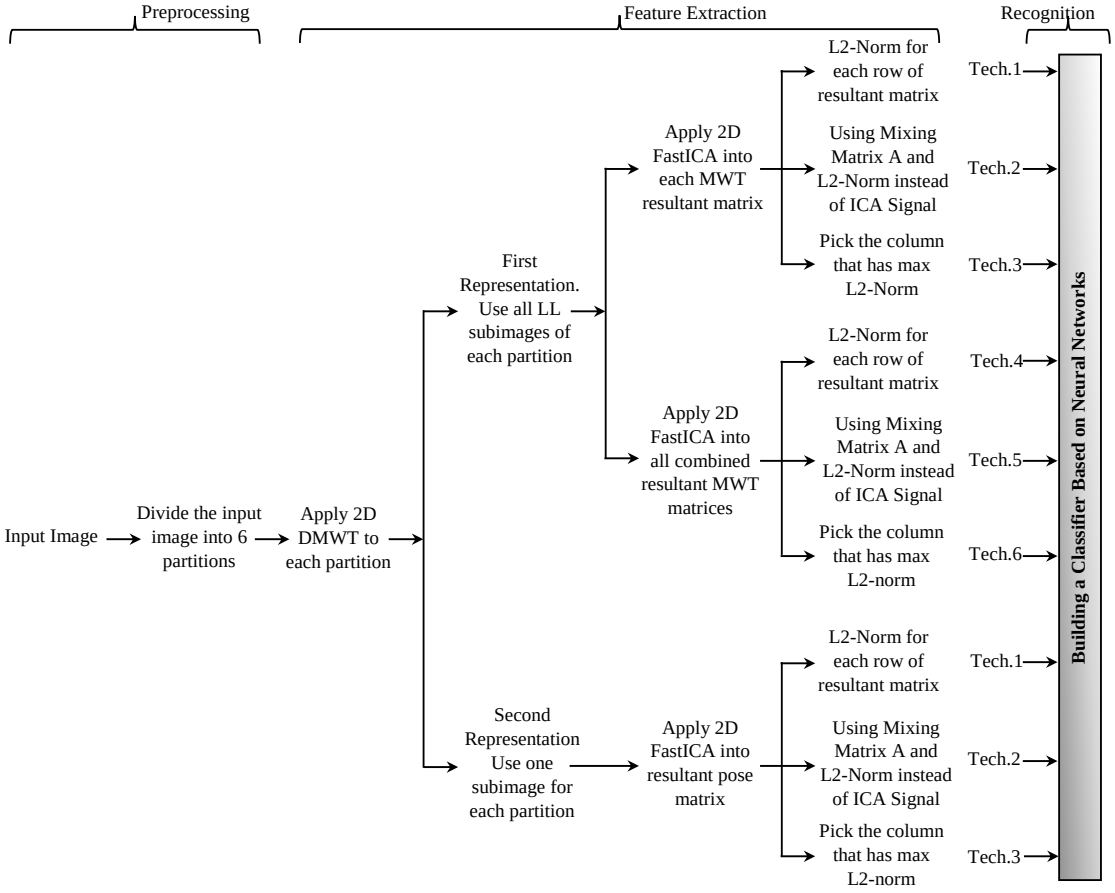


Figure 10.1: The Proposed System

10.1 Proposed System

The proposed system consists of three main phases as shown in figure 10.1. In the first phase, all poses of persons are divided into six parts to reduce the effect of the unnecessary features and to increase the effectiveness of the local features. The second phase aims to find the discriminating features from the set of different databases by employing the proposed transform and techniques. The determined discriminated features will enhance the recognition rates. Finally, NN based on BPTA is applied to the extracted features to built a classifier system. Five databases are used to evaluate the proposed techniques. Samples of each database are provided in figure 6.2.

Figure 10.2: Dividing the first pose of person one of ORL database into 6 parts.

10.1.1 Preprocessing

The preprocessing phase consists of several steps:

1) Apply cropping technique to the databases used as shown in figure 6.2. This step aims to remove most of the background information (unnecessary information) found in each pose.

2) Divide all poses of all persons in the databases to 6 partitions, see figure 10.2. The motivation from partitioning is to differentiate the local features in each part. This reduces the effect of the common features, which are mutual in some partitions between persons, on the overall system performance.

3) Convert dimensions of partitions of each pose into adequate dimensions as shown in Table 10.1. The proposed dimensions were chosen since the algorithm used in this chapter required dimensions that are power of two.

4) Apply a critically-sampled scheme Preprocessing (approximation based preprocessing) to input facial images [121, 122, 125].

Table 10.1: The Dimensions of The Databases

Databases	Actual Size	Dimensions After Cropping	Dimensions for each partition	Proposed Dimension for each partition
ORL	112×92	112×92	112×46 OR 56×92	64×64
YALE	243×320	180×150	180×75 OR 90×150	128×128
FERET	384×256	220×170	220×85 OR 110×170	128×128
FEI	480×640	320×290	320×145 OR 160×290	256×256
LFW	250×250	64×64	64×32 OR 32×64	32×32

10.1.2 Feature Extraction

Variations in the faces of the subjects in terms of facial expressions, illumination, partial obstacles, rotations, makeups, etc. lead facial recognition to be a challenging topic in several fields although there are many samples for each subject. Since the human faces are not solid bodies, there is a lot of unnecessary information that is not essential for recognition. Feature extraction, in which the data compression is employed, is considered to be an important step in every recognition system. Since the input facial image has large dimensions, which have lot of useless information, the recognition rate decreases and the computational complexity increases.

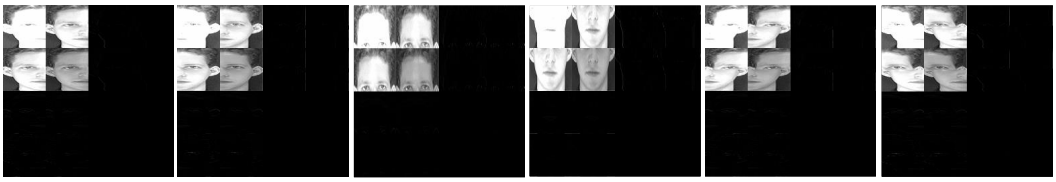
Therefore, the goal of feature extraction is to find the most pertinent features from the original images to best represent the faces and achieve high accuracy while fulfilling the low dimensionality and less storage requirements. Hence the following efficient tools are applied to get better representations of the images by extracting the most discriminating, and independent features:

1) The 2D DMWT based on MRA is used for:

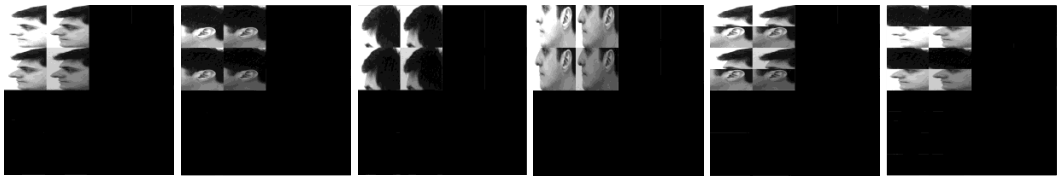
   i. Dimensionality reduction, Less storage requirement.

   ii. Noise alleviation.

   iii. Feature selection by localizing most of the energy in a single band.



(a) ORL



(b) FEI

Figure 10.3: An example of applying 2-D DMWT to ORL and FEI databases.

2) The FastICA is used for:

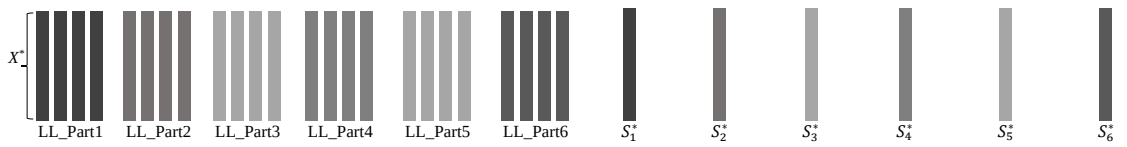
   i. Decorrelating the high order statistics beside decorrelating the second order moments. Most of the information about the local characteristics of the facial images is contained in the high order statistics. Hence, ICA basis is used for better representing the input facial images [64, 83, 67].

   ii. Improving the convergence rate and reducing the computational complexity [64, 144].

This leads to the resultant ICA features are:

a) Less sensitive to the facial variations arising from different facial expressions, illuminations, rotations, different poses, etc. [64, 83, 67, 144].

b) Independent. The ICA architecture-I produces spatially localized representation compared with PCA and ICA architecture-II that produce global representations, which are sensitive to any distortion in the faces. These traits of ICA architecture-I enhance the system performance [64, 75, 67].



(a) First Representation (b) Second Representation

Figure 10.4: Two representations. $S_1^* \dots S_6^*$ are denoted as sub-images with maximum ℓ_2-Norm.

The first step in the feature extraction is applying 2D DMWT to each partition of each pose in the databases. Samples of 2D DMWT results are shown in figure 10.3. As shown in figure 10.3, each partition is divided into four main sub-bands with 32×32 dimensions for ORL database, 64×64 dimensions for both YALE and FERET databases, 128×128 dimensions for FEI database, while each sub-band in LFW database has 16×16 dimensions. Moreover, each sub-band is further divided into four sub-images each with 16×16 dimensions for ORL database, 32×32 dimensions for both YALE and FERET databases, 64×64 dimensions for FEI database, and 8×8 dimensions for LFW database. From sub-figures (10.3-a), (10.3-b), most of the relevant information is localized in the upper left sub-band, which is low-low (LL) frequency sub-band. Therefore, the LL sub-band is maintained and all the other sub-bands are eliminated. Hence, the resultant feature tensor for each partition has dimensions of $16 \times 16 \times 4$, $32 \times 32 \times 4$, $64 \times 64 \times 4$, and $8 \times 8 \times 4$ for ORL, YALE and FERET, FEI, and LFW databases, respectively. The following procedures are applied to the retained feature matrices to obtain better facial representations:

a) Convert each sub-image in the retained LL sub-band of each partition to 1-D vector of 256×1, 1024×1, 4096×1, 64×1 for ORL, YALE and FERET, FEI, and LFW databases, respectively.

b) Repeat [1] to all sub-images in each partition of each pose in the databases.

There are two representations as shown in figure 10.4:

A. The first representation shown in figure 10.4-a is using all four sub-images of LL sub-band with $X^* \times 4 \times 6$ dimensions for each pose, where X^* is varied according to which database is involved.

B. The second representation, figure 10.4-b, uses only one sub-image that has maximum ℓ_2-Norm. Hence, each pose has $X^* \times 6$ dimensions.

Producing new features from the original one is a very beneficial step in order to reduce the dimensionality and achieve better facial representation. In order to achieve that, the following techniques are applied to the resultant matrices corresponding to the first and second representations:

A. First Representation shown in figure 10.4-a.

- *Technique 1*
  i. Apply 2D FastICA to each resultant LL sub-band matrix of each partition that has $X^* \times 4$ dimensions, see figure 10.4-a.
  ii. Find ℓ_2-Norm for each row of the resultant ICA matrix, independent features IF_l as shown in figure 10.5, to reduce the dimensionality of the data and constrain all the energy in a single column. Hence, the resultant features for each pose of each person has $X^* \times 6$ dimensions.
- *Technique 2* Is the same as *Tech.1* except that we used Mixing matrix A expressed in Eq. (3.26) instead of using *resultant Independent features* as an alternative way of facial representation [64, 75, 144].
- *Technique 3*
  i. Apply 2D FastICA to each LL sub-band matrix of each partition of each pose as in *Tech.1-i*.

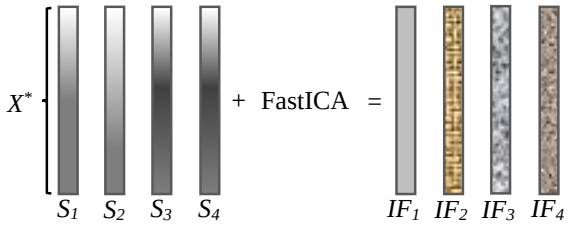


Figure 10.5: Applying FastICA to retained LL sub-band. S_l represents Sub-images of LL sub-band. IF_l represents Independent features. Where $l = 1, 2, 3, and\ 4$.

ii. Choose the IF_l column, see figure 10.5, that has the maximum ℓ_2-Norm for better feature representation and further dimensionality reduction. The dimensions of the resulting matrix features of each pose are the same as dimensions expressed in *Tech.1*.

- *Technique 4*

  i. Apply 2D FastICA to the combined feature matrix of all partitions of LL sub-band, figure 10.4-a, which has $X^* \times 24$ dimensions.

  ii. Find ℓ_2-Norm for each row of the resultant ICA pose matrix for further dimensionality reduction/feature compression. Here, the resultant features matrix for each pose has of $X^* \times 1$ dimensions.

- *Technique 5* It is the same as *Tech.4* except using Mixing matrix A presented in Eq.(3.26) instead of using *resultant Independent features*. As we mentioned, it is considered as an alternative way of representing the input signal [64, 75, 144].

- *Technique 6*

  i. Apply 2D FastICA to the combined feature matrix as performed in *Tech.4*.

  ii. Choose the IF_l column, figure 10.5, that has maximum ℓ_2-Norm for better facial representation and further data compaction. Thus, the resultant features for each pose have $X^* \times 1$ dimensions.

B. Second Representation shown in figure 10.4-b.

- *Technique 1*
  i. Apply 2D FastICA to the matrix resulted from second representation shown in figure 10.4-b that has $X^* \times 6$ dimensions.
  ii. Find ℓ_2-Norm for each row of the resultant features to reduce feature dimensions and concentrate the energy in one column. The resultant features for each pose have $X^* \times 1$ dimensions.
- *Technique 2*
  i. The mixing matrix A, used in Eq. (3.26), resulted form Tech.1 is used here instead of FastICA signal matrix.
  ii. Apply Tech.1-ii to mixing matrix. The dimensions of the resultant feature matrix of each pose are of $X^* \times 1$.
- *Technique 3*
  i. Apply the first step of Tech.1.
  ii. The IF_l column, shown in figure 10.5, that has maximum ℓ_2-Norm is selected. The dimensions of the output features resulted are the same as dimensions of Tech.1 $\&$ Tech.2.

c) Repeat step [1-2] for all poses and for all persons in the databases.

d) The final compacted independent feature, which has dimensions varied from Technique to another according to which database gets involved, is fed to the recognition step for classification, which is based on BPTA.

10.1.3 Recognition

In this phase, the unknown image will be compared with the training images (stored database). Therefore, building a good classifier is a very important step in order to anticipate a good performance and accuracy. Thus, the recognition phase consists of two important modes: Training and Testing Mode.

A NN is a very powerful tool used to classify the input signals. The NN based on BPTA is used in this chapter in both training and testing modes. The BPTA is a supervised learning algorithm, thus it is necessary to choose the desired output for each database used in this chapter. Each desired output for each person should be different from other desired outputs of other persons. However, for the same person the desired output should be the same for all poses used in the training mode and for all partitions. For example, the desired output for person one of YALE database is written as: $[1-1-1-1-1-1-1-1-1-1-1-1-1-1-1]$. Where number (1) represents the output (1) is active and number (-1) represents the output is inactive. Note that the output might not be equal to 1 or -1 but may reach to these values since the activation function used in the BPTA is Hyperbolic Tangent Sigmoid transfer function shown in figure 10.6 [149]. The configuration of the NN has three layers, namely, Input, Hidden, and Output layers. The number of the neurons in the input and the output layers is always fixed according to the dimensions of the input vectors and the number of the persons in the database, while the choice of the number of hidden neurons is somehow flexible.

In the training mode, the NN is configured and the classifier is built using the described training features. Each technique has a classifier different from other techniques since each one has its own resultant features and then each one will have its own results as shown in the experimental results section.

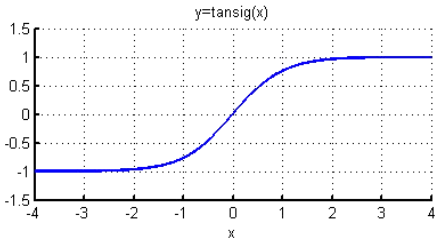


Figure 10.6: Hyperbolic tangent sigmoid transfer function. The activation function is expressed as $tansig(x) = \dfrac{2}{1+e^{-2x}} - 1$ and considered as $tansh(x)$. Its output range is between (-1,1).

In the testing mode, the same procedures of the training mode are employed. First, each test pose is passed through preprocessing steps. Second, 2D DMWT is applied to each partition. Then only LL sub-band is retained and the remaining sub-bands are eliminated. Next, convert all sub-bands to 1-D form. After that all discussed techniques of the two representations are applied to the resultant features. Finally, the resulting matrix features that have dimensions varies from technique to another are fed to NN for classification.

10.1.4 Decision

In this section, we will explain how the decision is made. The following example is presented according to the first representation. Let's assume we have an unknown pose (tested pose) corresponding to the person 1. Since we have 6 different features corresponding to 6 different partitions resulted from Tech.1-Tech.3, each one is tested individually with the training features and assigned to the specific person. Let's assume that the outputs from the NN for these six partitions are (111511), which mean the following:

- (1): The tested features of these partitions are matched with the corresponding trained features and assigned to **person 1**, **which is correct**.

- (5): The tested features of the fourth partition is matched with one of the trained partitions of person 5 and **person 5** is assigned, **which is incorrect**.

In general, if there are more than three tested partitions matched with the corresponding trained features of the same person, the testing pose is correctly matched. In the above example, the tested pose is correctly matched with **person 1**. If there is no majority matched, the testing pose is incorrectly matched. Also, if the partition 1-partition 4 are assigned to the same person, the program will stop and the decision is made since the majority is achieved. This leads to less computational complexity.

10.2 Experimental Results

The experimental results of the proposed techniques based two representations are presented and compared with a large number of the state-of-the-art methods. Five databases, namely, ORL, YALE, FERET, FEI, and LFW that have different facial variations, such as facial expressions, light conditions, rotations, etc. are used to test the proposed techniques. K-fold Cross Validation (CV) is employed to analyze the experimental results. Various values of K are chosen, $K = 2, K = 3, and\ K = 5$. The recognition rates reported in Tables 10.2-10.6 are the average of the rates obtained across the K-folds CV.The NN simulations of the proposed techniques are presented and compared with the existing approaches. The proposed techniques achieved 100% recognition rates when tested with the training poses.

10.2.1 Experimental Results for the ORL Database

The ORL database consists of 10 different poses of each 40 different persons, as mentioned in sub-subsection 1.3.1. Table 10.2 summarizes the results of the proposed techniques based two

representations compared with the state-of-the-art approaches. The proposed techniques improved the recognition accuracy compared with other methods. Also, the proposed techniques achieved higher recognition rates when compared with the results reported in [63, 91, 97, 34, 155, 38, 39].

Table 10.2: Experimental Results for the ORL Database

Proposed Techniques		Other Methods	
First Representation		PCA-CDA [45]	95%
Tech.1	**100%**	RV-LDA [55]	97%
Tech.2	99%	OLPP [62]	93.5%
Tech.3	98.5%	SNMFSPM [78]	98.15%
Tech.4	99.25%	DWT-PCA [81]	96.75%
Tech.5	98.75%	DWT-ICA [83]	97.5%
Tech.6	98.25%	LBPP-DCT [93]	95.5%
Second Representation		GWT-DCT [94]	97.25%
Tech.1	99%	DCT-PCA [95]	96%
Tech.2	98.75%	SADL [102]	97.5%
Tech.3	98.25%	DWT-MPSO [156]	98.33%

10.2.2 Experimental Results for the YALE Database

There are 15 persons in YALE database, each with 11 different poses, as discussed in sub-subsection 1.3.2. Table 10.3 shows the results of the proposed techniques based two representations and the

comparison with the other methods. As before, the results achieved using the proposed techniques were higher than the results reported in the above approaches and in [91, 97, 34, 38, 39].

Table 10.3: Experimental Results for the YALE Database

Proposed Techniques		Other Methods	
First Representation		Fast ℓ_1 PCA [54]	83%
Tech.1	**100%**	OLPP [62]	98.2%
Tech.2	99.17%	LPP [63]	97.14%
Tech.3	98.47%	SNMFSPM [78]	97.69%
Tech.4	99.58%	DWT-PCA [81]	96.55%
Tech.5	98.89%	DWT-ICA [83]	97.25%
Tech.6	98.1%	GWT-DCT [94]	96.95%
Second Representation		DCT-PCA [95]	96.34%
Tech.1	99.17%	SADL [102]	94.67%
Tech.2	98.75%	DWT-MPSO[156]	98.29%
Tech.3	97.18%	NFLS-II [150]	82.42%

10.2.3 Experimental Results for the FERET Database

The FERET database contains 200 individual, each with 11 different poses, as discussed in subsubsection 1.3.3. The results of both representations are summarizes in Table 10.4. The results exhibit the same behavior as in the previous databases when compared with the existing approaches.

Table 10.4: Experimental Results for the FERET Database

Proposed Techniques		Other Methods	
First Representation		OLPP [62]	92.1%
Tech.1	**99.6%**	DWT-PCA [81]	96.4%
Tech.2	98.86%	DWT-ICA [83]	97.33%
Tech.3	98.3%	GWT-DCT [94]	96.8%
Tech.4	99.2%	DCT-PCA [95]	95.7%
Tech.5	98.66%	LDA-SID [100]	98.5%
Tech.6	98.21%	SADL [102]	92.78%
Second Representation		SLF-RKR-ℓ_2 [109]	89.2%
Tech.1	99.05%	LFLM-SIFT [111]	92.4%
Tech.2	98.4%	SLBFLE [112]	97.7%
Tech.3	98%	FACE [157]	95%

Also, the proposed techniques accomplished higher accuracy in ,comparison with other approaches presented in [91, 97, 34, 38, 39].

10.2.4 Experimental Results for the FET Database

The FEI consists of 200 individuals, each with 14 different poses, as mentioned in sub-subsection 1.3.4. The results, which are presented in Table 10.5, of the proposed techniques based two representations demonstrate the same attitudes in comparison with the state-of-the-arts approaches.

Table 10.5: Experimental Results for the FEI Database

Proposed Techniques		Other Methods	
First Representation		OLPP [62]	98.9%
Tech.1	**99.61%**	DWT-PCA [81]	96.25%
Tech.2	99%	DWT-ICA [83]	97.21%
Tech.3	98.5%	DCT [91]	92%
Tech.4	99.4%	GWT-DCT [94]	96.65%
Tech.5	98.83%	DCT-PCA [95]	95.85%
Tech.6	98.23%	VQ-KFCG [97]	93.81%
Second Representation		LFLM-SIFT [111]	85.3%
Tech.1	99.2%	DWT-MPSO [156]	97.08%
Tech.2	98.6%	NFLS-II [150]	93%
Tech.3	98.2%	PCNC [151]	94.17%

10.2.5 Experimental Results for the LFW Database

Labeled Faces in the Wild (LFW) is a large database that contains 13233 images of 5749 persons, see sub-subsection 1.3.5 for more details. Table 10.6 summarizes the results of the proposed techniques compared with the state-of-the-arts methods.

Table 10.6: Experimental Results for the LFW Database

Proposed Techniques		Other Methods	
First Representation		DWT-PCA [81]	94.8%
Tech.1	**98.91%**	DWT-ICA [83]	95.72%
Tech.2	98.06%	GWT-DCT [94]	95.3%
Tech.3	97.33%	DCT-PCA [95]	94.5%
Tech.4	98.54%	CNN-SAE [99]	98.24%
Tech.5	97.87%	LDA-SID [100]	95.65%
Tech.6	97.11%	SLF-RKR-ℓ_2 [109]	81.9%
Second Representation		SLBFLE [112]	84.2%
Tech.1	98.1%	FACE [157]	61%
Tech.2	97.65%	GWT-PCA [158]	96%
Tech.3	96.8%	HPEN-HD-Gabor-JB [159]	95.25%

10.3 Discussion

As we mentioned that the purpose of applying partitioning step in the preprocessing phase was to reduce the effect of unnecessary information, which is common among all poses and persons and to highlight the local features. Also, partitioning faces helps to improve the recognition accuracy since each partition contained local facial features that are common with the other two partitions. As shown in the experimental results, this led to increase the matching probability for each pose and hence the overall recognition accuracy improved.

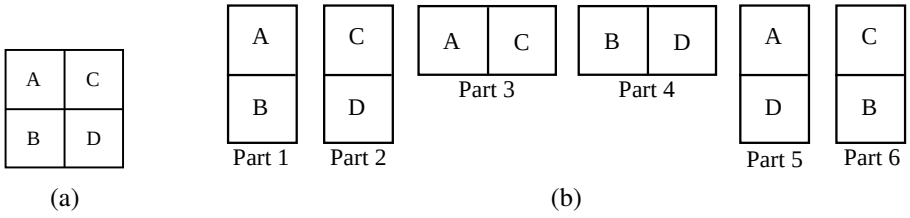


Figure 10.7: Shows input image division and partitioning construction.

In contrast to the DWT, the DMWT has several favorable features through achieving a good reconstruction (orthogonality), better performance (linear phase symmetry), high order of approximation (vanishing moments), and compact support. These desirable features can not be achieved at the same time in the scalar wavelet, while in MWT provide more degree of freedom and give perfect performance in signal and image applications [121]. Based on a number of decomposition levels required and the resulted sub-bands that mathematically expressed as $3 \times L^* + 1$ for DWT and $4 + 12 \times L^*$ for DMWT, DMWT achieves high dimensionality reduction than DWT. Where L^* is the number of decomposition levels required. Therefore, the proposed techniques accomplished less storage requirements compared with storage requirements reported in [81, 82, 83] that used DWT.

Also, the results of the proposed techniques based two representations are higher than the results reported in [111] even though they partitioned the input facial images to n overlapping partitions.These achievements are due to applying efficient integrated tools (2D DMWT $\&$ 2D FastICA) to the partitioned images. The dimensionality reduction and the most important features extracted from facial images were obtained through retaining only the LL sub-band of 2D DMWT. Furthermore, producing independent and more discriminating features were achieved through applying 2D FastICA to the resultant of DMWT representations.

As explained, FastICA has several advantages compared with ICA, such as FastICA has faster convergence than traditional ICA since it does not require to select a step size compared with ICA based gradient algorithm. Also, the independent non-Gaussian signal using FastICA is found by any arbitrary non-linear function $\theta(t)$, while other ICA algorithms require to evaluate the PDF of the selected non-linear function [75, 144]. Therefore, FastICA is more efficient of estimating the statistical components. Hence, the key of using FastICA in this chapter is to find the bases images that are statistically independent. These bases are maintaining the structural information of the facial images. The ICA representations are more robust to the different variations caused by different facial rotations, facial expressions, and illuminations, which can be considered as forms of noises with respect to the original images [64, 68, 67]. Therefore, combining these favorable properties of DMWT and the efficient representations of FastICA led us to achieve a notable dimensionality reduction as well as an improvement in the recognition rates.

Each partition has possessed its own features (local features). Since each partition is in common with other two partitions as seen in figure 10.7, the matching probability for each pose is significantly increased if features of each partition are participated in both training and testing phase. Since Tech.1-3 in the proposed first representation are applied to DMWT features of each partition, there are six resultant feature representations for each pose. After getting these six features of each pose tested with the trained stored features of whole database and after applying voting technique, we can see that the experimental results of Tech.1-3 outperformed other results of Tech.4-6 and Tech.1-3 in the first and second representations, respectively, which were employed to all partitions, whole image, at the same time.

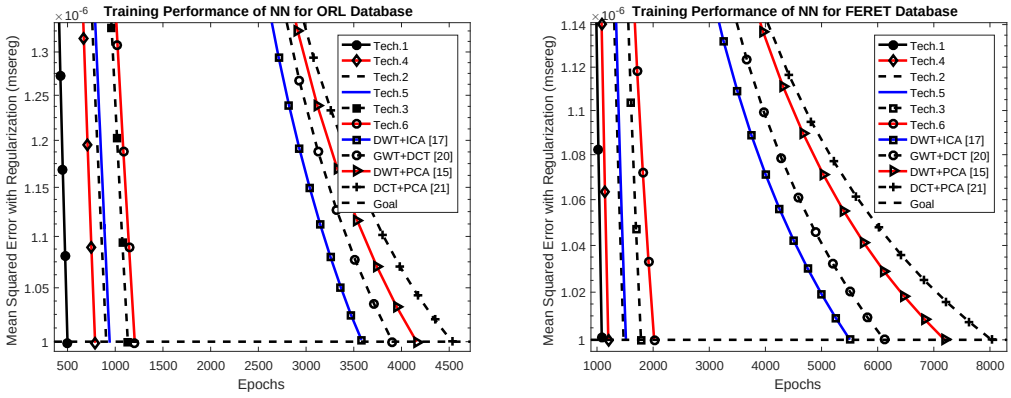
In other words, using the features extracted from each partition (6 features for each pose) in both training and testing mode led the proposed Techniques 1-3 in the first representation to accomplish higher recognition rates compared to those obtained using one matrix (features) to represent the whole pose as in Techniques 4-6 and Techniques 1-3 in the first and second representation, respec-

tively. Figure 10.7 shows how parts are participating with each other. Furthermore, the results of the first representation outperformed the results of the second representation. This is due to the fact that images of faces are efficiently described by the first representation as compared to the second representation. Although the Tech.4-6 and Tech.1-3 in the first and the second representation, respectively are applied to the all combined partitions (whole image), their results illustrate an improvement in the recognition rates when compared with the state-of-the-art approaches.

10.4 Performance of Neural Network

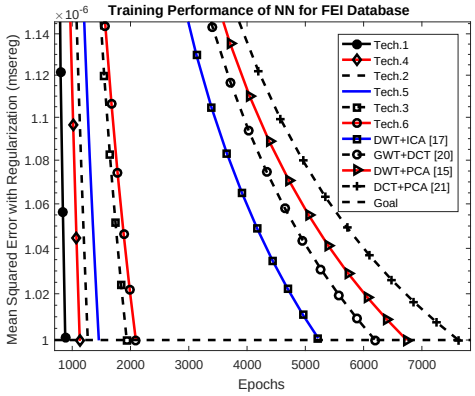
There are several factors that affect the performance of NN in the training mode and lead to impact the overall system accuracy. The first factor is the *Network Complexity*, which depends on the number of hidden layers, the number of neurons in each hidden layer, and the type of the activation functions used for each interconnection weights. The second one is *Learning Complexity*, which depends on choosing training algorithms, initialization parameters, and weight selection. Finally, the *Problem Complexity* that depends on how accurate, efficient, and sufficient the databases used in the training mode.

In this chapter, we used one hidden layer for all five databases. Hence, 256, 1024, 1024, 4069, and 64 neurons were selected for ORL, YALE, FERET, FEI, and LFW databases, respectively. The activation function used is hyperbolic tangent sigmoid shown in figure 10.6. The databases are trained and tested using BPTA. Mean square error with regularization (msereg) is used to calculate the error. Msereg is finding the mean sum of the square error between the actual output and the target. Figure 10.8 shows samples of the performance of the NN in the training mode of all discussed techniques based on the first representation in comparison with some of the state-of-the-arts approaches [81, 83, 94, 95].



(a) ORL

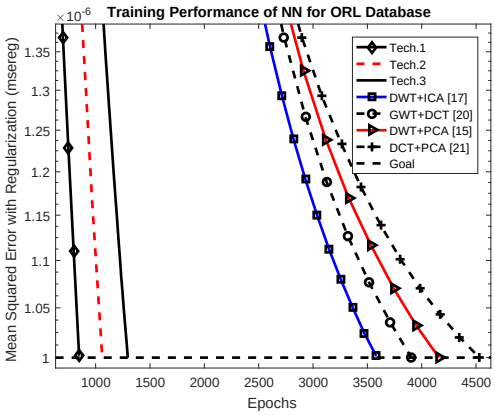
(b) FERET



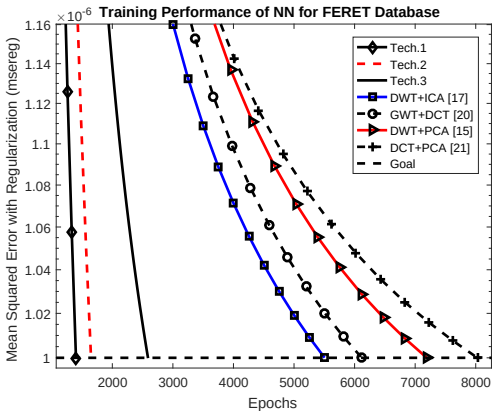
(c) FEI

Figure 10.8: NN performance for the proposed techniques compared with the other approaches based first representation.

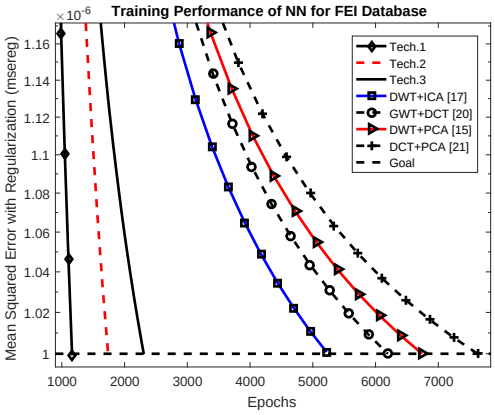
It is obvious from figure 10.8, the proposed Tech.1-3 outperformed and achieved at the goal faster than Tech.4-6. In addition, the performance of first representation is attained the goal faster than the second representation that shown in figure 10.9. In addition, all techniques based the two representations outperformed some existing approaches [81, 83, 94, 95]. This is due to the fact that partitioning step and employing the two efficient tools (DMWT and FastICA) in the feature extraction steps produced features more efficient than those obtained by the existing methods. Hence, the NN convergence is enhanced and the recognition rates are improved.



(a) ORL



(b) FERET



(c) FEI

Figure 10.9: NN performance for the proposed techniques compared with the other approaches based second representation.

10.5 Conclusion

A new approach applying 2D DMWT followed by FastICA to partitioned facial images for face recognition was proposed in this chapter. Dimensionality reduction, efficient feature extraction, and as a consequence, high recognition rates were accomplished in this contribution. The disadvantages and shortcomings of DWT and ICA and other methods have been taken into account

while considering the combination of DMWT and fastICA. The facial images were Partitioned into six parts in order to reduce the effect of the mutual information found among all persons in the databases on the overall system performance, and to highlight the local features. 2D DMWT was applied to each part and only the LL sub-band is retained and all other sub-bands were eliminated. As a consequence, the dimensionality reduction (less storage requirements) was accomplished. Two representations were constructed using DMWT features. Splitting the feature extraction into the two representation methods and the use of FastICA followed by the ℓ_2-norm helps to retrieve the useful information from the images and discards the non-discriminating features.

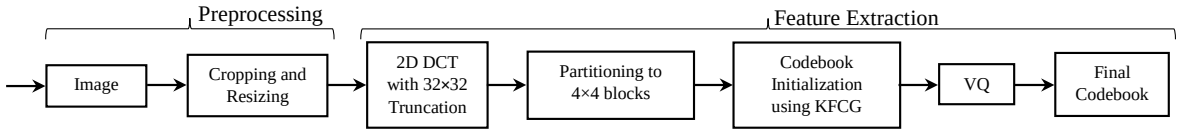
The dimensionality reduction proposed in the system helps in reducing the computational burden and improving the accuracy. The combination of the feature extraction methods (2D DMWT and FastICA followed by ℓ_2-norm) led to more discriminating features being extracted and further led to faster convergence. The final features extracted using the proposed techniques that fed to NN for training/testing have dimensions of $X^* \times Y^*$, where X^* is 256, 1024, 4096, and 64 dimensions for ORL, YALE $\&$ FERET, FEI, and LFW databases, respectively. Y^* is either 6 or 1 according to which technique gets involved and which representation was employed. The proposed system was extensively evaluated using five databases that have different facial variations, such as illuminations, facial expressions, rotations, facial details, etc. The experimental results were analyzed using different values of K-fold CV. The proposed techniques based on the two representations were not only achieving high recognition rates compared to the other methods, shown in Tables 10.2-10.6, but also having faster NN convergence than the other approaches, as shown in figure 10.8 and figure 10.9.

CHAPTER 11: EMPLOYING VECTOR QUANTIZATION ALGORITHM IN A TRANSFORM DOMAIN FOR FACIAL RECOGNITION

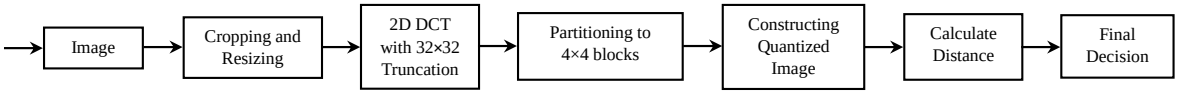
A Vector Quantization (VQ) algorithm in the Discrete Cosine Transform (DCT) domain is proposed for facial recognition in this chapter [42]. There are three main phases in the proposed system, namely, Preprocessing, Feature Extraction, and Recognition. Cropping and choosing an appropriate dimension are performed in the preprocessing step. Then, DCT with appropriate truncation dimensions is applied to the processed faces for dimensionality reduction. For further feature compaction, VQ algorithm employing Kekre Fast Codebook Generation (KFCG) approach for codebook initialization is applied to the transformed truncated features. Finally, the proposed system is extensively evaluated using four different databases, namely, ORL, YALE, FERET, and FEI that have different facial variations, such as illuminations, rotations, facial expressions, etc. Euclidean distance criterion is used to calculate the recognition rates. Then, the results are analyzed using K-fold Cross Validation (CV). The proposed approach is shown to improve the recognition rates as well as the storage requirements in comparison with some of the existing state-of-the arts approaches.

11.1 Proposed System

The proposed system, shown in figure 11.1, consists of three main phases: Preprocessing, Feature Extraction, and Classification. In this chapter, the design parameters chosen are: codeword (centroid) size 4×4 ($l \times w$), codebook size $L = 16$, MSE for distortion criterion, and KFCG method for codebook initialization.



(a) Preprocessing and Feature Extraction Phase



(b) Classification Phase

Figure 11.1: Proposed system. Figure 11.1-a shows the Preprocessing and the Feature extraction phases. Figure 11.1-b illustrates the classification phase.

11.1.1 Preprocessing

It consists of two steps:

a) All databases are cropped except ORL database in order to reduce the effect of irrelevant information (background) on the overall performance of the proposed system. Table 12.1 shows the dimensions of the databases after cropping. Samples of cropped databases are shown in figure 11.2-B.

Table 11.1: The Dimensions of the Databases

Databases	Actual Size	Dimensions After Cropping	Proposed Dimensions
ORL	112×92	112×92	128×128
YALE	243×320	160×150	128×128
FERET	384×256	220×170	128×128
FEI	480×640	320×290	128×128

b) In this step, the dimensions of the different databases used are converted to appropriate ones as shown in Table 11.1. This dimension is chosen since the algorithms used in this chapter required dimensions based on power of two.

11.1.2 Feature Extraction

Human faces have large dimensions that increase the computational complexity and hence affect the overall performance of the recognition system. Therefore, the goal of this step is to extract efficient facial features that lead to less storage requirements and thus improve the overall system performance.

First, the 2D DCT is applied to the processed facial image for better facial representation, shown in figure 11.3. Since most of the facial energy is concentrated in the low frequency region, a window of 32×32 dimensions is chosen for efficient feature extraction and dimensionality reduction. Thus, the extracted feature matrix has 32×32 dimensions for all databases. Then, the retained features are fed to VQ for further feature compaction and data concentration around the centroids.

To perform the VQ process, the extracted feature matrix of each pose that has 32×32 dimensions is partitioned into non-overlapping blocks, each with 4×4 ($l \times w$) dimensions. The first mean of the matrix, that also has dimensions of 4×4 as shown in figure 11.4, is calculated by averaging all blocks as expressed in the following Eqs.:

$$Y = (64) = \frac{proposed\ dimensions}{codeword\ dimensions} \tag{11.1}$$

$$X_i = \frac{x_i^y + x_i^y + \ldots + x_i^Y}{Y} \tag{11.2}$$

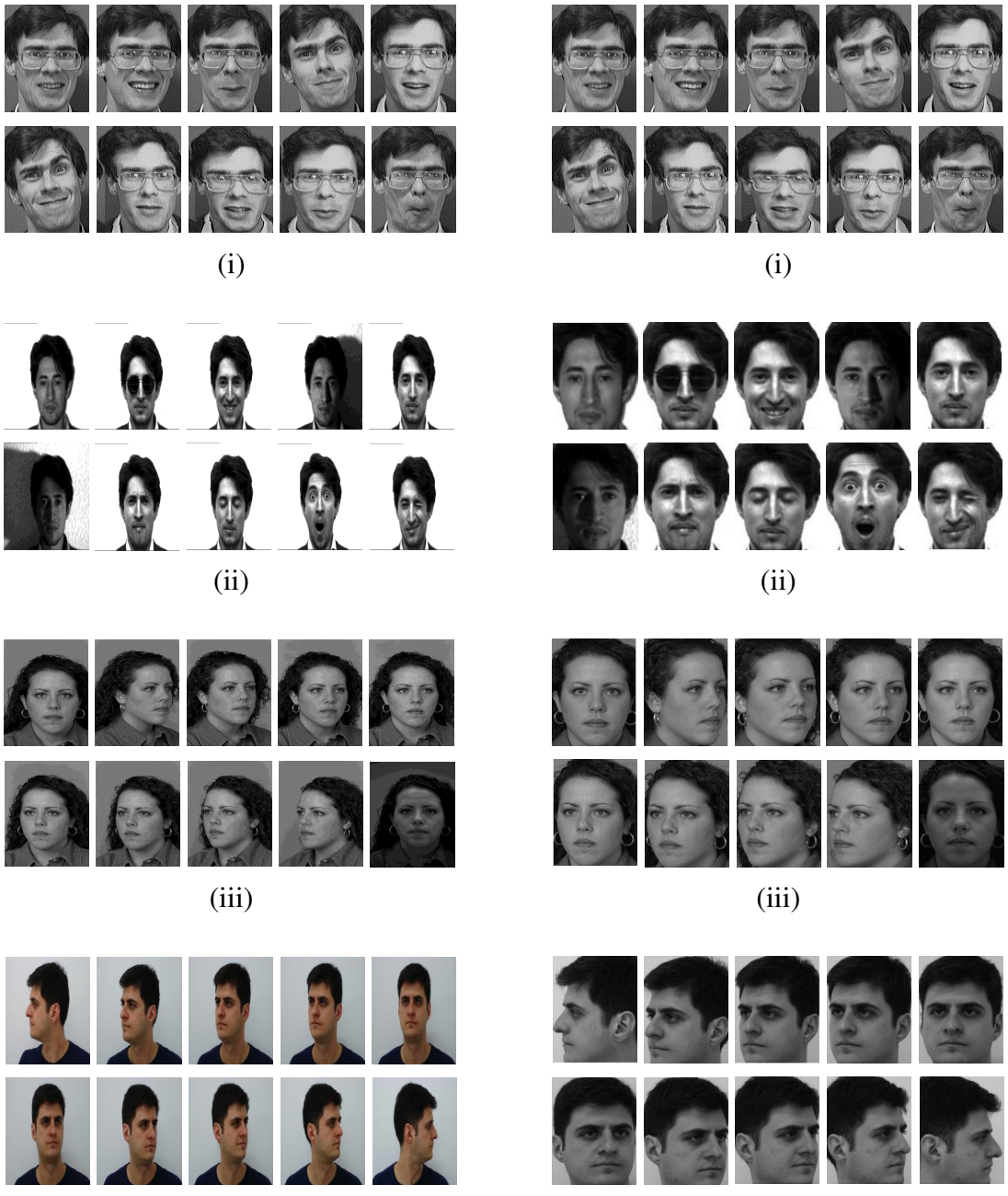


Figure 11.2: Figure 11.2-A shows 5 original samples of one person for all different databases. Figure 11.2-B shows samples after cropping.

(a) Original pose (b) DCT results

Figure 11.3: Shows the original pose and its DCT version

Where X_i is the i^{th} average of all corresponding elements across all blocks. $i \in Z : \{i = 1, 2, \ldots, l \times w\}$. x is the i^{th} element in the y^{th} block, $y \in \{1, 2, \ldots, Y\}$, Y is the total number of blocks.

Then, the first initial VQ codebook is generated using KFCG method from the first mean. All blocks of the facial image are divided into two groups according to their values as compared to each corresponding element in the mean. After $\log_2 L = 4$ splits, the initial codebook is constructed. Then, the LBG algorithm is applied to calculate the new codebook. Each new codeword (centroid) is the average of all the image blocks that were encoded using that codeword. The codebook calculation steps are repeated for all the training poses of each subject in the database. The final feature matrix extracted from each pose of each person and for all databases used has $4 \times 4 \times 16$ dimensions (16 centroids for each pose).

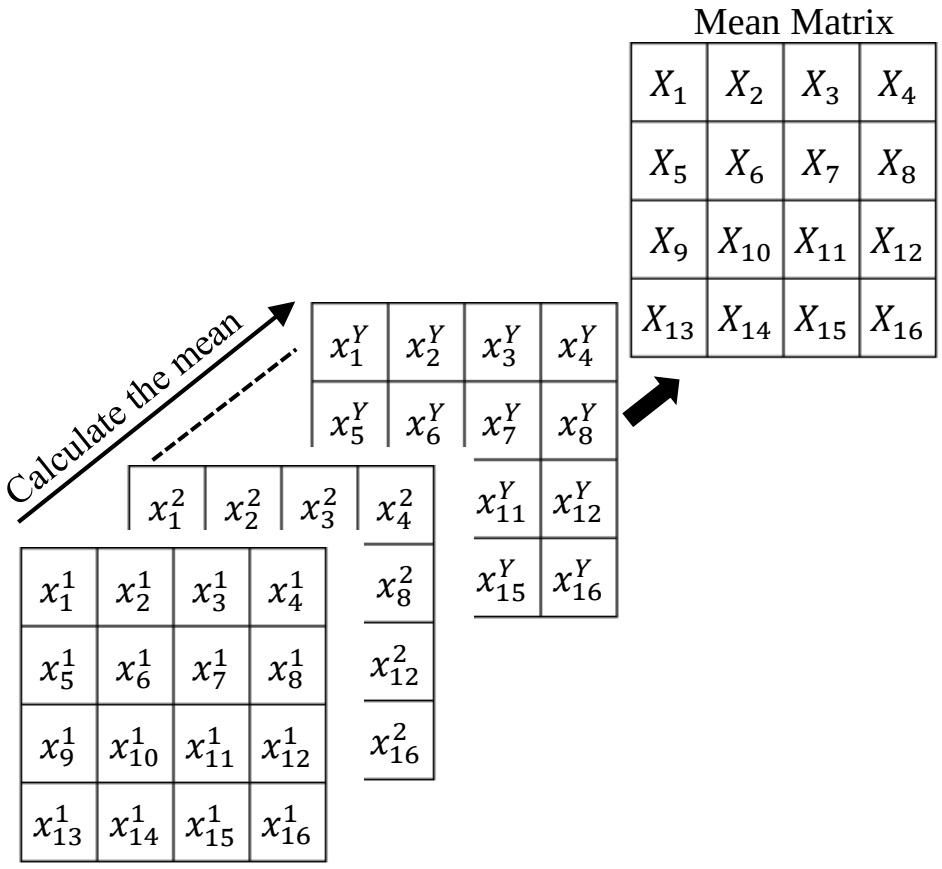


Figure 11.4: Calculating the first mean of the retained matrix of each pose.

11.1.3 Classification

The same preprocessing steps, cropping technique and choosing an appropriate dimensions, are applied to the test image. All codebooks from all subjects are used to reconstruct the quantized versions of the processed test image. The Euclidean distances, between input image and the reconstructed ones, are calculated. The person whose codebook has the minimum distance is declared as the correct one. All poses that are correctly identifying each person constitute the final recognition rate. The recognition rate is measured as $\frac{M}{T} \times 100\ \%$, where M is the total number of poses correctly matched and T is the total number of poses in the database.

11.2 Experimental Results

The results of the proposed system are presented in this section. The proposed approach is evaluated using four databases, namely, ORL, YALE, FERET, and FEI that have different facial configurations, such as facial expressions, illuminations, rotations, etc. The results are analyzed using K-fold CV. The recognition rates reported in the tables are the average of the rates obtained across the K-folds CV. The average of the recognition rates obtained across the k-folds CV is reported. The proposed system achieved 100% recognition rates when tested with the training poses. The proposed approach is compared with the some of the existing methods. This work was done using MATLAB 2016a Student Version. Samples for each database are shown in figure 11.2.

11.2.1 Experimental Results for the ORL Database

The ORL database is widely used in face recognition. There are 40 persons each with 10 different poses [28]. Table 11.2 shows the results of the proposed system compared with the other approaches. As shown in Table 11.2, increasing the value of K results in improving the recognition rates.

Table 11.2: Experimental Results for the ORL Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95%**	89%	88%	92.5%
K=3	**96.81%**	92.3%	91.88%	94.75%
K=5	**98.25%**	94.75%	94.25%	96.5%

11.2.2 Experimental Results for the YALE Database

In the YALE database, there are 15 subjects each with 11 different poses [29]. As in ORL database, the results of the proposed approach shown in Table 11.3 exhibit the same behavior.

Table 11.3: Experimental Results for the YALE Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.78%**	88.33%	87.33%	92.12%
K=3	**96.81%**	91.4%	90.56%	93.94%
K=5	**98.47%**	94.17%	93.1%	95.15%

11.2.3 Experimental Results for the FERET Database

It consists of 200 persons each with 11 different poses [30, 31]. As shown in Table 11.4, the proposed approach accomplished high recognition rates compared with other approaches.

11.2.4 Experimental Results for the FEI Database

This is a colored database. The gray-scale version of this database is used in this contribution. It has 200 subjects with 14 different poses for each subject. The facial images have 180 degree rotation with distinct appearance, hairstyle, and garnishes [32]. The proposed method in Table 11.5 demonstrate the same performance as the preceding databases.

Table 11.4: Experimental Results for the FERET Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.1%**	88.21%	87.5%	92.5%
K=3	**96.47%**	91.23%	90.51%	94.86%
K=5	**97.91%**	93.57%	92.21%	95.68%

Table 11.5: Experimental Results for the FEI Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.18%**	87.5%	87.32%	92.25%
K=3	**96.32%**	90.31%	90.33%	93.92%
K=5	**97.81%**	93.81%	92%	95.58%

11.3 Discussion

The recognition rates of the proposed system, shown in Tables 11.2, 11.3, 11.4, and 11.5 are higher than the rates obtained in [90], [97], and [91]. In this contribution, integrating 2D DCT and VQ in the manner proposed results in an enhancements in image feature selection and, thus, the recognition rates compared with [90], [97], and [91].

11.4 Conclusion

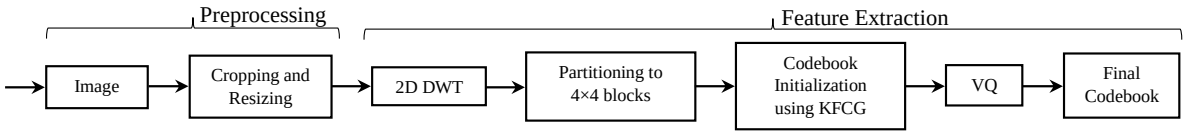
A face recognition system based on combining 2D DCT/VQ was proposed. Each pose in the database was represented more efficiently in the transform domain than in the spatial domain. A window of size 32×32 was selected for dimensionality reduction and feature selection. Then, the VQ based on KFCG for codebook initialization was employed to the retained features to accomplish further data compaction. The final feature matrices (centroids) were of size $4 \times 4 \times 16$ (256), while the processed facial image were 128×128 (16384) dimensions. The recognition rates accomplished using the proposed approach were higher than those reported in [97, 91], and [90]. The results were analyzed using K-fold CV. The proposed approach was extensively evaluated using four databases, namely, ORL, YALE, FERET, and FEI. The recognition rates accomplished, for example K=5, were **98.25%**, **98.47%**, **97.91%**, and **97.81%** for ORL, YALE, FERET, and FEI databases, respectively.

CHAPTER 12: EFFICIENT FACIAL RECOGNITION USING VECTOR QUANTIZATION OF 2D DWT FEATURES

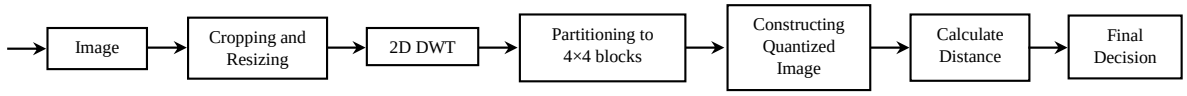
A new approach for facial recognition employing Two Dimensional Discrete Wavelet Transform (2D DWT) and Vector Quantization (VQ) is proposed in this chapter [43]. preprocessing, feature extraction, and classification are the three main phases in this chapter. A cropping technique and appropriate dimensions selection are employed in the preprocessing step. In the feature extraction step, 2D DWT is applied to the processed facial images for dimensionality reduction and feature extraction. Then the VQ algorithm using Kekre Fast Codebook Generation (KFCG) for codebook initialization is implemented to the resultant DWT features for further feature compression and better facial representation. The proposed algorithm is evaluated using four databases, namely, ORL, YALE, FERET, and FEI that have different facial variations, such as facial expressions, illumination, rotations, etc. Then the experimental results are analyzed using K-fold Cross Validation (CV). The results show that the proposed approach improves the recognition rates and reduces the storage requirements compared with existing methods.

12.1 Proposed System

There are three main phases in the proposed system, namely, Preprocessing, Feature Extraction, and Recognition as shown in figure 12.1. The First phase uses a cropping technique and then resizing the facial images to the appropriate dimensions. Dimensionality reduction and feature selections are accomplished in the second phase. Finally, the Euclidean distance is used to recognize the identity of the persons in the set of different databases. The databases used to evaluate the proposed system are ORL, YALE, FERET, and FEI that have different facial variations, such as illuminations, rotations, facial expressions, etc. Samples of all databases are shown in figure 11.2-A.



(a) Preprocessing and Feature Extraction Phase



(b) Classification Phase

Figure 12.1: Proposed system. Figure 12.1-a shows the Preprocessing and the Feature extraction phases. Figure 12.1-b illustrates the classification phase.

An exhaustive search had been done to select VQ parameters $C, q, \& p$ since the value of each one is totally data dependent. In this chapter, the design parameters chosen are: $C = 16$, $q = p = 4$, distortion criterion used is MSE, and codebook initialization method is based on KFCG.

12.1.1 Preprocessing

It consists of two steps:

a) All databases are cropped except ORL database to reduce the effect of irrelevant information (background) on the overall performance of the proposed system. Table 12.1 shows the dimensions of the databases.Samples of cropped databases are shown in figure 11.2-B.

b) Images are resized to power of two dimensions, see Table 12.1. This dimension is chosen since the algorithms used required dimensions based on power of two.

Table 12.1: The Dimensions of the Databases

Databases	Actual Size	Dimensions After Cropping	Proposed Dimensions
ORL	112×92	112×92	128×128
YALE	243×320	160×150	128×128
FERET	384×256	220×170	128×128
FEI	480×640	320×290	256×256

12.1.2 Feature Extraction

Despite the fact that there are several samples for each subject, the facial recognition considered as a challenging task in several areas due to diversities of the facial images of the subjects in facial expressions, illuminations, obstacles, makeups, rotations, etc. There is lot of irrelevant information in the facial images since the human faces are not rigid bodies. Furthermore, human faces have large dimensions that increase the computational complexity, and hence affect the overall performance of the recognition system.Therefore, Feature Extraction plays a crucial role in the face recognition system. Ideally, the goal of the feature extraction step is to eliminate unnecessary information/dimensionality reduction without affecting the system performance. This leads to reduce the computational complexity, achieve less storage requirements, and improve the recognition rates. As a consequence, the subsequent algorithms are employed to obtain distinct and

a) The 2D DWT based on Multi-resolution Analysis (MRA) is used for:

   A. Dimensionality compaction.

   B. Noise reduction.

C. localizing most of the facial energy in one single band called LL sub-band.

b) The VQ algorithm is used for:

A. Data Compression [134] and [160].

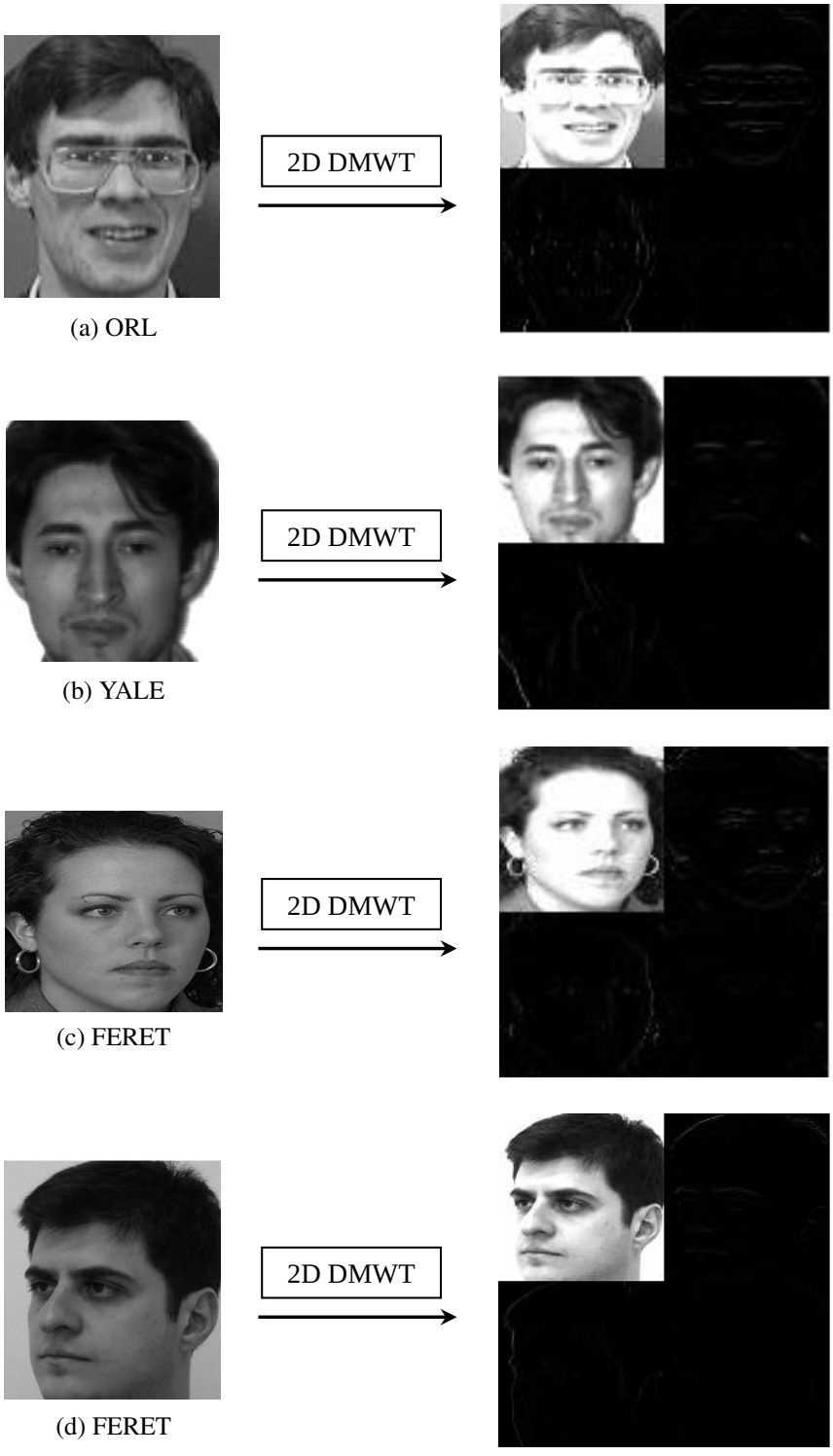
B. facial image representation by using its codebook as a feature.

First, the one level of 2D DWT decomposition is applied to the processed facial images. The image is divided into four frequency subbands, namely LL, LH, HL, and HH as shown in figure 3.1. As mentioned, the L and H are corresponding to the low pass and high pass filters, respectively. Hence, figure 12.2 is shown the resultant of applying the first level of DWT decomposition onto the input facial image. Each sub-band has 64×64 dimensions for ORL, YALE, and FERET databases and 128×128 dimensions for FEI database. As shown in figure 12.2, most of the face energy is localized in the upper left sub-band, which corresponds to the low-low (LL) frequency sub-band of the 2D DWT. Therefore, the LL frequency sub-band is retained, and all other sub-bands are eliminated. Thus, the resultant extracted feature matrix has 64×64 dimensions for ORL, YALE, and FERET databases and 128×128 dimensions for FEI database. The retained features are fed to VQ for further feature compression and data concentration around the centroids.

Second, to perform the VQ process, the extracted feature matrix is partitioned into non-overlapping blocks, each with 4×4 ($p \times q$) dimensions. The first mean of the matrix, that also has dimensions of 4×4, is calculated by averaging all blocks as shown in figure 11.4 and expressed in the following Eqs.:

$$Y = (1024) = \frac{proposed\ dimensions}{codeword\ dimensions} \tag{12.1}$$

$$X_i = \frac{x_i^y + x_i^y + \ldots + x_i^Y}{Y} \tag{12.2}$$



(a) ORL

(b) YALE

(c) FERET

(d) FERET

Figure 12.2: Application of 2D DMWT on different databases.

Where X_i is the i^{th} average of all corresponding elements across all blocks. $i \in Z : \{i = 1, 2, \ldots, p \times q\}$. x is the i^{th} element in the y^{th} block, $y \in \{1, 2, \ldots, Y\}$, Y is the total number of blocks.

Finally, the KFCG method is used to generate the initial VQ codebook from the first mean. All blocks are divided into two groups according to their values as compared to each corresponding element in the mean. After $\log_2 C = 4$ splits, the initial codebook is constructed. Then, the LBG algorithm is applied to calculate the new codebook. Each new codeword is the average of all the image blocks that were encoded using that codeword. The codebook calculation steps are repeated for all the training poses for each individual in the database. The final feature matrix extracted from each pose of each person and for all databases used has $4 \times 4 \times 16$ dimensions.

12.1.3 Classification

The same preprocessing steps, cropping technique and choosing appropriate dimensions, are applied to the test image. All codebooks from all subjects are used to reconstruct quantized versions of the processed test image. The Euclidean distances, between input image and the reconstructed ones, are calculated. The person whose codebook has the minimum distance is declared as the correct one. All poses that are correctly identifying each person are constitute the final recognition rate. The recognition rates are measured as:

$$Recognition\ Rates = \frac{Total\ Number\ of\ Poses\ Correctly\ Matched}{Total\ Number\ of\ Poses\ in\ The\ Dtatabase} \times 100\% \qquad (12.3)$$

12.2 Experimental Results

The results of the proposed algorithm are presented in this section. The proposed system is compared with the algorithms recently reported in [97], [80], and [81]. K-fold Cross Validation (CV) is used to analyze the results. As shown in Tables 12.2, 12.3, 12.4, and 12.5, various values of K are selected. The proposed approach is extensively evaluated using four databases that have different facial variations, such as facial expressions, illuminations, rotations, etc. Samples for each database are shown in figure 11.2.

12.2.1 Experimental Results for the ORL Database

The ORL database is widely used in the area of face recognition. There are 40 persons each with 10 different poses [28]. Table 12.2 summarizes the results of the proposed system with the comparison with the other methods.

Table 12.2: Experimental Results for the ORL Database

K-Fold	Rates of the Proposed System	VQ-KFCG [97]	DWT [80]	DWT&PCA [81]
K=2	**97.5%**	89%	90.75%	92.5%
K=3	**98.17%**	92.3%	92.3%	94.75%
K=5	**98.9%**	94.75%	94%	96.75%

As shown in Table 12.2, even when number of poses used in the raining mode is small (K=2), the proposed approach achieved higher recognition rates than those reported in other approaches.

12.2.2 Experimental Results for the YALE Database

In the YALE database, there are 15 subjects each with 11 different poses [29]. As previously, the results of the proposed system shown in Table 12.3 outperformed the other techniques.

Table 12.3: Experimental Results for the YALE Database

K-Fold	Rates of the Proposed System	VQ-KFCG [97]	DWT [80]	DWT&PCA [81]
K=2	**96.95%**	88.33%	90.33%	92.12%
K=3	**97.52%**	91.4%	92.22%	93.94%
K=5	**98.78%**	94.17%	93.89%	96.55%

12.2.3 Experimental Results for the FERET Database

There are 200 persons in FERET database, each with 11 different poses [30, 31]. In Table 12.4, the proposed approach accomplished higher results than the other approaches.

Table 12.4: Experimental Results for the FERET Database

K-Fold	Rates of the Proposed System	VQ-KFCG [97]	DWT [80]	DWT&PCA [81]
K=2	**96.1%**	88.21%	88.92%	92.5%
K=3	**97.12%**	91.23%	91.3%	94.86%
K=5	**98.45%**	93.57%	93.33%	96.4%

12.2.4 Experimental Results for the FEI Database

This is a colored database. It has 200 subjects, each with 14 poses [32]. As the preceding results, the discussed approach achieved high accuracies compared with the other methods, see Table 12.5.

Table 12.5: Experimental Results for the FEI Database

K-Fold	Rates of the Proposed System	VQ-KFCG [97]	DWT [80]	DWT&PCA [81]
K=2	**96.38%**	87.5%	88.21%	92.25%
K=3	**97.1%**	90.31%	90.78%	93.92%
K=5	**98.32%**	93.81%	93.19%	96.25%

12.3 Discussion

As shown in Tables 12.2-12.5, recognition rates of the proposed approach are higher than the accuracies obtained using [97], [80], and [81]. This is due to the wavelet transform concentrate most of the facial image energy in one single subband corresponding to the LL frequency subband. Hence, the wavelet domain features are efficiently representing the facial images than the spatial features. In this work, it was shown that combining the DWT and VQ in the manner proposed results in an overall enhancement in image feature extraction and, consequently, the recognition accuracy compared with [97], [80], and [81].

12.4 Conclusion

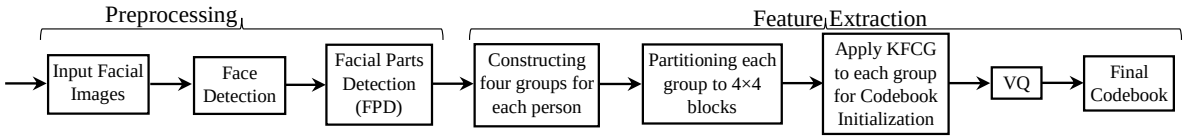
A face recognition system based on combining 2D DWT and VQ was discussed. 1-level of 2D DWT decomposition was applied to the input facial images. Only LL frequency sub-band is retained and all other sub-bands are eliminated. Then, VQ, based on KFCG for codebook initialization, was employed on the retained LL sub-band to achieve further feature compression, which led to achieve less storage requirements. While the processed facial images were either 128×128 (16384) dimensions or 256×256 (65536) dimensions depend on which database is employed , the final feature matrices (centroids) were of dimensions $4 \times 4 \times 16$ (256). The recognition accuracies achieved using the proposed approach were improved in comparison with those reported in [97], [80], and [81]. The performance of the proposed algorithm was analyzed using K-fold CV. Recognition accuracies, thus realized, were 98.9%, 99.1%, 98.45%, and 98.31% for ORL, YALE, FERET, and FEI databases, respectively.

CHAPTER 13: EMPLOYING VECTOR QUANTIZATION ON DETECTED FACIAL PARTS FOR FACE RECOGNITION

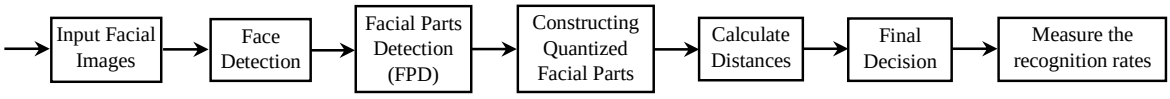
Facial Parts Detection (FPD) approach in conjunction with Vector Quantization (VQ) algorithm are proposed in this chapter for face recognition [44]. There are three phases in the proposed system, namely, Preprocessing, Feature Extraction, and Classification. Detecting facial parts, which are nose, both eyes, and mouth, and choosing appropriate dimensions for each part are done in the preprocessing phase. In the feature extraction phase, four groups for each person, one group for each detected part, are constructed for dimensionality reduction and feature discrimination by considering all parts of all training poses. For further data compaction, VQ algorithm employing Kekre Fast Codebook Generation (KFCG) approach for codebook initialization is applied to each of the four groups. Finally, Euclidean distance criterion is used to obtain the recognition rates. Four databases, namely, ORL, YALE, FERET, and FEI that have different facial variations, such as illuminations, rotations, facial expressions, etc. are used to evaluate the proposed system. Experimental work is performed to evaluate the performance of the proposed technique and the other approaches. Then, K-Fold Cross Validation (CV) is used to analyze the results. The proposed system consistently improved the recognition rates as well as the storage requirements.

13.1 Proposed System

The proposed system, shown in figure 13.1, consists of three main phases: preprocessing, feature extraction, and classification. The design parameters chosen are: codeword (centroid) size 4×4 ($l \times w$), codebook size $L = 16$, MSE for distortion criterion, and KFCG method for codebook initialization. The proposed system is evaluated using four databases, namely, ORL, YALE, FERET, and FEI. Samples of the databases are shown in figure 11.2.



(a) Preprocessing and Feature Extraction Phase



(b) Classification Phase

Figure 13.1: Proposed system. Figure 13.1-a shows the Preprocessing and the Feature extraction phases. Figure 13.1-b illustrates the classification phase.

13.1.1 Preprocessing

It consists of four steps:

a) The first step aims to detect faces from the images using Face Detection Algorithm that employs Viola-Jones algorithm. This step is done using *vision.CascadeObjectDetector* MATLAB 2016a. Samples of detected faces are shown in figure 13.2-(c & g).

b) FPD algorithm is applied to the resultant of step 1 to detect facial parts, which are Left Eye, Right Eye, Mouth, and Nose. Figure 13.2-(d & h) show an example of detected facial parts.

c) It is not necessary that the detected facial parts of the same person have equal dimensions. Table 13.1 shows dimensions of the detected faces and facial parts for the two poses of the first person of YALE database shown in figure 13.2.

Figure 13.2: Figure 13.2-(a & e) show the original poses. Figure 13.2-(b & f) illustrate the detected faces and facial parts. Figure 13.2-(c & g) represent the detected faces of the two poses. Detected facial parts of the two poses are displayed in figure 13.2-(d & h).

d) The fourth step aims to Unify the dimensions among all parts. The dimensions of the detected face and parts (Left Eye, Right Eye, Mouth, and Nose) are converted to the appropriate one as shown in the Table 13.2. These dimensions have been chosen after examining all dimensions of the detected parts of all poses in the databases.

Table 13.1: The Dimensions of the Detected Faces/Facial Parts

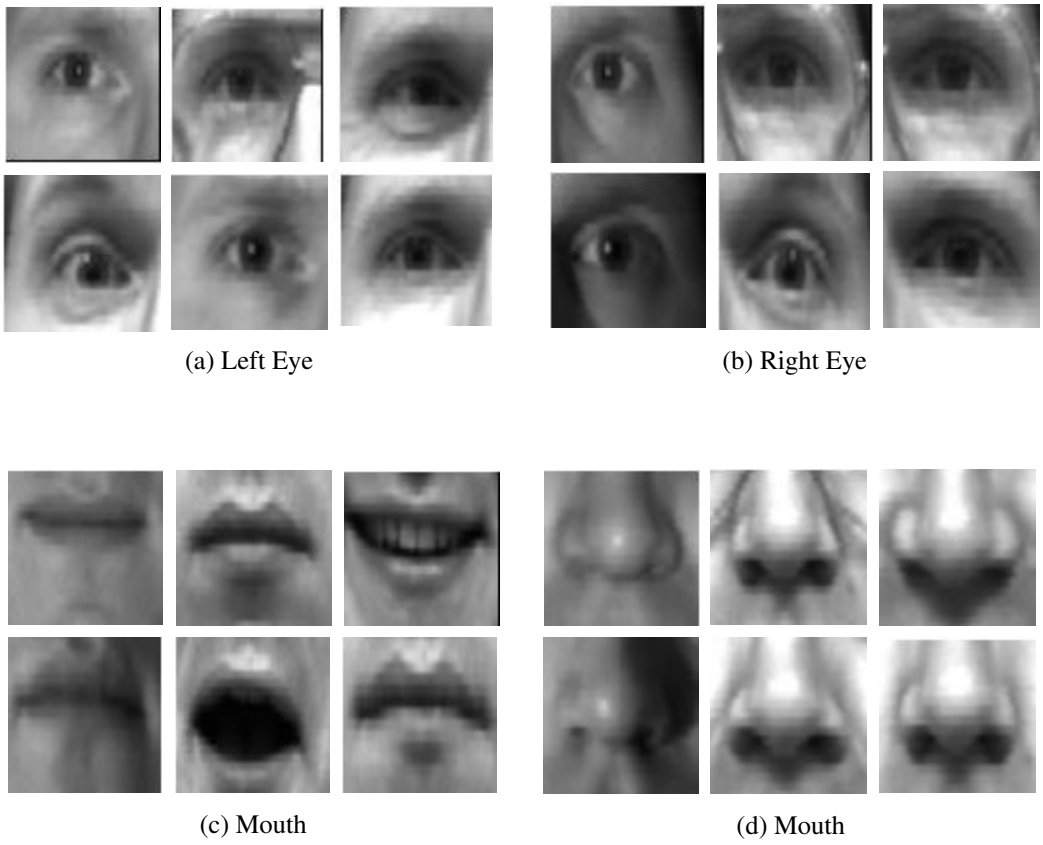
Person 1	Faces and Facial Parts Detected				
	Face	Left Eye	Right Eye	Mouth	Nose
Pose 1	146×146	38×55	36×53	39×64	37×44
Pose 2	153×153	38×56	38×56	40×65	41×49

Table 13.2: Proposed Dimensions of the Detected Faces/Facial Parts

Databases	Actual Size	Faces and Facial Parts Detected				
		Face	Left Eye	Right Eye	Mouth	Nose
ORL	112×92	64×64	32×32	32×32	32×32	32×32
YALE	243×320	128×128	32×32	32×32	32×32	32×32
FERET	384×256	128×128	32×32	32×32	32×32	32×32
FEI	480×640	256×256	64×64	64×64	64×64	64×64

13.1.2 Feature Extraction

Human faces have large dimensions that lead to increase the computational complexity and hence affect the overall performance of the recognition system. Therefore, the goal of the feature extraction is to extract efficient facial features by eliminating the irrelevant information from the facial images. This leads to achieve data compaction, less computational complexity, and hence improve the recognition rates.



(a) Left Eye (b) Right Eye

(c) Mouth (d) Mouth

Figure 13.3: Constructing four groups using six training poses of the first person of YALE database.

In the training mode, figure 13.1-a, 4 groups for each person, one group for each facial part, are established using different poses. Figure 13.3 shows an example of constructing 4 groups using *6 Poses* of *Person 1* of YALE database. As shown in Table 13.2, each detected part has proposed dimensions of 32×32. In each group shown in figure 13.3, there are 6 facial segments corresponding to 6 different poses. Each group is belonging to one facial part. The groups are arranged as follow: Let assume the number of the poses used in the training mode is P, and $D = \sqrt{P}$. If D is an integer number, the radicand P is called a perfect square, the groups are arranged as a square, otherwise they arranged as rectangular. Therefore, each group in figure 13.3

has 96×128 dimensions. In general, each group has dimensions of $N \times M$, where N and M depend on the number of training poses used, and N=M when P is a perfect square.

Obtaining new efficient features from the original one (groups) further reduces the dimensions of the extracted features, gets efficient representation of the facial parts, achieves good performance, and then obtains better results. Therefore, VQ is employed on each group for further feature compaction and data concentration around the centroids.

VQ algorithm is applied to each group individually to obtain one centroid. To perform VQ algorithm on one group, the extracted feature matrix representing in each group that has $N \times M$ dimensions is partitioned into non-overlapping blocks, each with 4×4 ($l \times w$) dimensions. The first mean of the matrix, that also has dimensions of 4×4 as shown in figure 11.4, is calculated by averaging all blocks as expressed in Eqs.:

$$Y = \frac{group\ dimensions}{codeword\ dimensions} \tag{13.1}$$

$$X_i = \frac{x_i^y + x_i^y + \ldots + x_i^Y}{Y} \tag{13.2}$$

Where X_i is the i^{th} average of all corresponding elements across all blocks. $i \in Z : \{i = 1, 2, \ldots, l \times w\}$. x is the i^{th} element in the y^{th} block, $y \in \{1, 2, \ldots, Y\}$, Y is the total number of blocks.

Then, KFCG method is used to generate the first initial VQ codebook form the first mean. All blocks of each group are divided into two groups according to their values as compared to each corresponding element in the mean. After $\log_2 L = 4$ splits, the initial codebook is constructed. Then, the LBG algorithm is applied to calculate the new codebook. Each new codeword (centroid)

is the average of all the image blocks that were encoded using that codeword. The codebook calculation steps are repeated for all groups of each subject in the database. Since each person in all databases used has 4 groups as shown in figure 13.3, there are 4 *Centroids* for each person regardless the number of poses used in the training mode. Therefore, the final feature matrix extracted from each person has dimensions of $4 \times Centroids$ $(4 \times 4 \times 16)$. $L = 16$ centroids (codewords) for each group, and each centroid has 4×4 dimensions.

In the example above, figure 13.3, person 1 is efficiently represented using $4 \times Centroids$. Thus, the dimensions of the training features are reduced by $\approx 98.96\%$ $(1 - \frac{4 \times Centroids}{6 \times 128 \times 128}) \times 100\%$. Therefore, the storage requirements are reduced. In general, the dimensionality reduction is expressed in the following equation:

$$Dim. = (1 - \frac{4 \times c}{n \times o}) \times 100\% \tag{13.3}$$

where Dim. is a dimensionality reduction. c is the Centroids dimensions. n is the number of trained poses. o is the detected face dimensions.

13.1.3 Classification

In the testing phase, figure 13.1-b, the same preprocessing steps of detecting face, facial parts, and resizing to appropriate dimensions are applied to the test facial image. Each facial part detected is tested separately. All codebooks of all groups belonging to the same class of the detected part of all subjects in the database are used to reconstruct the quantized versions of the detected part. The Euclidean distances, between detected part and the reconstructed ones, are calculated. Find the minimum distance among all the results. Then, the detected part will match the person whose codebook has the minimum distance. Repeat the process for other parts. Since each test image

of each person has 4 facial parts detected (Left Eye, Right Eye, Mouth, and Nose), there are 4 results. If 3 out of 4 detected facial parts are matched to the same person, that test pose will declare as a correctly matched. All poses that are correctly identifying each person constitute the final recognition rate. The recognition rate is measured as $\frac{C}{T} \times 100\ \%$, where C is the total number of poses correctly matched and T is the total number of poses in the database, which are used in the testing mode.

13.2 Experimental Results

The results of the proposed system are presented compared with those obtained in [97, 91], and [90]. The proposed approach is evaluated using four databases, namely, ORL, YALE, FERET, and FEI that have different facial variations, such as illuminations, rotations, makeups, facial expressions, etc. K-fold CV is used to analyze the results of the proposed system. The proposed system achieved 100% recognition rates when tested with the training poses.

13.2.1 Experimental Results for the ORL Database

ORL database has 40 persons each with 10 different poses [28]. Table 13.3 summarizes the results of the proposed system compared with the other approaches. As shown in Table 13.3, increasing the value of K results in improving the recognition rates.

13.2.2 Experimental Results for the YALE Database

In the YALE database, there are 15 subjects each with 11 different poses [29]. As before in ORL database, the results of the proposed approach shown in Table 13.4 are higher than other methods.

Table 13.3: Experimental Results for the ORL Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.33%**	89%	88%	92.5%
K=3	**97.1%**	92.3%	91.88%	94.75%
K=5	**98.25%**	94.75%	94.25%	96.5%

Table 13.4: Experimental Results for the YALE Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.75%**	88.33%	87.33%	92.12%
K=3	**97.6%**	91.4%	90.56%	93.94%
K=5	**98.51%**	94.17%	93.1%	95.15%

13.2.3 Experimental Results for the FERET Database

FERET database consists of 200 persons each with 11 different poses [30, 31]. The results of the proposed approach shown in Table 13.5 exhibit the same behavior as preceding databases.

Table 13.5: Experimental Results for the FERET Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.01%**	88.21%	87.5%	92.5%
K=3	**96.83%**	91.23%	90.51%	94.86%
K=5	**97.98%**	93.57%	92.21%	95.68%

13.2.4 Experimental Results for the FEI Database

It has 200 subjects each with 14 different poses. This is a colored database. The gray-scale version of this database is used in this contribution [32]. As preceding databases, the results of the proposed approach shown in Table 13.6 achieved higher recognition rates in comparison with the other approaches.

Table 13.6: Experimental Results for the FEI Database

K-Fold	Recognition rates of the proposed system (DCT/VQ)	VQ-KFCG [97]	DCT [91]	DWT and PCA [90]
K=2	**95.18%**	87.5%	87.32%	92.25%
K=3	**96.54%**	90.31%	90.33%	93.92%
K=5	**97.92%**	93.81%	92%	95.58%

13.3 Discussion

The recognition rates of the proposed system, shown in Tables 13.3-13.6, are higher than the rates obtained in [97, 91], and [90]. Also, recognition rates of the proposed approach were 100% for all databases used when the system tested with the training poses. In this contribution, FPD in conjunction with VQ in the manner proposed results in an enhancements in the image feature selection and, thus, the recognition rates compared with [97, 91], and [90].

13.4 Conclusion

A face recognition system based on FPD integrated with the VQ algorithm was proposed in this chapter. Both low storage requirements and high recognition rates were accomplished in this contribution. FPD were employed to detect Left Eye, Right Eye, Mouth, and Nose. Four groups for each person, one group for each detected part, were established from the detected facial parts using all training poses that led to reduce the training features dimensions. Then, VQ algorithm employing KFCG was applied to each group for further feature compaction and data discrimination. Each person was efficiently presented using 4 centroids of 4 groups regardless the number of poses used in the training mode. The final feature matrices (cenroids) for each person were of size $4 \times Centroids$ that have dimensions of $4 \times 4 \times 16$ (256), while the input facial images have different dimensions shown in Table 13.2. As stated in the Tables 13.3-13.6, the recognition rates of the proposed approach were higher than the results reported in [97, 91], and [90]. The results were analyzed using K-Fold CV. Four databases, namely, ORL, YALE, FERET, and FEI were used to evaluate the system. For example $K = 5$, the recognition rates realized were **98.25%**, **98.51%**, **97.98%**, and **97.92%** for ORL, YALE, FERET, and FEI databases, respectively.

CHAPTER 14: CONCLUSION AND FUTURE WORKS

The performance of any face recognition system depends on three factors: (1) the recognition accuracy, (2) the storage requirements, and (3) the computational complexity. Some of the researchers in this field focused on achieving high recognition rates while de-emphasizing reduction of the storage requirements and/or the computational complexity.

In this dissertation, two different recognition system families were developed, presented, and analyzed. Each family was composed of several face recognition systems. Each system was implemented in three steps, namely, preprocessing, feature extraction, and classification.

Five databases were used in this dissertation to evaluate the proposed systems. ORL, YALE, and FERET were used to test the systems presented in chapters 4, 5, 7, and 8. The systems explained in chapters 9, 11, 12, and 13 were evaluated using ORL, YALE, FERET, and FEI databases.The five databases discussed in sub-subsection 1.3.1-1.3.5 were employed to evaluate the systems presented in chapters 6 and 10. These databases contain different facial variations, such as light conditions, rotations, facial expressions, make-up, facial details, etc. Also, for classification purposes, this dissertation used either the NN based classifier employing BPTA for training and testing or the Euclidean distance. Finally, the experimental results were analyzed by K-folds CV.

Three integrated tools, namely, 2D DMWT, 2D RT, and 3D DWT were proposed for face recognition in Chapter 5. The proposed system accomplished higher recognition rates and less storage requirements compared to those accomplished by systems presented in chapter 4, which were based on the integrated tools: 2D DMWT, 2D RT, and 2D DWT.

Two powerful tools, 2D DMWT and FastICA, were proposed in chapter 7 for face recognition. In chapter 7, higher recognition rates, less computational complexity, and less storage requirements

were accomplished compared to those achieved by the systems presented in chapters 4, 5, and 6. This is due to the fact that the facial images were efficiently represented by the eigenfaces compared to represented by the systems presented in chapters 4, 5, and 6.

The same tools, 2D DMWT and FastICA, were used by the system presented in chapter 8. In this chapter, discriminating and independent features were obtained by applying FastICA to different combinations of poses corresponding to the sub-images of LL sub-band of DMWT. The system presented in chapter 8 accomplished less storage requirements and lower computational complexity with comparable recognition rates relative to those three factors obtained by the systems reported in chapters 4, 5, 6, and 7. Each person in chapter 8 was efficiently represented by four feature vectors regardless of the number of poses used in the training mode.

In chapter 9, the facial images were partitioned into six parts. Each pose was represented by six feature vectors. In chapter 9, the proposed system accomplished comparable results to those achieved by the other methods. The system presented in chapter 10 was the extended version of the system presented in chapter 9. In chapter 10, two facial representations were obtained after applying 2D DMWT to each part of each pose. Then, several techniques were applied to each representation. In chapter 10, *Techniques 1-3* in the first representation achieved higher rates than those reported by *techniques 1-3* in the second representation. Among all techniques, *technique 1* of the first representation, applying FastICA/ 2D DMWT with ℓ_2-norm to each part, accomplished the highest accuracies. In fact, this extended system achieved the highest recognition rates compared to those obtained by all other systems presented in this dissertation.

Another facial representation was obtained by the systems proposed in the second family. The VQ algorithm in the DCT domain was proposed for face recognition in chapter 11. In chapter 11, the dimensionality reduction was achieved by applying 2D DCT to the facial images. Then, VQ algorithm employing KFCG for codebook initialization was applied to the resulting features

for further dimensionality reduction and better facial representation. Each pose in chapter 11 was represented by one centroid. In chapter 11, we achieved less storage requirements and high recognition rates compared to the state-of-the-art approaches.

VQ algorithm in the DWT domain was proposed in chapter 12 to improve the performance of the recognition system presented in chapter 11. In chapter 12, the proposed system achieved less storage requirements and higher recognition rates compared to those obtained in chapter 11 and the other state-of-the-art approaches.

The FPD and VQ algorithm were proposed in chapter 13 for face recognition. Each person in chapter 13 was represented by 4 centroids regardless of the number of poses used in the training mode. Therefore, this system achieved less storage requirements and comparable results to those reported by the systems presented in chapters 11, 12, and the state-of-the-art methods. Therefore, applying FPD to the facial images in the system presented in chapter 13 led to reduce the storage required for each person.

The systems in family-2 were proposed to further reduce the storage requirements compared to those accomplished by the systems in family-1 while attaining comparable results. For example, in family-1, the integrated tools (FastICA/ 2D DMWT), applied to different combinations of sub-images in the FERET database with K-fold=5 (9 different poses used in the training mode), reduce the dimensions of the database by 97.22% and achieve 99% accuracy. In contrast, the integrated tools, VQ/ FPD, in the family-2 reduce the dimensions of the data by 99.31% and achieve 97.98% accuracy. In this example, the integrated tools, VQ/ FPD, achieved further data compression and lower accuracy compared to those reported by FastICA/ 2D DMWT tools.

14.1 Future Works

In this section, different ideas for future works will be discussed.

- Most of the databases used in this dissertation were noise free. We would like to apply the proposed techniques to noisy databases, big databases that have large number of subjects, low resolution databases, and databases that have unavailability of whole face. Also, we would like to examine the impact of adding noise to the training database on the performance of our recognition systems.

- Although the dimensionality of the trained features of each person was reduced as expressed in Eq. 13.3, the computational complexity and the recognition rates were either less or comparable to those obtained by other systems and the state-of-the-art approaches. We would like to deeply investigate other methods to obtain more sparse facial representations while accomplishing high recognition rates and low computational complexity.

- We would like to examine the performance of the proposed systems in different disciplines. We would like to apply the proposed techniques in Civil Engineering/ Earthquake Engineering. The goal is to recognize the types and the degrees of the earthquakes by analyzing the earthquake signals through our methods. Also, we would like to examine the behavior of the proposed techniques in the Medical field, where we believe that the proposed techniques are a powerful way of recognizing different types of tumors.

LIST OF REFERENCES

[1] Nikolaos V. Boulgouris, Konstantinos N. Plataniotis, and Evangelia Micheli-Tzanakou. *Biometrics. [electronic resource] : theory, methods, and applications.* IEEE Press series on computational intelligence. Piscataway, N.J. : IEEE Press ; Hoboken, N.J. : Wiley, 2010., 2010. ISBN 9780470522356.

[2] Rabia Jafri and Hamid R Arabnia. A survey of face recognition techniques. *journal of information processing systems*, 5(2):41–68, (2009).

[3] Smita S Mudholkar, Pradnya M Shende, and M Sarode. Biometrics authentication technique for intrusion detection system using fingerprint recognition. *International Journal of Computer Science, Engineering and Information Technology*, 2(1):57–65, (2012).

[4] Asit Kumar Datta, Madhura Datta, and Pradipta Kumar Banerjee. *Face detection and recognition : theory and practice.* Boca Raton : CRC Press / Taylor and Francis Group, 2016., 2016.

[5] Alexandros Eleftheriadis and Arnaud Jacquin. Automatic face location detection for model-assisted rate control in h.261-compatible coding of video. *Signal Processing: Image Communication*, 7:435 – 455, (1995).

[6] Yu-Bu Lee and Sukhan Lee. Robust face detection based on knowledge-directed specification of bottom-up saliency. *Etri Journal*, 33(4):600–610, (2011).

[7] Wenkai Xu and Eung-Joo Lee. A novel multi-view face detection method based on improved real adaboost algorithm. *KSII Transactions on Internet and Information Systems (TIIS)*, 7 (11):2720–2736, (2013).

[8] Y. Hongtao, Q. Jiaqing, and F. Ping. Face recognition with discrete cosine transform. In

IEEE Second International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC), pages 802–805, Dec 2012.

[9] Erik Hjelms and Boon Kee Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236 – 274, 2001. ISSN 1077-3142.

[10] S. Z. Li and Zhenqiu Zhang. Floatboost learning and statistical face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1112–1123, Sept 2004.

[11] Anil K. Jain and S. Z. Li. *Handbook of face recognition. [electronic resource].* London ; New York : Springer, c2011., 2011.

[12] Takeo Kanade. Picture processing system by computer complex and recognition of human faces. In *doctoral dissertation, Kyoto University*. November 1973.

[13] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. Am. A*, 4(3):519–524, Mar 1987.

[14] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):103–108, Jan 1990.

[15] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3 (1):71–86, Jan 1991.

[16] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *IEEE Proceedings Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 586–591, Jun 1991.

[17] MorphoTrust Technology A Safaran Company. Face recognition. `http://www.morphotrust.com/Technology/FaceRecognition.aspx`, USA.

[18] Cognitec systems GmbH. Face recognition. http://www.cognitec.com/, Germany.

[19] Facekey corporation. Face recognition. http://www.facekey.com/, USA.

[20] LATHEM. Face recognition. https://www.lathem.com/time-clocks/face-recognition-biometric-time-clocks, USA.

[21] FINGERTEC. Face recognition. http://www.fingertec.com/images/w_brochure/facerecognition_e.html, USA.

[22] Yong Xu, Xiaozhao Fang, Xuelong Li, Jiang Yang, Jie You, Hong Liu, and Shaohua Teng. Data uncertainty in face recognition. *IEEE Transactions on Cybernetics*, 44(10):1950–1961, (2014).

[23] Shuicheng Yan, Dong Xu, Benyu Zhang, Hong-Jiang Zhang, Qiang Yang, and Stephen Lin. Graph embedding and extensions: a general framework for dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1):40–51, (2007).

[24] Jian Yang, David Zhang, Alejandro F Frangi, and Jing-yu Yang. Two-dimensional pca: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):131–137, (2004).

[25] Ripal Patel, Nidhi Rathod, and Ami Shah. Comparative analysis of face recognition approaches: a survey. *International Journal of Computer Applications*, 57(17):50–69, (2012).

[26] Mohammad Mahdi Karimi and Hamid Soltanian-Zadeh. Face recognition: A sparse representation-based classification using independent component analysis. In *IEEE Sixth International Symposium on Telecommunications (IST)*, pages 1170–1174, (2012).

[27] Thai Hoang Le. Applying artificial neural networks for face recognition. *Advances in Artificial Neural Systems*, 2011:15, 2011.

[28] ORL Database. At&t laboratories cambridge database of faces. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html, April (1992-1994).

[29] YALE Database. Ucsd computer vision. http://vision.ucsd.edu/content/yale-face-database.

[30] P Jonathon Phillips, Harry Wechsler, Jeffery Huang, and Patrick J Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16 (5):295–306, (1998).

[31] P Jonathon Phillips, Hyeonjoon Moon, Syed Rizvi, Patrick J Rauss, et al. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1090–1104, (2000).

[32] Carlos Eduardo Thomaz and Gilson Antonio Giraldi. A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing*, 28(6):902 – 913, 2010. Avialable at: http://fei.edu.br/ cet/facedatabase.html.

[33] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[34] Ahmed Aldhahab, George Atia, and Wasfy B Mikhael. Supervised facial recognition based on multi-resolution analysis and feature alignment. In *2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 137–140, Aug (2014). doi: 10.1109/MWSCAS.2014.6908371.

[35] A. Aldhahab, G. Atia, and W. B. Mikhael. Supervised facial recognition based on multiresolution analysis with radon transform. In *IEEE 48th Asilomar Conference on Signals, Systems and Computers*, pages 928–932, Nov 2014. doi: 10.1109/ACSSC.2014.7094589.

[36] Ahmed Aldhahab and Wasfy B. Mikhael. Three supervised facial recognition techniques based on fastica/dmwt. *Multidimensional Systems and Signal Processing*, Submitted on August 2016 and revised on February 2017, .

[37] A. Aldhahab and W. B. Mikhael. Employing efficient techniques based on 2d dmwt/fastica for supervised facial recognition. In *2016 Future Technologies Conference (FTC)*, pages 1316–1323, Dec 2016. doi: 10.1109/FTC.2016.7821773.

[38] A. Aldhahab, G. Atia, and W. B. Mikhael. Supervised facial recognition based on eigenanalysis of multiresolution and independent features. In *2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1–4, Aug 2015. doi: 10.1109/MWSCAS.2015.7282087.

[39] Ahmed Aldhahab, George Atia, and Wasfy B. Mikhael. *High Performance and Efficient Facial Recognition Using Norm of ICA/Multiwavelet Features*, pages 672–681. Springer International Publishing, Cham, 2015. ISBN 978-3-319-27863-6. doi: 10.1007/978-3-319-27863-6_63.

[40] A. Aldhahab and W. B. Mikhael. A facial recognition method based on dmw transformed partitioned images. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Accepted, Aug 2017.

[41] Ahmed Aldhahab and Wasfy B. Mikhael. Face recognition employing dmwt followed by fastica. *Circuits, Systems, and Signal Processing*, Accepted in June 2017, .

[42] A. Aldhahab, T. A. Obaidi, and W. B. Mikhael. Employing vector quantization algorithm in a transform domain for facial recognition. In *2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1–4, Oct 2016. doi: 10.1109/MWSCAS.2016.7869957.

[43] A. Aldhahab, T. Alobaidi, and W. B. Mikhael. Efficient facial recognition using vector quantization of 2d dwt features. In *2016 50th Asilomar Conference on Signals, Systems and Computers*, pages 439–443, Nov 2016. doi: 10.1109/ACSSC.2016.7869077.

[44] T. Alobaidi A. Aldhahab and Wasfy B. Mikhael. Employing vector quantization on detected facial parts for face recognition. In *IEEE 4th Global Conference on Signal and Information Processing*. IEEE, Dec 2016.

[45] K. Papachristou, A. Tefas, and I. Pitas. Symmetric subspace learning for image analysis. *IEEE Transactions on Image Processing*, 23(12):5683–5697, Dec 2014.

[46] P. C. Yuen, G. C. Feng, and D. Q. Dai. Human face image retrieval system for large database. In *IEEE Proceedings Fourteenth International Conference on Pattern Recognitionn*, volume 2, pages 1585–1588 vol.2, Aug 1998.

[47] Laurenz Wiskott, J-M Fellous, N Kuiger, and Christoph Von Der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7):775–779, 1997.

[48] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, Jul 1997.

[49] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696–710, Jul 1997.

[50] Thomas Vetter. Synthesis of novel views from a single face image. *International journal of computer vision*, 28(2):103–116, 1998.

[51] T. Shakunaga and K. Shigenari. Decomposed eigenface for face recognition under various

lighting conditions. In *IEEE Proceedings in Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages I–864–I–871 vol.1, 2001.

[52] Fujin Zhong. L1-norm-based (2d)ˆ 2pca. In *Conference on Computer Science and Electrical Engineering*, pages 1293–1296, 2013.

[53] S. Kundu, P. P. Markopoulos, and D. A. Pados. Fast computation of the l1-principal component of real-valued data. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8028–8032, May 2014.

[54] M. Johnson and A. Savakis. Fast l1-eigenfaces for robust face recogntion. In *IEEE Western New York Image and Signal Processing Workshop (WNYISPW)*, pages 1–5, Nov 2014.

[55] A. Iosifidis, A. Tefas, and I. Pitas. On the optimal class representation in linear discriminant analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 24(9):1491–1497, Sept 2013.

[56] H. M. Moon, Dongjin Choi, Pankoo Kim, and S. B. Pan. Lda-based face recognition using multiple distance training face images with low user cooperation. In *2015 IEEE International Conference on Consumer Electronics (ICCE)*, pages 7–8, Jan 2015.

[57] Salient feature and reliable classifier selection for facial expression classification. *Pattern Recognition*, 43(3):972 – 986, 2010.

[58] Yu Ting Zhang, Jia Jun and Shi. Face recognition systems based on independent component analysis and support vector machine. In *IEEE International Conference on Audio, Language and Image Processing (ICALIP)*, pages 296–300, (2014).

[59] M. M. Prasad. 1d-lda verses 2d-lda in online handwriting recognition. In *IEEE International Conference on Circuits, Communication, Control and Computing (I4C)*, pages 431–433, Nov 2014.

[60] Alok Sharma, Kuldip K. Paliwal, and Godfrey C. Onwubolu. Class-dependent pca, {MDC} and lda: A combined classifier for pattern classification. *Pattern Recognition*, 39(7):1215 – 1229, 2006.

[61] Xiaofei He and Partha Niyogi. Locality preserving projections. In S. Thrun, L. K. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*, pages 153–160. MIT Press, 2004.

[62] J. Soldera, C. Alberto Ramirez Behaine, and J. Scharcanski. Customized orthogonal locality preserving projections with soft-margin maximization for face recognition. *IEEE Transactions on Instrumentation and Measurement*, 64(9):2417–2426, Sept 2015.

[63] M. E. Ashalatha, M. S. Holi, and P. R. Mirajkar. Face recognition using local features by lpp approach. In *IEEE International Conference on Circuits, Communication, Control and Computing (I4C)*, pages 382–386, Nov 2014.

[64] M.S. Bartlett, Javier R. Movellan, and T.J. Sejnowski. Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, 13(6):1450–1464, Nov (2002). ISSN 1045-9227.

[65] K.S. Kinage and S.G. Bhirud. Face recognition based on independent component analysis on wavelet subband. In *IEEE 3rd International Conference on Computer Science and Information Technology (ICCSIT)*, volume 9, pages 436–440, July 2010.

[66] Pierre Comon. Independent component analysis, a new concept? *Signal processing*, 36(3): 287–314, (1994).

[67] Bruce A Draper, Kyungim Baek, Marian Stewart Bartlett, and J Ross Beveridge. Recognizing faces with pca and ica. *Computer vision and image understanding*, 91(1):115–137, 2003.

[68] Pong C Yuen and Jian-Huang Lai. Face representation using independent component analysis. *Pattern recognition*, 35(6):1247–1257, (2002).

[69] Goutam Majumder and Mrinal Kanti Bhowmik. Gabor-fast ica feature extraction for thermal face recognition using linear kernel support vector machine. In *IEEE International Conference on Computational Intelligence and Networks (CINE)*, pages 21–25, (2015).

[70] Xingfu Zhang and Xiangmin Ren. Two dimensional principal component analysis based independent component analysis for face recognition. In *IEEE International Conference on Multimedia Technology (ICMT)*, pages 934–936, (2011).

[71] Yongguo Liu, Gang Chen, Jiwen Lu, and Wanjun Chen. Face recognition based on independent component analysis and fuzzy support vector machine. In *2006 6th World Congress on Intelligent Control and Automation*, volume 2, pages 9889–9892, 2006.

[72] M. Li, F. Wu, and X. Liu. Face recognition based on wt, fastica and rbf neural network. In *Third International Conference on Natural Computation (ICNC 2007)*, volume 2, pages 3–7, Aug 2007.

[73] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[74] Daniel D. Lee and H. Sebastian Seung. Algorithms for non-negative matrix factorization. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13*, pages 556–562. MIT Press, 2001.

[75] Xianchuan Yu, Dan Hu, and Jindong Xu. *Blind source separation : theory and applications.* Singapore : Wiley, 2014. ISBN 9781118679869.

[76] S. Z. Li, Xin Wen Hou, Hong Jiang Zhang, and Qian Sheng Cheng. Learning spatially localized, parts-based representation. In *Proceedings of the IEEE Computer Society Conference*

on Computer Vision and Pattern Recognition (CVPR), volume 1, pages I–207–I–212 vol.1, 2001.

[77] S. Zafeiriou, A. Tefas, I. Buciu, and I. Pitas. Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification. *IEEE Transactions on Neural Networks*, 17(3):683–695, May 2006.

[78] X. Long, H. Lu, and Y. Peng. Sparse non-negative matrix factorization based on spatial pyramid matching for face recognition. In *IEEE 5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 1, pages 82–85, Aug 2013.

[79] F. Purnomo, D. Suhartono, M. Shodiq, A. Susanto, S. Raharja, and R. W. Kurniawan. Face recognition using gabor wavelet and non-negative matrix factorization. In *IEEE in SAI Intelligent Systems Conference (IntelliSys)*, pages 788–792, Nov 2015.

[80] Pallavi D Wadkar and Megha Wankhade. Face recognition using discrete wavelet transforms. *International Journal of Advanced Engineering Technology*, 3:1–3, 2012.

[81] M. M. Mukhedkar and S. B. Powalkar. Fast face recognition based on wavelet transform on pca. In *IEEE International Conference on Energy Systems and Applications*, pages 761–764, Oct 2015.

[82] A. S. B. Mahajan and K. J. Karande. Pca and dwt based multimodal biometric recognition system. In *IEEE International Conference on Pervasive Computing (ICPC)*, pages 1–4, Jan 2015.

[83] Min Luo, Liu Song, and Shi-dong Li. An improved face recognition based on ica and wt. In *IEEE Asia-Pacific Services Computing Conference (APSCC)*, pages 315–318, (2012).

[84] M. P. Satone and G. K. Kharate. Face recognition based on pca on wavelet subband. In *2012*

IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS),, pages 1–4, March 2012.

[85] Tarik Z Ismaeel, Aya A Kamil, and Ahkam K Naji. Human face recognition using stationary multiwavelet transform. *International Journal of Computer Applications*, 72(1), (2013).

[86] Xie Zhihua and Liu Guodong. Weighted infrared face recognition in multiwavelet domain. In *IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 70–74, Oct (2013).

[87] A. A. G. Azzawi and M. A. H. Al-Saedi. Face recognition based on mixed between selected feature by multiwavelet and particle swarm optimization. In *2010 Developments in E-systems Engineering*, pages 199–204, Sept 2010.

[88] S. K. Dandpat and S. Meher. Performance improvement for face recognition using pca and two-dimensional pca. In *2013 IEEE International Conference on Computer Communication and Informatics*, pages 1–5, Jan 2013.

[89] M. R. Mulla, R. P. Patil, and S. K. Shah. Facial image based security system using pca. In *2015 IEEE International Conference on Information Processing (ICIP)*, pages 548–553, Dec 2015.

[90] Venus AlEnzi, Mohanad Alfiras, and Falah Alsaqre. Face recognition algorithm using two dimensional principal component analysis based on discrete wavelet transform. In *Digital Information Processing and Communications*, pages 426–438. Springer, 2011.

[91] Z. Karhan and B. Ergen. Classification of face images using discrete cosine transform. In *Conference in Signal Processing and Communications Applications*, pages 1–4, April 2013.

[92] D. Sisodia, L. Singh, and S. Sisodia. Incremental learning algorithm for face recognition

using dct. In *IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN)*, pages 282–286, March 2013.

[93] A. Dahmouni, N. Aharrane, K. El Moutaouakil, and K. Satori. Face recognition using local binary probabilistic pattern (lbpp) and 2d-dct frequency decomposition. In *IEEE 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, pages 73–77, March 2016.

[94] S. Ajitha, A. Annis Fathima, V. Vaidehi, M. Hemalatha, and R. Karthigaiveni. Face recognition system using combined gabor wavelet and dct approach. In *IEEE International Conference on Recent Trends in Information Technology (ICRTIT)*, pages 1–6, April 2014.

[95] F. Z. Chelali, A. Djeradi, and N. Cherabit. Investigation of dct/pca combined with kohonen classifier for human identification. In *IEEE 4th International Conference on Electrical Engineering (ICEE)*, pages 1–7, Dec 2015.

[96] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28(1):84–95, Jan 1980.

[97] Shachi J Natu, Prachi J Natu, Tanuja K Sarode, and HB Kekre. Performance comparison of face recognition using dct against face recognition using vector quantization algorithms lbg, kpe, kmcg, kfcg. *International Journal of Image Processing (IJIP)*, 4(4):377 – 389, Octobor 2010.

[98] Q. Chen, K. Kotani, F. Lee, and T. Ohmi. Face recognition using markov stationary features and vector quantization histogram. In *IEEE 17th International Conference on Computational Science and Engineering (CSE)*, pages 1934–1938, Dec 2014.

[99] C. Ding and D. Tao. Robust face recognition via multimodal deep face representation. *IEEE Transactions on Multimedia*, 17(11):2049–2058, Nov 2015.

[100] Z. Lei, D. Yi, and S. Z. Li. Learning stacked image descriptor for face recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, PP(99):1–1, 2015.

[101] W. Ouarda, H. Trichili, A. M. Alimi, and B. Solaiman. Mlp neural network for face recognition based on gabor features and dimensionality reduction techniques. In *IEEE International Conference on Multimedia Computing and Systems (ICMCS)*, pages 127–134, April 2014.

[102] Z. Zhang, J. Li, and R. Zhu. Deep neural network for face recognition based on sparse autoencoder. In *IEEE 8th International Congress on Image and Signal Processing (CISP)*, pages 594–598, Oct 2015.

[103] V. E. Liong, J. Lu, and G. Wang. Face recognition using deep pca. In *IEEE 9th International Conference on Information, Communications and Signal Processing (ICICS)*, pages 1–5, Dec 2013.

[104] D. Li, H. Zhou, and K. M. Lam. High-resolution face verification using pore-scale facial features. *IEEE Transactions on Image Processing*, 24(8):2317–2327, Aug 2015.

[105] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, Jul 2002.

[106] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, Dec 2006.

[107] L. Wiskott, J. M. Fellous, N. Kuiger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, Jul 1997.

[108] N. K. Patil, S. Vasudha, and L. R. Boregowda. Performance improvement of face recognition system by decomposition of local features using discrete wavelet transforms. In *IEEE International Symposium on Electronic System Design (ISED)*, pages 172–176, Dec 2013.

[109] M. Yang, L. Zhang, S. C. K. Shiu, and D. Zhang. Robust kernel representation with statistical local features for face recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 24(6):900–912, June 2013.

[110] S. Nazari and M. S. Moin. Face recognition using global and local gabor features. In *IEEE 21st Iranian Conference on Electrical Engineering (ICEE)*, pages 1–4, May 2013.

[111] X. Duan and Z. H. Tan. Local feature learning for face recognition under varying poses. In *IEEE International Conference on Image Processing (ICIP)*, pages 2905–2909, Sept 2015.

[112] J. Lu, V. E. Liong, and J. Zhou. Simultaneous local binary feature learning and encoding for face recognition. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3721–3729, Dec 2015.

[113] Semih Ergin and M Bilginer Gulmezoglu. A novel framework for partition-based face recognition. *International Journal of Innovative Computing Information and Control*, 9 (5):1819–1834, (2013).

[114] Guanghui He, Yuanyan Tang, Bin Fang, and Taiping Zhang. Weightiness image partition in 3d face recognition. In *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 5068–5071, (2009).

[115] Wilhelm Burger and Mark J Burge. *Digital image processing: an algorithmic introduction using Java*. Springer Science & Business Media, 2009.

[116] Charles K Chui. *An introduction to wavelets*, volume 1. Academic press, 2014.

[117] Ming Li, Fuwen Wu, and Xueyan Liu. Face recognition based on wt, fastica and rbf neural network. In *IEEE Third International Conference on Natural Computation (ICNC)*, volume 2, pages 3–7, 2007.

[118] Ingram J. Brown. A wavelet tour of signal processing: the sparse way. *Investigacion Operacional*, (1):85, 2009. ISSN 0257-4306.

[119] Rafael C. Gonzalez and Richard E. Woods. *Digital image processing.* Upper Saddle River, N.J. : Prentice Hall., 2002. ISBN 0201180758.

[120] Jun-Hai Zhai, Su-Fang Zhang, and Li-Juan Liu. Image recognition based on wavelet transform and artificial neural networks. In *2008 International Conference on Machine Learning and Cybernetics*, volume 2, pages 789–793, July 2008.

[121] V. Strela, P.N. Heller, G. Strang, P. Topiwala, and C. Heil. The application of multiwavelet filterbanks to image processing. *IEEE Transactions on Image Processing*, 8(4):548–563, Apr (1999). ISSN 1057-7149.

[122] Vasily Strela and Andrew T Walden. Orthogonal and biorthogonal multiwavelets for signal denoising and image compression. In *Aerospace/Defense Sensing and Controls*, pages 96–107. International Society for Optics and Photonics, (1998).

[123] Jeffrey S Geronimo, Douglas P Hardin, and Peter R Massopust. Fractal functions and wavelet expansions based on several scaling functions. *Journal of approximation theory*, 78(3):373–401, (1994).

[124] Vasily Strela. *Multiwavelets: theory and applications*. PhD thesis, Citeseer, 1996.

[125] Kwok-Wai Cheung and Lai-Man Po. Preprocessing for discrete multiwavelet transform of two-dimensional signals. In *Proceedings IEEE International Conference on Image Processing*, volume 2, pages 350–353. IEEE, 1997.

[126] Douglas P Hardin and David W Roach. Multiwavelet prefilters. 1. orthogonal prefilters preserving approximation order p⩽ 2. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(8):1106–1112, 1998.

[127] James T Miller and Ching-Chung Li. Adaptive multiwavelet initialization. *IEEE Transactions on Signal Processing*, 46(12):3282–3291, 1998.

[128] Mariantonia Cotronei, Laura B Montefusco, and Luigia Puccio. Multiwavelet analysis and signal processing. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(8):970–987, 1998.

[129] Xiang-Gen Xia, Jeffrey S Geronimo, Douglas P Hardin, and Bruce W Suter. Design of prefilters for discrete multiwavelet transforms. *IEEE Transactions on Signal Processing*, 44 (1):25–35, 1996.

[130] Xiang-Gen Xia. A new prefilter design for discrete multiwavelet transforms. *IEEE Transactions on Signal Processing*, 46(6):1558–1570, 1998.

[131] Wafi N. K., Saleh Z. J. M., and Mahmoud W. A. The determination of critical-sampling scheme of preprocessing for multiwavelets decomposition as 1st and 2nd orders of approximations. *Al-Khawarizmi Engineering Journal*, (1):26, 2005. ISSN 18181171.

[132] Lian Cai and Sidan Du. Rotation, scale and translation invariant image watermarking using radon transform and fourier transform. In *Proceedings of the IEEE 6th Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication (IEEE Cat. No.04EX710)*, volume 1, pages 281–284 Vol.1, May 2004.

[133] P. Rai and P. Khanna. Gender classification using radon and wavelet transforms. In *2010 5th International Conference on Industrial and Information Systems*, pages 448–451, July 2010.

[134] Allen Gersho and Robert M. Gray. *Vector quantization and signal compression.* The Kluwer international series in engineering and computer science: SECS 159. Boston, 1992.

[135] HB Kekre and Tanuja K Sarode. New fast improved codebook generation algorithm for color images using vector quantization. *International Journal of Engineering and Technology*, 1(1):67–77, 2008.

[136] Juha Karhunen, Petteri Pajunen, and Erkki Oja. The nonlinear pca criterion in blind source separation: Relations with other approaches. *Neurocomputing*, 22(1):5–20, 1998.

[137] M Girolami and C Fyfe. Blind separation of sources using exploratory projection pursuit networks. In *International Conference on the Speech and Signal Processing Engineering Applications of Neural Networks*, pages 249–252, 1996.

[138] A Bell and T Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, Nov (1995). ISSN 0899-7667.

[139] Gilles Burel. Blind separation of sources: A nonlinear neural algorithm. *Neural networks*, 5(6):937–947, 1992.

[140] Aapo Hyvärinen and Erkki Oja. A fast fixed-point algorithm for independent component analysis. *Neural computation*, 9(7):1483–1492, 1997.

[141] Francis R Bach and Michael I Jordan. Kernel independent component analysis. *The Journal of Machine Learning Research*, 3:1–48, 2003.

[142] Erkki Oja. The nonlinear pca learning rule in independent component analysis. *Neurocomputing*, 17(1):25–45, 1997.

[143] Z. Lihong, W. Ye, and T. Hongfeng. Face recognition based on independent component analysis. In *IEEE Chinese Control and Decision Conference (CCDC)*, pages 426–429, May 2011.

[144] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, May (1999). ISSN 1045-9227.

[145] Aapo Hyvarinen. New approximations of differential entropy for independent component analysis and projection pursuit. *Advances in neural information processing systems*, 10(2): 273–279, 1998.

[146] Aapo Hyvärinen. A family of fixed-point algorithms for independent component analysis. In *IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP*, volume 5, pages 3917–3920. IEEE, 1997.

[147] Al-Jouhar W. A. and Abbas T. M. Feature combination and mapping using multiwavelet transform. *Iraq Academic Scientific Journals*, 3(19):13–34, 2006. ISSN 16816870.

[148] Zhao Lihong, Wang Ye, and Teng Hongfeng. Face recognition based on independent component analysis. In *Control and Decision Conference (CCDC), Chinese*, pages 426–429, May 2011.

[149] Peter de B Harrington. Sigmoid transfer functions in backpropagation neural networks. *Analytical Chemistry*, 65(15):2167–2168, (1993).

[150] J. S. Pan, Q. Feng, L. Yan, and J. F. Yang. Neighborhood feature line segment for image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(3): 387–398, March 2015.

[151] E. Kussul and T. Baydyk. Face recognition using special neural networks. In *IEEE International Joint Conference on Neural Networks (IJCNN)*, pages 1–7, July 2015.

[152] Samarjeet Powalkar and Moresh M Mukhedkar. Fast face recognition based on wavelet transform on pca. *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, 1(4):21–24, 2015.

[153] M. R. Y. Darestani, M. Sheikhan, and M. Khademi. Face recognition using contourlet-based features and hybrid pso-neural model. In *IEEE 5th Conference on Information and Knowledge Technology (IKT)*, pages 1818–186, May 2013.

[154] Sylvain Arlot, Alain Celisse, et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79, (2010).

[155] A. A. G. Azzawi and M. A. H. Al-Saedi. Face recognition based on mixed between selected feature by multiwavelet and particle swarm optimization. In *Developments in E-systems Engineering (DESE)*, pages 199–204, Sept 2010.

[156] T. Khadhraoui, S. Ktata, F. Benzarti, and H. Amiri. Features selection based on modified pso algorithm for 2d face recognition. In *IEEE 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, pages 99–104, March 2016.

[157] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler. Robust face recognition for uncontrolled pose and illumination changes. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(1):149–163, Jan 2013.

[158] B. Dhivakar, C. Sridevi, S. Selvakumar, and P. Guhan. Face detection and recognition using skin color. In *IEEE 3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, pages 1–7, March 2015.

[159] Xiangyu Zhu, Z. Lei, Junjie Yan, D. Yi, and S. Z. Li. High-fidelity pose and expression normalization for face recognition in the wild. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 787–796, June 2015.

[160] Silvia Ferrari, Iuri Frosio, Vincenzo Piuri, and N Alberto Borghese. Enhanced vector quantization for data reduction and filtering. In *IEEE 2nd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, pages 470–477, 2004.