

University of Windsor

Scholarship at UWindor

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

2011

Human Face Recognition

Amirhosein Nabatchian

University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

Recommended Citation

Nabatchian, Amirhosein, "Human Face Recognition" (2011). *Electronic Theses and Dissertations*. 437. <https://scholar.uwindsor.ca/etd/437>

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

Human Face Recognition

by

Amirhosein Nabatchian

A Dissertation

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

Windsor, Ontario, Canada
2011

© 2011 Amirhosein Nabatchian

All Rights Reserved. No part of this document may be reproduced, stored or otherwise retained in a retrieval system or transmitted in any form, on any medium, by any means without prior written permission of the author.

Declaration of Co-Authorship and Previous Publication

I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research with Dr. Majid Ahmadi and Dr. Esam Abdel-Rahem and under their supervision. I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-authors to include the above materials in my thesis.

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

II. Declaration of Previous Publication

This thesis includes six original papers that have been previously published/submitted for publication in peer reviewed journals, as follows:

Thesis Chapter	Publication title and full citation	Publication Status
Chapter 3	A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, Human Face Recognition in Presence of Noise and Misalignment in the Detected Faces, Digital Signal Processing.	Submitted
Chapter 4 and Chapter 5	A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, Illumination Invariant Feature Extraction and Mutual-Information-based Local Matching for Face Recognition under Illumination Variation and Occlusion, Pattern Recognition.	Submitted
Chapter 3	A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, Human Face Recognition Using Different Moment Invariants: A Comparative Study, Image and Signal Processing, 2008. CISP '08. Congress on, 27-30 May 2008, China	Published
Chapter 3	A. Nabatchian, I. Makaremi, E. Abdel-Raheem, M. Ahmadi, Pseudo Zernike Moment Invariants for Recognition of Faces Using Different Classifiers in FERET Database, Convergence and Hybrid Information Technology, 2008. ICCIT '08. Third International Conference on, 11-13 Nov. 2008, Busan, Korea.	Published
Chapter 4	A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, An efficient method for face recognition under illumination variations, High Performance Computing and Simulation (HPCS), 2010 International Conference on, June 28 -July 2, 2010, Caen, France.	Published
Chapter 4 and Chapter 5	A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, A Weighted Voting Scheme for Recognition of Faces with Illumination Variation, Control, Automation, Robotics and Vision, ICARCV 2010, 11 th International Conference on, 7-10 December 2010, Singapore.	Published

I certify that I have obtained a written permission from the copyright owners to include the above published materials in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owners to include such materials in my thesis.

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

Abstract

Face recognition, as the main biometric used by human beings, has become more popular for the last twenty years. Automatic recognition of human faces has many commercial and security applications in identity validation and recognition and has become one of the hottest topics in the area of image processing and pattern recognition since 1990. Availability of feasible technologies as well as the increasing request for reliable security systems in today's world has been a motivation for many researchers to develop new methods for face recognition.

In automatic face recognition we desire to either identify or verify one or more persons in still or video images of a scene by means of a stored database of faces. One of the important features of face recognition is its non-intrusive and non-contact property that distinguishes it from other biometrics like iris or finger print recognition that require subjects' participation.

During the last two decades several face recognition algorithms and systems have been proposed and some major advances have been achieved. As a result, the performance of face recognition systems under controlled conditions has now reached a satisfactory level. These systems, however, face some challenges in environments with variations in illumination, pose, expression, *etc.* The objective of this research is designing a reliable automated face recognition system which is robust under varying conditions of noise level, illumination and occlusion. A new method for illumination invariant feature extraction based on the illumination-reflectance model is proposed which is computationally efficient and does not require any prior information about the face model or illumination. A weighted voting scheme is also proposed to enhance the performance under illumination variations and also cancel occlusions. The proposed method uses mutual information and entropy of the images to generate different weights for a group of ensemble classifiers based on the input image quality. The method yields outstanding results by reducing the effect of both illumination and occlusion variations in the input face images.

To my love, my life and my inspiration,
To my wife and my best friend,
Behjat,
For her patience, love and support.

Contents

Declaration of Co-Authorship and Previous Publication	iv
Abstract	vii
Dedication	ix
List of Figures	xiv
List of Tables	xvii
List of Abbreviations	xix
1. Introduction	1
1.1. Face Recognition	2
1.2. Identification vs. Verification	3
1.3. Applications	4
1.4. Challenges	5
1.4.1. Illumination	5
1.4.2. Head Pose	6
1.4.3. Facial Expressions	6
1.4.4. Occlusion	7
1.4.5. Inter-Class Similarities	7

1.5. Research Objective and Contributions	8
1.6. Outline of the Thesis	8
2. Face recognition Methods	10
2.1. A Face Recognition System	11
2.2. Preprocessing	12
2.3. Face Detection	13
2.4. Feature Extraction	13
2.4.1. Correlation	17
2.4.2. Principal Component Analysis	17
2.4.3. Linear Discriminant Analysis	18
2.4.4. Independent Component Analysis	19
2.4.5. Kernel Principal Component Analysis	20
2.4.6. Wavelets	20
2.5. Classification	21
2.5.1. k Nearest Neighbors	21
2.5.2. Neural Networks	21
2.5.3. Support Vector Machines	22
2.5.4. Hidden Markov Model	23
2.6. Databases	24
2.6.1. The AT&T Database	24
2.6.2. The FERET Database	25
2.6.3. The Yale B Database	25
2.6.4. The Extended Yale B Database	26
2.6.5. The CMU-PIE Database	27

3. Moment Invariants for Human Face Recognition	28
3.1. Moment Invariants	28
3.1.1.Regular Moment Invariants	29
3.1.2.Hu Moment Invariants	30
3.1.3.Bamieh Moment Invariants	30
3.1.4.Zernike Moment Invariants	31
3.1.5.Pseudo Zernike Moment Invariants	32
3.1.6.Teague - Zernike Moment Invariants	33
3.1.7.Normalized Zernike and pseudo Zernike Moment Invariants	34
3.1.8.Legendre Zernike	34
3.2. Locating the Face	35
3.3. Experimental Results	36
3.3.1.Comparing Different Moment Invariants	36
3.3.2.Performance of Different Classifiers	43
3.3.3.Effect of Misalignment	48
3.3.4.Robustness to Noise	58
3.4. Conclusion	61
4. Illumination Invariant Feature Extraction	63
4.1. Illumination Invariant Methods	64
4.2. Illumination Invariants	68
4.2.1.The Illumination Model	68
4.2.2.The Maximum Filter	71
4.2.3.The Proposed Algorithm	75

4.3. Experimental Results	85
4.4. Conclusion	88
5. Local Matching and Voting Systems	89
5.1. Democratic Voting Scheme	90
5.2. Weighted Voting Scheme	94
5.2.1. Mean Grayscale Level	94
5.2.2. Image Entropy	96
5.2.3. Mutual Information between Corresponding Sub-Images	97
5.3. Adaptive Weight Generator	98
5.4. Experimental Results	100
5.4.1. Face recognition	101
5.4.2. Face Verification and ROC Analysis	104
5.4.3. Effect of Reducing the Training Samples	107
5.4.4. Effect of Partial Image Occlusion	109
5.4.5. Effect of a Larger Database	115
5.5. Conclusion	116
6. Concluding Remarks	118
6.1. Summary of Contributions	118
6.2. Conclusion	119
6.3. Future Work	122
References	123
VITA AUCTORIS	134

List of Figures

1.1	An example of authentication vs. identification	4
1.2	Bad illumination can cause between-class similarities overcome the inter-class similarities. Images in each row belong to the same subject but with different illumination	6
2.1	Flow diagram of face recognition process	12
2.2	PCA maps the samples to the direction with maximum variation where LDA finds the direction with maximum between-class variation to within-class variation	18
2.3	Three layer perceptron	22
2.4	Maximum-margin hyperplane for a SVM classifier. Samples on the margin are called the support vectors	23
2.5	Sample images of two subjects the in the AT&T database	24
2.6	One of the samples of the FERET database in frontal position	25
2.7	One of the subjects in the Extended Yale B database with its 5 subsets	26
2.8	One of the samples in CMU-PIE database (a) lights-off group (b) lights-on group	27
3.1	Two of the samples in the database and the result of applying face localization on them	36
3.2	Some wrong detections using Viola-Jones face detector on FERET database	50
3.3	(a) Some samples from FERET database. (b) The result of the Viola-Jones face detection. (c) Images in (b) after histogram equalization	51

3.4	Effect of translation on images in the FERET database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with half of the database without any translations and tested with the other half with the translation	54
3.5	Effect of translation on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the second subset with shifting	55
3.6	Effect of rotation θ degrees on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the translated subsets	56
3.7	Effect of rotation with the maximum of θ degrees on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the translated subsets	57
3.8	Result of the rotation test for the FERET database using 4 methods (a) effect of rotation (b) effect of random rotation with maximum of θ degrees	57
3.9	Effect of Gaussian Noise on images in the Yale B database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with the first subset without Noise and tested with all the subsets after adding noise	59
3.10	Effect of Salt & Pepper Noise on images in the Yale B database using 4 different methods: (a) PCA (b) FLD (c) ZMI (d) LMI. The algorithms were trained with the first subset without Noise and tested with all the subsets after adding noise	60
3.11	Result of the noise test for the FERET database using 4 methods (a) effect of Gaussian noise (b) effect of Salt and Pepper noise	60
4.1	Bad illumination can cause between class similarities overcome the inter-class similarities. Images in each row belong to the same subject but with different illumination	65
4.2	Extracting the illumination invariants based on the illumination-reflection model by applying a HPF on the logarithm of the image	69
4.3	Extracting the illumination invariants by using LPFs (maximum filter) ...	70
4.4	Recognition results using different lowpass and highpass filters in the two structures. (Extended Yale B database is used with PCA as feature extractor and SVM for classification)	70
4.5	The results of applying 3×3 and 5×5 maximum filters on Lena image	71
4.6	Illumination invariants achieved by different lowpass and highpass filters	73
4.7	Edge enhancement property of the Maximum filter. The right image is the	74

	result of applying Maximum filter on the left image	
4.8	Different steps in extracting the illumination invariants	75
4.9	One of the subjects in the Yale B database with 5 corresponding subsets and the illumination invariants generated using the proposed method: (a) subset 1 (θ upto 12°) (b) subset 2 (θ upto 25°) (c) subset 3 (θ upto 50°) (d) subset 4 (θ upto 77°)	83
4.10	One of the samples in CMU-PIE database and its illumination invariants (a) lights-off group (b) lights-on group (c) and (d) are the corresponding illumination invariants to (a) and (b)	84
5.1	Samples of images with different angle between the light source and camera axes, θ , (a): $\theta = 0^\circ$, (b) $\theta = 40^\circ$ (c) $\theta = 90^\circ$, (d), (e) and (f) are corresponding illumination invariant images, respectively	90
5.2	Democratic Voting Scheme (DVS)	93
5.3	The Weighted Voting Scheme (WVS)	95
5.4	ROC curve for face verification where different methods of WVS, DVS and mPCA are compared. 5×5 grids are used for all methods	105
5.5	Performance of different methods with different number of samples of subset 1 in train set varying from 1 to 7 and different subsets as test probes. (a) subset 2 (b) subset 3 (c) subset 4 (d) subset 5. Extended Yale B database is used with SVM classifiers	108
5.6	Occluded images and their illumination invariants. Input images are occluded with a pair of sunglasses, striped and black scarves, wall, another face and back of another person's head	110
5.7	Images in different subsets of Yale B database occluded using random circles and their corresponding illumination invariants	112
5.8	Performance of different recognition methods under varying amount of occlusion. (a) back of head (b) another face (c) wall overlapping from left and right (d) wall overlapping from top and bottom	113

List of Tables

3.1	Bamieh Moment Invariants	31
3.2	Classification parameters for different moment invariants	40
3.3	Recognition time for different methods	41
3.4	Recognition Accuracy for PZMI method using different orders of the moment invariants	42
3.5	Classification results for different classifiers	47
4.1	A comparison between different filters. Values show the average Euclidian distance between the images with illumination variations and the image with frontal lightning. Maximum filter shows to have the best results.	74
4.2	Recognition rates for different filters and SVM as classifier	76
4.3	Recognition rates for different filters and k NN with $k = 1$ as classifier	77
4.4	Recognition rates for different filters and k NN with $k = 3$ as classifier	78
4.5	Recognition rates for different filters and k NN with $k = 5$ as classifier	79
4.6	Recognition rates for different filters and fuzzy k NN with $k = 1$ as classifier	80
4.7	Recognition rates for different filters and fuzzy k NN with $k = 3$ as classifier	81
4.8	Recognition rates for different filters and fuzzy k NN with $k = 5$ as classifier	82
4.9	Recognition results for 200 simulations on CMU-PIE database	86
4.10	Recognition rates using different methods for CMU-PIE database	86

4.11	Recognition rates using different methods for Yale B database (10 subjects)	87
4.12	Recognition rates using the Extended Yale B database (38 subjects)	87
5.1	Recognition rates using different methods for Yale B database (10 subjects)	103
5.2	Recognition rates using the Extended Yale B database (38 subjects)	103
5.3	Recognition rates using different methods for CMU-PIE database	103
5.4	Threshold values and verification rates for different FAR and FRR conditions	106
5.5	Recognition rate vs. number of training samples from light on and off groups for CMU-PIE database	107
5.6	Recognition rates for images occluded with sunglasses, striped scarf and black scarf	112
5.7	Effect of occlusion on recognition results using different methods for Ext Yale B database (38 subjects)	114
5.8	Recognition results for combination of Ext. Yale B and CMU-PIE databases	115

List of Abbreviations

2D	Two Dimensional
3D	Two Dimensional
HE	Histogram Equalization
GIC	Gamma Intensity Correction
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
FLD	Fisher's Linear Discriminant
ICA	Independent Component Analysis
LS-ICA	Locally Salient Independent Component Analysis
KPCA	Kernel Principal Component Analysis
KFD	Kernel Fisher's Discriminant
HMI	Hu Moment Invariant
RMI	Regular Moment Invariant
BMI	Bamieh Moment Invariant
ZMI	Zernike Moment Invariant
PZMI	Pseudo Zernike Moment Invariant
TZMI	Teague Zernike Moment Invariant
NZMI	Normalized Zernike Moment Invariant
NPZMI	Normalized Pseudo Zernike Moment Invariant
LMI	Legendre Moment Invariant
DLA	Dynamic Link Architecture

EBGM	Elastic Bunch Graph Matching
ANN	Artificial Neural Networks
k NN	K Nearest Neighbors
RBF	Radial Basis Function
SVM	Support Vector Machine
HMM	Hidden Markov Model
FV	Feature Vector
VJ	Viola and Jones face detector
II	Illumination Invariants
QI	Quotient Image
QIR	Quotient Illumination Relighting
SQI	Self-Quotient Image
MQI	Morphological Quotient Image
mPCA	Modular PCA
mLDA	Modular LDA
DVS	Democratic Voting Scheme
WVS	Weighted Voting Scheme
AWG	Adaptive Weight Generator
ROC	Receiver Operating Characteristic
FAR	False Accept Ratio
FRR	False Reject Ratio

Chapter 1

Introduction

Among all biometric methods for human recognition and human authentication, the fingerprint and face recognition have been used more frequently. Finger print recognition is very old and has been used for long, but face recognition has become more popular for the last twenty years. Face recognition is the main biometric used by human beings. When two people meet each other, their brains run a variety of biometrics based on height, age, hair color and style, skin color, *etc.* However, the final decision of the other person's identity is made based on his/her face; hence face is assumed to be the part of the body that carries more information than other parts.

Automatic face recognition has many commercial and security applications in identity validation and recognition and has become one of the hottest topics in the area of image

processing and pattern recognition since 1990. Availability of feasible technologies as well as the increasing request for reliable security systems in today's world has been a motivation for many researchers to develop new methods for face recognition. Automatic recognition of human faces continues to attract researchers from different areas such as computer vision, image processing, pattern recognition, neural networks, computer graphics, and psychology and also has been increasingly accepted by the general public to be used in authentication, security and law enforcement.

1.1 Face Recognition

In January 2001 police in Tampa Bay, Florida, used a face recognition software at Super Bowl XXXV to search for potential criminals and terrorists attending the event [1]. The system found 19 people with pending arrest warrants. This is a practical example of using a face recognition system in security applications. In automatic face recognition we desire to either identify or verify one or more persons in still or video images of a scene by means of a stored database of faces [2]. One of the important features of face recognition is its non-intrusive and non-contact property that distinguishes it from other biometrics like iris or finger print recognition that require subjects' participation.

An automatic face recognition system is usually a procedure of four main stages. In most cases these four stages or blocks are namely: pre-processing, face detection, feature extraction and finally classification. The input images obtained from image acquisition devices *e.g.* camera, might not be suitable for the application due to noise or illumination conditions. Therefore, first step is the preprocessing stage to remove noise, normalize the

color, and fix the illumination. Then we must find the face in the image. Some face detection algorithms are presented in chapter 2 of this thesis. Next step would be to extract some predefined features in order to make a feature vector. These features must include distinctive information about each person in the database to enable the recognition of the individual based on these features. Finally, the last stage is the classifier where we intend to recognize an unknown sample by assigning a class to its feature vector based on the database of features that we have gathered from previously seen samples.

1.2 Identification vs. Verification

Biometric methods are used for either one of two functions, identification or verification. The selection and implementation of the technology are closely related to this objective. In identification the goal is to identify an individual based on comparison of features collected against a database of previously collected samples. In other words, systems which are designed for the purpose of identification will answer the question: “Who the subject is?” (Figure 1.1). In the verification applications, on the other hand, we desire to verify whether the subject is the person that they claim to be. This is done by validating the collected features against a previously collected feature sample for the individual in our library. Biometric methods designed for verification purposes answer the question: “Is the subject who he says he is?”

Face Authentication/Verification (1:1 matching)



Face Identification/Recognition (1:N matching)

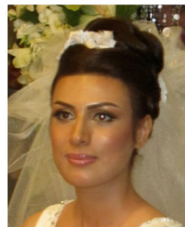


Figure 1.1: An example of authentication vs. identification.

1.3 Applications

Today's technologies have made face recognition a possible reliable solution in applications like driver's license, passport or national ID verification or security access to ATM machines, database or medical records [2]. Some of the applications of automatic face recognition can be categorized as follows.

Entertainment: Video games, virtual reality, training programs, human robot interaction, human computer interaction

Smart cards: Drivers' licenses, entitlement programs, immigration, national ID, passports, voter registration, welfare fraud

Information security: TV parental control, personal device logon, desktop logon, application security, database security, file encryption, intranet security, internet access, medical records, secure trading terminals

Law enforcement and surveillance: Advanced video surveillance, CCTV control, portal control, post event analysis, shoplifting, suspect tracking and investigation

1.4 Challenges

During the last two decades several face recognition algorithms and systems have been proposed and some major advances have been achieved. As a result, the performance of face recognition systems under controlled conditions has now reached a satisfactory level. These systems, however, face some challenges in environments with variations in illumination, pose, expression, *etc.* The performance of a face recognition system is directly related to the amount of variation observed in face images. The better a system can eliminate the effect of these variations, it yields better recognition results and hence it would be a more reliable system. The major variations that challenge the face recognition algorithms are as follows.

1.4.1 Illumination

Whether the face is indoor or outdoor or under certain illumination condition, it has been observed that changes in lighting condition shape different shading and shadows on the face due to the 3D shape of human faces. This may weaken some face features or may result in too bright or too dark parts in images. These variations can be larger than the variations due to personal identity and hence yield a lower recognition rate. (Figure 1.2)

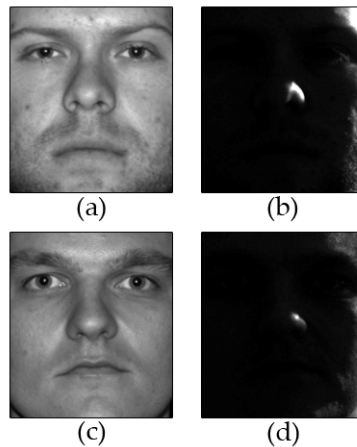


Figure 1.2: Bad illumination can cause between-class similarities overcome the inter-class similarities. Images in each row belong to the same subject but with different illumination.

1.4.2 Head Pose

In most cases, the face recognition systems are trained with a group of images under supervised conditions. One of these conditions would be the angle between the axis of camera and the orientation of the face which we desire to be minimum or zero. In other words we desire the face images to be taken in frontal pose where the person is looking at the camera. In this position maximum information of the face is present, therefore, a better recognition rate is expected. This ideal head position would be a valid assumption in applications like access to a specific facility, but not in a subway surveillance system. The three dimensional structure of human faces will occlude some features resulting an information loss which leads to classification errors in the face recognition system.

1.4.3 Facial Expressions

As it was mentioned, most face recognition algorithms are trained with normal and neutral face images. Facial accessories like glasses, facial hair (mustache and bear), and

also emotional expressions like laughing, smiling or frowning will change some of the features in the face and might affect the result of the classification. An ideal automated face recognition system is desired to overcome this problem, for example, by proper feature selection.

1.4.4 Occlusion

Generally human beings are able to recognize another person even if they are wearing sunglasses or a scarf. This is a challenge for the automatic face recognition systems that are designed to replace human brains. Partial occlusion by another object, another person's face, sunglasses or scarves is a common problem in many face recognition applications. These occlusions will affect some of the face features due to covering parts of the face image and might degrade the recognition performance. Many methods tend to solve this problem by segmenting the occluded parts of the image and disregarding their corresponding features.

1.4.5 Inter-Class Similarities

Distinguishing between two identical twins is not impossible but still sometimes a hard job even for their parents. So inter-class similarities are not only a challenge for biometrics technologies, but they also confront human brains. It is the problem of differentiating between two different subjects that have very similar features. In many cases multi-biometric methods like combining face and fingerprint recognition will improve the performance.

1.5 Research Objective and Contributions

The objective of this research is designing a reliable automated face recognition system which is robust under varying conditions of noise level, illumination and occlusion. This work's original contribution is in unique combination of illumination invariant feature extraction and weighted local matching to achieve promising recognition results in presence of both illumination variations and occlusions.

A new method for illumination invariant feature extraction is proposed and the results are compared to similar algorithms. The proposed technique is computationally efficient and does not require any prior information about the face model or illumination. Unlike many methods in the literature, our method does not need many training samples to achieve illumination invariants and can be applied on each single image to obtain the illumination invariants.

A weighted voting scheme is proposed to enhance the performance under illumination variations and also cancel occlusions. The proposed method uses mutual information and entropy of the images to generate different weights for a group of ensemble classifiers based on the input image quality. The method yields outstanding results by reducing the effect of both illumination and occlusion variations in the input face images.

1.6 Outline of the Thesis

Different building blocks of an automated face recognition system as well as a brief review about available methods and algorithms for each block are presented in chapter 2

of the thesis. Many different face image databases have been developed by researchers in the area. Each of these databases accentuates one or more of the condition variations. Some of these databases are also introduced in this chapter.

Moment invariants are pattern sensitive features and are used in pattern recognition applications. In the third chapter different moment invariants are used to extract features from human face images. A comparison of the performance of different moment invariants with different classifiers and databases is presented in this chapter. The effect of noise and misalignment in the detected faces is also studied.

An efficient illumination invariant feature extraction method is proposed in chapter 4 of this thesis and its recognition results are compared with similar methods for cancelling illumination available in the literature.

In chapter 5 a new weighted voting scheme based on mutual information and image entropy is proposed. The proposed method makes a decision based on the different parts of the face image and disregards the features corresponding to occluded or poorly illuminated parts.

Finally, conclusion remarks and future work are provided in chapter six of the thesis.

Chapter 2

Face Recognition Methods

While we are almost sure that most of the human skills for face perception do not exist in babies, there are studies showing an instinctive affinity to pay attention to faces from birth [3]. This is the first step towards the development of visual perception skills such as identifying friendly others and pre-verbal communications. As the person grows up, understanding faces develops from basic understanding of sensory information to receive age, gender details about the individual, to the ability to remember significant information such as their name or past familiarity with the person. Face perception plays an important role in the human perception system and is an ordinary task for human brains, but designing a comparable fully automated computer system is still an open research area. It continues to attract researchers from different areas such as computer

vision, image processing, pattern recognition, neural networks, computer graphics, and psychology and several methods and techniques have been proposed to fulfill this task.

2.1 A Face Recognition System

An automatic face recognition system is usually a procedure of four main stages. In most cases these four stages or blocks are namely: pre-processing, face detection, feature extraction and finally classification. The flow diagram of a face recognition system is presented in Figure 2.1. The input images obtained from image acquisition devices *e.g.* cameras, might not be suitable for recognition due to noise or illumination conditions. Therefore, first step is the preprocessing stage to remove noise, fix the illumination and normalize the color. Then faces should be located and segmented in input images. Some face detection algorithms are presented in this chapter. Next step would be to extract some predefined features in order to make a feature vector. These features must include distinctive information about each person in the database so we can recognize the individual based on these features. And finally the last stage is the classifier where we intend to recognize an unknown sample by assigning a class to its feature vector based on the database of features that we have from previously seen samples.

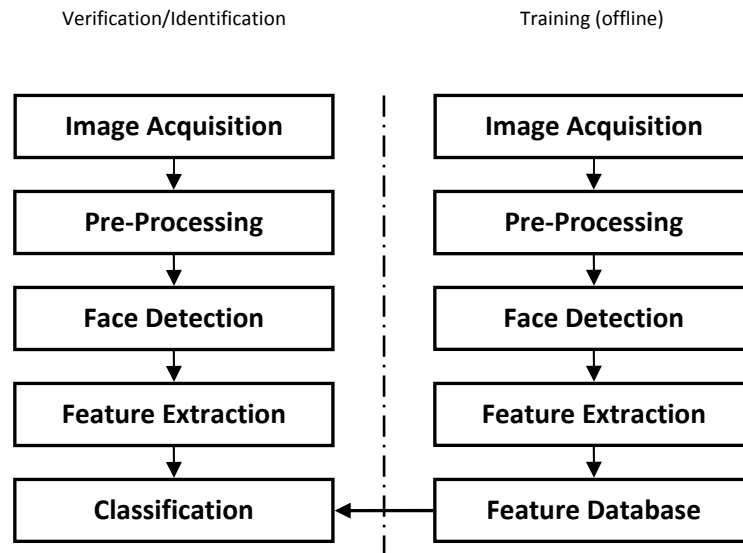


Figure 2.1: Flow diagram of face recognition process.

2.2 Preprocessing

The first step in most face recognition systems as well as many other computer vision applications is preprocessing or image quality enhancement. The input images acquired via still or video cameras might suffer from noise, bad illumination or unrealistic color. Therefore, noise removal might be a necessary block in some cases. Histogram equalization is the most common method used for image enhancement when images have illumination variations [4]. Even for images under controlled illumination, histogram equalization improves the recognition results by flattening the histogram of pixel intensities of the images. Shan *et al.* proposed Gamma intensity correction method in [5] and used it for illumination normalization along with histogram equalization. They also presented a region-based method for equalizing histogram and gamma locally in small portions of an image in their work.

2.3 Face Detection

Another step in face recognition process is to locate the face in an input image. This requires the algorithm to locate a single face with a known scale and orientation [2]. Face detection at frame rate is an impressive goal that has an apparent application to practical face tracking and real time face recognition. Papageorgiou *et al.* [6] proposed a general object detection scheme which uses a wavelet representation and statistical learning techniques. Osuna *et al.* apply Vapnik's support vector machine technique to face detection in [7], and Romdhani *et al.* improved it by creating reduced training vector sets for their classifier [8]. Schneiderman and Kanade's statistical method considered profile detection [9], but their method considered only three face orientations, and each orientation is treated as a different object, Rowley and Kanade used neural network-based filters [10], while Fleuret and Geman attempted a coarse-to-fine approach to face detection, focusing on minimizing computation in their approach [11]. Viola and Jones [12] used the concept of an integral image, along with a rectangular feature representation and a boosting algorithm as its learning method. They obtained good result of detecting faces at 15 frames per second. Their method minimized computation time while achieving high detection accuracy and is approximately 15 times faster than previous approaches based on similarity functions.

2.4 Feature Extraction

Extraction of applicable features from the human face images is an important part of the recognition. Therefore choosing the proper feature extractor is crucial when designing

a face recognition system with high recognition rate. There are two main approaches for the feature extraction problem [2]. The first one is based on extracting geometrical and structural facial features like the shapes of the eyes, nose, mouth and the distance between these. These methods are not affected by irrelevant information in the images, but are sensitive to unpredictability of face appearance and environmental conditions. In the second method which is the holistic approach, features from the whole image are extracted. Since the features are global in the whole image, irrelevant elements like background and hair might affect the feature vectors and cause erroneous recognition results.

Eigenfaces method by Turk and Pentland [13] and Fisherfaces method by Belhumeur *et al.* [14] are the most referred algorithms among appearance-based holistic approaches and have proven to be effective in experiments with large databases. The basic idea behind these algorithms is the principal component analysis that can be used to decrease the feature space dimension while keeping the characteristics of the images. In principal component analysis (PCA), the mean of the training images is calculated and is subtracted from the whole training set, and then eigenvalues of the covariance matrix are calculated. n eigenvectors regarding n greatest eigenvalues are calculated and called eigenfaces. Any image in the image space will be projected on those n basic eigenfaces. Those directions or eigenfaces are assumed to be the directions or planes with the maximum data scattering. In Fisher's Linear Discriminant [14] however, the goal is to find directions with the maximum ratio of between-class scatter to within-class scatter. This can be done by finding the generalized eigenvalues and eigenvectors of the between-

class scattering and within-class scattering matrices. Images are projected on the calculated hyperplanes called fisherfaces, which leads to better classification and higher recognition rates in most of the experiments. Independent component analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents by assuming the mutual statistical independence of the non-Gaussian source signals. This technique has been proposed by Liu and Wechsler [15] to minimize both second-order and higher-order dependencies and also reduce dimensionality. Kim *et al.* proposed an effective part-based local representation method named locally salient ICA (LS-ICA) method for face recognition that is robust to partial occlusion. LS-ICA only employs locally salient information from important facial parts in order to maximize the benefit of applying the idea of recognition by parts [16].

The face manifold in subspace is not restricted to be linear. Kernel methods [17-19] can be applied to extract features from images with large variations. This is based on non-linear mapping of the input image, transferring it from the input space to a higher dimensional feature space and applying the PCA, ICA or linear discriminant analysis method in the feature space. Bach and Jordan [20] use contrast functions based on canonical correlations in a reproducing kernel Hilbert space for ICA. Kernel PCA (KPCA) is developed by Scholkopf [17], and Kernel Fisherfaces (KFD) is later proposed by Mika [18] and Baudat [19]. Mika's work is mainly focused on two-class problems, while Baudat's algorithm is applicable for multi-class problems. Lu *et al* [21] use kernel direct discriminant analysis as the nonlinear substitute for Fisherface method. Yang studied kernel eigenfaces and kernel Fisherfaces in [22]. Yang *et al* combined KPCA

with FLD in [23]. Wiskott *et al.* [24] used a feature-based graph matching approach which is quite successful.

Compared to holistic approaches, feature-based methods are assumed to be less sensitive to variations in illumination and viewpoint as well as inaccuracy in face localization. However, finding a proper feature extraction technique needed for this type of approach is crucial because some methods might not be as reliable or accurate for our application [2]. Most eye localization techniques for example, assume some geometric and textural models and do not work if the eye is closed.

An image moment in general is a particular weighted average of the image pixels' intensities that can be used to describe the image based on its properties. Moment features are invariant under scaling, translation rotation and reflection. Moment invariants have been proven to be useful in pattern recognition applications due to their sensitivity to the pattern features. In 1961 Hu [25] introduced the first set of algebraic moment invariants (HMI). Regular moment invariants (RMI) [26] are the simplest and perhaps the most famous moment invariants. Bamieh and De Figueiredo [27] derived Bamieh moment invariants (BMI) which have small feature vectors so that they are computationally efficient. Zernike and pseudo Zernike orthogonal polynomials [28] are the basis of the Zernike moment invariant (ZMI) [29] and pseudo Zernike moment invariant (PZMI) [30]. Teague-Zernike moment invariants (TZMI) [31] were proposed by Teague *et al.* as a method for calculating ZMIs. Orthogonal moments like Legendre moments (LMI) [32] and ZMI can be used to represent an image with the minimum amount of information redundancy. Haddadnia *et al.* [33] used PZMI with an RBF neural

network classifier for face recognition. A comparative study on different moment invariants in face recognition is presented in [34]. Rashidy Kanan *et al.* proposed a novel face recognition approach based on adaptively weighted patch pseudo Zernike moment array (AWPPZMA) when only one exemplar image per person is available [35]. Details of some of the approaches are described in the following.

2.4.1 Correlation

Correlation-based recognition is one of the simplest methods for identifying face images. Assuming that the unknown input image and all the gallery images are perfectly aligned, the class that maximizes the normalized correlation between the input and gallery images will be assigned to the subject. It can also be used as a face detector where the cropped gallery image will be shifted as a small window over the test image and the maximum normalized correlation is calculated. This works best where faces in the test and train images have equal sizes and head pose, illumination and facial expressions are similar during test and train.

2.4.2 Principal Component Analysis

PCA is used for reducing the dimensionality of the feature space by transforming a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. It is based on the assumption that most of the information about classes is in the directions along which the variations or scatterings are maximum. For a given database of images with N images, the first step is to find the mean of the images in the database and subtract this mean image, μ , from all the images, x_i ($i = 1, 2,$

..., N). Then the covariance matrix, C , is calculated as:

$$C = \frac{1}{N} \sum_1^N (x_i - \mu)(x_i - \mu)^T \quad (2.1)$$

For M largest eigenvalues of C we have:

$$CT_i = \lambda_i T_i \quad i = 1, \dots, M \quad (2.2)$$

where, T_i 's are the eigenvectors or eigenfaces. Input images are projected on the direction of eigenfaces and then classified based on minimum distance or support vector machines.

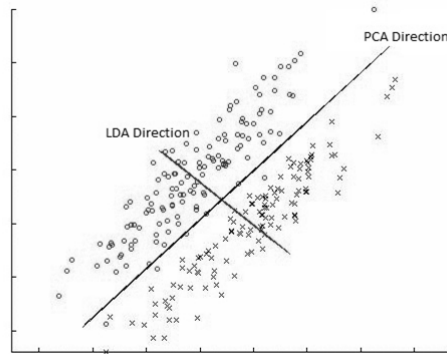


Figure 2.2: PCA maps the samples to the direction with maximum variation where LDA finds the direction with maximum between-class variation to within-class variation.

2.4.3 Linear Discriminant Analysis

In PCA we emphasize on directions related to large eigenvalues or in other words large variation. In some cases, some of the important information might be in directions with small variation. What Fisher linear discriminant (FLD) does is maximizing this objective (Figure 2.2). as it can be seen in Figure 2.2 both classes are scattered in the same direction which is found by PCA. LDA tries to maximize the direction with

maximum between-class scattering, hence better result is achieved.

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (2.3)$$

where, S_B is the between-classes scatter matrix and S_W is the within-classes scatter matrix:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T \quad (2.4)$$

$$S_W = \sum_{i=1}^c \sum_{x_k \in \chi_i} (x_k - \mu_i) (x_k - \mu_i)^T \quad (2.5)$$

where, χ_i is the i th class and N_i is the number of samples in it.

2.4.4 Independent Component Analysis

Independent component analysis (ICA) can be viewed as a generalization of PCA. Given a multivariate signal of statistical independent non-Gaussian sources, ICA can be used to separate the input signal into its additive subcomponents. As mentioned before, PCA tries to minimize the training covariance. In the test set this is done by mapping the samples by an orthogonal transformation. ICA is not satisfied with orthogonality of the outputs and transforms the data into a set of statistically independent signals. The advantage of ICA is its ability to minimize both second order and higher order dependencies in the input, while PCA can decorrelate data only based on second-order statistic.

2.4.5 Kernel Principal Component Analysis

Kernel principal component analysis (KPCA) is a nonlinear extension of PCA that aims to extract the nonlinear manifolds in feature space by using kernel methods. In a d -dimensional feature space N points cannot be linearly separated in general where $d < N$, while they can almost always be linearly separated in $d > N$ dimensions by creating a hyperplane that divides the points into different clusters. In kernel PCA, we make a kernel which represents the mapping into higher dimension and coming back to the original dimension, only we do not calculate the mapping function. KPCA computes not the principal components themselves, but the projections of the input data onto those component directions.

2.4.6 Wavelets

Jones and Palmer [36] showed that the real part of the complex Gabor function is similar to the receptive fields of simple cells found in a cat's cortex. This biological motivation as well as its ability to yield distortion tolerant feature spaces has made Gabor wavelet popular solution in pattern recognition applications such as fingerprint recognition, handwritten recognition and texture segmentation [37-39]. The Gabor space is very useful in image processing tasks due to its optimized resolution in both the spatial and frequency domains.

Gabor wavelets were used for face recognition first as in dynamic link architecture (DLA) proposed by Lades *et al.* [40]. In DLA, local features are extracted at deformable nodes using Gabor wavelets making a rectangular graph representing the faces which is

called Gabor jets. Elastic Bunch Graph Matching (EBGM) is an extension of DLA proposed by Wiskott *et al.* [24]. In EBGM graph nodes are located at some facial landmarks. Gabor wavelets have also been used as holistic feature extractors for face recognition [41-43] where the wavelet is applied on the whole image for feature representation.

2.5 Classification

Having the feature vectors for all the samples in the train and test sets, the next step would be to design a classifier. Many different classifiers have been used in face recognition systems. Here we have a brief look at some of the most popular ones.

2.5.1 k -Nearest Neighbors

The k -nearest neighbors algorithm is one of the simplest machine learning algorithms used by many researchers as the classifier in their face recognition systems. A new object is classified and assigned to a class by a majority vote of its k nearest neighbors, where k is a typically small positive integer. If $k = 1$, then the object is simply assigned to the class of its nearest neighbor. Euclidian distance is usually the measure calculated in order to classify the new object. A comparison of different distance measures for PCA based face recognition is available in [44].

2.5.2 Neural Networks

Inspired by the structure of biological neural networks, artificial neural networks (ANN) have been widely used as the classifier in many pattern recognition applications

and are very common in face recognition systems. In an artificial neural network, a group of artificial neurons are connected together and form a network of neurons imitating the biological neural networks. Many different architectures for neural networks have been proposed by researchers such as multilayer perceptrons, back propagation, radial basis function (RBF) neural networks and *etc.* Dynamic link architecture (DLA) is one of the most common methods based on neural networks that is applied on the features extracted from face images to dynamically group the neurons to higher order and results in more robust face recognition against environment variations [40,45].

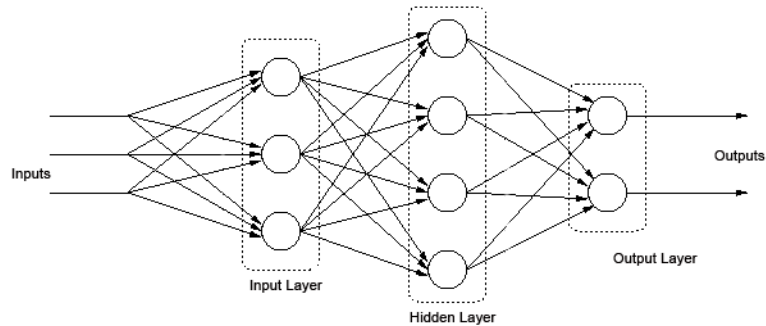


Figure 2.3: Three layer perceptron

2.5.3 Support Vector Machines

Support vector machines (SVM) simultaneously minimize the classification error and maximize the geometric margin; so they are known as maximum margin classifiers [46]. An SVM makes a separating hyperplane in the feature space which maximizes the margin between the data sets. Two parallel hyperplanes are constructed to calculate the margin, one on each side of the separating one. These hyperplanes are pushed up against the two data sets, so that a good separation is achieved by the hyperplane that has the

largest distance to the neighboring support vectors of both classes (Figure 2.4). The larger the margin or distance between these parallel hyperplanes is, the larger is the probability that unknown samples are classified correctly [47].

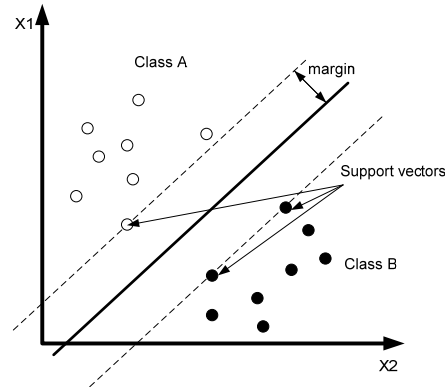


Figure 2.4: Maximum-margin hyperplane for a SVM classifier. Samples on the margin are called the support vectors.

2.5.4 Hidden Markov Model

The hidden Markov model as a double stochastic process can efficiently model the generation of sequential data. HMM has successfully been used in speech, face and handwritten recognition. There are basically two different approaches in using HMM as a classifier. In the first approach, a single HMM models all classes and based on the different paths, classes are distinguished. In the second one, known as model discriminant approach, a separate HMM is used for each class. 2D embedded HMM proposed by Nefian [48] contains a set of super states and each super state is related to a set of embedded states. Super states represent main facial regions and embedded states within each super state describe more detail about their corresponding facial region. The observation sequence of an unknown image is fed to all of the trained HMMs each

belonging to a class. Each HMM produces a conditional probability, and the HMM with the highest probability decides the class that the person belongs to.

2.6 Databases

Several face image databases have been produced by researchers in the field. Each of these databases emphasizes on some of the factors that exist in real application situations. Some of the databases that have been used in our work are described in this section.

2.6.1 The AT&T Database

The AT&T face image database (formerly known as the ORL database) contains a set of face images taken between April 1992 and April 1994 from 40 people [49]. There are 10 different images of each person taken at different times. Some of the images have varying illumination, facial expressions like open or closed eyes or glasses, or have different emotion expressions like smiling or frowning. Some of the images may also have rotations up to 20 degrees. All of the images have a dark homogeneous background and the subjects are in an upright, frontal position. The size of each image is 92×112 pixels, with 256 grey levels per pixel. Some samples of the images in this database are shown in Figure 2.5.



Figure 2.5: Sample images of two subjects the in the AT&T database.

2.6.2 The FERET Database

The FERET database of images [50] includes a large number of images and subjects with different variations. The FERET program ran from 1993 through 1997 to develop automatic face recognition capabilities that could be employed to assist security, intelligence and law enforcement personnel in the performance of their duties. The FERET database of images consists of 14051 eight bit grayscale images of human heads with views ranging from frontal to left and right profiles. The size of the pictures is 256×384 pixels. The pictures are from 1209 people, taken at different time, illuminations, and facial expressions.



Figure 2.6: One of the samples of the FERET database in frontal position.

2.6.3 The Yale B Database

The Yale B Face database [51] contains 5760 images from 10 individuals. Each subject has been pictured under 576 viewing conditions which are nine poses and 64 illuminations per pose. Single light sources have been used in different angles for the illumination variations.

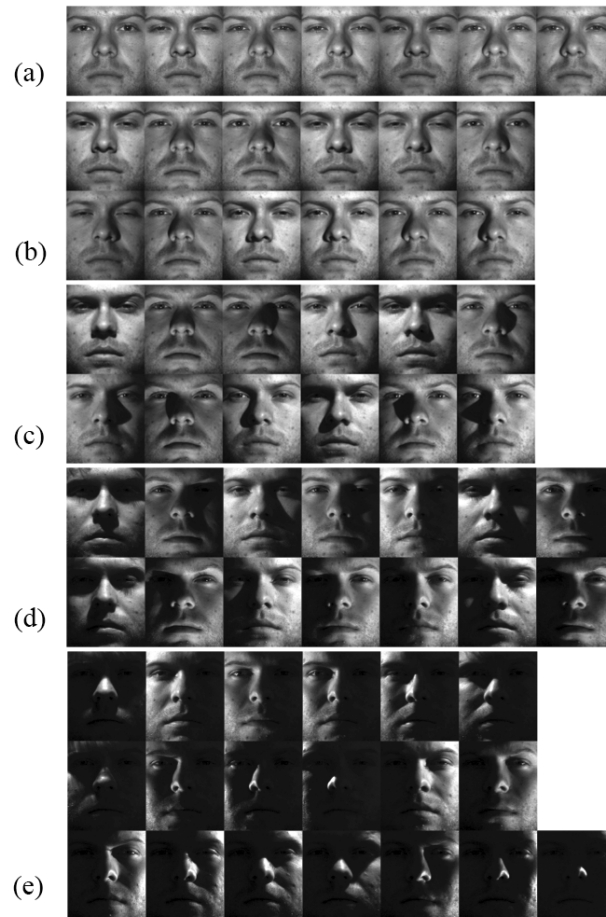


Figure 2.7: One of the subjects in the Extended Yale B database with its 5 subsets.

2.6.4 Extended Yale B Database

The Extended Yale B database [52] is an expanded version of the Yale B Database with 28 more subjects and contains 21888 single light source images of 38 subjects each seen under 576 viewing conditions similar to the Yale B database. For every subject in a particular pose, an image with background illumination was also captured. Both Yale B and Extended Yale B databases are divided into five different subsets according to the

angle between the light source direction and the camera axis. The five subsets for one of the subjects in the database are presented in Figure 2.7.

2.6.5 CMU-PIE Database

The CMU Pose, Illumination, and Expression (CMU-PIE) database [53] consists of 41368 images of 68 people. Each person has been imaged under 13 different poses, 43 different illumination conditions with lights ON and lights OFF, and with 4 different expressions.



Figure 2.8: One of the samples in CMU-PIE database (a) lights-off group (b) lights-on group.

Chapter 3

Moment Invariants for Human Face Recognition

Moment Invariants are pattern sensitive features and are used in pattern recognition applications. In this chapter different moment invariants have been used to extract features from human face images for recognition application. Moment invariants of Hu (HMI), Bamieh (BMI), Zernike (ZMI), Pseudo Zernike (PZMI), Teague-Zernike (TZMI), Normalized Zernike (NZMI), Normalized Pseudo Zernike (NPZMI), Legendre (LMI) and also regular Moment Invariant (RMI) have been applied to different face databases and the results have been compared.

3.1 Moment Invariants

Moment features are invariant under scaling, translation rotation and reflection. Moment Invariants have been proven to be useful in pattern recognition applications due

to their sensitivity to the pattern features. Some of the most famous moment invariants are briefly described in this section.

3.1.1 Regular Moment Invariants

RMI [26] is defined as:

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad (3.1)$$

where $f(x,y)$ is the image and M_{pq} is its two dimensional moment of order $(p+q)$. The center of an image can be located using first order moments as:

$$\bar{x} = \frac{M_{10}}{M_{00}}, \quad \bar{y} = \frac{M_{01}}{M_{00}} \quad (3.2)$$

Central moment is defined as:

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (3.3)$$

If we divide M_{pq} by $[M_{20} + M_{02}]^\gamma$ it would be invariant to changing in the size or scaling. γ is $[p+q+2]/4$. Therefore, we can write scale-invariant moments as:

$$\mu_{pq} = \frac{1}{[M_{20} + M_{02}]^\gamma} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (3.4)$$

In order to make μ_{pq} invariant under rotation and reflection we estimate the angle under which the image has been rotated using second order moments as:

$$\theta = \frac{1}{2} \tan^{-1} \left[\frac{2M_{11}}{M_{20} - M_{02}} \right] \quad (3.5)$$

Then we can calculate regular moment invariants (RMI) as the following equation:

$$(RMI)_{jk} = \sum_{r=0}^j \sum_{s=0}^k (-1)^{k-s} \binom{j}{r} \binom{k}{s} (\cos \theta)^{j-r+s} \times (\cos \theta)^{k+r-s} \mu_{(j+k-r-s), (r+s)} \quad (3.6)$$

where μ_{pq} is the scale invariant central moment and θ is the rotation angle.

3.1.2 Hu Moment Invariants

Hu [25] derived the following set of moment polynomials using algebraic invariants theory:

$$I_{p-r,r} = \sum_{l=0}^r (-j)^l \binom{p-2l}{l} \sum_{k=0}^r \binom{r}{k} \mu_{p-2k-l, 2k+l} \quad (3.7)$$

$$p-2r > 0, \quad j = \sqrt{-1}$$

or

$$I_{p-r,r} = |I_{p-r,r}| e^{j\varphi_{p-r,r}} \quad (3.8)$$

where

$$\varphi_{p-r,r} = \tan^{-1} \left(\frac{\operatorname{Re}(I_{p-r,r})}{\operatorname{Im}(I_{p-r,r})} \right) \quad (3.9)$$

HMIs can be defined as:

$$(HMI)_0 = |I_{p0}|^2 \quad (3.10)$$

$$(HMI)_r = |I_{p-r,r}|^2, \quad r = 1, 2, 3, \dots, p-2r \quad (3.11)$$

3.1.3 Bamieh Moment Invariants

Another set of algebraic moment invariants are Bamieh moment invariants presented in [27]. They have small-sized feature vector that makes the classification part much

faster. BMIs are based on moment tensors as following where i, j, k can take the values of 1 or 2:

$$M^{ijk} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^i x^j x^k \dots f(x^1, x^2) dx^1 dx^2 \quad (3.12)$$

The BMIs are then calculated based on these moment tensors. BMIs up to 4th order are listed in table 1 using scale-invariant moments μ_{pq} .

Table 3.1: Bamieh Moment Invariants

2nd order	$(BMI)_1 = \mu_{20}\mu_{02} - \mu_{11}^2$
3rd order	$(BMI)_2 = (\mu_{30}\mu_{03} - \mu_{12}\mu_{21})^2 - 4(\mu_{03}\mu_{12} - \mu_{21}^2)(\mu_{21}\mu_{30} - \mu_{12}^2)$
4th order	$(BMI)_3 = \mu_{40}\mu_{04} - 4\mu_{13}\mu_{31} + 3\mu_{22}^2$ $(BMI)_4 = \mu_{40}\mu_{22}\mu_{04} - 2\mu_{13}\mu_{22}\mu_{31} - \mu_{13}^2\mu_{40} - \mu_{04}\mu_{31}^2 - \mu_{22}^3$

3.1.4 Zernike Moment Invariants

Zernike moment invariants [29] are based on Zernike polynomials [28]. These polynomials are set of complex valued polynomials that can be written as:

$$V_{nL}(x, y) = V_{nL}(r \cos \varphi, r \sin \varphi) = R_{nL}(r) e^{jL\varphi} \quad (3.13)$$

where $x^2 + y^2 = r^2$ and n is the degree of the polynomial and L is the angular dependence. r and φ are the scale and rotation in a cylinder representation. Zernike moments defined as:

$$A_{nL} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^\infty f(r \cos \varphi, r \sin \varphi) R_{nL}(r) \times e^{-jL\varphi} r dr d\varphi \quad (3.14)$$

where $R_{nL}(r)$ can be written as:

$$R_{nL}(r) = \sum_{k=L}^n B_{nLk} r^k \quad (3.15)$$

and

$$B_{nLk} = \frac{(-1)^{\frac{n-k}{2}} \times \left(\frac{n+k}{2}\right)!}{\left(\frac{n-k}{2}\right)! \times \left(\frac{L+k}{2}\right)! \times \left(\frac{k-L}{2}\right)!} \quad (3.16)$$

3.1.5 Pseudo Zernike Moment Invariants

Pseudo Zernike polynomials [30] can be defined as:

$$V_{nm}(x, y) = R_{nm}(x, y) e^{jm \tan^{-1}\left(\frac{y}{x}\right)} \quad (3.17)$$

where

$$x^2 + y^2 \leq 1, \quad n \geq 0, \quad |m| \leq n \quad (3.18)$$

and

$$R_{nm}(x, y) = \sum_{s=0}^{n-|m|} D_{n,|m|,s} (x^2 + y^2)^{\frac{n-s}{2}} \quad (3.19)$$

and

$$D_{n,|m|,s} = \frac{(-1)^s \times (2n+1-s)!}{s! \times (n-|m|-s)! \times (n-|m|-s+1)!} \quad (3.20)$$

If we assume

$$CM_{pq} = \frac{\mu_{pq}}{M_{00}^{(p+q+2)/2}} \quad (3.21)$$

and

$$RM_{pq} = \frac{\sum_x \sum_y f(x, y) \hat{x}^p \hat{y}^q \sqrt{\hat{x}^2 + \hat{y}^2}}{M_{00}^{(p+q+2)/2}} \quad (3.22)$$

where $\hat{x} = x - x_0$, $\hat{y} = y - y_0$, and by assuming $k = (n - s - m)/2$, $d = (n - s - m + 1)/2$, then we can compute pseudo Zernike moment invariants (PZMI) as:

$$\begin{aligned} PZMI_{nm} = & \\ & \frac{n+1}{\pi} \sum_{\substack{s=0 \\ (n-m-s)\text{even}}}^{n-|m|} D_{n,|m|,s} \\ & \times \sum_{a=0}^k \sum_{b=0}^m \binom{k}{a} \binom{m}{b} (-j)^b CM_{2k+m-2a-b, 2a+b} \\ & + \frac{n+1}{\pi} \sum_{\substack{s=0 \\ (n-m-s)\text{odd}}}^{n-|m|} D_{n,|m|,s} \\ & \times \sum_{a=0}^d \sum_{b=0}^m \binom{d}{a} \binom{m}{b} (-j)^b RM_{2d+m-2a-b, 2a+b} \end{aligned} \quad (3.23)$$

3.1.6 Teague-Zernike Moment Invariants

Teague-Zernike Moment Invariants (TZMI) [31] are a set of invariants based on the Zernike moment invariants. TZMIs can be written as:

$$\begin{aligned} (TZMI)_{n0} &= A_{n0} \\ (TZMI)_{nl} &= |A_{nl}|^2 \\ (TZMI)_{nz} &= \left[A_{nl}^* (A_{mh})^p \right] + \left[A_{nl}^* (A_{mh})^p \right]^* \end{aligned} \quad (3.24)$$

where

$$\begin{aligned} h &\leq l, p = l/h, p \geq 1 \\ l \bmod h &= 0, z = p + l + h \end{aligned} \quad (3.25)$$

3.1.7 Normalized Zernike and Pseudo Zernike Moment Invariants

Normalizing Zernike and pseudo Zernike moment invariants will reduce the dynamic range of the invariants. NZMIs and NPZMIs can be defined as:

$$\begin{aligned} NZMI_{nl} &= \frac{A_{nl}}{A_{n-2,l}} & l < n, A_{n-2,l} \neq 0 \\ NZMI_{nl} &= A_{nl} & l = n, \text{ or } A_{n-2,l} = 0 \end{aligned} \quad (3.26)$$

and

$$\begin{aligned} NPZMI_{nl} &= \frac{PZMI_{nl}}{PZMI_{n-1,l}} & l < n, A_{n-2,l} \neq 0 \\ NPZMI_{nl} &= PZMI_{nl} & l = n, \text{ or } A_{n-2,l} = 0 \end{aligned} \quad (3.27)$$

3.1.8 Legendre Moment Invariants

The Legendre moments [32] of order $(n+m)$ with image intensity function $f(x,y)$ are defined as

$$L_{nm} = \lambda_{nm} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P_n(x_i) P_m(y_j) f(i, j) \quad (3.28)$$

where Legendre polynomial of order n , $P_n(x)$ is given by:

$$P_n(x) = \sum_{k=0}^n \left[(-1)^{\frac{n-k}{2}} \times \frac{1}{2^n} \frac{(n+k)! x^k}{\left(\frac{n-k}{2}\right)! \cdot \left(\frac{n+k}{2}\right)! \cdot k!} \right]_{(n-k)\text{even}} \quad (3.29)$$

and the normalizing constant is as follows:

$$\lambda_{nm} = \frac{(2n+1)(2m+1)}{N \times M} \quad (3.30)$$

where N and M are width and height of the image. x_i and y_j are the normalized pixel coordinates in the range of $[-1,1]$:

$$x_i = \frac{2i}{N-1} - 1, y_j = \frac{2j}{M-1} - 1 \quad (3.30)$$

3.2 Locating the Face

In order to locate the face in the image, we use second order moment invariants. We can locate the face with an ellipse around the face with the center of (x_0, y_0) which is the centroid, and the major and minor axes of 2α and 2β [55]. The orientation of the face according to the major axis would be θ where x_0 , y_0 and θ are calculated by equations (3.2) and (3.5). α and β are defined as:

$$\alpha = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{I_{\max}^3}{I_{\min}}\right)^{\frac{1}{8}}, \beta = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{I_{\min}^3}{I_{\max}}\right)^{\frac{1}{8}} \quad (3.32)$$

where

$$\begin{aligned} I_{\min} &= \sum_x \sum_y [(x-x_0)\cos\theta - (y-y_0)\sin\theta]^2 \\ I_{\max} &= \sum_x \sum_y [(x-x_0)\sin\theta - (y-y_0)\cos\theta]^2 \end{aligned} \quad (3.33)$$

We can eliminate the background by superimposing this ellipse on the image and setting pixels outside the ellipse to zero. It would be better if we modify the ellipse in a way that it can also omit some of the facial features like hair or beard that can change with time. We use the parameter ρ to adjust the ellipse as:

$$A = \rho\alpha, \quad B = \rho\beta \quad (3.34)$$

where $2A$ and $2B$ are the axes of the new ellipse and ρ varies between 0 and 1. The new ellipse will have the same orientation as the best fit ellipse calculated using equations (3.32) and (3.33) but with shorter axes. This will eliminate the unnecessary information from the face image. Our experiments show that the optimum value for ρ is 0.7 for the AT & T database.

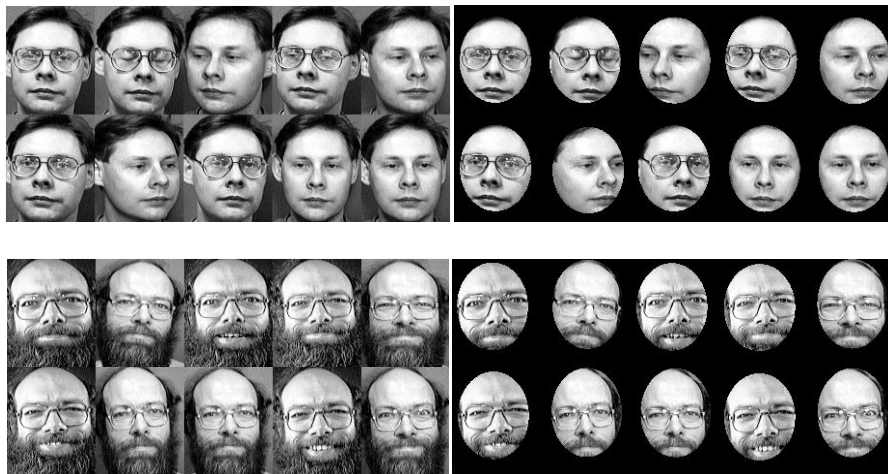


Figure 3.1: Two of the samples in the database and the result of applying face localization on them.

3.3 Experimental Results

3.3.1 Comparing Different Moment Invariants

Extraction of applicable features from the human face images is an important part of the recognition. Therefore choosing the proper feature extractor is very crucial when designing a face recognition system with high recognition rate. In this section different moment invariants have been used as feature extractor and the performance of these methods has been compared. The AT&T database of images is used which contains 400 images of 40 different people.

3.3.1.1 Making the Feature Vectors

The moment invariants described in section 3.1 have been used for extraction of the pertinent features in the human face image. A feature vector will be formed by calculating different orders of each moment invariant. For example for Bamieh moment invariants we have:

$$FV_{BMI} = [BMI_1, BMI_2, BMI_3, BMI_4] \quad (3.35)$$

or in general we have

$$FV_{MI,n} = [MI_1, MI_2, MI_3, \dots, MI_n] \quad (3.36)$$

where MI_j s are moments of j^{th} order. For most of the moment invariants studied in this experiment there are usually more than one moment for each fixed order. Therefore the size of feature vector would be the number of all these possible moments. Calculating this vector for all the images in our training database will result in a $M \times n$ feature space matrix where M is the number of images in the training database and n is the number of the moments that we use for each type of moment invariants.

3.3.1.2 Classification

Having the feature vectors for all the samples in the train and test sets, the next step would be to design a classifier. Artificial neural networks (ANN) have been widely used as the classifier in many face recognition systems [56]. We used a three layer perceptron neural network. The number of neurons in input layer is equal to the size of related feature vector for each experiment and the number of output neurons is equal to the number of classes which is 40. Different number of neurons for hidden layer have been tested and the one producing best results has been reported.

The whole database was divided into two training and test sets by choosing 5 random images of each person for the training set and the rest for the test set. This makes the training and test sets with 200 images in each. There is no overlap between train and test sets. Histogram equalization was applied over all the images. In real applications where there is a camera and a recognition system for example in an automated banking machine, pre-processing techniques like noise removal and histogram normalization are important in order to get good feature vectors and also good classification results. For this database we did not need any noise removal due to the good quality of images.

Face location was determined using equations (3.32) and (3.33). Irrelevant parts like background, hair and shoulders were omitted using equation (3.34). We tested different values of ρ and $\rho = 0.7$ was used the in this study for best results.

In this experiment moment invariants were calculated for different orders. We used Regular moment invariants (RMI), Hu moment invariants (HMI), Bamieh moment invariants (BMI), Zernike moment invariants (ZMI), pseudo Zernike moment invariants (PZMI), normalized Zernike moment invariants (NZMI) and normalized pseudo Zernike moment invariants (NPZMI). The moment invariants were calculated for all the images in the test and train set and were also mean normalized in order to have the mean of zero and variance of one as following:

$$WFV_{MI,n} = \frac{FV_{MI,n} - \text{mean}(FV_{MI,n})}{\text{var}(FV_{MI,n})} \quad (3.37)$$

where $\text{mean}(FV_{MI,n})$ and $\text{var}(FV_{MI,n})$ are the mean and variance of the moment invariant values for all the 400 images in test and train sets. The new moment invariants based on

the above modification have been formed in two feature space matrixes and then were classified using neural network classifier.

3.3.1.3 Classification Accuracy

The number of neurons in the input layer of the perceptron neural network is equal to the size of the feature vector. For different moment invariants we have different number of features, therefore the size of the neural network is dependent on the moment being used. ZMI, PZMI, NZMI and NPZMI are complex valued moments. In order to make the feature vector, the real and imaginary part of these invariants were both used as features which made the size of feature vectors almost twice larger. For some orders of the moment invariants the imaginary part was found to be equal to zero for the whole matrix. These imaginary components of the moments did not carry any information and were omitted from the vector. The number of neurons in the hidden layer is also changed in the experiment regarding the type of the moment invariant in use. The results of the recognition are shown in Table 3.2. R is the ratio of the number of neurons in hidden layer to the number of neurons in input layer and varies between 0.9 to 3 for different methods:

$$R = \frac{\text{number of elements in hidden layer}}{\text{number of input features}} \quad (3.38)$$

Results in Table 3.2 indicate, pseudo Zernike, normalized Zernike and normalized pseudo Zernike have yielded the best recognition accuracy. NZMI method has very good results despite its fewer features. It can also be seen that for optimum results PZMI and

NPZMI methods require smaller R and will need less training epochs while other methods require more epochs to reach similar classification accuracy.

Table 3.2: Classification parameters for different moment invariants for AT&T database. The results reported for each method are the best results for different orders.

Method	Classification accuracy	Number of features	order	Number of neurons in input layer	Number of neurons in hidden layer	R	Number of epochs
PZMI	94 %	120	10	120	108	0.9	1000
NZMI	91 %	34	7	34	68	2.0	6000
NPZMI (10)	90.5 %	118	10	118	106	0.9	1000
ZMI	87 %	76	11	76	152	2.0	4000
NPZMI (7)	86 %	61	7	61	122	2.0	2000
RMI	85.5 %	75	11	75	150	2.0	2000
TZMI	67.5 %	32	7	32	96	3.0	10000
HMI	61 %	32	7	32	96	3.0	6000
BMI	22 %	4	4	4	12	3.0	4000

3.3.1.4 Recognition Speed

Recognition time for different methods has been measured and shown in Table 3.3. Although the BMI is the fastest method, its poor recognition makes it improper for face recognition applications. RMI method has a tolerable computation cost and recognition time of 4.53 seconds which looks good when compared to other methods with similar recognition accuracy. ZMI method is the most time consuming one and takes almost 10 times the PZMI method. Since the TZMI, NZMI and NPZMIs are calculated from ZMI and PZMI features, their computation costs were not measured here. All the algorithms were implemented using Matlab on a PC with a 2.1 GHz Pentium Duo Core Processor.

Table 3.3: Recognition time for different methods

Moment Invariant	Recognition time (s)	order	Number of Features
BMI	0.15	4	4
HMI	0.62	7	32
RMI	4.53	7	33
PZMI	10.75	7	64 (36 Re+28 Im)
ZMI	124	7	33 (18 Re+15 Im)

3.3.1.5 The effect of order and number of features

As mentioned before, experimental results show that pseudo Zernike moment invariants lead to the best classification accuracy. The effect of using different orders of the moment invariant on the classification results has been studied. The classification part was repeated with different feature vectors containing the moment invariants for only specific orders and the results are reported in Table 3.4. The number of training epochs for the neural network classifier differs in each row.

Table 3.4: Recognition Accuracy for PZMI method using different orders of the moment invariants

Orders	Number of features	Recognition accuracy for different number of epochs				
		1000	2000	3000	4000	5000
1-5	21	81.5	84.5	89	91	93.5
1-6	28	90.5	91.5	92.5	93	94
1-7	36	94	93	92.5	94.5	95
1-8	45	91	92.5	92.5	93	91.5
1-9	55	93	90	91.5	92.5	92.5
1-10	65	91.5	93	92	92.5	92
4-10	56	90.5	93.5	93	91	91.5
5-10	51	91.5	92	93	92.5	92.5
6-10	45	87.5	90	93	93	91.5
7-10	38	85.5	89	87.5	87.5	86
8-10	30	78.5	86.5	86.5	87.5	87
9-10	21	59.5	72.5	74	77	82

The performance of the system can be improved by using proper value for ρ to remove background and hair and also by choosing proper orders of the moment invariants. High order PZMI's contain useful information, therefore lead to the best results. BMI's were the fastest to be computed but had very poor recognition accuracy for face images while calculating ZMI's was the most time consuming. The best recognition rate of 95% was achieved with PZMI of order 1 to 7 with 5000 epochs training.

3.3.2 Performance of Different Classifiers

As the results of last set of experiments show, pseudo Zernike moment invariants had the best results among other moment invariants so they are chosen as the feature extractor for this experiment. In this section pseudo Zernike moment invariant is applied on the FERET database of faces which includes a variety of images with different illuminations and facial expression. Performance of k NN, SVM, and HMM classifiers for this application is studied.

The FERET database of images has been used in this study which consists of 14051 eight-bit grayscale images of human heads with views ranging from frontal to left and right profiles. The size of the pictures is 256x384 pixels. The pictures are from 1209 people, taken at different time, illuminations, and facial expressions.

3.3.2.1 Face detection & Feature Extraction

First step in our face recognition process is to locate the face in input images. For this reason Viola-Jones algorithm [12] has been used which minimizes computation time while achieving high detection accuracy. This approach has been used to construct a face

detection system which is approximately 15 times faster than other face detection approaches based on similarity functions.

Pseudo Zernike moment invariants were used in this study as feature extractor. Just like last experiment, a feature vector will be formed by calculating different orders of PZMI as we have

$$FV_{MI,n} = [MI_1, MI_2, MI_3, \dots, MI_n] \quad (3.38)$$

where MI_j s are PZMIs of j^{th} order. The size of feature vector, n , is the number of all the possible moments that we are going to use in the experiment.

3.3.2.2 Recognition

The next step is to design a classifier to classify the feature vectors previously produced. Three different classifiers have been used to do this job: k -NN, SVM and HMM.

A. k Nearest Neighbors

The k nearest neighbor algorithm is one of the easiest machine learning algorithms. A new object is classified and assigned to a class by a majority vote of its neighbors. The class that the object is assigned to is the most common class amongst its k nearest neighbors. k is a typically small positive integer. If $k = 1$, then the object is simply assigned to the class of its nearest neighbor.

B. Support Vector Machines

Support vector machines simultaneously minimize the classification error and maximize the geometric margin; and are also known as maximum margin classifiers [46].

An SVM will make a separating hyperplane in the feature space which maximizes the margin between the data sets. Two parallel hyperplanes are constructed to calculate the margin, one on each side of the separating one. These hyperplanes are pushed up against the two data sets, so a good separation is achieved by the hyperplane that has the largest distance to the neighboring support vectors of both classes. The larger the margin or distance between these parallel hyperplanes is, there would be more hope that unknown image samples will be classified correctly [47].

C. HMM

The hidden Markov model as a double stochastic process can efficiently model the generation of sequential data. HMM has successfully been used in many pattern recognition applications. There are basically two different approaches in using HMM as a classifier. In the first approach, a single HMM models all classes and based on the different paths, classes are distinguished. In the second one, a separate HMM is used for each class. Since we have a limited numbers of classes, we used the second one which is known as model discriminant approach [57]. Therefore, for each class (ω_c) is modeled by a single HMM (λ_c). Each face is represented as a sequence of Pseudo-Zernike moments. At first, we quantized the sequence with a codebook. But the result was very poor. So we continued with the continuous HMM. In the following, the notations for manipulating a continuous HMM are briefly described [58].

- The observation set $O = \{x_1, x_2, \dots, x_T\}$ where T is length of sequence and equals to the number of pseudo Zernike features in this work.

- The set of states in the model $S = \{S_1, S_2, \dots, S_N\}$ where N is the number of states.

We tried different number of states to find the best result.

- The matrix A of state transition probabilities

$$A = \{a_{ij} | a_{ij} = P(S_t = j | S_{t-1} = i)\} \quad (3.39)$$

- The initial state distribution $\Pi = \{\pi_i | \pi_i = P(S_1 = i)\}$.
- The observation probability distribution

$$b_j(x_t) = p(x_t | S_t = j) = \sum_{m=1}^M c_{jm} \mathcal{N}(x_t, \mu_{jm}, \Sigma_{jm}). \quad (3.40)$$

where $\sum_{m=1}^M c_{jm} = 1$.

We trained the HMMs with Baum-Welch algorithm which is a standard algorithm. With C different HMMS, we classified the samples in test set to classes that the probability $P(O|\lambda_c)$ was higher.

3.3.2.3 Results and Comparison

The face detection and recognition methods described above were applied on FERET database of images. In this experiment, only frontal images (FA & FB) and only classes with more than 8 images were used. Half of the images in each class were used as training set and the rest were assigned to the test set. There is no overlap between training and test sets. Vila-Jones [12] face locator was applied on these images and histogram equalization was applied on the results. The outputs of the face locator were all resized to 100×100 pixels to reduce computation.

In the next step Pseudo Zernike moment invariant feature extractor was applied and 63 moments for each image were extracted and also normalized in order to have the mean of zero and variance of one.

These features were presented to the described classifiers: k -NN, SVM and HMM. The results of the recognition are shown in Table 3.5. In k NN method, two different numbers of neighborhoods were tried. By increasing the number of the neighborhood the classification rate decreases. SVM classifier was used with different kernel functions. Linear kernel has the poorest result and RBF gave the highest accuracy.

Continuous HMM also was used. Four mixture density functions were used to make the observation probability distribution [59]. Different numbers of states were tried and 35 states gave us the best answer.

Table 3.5: Classification results for different classifiers.

Classifier	Acc. Rate -Train	Acc. Rate -Test
kNN-k=1	-	0.89
kNN-k=3	-	0.77
SVM-RBF	0.97	0.91
SVM-Polynomial	0.97	0.87
SVM-Linear	1.00	0.81
HMM, N=25	1.00	0.88
HMM, N=35	0.97	0.90
HMM, N=45	0.93	0.80

As it can be seen in Table 3.5, the best recognition rate of 91% was achieved by SVM with RBF kernel.

3.3.3 Effect of Misalignment

Before starting to extract features from the input image, we should find the face in the image. The face is cropped using a face locator and the output is sent to the feature extractor. This output however is not always cropped with the same size or perfect orientation for even the best face detectors. The output of the face detector might make angles with the ideal orientation of the face or might be shifted from the center. Faces of one person, detected in different images might also have different sizes in the cropped image. In this research we use the word “misalignment” for all of these distortions. This leads to the images of the faces not being perfectly aligned with each other, meaning the eyes and noses of different images are not in the same position. This misalignment might affect the methods that use averaging and normalizing as part of the recognition process. When calculating the mean of the images in PCA method for example, this will cause problem. A comparison between the performances of some face recognition systems in presence of image misalignments is presented in this section. The aim of this experiment is to investigate the effect of shifting or rotation misalignment in a practical face recognition system on the recognition rates for some moment invariant based methods and then to compare them to other well known methods. In this test, the performance of pseudo Zernike and Legendre moment invariants is compared to two best known face recognition systems, PCA and FLD. FERET and extended Yale B databases are used and recognition results are presented for different transformations.

PCA is used for reducing the dimensionality of the feature space by transforming a number of possibly correlated variables into a smaller number of uncorrelated variables

called principal components. It is based on the assumption that most of the information about classes is in the directions along which the variations or scatterings are maximum. For a given database of images with N images, the first step is to find the mean of the images in the database and subtract this mean image from all the images. Then the covariance matrix and its eigenvalues and eigenvectors are calculated. Input images are projected on the direction of eigenfaces and then classified based on minimum distance or support vector machines.

In PCA we emphasize on directions related to large eigenvalues or in other words large variation. In some cases, some of the important information might be in directions with small variation. What Fisher linear discriminant (FLD) does is finding a direction which maximizes the scatter between classes over the scatter within classes.

Flusser proved in [60] that Hu moment invariants are incomplete and are not rotation invariant. He also has shown that Wong [61] and Li moments [62] are not complete either and need phase cancellation to be rotation invariant. He also mentions that Zernike moments presented by Teague [31] and Wallin [63] are not invariant. The magnitude of the complex Zernike moments which is known to be rotation invariant is actually varying with in-plane rotation. This problem appears when using the quantized images and replacing integration in the moments formula with a simple summation over the digital image where the values of x and y are quantized.

The four methods of PCA, FLD, PZMI and LMI are applied on FERET and the extended Yale B database of face images. The Extended Yale B database contains 21888 single light source images of 38 subjects each seen under 576 viewing conditions: 9 poses

and 64 illumination conditions. The database is divided into five different subsets according to the angle between the light source direction and the camera axis. For the FERET database the subjects that had more than seven frontal images in the dataset are selected. Half of the images in each class are used as training set and the rest are assigned to test set. There is no overlap between training and test sets. Viola-Jones [12] face detector (VJ) is applied on these images and histogram equalization is applied on the results. Then the faces were normalized by the position of the eyes and mouth. The outputs of the face detector which all are squares but in different sizes are then resized to 50×50 pixels. This reduces the size of the images and as a result will reduce the computation cost. The output of VJ face detector had some wrong detections that are omitted manually from the dataset. The reason to do this is that we did not want the error of the VJ detector to be added to the error generated by our own rotation and translation experiments and we wanted these experiments to be independent. Some of these images with detection problems can be seen in Figure 3.2. As it can be seen in this figure, some of the detected faces include only a part of the face and some show the whole head with hair and neck. One of the subjects in the FERET database and the detected images of this



Figure 3.2: Some wrong detections using Viola-Jones face detector on FERET database.



Figure 3.3: (a) Some samples from FERET database. (b) The result of the Viola-Jones face detection. (c) Images in (b) after histogram equalization.

subject using VJ face detector are presented in Figure 3.3.

To test the effect of input image rotation on recognition rate of the face recognition systems, a rotation of θ degrees is applied to the images in the test set, where θ varies from 0 to 21 degrees with the steps of 3 degrees. This transformation is applied on half of the images in each class (test set) and the recognition rate is obtained for each value of θ . This task has been repeated for all four recognition algorithms of PCA, FLD, ZMI and LMI and results are presented in Figure 3.8.a. In another test a random rotation with $|\theta| < \theta_{max}$ is applied where θ_{max} has the values of 0 to 21 degrees with the steps of 3 degrees, as in practice we can come up with a range or boundary for the rotation of the camera and the value of the rotation will be a random number. The results of the second test are presented in Figure 3.8.b. as it can be seen in both Figures 3.8.a and 3.8.b, moment invariant based methods, ZMI and LMI, yield better results dealing with rotation rather than statistical methods. ZMI has best results for both tests. As can be seen the

recognition rates for the second test with a random amount of rotation are a little bit higher than the first one where all the images have been rotated with the maximum value.

Another test is investigating the effect of a translation or shift in the detected image (output of VJ detector) on recognition rate. Independent translations in x and y directions are applied on the images from 0 to 9 pixel shift in each direction resulting in 100 different situations per test image. In the worst case, i.e. 9 pixels shift in both directions, the result image has about 64% of the information of the original cropped face in the training set. Just like the rotation test, these translations are applied on half of the data base while the algorithms are trained with the other half without any translation. Recognition results for the four methods are presented in Figure 3.4. For each graph, x and y axes show the shift in x and y directions which varies from 0 to 9 pixels and z axis is the recognition rate. As it can be seen, recognition rate for PCA and FLD drops abruptly after 3 to 5 pixels of shift but for ZMI and LMI the decrease in the recognition rate happens very smoothly.

For PCA as well as the Fisherfaces method, the number of eigenfaces is equal to the number of real positive eigenvalues of the covariance matrix. For pseudo Zernike method we have calculated the moments to the order of 10 with all the repetitions resulting in a total of 120 moments per image. The pseudo Zernike moments are complex numbers and their absolute values are used as the features to have the maximum invariance due to rotation. The moment values are whitened in order to have the mean of zero and variance of one. Finally for Legendre moments, we have calculated the moments to the order of 20, and formed them as the feature vector for classification.

The same process is applied on the images from the Extended Yale B database. For this database the cropped version of the images is provided with the database. These images have been manually cropped and the positions of eyes and lips have been perfectly aligned, so we did not use VJ face detector for this part. Each subject in this database has been pictured under 64 different illuminations and 9 poses. We used only images with frontal position just like the images in the FERET database. The images in the database are grouped into five subsets regarding their illumination condition. For a given image if the angle between the axis of the camera and the light source, θ_c is less than 12 degrees it would be classified in subset 1. If $20^\circ < \theta_c < 25^\circ$ it would be in subset 2, and for $35^\circ < \theta_c < 50^\circ$, $60^\circ < \theta_c < 77^\circ$, and finally $\theta_c > 78^\circ$ the image would be put in subset 3, subset 4 and subset 5. We used the first subset as training set and all five subsets as the test images in separate experiments. For each subset rotation and, random rotation applied and the recognition rates are calculated. Figures 3.6, 3.7, 3.8 and 3.9 show the results for rotation and random rotation for different subsets respectively. We keep the first subset in the test set as well to figure out how well the four methods of PCA, FLD, ZMI and LMI will recognize the same images with no illumination change after misalignments. For the translation test, the recognition algorithms are trained using subset 1 and tested with subset 2 after translation. Recognition results for this test are presented in Figure 3.5. Since we did not apply any illumination invariant method on the images, the results for subsets 4 and 5 are very poor due to bad illumination.

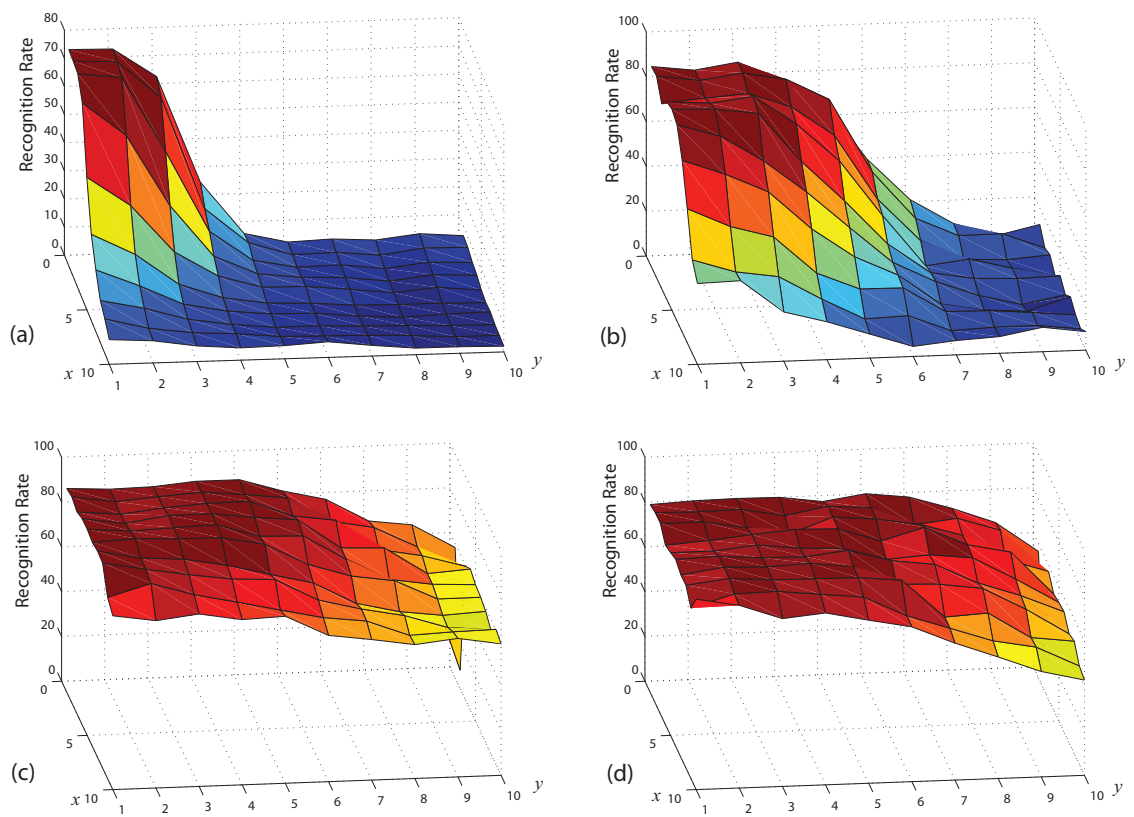


Figure 3.4: Effect of translation on images in the FERET database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with half of the database without any translations and tested with the other half with the translation.

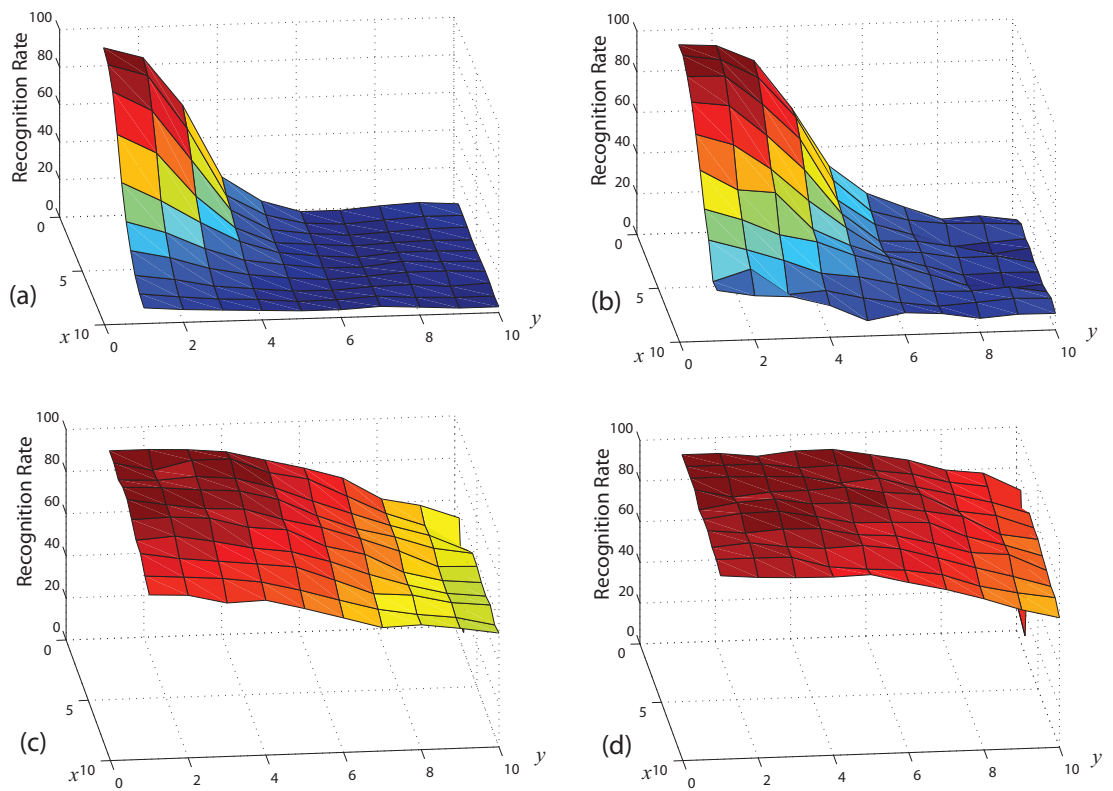


Figure 3.5: Effect of translation on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the second subset with shifting.

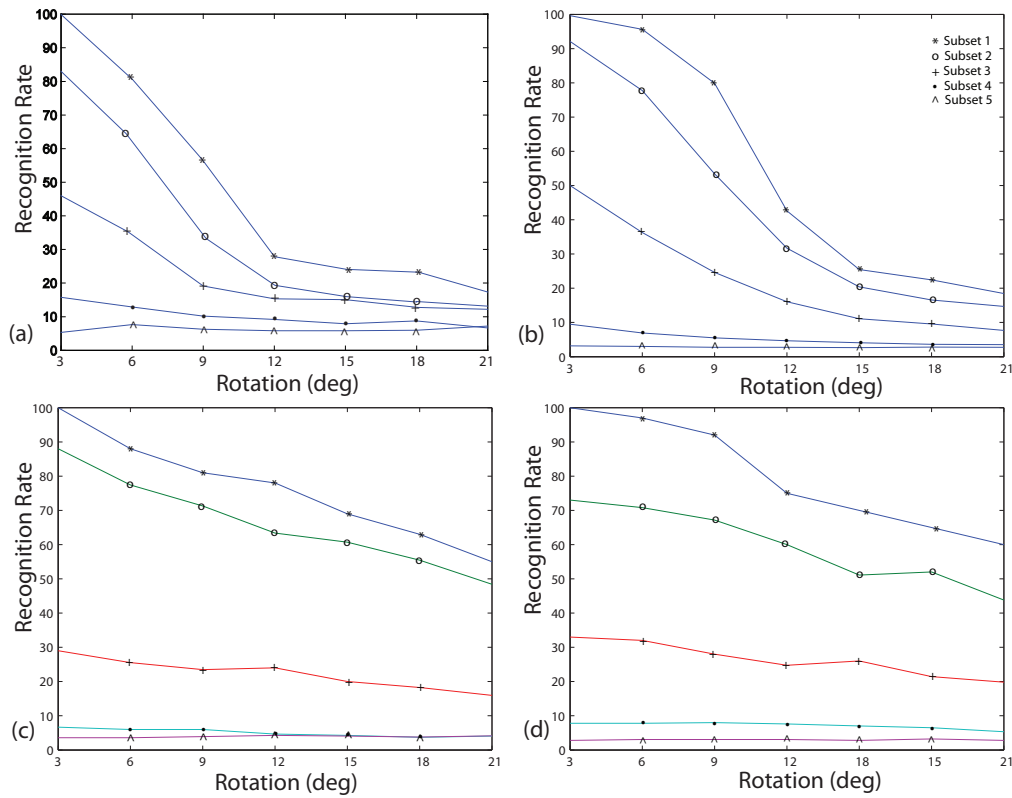


Figure 3.6: Effect of rotation θ degrees on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the translated subsets. Since we did not apply any illumination invariant method the results for subsets 4 and 5 are very poor.

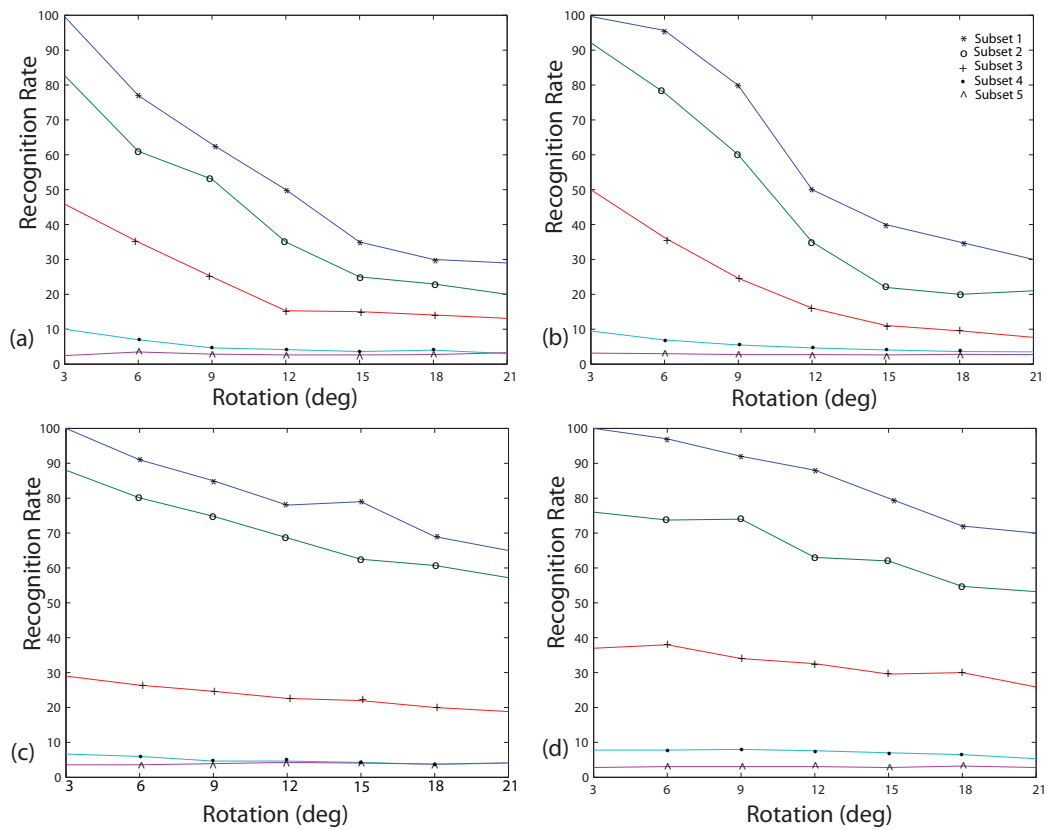


Figure 3.7: Effect of rotation with the maximum of θ degrees on images in the extended Yale B database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with the first subset without translation and tested with the translated subsets.

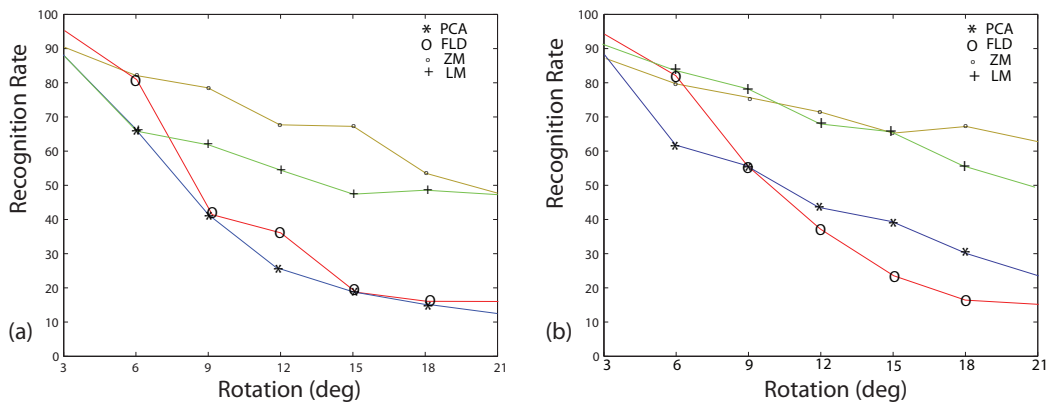


Figure 3.8: Result of the rotation test for the FERET database using 4 methods (a) effect of rotation (b) effect of random rotation with maximum of θ degrees.

3.3.4 Robustness to Noise

Image noise also might be another factor affecting the recognition process. In a face recognition system, if the input image is noisy due to any reason, there is a loss of information which affects correct feature extraction. Impulsive noise, also called salt and pepper noise, can be caused by dead pixels, analog to digital converter errors or bit errors in transmission [64]. Gaussian noise, however, is mostly the result of signal amplification in the image sensor. The standard model of amplifier noise is additive, Gaussian noise, which is independent at each pixel and independent of the signal intensity. Amplifier noise is primarily caused by Johnson–Nyquist noise (thermal noise), including the noise that comes from the reset noise of capacitors (kTC noise) [65]. Amplifier noise is a major part of the read noise of an image sensor, *i.e.*, of the constant noise level in dark areas of the image [66]. Effects of these two main sources of image noise on face recognition results are tested in this section. Gaussian noise and salt and pepper noise are added to the images from low levels to high levels and the robustness of different methods with respect to increasing noise level is tested.

Just like the previous experiment pseudo Zernike and Legendre moment invariants are compared with principal component analysis and Fisher discriminant methods. The FERET database of faces and the extended Yale B database are used in this experiment. Faces are cropped and normalized just like the last experiment.

A Gaussian noise with zero mean and the variance of σ is added to the test samples, where σ varies between 0 and 0.5. Salt and pepper noise also is added to the test images in another test with the noise density varying from 0 to 1. These noisy images are fed to

the PZMI, LMI, PCA and FLD systems which are trained with images without noise. Recognition results for Yale B database are presented in Figures 3.9 and 3.10 for Gaussian noise and Salt and Pepper noise respectively. The results for the FERET database are presented in Figure 3.11. As the results show, statistical methods, PCA and FLD have better recognition rates rather than moment invariant based methods. Zernike method is the one which is affected by the noise level more than others while PCA yields best results and seems to have the best noise stability.

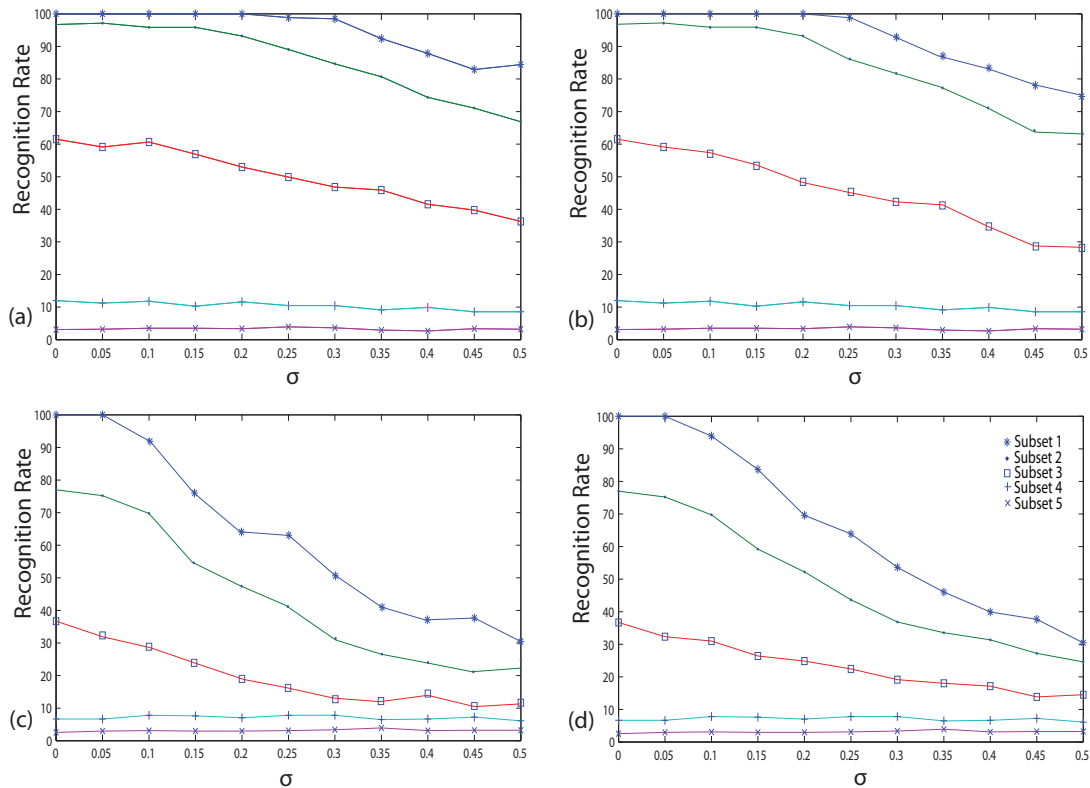


Figure 3.9: Effect of Gaussian Noise on images in the Yale B database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with the first subset without Noise and tested with all the subsets after adding noise.

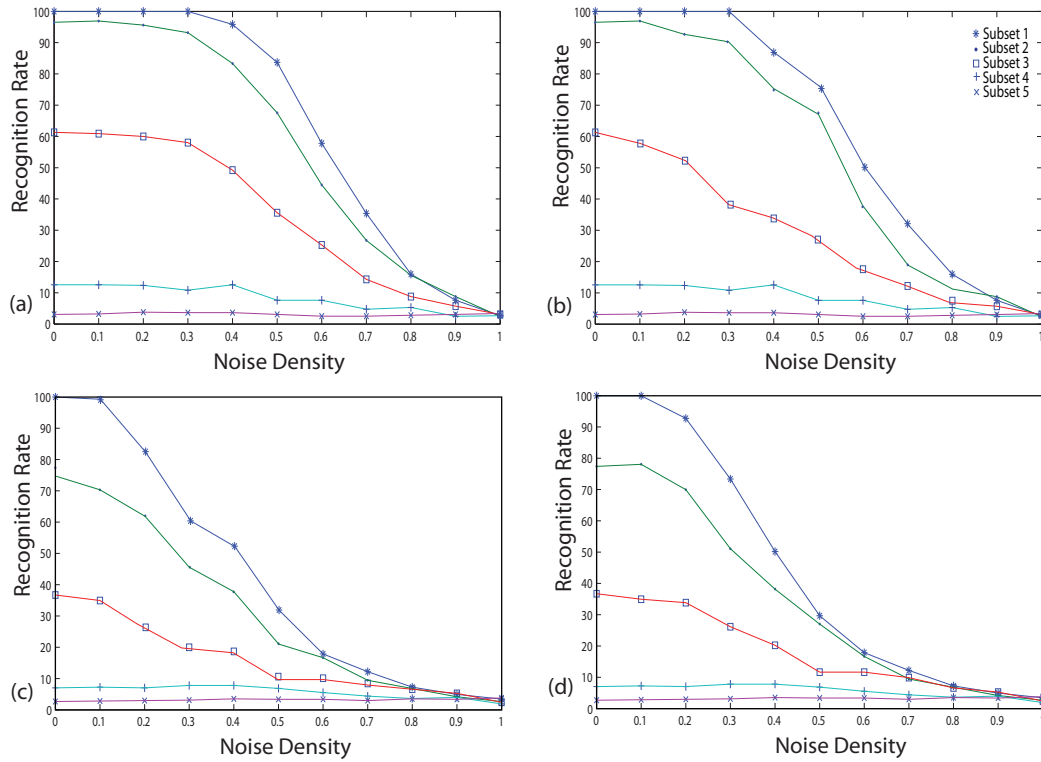


Figure 3.10: Effect of Salt & Pepper Noise on images in the Yale B database using 4 different methods: (a) PCA (b) FLD (c) PZMI (d) LMI. The algorithms were trained with the first subset without Noise and tested with all the subsets after adding noise.

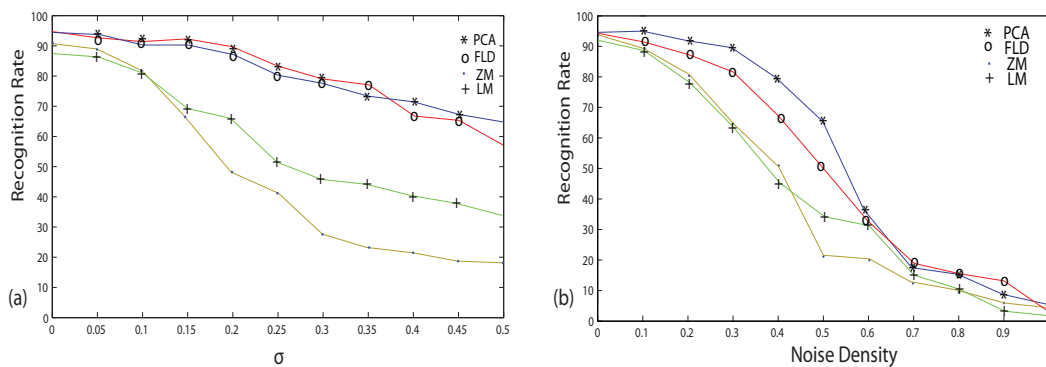


Figure 3.11: Result of the noise test for the FERET database using 4 methods (a) effect of Gaussian noise (b) effect of Salt and Pepper noise.

3.4 Conclusions

Performance of different moment invariants for feature extraction in face recognition was studied in this chapter. The AT&T database, FERET and Extended Yale B were used in different experiments performed in this study. The moment invariants used in this study were HMI, BMI, RMI, ZMI, TZMI, PZMI, NZMI, NPZMI and LMI. A three layer perceptron neural network was used as the classifier in the first experiment and the best recognition rate of 95% was achieved with PZMI of order 1 to 7 with 5000 epochs training. The performance of this system can be optimized by using proper value for ρ and also by choosing proper orders of the moment invariants. High order PZMIs contain useful information, therefore lead to the best results. BMIs were the fastest to be computed but had very poor recognition accuracy for face images while calculating ZMIs was the most time consuming.

In another experiment performance of pseudo Zernike moment invariants for feature extraction along with different classifiers was studied in this chapter. The FERET database of images was used to perform this study. Viola and Jones face detector was used for locating the face in the images. k -NN, SVM and HMM algorithms were used for the classification step. They had the best results of 89%, 91% and 90% respectively. The best recognition rate of 91% was achieved by SVM with RBF kernel.

Misalignment and noise might exist in the output of the face detector in a face recognition system. These factors affect the performance of the system. Performance of different face recognition methods under different conditions was studied in another experiment. Performance of eigenfaces, Fisherfaces, pseudo Zernike moment invariant

and Legendre moment invariants were tested on images with rotation, translation, Gaussian noise and salt and pepper noise. FERET and Extended Yale B databases were used as face image sources. Viola and Jones face detector, was used for locating the face in the FERET database of images. For both of the databases, statistical methods, *i.e.* PCA and FLD had better results in presence of noise while moment invariant based methods performed better where there were geometrical misalignments. PCA had the best results among the four methods for noise stability, but was less steady when images were rotated or shifted. PCA had the recognition rate of 65% and 68% in presence of the Gaussian noise with $\sigma = 0.5$ and the salt and pepper noise with noise density of 0.5, respectively. Zernike moments showed to have the best performance for the shift test as well as the rotation test with the recognition rate of 61% for FERET database and 67% for the Yale B database for the maximum shift of 9 pixels in both x and y directions. The recognition rates for PCA and FLD dropped quickly after shifting 3 or 4 pixels.

Chapter 4

Illumination Invariant Feature Extraction

Face recognition has been one of the most popular research topics in computer science and information technology since 1990. During the last two decades several face recognition algorithms and systems have been proposed and some major advances have been achieved. As a result, the performance of face recognition systems under controlled conditions has now reached a satisfactory level. These systems, however, face some challenges in environments with variations in illumination, pose, expression, *etc.* Illumination along with pose is the most significant factor that affects face recognition results. Whether the face is indoor or outdoor or under certain illumination condition, it is observed that changes in lighting condition shape different shading and shadows on the face because of the 3D shape of human faces. This may weaken some face features or

may result in too bright or too dark parts in images. These variations can be larger than the variations due to personal identity and hence yield a lower recognition rate (Figure 4.1).

An efficient method for face recognition which is robust under illumination variations is proposed in this chapter. The proposed method achieves the illumination invariants based on the illumination-reflection model. Different filters have been tested to achieve the reflectance part of the image which is illumination invariant and maximum filter is suggested as the best method for this purpose. The proposed method does not need any prior information about the face shape or illumination and can be applied on each image separately. Unlike most available methods, our method does not need multiple images in training stage to get the illumination invariants. Support vector machines and k -nearest neighbors methods are used as classifier. Several experiments are performed on Yale B, extended Yale B and CMU-PIE databases. Recognition results show that the proposed method is suitable for efficient face recognition under illumination variations.

4.1 Illumination Invariant Methods

The available methods for illumination invariant face recognition can be categorized into two main groups, namely passive and active methods [67]. Passive approaches try to solve the problem using the common visible spectrum images. Active approaches, however, try to achieve illumination invariant face images by using 3D cameras, thermal infrared imaging, near infrared illumination, differential imaging, *etc.* Passive methods are more common and can generally be grouped in three main categories, *i.e.*, face

illumination modeling, illumination invariant feature extraction, and face image normalization [44].

Face illumination modeling methods use statistical or physical information models. Statistical techniques like principal component analysis (PCA) [13] and linear discriminant analysis (LDA) [14] need no surface property information and just take the input images to a low dimensional linear subspace. Hallinan showed in [68] that face images with illumination variance can be represented with only five eigenfaces. Belhumeur's 3D linear subspace method [69] used three or more images of the same face which are taken under different illumination and makes a 3D basis for the linear subspace. Belhumeur and Kriegman proved in [70] that images of an object with Lambertian surface under different illumination but the same pose, form an illumination convex cone. Basri and Jacob [71] suggested a method to model the intensity of object surface under illumination with a 9 dimensional linear subspace based on spherical harmonics. Lee *et al.* [72] showed that a set of nine directions can be found in order to

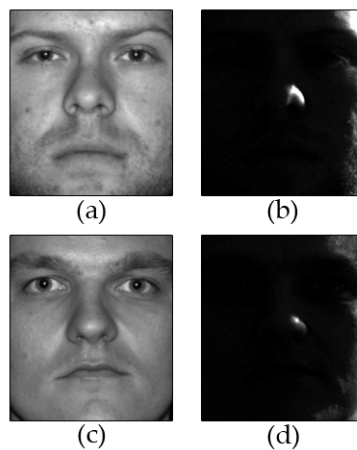


Figure 4.1: Bad illumination can cause between class similarities overcome the inter-class similarities. Images in each row belong to the same subject but with different illumination.

make an effective subspace based on nine images with lighting sources in these nine directions. This subspace can be used for recognition of images under larger illumination variations.

Gao and Leung proposed line edge map in [73] to extract illumination invariant features. Zhao and Chellappa [74] used the symmetry of face and shape similarity among all faces and presented a method based on symmetric shape from shading. Sim and Kanade [75] obtained face shape from a single image and synthesized the same face under different illuminations. Shashua and Riklin-Raviv proposed quotient image method (QI) in [76] defined as the ratio between a test image and a linear combination of three images of the same face with different illuminations. Wang *et al.* suggested self-quotient image (SQI) method [77]. SQI is a multi-scale retinex approach using anisotropic smoothing with different scales.

Histogram equalization (HE) is the most common method used for image enhancement when images have illumination variations [4]. Even for images under controlled illumination, HE improves the recognition results by flattening the histogram of pixel intensities of the images. Shan *et al.* proposed Gamma intensity correction (GIC) method in [5] and used it for illumination normalization along with HE. They also presented a region-based method for equalizing histogram and gamma locally in small portions of an image. Chen *et al.* [78] used DCT to reduce the effect of illumination on recognition results. They discarded a number of DCT coefficients corresponding to low frequencies. Non-uniform illumination is removed in the image made by inverse DCT of

the remaining coefficients. A complete survey of methods used for illumination invariant face recognition can be found in [67].

As mentioned earlier, most of the methods in the literature need prior information of the face or illumination model. Some of the algorithms need three or more images of the same person with the same pose and different illuminations. These methods encounter the alignment problem that makes them more complicated. This problem has been solved in methods like SQI [77]. Although SQI method has simple structure, the empirical parameter selection for the weighted Gaussian filter used is complicated.

Another approach toward illumination variations can be grouped as local statistical methods [79-82] which add more robustness against lighting conditions or occlusion. Modular PCA (mPCA) [79] divides the face images into small patches and applies PCA on each set of patches. Modular LDA [80] uses a set of independently trained observers on different parts of faces. Each observer performs LDA independently by projecting faces to a lower dimensional subspace and performing recognition. The final result is achieved by using a simple sum rule on the recognition results.

In this chapter, we propose a method based on illumination-reflection model. The method approximates image illumination by using maximum filter for making a smooth form of the face image. Unlike other methods in the literature, it does not need any prior information about the face models or illumination of the samples and can be applied on each image separately. The maximum filter that is used is very simple and does not need any manual parameter selection or adjustment. It obtains good recognition results on the available face databases, and due to its simplicity, it is also computationally efficient. In

next chapter, we use a weighted modular voting scheme to further improve the performance of the system.

4.2 Illumination Invariants

4.2.1 The Illumination Model

Named after Johann Heinrich Lambert, a Lambertian surface is one that exhibits Lambertian reflectance which means that the surface luminance is isotropic. If the surface is illuminated, light is scattered in a way that regardless of the observer's angle of view, the perceptible brightness of the surface to an observer is the same. For example, cotton fabric or wood have Lambertian reflectance, but waxed leather or glossy coated wood do not, because based on the viewers angle, some highlight spots may appear on the surface at different locations. Face skin can be assumed a Lambertian surface provided not being wet or oily. For a Lambertian object, it's assumed that the image obtained from it is the pixel wise multiplication of the illumination and a retinex reflectance which is the property of the object itself and does not change with illumination.

Assuming faces to be Lambertian surfaces, intensity level of a given face image I under certain illumination condition can be presented as product of the reflectance and illumination [83]. That is

$$I(x, y) = L(x, y).R(x, y), \quad (4.1)$$

where $L(x,y)$ and $R(x,y)$ are the illuminance and reflectance at point (x,y) , respectively. The goal in this section is to find $R(x,y)$ which is the illumination invariant part of the

image. It is assumed that the illumination lies in the low frequency part of the image while reflectance, on the other hand, is the high frequency part of the image [44]. Physiological evidence shows that the response of retina cells to the intensity of the image on retina is a nonlinear function. This function can be approximated as the logarithm of the intensity of each pixel [84]. Applying the logarithm on an image will compress the range of bright pixel values, and conversely, expand the range of dark ones. By taking the logarithm of I , we have:

$$I' = \log(I) = \log(L) + \log(R) = L' + R', \quad (4.2)$$

where R' and L' are logarithms of reflectance and illumination, respectively. Now assuming that illumination lies in the low frequency part of the image I , we can obtain the illumination invariant information of the image by applying a 2D high-pass filter on I' . The output of the high-pass filter is illumination invariant and can be used for the recognition step, *e.g.*, using PCA or LDA. Figure 4.2 shows the flowchart of this system.

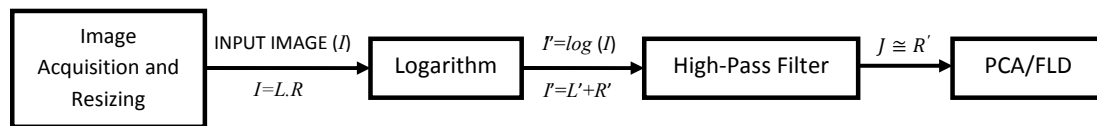


Figure 4.2: Extracting the illumination invariants based on the illumination-reflection model by applying a HPF on the logarithm of the image.

R' can be presented as $I' - L'$. Since the illumination invariant part of the image, R' , is sought, it can be determined by applying a 2D low-pass filter on the logarithm of the

input image and then subtracting the output, L' , from I' as shown in Figure 4.3. The result image, J , will be an approximation of R' .

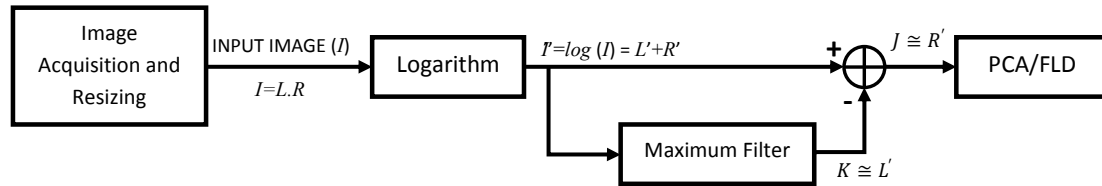


Figure 4.3: Extracting the illumination invariants by using LPFs (maximum filter).

In order to obtain the illumination invariants, 17 different high-pass and 31 low-pass linear or nonlinear 2D filters which are common in image processing are used in the two structures discussed in Figures 4.2 and 4.3. For classification task we used support vector machines, k nearest neighbors method and fuzzy k nearest neighbors method each with $k=1, 3$ and 5. Yale B and extended Yale B databases were used in this experiment. The maximum, Median, Bilateral and Wiener filters showed to have the best results. The

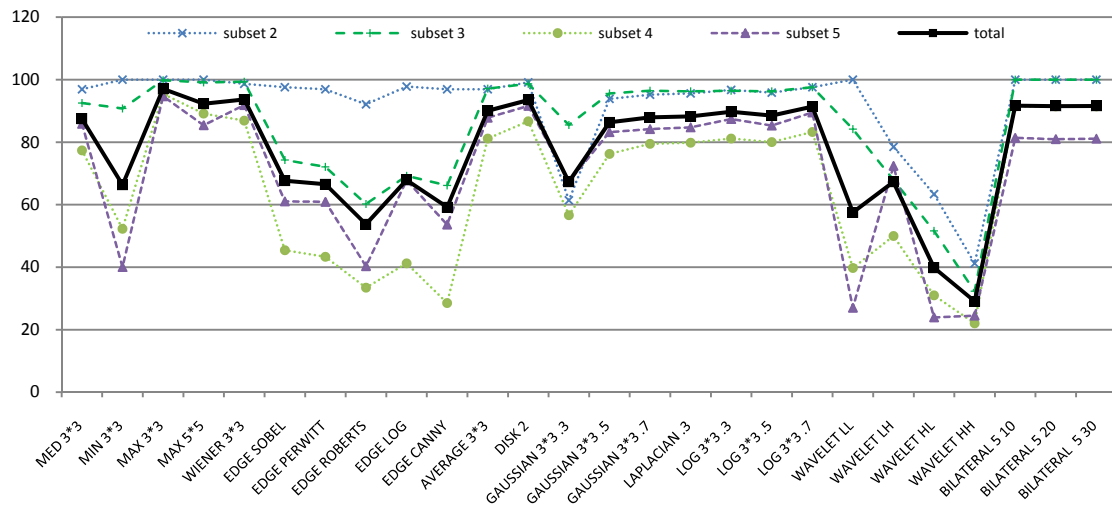


Figure 4.4: Recognition results using different lowpass and highpass filters in the two structures. (Extended Yale B database is used with PCA as feature extractor and SVM for classification).

results for all the filters used in the test with different classifiers and data bases are presented in tables 4.1 to 4.7

4.2.2 The Maximum Filter

Maximum filters, also known as dilation filters [85], are morphological filters that work by considering a neighborhood around each pixel. They are nonlinear 2D low-pass filters which are typically used in image processing to remove negative outlier noise. From the list of neighbor pixels, the maximum value is found and stored as the corresponding resulting value. Finally, each pixel in the image is replaced by the resulting value generated for its associated neighborhood. The Maximum filter is supposed to enhance bright values in the image by increasing its area. For each pixel in the output image, B , the intensity value is the maximum value in an $m \times n$ neighborhood in the input A .

$$B(x, y) = \max_{(i, j) \in \{m \times n \text{ neighborhood of } (x, y)\}} A(i, j) \quad (4.3)$$

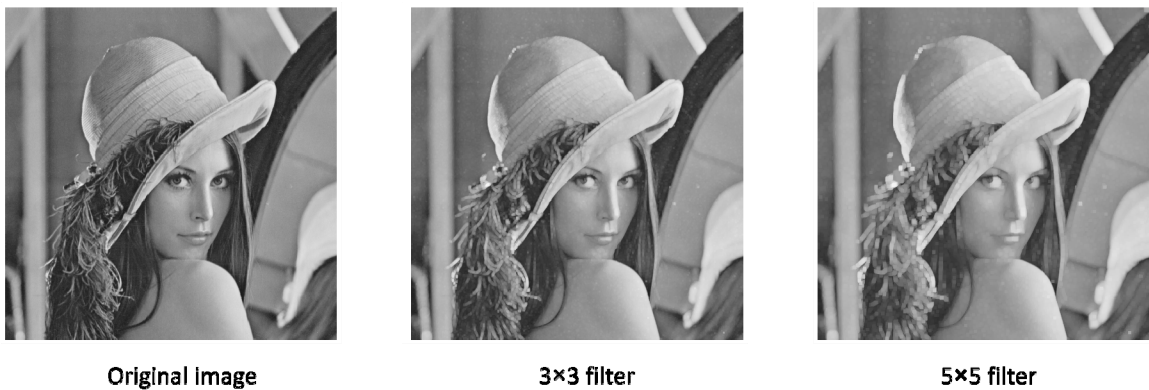


Figure 4.5: The results of applying 3×3 and 5×5 maximum filters on Lena image.

Combination of a dilation and subtraction in the images will produce bright edges which are the high frequency components in the image. In our experiment the maximum filter has the best results amongst all filters with different classifiers of support vector machine (SVM), k -nearest neighbors (k NN) and fuzzy k NN applied on Yale B, Extended Yale B and CMU-PIE databases.

Figure 4.6 shows the illumination invariants achieved by using Median, Gaussian and Maximum filters and also Sobel and Canny edge detectors. As it can be seen Maximum filter has successfully eliminated the effect of illumination in different subsets, and also maintains the details of the face while emphasizing the edges as the important features of the face. The average of the Euclidian distance between the illumination invariants of the image with good frontal illumination, and the ones with illumination variations has been reported in table 4.1 for these filters. Again we can see that the illumination invariants produced by Maximum filter show more similarity and have less distance so can be more accurate when used for feature extraction.

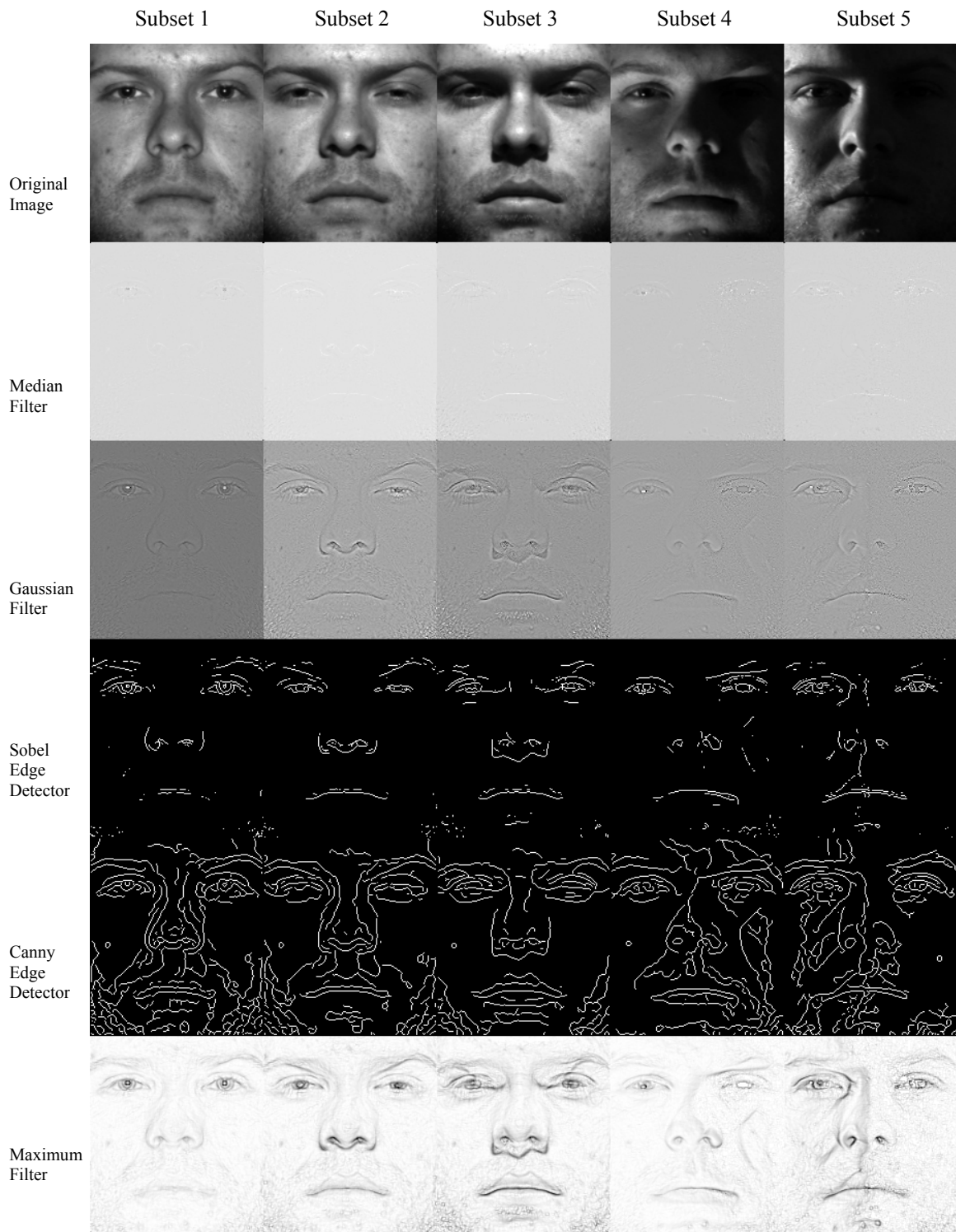


Figure 4.6: Illumination invariants achieved by different lowpass and highpass filters.

Table 4.1: A comparison between different filters. Values show the average Euclidian distance between the images with illumination variations and the image with frontal lightning. Maximum filter shows to have the best results.

	Subset 2	Subset 3	Subset 4	Subset 5
Median Filter	9.56	12.24	16.73	17.21
Gaussian Filter	12.30	20.06	19.86	19.51
Sobel Edge Detector	31.68	34.03	35.22	35.53
Canny Edge Detector	62.26	65.31	66.28	66.67
Maximum Filter	8.55	10.92	14.59	14.87

In our method we apply the filters on $\log(I)$ to achieve the illumination invariants and we use these illumination invariants for feature extraction. One of the important parts of the images that are used in feature extraction are edges. So the desired filter should have two properties. 1) it should be a good lowpass filter 2) it should emphasize the edges. While all other filters used are either good lowpass filters or good edge detectors, the Maximum filter has both properties at the same time and this is the key feature that leads to good results achieved by using the Maximum filter for canceling illumination. The lowpass character of the Maximum filter is shown in Figure 4.5 and also more described in [4]. Figure 4.7 shows edge enhancement property of Maximum filter that makes it unique among other lowpass filters. When applied on an edge, the maximum filter adds more pixels to the edge as dilation. As shown in Figure 4.6 maximum filter can eliminate the effect of illumination on the images while keeping the face features and also highlighting the edges.



Figure 4.7: Edge enhancement property of the Maximum filter. The right image is the result of applying Maximum filter on the left image

4.2.3 The Proposed Algorithm

The proposed method for extracting illumination invariants is described as follows:

- a) The face in the input image is detected, cropped and then is resized to $W_1 \times W_2$ image I to reduce the computation cost.
- b) The logarithm of the image I , I' , is calculated.
- c) The low frequency part of I' , K , is calculated by applying a maximum filter on I' .
- d) K is subtracted from I' to produce enhanced image J .

$$J = I' - K \cong R' \quad (4.4)$$

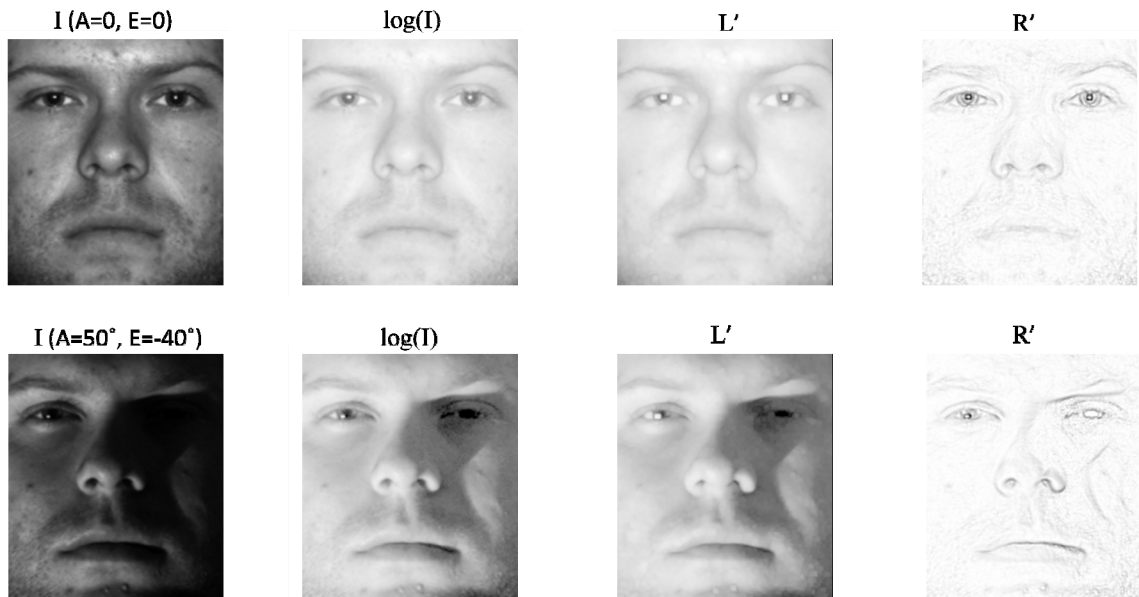


Figure 4.8: Different steps in extracting the illumination invariants. L' is the lowpass part of $\log(I)$. A and E are azimuth and elevation of the light source.

Figure 4.8 presents the images in each step of the algorithm for two different illuminations. The results of applying the proposed method on some samples in the Yale B and CMU-PIE databases are shown in Figures 4.9 and 4.10.

Table 4.2: Recognition rate for different filters and SVM as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	96.93	92.53	77.38	85.85
Median 5x5	100.00	98.90	93.92	94.82
MAX 3x3	100.00	99.78	95.44	94.68
MIN 3x3	100.00	90.77	52.28	40.06
MAX 5x5	100.00	99.12	89.16	85.43
MIN 5x5	100.00	93.85	63.69	47.06
WIENER 3x3	98.68	99.34	86.88	91.88
WIENER 5x5	99.56	99.12	91.25	94.82
EDGE SOBEL	97.59	74.29	45.44	61.06
EDGE PERWITT	96.93	72.09	43.35	60.92
EDGE ROBERTS	92.11	60.22	33.46	40.34
EDGE LOG	97.81	69.23	41.25	67.79
EDGE CANNY	96.93	66.15	28.52	53.64
AVERAGE 3x3	96.93	97.14	81.18	87.82
AVERAGE 5x5	99.56	98.90	90.68	93.84
DISK r=2	99.12	98.68	86.69	91.60
DISK r=3	100.00	99.34	92.02	93.28
DISK r=4	100.00	99.56	92.78	92.72
DISK r=5	100.00	99.56	92.97	91.74
GAUSSIAN 3x3 .3	61.40	85.49	56.65	67.37
GAUSSIAN 5x5 .3	61.40	85.27	56.84	67.37
GAUSSIAN 3x3 .5	93.86	95.60	76.24	83.19
GAUSSIAN 5x5 .5	93.86	95.82	76.24	83.47
GAUSSIAN 3x3 .7	95.18	96.48	79.47	84.17
GAUSSIAN 5x5 .7	96.05	96.48	80.23	86.13
LAPLACIAN .1	93.86	95.60	75.67	82.77
LAPLACIAN .2	94.74	96.26	78.14	83.89
LAPLACIAN .3	95.61	96.26	79.85	84.73
LOG 3x3 .3	96.71	96.48	81.18	87.39
LOG 5x5 .3	99.56	99.12	90.68	93.70
LOG 3x3 .5	95.83	96.26	80.04	85.29
LOG 5x5 .5	98.25	97.80	84.41	89.78
LOG 3x3 .7	97.59	97.58	83.27	89.50
LOG 5x5 .7	99.12	99.12	88.59	92.44
UNSHARP 0	93.64	95.16	72.43	81.79
UNSHARP .2	94.74	96.26	78.33	83.89
UNSHARP 1	98.03	98.02	85.17	90.76
DB4 WAVELET LL	100.00	84.18	39.73	27.03
DB4 WAVELET LH	78.51	67.91	50.00	72.41
DB4 WAVELET HL	63.38	51.65	30.99	23.95
DB4 WAVELET HH	41.23	32.31	22.05	24.51
BILATERAL 3 10	100.00	100.00	93.73	90.62
BILATERAL 5 10	100.00	100.00	91.44	81.37
BILATERAL 3 20	100.00	100.00	93.73	90.62
BILATERAL 5 20	100.00	100.00	91.44	80.95
BILATERAL 3 30	100.00	100.00	93.73	90.62
BILATERAL 5 30	100.00	100.00	91.44	81.09

Table 4.3: Recognition rate for different filters and k NN with $k=1$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	90.57	93.85	60.65	74.09
Median 5x5	99.34	100.00	88.02	89.50
MAX 3x3	100.00	100.00	96.01	92.44
MIN 3x3	100.00	93.19	54.94	42.02
MAX 5x5	100.00	99.56	90.87	87.68
MIN 5x5	100.00	95.16	62.93	50.56
WIENER 3x3	96.49	99.56	79.47	82.35
WIENER 5x5	99.56	100.00	85.93	89.36
EDGE SOBEL	93.64	63.52	29.66	41.46
EDGE PERWITT	95.18	63.74	32.13	42.44
EDGE ROBERTS	86.62	48.79	25.86	32.07
EDGE LOG	98.25	87.03	43.35	68.49
EDGE CANNY	95.39	83.30	32.89	55.74
AVERAGE 3x3	93.42	98.68	75.29	79.69
AVERAGE 5x5	98.68	100.00	88.02	87.82
DISK r=2	97.15	99.34	81.75	83.89
DISK r=3	99.56	100.00	88.21	88.52
DISK r=4	100.00	100.00	92.59	90.06
DISK r=5	100.00	100.00	93.54	89.36
GAUSSIAN 3x3 .3	89.47	95.38	64.64	73.81
GAUSSIAN 5x5 .3	89.47	95.38	64.64	74.37
GAUSSIAN 3x3 .5	90.57	96.70	68.82	76.33
GAUSSIAN 5x5 .5	90.79	96.70	69.01	76.19
GAUSSIAN 3x3 .7	92.32	97.14	71.10	78.29
GAUSSIAN 5x5 .7	92.76	98.02	73.38	79.27
LAPLACIAN .1	90.35	96.48	68.44	76.05
LAPLACIAN .2	91.67	96.92	70.53	77.59
LAPLACIAN .3	92.54	97.14	72.05	78.43
LOG 3x3 .3	93.42	98.02	75.10	79.55
LOG 5x5 .3	98.46	100.00	87.64	87.54
LOG 3x3 .5	92.54	97.58	72.24	78.99
LOG 5x5 .5	95.83	99.34	79.66	82.07
LOG 3x3 .7	94.74	98.90	77.95	81.09
LOG 5x5 .7	97.59	99.78	84.60	85.01
UNSHARP 0	89.69	95.38	64.83	73.95
UNSHARP .2	91.67	96.92	70.34	77.45
UNSHARP 1	95.61	98.90	79.28	81.93
DB4 WAVELET LL	99.78	85.05	40.87	26.75
DB4 WAVELET LH	57.02	55.38	29.47	48.32
DB4 WAVELET HL	35.75	35.82	18.63	15.97
DB4 WAVELET HH	23.68	19.78	13.31	13.59
BILATERAL 3 10	100.00	100.00	94.11	89.92
BILATERAL 5 10	100.00	100.00	94.11	86.55
BILATERAL 3 20	100.00	100.00	94.11	89.78
BILATERAL 5 20	100.00	100.00	94.30	86.27
BILATERAL 3 30	100.00	100.00	94.11	89.78
BILATERAL 5 30	100.00	100.00	94.30	86.41

Table 4.4: Recognition rates for different filters and k NN with $k=3$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	70.39	90.99	56.08	71.15
Median 5x5	88.38	99.78	83.84	87.82
MAX 3x3	96.49	100.00	93.35	93.00
MIN 3x3	96.71	89.67	52.09	42.44
MAX 5x5	98.68	99.78	89.73	88.66
MIN 5x5	98.46	91.65	61.22	48.18
WIENER 3x3	74.34	99.34	67.68	76.33
WIENER 5x5	85.53	100.00	81.18	86.41
EDGE SOBEL	80.48	59.34	28.14	38.80
EDGE PERWITT	81.80	62.42	31.94	41.46
EDGE ROBERTS	71.05	46.15	22.24	28.29
EDGE LOG	85.75	82.42	40.49	67.79
EDGE CANNY	82.02	77.58	29.09	54.48
AVERAGE 3x3	69.52	97.14	70.34	76.75
AVERAGE 5x5	85.96	100.00	82.13	85.71
DISK $r=2$	79.61	99.34	75.29	80.53
DISK $r=3$	87.28	100.00	84.22	86.13
DISK $r=4$	93.42	100.00	87.64	88.80
DISK $r=5$	95.83	100.00	88.02	88.52
GAUSSIAN 3x3 .3	65.79	94.07	57.98	68.91
GAUSSIAN 5x5 .3	65.79	94.07	58.37	68.63
GAUSSIAN 3x3 .5	66.89	95.60	62.36	71.15
GAUSSIAN 5x5 .5	66.89	95.60	62.93	71.15
GAUSSIAN 3x3 .7	68.20	96.26	64.83	74.51
GAUSSIAN 5x5 .7	69.74	97.14	68.06	75.91
LAPLACIAN .1	66.67	95.60	61.41	70.87
LAPLACIAN .2	67.11	96.04	64.07	73.25
LAPLACIAN .3	69.08	96.48	66.54	75.35
LOG 3x3 .3	69.30	97.14	69.96	76.61
LOG 5x5 .3	84.87	100.00	81.56	85.57
LOG 3x3 .5	69.30	96.70	67.87	75.49
LOG 5x5 .5	74.34	99.12	73.00	78.99
LOG 3x3 .7	73.03	97.80	72.81	78.57
LOG 5x5 .7	81.80	99.56	78.33	82.63
UNSHARP 0	65.79	93.85	57.79	68.63
UNSHARP .2	67.11	96.26	64.07	73.25
UNSHARP 1	75.22	98.02	73.95	78.99
DB4 WAVELET LL	94.96	81.76	37.64	25.35
DB4 WAVELET LH	41.89	42.64	19.20	36.69
DB4 WAVELET HL	24.78	29.01	11.79	12.46
DB4 WAVELET HH	18.64	19.34	12.36	13.31
BILATERAL 3 10	98.25	100.00	89.92	88.66
BILATERAL 5 10	98.90	99.78	88.59	85.43
BILATERAL 3 20	98.25	100.00	89.54	88.52
BILATERAL 5 20	98.90	99.78	88.59	85.15
BILATERAL 3 30	98.03	100.00	89.54	88.52
BILATERAL 5 30	98.90	99.78	88.59	85.01

Table 4.5: Recognition rates for different filters and k NN with $k=5$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	68.20	83.52	53.80	67.79
Median 5x5	87.72	98.24	80.42	85.29
MAX 3x3	96.05	100.00	92.78	91.04
MIN 3x3	96.05	88.79	50.57	40.90
MAX 5x5	98.25	98.68	88.78	87.54
MIN 5x5	98.25	90.55	57.79	45.94
WIENER 3x3	73.25	98.68	66.35	76.89
WIENER 5x5	84.65	99.34	77.95	85.43
EDGE SOBEL	77.19	59.78	28.33	39.36
EDGE PERWITT	77.63	59.12	30.80	41.88
EDGE ROBERTS	68.64	47.69	24.71	30.11
EDGE LOG	83.99	76.26	40.30	66.81
EDGE CANNY	78.95	76.04	30.99	54.48
AVERAGE 3x3	69.08	95.82	66.73	74.93
AVERAGE 5x5	84.87	99.34	80.04	84.59
DISK r=2	76.97	98.46	72.81	79.13
DISK r=3	86.18	99.34	81.37	85.57
DISK r=4	92.54	99.78	84.41	86.83
DISK r=5	94.96	99.78	85.93	87.11
GAUSSIAN 3x3 .3	62.50	90.33	53.42	65.83
GAUSSIAN 5x5 .3	62.28	90.33	53.42	65.83
GAUSSIAN 3x3 .5	64.25	91.65	58.17	69.61
GAUSSIAN 5x5 .5	64.69	91.65	58.17	69.61
GAUSSIAN 3x3 .7	66.89	93.19	61.22	73.25
GAUSSIAN 5x5 .7	68.42	94.73	64.26	74.79
LAPLACIAN .1	64.04	91.65	57.79	69.05
LAPLACIAN .2	65.79	92.75	60.46	71.57
LAPLACIAN .3	67.32	94.07	62.55	73.53
LOG 3x3 .3	68.64	95.82	66.16	74.79
LOG 5x5 .3	83.99	99.34	79.47	84.17
LOG 3x3 .5	67.98	94.29	63.31	73.81
LOG 5x5 .5	72.81	96.48	70.53	76.75
LOG 3x3 .7	72.15	95.82	70.72	76.89
LOG 5x5 .7	80.48	98.90	75.86	81.23
UNSHARP 0	62.50	89.89	52.85	65.55
UNSHARP .2	65.79	92.97	59.89	71.71
UNSHARP 1	73.46	96.04	72.43	77.73
DB4 WAVELET LL	94.96	78.68	33.46	22.83
DB4 WAVELET LH	39.25	39.12	17.68	32.49
DB4 WAVELET HL	21.71	23.52	10.84	12.75
DB4 WAVELET HH	16.23	18.46	11.98	13.31
BILATERAL 3 10	96.49	99.78	87.45	87.25
BILATERAL 5 10	98.46	99.34	86.31	82.63
BILATERAL 3 20	96.49	99.78	87.45	86.83
BILATERAL 5 20	98.46	99.34	86.12	82.63
BILATERAL 3 30	96.49	99.78	87.45	87.11
BILATERAL 5 30	98.46	99.34	86.12	82.63

Table 4.6: Recognition rates for different filters and fuzzy k NN with $k=1$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	90.57	93.85	60.65	74.09
Median 5x5	99.34	100.00	88.02	89.50
MAX 3x3	100.00	100.00	96.01	92.44
MIN 3x3	100.00	93.19	54.94	42.02
MAX 5x5	100.00	99.56	90.87	87.68
MIN 5x5	100.00	95.16	62.93	50.56
WIENER 3x3	96.49	99.56	79.47	82.35
WIENER 5x5	99.56	100.00	85.93	89.36
EDGE SOBEL	93.64	63.52	29.66	41.46
EDGE PERWITT	95.18	63.74	32.13	42.44
EDGE ROBERTS	86.62	48.79	25.86	32.07
EDGE LOG	98.25	87.03	43.35	68.49
EDGE CANNY	95.39	83.30	32.89	55.74
AVERAGE 3x3	93.42	98.68	75.29	79.69
AVERAGE 5x5	98.68	100.00	88.02	87.82
DISK r=2	97.15	99.34	81.75	83.89
DISK r=3	99.56	100.00	88.21	88.52
DISK r=4	100.00	100.00	92.59	90.06
DISK r=5	100.00	100.00	93.54	89.36
GAUSSIAN 3x3 .3	89.47	95.38	64.64	73.81
GAUSSIAN 5x5 .3	89.47	95.38	64.64	74.37
GAUSSIAN 3x3 .5	90.57	96.70	68.82	76.33
GAUSSIAN 5x5 .5	90.79	96.70	69.01	76.19
GAUSSIAN 3x3 .7	92.32	97.14	71.10	78.29
GAUSSIAN 5x5 .7	92.76	98.02	73.38	79.27
LAPLACIAN .1	90.35	96.48	68.44	76.05
LAPLACIAN .2	91.67	96.92	70.53	77.59
LAPLACIAN .3	92.54	97.14	72.05	78.43
LOG 3x3 .3	93.42	98.02	75.10	79.55
LOG 5x5 .3	98.46	100.00	87.64	87.54
LOG 3x3 .5	92.54	97.58	72.24	78.99
LOG 5x5 .5	95.83	99.34	79.66	82.07
LOG 3x3 .7	94.74	98.90	77.95	81.09
LOG 5x5 .7	97.59	99.78	84.60	85.01
UNSHARP 0	89.69	95.38	64.83	73.95
UNSHARP .2	91.67	96.92	70.34	77.45
UNSHARP 1	95.61	98.90	79.28	81.93
DB4 WAVELET LL	99.78	85.05	40.87	26.75
DB4 WAVELET LH	57.02	55.38	29.47	48.32
DB4 WAVELET HL	35.75	35.82	18.63	15.97
DB4 WAVELET HH	23.68	19.78	13.31	13.59
BILATERAL 3 10	100.00	100.00	94.11	89.92
BILATERAL 5 10	100.00	100.00	94.11	86.55
BILATERAL 3 20	100.00	100.00	94.11	89.78
BILATERAL 5 20	100.00	100.00	94.30	86.27
BILATERAL 3 30	100.00	100.00	94.11	89.78
BILATERAL 5 30	100.00	100.00	94.30	86.41

Table 4.7: Recognition rates for different filters and fuzzy k NN with $k=3$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	84.21	92.97	58.75	72.69
Median 5x5	96.05	99.78	84.98	88.52
MAX 3x3	100.00	100.00	94.87	93.00
MIN 3x3	99.34	91.87	55.13	42.16
MAX 5x5	100.00	99.78	90.11	89.22
MIN 5x5	99.34	94.73	62.36	50.28
WIENER 3x3	92.76	99.56	70.72	77.03
WIENER 5x5	96.27	100.00	81.75	86.83
EDGE SOBEL	89.47	60.88	28.71	41.32
EDGE PERWITT	91.23	63.30	31.18	43.00
EDGE ROBERTS	85.31	50.33	25.67	33.19
EDGE LOG	92.98	85.27	42.21	69.33
EDGE CANNY	87.28	79.12	29.09	55.18
AVERAGE 3x3	88.16	97.58	72.43	77.73
AVERAGE 5x5	94.74	100.00	82.89	86.13
DISK r=2	93.64	99.34	76.62	81.51
DISK r=3	95.83	100.00	84.79	86.69
DISK r=4	98.25	100.00	88.02	89.22
DISK r=5	99.34	100.00	89.16	88.66
GAUSSIAN 3x3 .3	82.46	94.95	60.08	70.31
GAUSSIAN 5x5 .3	82.46	94.95	60.08	70.45
GAUSSIAN 3x3 .5	84.43	96.04	64.07	72.69
GAUSSIAN 5x5 .5	84.65	96.04	64.45	72.69
GAUSSIAN 3x3 .7	85.09	96.48	66.35	75.35
GAUSSIAN 5x5 .7	87.28	97.36	70.34	77.03
LAPLACIAN .1	83.99	96.04	63.69	72.27
LAPLACIAN .2	85.09	96.26	65.40	74.37
LAPLACIAN .3	86.18	96.70	68.63	76.19
LOG 3x3 .3	87.94	97.58	72.05	77.59
LOG 5x5 .3	93.64	100.00	82.32	85.99
LOG 3x3 .5	86.40	97.14	69.39	76.75
LOG 5x5 .5	91.45	99.12	74.71	79.69
LOG 3x3 .7	89.47	98.02	75.48	79.69
LOG 5x5 .7	93.64	99.56	80.04	83.19
UNSHARP 0	82.46	94.95	60.08	70.31
UNSHARP .2	84.43	96.26	65.21	74.37
UNSHARP 1	91.01	98.24	76.81	80.11
DB4 WAVELET LL	99.34	82.42	38.40	25.77
DB4 WAVELET LH	53.07	50.33	24.90	44.26
DB4 WAVELET HL	33.11	33.19	16.73	14.71
DB4 WAVELET HH	22.59	19.78	13.31	13.59
BILATERAL 3 10	100.00	100.00	90.87	88.80
BILATERAL 5 10	99.78	99.78	90.30	85.85
BILATERAL 3 20	100.00	100.00	90.49	88.66
BILATERAL 5 20	99.78	99.78	90.49	85.57
BILATERAL 3 30	100.00	100.00	90.49	88.66
BILATERAL 5 30	99.78	99.78	90.49	85.57

Table 4.8: Recognition rates for different filters and fuzzy k NN with $k=5$ as classifier.

Filter Type	Extended Yale B Database			
	Subset2	Subset3	Subset4	Subset5
Median 3x3	78.07	90.77	58.94	72.69
Median 5x5	91.89	98.68	84.03	86.83
MAX 3x3	99.12	100.00	93.16	92.58
MIN 3x3	98.25	91.87	54.56	41.04
MAX 5x5	100.00	99.56	89.54	88.10
MIN 5x5	99.56	94.29	61.41	48.60
WIENER 3x3	87.28	99.34	69.58	78.85
WIENER 5x5	91.67	99.56	80.04	85.85
EDGE SOBEL	84.21	60.00	29.09	39.92
EDGE PERWITT	84.21	59.56	31.18	42.16
EDGE ROBERTS	78.29	50.33	26.81	32.63
EDGE LOG	88.82	80.22	42.59	68.49
EDGE CANNY	82.46	78.46	29.66	53.92
AVERAGE 3x3	82.46	97.14	70.53	76.33
AVERAGE 5x5	90.13	99.56	81.18	85.85
DISK r=2	86.84	99.56	75.10	80.53
DISK r=3	91.45	99.56	81.75	86.41
DISK r=4	96.93	99.78	85.36	87.54
DISK r=5	98.68	99.78	86.88	87.82
GAUSSIAN 3x3 .3	73.68	93.41	56.46	68.91
GAUSSIAN 5x5 .3	73.46	93.41	56.27	68.91
GAUSSIAN 3x3 .5	75.66	95.16	61.41	71.85
GAUSSIAN 5x5 .5	76.10	94.95	61.79	71.85
GAUSSIAN 3x3 .7	78.29	96.04	65.02	74.79
GAUSSIAN 5x5 .7	80.70	97.14	68.25	75.91
LAPLACIAN .1	75.00	94.95	61.03	71.57
LAPLACIAN .2	77.41	95.82	63.50	73.25
LAPLACIAN .3	79.39	96.26	66.35	74.93
LOG 3x3 .3	81.36	97.14	69.96	76.19
LOG 5x5 .3	89.25	99.56	80.42	85.43
LOG 3x3 .5	79.61	96.26	66.73	75.21
LOG 5x5 .5	83.33	98.90	73.00	78.85
LOG 3x3 .7	82.68	97.80	73.57	78.29
LOG 5x5 .7	88.82	99.56	77.00	82.07
UNSHARP 0	73.03	93.41	56.08	68.63
UNSHARP .2	77.41	95.82	63.50	73.25
UNSHARP 1	84.21	98.46	74.52	78.43
DB4 WAVELET LL	98.25	80.44	35.17	22.83
DB4 WAVELET LH	48.03	48.57	21.48	41.04
DB4 WAVELET HL	28.73	29.67	13.88	13.87
DB4 WAVELET HH	21.27	20.22	13.12	13.03
BILATERAL 3 10	99.12	99.78	88.97	87.96
BILATERAL 5 10	99.56	99.56	89.54	85.15
BILATERAL 3 20	99.12	99.78	88.97	87.82
BILATERAL 5 20	99.34	99.56	89.54	85.01
BILATERAL 3 30	99.12	99.78	88.97	87.96
BILATERAL 5 30	99.34	99.56	89.35	85.01



Figure 4.9: One of the subjects in the Yale B database with 5 corresponding subsets and the illumination invariants generated using the proposed method: (a) subset 1 (θ upto 12°) (b) subset 2 (θ upto 25°) (c) subset 3 (θ upto 50°) (d) subset 4 (θ upto 77°).



Figure 4.10: One of the samples in CMU-PIE database and its illumination invariants (a) lights-off group (b) lights-on group (c) and (d) are the corresponding illumination invariants to (a) and (b).

4.3 Experimental results

The results of the proposed method are compared with similar illumination invariant methods in tables 4.8 to 4.11. Since studying illumination invariance is the focus of this work, for Yale B, Extended Yale B and CMU-PIE databases, we use the images in frontal pose only which are 64 per person. The Yale B and Extended Yale B databases are divided into five different subsets according to the angle between the light source direction and the camera axis, θ . Each subset contains images with specific range of θ . Subset 1 considers $\theta < 12^\circ$, subset 2 considers $20^\circ < \theta < 25^\circ$, subset 3 considers $35^\circ < \theta < 50^\circ$, subset 4 considers $60^\circ < \theta < 77^\circ$, and finally subset 5 contains face images with $\theta > 78^\circ$.

CMU Pose, Illumination, and Expression (CMU-PIE) database [53] consists of 41368 images of 68 people. Each person has been imaged under 13 different poses, 43 different illumination conditions with lights ON and lights OFF, and with 4 different expressions. We used the images with both lights ON and OFF in frontal pose (camera27).

For all the images in the databases, faces are cropped into 192×168 pixel images. The position of the face in the images is normalized in a way that the eyes and the lips have almost the same position in all the images. For the Yale B and Extended Yale B databases, we use images in subset 1 which are taken under small illumination variations for training and other images in four other subsets for testing. Subset 1 has 7 images per person which make our training gallery and subsets 2 to 5 have 12, 12, 14, and 19 images per person, respectively, forming the test sets. For the CMU-PIE database, for each subject, there are 43 images available where we selected a random group of 3 to 7 images

out of these 43 for the training set and the rest formed the testing set. We repeated the algorithm and the averages of the results of 200 simulations for this experiment are shown in Table 4.9.

As it can be seen in Tables 4.9 to 4.12, the proposed method outperforms other methods achieving superior results in comparison with all other approaches studied here.

Table 4.9. Recognition results for 200 simulations on CMU-PIE database.

Train set	Test set	# train samples	min	max	average
Lights on	Lights on	3/21	91.5	97.63	94.85
Lights off	Lights off	4/24	99.70	100	99.99
Lights on+off	Lights on+off	7/45	98.33	99.88	99.74
Lights on+off	Lights on+off	6/45	98.45	99.81	99.55
Lights on+off	Lights on+off	5/45	96.73	99.12	98.99

Table 4.10. Recognition rates using different methods for CMU-PIE database.

Method	CMU-PIE
PCA [13]	54
QI [76]	84
SQI [77]	99.33
Proposed Illum. Inv + PCA	99.35

Table 4.11. Recognition rates using different methods for Yale B database (10 subjects).

Method	Subset 2	Subset 3	Subset 4	Subset 5
PCA [13]	98.3	79.2	30.0	15.8
Histogram Equalization[5]	100	89.0	55.1	44.4
GIC[5]	100	88.1	39.9	27.5
Linear subspace [71]	100	100	85.0	-
Cones attached [72]	100	100	91.4	-
Cones cast [72]	100	100	100	-
QI [76]	99.3	61.9	34.1	23.3
QIR [5]	100	100	90.6	78.8
SQI [77]	100	100	96.4	97.9
MQI [86]	100	100	100	98.4
Proposed Illum. Inv + PCA	100	100	98.6	98.9
(#misclassified/#total)	0/120	0/120	2/140	2/190

Table 4.12. Recognition rates using the Extended Yale B database (38 subjects).

Method	Subset 2	Subset 3	Subset 4	Subset 5
PCA[13]	90	41	6	3
LDA[14]	100	98	38	5
Proposed Illum. Inv +PCA+SVM	100	99.78	95.44	94.68
(#misclassified/#total)	0/456	1/455	24/526	38/714
Proposed Illum. Inv +PCA+INN	100	100	96.01	92.44
(#misclassified/#total)	0/456	0/455	21/526	54/714

4.4 Conclusion

An efficient method for illumination invariant face recognition was proposed in this chapter. The proposed method achieved the illumination invariants based on the illumination-reflection model. Different filters were tested to achieve the reflectance part of the image which is illumination invariant and maximum filter was suggested as the best method for this purpose. We showed that the illumination invariants gained by this method result in better recognition comparing to other methods like QI, SQI or image enhancement methods. The proposed method does not need any prior information about the face shape or illumination and can be applied on each image separately. Unlike most available methods, our method does not need multiple images in training stage to get the illumination invariants. Support vector machines and k -nearest neighbors methods were used as classifiers. Several experiments were performed on Yale B, extended Yale B and CMU-PIE databases. Recognition results show that the proposed method is suitable for efficient face recognition under illumination variations and out performs other similar methods.

Chapter 5

Local Matching and Voting Systems

Although the results of the method described in chapter 4 are better than most of the algorithms for illumination invariant face recognition, however, there are still some artifacts resulted by illumination changes in the retinex images. These artifacts are not global and just appear as dark lines in small parts of the face images as shown in Figure 5.1 (e and f). In this chapter we try to eliminate the effect of these artifacts and, therefore, improve the overall recognition rate by using local matching methods [79-82] which apply recognition on sub-images of the faces. By reducing the effect of the corrupted sub-images the system can make more accurate decision for classification. For each input image a set of adaptively weighted classifiers vote on different sub-images and a decision is made based on their votes. Our voting scheme makes a decision based on the results of

different classifiers applied on sub-images of the input face image. The proposed method combines the local and global information of the images by using the original image as well as the sub-images. We use image entropy and mutual information between sub-images to generate the weights. Unlike other modular methods, weights are not applied on the distance or similarity functions and are simply applied on the output of the classifiers.

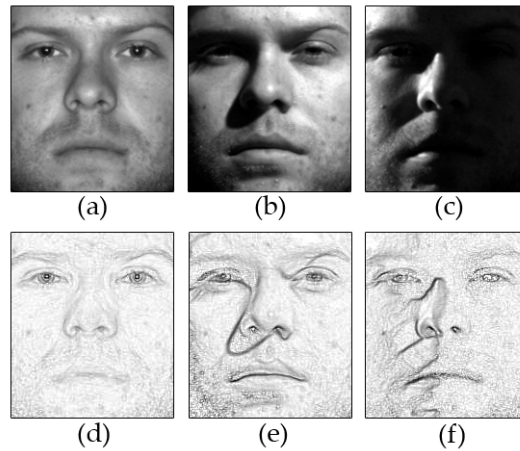


Figure 5.1: Samples of images with different angle between the light source and camera axes, θ , (a): $\theta = 0^\circ$, (b) $\theta = 40^\circ$ (c) $\theta = 90^\circ$, (d), (e) and (f) are corresponding illumination invariant images, respectively.

5.1 Democratic Voting Scheme

Holistic feature extraction methods such as PCA use the global information of the face images. This limits their performance in environments with facial expression, illumination variation or occlusion. One approach toward solving this problem is to utilize face's local information in the feature extraction by using a group of methods that we call them here as democratic voting methods. The main idea is to divide the face

images into l smaller sub-images forming a group of l datasets for each part of the face. A conventional feature extraction and classification process is applied on each of these l datasets and recognition results are achieved as an $l \times 1$ vector of classes. The final decision is made based on the classes in the output vector with equal weight values as presented in Figure 5.2. Therefore, in this structure, face expression or illumination variations in the image will only affect some sub-images rather than the whole image in global methods. This local representation leads to better overall recognition in poor illumination conditions or occlusion. In modular PCA (mPCA) [79], feature extraction and classification are PCA, while in modular LDA (mLDA) [80], Fisher Discriminant analysis is used for feature extraction and classification. In our method, we use PCA as feature extractor and SVM as classifier. Using LDA in feature extraction reduces the effect of illumination variations, however our results show that by employing the method described in chapter 4, these variations are mostly canceled and LDA cannot increase the robustness toward unequal illumination conditions. As it can be seen in Figure 5.1, only minor artifacts exist in the illumination invariant images generated from our filtering method and by dividing the images into several sub-images and performing the classification, the parts that are not affected by these artifacts contribute to a correct recognition. Since the number of these sub-images is normally more than the corrupt ones, suitable results are obtained by democratic voting scheme (DVS), which will eliminate the votes of the corrupt sub-images by considering the class with maximum votes as the output class.

In our algorithm, illumination invariant images are achieved using the method described in chapter 4. Each illumination invariant image is divided into a set of $l = m^2$ equally sized non-overlapping sub-images. Each image has a size of $W_1 \times W_2$ while the size of each sub-image is $w_1 \times w_2$, where $w_1 = W_1/m$ and $w_2 = W_2/m$, respectively. Assume that we have M persons and each of them have N images available in the training set. Each image is divided into l sub-images and concatenated into $M \times N \times l$ corresponding vectors with $w_1 \times w_2$ elements. The vectors related to the same position of all face images form a $(w_1 \times w_2) \times (M \times N)$ training matrix representing that specific position of the faces. Accordingly, l matrices corresponding to l positions are generated.

In many cases identifying the subject based on small sub-images is harder than recognition using original image. The best result is achieved when both local information of sub-images and the global information of the original image are used together. As a combination of the global and local image information, we added the original image into the decision making structure (Figure. 5.2) which did not exist in previous methods like mPCA. Since the global image is holding more information than the other sub-images, we allocate a weight W_0 to the result of the recognition based on the global image and combine it with the results of the classification of sub-images. For each input image in the test dataset, the results of the l classifiers along with the result of the global classifier are sent into a simple majority vote counter where each sub-image has one vote and the main image has W_0 votes. For best results, W_0 is chosen around m . The person with maximum votes will be recognized as the desired subject.

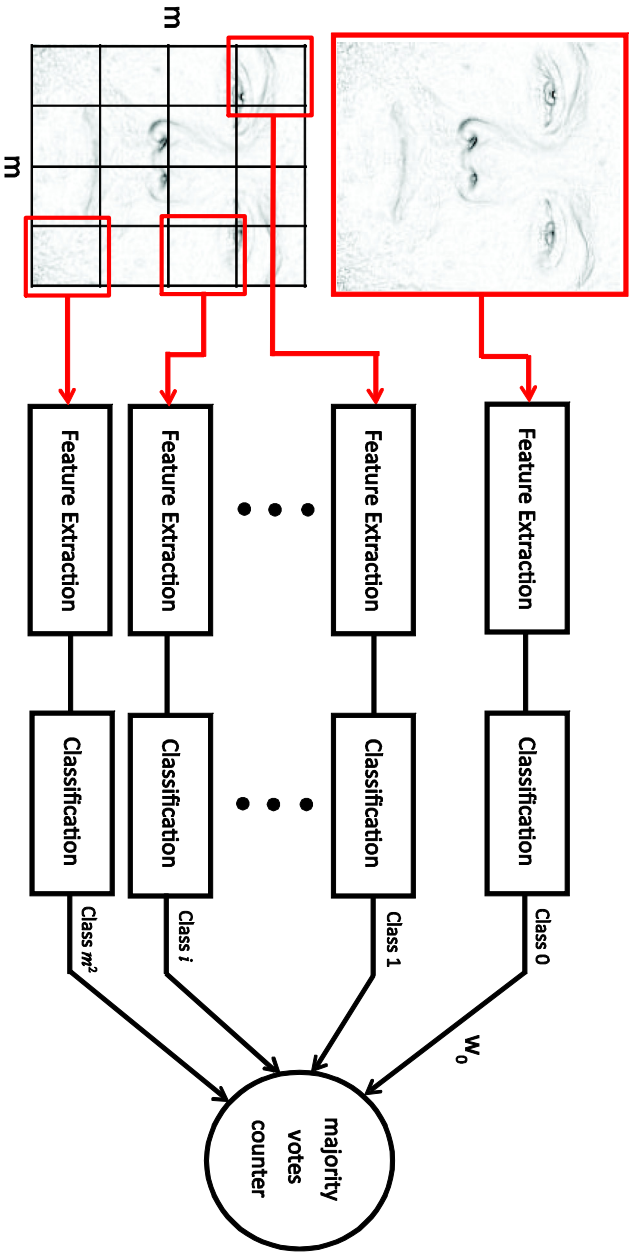


Figure 5.2: Democratic Voting Scheme (DVS).

5.2 Weighted Voting Scheme

The DVS method described in section 5.1 reduces the effect of illumination artifacts in the images. All sub-images in DVS have equal weights in the majority vote counter unit. For a given sub-image, poor illumination, generally, leads to poor recognition rate. Also some parts of the face image like eyes, for example, have more important information and hence can be more useful in the recognition process. The question here is why a sub-image with bad lighting condition or low spatial information should have equal voting right as the sub-images with good illumination condition or better face detail information. As a result we propose the Weighted Voting Scheme (WVS), presented in Figure 5.3. In this system, an adaptive weight generator calculates proper weights for each sub-image based on three main factors, sub-image mean, sub-image entropy and mutual information between corresponding sub-images.

5.2.1 Mean Grayscale Level

The mean of the gray level values in each sub-image, represents the illumination condition in that position. If this mean value is too low, it means that sub-image is located in the shadow of the face where the pixel values are lower. Sub-images with low illumination are more probable to result in a wrong recognition, and therefore, should have low voting weight. For a given sub-image position in order to achieve a factor showing the illumination condition in that position, we find the mean of the gray-pixel values of the original image in the window of that particular sub-image and compare it to the mean of the same window in the canonical image. The canonical image is the average

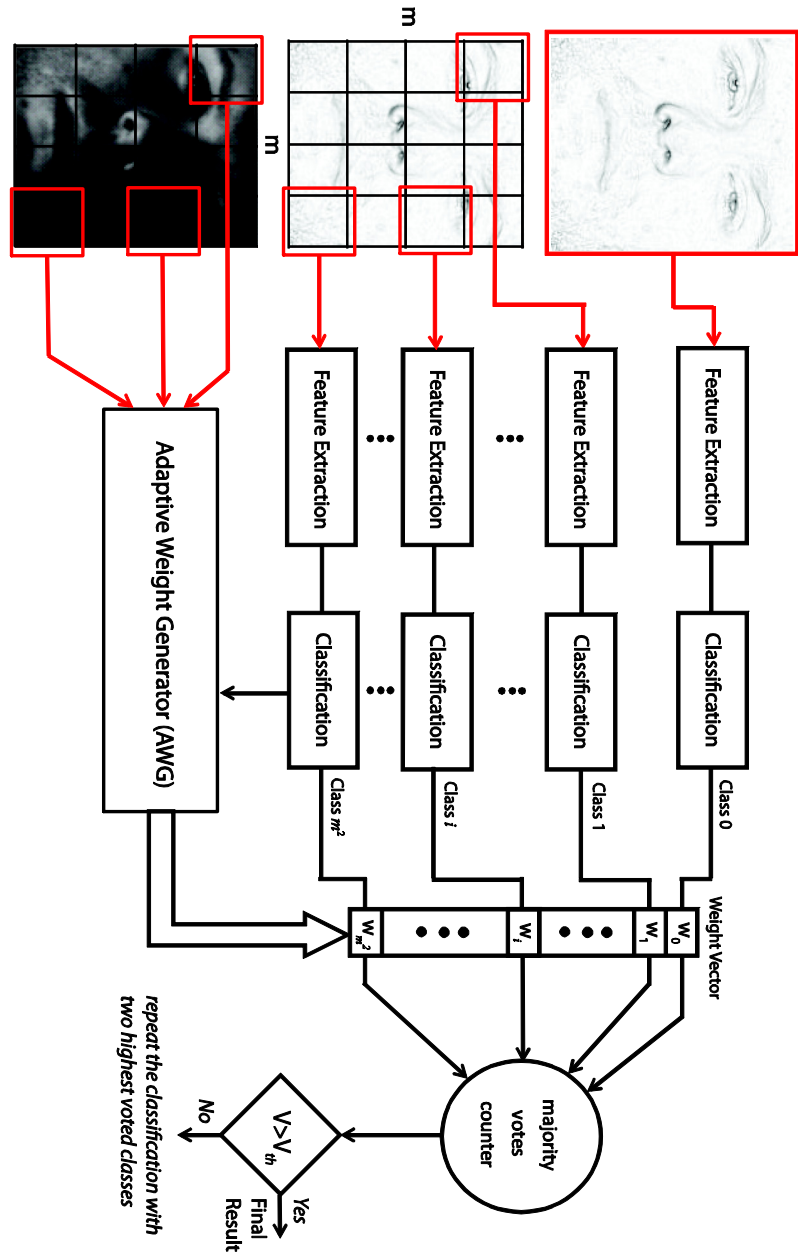


Figure 5.3: The Weighted Voting Scheme (WVS).

of all the images in the training set which are taken under good lighting conditions. The mean of each sub-image in the canonical image will be a reference value for good illumination. For a given sub-image, if the mean is closer to this reference value, it means it has better illumination and will have a better chance for correct recognition; as a result, this sub-image will have a higher weight.

$$d_i = |mean_i - mean_{i_{canonical}}| \quad (5.1)$$

d_i generated in equation 5.1 is low for sub-images with good illumination and high for the ones with bad illumination. For our work we considered a maximum d_i that can be tolerated by the system, d_{max} . If d_i is more than d_{max} then it will be replaced by d_{max} . Maximum tolerable difference of 80 out of the 256 grayscale values was used in our work.

5.2.2 Image Entropy

For a given image $I(x,y)$, entropy can be defined as [87]:

$$H[I(x,y)] = -\sum_{i=1}^{N_g} p_i \log(p_i), \quad (5.2)$$

where p_i is the probability of the i th gray level value and N_g is the total number of gray level values. Image entropy is related to the amount of information in the image and in our work it provides a good measure for the amount of face texture information contained in each sub-image. If a given sub-image is flat or has low texture information, it is more probable that it results in a wrong recognition answer. Therefore, if the entropy of a sub image is low, it should lead to less voting weight for that sub-image and vice versa.

5.2.3 Mutual information between corresponding sub-images

The mutual information of two discrete random variables X and Y quantifies the dependence between the joint distribution of X and Y and can be defined as [88]:

$$MI(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left(\frac{p(x, y)}{p_1(x)p_2(y)} \right), \quad (5.3)$$

where $p(x, y)$ is the joint probability distribution function of X and Y , and $p_1(x)$ and $p_2(y)$ are the marginal probability distribution functions of X and Y , respectively. Mutual information can also be expressed based on the joint entropy and the conditional entropy. The joint entropy measures how much entropy is contained in a joint system of two random variables. If the random variables are X and Y , the joint entropy is written as [89]:

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log(p(x, y)) \quad (5.4)$$

The conditional entropy is a statistics that measures the randomness of Y given knowledge of X . It is defined by [90]:

$$H(X|Y) = - \sum_{x \in X} \sum_{y \in Y} p(x|y) \log(p(x|y)) \quad (5.5)$$

For mutual information we have:

$$\begin{aligned} MI(X, Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X, Y) \\ &= H(X, Y) - H(X|Y) - H(Y|X) \end{aligned} \quad (5.6)$$

where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, and $H(X, Y)$ is the joint entropy of X and Y .

Pixel values of an image can be considered as outcomes of a random variable. Pixels with common pixel values contain little information and pixels with uncommon values are regarded as more information. While entropy for an image remains fixed, joint entropy and mutual information of two images vary between different images. If two images are identical, joint entropy is minimized and mutual information is maximized. In our proposed method we use the mutual information between the input sub-images and their corresponding sub-images in the training set to generate the weights. In the structure presented in Figure 5.3, assume that the p th classifier representing the p th sub-image window position recognizes the input image to be the q th person in the gallery. Also assume that $I_{\text{mean-}q}$ is the average of the N images of the q th person that are available in the training set. In the window related to the p th classifier, we find the mutual information between the input illumination invariant sub-image and $I_{\text{mean-}q}$. This will represent how much that sub-image is close to the corresponding sub-image in the training set and is a factor used to generate the weight for that sub-image position.

5.3 Adaptive Weight Generator

In the weighted voting system presented in Figure 5.3, the adaptive weight generator (AWG) block generates weights based on the three factors previously described. The inputs of this block are the canonical image and the mean of training classes as well as the input image and the output of the l classifiers. For the p th sub-image window

position, the difference between mean of gray level values, d_p , represents the quality of illumination, the image entropy, H_p , measures the amount of information in that sub-image and the mutual information, MI_p , is used to measure how similar the recognized sub-image is to the original sample.

In order to generate weights based on these factors, first they are normalized to values between 0 and 1. The weight for the p th sub-image window location is then generated as:

$$W_p = \left| \frac{d_{max} - d_p}{d_{max}} \right| \times H_p^2 \times MI_p \quad (5.7)$$

For $d_p = d_{max}$ the weight W_p will be zero and for an image with good quality or $d_p = 0$ the first factor will be 1. In the proposed method the weight assigned to a sub-image with good illumination and image entropy is 1 just like DVS. In an $m \times m$ WVS we have m^2 different classifiers voting for the subject's identity and we hope that the number of corrupted sub-images is less than the ones in good condition. Our goal is to find the sub-images with bad illumination and information in order to eliminate their effect on the overall performance of the system. Therefore we generated our weights in multiplicative form so that if one of the factors is poor we can reduce the total weight abruptly. In other words we don't want to trust a sub-image's vote if it has one of the factors very low but the two other are high. If we used the weight as the addition of the factors, for example, this sub image with one low factor and two good ones would still have two third of its vote power which is not acceptable. We optimized the weight generation with different powers for each factor and the weight generated by equation 5.7 had the best results.

For each subject the total number of the votes will be the sum of the weights of the classifiers that voted for that subject. In some cases with extremely poor illumination, most sub-images do not have enough useful information, and therefore, result in erroneous recognition. By assuming a minimum acceptable number of votes, V_{th} , in the system, the chance of this false recognition is reduced. The subject with the maximum votes will be recognized as the target person only if it achieves more than V_{th} votes. If none of the subjects reaches this threshold value, then k subjects with the most votes go for a secondary recognition with k classes only. The secondary recognition results, with k much smaller than M , converge much faster and one of the classes is selected as the final result.

5.4 Experimental Results

The performance of the proposed method is evaluated on Yale B, Extended Yale B and CMU-PIE databases and the results have been compared to several existing methods in the literature. The Yale B Face database consists of images from 10 individuals while the Extended Yale B database is expanded with 28 more subjects and contains images of 38 subjects. Each subject has been pictured under 576 viewing conditions which are nine poses and 64 illuminations per pose. Single light sources have been used in different angles for the illumination variations. Since studying illumination invariance is the focus of this work, for both Yale B and Extended Yale B databases, we only use 64 images of each person that are in frontal pose. The databases are divided into five different subsets according to the angle between the light source direction and the camera axis, θ . The

CMU-PIE database images of 68 people under 13 different poses and 43 different illuminations with lights ON and lights OFF, and also 4 different expressions. We used the images with both lights ON and OFF in frontal pose.

For all the images in the databases, faces are cropped into 192×168 images. The position of the face in the images is normalized in a way that the eyes and the lips have almost the same position in all the images.

5.4.1. Face Recognition

As noted above all images in databases are eye aligned and cropped into 192×168 pixels. The illumination invariants are achieved using the method proposed in chapter 4. The natural logarithm of the images is calculated and a 3×3 maximum filter is applied. The output of the filter is then subtracted from the logarithm of the images to get the image invariants. This process is applied to all the images in the databases and the illumination invariant images are obtained. In our test, each of these images is divided into 9 sub-images of 64×56 pixels for DVS with $m = 3$ or 25 sub-images of 38×33 pixels for DVS with $m = 5$. The output results are provided to DVS and WVS systems described in Sections 5.1 and 5.2. PCA is used for feature extraction and eigenfaces are calculated for each sub-image position. SVM with RBF kernel is used for classification. The corresponding weight for each sub-image classifier is generated based on the entropy of the sub-image, the illumination level of the image and the mutual information between the input and selected images.

For the Yale B and Extended Yale B databases, we use images in subset 1 which are taken under small illumination variations for training and other images in four other

subsets for testing. Subset 1 has 7 images per person which make our training gallery and subsets 2 to 5 have 12, 12, 14, and 19 images per person, respectively, forming the test sets. The results of rank one recognition for illumination invariant method, DVS and WVS are presented in Tables 5.1 and 5.2 along with the most common methods for illumination invariant face recognition. As it can be seen in these tables, the proposed method outperforms other available methods and achieves very good recognition rates. The results are further improved by eliminating the light artifacts in the illumination invariants by using local matching methods of DVS and WVS. For the Yale B database, illumination invariant method (II) misclassifies four subjects, two in subset 4 and two in subset 5. The combination of II and DVS misclassifies two subjects, one in subset 4 and one in subset 5. By using II and WVS, only one image in the whole database is classified in a wrong group. This single image is in subset 5 with very poor illumination and has the grayscale levels of 0 to 6, covering less than three percent of the whole 0-255 grayscale level range.

The same experiment is repeated for the CMU-PIE database. The input images have been cropped and resized to 192×168 pixels and then divided to 9 sub-images of 64×56 pixels. For each subject, there are 43 images available where we selected a random group of 3 to 7 images out of these 43 for training set and the rest formed the testing set. We repeated the algorithm and the averages of the results of 200 simulations for this experiment are shown in Table 5.3.

Table 5.1. Recognition rates using different methods for Yale B database (10 subjects).

Method	Subset 2	Subset 3	Subset 4	Subset 5
PCA [13]	98.3	79.2	30.0	15.8
Histogram Equalization[5]	100	89.0	55.1	44.4
GIC[5]	100	88.1	39.9	27.5
Linear subspace [71]	100	100	85.0	-
Cones attached [72]	100	100	91.4	-
Cones cast [72]	100	100	100	-
QI [76]	99.3	61.9	34.1	23.3
QIR [5]	100	100	90.6	78.8
SQI [77]	100	100	96.4	97.9
MQI [86]	100	100	100	98.4
Proposed Illum. Inv + PCA [91]	100	100	98.6	98.9
Proposed Illum. Inv +PCA+DVS [92]	100	100	99.29	99.47
Proposed Illum. Inv +PCA+WVS	100	100	100	99.47

Table 5.2. Recognition rates using the Extended Yale B database (38 subjects).

Method	Subset 2	Subset 3	Subset 4	Subset 5
PCA[13]	90	41	6	3
LDA[14]	100	98	38	5
Proposed Illum. Inv +PCA+SVM	100	99.78	95.44	94.68
Proposed Illum. Inv +PCA+1NN	100	100	96.01	92.44
Proposed Illum. Inv +PCA+DVS	100	100	96.6	95.4
Proposed Illum. Inv +PCA+WVS	100	100	97.91	96.5

Table 5.3. Recognition rates using different methods for CMU-PIE database.

Method	CMU-PIE
PCA [13]	54
QI [76]	84
SQI [77]	99.33
Proposed Illum. Inv + PCA	99.35
Proposed Illum. Inv +PCA+DVS	99.61
Proposed Illum. Inv +PCA+WVS	99.74

5.4.2 Face Verification and ROC Analysis

In this section, performance of the proposed method is tested as a face verification/authentication system. Face verification is to either verify or reject a claimed identity by comparing an unknown input face image of the person with other images belonging to the individual in a database. A similarity metric is needed to verify the claimed identities. If this metric or score is more than a specific threshold, the input identity is approved, otherwise rejected.

In our work for sample i , the total of the votes of the classifiers for class j shows how similar sample i is to the stored images of class j , therefore this value can be used as the similarity score. The chance of a correct acceptance of sample i with claimed identity of class j increases with more classifiers voting sample i to belong to class j . this will result in a similarity matrix that describes how close one test sample is to each class. The size of the similarity matrix is number of test samples by number of classes.

ROC curve [93] for the face verification using the proposed method is presented in Figure 5.4. The results of WVS are compared to DVS and mPCA that use number of votes as the similarity score. “Positive” or “accept” will be when the claimed identity and the result identity match, otherwise, we have “negative” or “reject”. The characteristic of the system depends on the threshold value we use to compare the votes, V_{th} . A large threshold value will set up our verification system in a conservative status where only samples with a large number of votes will be accepted. On the other hand a small value for V_{th} will lead to a more liberal status that more samples will be accepted but a number of false accepts should be tolerated. The verification rates and their corresponding

threshold values of V_{th} are presented in Table 5.4 for different error tolerance strategies. “Min error” is where the total of false accepts and false rejects is minimum. Equal error rate is the point where false accept ratio (FAR) and false reject ratio (FRR) are equal. V_{th} can be selected based on the FAR and FRR than can be tolerated in the desired application.

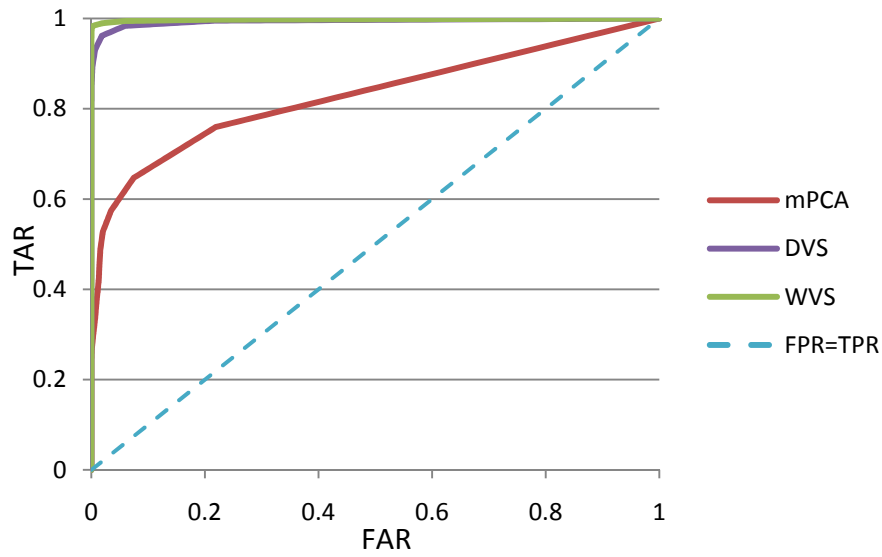


Figure 5.41: ROC curve for face verification where different methods of WVS, DVS and mPCA are compared. 5×5 grids are used for all methods.

Table 5.4: Threshold values and verification rates for different FAR and FRR conditions.

		mPCA		DVS		WVS	
		V_{th}	Verification rate	V_{th}	Verification rate	V_{th}	Verification rate
3x3 grid	Min Error	0.9	68.93	1.9	98.30	1.27	99.01
	EER	0.9	68.93	1.9	98.76	0.75	99.01
	FRR=0.1FAR	0	100	0.9	99.46	0.26	99.54
	FRR=10FAR	1.9	54.47	2.9	98.30	1.25	98.05
	FAR=0.01	7.9	24.15	4.9	77.84	8.99	93.95
	FAR=0.1	6.9	28.62	3.9	98.01	1.47	96.77
	FAR=1	2.9	46.40	1.9	98.67	0.77	99.01
5x5 grid	Min Error	1.9	64.71	4.9	98.09	2.67	98.34
	EER	0.9	75.97	3.9	98.63	1.47	98.55
	FRR=0.1FAR	0	100	1.9	99.25	0.71	99.46
	FRR=10FAR	2.9	57.46	4.9	98.09	2.48	98.34
	FAR=0.01	16.9	24.90	7.9	80.86	20.95	94.82
	FAR=0.1	15.9	26.93	5.9	97.97	2.95	97.89
	FAR=1	8.9	38.28	3.9	98.59	1.58	98.55

5.4.3 Effect of reducing the training samples

Face recognition systems are designed to identify or verify one or more persons from still or videos images, based on an available database. In many applications like driver license for example, the database might contain very few or even only one image per person. This might happen due to difficulties of collecting images or storage limitations [94]. For every face recognition system, it is important to know how the recognition rate will change by reducing the images each person has in the gallery. In this experiment we have studied the effect of reducing the training samples on our illumination invariant method, DVS and WVS. For Yale B and Ext Yale B, a random number of images of each person in the first subset were selected for training and other four subsets were used for testing. Each test has been repeated for 20 times and the average of the recognition results for each subset is presented in Figure 5.5. For CMU-PIE database there are 43 images per person in the database. We selected different number of samples from 1 to 7 for training and the rest were used for testing. Recognition results for the CMU-PIE database using different methods are available in Table 5.5.

Table 5.5. Recognition rate vs. number of training samples from light on and off groups for CMU-PIE database.

		training samples with lights on							
		0	1	2	3	4	5	6	7
training samples with lights off	0	0	57.14	80.05	89.68	94.19	96.58	97.83	98.72
	1	57.73	79.92	89.67	93.98	96.44	97.77	98.65	99.14
	2	79.63	89.46	94.22	96.47	97.73	98.69	99.15	99.39
	3	89.29	93.85	96.46	97.97	98.68	99.18	99.45	99.64
	4	93.84	96.44	97.88	98.67	99.13	99.49	99.66	99.73
	5	96.41	97.87	98.62	99.05	99.40	99.64	99.76	99.78
	6	98.00	98.69	99.19	99.45	99.58	99.77	99.85	99.89
	7	98.58	99.07	99.44	99.66	99.76	99.84	99.88	99.91

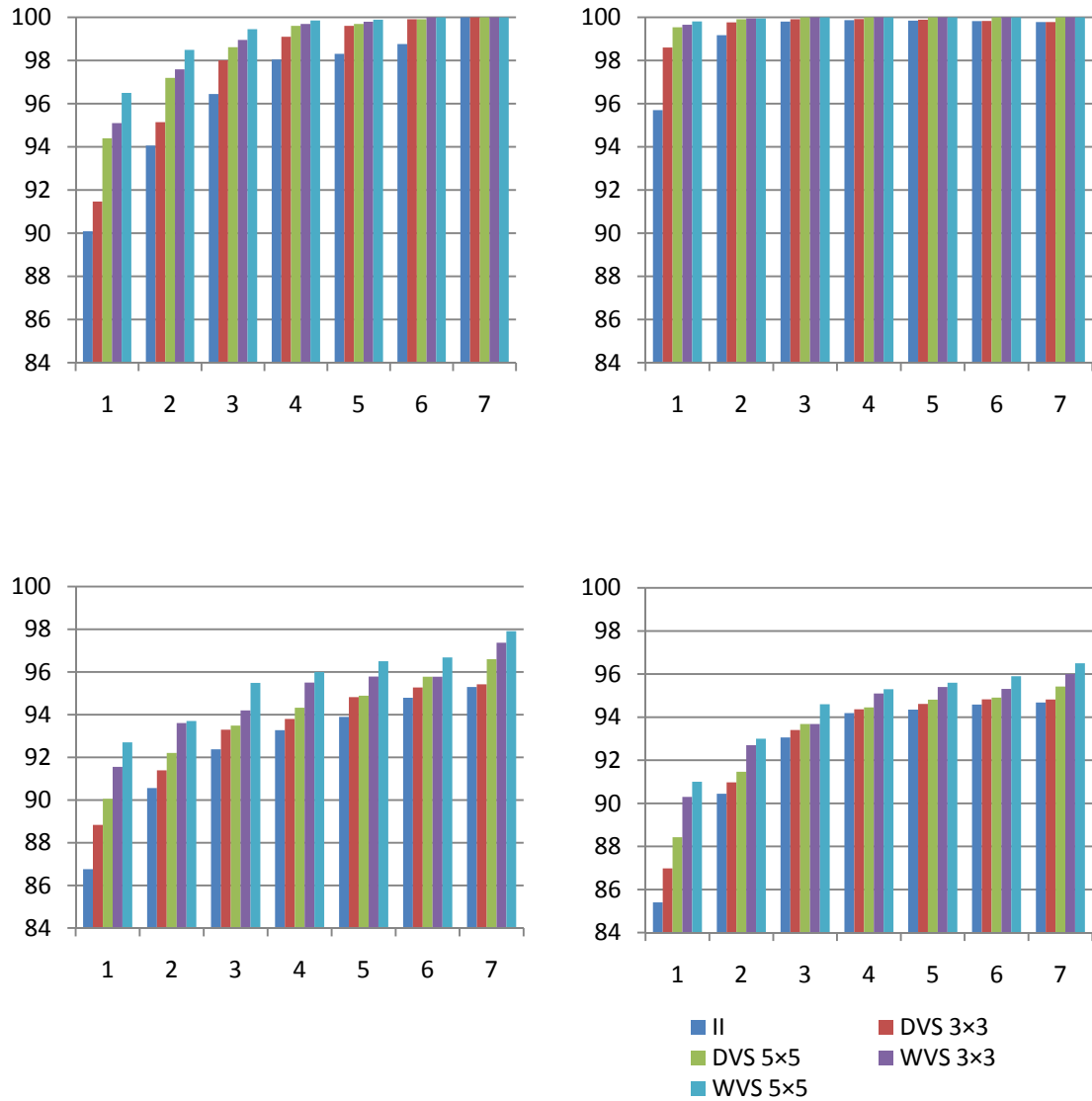


Figure 5.5: Performance of different methods with different number of samples of subset 1 in train set varying from 1 to 7 and different subsets as test probes. (a) subset 2 (b) subset 3 (c) subset 4 (d) subset 5. Extended Yale B database is used with SVM classifiers.

5.4.4 Effect of Partial Image Occlusion

A bonus point of using the local matching methods is that they have better tolerance to image occlusion as well as illumination variations. Since this type of degradation only affects some parts of the images, other sub-images that are not corrupted can do correct identification. In this section the effect of partial occlusion on the recognition rate of the illumination invariant method, DVS and WVS is studied. We wanted to examine the effect of occlusion along with the illumination variation in the images and compare the results with the same images without occlusion. For this purpose the images from the Extended Yale B database are manually occluded with different possible obstacles that might cause occlusion in real applications for example subway surveillance. Sun glasses, two scarves with different textures, a wall, face of another person and back of the head of another person were used as well as black circles of random sizes and location to model occlusion on the input images. The resulting images were fed to the face recognition algorithms and results were compared. A number of samples of images with manual occlusion can be found in Figures 5.6 and 5.7.

The effect of applying sun glasses and scarves occlusions on the methods discussed in this paper is presented in Table 5.6. As it can be seen WVS outperforms all other methods. It is basically because WVS is using information about entropy and mutual information between images and can prevent occluded parts of the image to affect the final decision. Using entropy as a factor for weight generation eliminates the sub-images with low texture and thus black occlusions are omitted. Even if the occluding object contains some texture, the mutual information unit double checks the results with the

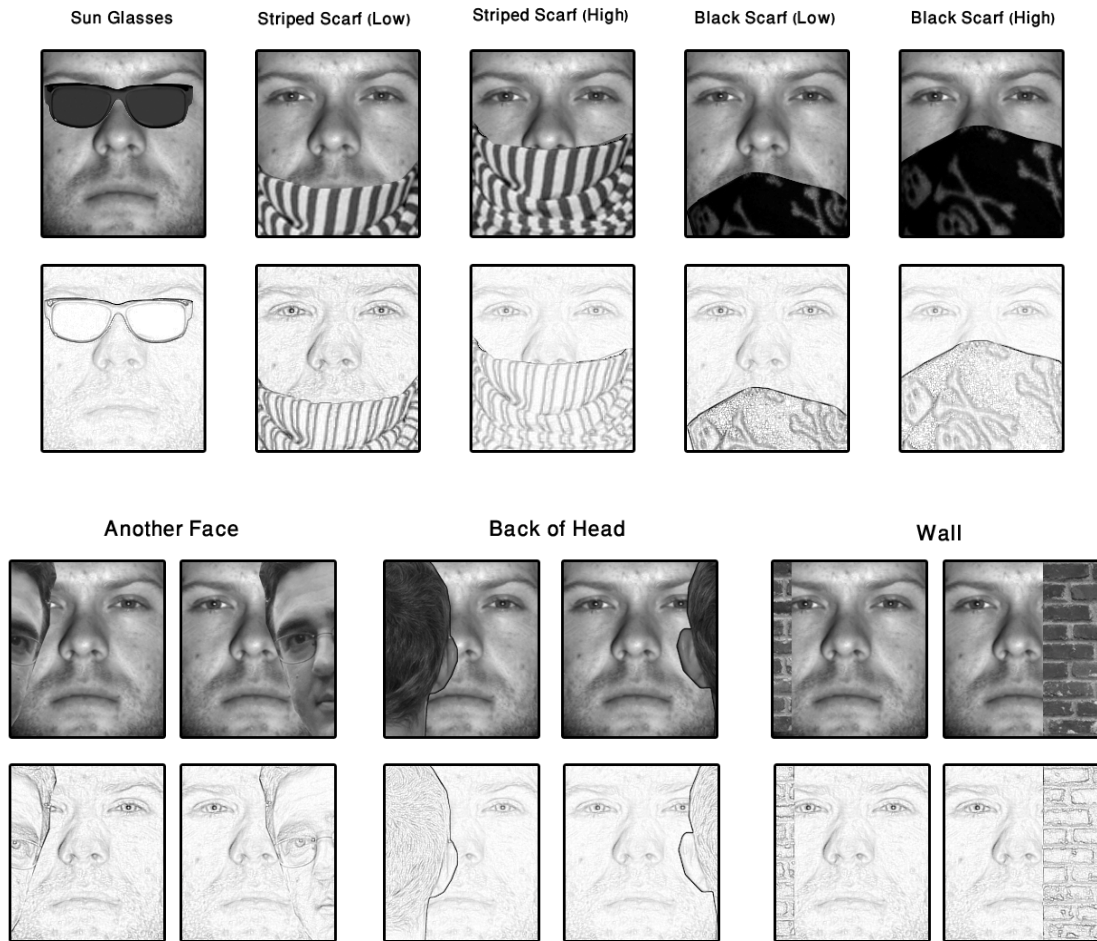


Figure 5.6: Occluded images and their illumination invariants. Input images are occluded with a pair of sunglasses, striped and black scarves, wall, another face and back of another person's head.

original face image and if it is an occluded sub-image its weight will be highly reduced. PCA and mPCA have recognition rates of about 50% less than our proposed Illumination invariant method and voting methods of DVS and WVS.

In figures 5.8(a) to 5.8(d) the effect of occlusion using human obstacles or walls has been investigated. The horizontal axis in these figures presents the number of pixels that the occluding image (*i.e.* the face, the wall or back of the head) has been shifted in to the original face image. The size of original image and final image are both 192×168. As these graphs show while the amount of occlusion is increasing, performance of the local matching methods like DVS and WVS are more robust.

In order to have an estimate of the performance of the system versus the amount of occlusion, black circles of random size and location are added to the faces. This is equal to losing that portion of the image information. This test investigates the performance of the proposed system versus the amount of occlusion. The recognition results for the methods discussed can be found in Table 5.7. The results are average of several iterations with different occlusions and are categorized based on the parameter β . β is the ratio of the occluded pixels to the whole pixels. The results show that the proposed WVS system with entropy and mutual information weights has reached very satisfactory recognition rates.

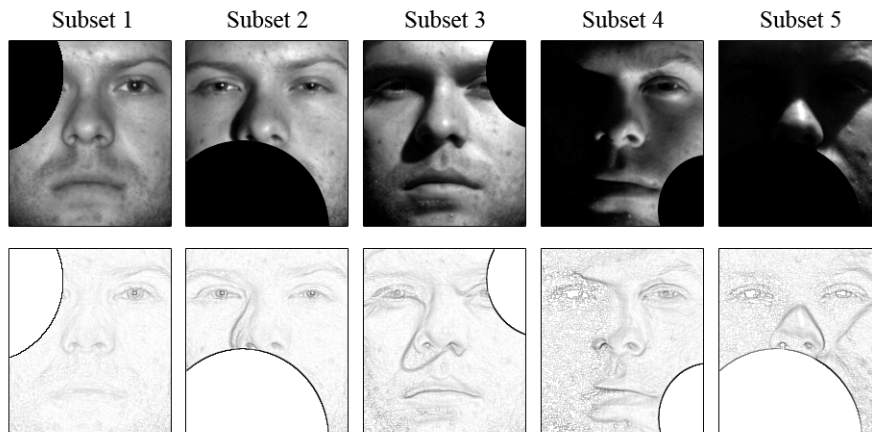


Figure 5.7: Images in different subsets of Yale B database occluded using random circles and their corresponding illumination invariants.

Table 5.6: Recognition rates for images occluded with sunglasses, striped scarf and black scarf.

		PCA	mPCA	II	DVS	WVS
Normal		44.61	46.06	97.35	97.85	98.51
Sun Glasses		31.61	40.85	47.97	79.37	91.42
Scarf 1	L	35.87	41.59	85.17	91.76	96.97
	H	29.33	43.87	70.88	84.96	94.61
Scarf 2	L	30.32	41.22	79.74	88.57	96.14
	H	11.64	24.48	56.92	77.55	92.33

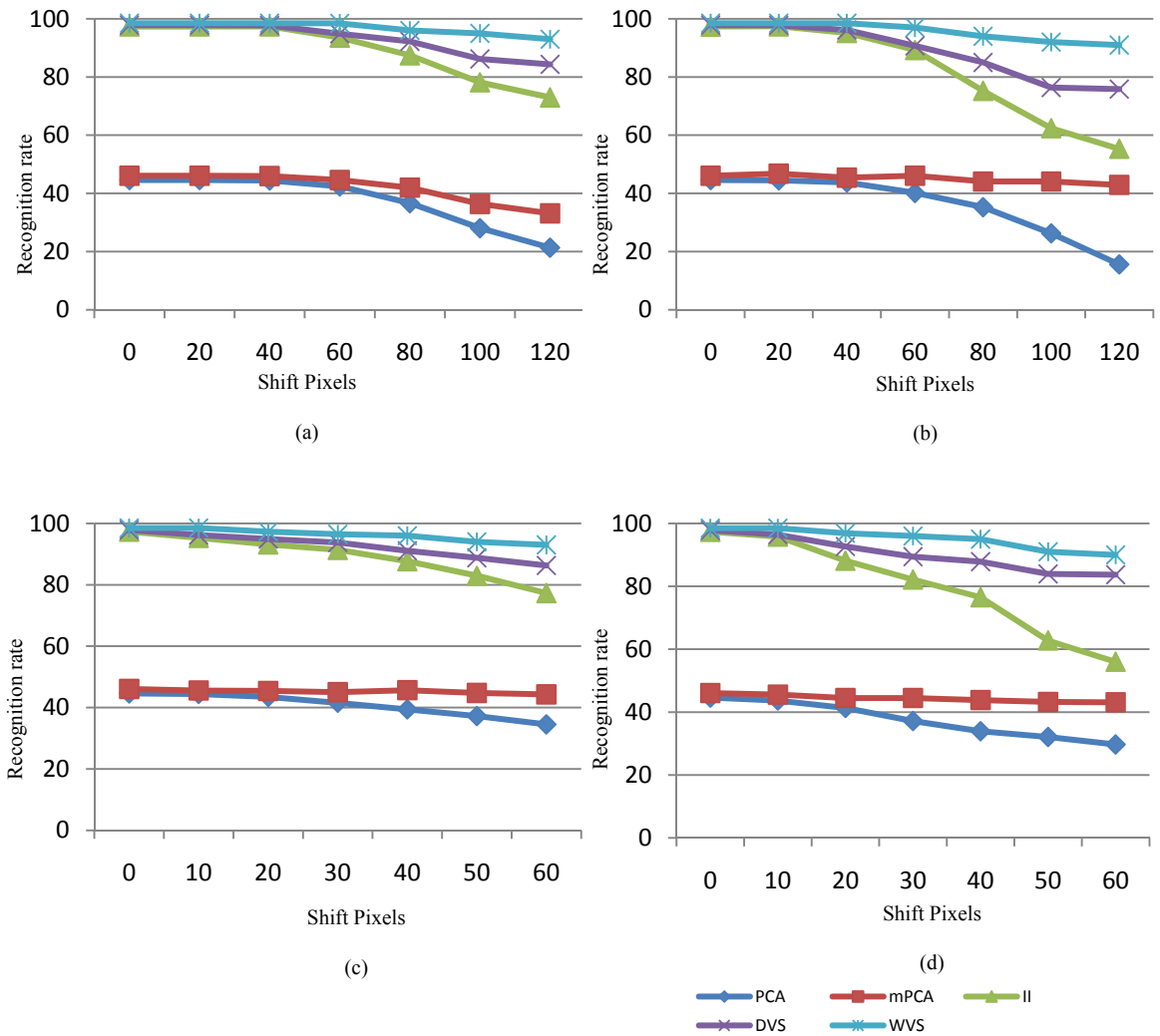


Figure 5.8: Performance of different recognition methods under varying amount of occlusion. (a) back of head (b) another face (c) wall overlapping from left and right (d) wall overlapping from top and bottom.

Table 5.7: Effect of occlusion on recognition results using different methods for Ext Yale B database (38 subjects).

		$\beta < 0.1$	$0.1 < \beta < 0.2$	$0.2 < \beta < 0.3$	$0.3 < \beta < 0.4$	$0.4 < \beta < 0.5$
Subset 1	PCA	98.73	79.40	49.94	29.92	18.82
	II	99.85	96.15	91.67	80.80	68.00
	mPCA	100.00	100.00	100.00	99.96	77.63
	II+DVS	100.00	100.00	100.00	99.92	99.68
	II+WVS	100.00	100.00	100.00	100.00	100.00
Subset 2	PCA	90.50	60.53	39.11	24.98	15.61
	II	99.30	91.98	86.86	74.82	59.72
	mPCA	97.02	94.15	89.82	80.90	49.96
	II+DVS	99.94	98.62	97.35	94.80	84.58
	II+WVS	100.00	100.00	99.53	99.21	97.85
Subset 3	PCA	57.14	37.17	24.84	16.26	10.37
	II	97.77	86.73	82.44	66.66	53.66
	mPCA	64.81	49.48	36.32	23.01	11.61
	II+DVS	99.90	99.30	97.93	94.07	81.98
	II+WVS	100.00	99.96	99.82	98.30	97.11
Subset 4	PCA	11.49	8.20	6.67	5.13	4.56
	II	89.37	70.56	66.79	54.03	41.38
	mPCA	9.67	6.42	5.15	3.88	3.20
	II+DVS	92.76	81.30	76.96	67.41	53.17
	II+WVS	97.90	97.73	95.20	92.12	89.74
Subset 5	PCA	3.07	2.91	2.89	2.83	2.82
	II	90.59	75.19	70.23	57.97	45.31
	mPCA	2.80	2.75	2.71	2.69	2.68
	II+DVS	92.69	82.49	78.20	68.52	57.38
	II+WVS	96.47	96.19	92.73	89.97	86.23

5.4.5 Effect of a Larger Database

Performance of a face recognition system is reliable and promising when tested on a large group of subjects. Having more subjects in the dataset causes more inter-class similarities and can challenge the feature extraction and classification steps more severely. In this experiment we combined the Extended Yale B and CMU-PIE databases to make a larger database with 106 subjects. For a fair comparison we choose 45 images per person which are subsets 1 to 4 for the Extended Yale B database and all the images of each person for CMU-PIE. For each individual, 7 images with good illumination are selected as train set and 38 images are in the test set. Due to unequal image size for the cropped faces in two databases, all the images are resized to 112×128 pixels and are aligned in a way that position of the eyes and mouth is almost the same for all 106 subjects. The total number of images in train set is 739 images aligned with 4019 images in the test set. Large number of train and test samples results in larger covariance matrixes and more eigenvectors, thus training process requires more time and memory resources. Recognition results for the combination of extended Yale B and CMU-PIE databases are provided in Table 5.8.

Table 5.8: Recognition results for combination of Ext. Yale B and CMU-PIE databases.

	Recognition Rates (%)	
	m = 3	m = 5
PCA	42.74	
mPCA	47.67	50.51
Illum. Inv+ PCA	97.83	
Illum. Inv+ DVS	98.36	98.43
Illum. Inv+ WVS	98.91	99.03

Illumination variations in the CMU-PIE database are not as severe as the Extended Yale B database; therefore, the distribution of the classification errors is not equal. For example for DVS with $m = 5$, total number of errors is 63 out of 4019 test samples resulting in 98.43% recognition accuracy. Among these 63 misclassified samples, 52 out of 1437 test images are from Yale B and 11 out of 2582 samples are from the CMU-PIE database, resulting in 96.38% and 99.57% correct recognition for Yale B and CMU-PIE databases, respectively.

5.5 Conclusion

In this chapter we proposed a novel method for illumination invariant face recognition based on the reflectance-luminance model combined with local matching using our proposed weighted voting method to eliminate the artifacts in the retinex images. We used the maximum filter as a simple filter with low computation costs to extract the illumination invariants. We used both global and local image information which yielded better recognition results. Our proposed weighted voting scheme performs local matching with a vector of sub-image weights based on three factors: sub-image illumination level, image entropy showing the amount of information in that sub-image, and also the mutual information between corresponding sub-images of the selected subject and the original sample. Generating the weights based on these three factors also grants robustness to the face recognition system under image occlusion. Different occlusions were added to the images and the performance of the proposed method was compared to similar face recognition methods. The proposed method does not need any

kind of parameter selection to generate the illumination invariants. It does not need any prior information about face models or the illumination condition and needs only one single image unlike most of the methods found in the literature. Experimental results show that the proposed method had high recognition rates of 99.82% in the Yale B and 99.81% in CMU-PIE databases.

Chapter 6

Concluding Remarks

6.1 Summary of Contributions

Significant research is done in this work with an aim of designing a reliable automated face recognition system which is robust under varying conditions of noise level, illumination and occlusion. This work's original contribution is in unique combination of illumination invariant feature extraction and weighted local matching to achieve promising recognition results in presence of both illumination variations and occlusions.

A new method for illumination invariant feature extraction is proposed and the results are compared to similar algorithms. The proposed technique is computationally efficient and does not require any prior information about the face model or illumination. Unlike

many methods in the literature, our method does not need many training samples and can be applied on each single image to obtain the illumination invariants.

A classification with a weighted voting scheme is proposed to enhance the performance under illumination variations and also cancel occlusions. The proposed method used mutual information and entropy of the images to generate different weights for a group of ensemble classifiers based on the input image quality. The method yields outstanding results by reducing the effect of both illumination and occlusion variations in the input face images.

6.2 Conclusion

In this work we proposed a novel method for illumination invariant face recognition based on the reflectance-luminance model combined with local matching using our proposed weighted voting method to eliminate the artifacts in the retinex images. We tested 37 different linear or nonlinear, highpass and lowpass filters to achieve the reflectance part of the image which is illumination invariant. The maximum filter, as a simple filter with low computation complexity, showed the best results for extracting the illumination invariants. We showed that the illumination invariants gained by this method result in better recognition comparing to other methods like QI, SQI or image enhancement methods. The proposed method does not need any prior information about the face shape or illumination and can be applied on each image separately. Unlike most available methods, our method does not need multiple images in training stage and does

not need any kind of parameter selection to generate the illumination invariants. Several experiments were performed on Yale B, extended Yale B and CMU-PIE databases.

We also proposed a weighted voting system to further improve the illumination robustness. Some parts of the image are likely to affect the recognition result due to their bad illumination, occlusion, noise, or lack of distinctive information. The rule of the proposed weighted voting scheme is to find these parts of the image and then reduce their effect based on some predefined factors. We used gray scale mean, image entropy and mutual information as the selective factors to generate a weight for each part of the image. We used both global and local image information which yielded better recognition results. Combining these two will give more robustness to shift and occlusion because in many cases the global image can correct the sub-image decisions if the faces are not perfectly aligned. On the other hand the sub-images have better decisions when there is occlusion or illumination variation. We tested different values for the weight of the global classifier, W_0 , and found out that the best performance is achieved when W_0 is equal to m , where m is the number of classifiers in each column or row. Different occlusions were added to the images and the performance of the proposed method was compared to similar face recognition methods. The combination of the illumination invariant method and the voting scheme achieves promising results in presence of illumination, occlusion and misalignment.

Performance of different moment invariants for feature extraction in face recognition was also studied in this work. The moment invariants used in this study were HMI, BMI, RMI, ZMI, TZMI, PZMI, NZMI, NPZMI and LMI. the best recognition rate of 95% for

the AT &T database was achieved by PZMI of order 1 to 7 with 5000 epochs training. We used second order moments for face localization. The performance of the system can be optimized by using proper value for the scaling factor, ρ , and also by choosing proper orders of the moment invariants. It was concluded that high order PZMIs contain useful information, therefore lead to the best results. BMIs were the fastest to be computed but had very poor recognition accuracy for face images while calculating ZMIs was the most time consuming. In another experiment performance of pseudo Zernike moment invariants for feature extraction along with different classifiers was studied. k -NN, SVM and HMM algorithms were used for classification where the best recognition rate of 91% was achieved by SVM with RBF kernel.

Misalignment and noise might exist in the output of the face detector in a face recognition system and could possibly affect the performance of the system. In another experiment, performance of eigenfaces, Fisherfaces, Zernike moment invariant and Legendre moment invariants were tested on images with rotation, translation, Gaussian noise and salt and pepper noise. For both of FERET and Yale B databases, statistical methods, *i.e.* PCA and FLD showed better results in presence of noise while moment invariant based methods performed better where there were geometrical misalignments. PCA had the best results among the four methods for noise stability, but was less steady when images were rotated or shifted. PCA had the recognition rate of 65% and 68% in presence of the Gaussian noise with $\sigma = 0.5$ and the salt and pepper noise with noise density of 0.5, respectively. Zernike moments showed to have the best performance for the shift test as well as the rotation test with the recognition rate of 61% for FERET

database and 67% for the Yale B database for the maximum shift of 9 pixels in both x and y directions. The recognition rates for PCA and FLD dropped quickly after shifting 3 or 4 pixels.

6.3 Future Work

The main goal of this work was on illumination invariant face recognition. In many real applications pose is also a factor that challenges the designed system just like illumination. Investigating pose invariance as well as illumination and occlusion robustness will lead to a good performance for the face recognition system.

In many security and surveillance applications real time face recognition is an aim. A fast and reliable recognition system can be produced by designing in hardware. FPGA implementation of the proposed method will generate a promising real time face recognition system.

All the image databases in this work contain grayscale images. Compared to grayscale levels, color of the face changes more severely with illumination variation. Therefore we used only grayscale images in our research, but by discarding the color data we are apparently losing some information. As a result, a study on illumination invariant feature extraction for color images is necessary.

And finally, the face recognition system designed in this work considers only still images while in many surveillance applications the input data is a video sequence. Extracting illumination invariant features from a video sequence is a challenge for the face recognition system.

References

- [1] K.W. Bowyer, Face recognition technology: security versus privacy, *Technology and Society Magazine, IEEE*, Vol 23-1 (2004), 9-19 .
- [2] W. Zhao, R. Chellappa, P.J. Phillips, A. Rosenfeld, Face recognition: A literature survey, *ACM Comput.Surv.* Vol 35 (2003) 399-458.
- [3] C.A. Nelson, The Development and Neural Bases of Face Recognition, *Infant and Child Development*, vol 10, (2001), 3–18.
- [4] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, second ed., Prentice Hall, 1992.
- [5] S. Shan, W. Gao, B. Cao, D. Zhao, Illumination Normalization for Robust Face Recognition Against Varying Lighting Conditions, *IEEE International Workshop AMFG*, (2003) 157-164.
- [6] C.P. Papageorgiou, M. Oren, T. Poggio, A general framework for object detection, *Sixth International Conference on Computer Vision*, (1998) 555-562.

-
- [7] E. Osuna, R. Freund, F. Girosit, Training support vector machines: an application to face detection, Proceedings of IEEE Computer Society Conference on. Computer Vision and Pattern Recognition (1997) 130-136.
- [8] S. Romdhani, P. Torr, B. Scholkopf, A. Blake, Computationally efficient face detection, Proceedings of eighth IEEE International Conference on Computer Vision, ICCV2001. (2001) 695-700 vol.2.
- [9] H. Schneiderman, T. Kanade, A statistical method for 3D object detection applied to faces and cars, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, CVPR2000 (2000) 746-751 vol.1.
- [10] H.A. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, CVPR'96 (1996) 203-208.
- [11] F. Fleuret, D. Geman, Coarse-to-Fine Face Detection, International Journal of Computer Vision, 41 (2001) 85-107.
- [12] P. Viola, M.J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision, 57 (2004) 137-154.
- [13] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, CVPR'91, (1991) 586-591.
- [14] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19 (1997) 711-720.
- [15] C. Liu, H. Wechsler, Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition, International Conference on Audio and Video Based Biometric Person Authentication. (1999) 22-24.
- [16] J. Kim, J. Choi, J. Yi, M. Turk, Effective representation using ICA for face recognition robust to local distortion and partial occlusion, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005) 1977-1981.

-
- [17] B. Schölkopf, A. Smola, K. Müller, Nonlinear Component Analysis as a Kernel Eigenvalue Problem, *Neural Computation*, 10 (1998) 1299-1319.
- [18] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, K.R. Mullers, Fisher discriminant analysis with kernels, *Proceedings of the 1999 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing IX*, (1999) 41-48.
- [19] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach, *Neural Computation*, 12 (2000) 2385-2404.
- [20] F.R. Bach, M.I. Jordan, Kernel Independent Component Analysis, *Journal of Machine Learning Research*, 3 (2002) 1-48.
- [21] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, Face recognition using kernel direct discriminant analysis algorithms, *IEEE Transactions on Neural Networks*, 14 (2003) 117-126.
- [22] M.H. Yang, Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using kernel methods, *Proceedings. Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, (2002) 215-220.
- [23] J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 (2005) 230-244.
- [24] L. Wiskott, J. Fellous, N. Krüger, C. Malsburg, Face Recognition by Elastic Bunch Graph Matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (1997) 775-779.
- [25] M.K. Hu, Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory*, 8 (1962) 179-187.
- [26] S.S. Reddi, Radial and Angular Moment Invariants for Image Identification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3 (1981) 240-242.

-
- [27] B. Bamieh, R. De Figueiredo, A general moment-invariants/attributed-graph method for three-dimensional object recognition from a single image, *IEEE Journal of Robotics and Automation*, 2 (1986) 31-41.
- [28] C. Teh, R.T. Chin, On image analysis by the methods of moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10 (1988) 496-513.
- [29] L. Kotoulas, I. Andreadis, Real-time computation of Zernike moments, *IEEE Transactions on Circuits and Systems for Video Technology*, 15 (2005) 801-809.
- [30] A. Wallin, O. Kubler, Complete sets of complex Zernike moment invariants and the role of the pseudo invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 (1995) 1106-1110.
- [31] M.R. Teague, Image analysis via the general theory of moments, *Journal of the Optical Society of America* 70 (1980) 920-930.
- [32] K.M. Hosny, Exact Legendre moment computation for gray level images, *Pattern Recognition*. 40 (2007) 3597-3605.
- [33] J. Haddadnia, M. Ahmadi, K. Faez, An efficient feature extraction method with pseudo-Zernike moment in RBF neural network-based human face recognition system, *EURASIP Journal on Applied Signal Processing*, (2003) 890-901.
- [34] A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, Human Face Recognition Using Different Moment Invariants: A Comparative Study, 2008 Congress on Image and Signal Processing, CISP '08, 3 (2008) 661-666.
- [35] H.R. Kanan, K. Faez, Y. Gao, Face recognition using adaptively weighted patch PZM array from a single exemplar image per person, *Pattern Recognition*, 41 (2008) 3799-3812.
- [36] J.P. Jones and L.A. Palmer, An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex, *Journal of Neurophysiology*, 58-6,(1987), 1233-1258.

-
- [37] C.J. Lee; S.D. Wang, Fingerprint feature extraction using Gabor filters, *Electronics Letters*, 35-4, (1999) 288-290.
- [38] R. Buse, Z.Q. Liu, Feature extraction and analysis of handwritten words in grey-scale images using Gabor filters, *Image Processing, Proceedings of IEEE International Conference on Image Processing, ICIP'94*, 1 (1994) 164-168.
- [39] J. Daugman, Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression, *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 36, 7, (1988) 1169–1179.
- [40] M. Lades, J.C. Vorbriiggen, J. Buhmann, J. Lange, C. Malsburg, R.P. Wiirtz, and Wolfgang Konen, Distortion Invariant Object Recognition in the Dynamic Link Architecture, *IEEE Transactions on Computers*, 42, 3, (1993) 300-311.
- [41] C.J. Liu, H. Wechsler, Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, *IEEE Transactions on Image Processing*, 11, (2002) 467–476.
- [42] L. Shen, L. Bai, M. Fairhurst, Gabor wavelets and General Discriminant Analysis for face identification and verification. *Image and Vision Computing*, 25, (2007) 553–563.
- [43] L. Shen, L. Bai, A review on Gabor wavelets for face recognition. *Pattern Analysis and Applications*, 9, (2006) 273–292.
- [44] J. Ruiz-del Solar, J. Quinteros, Illumination compensation and normalization in eigenspace-based face recognition: a comparative study of different pre-processing approaches, *Pattern Recognition Letters*, 29 (2008) 1966-1979.
- [45] A. Tefas, C. Kotropoulos, I. Pitas, Variants of dynamic link architecture based on mathematical morphology for frontal face authentication, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (1998) 23-25
- [46] N. Cristianini, J. Shawe-Taylor, *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.

-
- [47] V. Kecman, Learning and Soft Computing -Support Vector Machines, Neural Networks, Fuzzy Logic Systems, The MIT Press, Cambridge, MA, 2001.
- [48] A. Nefian and M. Hayes. An embedded HMM-based approach for face detection and recognition. IEEE International Conference on Acoustics, Speech and Signal Processing, Phoenix, AZ, (1999) 3553–3556.
- [49] <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [50] <http://www.itl.nist.gov/iad/humanid/feret/>.
- [51] <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>
- [52] <http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>
- [53] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression database, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (2003) 1615–1618.
- [54] S.O. Belkasim, M. Shridhar, M. Ahmadi, Pattern recognition with moment invariants: A comparative study and new results, Pattern Recognition, 24 (1991) 1117-1138.
- [55] J. Wang and T. Tan, A new face detection method based on shape information, Pattern Recognition Letters, 21, 6-7, pp. 463–471, 2000.
- [56] E. Hjelmås and B. K. Low, Face detection: a survey, Computer Vision and Image Understanding, 83, 3 (2001) 236– 274.
- [57] M.Y. Chen, A. Kundu, S.N. Srihari, Variable duration hidden Markov and morphological segmentation for handwritten word recognition, IEEE Transactions on Image Processing, 4, 12 (1995) 1675-1688.
- [58] Rabiner 1989. L.R. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE, 77, 2 (1989), 257–285.
- [59] A. Nabatchian, I. Makaremi, E. Abdel-Raheem, M. Ahmadi, Pseudo-Zernike Moment Invariants for Recognition of Faces Using Different Classifiers in FERET Database, Third

-
- International Conference on Convergence and Hybrid Information Technology, ICCIT '08, Busan, Korea, (2008).
- [60] J. Flusser, Moment Invariants in Image Analysis, Proceedings of World Academy of Science, Engineering and Technology. 11 (2006) 196-201.
- [61] W. Wong, W. Siu, K. Lam, Generation of moment invariants and their uses for character recognition, Pattern Recognition Letters, 16 (1995) 115-123.
- [62] Y. Li, Reforming the theory of invariant moments for pattern recognition, Pattern Recognition, 25 (1992) 723-730.
- [63] A. Wallin, 3D surface detection and pattern recognition for surgical planning in orthopedics, Proceedings of. Computer Assisted Radiology, CAR '91, Berlin: Springer-Verlag, (1991) 323-328.
- [64] L.G. Shapiro, G.C. Stockman, Computer Vision, Prentice-Hall, 2001.
- [65] J. Ohta, Smart CMOS Image Sensors and Applications, CRC Press, 2008.
- [66] J. Nakamura, Image Sensors and Signal Processing for Digital Still Cameras. CRC Press, 2005
- [67] Z. Xuan, J. Kittler, K. Messer, Illumination Invariant Face Recognition: A Survey. Proceedings of IEEE Conference on Biometrics: Theory, Applications, and Systems, BTAS (2007) 1-8.
- [68] P.W. Hallinan, A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions. Proceedings of IEEE Computer Society Conference on. Computer Vision and Pattern Recognition, CVPR'94 (1994) 995-999.
- [69] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23 (2001) 643-660.

-
- [70] P.N. Belhumeur, D.J. Kriegman, What is the Set of Images of an Object Under all Possible Lighting Conditions? Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'96 (1996) 270-277.
- [71] R. Basri, D.W. Jacobs, Lambertian Reflectance and Linear Subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (2003) 218-233.
- [72] J.C. Lee, J. Ho, D. Kriegman, Nine points of light: acquiring subspaces for face recognition under variable lighting, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1 (2001) 519-526.
- [73] Y. Gao, M.K.H. Leung, Face Recognition using Line Edge Map, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (2002) 764-779.
- [74] W. Zhao, R. Chellappa, Illumination-Insensitive Face Recognition using Symmetric Shape-from-Shading, Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1 (2000) 286-293.
- [75] T. Sim, J. Kanade, Combining Models and Exemplars for Face Recognition: An Illuminating Example, Workshop on Models vs. Exemplars, CVPR (2001).
- [76] A. Shashua, T. Riklin-Raviv, The Quotient Image: Class-Based Re-Rendering and Recognition with Varying Illuminations, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23 (2001) 129-139.
- [77] H. Wang, S.Z. Li, Y. Wang, Face Recognition Under Varying Lighting Conditions using Self Quotient Image, Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, (2004) 819-824.
- [78] W. Chen, M.J. Er, S. Wu, Illumination Compensation and Normalization for Robust Face Recognition using Discrete Cosine Transform in Logarithm Domain, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 36 (2006) 458-466
- [79] R. Gottumukkal, V.K. Asari, An improved face recognition technique based on modular PCA approach, Pattern Recognition Letters, 25 (2004) 429-436.

-
- [80] J.R. Price, T.F. Gee, Face recognition using direct, weighted linear discriminant analysis and modular subspaces, *Pattern Recognition*, 38 (2005) 209-219.
- [81] B. Heisele, P. Ho, J. Wu, T. Poggio, Face recognition: component-based versus global approaches, *Computer Vision and Image Understanding*, 91 (2003) 6-21.
- [82] K. Tan, S. Chen, Adaptively weighted sub-pattern PCA for face recognition, *Neurocomputing*, 64 (2005) 505-511.
- [83] B.K.P. Horn, *Robot Vision*, MIT Press, Cambridge, 1986.
- [84] Y. Adini, Y. Moses, S. Ullman, Face Recognition: The Problem of Compensating for Changes in Illumination Direction, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (1997) 721-732.
- [85] B. Saether, H. Rueslatten, A. Gronlie, Application of the Hough transform for automated interpretation of linear features in imageries, *IEEE International Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation*, 2 (1994) 847-850.
- [86] Y. Zhang, J. Tian, X. He, X. Yang, MQI Based Face Recognition under Uneven Illumination. *Proceedings of International Conference on Advances in Biometrics, ICB*, (2007) 290-298.
- [87] C.E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal*, 27 (1948) 379-423 623-656.
- [88] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Chapter 15, 2nd ed., McGraw Hill, 1991.
- [89] G.A. Korn, T.M. Korn, *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*, 2nd ed., Dover Publications, New York, 2000.
- [90] C. Arndt, *Information Measures: Information and its description in Science and Engineering*, Springer, Berlin, 2001.

-
- [91] A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, An efficient method for face recognition under illumination variations, International Conference on High Performance Computing and Simulation, HPCS2010, June 28 -July, Caen, France (2010).
- [92] A. Nabatchian, E. Abdel-Raheem, M. Ahmadi, A Weighted Voting Scheme for Recognition of Faces with Illumination Variation, 11th International Conference on Control, Automation, Robotics and Vision, ICARCV2010, Singapore, 7-10 December, (2010).
- [93] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters, 27, 8, (2006), 861-874.
- [94] X.Y. Tan, S.C. Chen, Z.H. Zhou, Face Recognition from a Single Image per Person: A Survey. Pattern Recognition, 39 (2006) 1725-45.
- [95] C.J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery, 2 (1998)121-167.
- [96] B.V. Dasarathy, Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos, CA, (1991).
- [97] G. Shakhnarovich, T. Darrell, P. Indyk, Nearest-Neighbor Methods in Learning and Vision Theory and Practice, MIT Press, Cambridge, 2006.
- [98] S. Lawrence, C.L. Giles, A.h. Chung Tsoi, A.D. Back, Face recognition: a convolutional neural-network approach, IEEE Transactions on Neural Networks, 8 (1997) 98-113.
- [99] S.K. Singh, M. Vatsa, R. Singh, D.S. Chauhan, A comparison of face recognition algorithms neural network based & line based approaches, 2002 IEEE International Conference on Systems, Man and Cybernetics,.6 (2002)
- [100] K.C. Kwak, W. Pedrycz, Face recognition using fuzzy Integral and wavelet decomposition method, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 34 (2004) 1666-1675.

- [101] X. Xie, K.M. Lam, Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image, *IEEE Transactions on Image Processing*, 15 (2006) 2481-2492.

- [102] V. Perlibakas, Distance measures for PCA-based face recognition, *Pattern Recognition Letters*, 25 (2004) 711-724.

VITA AUCTORIS

NAME: Amirhosein Nabatchian

PLACE OF BIRTH: Tehran, Iran

YEAR OF BIRTH: 1981

EDUCATION Iran University of Science and Technology, Tehran, Iran
1999-2003 B.Sc. in Electrical Engineering

Iran University of Science and Technology, Tehran, Iran
2003-2006 M.Sc. in Electrical Engineering

University of Windsor, Windsor, Ontario
2006-2010 Ph.D. in Electrical Engineering