

**THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY**

**A COMPARATIVE STUDY ON THE EFFECTIVENESS OF
GABOR WAVELET ON COMMON FACE RECOGNITION
METHODS**

Master's Thesis

Enji Issa Zainalabdin

Advisor: Prof. Dr. Adem Karahoca

ISTANBUL, 2017

**THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
INFORMATION TECHNOLOGIES**

**A COMPARATIVE STUDY ON THE
EFFECTIVENESS OF GABOR WAVELET ON
COMMON FACE RECOGNITION METHODS**

Master's Thesis

Enji Issa Zainalabdin

Advisor: Prof. Dr. Adem Karahoca

ISTANBUL, 2017

**THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
INFORMATION TECHNOLOGIES**

Thesis Title: A COMPARATIVE STUDY ON THE EFFECTIVENESS OF GABOR WAVELET ON COMMON FACE RECOGNITION METHODS

Student Name: Enji Issa Zainalabdin

Date of Thesis Defense: 06.01.2017

This Thesis has been approved by the Institute of The Graduate School Natural and Applied Sciences.

Prof. Dr. Nafiz ARICA
The Director of Graduate School

This is to certify that we have read this thesis and we find it fully adequate in scope, quality and content, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Mehmet Alper TUNGA
Program Coordinator

Examining Committee Members

Signature

Thesis Supervisor
Prof. Dr. Adem KARAHOCA

.....

Member
Prof. Dr. İbrahim PINAR

.....

Member
Asst. Prof. Dr. Dilek KARAHOCA

.....

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my Supervisor Prof. Dr. ADEM KARAHOCA for guiding me throughout my pursuit of master degree. His useful notes and comments have made this work possible. Also I would like to thank Asst. Prof. Dr. Dilek Karahoca for her helpful comments.

Also I would like to thank my family for the continuous support they gave me during my study. Their faith in me and nonstop encouragement is what made me keep going and never give up. Without my family my success has no meaning.

Istanbul, 2017

Enji Issa ZAINALABDIN

ABSTRACT

A COMPARATIVE STUDY ON THE EFFECTIVENESS OF GABOR WAVELET ON COMMON FACE RECOGNITION METHODS

ENJI ISSA ZAINALABDIN

MSc in Information Technology

Thesis Advisor: Prof. Dr. Adem Karahoca

January 2017, 53 pages

Analyzing an image and being able to recognize some or all of the parts in it still remains one of the most challenging tasks that a computer can achieve. Face recognition is a very important problem in the field of Computer Vision. A human brain is very powerful and is able to recognize any faces within seconds even if they have dramatically changed due to aging, disguise, makeup, facial hair, etc, unfortunately a computer cannot achieve the same thing as fast and easily. What makes face recognition a challenging problem is the fact that until now there is no single method that offers a successful solution to all of the different encountered situations. Another reason that makes face recognition a difficult problem is the variability of different facial appearances and background images. Face recognition is relatively a new concept yet it has developed very fast and became a popular topic both in the areas of research and commercials. The availability of feasible technologies is a key reason behind the development of face recognition. Also the demand for the use of a face recognition system in public areas for security reasons has also accelerated the development process.

In this thesis, several popular face recognition methods have been introduced and discussed with and without using Gabor Wavelet filters. The introduced methods are Principle Component Analysis, Linear Discriminant Analysis, Kernel Principle Component Analysis, Kernel Linear Discriminant Analysis, Back Propagation Neural Network, and Hidden Markov Model. To show the effectiveness of using Gabor filters, test results from both experiments have been compared. The comparison shows that the recognition rate of the used methods have increased with the use of Gabor Wavelet filters.

Keywords: Face Recognition, PCA, LDA, BPNN, HMM.

ÖZET

GABOR WAVELET'İN ORTAK YÜZ TANIMLANMASI YÖNTEMLERİ ÜZERİNE ETKİNLİĞİNE İLİŞKİN KARŞILAŞTIRMALI İNCELEME

ENJI ISSA ZAINALABDIN

Bilgi Teknolojileri Yüksek Lisans Programı

Tez Danışmanı: Prof. Dr. Adem Karahoca

Ocak 2017, 53 sayfa

Bir görüntüyü analiz etmek ve içindeki parçaların bir kısmını veya tamamını tanıyabilmek, halen bir bilgisayarın gerçekleştirmesini istemek için en zor ve zorlu görevlerden biri olmayı sürdürüyor. Yüz tanıma, bilgisayar görüşünün önemli sorunlarından biridir. Bir insan beyni çok güçlüdür ve yaşlanma, kılık değiştirme, makyaj, yüz kılları vb. nedeniyle dramatik bir şekilde değişmiş olsa bile saniyeler içinde herhangi bir yüzü tanıyabilmektedir, maalesef bir bilgisayar aynı şeyi hızlı ve kolay bir şekilde başaramıyor. Yüz tanımayı zorlu bir sorun haline getiren şey şu ana kadar karşılaşılan tüm durumlara başarılı bir çözüm sunan tek bir yöntem bulunmamasıdır. Yüz tanımayı zorlu bir sorun haline getiren diğer bir neden, farklı yüz görünüşlerinin ve arka plan görüntülerinin değişkenliğidir. Yüz tanıma göreceli olarak yeni bir konsept, ancak çok hızlı gelişti ve hem araştırma hem de reklam alanlarında popüler bir konu haline geldi. Uygulanabilir teknolojilerin bulunması, yüz tanımanın geliştirilmesinin arkasındaki kilit nedenlerden biridir. Ayrıca yüz tanıma sisteminin kullanılması talebi kamusal alanlarda güvenlik nedenleriyle gelişim sürecini hızlandırmıştır.

Bu tezde, çeşitli popüler yüz tanıma yöntemleri tanıtıldı ve Gabor Wavelet filtrelerini kullanmadan ve tartışmadan tartıştı. Giriş yöntemi Prensip Bileşen Analizi, Doğrusal Ayırt Etme Analizi, Çekirdek Prensip Bileşen Analizi, Çekirdek Doğrusal Ayırt Etme Analizi, Geri Yayılım Sinir Ağı ve Gizli Markov Modeli'dir. Gabor filtrelerinin etkinliğini göstermek için, her iki deneyden alınan test sonuçları karşılaştırılmıştır. Karşılaştırma, kullanılan yöntemlerin Gabor Wavelet filtrelerinin kullanımı ile tanıma oranındaki artışı göstermektedir.

Anahtar kelimeler: Yüz Tanıma, PCA, LDA, BPNN, HMM.

CONTENTS

TABLES	viii
FIGURES	ix
ABBREVIATIONS	xi
1. INTRODUCTION	1
1.1 BACKGROUND	1
1.2 WHY DO WE NEED FACE RECOGNITION	2
1.3 FACE RECOGNITION PROBLEM	4
1.4 STAGES OF FACE RECOGNITION	10
1.4.1 Face Detection	10
1.4.1 Feature Extraction	12
1.4.1 Face Recognition	13
1.4 APPLICATIONS OF FACE RECOGNITION	15
1.4 FACE RECOGNITION HISTORY	15
1.4 FACE RECOGNITION TODAY	16
2. RELATED WORK	17
2.1 PCA BASED WORKS	17
2.2 LDA BASED WORKS	17
2.3 NN BASED WORKS	18
2.4 HMM BASED WORKS	18
2.5 GWT BASED WORKS	18
3. FACE RECOGNITION METHODS	19
3.1 PCA	19
3.1.1 PCA Advantages	20
3.1.2 PCA Disadvantages	20
3.1.3 Mathematical representation of PCA	21
3.2 LDA	22
3.2.1 LDA Advantages	23
3.2.2 LDA Disadvantages	23
3.2.3 Mathematical representation of LDA	23

3.3 KPCA	25
3.3.1 Mathematical representation of KPCA	25
3.4 KFA	27
3.4.1 Mathematical representation of KFA	27
3.5 NEURAL NETWORKS	29
3.5.1 BPN	30
3.5.1.1 BPN Advantages	31
3.5.1.2 BPN Disadvantages	31
3.5.1.3 Mathematical representation of BPN	31
3.6 HMM	33
3.6.1 HMM Advantages	34
3.6.2 HMM Disadvantages	34
3.6.2 Mathematical representation of HMM	34
3.7 GABOR WAVELET TRANSFORM	36
3.7.1 GWT Advantages	37
3.7.2 GWT Disadvantages	37
3.7.2 Mathematical representation of GWT	37
4. RESULTS AND DISCUSSION	39
4.1 THE ORL DATABASE	39
4.2 RESULTS	40
4.2.1 Matlab	40
4.2.2 Results of face recognition methods without Gabor Wavelet	40
4.2.3 Results of face recognition methods using Gabor Wavelet	41
4.3 DISCUSSION	44
5. CONCLUSIN	46
REFERENCES	47

TABLES

Table 4.1: Test results of face recognition methods	41
Table 4.2: Test results of face recognition methods with Gabor wavelet	43
Table 4.3: Comparing test results of face recognition methods with and without Gabor wavelet filters	45

FIGURES

Figure 1.1: Popular biometrics	2
Figure 1.2: Example of some biometrics	3
Figure 1.3: An example of age difference from the FG-NET database	5
Figure 1.4: An example of color differentiation of same images	5
Figure 1.5: An example of different gender	6
Figure 1.6: An example of different facial expression	6
Figure 1.7: An example of Audrey Hepburn with different facial looks	7
Figure 1.8: Images of faces taken under different lighting	8
Figure 1.9: Images of a face with different poses	8
Figure 1.10: An example of occlusion	9
Figure 1.11: An example ethnicity	9
Figure 1.12: A common structure of face recognition system	10
Figure 1.13: The procedure of face detection	11
Figure 1.14: Face detection example	12
Figure 1.15: Feature extraction procedure	13
Figure 1.16: Fiducial points of the face	13
Figure 1.17: Categories of face recognition methods	14
Figure 3.1: PCA based face recognition system diagram	20
Figure 3.2: LDA based face recognition system diagram	22
Figure 3.3: An illustration of a biological neuron and its mathematical model	29
Figure 3.4: Multilayered neural network	30
Figure 3.5: Training of neural networks	30
Figure 3.6: Three layered BPN	31
Figure 3.7: HMM for face recognition	33

Figure 3.8: HMM based face recognition system	34
Figure 3.9: The GW representation of an image.....	36
Figure 4.1: ORL Database.....	39
Figure 4.2: Recognition rate without Gabor wavelet	41
Figure 4.3: Face recognition system with Gabor wavelet.....	42
Figure 4.4: Recognition rate using Gabor wavelet.....	43
Figure 4.5: A comparison between recognition rates.....	45

ABBREVIATIONS

BPN	: Back Propagation Network
BPNN	: Back Propagation Neural Network
CV	: Computer Vision
CNN	: Convolutional Neural Network
DNA	: Deoxyribonucleic Acid
DC	: Discrete Cosine
DCT	: Discrete Cosine Transform
FLD	: Fisher's Linear Discriminant
GWT	: Gabor Wavelet Transform
HMM	: Hidden Markov Model
HMMs	: Hidden Markov Models
ID	: Identification
KPCA	: Kernel PCA
KFA	: Kernel Fisher's Analysis
LDA	: Linear Discriminant Analysis
MLP	: Multi-Layer Perceptrons
MATLAB	: Matrix Laboratory
NN	: Neural Networks
ORL	: Olivetti Oracle Research Lab
PCA	: Principal Component Analysis
PhD	: Pretty helpful Development
RR	: Recognition Rate
SOM	: Self Organize Map
SSS	: Small Sample Size

1. INTRODUCTION

1.1 BACKGROUND

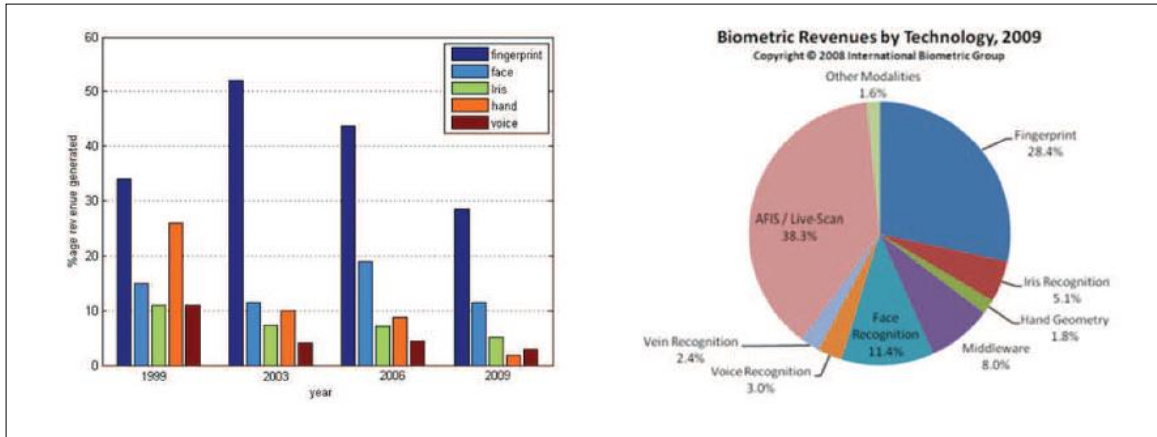
One of the most remarkable applications in the areas of Pattern Recognition and Image analysis is Face Classification or Recognition. Vision based processes between the human and the computer and vice versa depend hugely on the obtained facial images [1].

Countless researches and studies of face recognition have been made especially in the past two decades. An individual's information has become very important in today's new information age era since it can be used for social and security issues. Nowadays we can find face recognition in several areas like credit card verification, security areas, and lots of human computer interaction. Due to the major development in this area, face recognition has been used in numerous law enforcement and commercial sectors such as driver's license, passport controls, voter registration card, safety management, and database security [2].

Many biometrics systems exist, and some of them are quite popular and efficient. Yet majority of these used methods have disadvantages with their identification and verification techniques that are used in their applications, except for voice and human recognition. The reason is because when using these methods one is obliged to memorize a certain password or human action which can be lost or forgotten. On the other hand, this cannot occur to fingerprint, retina, or facial features [3]. Figure (1.1) shows the popular biometrics and the percentage of their usage.

Face recognition has shown a lot of improvement in huge databases with good pose and lighting conditions and has acquired a recognition rate of 90 percent. This is a good proof that Face Recognition Systems can be used as a replacement for a lower security environment and that it can be successfully used in different kinds of issues, i.e. multimedia [3].

Figure 1.1: Popular biometrics



Source: Z. Riaz, M. S. Sarfraz and M. Beetz, "Towards Unconstrained Face Recognition Using 3D Face Model, Biometric Systems".

1.2 WHY DO WE NEED FACE RECOGNITION?

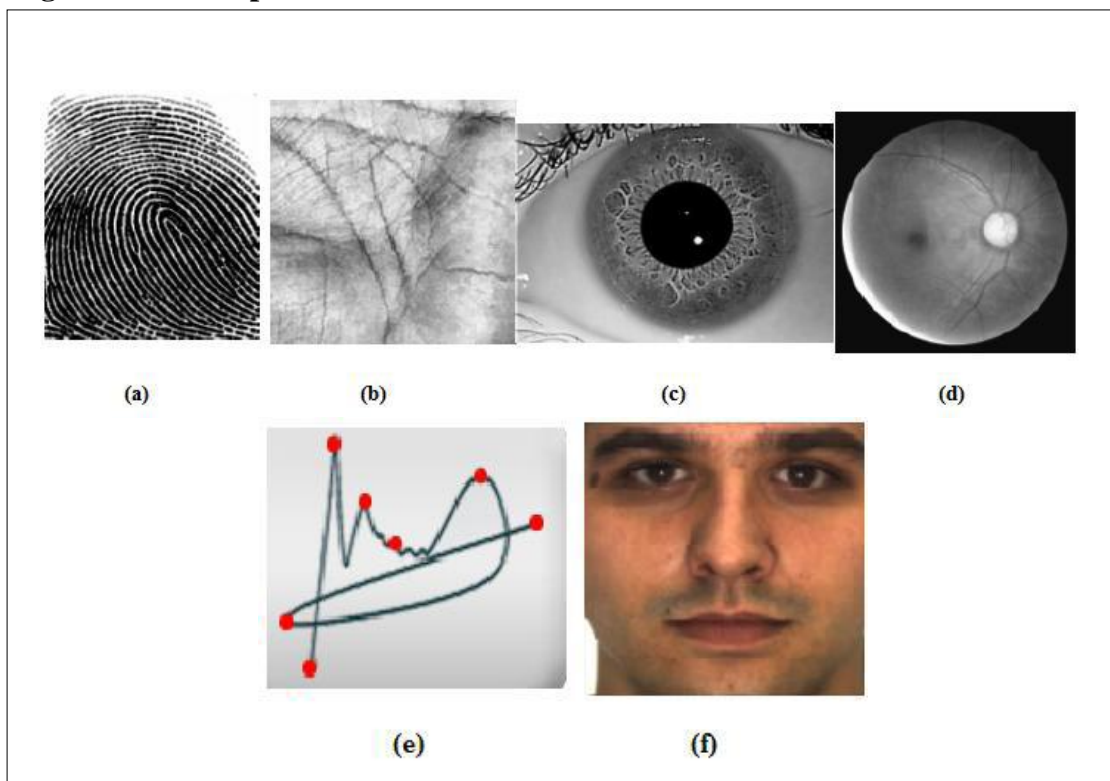
The necessity to keep the security of data and physical belongings is becoming both very difficult and very important in today's connected world. We have heard many times about crimes regarding credit card fraud or security and computer break-in's whether they occur in companies or government buildings. An example can be what happened in 1998 where cyber thieves have managed to steal over 100 million dollars. By taking an advantage of the flaws that the common access methods have, the criminals easily breached those systems. The main flaw of these systems is that they do not give the user access permission by the mean of "who we are" but by the fact of "what we have", an example would be, ID cards, PIN codes, keys, passwords, mother's maiden name, etc. If we take a good look at these ways we notice that none of them really describe us. On the contrary these ways are only meant to give us authentication for access or usage. This means that anyone who is able to get these methods of access will be able to use the data they contain whenever they want [5].

Lately the development of technology has made the availability of "true" individual identity verification possible. The technology that made it possible is called "biometrics". Biometrics grant individual's access through automated methods of recognition or verification based on some specific physiological characteristics, like fingerprints or facial features, or some behavioral features, like writing style or

keystroke pattern. What makes verification methods hard to fake is that they recognize a person by his/her biological characteristics [5].

By comparing the different biometric methods, the physiological methods like fingerprints, face, and DNA have proven to be more stable and secure than the behavioral methods like voice print and keystroke. The reason is because physiological characteristics are usually unchangeable except by a severe injury, whereas, behavioral patterns can be easily affected by stress, exhaustion, and sickness. On the other hand, people would prefer signing as a verification method than giving a blood sample [5]. Figure (1.2) shows examples of different biometrics.

Figure 1.2: Example of some biometrics



Source: R. Bhatia, " Biometrics and Face Recognition Techniques".

Face recognition has drawn the attention of a lot of researchers since 1970's especially in the fields of security, psychology, and image processing. The reason is because face recognition techniques have the characteristics of being very accurate without causing any kind of disturbance for the user. In other words, face recognition techniques have the qualities of both physiological and behavioral approach [5].

Face recognition has proven to have very successful results especially when the used face images are full-frontal and are under similar conditions in terms of lightening and pose. Yet there are many large databases that contain images with different poses and illumination conditions. Despite the fact that computers are not able to choose a person amongst thousands of people who are passing in front of a camera (people also cannot differentiate between similar people whom they have not met before [7, 8], yet applications like iPhoto, Picasa or Facebook have developed a face recognition ability in order to at least differentiate between a small number of people [9].

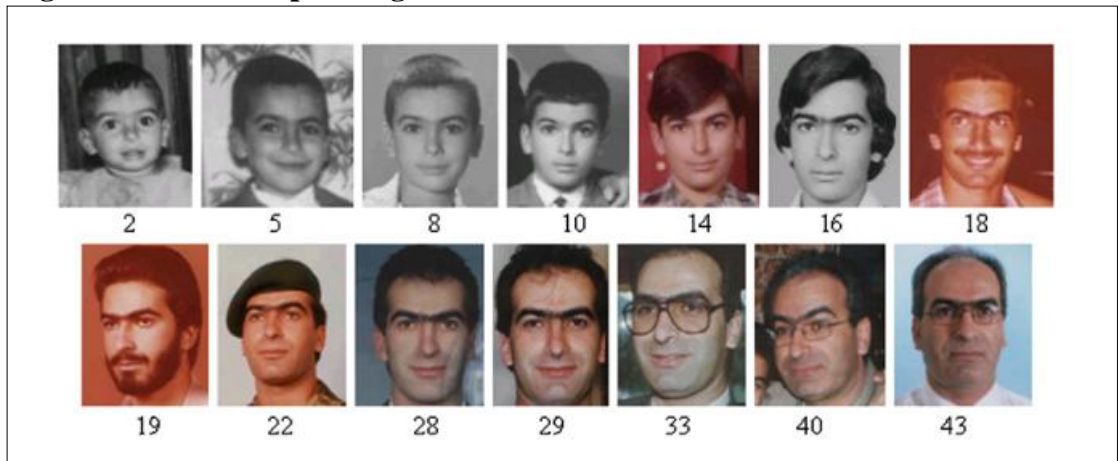
A few number of the earliest face recognition systems have included finding the location of some unique facial features like the eyes, nose, ears or mouth and use the distance measurements between those features to recognize faces. Later other approaches used different techniques like grey-level images (Eigenfaces) and shape and appearance variations [9].

1.3 FACE RECOGNITION PROBLEM

The purpose of face recognition is to detect a certain face from an image and be able to match it with the faces from a stored database. A small child is able to easily recognize a human face, yet it is difficult for a computer to do so. That is why the main goal has always been to build a face recognition system that has the ability to recognize a human face perfectly and correctly just like a small child does. Despite the huge complication, the human visual system is able to differentiate between the human faces and as a result completing the face recognition efficiently. However the same thing does not apply to computer because any change in the facial appearance or the environment of the person will have a great affect on the work of the face recognition system [2]. Here are a number of the issues that will have an affect the work of a face recognition system:

Age: human face gets very different throughout his life which makes it difficult for the computer to match someone who is old with the correct image of them when they were young. Figure (1.3) shows an example of how a person can look differently throughout different age phases.

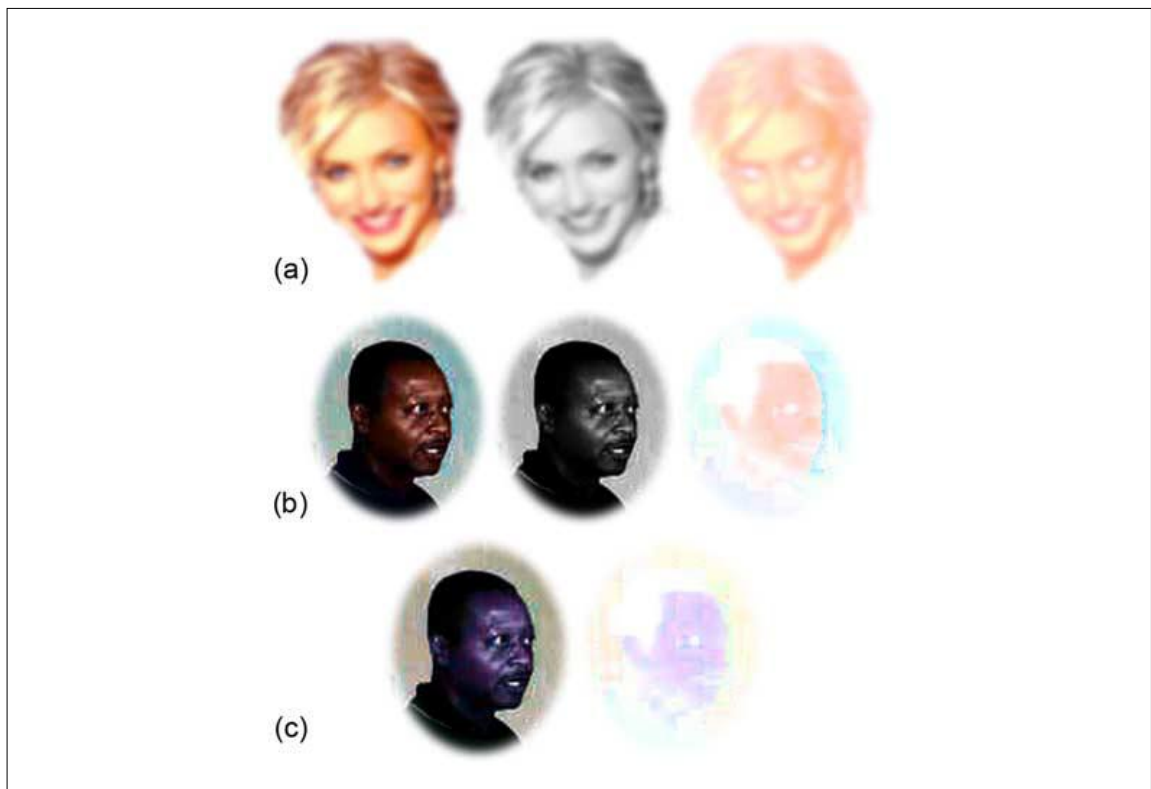
Figure 1.3: An example of age difference



Source: The FG-NET database.

Color: skin color affects the recognition process because the system might consider the same person as two different people. Figure (1.4) shows how different color conditions can have an effect on the image.

Figure1.4: An example of color differentiation of same images



Source: Sinha, P., Balas, B., Ostrovsky, Y., and Russell, R., "Face recognition by humans: Nineteen results all computer vision researchers should know about".

Gender: a computer cannot differentiate between man and women unless there is a specific application designed for that. Figure (1.5) shows gender differentiation.

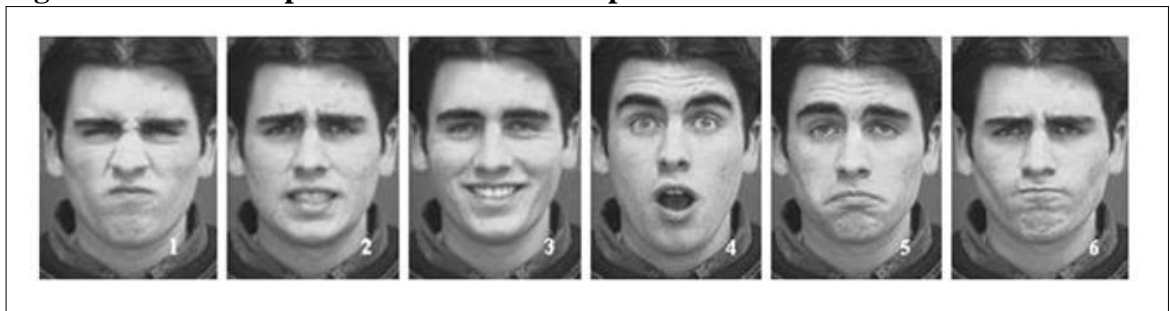
Figure1.5: An example of different gender



Source: <http://journal.frontiersin.org/article/10.3389/fpsyg.2013.00029/full>.

Facial expression: appearance of people while they are laughing or sad or angry can be very different and they affect the shape of the face in an image. Figure (1.6) shows different facial expressions.

Figure 1.6: An example of different facial expression



Source: Kanade, T., Cohn, J., Tian, Y.-L. "Comprehensive database for facial expression analysis".

Facial feature: wearing glasses or having beard or makeup or wearing a hat, scarf, or glasses make the faces look different. Figure (1.7) shows how a person looks different wearing different looks.

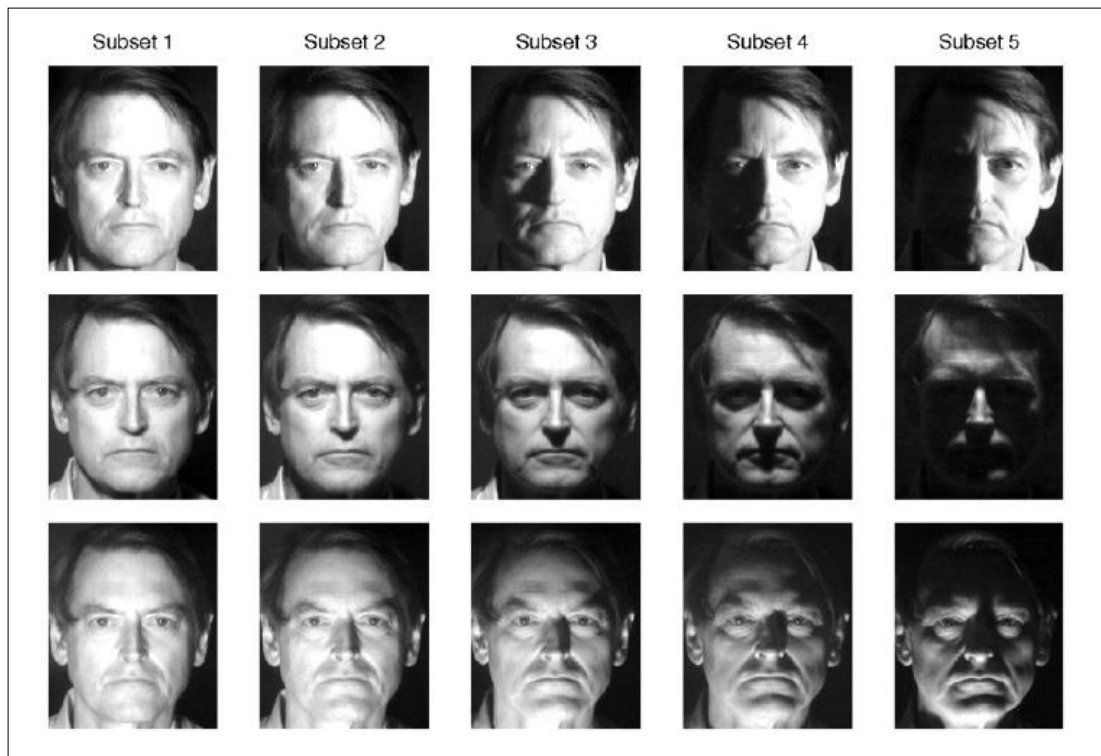
Figure 1.7: An example of Audrey Hepburn with different facial looks



Source:https://www.google.com.tr/search?hl=ar&site=imghp&tbm=isch&source=hp&biw=1366&bih=662&q=audrey+hepburn&oq=audr&gs_l=img.3.0.0110.2077.3374.0.5150.4.4.0.0.0.118.436.0j4.4.0....0...1ac.1.64.img..0.4.434.VxnBtxJtF1g.

Illumination condition: lighting makes faces look different because faces in bright light appear different than faces in dark light. Figure (1.8) shows the same face under different illumination conditions.

Figure 1.8: Images of faces taken under different lighting



Source: P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,”.

Pose variation: different poses of the face whether it is frontal or side view affect the process of face recognition. Figure (1.9) shows different poses of the same face.

Figure 1.9: Images of a face with different poses



Source: J. Yang D. Zhang, A. F. Frangi, and J. Y. Yang, “Two-dimensional PCA: A new approach to appearance-based face representation and recognition,”.

Occlusion: faces that are being partially covered by another person's face or hand are difficult to recognize. Figure (1.10) shows occlusion in a crowded environment.

Figure 1.10: An example of occlusion



Source: Luis Torres, "Is There Any Hope For Face Recognition?", Technical University of Catalonia, Barcelona, Spain.

Ethnicity: being part of a specific ethnicity could have an effect of the performance of face recognition system as well [15]. Figure (1.11) shows people from different ethnic groups.

Figure 1.11: An example ethnicity

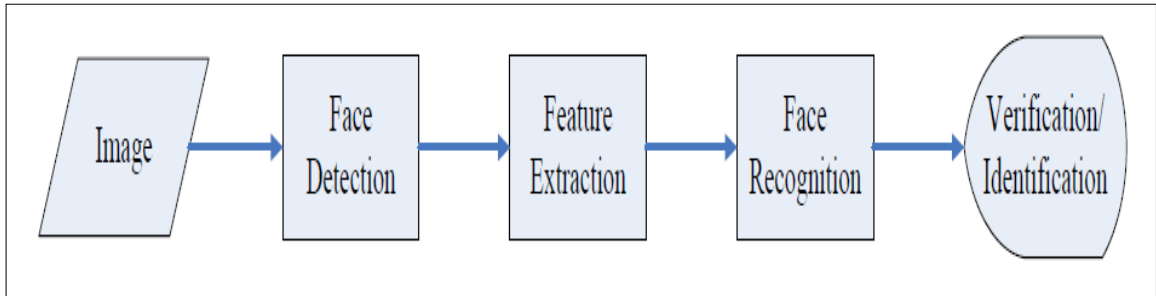


Source: <http://biology.about.com/od/genetics/ss/Polygenic-Inheritance.htm#step3>.

1.4 STAGES OF FACE RECOGNITION

Face recognition system usually works in three known phases. These phases are face detection, feature extraction, and face recognition. The way a face recognition system performs is by extracting all the needed information from an image in order to recognize the person's identity. Figure (1.12) shows a face recognition structure.

Figure 1.12: A common structure of face recognition system

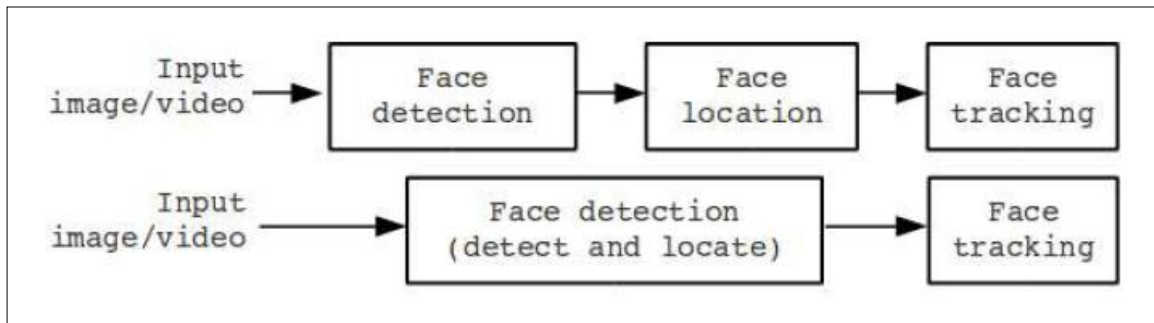


Source: M. D. Hegde, Sayeesh, "Face expression recognition using score level fusion method".

1.4.1 Face Detection

The act of locating the face part in an image is known as face detection. Face detection is not a simple task for a computer even though it looks like a simple task for the human. Verifying if human faces are shown in a given image and where they are located at is a key part of this phase. Patches that hold the face parts from the input image or video are usually the estimated result of the face detection phase. Face detection can be used also for other purposes besides being utilized as the pre-processing tool for face recognition system. Two important data results should be considered in every detection problem, first one is called true positives, and the second is called false positives [17]. A perfectly working face detection system should have high true positives and low false positives [18]. Figure (1.13) shows the steps of a face detection procedure, it also shows that in some systems the face detection and location are performed at the same time whereas in other systems they are performed separately.

Figure 1.13: The procedure of face detection



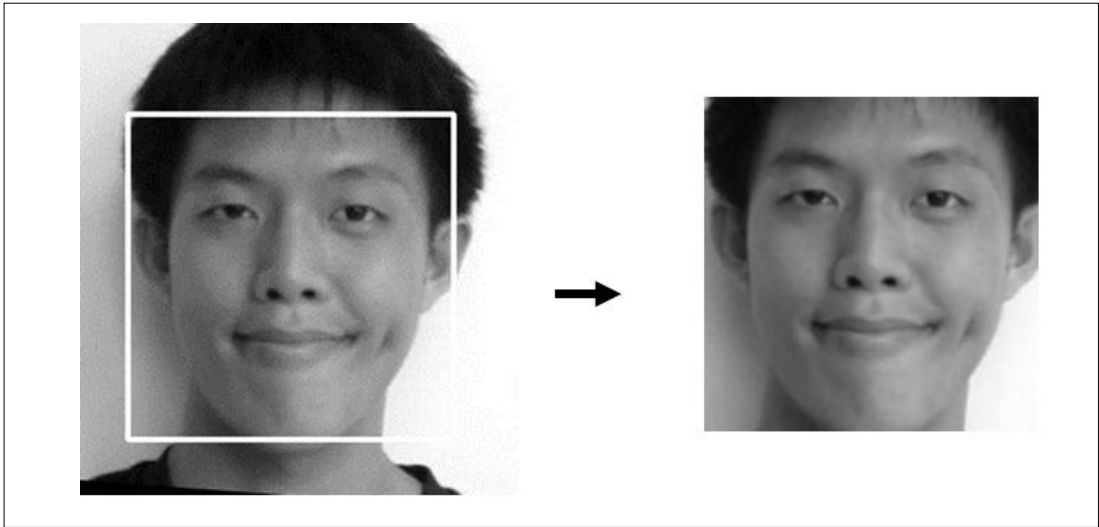
Source: I. Marques, "Face Recognition Algorithms", June 16, 2010.

Face detection methods are divided into four types:

- a) **Knowledge-based methods:** with these methods, the familiarity of a usual face representation by the human is trained and used to find faces by computing the correlation between facial features. They are mostly used for face localization.
- b) **Feature-based methods:** with these methods, some distinct facial appearances like the eyes, nose, and mouth are found and used despite their pose or lighting conditions to find the location of the face. These methods are also used mostly for face localization.
- c) **Template-based methods:** with these methods, a stored description of some patterns of a face as a whole or separated is used to detect faces by calculating the relationship with the input image. These methods are used for both detection, and localization.
- d) **Appearance-based methods:** with these methods, learned models from a collection of trained images, these methods are mostly used for detection.

Figure (1.14) shows face detection.

Figure 1.14: Face detection example

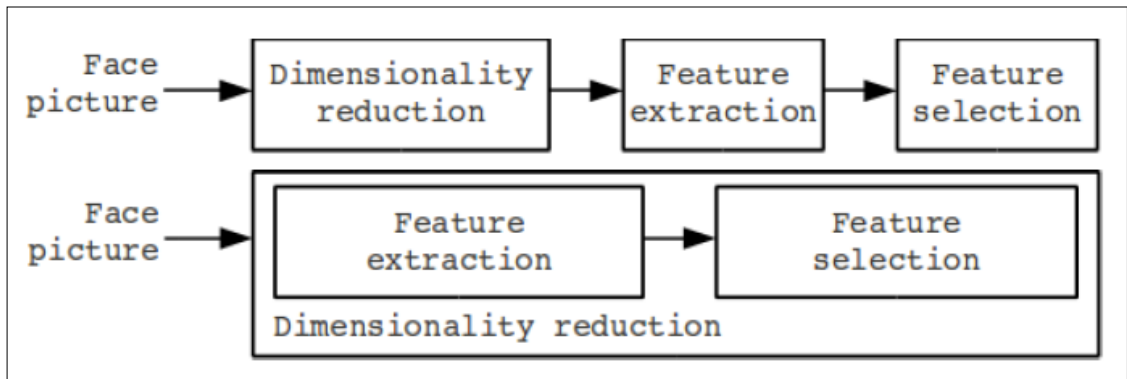


Source: <http://robinhsieh.com/?p=156>.

1.4.2 Feature extraction

Once the human face has been detected in the face detection phase, the role of feature extraction comes. The patches estimated from the face detection phase cannot be used directly in the recognition phase since the size will be big and the detected face image might suffer from occlusion. The act of extracting the important information from the detected face part of an image is known as feature extraction [17]. Feature extraction phase goes through three steps: dimensionality reduction, feature extraction, and feature selection. The first phase is an important task to do in order to reduce the size of the extracted face image. Feature extraction and selection are used interchangeably most of the time, but a distinction between the two must be made. Feature extraction as the name indicates extracts the feature of the face from the input data. Whereas, feature selection chooses the best parts of the input feature from the information set. Figure (1.15) shows the steps of feature extraction phase, it also shows how feature extraction and selection are sometimes considered and used as being the same.

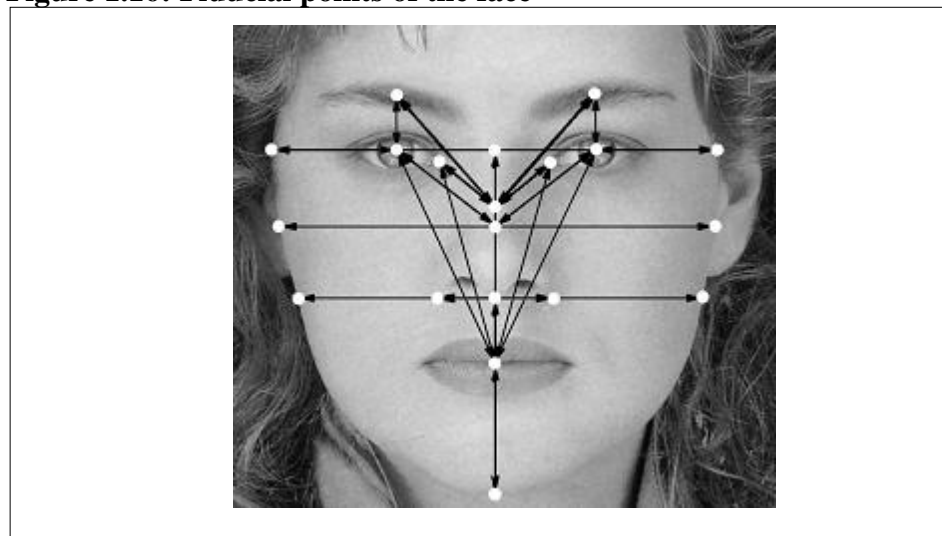
Figure 1.15: Feature extraction procedure



Source: I. Marques, "Face Recognition Algorithms", June 16, 2010.

At the end of this phase, the patches estimated from the detection phase are transformed into vectors. These vectors are going to have a fixed size and some fiducial points along with their positions. Figure (1.16) shows the fiducial points of the face.

Figure 1.16: Fiducial points of the face



Source: V.V. Starovoitov, D.I Samal1, D.V. Briliuk," Three Approaches for Face Recognition".

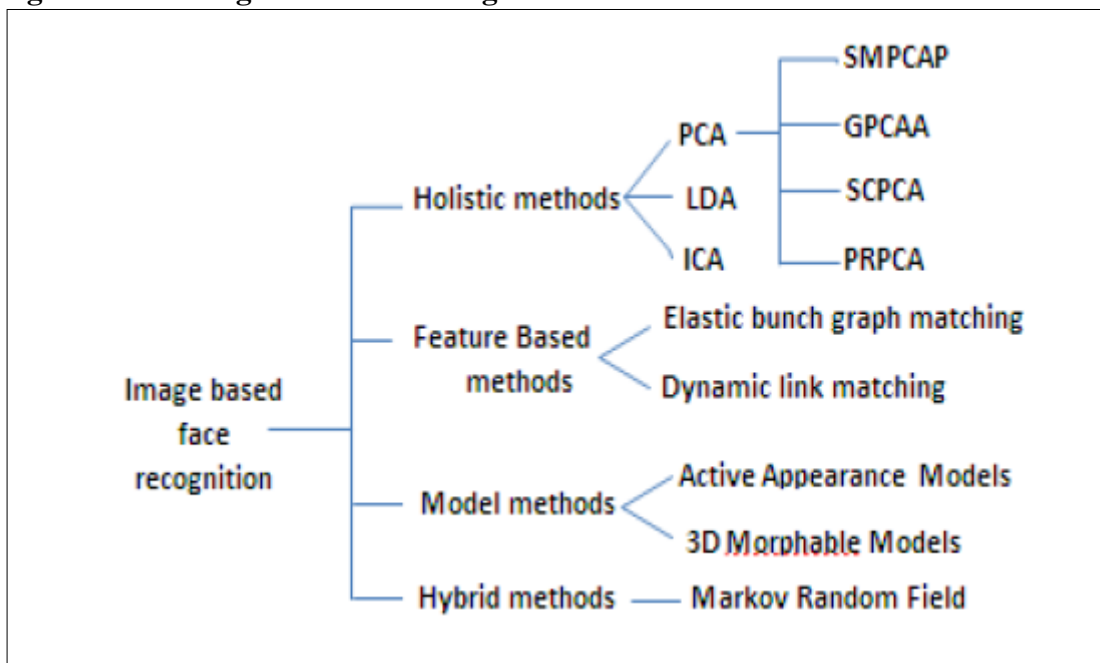
1.4.3 Face recognition

The last phase in the face recognition system is the face recognition phase. After the features have been extracted in the previous phase, we now need to identify the face in the image. In order to perform an automated recognition system, we need to build a face

database first. The database contains a number of images for every individual along with their extracted features. The way how face recognition works is by comparing the features of the input image with the features of images that are stored in the database and getting the closest match as a result [17]. Face recognition methods are usually divided into three types:

- a) **Holistic Methods:** with these methods, the entire face area is defined as a vector with high dimensions and it is used like a classifier input.
- b) **Local Methods:** with these methods, specific features of the face (nose, eyes, and mouth) are extracted and used to identify the face.
- c) **Hybrid Methods:** with these methods, a combination of both holistic and local methods is used to identify the face. These methods are claimed to be better than the previous two methods [20] because they keep the strengths of them yet they overcome their weaknesses. Figure (1.17) shows the different categories of face recognition methods.

Figure 1.17: Categories of face recognition methods



Source: B. S. Bargavi, C. Santhi, " Global and Local Facial Feature Extraction using Gabor Filters".

1.5 APPLICATION OF FACE RECOGNITION

Applications of face recognition are divided into two types, one that is used for commercial purposes and the other for law enforcement purposes.

Face identification: face identification systems identify people with images of their faces. This means that the attendance of an authorized person is needed instead of an identification card or a password.

Access control: face recognition systems used for access control in an office or computer logon achieve a high accuracy with much co-operation for the user.

Security: security has become a primary concern now more than ever in everywhere. Face recognition system have been implemented in many areas to alert the police whenever a person that matches the appearance of a terrorists enters.

Investigation: search image database for immigrants, missing children and licensed drivers.

Identity verification: to minimize fraud during electoral registration, bank transfers, electronic commerce, national IDs, passports, and healthcare.

Surveillance: using surveillance footage of a crime to search for the suspect.

1.6 FACE RECOGNITION HISTORY

Automatic face recognition system is a new concept to some extent. The setback of the early recognition systems was that the locations of features and the measurements of them were calculated manually. A company called Panoramic Research in California has started the earliest machine learning researches back in 1960's. This company was amongst lots of companies that were doing research in this area of study, which later on was known as artificial intelligence, arising during that time. Since the Department of Defense of United States was funding majority of these research, they inevitably became part of the technological power in the United State fight during the Cold War. A man named Woodrow Wilson Bledsoe, Woody Bledsoe, who was confounders of the

Panoramic Research, is known as being one of the first researchers who worked in the field of face recognition system by machines [22]. Other researches back in the 1970's have used unique characteristics of the individual's face for recognition like the color of the hair or the thickness of the lips [23]. Another research done in 1988, which was considered to be the landmark of face recognition problem because Principal Component Analysis was applied and used for coding the system. The results proved that in order to code the system correctly so that it aligns and normalizes a face image, less than a 100 value is required [24]. Turk and Pentland have discovered in 1991, that the remaining error can be used for detecting faces in an image while they were doing experiments using Eigenfaces techniques. Their discovery has made it possible to do a successful real time face recognition system [25]. It has also created a lot of interest in the field of automated face recognition despite the constraints of environmental factors.

1.7 FACE RECOGNITION TODAY

Face recognition has evolved a lot since 1960's with the availability of new technologies nowadays. Face recognition is being used in different areas today from governmental to commercial. Today, the technology of face recognition has been implemented and being used to counteract terrorism, find missing children, investigate crimes, combat passport fraud, and many more.

2. RELATED WORK

In this thesis, we are focusing on the problem of face recognition and the different methods used for it. Lots of researches have been done in order to build the best recognition system that will overcome majority of the obstacles especially in the recent years. In this section a number of related works in the area of face recognition are listed below:

2.1 PCA BASED WORK

Many researchers have used PCA in their works. (Bakshi and Singhal, 2014) introduce a novel face recognition technique by using PCA, DCT, and SOM neural networks. In their system, PCA was used for dimensionality reduction, DCT was used for compression, and SOM was used for face representation. They have implemented and evaluated their work in MATLAB by taking images of people with different facial expressions. Their system has acquired a recognition rate of 97.5% which is very high. Besides the high recognition rate, they say that their system does not require a lot of computation and that its execution time is very fast [26].

In another work (Jayanthi and Aji, 2014), KPCA has been used for face recognition. General features of the faces (like beard, eyebrows, and glasses) were excluded from the images by using the skin segmentation method of the Gaussian model. Results have shown that the system works efficiently for recognizing faces [27].

Combining both PCA and eigenfaces for face recognition is also introduced in a different work (NichalI, Jagtap, Ingale, Patil). Here, face images are changed into face patches and these patches are identified by the eigen face and after that PCA is used for the recognition of the face [28].

2.2 LDA BASED WORKS

LDA is a well-liked face recognition method as well. Here (Chen, Liao, Ko, Lin, and Yu, 2000) a new LDA based face recognition system that will overcome the problem of a small size is introduced. The test results confirm that a great improvement happened in the performance of face recognition by using the LDA new method [29].

A fusion of LDA and PCA (Dargham, Chekima, and Mounq, 2012) is introduced. The results show an increase of the recognition rate of the new method by 8% more than the PCA based method, and 3% more than the LDA based method [30].

2.3 NN BASED WORKS

Neural Networks are popular methods for face recognition. Another hybrid method (Dhoke, Parsai, 2014) is introduced using BPN with PCA. For dimensionality reduction PCA is used, and for face recognition BPN is used. Results were quite high and the performance was very efficient according them [31].

A face recognition method (Parkhi, Vedaldi, Zisserman, 2015) was introduced which is known as Convolutional Neural Network (CNN) or Deep recognition. The obtained results show that with the correct training of this method, a really high recognition rate can be got [32].

2.4 HMM BASED WORKS

HMM is another useful tool for face recognition. (S. K. Jameel, 2015) introduces a method using DCT for feature extraction, PCA for dimensionality reduction, and HMM for face recognition. The results gave a recognition rate of 95.211% using the ORL database [33].

Other work (Raut and Patil, 2012) introduces a MC-HMM based system by using discriminative feature extraction method for face recognition. The experiments show higher results compared to the conventional HMM based systems [34].

2.5 GWT BASED WORKS

Many works have been done using GWT for face recognition. (Cho, Roberts, Jung, Choi, and Moon, 2014) have used a hybrid method that combined PCA and GWT. The results showed efficiency and higher recognition rate than using GWT algorithm by its own [35].

Another work based on GWT filters (Jin and Ruan, 2009) is introduced with Improved Supervised Locality Preserving Projections. Results have shown the effectiveness of the new method in terms of great performance and high recognition rate [36].

3. FACE RECOGNITION METHODS

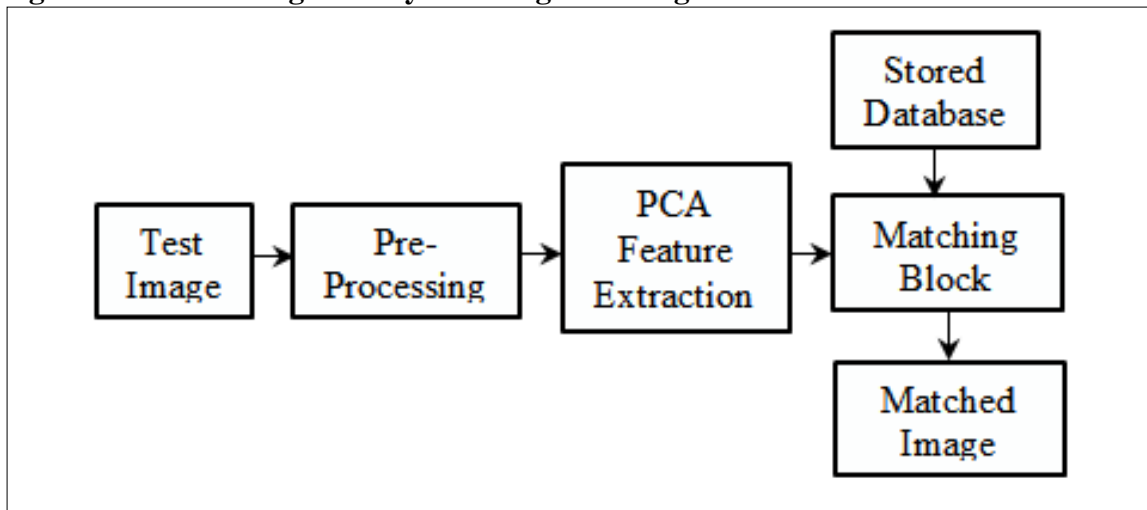
Numerous researchers from various backgrounds were attracted by the problem of face recognition for being both challenging and interesting problem. When face recognition searches began back in the 1970's, it used to be treated like a two-dimensional problem. The distances between important points of the face were used as a way to identify faces (i.e. distance between the eye, ears, or mouth and nose). The problem of face recognition is known as having two sets of images that are saved in the database. One of the image sets is defined as the training database and the other is defined as the testing database. The problem is being able to identify a person in the testing database when matching with the training database correctly.

In this section, popular face recognition methods are described in detail along with their advantages and disadvantages.

3.1 PCA (EIGENFACES)

One of the most commonly used methods for face recognition is Principal Component Analysis (PCA). The core vectors of PCA are calculated from a set of images that are made for training [37]. PCA usually uses eigenfaces in which both the image that is being searched for and the images in the database are the same size. In order to provide a structure of the face feature with low dimensions, dimensionality reduction approach is used. During the reduction of the face size useless information are taken out and the face parts are broken into a set of orthogonal pieces called eigenfaces. These pieces are then stored in a one-dimensional array [38]. PCA is not just used for face recognition; it also has other different applications like voice recognition, medical imaging analysis, lip reading, and gesture interpretation. Figure (3.1) shows a diagram of a face recognition system using PCA.

Figure 3.1: Face recognition system diagram using PCA



Source: A. Nichal, H. Jagtap, S. Ingale, N. Patil, " Face Recognition using PCA and Eigen Face Approach".

3.1.1 PCA Advantages

- 1) Easy and efficient for recognition.
- 2) The compression of information is done by the low-dimensional subspace representation.
- 3) Further processing is not required for the gained raw data and they can be used directly for learning and recognition.
- 4) Reproduction of faces and being familiar with their geometrical structure is not necessary.

3.1.2 PCA Disadvantages

- 1) For scale normalization, a low-level preprocessing is used because PCA is quite sensitive to scaling.
- 2) Recognition rate decreases under different pose, illumination, expression and disguise.
- 3) Updating the database is hard because data learning takes a lot of time.

3.1.3 Mathematical Representation of PCA

In Computer Vision, especially face recognition field, PCA method is a generally used for dimensionality reduction and it is also known as the Karhunen-Loeve methods. In order to increase the spread out of the projected samples, PCA uses a linear projection method for dimensionality reduction [40].

In order to understand the work of a PCA method more, a number of certain sample images N is set $\{x_1, x_2, \dots, x_N\}$ with N -dimensional image space values, also let us presume that every image is a part of c classes $\{X_1, X_2, \dots, X_c\}$. Furthermore, we assume the change of the original n -dimensional image space into a m -dimensional feature space happens by a linear transformation, where m is less than n . A matrix with orthonormal columns $W \in \mathbb{R}^{n \times m}$, and the new feature vectors $y_k \in \mathbb{R}^m$ are both defined by the following:

$$y_k = W^T x_k, \quad k = 1, 2, \dots, N \quad (3.1)$$

S_T which is the total distribution of matrix is defined as:

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (3.2)$$

Where the mean of all the samples images is $\mu \in \mathbb{R}^n$, and W^T is the linear transformation, $W^T S_T W$ is the distribution of the changed feature vectors $\{y_1, y_2, \dots, y_N\}$. In order to increase the determinant of the total distribution matrix of the projected samples we choose W_{opt} , i.e.

$$W_{opt} = \arg \max_W |W^T S_T W| = [w_1 \ w_2 \ \dots \ w_m] \quad (3.3)$$

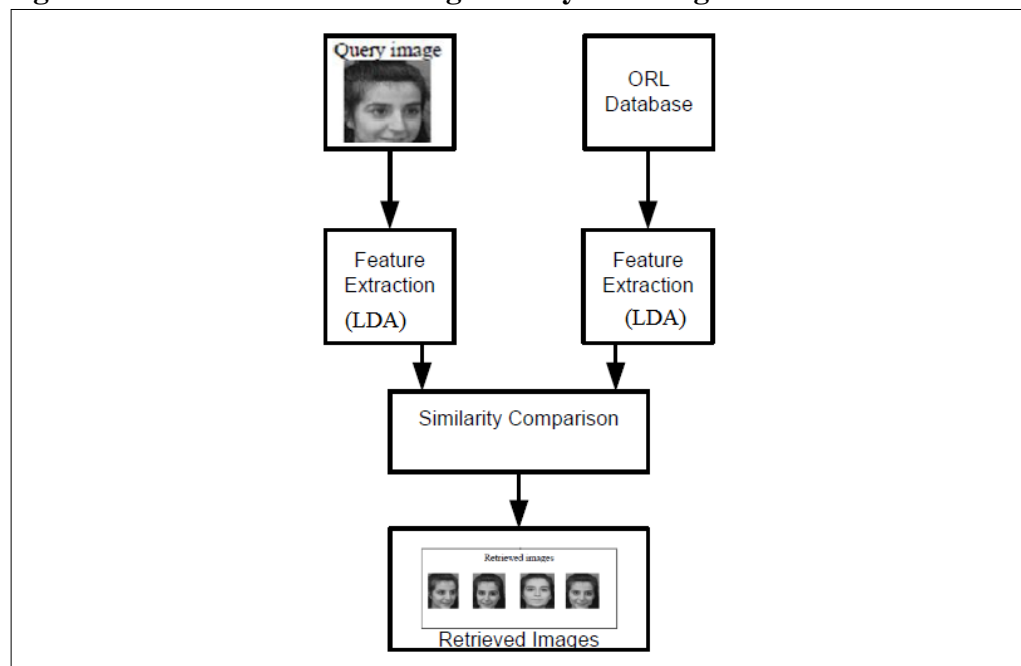
$\{w_i \mid i = 1, 2, \dots, m\}$ is S_T with a set of eigenvectors with N -dimensional and it should be matching the eigenvalues with the largest m . Eigenvectors are usually known as Eigenfaces because they have the same sizes the original images have [11].

3.2 LDA (FISHERFACES)

Another popular technique for both feature extraction and dimensionality reduction is known as Linear Discriminant Analysis (LDA), or Fisher's Discriminant Analysis. LDA is able to easily discover the inequality of the within-class frequencies and examine the performance of the randomly generated data. Maximum separation of data is guaranteed with LDA as it increases the percentage of the between-class difference to the within-class difference in any group of data. LDA has been used in other applications besides face recognition like recovery of images and speech recognition [11].

By applying LDA principle, images of the same class are put together in one group and are separated from images of the other classes. Just like PCA method, images used for training are projected into a subspace and by using similarity measures the same projected subspace from test images are identified accordingly [36]. However, LDA methods try to find the subspace that separates the different face classes in the best way, whereas PCA methods extract the facial features in order to represent faces. Figure (3.2) shows a diagram of an LDA based face recognition system.

Figure 3.2: LDA based face recognition system diagram



Source: Y. Jin, Q. Ruan, "Face Recognition Using Gabor-Based Improved Supervised Locality Preserving Projections".

3.2.1 LDA Advantages

- 1) Classification is easier.
- 2) The illumination problem is solved by increasing the percentage of the between-class distribution to within-class distribution.
- 3) When increasing the image's low-dimensional representation, the most discriminant feature is focused on and this gives LDA an advantage over PCA.

3.2.2 LDA Disadvantages

- 1) A singularity problem, which is a built in limitation of LDA, makes it fail when all distributed matrices are singular.
- 2) The problem of the Small Sample Size (SSS) happens in practice due to the availability of large number of pixels, yet the feature space dimensions are bigger than the total amount of training samples, LDA method fails. The reason is because this means that all the distributed metrics are singular which takes us to disadvantage 1.

3.2.3 Mathematical Representation of LDA

Linear data separation can be achieved with the use of linear projection for dimensionality reduction. This is a powerful statement that agrees with the use of linear based methods for dimensionality reduction for face recognition, especially if we are looking for insensitivity to different lighting conditions.

With LDA, a better dimensionality reduction method of the feature spaces is constructed by using the information from the learning set is used. By using LDA method we choose W in equation (1) is so that the percentage of the between-class distribution and the within-class distribution is increases.

By defining the between-class distribution matrix as

$$S_B = \sum_{i=1}^c |x_i| (\mu_i - \mu)(\mu_i - \mu)^T \quad (3.4)$$

the within-class distribution matrix is defined as

$$S_W = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i) (x_k - \mu_i)^T \quad (3.5)$$

where the mean image of the class X_i is μ_i , and the amount of the samples in class X_i is $|x_i|$. Assuming S_W is non-singular, as the matrix with orthonormal columns we choose the best projection represented by W_{opt} . This will increase the percentage of the between-class distribution matrix determinant to the within-class distribution matrix determinant of the estimated samples:

$$W_{opt} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1 w_2 \dots w_m] \quad (3.6)$$

$\{w_i | i = 1, 2, \dots, m\}$ represents the group of generalized eigenvectors of S_B and S_W and they are equal to the largest generalized eigenvalues of m , $\{\lambda_i | i = 1, 2, \dots, m\}$,

$$S_B w_i = \lambda_i S_W w_i, \quad i = 1, 2, \dots, m \quad (3.7)$$

It is noted that the generalized eigenvalues are mainly $c-1$ non-zero, which is the same as a higher bound on m , c is, as stated in the previous section, the number of classes.

Due to the fact that the order of S_W is generally $N-c$, and the amount of pixels in every image n is a lot bigger than the amount of images in the learning database N , the within-class distribution matrix $S_W \in \mathbb{R}^{n \times m}$ is constantly singular.

LDA overcomes the problem of a singular S_W by moving the set of input images to a low-dimensional location which changes the within-class distribution matrix S_W into non-singular. PCA is used for dimensionality reduction to $N-c$ from the feature space and after that the classic Fisher's LDA is applied which is identified by equation (3) so that the dimension is decreased to $c-1$. As a result W_{opt} is specified by:

$$W_{opt} = W_{fld} W_{pca} \quad (3.8)$$

Where

$$W_{pca} = \arg \max_W W^T S_B W \quad (3.9)$$

$$W_{fld} = \arg \max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|} \quad (3.10)$$

As we see, both W_{pca} and W_{fld} are completed using orthogonal columns with the difference that W_{pca} optimization is done with $n \times (N - c)$ matrices, while W_{fld} optimization is done with $(N - c) \times m$ matrices. The smallest c principal component was thrown away by calculating W_{pca} [11].

3.3 KPCA

Kernel principal component analysis (Kernel PCA) is the use of PCA with kernel methods.

3.3.1 Mathematical Representation of KPCA

By giving a set of m centered, where the mean is equal to zero and the variance is the same, for samples X_k , $X_k = \{x_{k1}, \dots, x_{kn}\}^T \in R^n$, increasing the variance, C , by finding the projection direction using PCA is the same as finding the eigenvalues from the covariance matrix

$$\lambda \mathbf{w} = C \mathbf{w} \quad (3.11)$$

For eigenvalues $\lambda \geq 0$, and eigenvectors $\mathbf{w} \in R^n$. In Kernel PCA, a non-linear diagram function $\Phi : R^n \rightarrow R^f$, $f \gg n$ is used in order to change the location of the \mathbf{x} vector of the input space, R^n , to feature space with higher dimensions, R^f . In R^f , the equivalent eigenvalue problem is

$$\lambda \mathbf{w}^\Phi = C^\Phi \mathbf{w}^\Phi \quad (3.12)$$

in which a covariance matrix is represented by C^Φ . And all of the solutions \mathbf{w}^Φ with $\lambda \neq 0$ are positioned in the distance of $\Phi(x_i), \dots, \Phi(x_m)$, and coefficients α_i exists there as:

$$\mathbf{w}^\Phi = \sum_{i=1}^m \alpha_i \Phi(x_i) \quad (3.13)$$

A matrix K with the dimensions $m \times m$ is represented by:

$$K_{i,j} = k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (3.14)$$

, this represents the Kernel PCA by:

$$m\lambda K\alpha = K^2\alpha \quad (3.15)$$

$$m\lambda\alpha = K\alpha \quad (3.16)$$

Where α indicates a column vector with entries $\alpha_1, \dots, \alpha_m$. The above sources suppose that all of the estimated samples $\Phi(x)$ are located in R^f (For more information on a method about how vectors $\Phi(x)$ are centered in R^f check [41]).

Kernel PCA is an overview of the classic PCA because for different non-linear projections, different kernels can be used.

By assigning the vectors in R^f to location that is surrounded by the eigenvectors \mathbf{w}^Φ with low dimensions, let us consider that R^f has a test sample x with the projection $\Phi(x)$ in the same location, then the non-linear principle components is the projection of $\Phi(x)$ towards the eigenvectors \mathbf{w}^Φ and is equal to Φ :

$$\mathbf{w}^\Phi \cdot \Phi(x) = \sum_{i=1}^m \alpha_i (\Phi(x_i) \cdot \Phi(x)) \sum_{i=1}^m \alpha_i k(x_i, x) \quad (3.17)$$

This means that, without using costly operations that only assign the samples to a R^f location with high dimensions, we can take out the first non-linear principal components q ($1 \leq q \leq m$) with the kernel function. This results in the match of the first q non-increasing eigenvalues of (10) and the first q components. The extracted non-linear

principal components where a face image is trained by every x , are called Kernel Eigenfaces or KPCA in a face recognition system [42].

3.4 KFA

Achieving a non-linear discriminant analysis in the higher space is the idea of Kernel FLD. KFA and KPCA both used the same procedure with the exception that the former uses FLD and the latter uses PCA.

3.4.1 Mathematical Representation of KFA

Just like the sources in Kernel PCA, $\Phi(x)$ which is the estimated samples are presumably located in R^f (For more information on a method about how vectors $\Phi(x)$ are centered in R^f check [42]), we make the dot products in the equations so that they are only used for FLD. \mathbf{S}_W^Φ and \mathbf{S}_B^Φ represent the within-class and between-class distribution matrices repeatedly, also in order to locate the eigenvalues λ and the eigenvectors \mathbf{w}^Φ of (12) through the application of FLD

$$\lambda \mathbf{S}_W^\Phi \mathbf{w}^\Phi = \mathbf{S}_B^\Phi \mathbf{w}^\Phi \quad (3.18)$$

, that can be found with

$$W_{OPT}^\Phi = \arg \max_{W^\Phi} \frac{|(W^\Phi)^T \mathbf{S}_B^\Phi W^\Phi|}{|(W^\Phi)^T \mathbf{S}_W^\Phi W^\Phi|} = [\mathbf{w}_1^\Phi, \mathbf{w}_2^\Phi, \dots, \mathbf{w}_m^\Phi] \quad (3.19)$$

In which the set of generalized eigenvectors is identified by $\{\mathbf{w}_i^\Phi | i = 1, 2, \dots, m\}$ and equivalent to $\{\lambda_i^\Phi | i = 1, 2, \dots, m\}$ which is the m largest generalized eigenvalues.

The kernel function is defined by giving the classes t and u and their samples as:

$$(k_{rs})_{tu} = k(x_{tr}, x_{us}) = \Phi(x_{tr}) \cdot \Phi(x_{us}) = \Phi(x_{tr})^T \Phi(x_{us}) \quad (3.20)$$

Let the elements $(K_{tu})_{u=1, \dots, c}^{t=1, \dots, c}$ be the representation of an $m \times m$ matrix K , where a dot products that are part of a matrix collection located in the feature space R^f defined by K_{tu} , as

$$K = (K_{tu})_{u=1, \dots, c}^{t=1, \dots, c} \text{ where } K_{tu} = (K_{rs})_{s=1, \dots, l_u}^{r=1, \dots, l_t} \quad (3.21)$$

Note that a $l_t \times l_u$ matrix defines K_{tu} , and a $m \times m$ symmetric matrix defines K . Matrix Z is then defined as:

$$Z = (Z_t)_{t=1, \dots, c} \quad (3.22)$$

In which a $l_t \times l_t$ matrix defines Z_t with terms equal to $\frac{1}{l_t}$, in other words, Z is a $m \times m$ block diagonal matrix. In a high-dimensional feature space R^f , the between-class and within-class distribution matrices are defined as

$$\mathbf{S}_B^\Phi = \sum_{i=1}^c l_i \boldsymbol{\mu}_i^\Phi (\boldsymbol{\mu}_i^\Phi)^T \quad (3.23)$$

$$\mathbf{S}_W^\Phi = \sum_{i=1}^c \sum_{j=1}^{l_i} \Phi(\mathbf{x}_{i,j}) \Phi(\mathbf{x}_{i,j})^T \quad (3.24)$$

In which the mean of class i in R^f is represented by $\boldsymbol{\mu}_i^\Phi$ and the number of samples l_i belongs to the class i . From the theory of reproducing kernels (Aronszajn, 1950), every solution $\mathbf{w}^\Phi \in R^f$ must be in the distance of all the samples being trained in R^f , meaning,

$$\mathbf{w}^\Phi = \sum_{p=1}^c \sum_{q=1}^{l_p} \alpha_{pq} \Phi(\mathbf{x}_{pq}) \quad (3.25)$$

We can obtain the solution for (19) by solving:

$$\lambda K K \alpha = K Z K \alpha \quad (3.26)$$

As a result, we can write (13) as

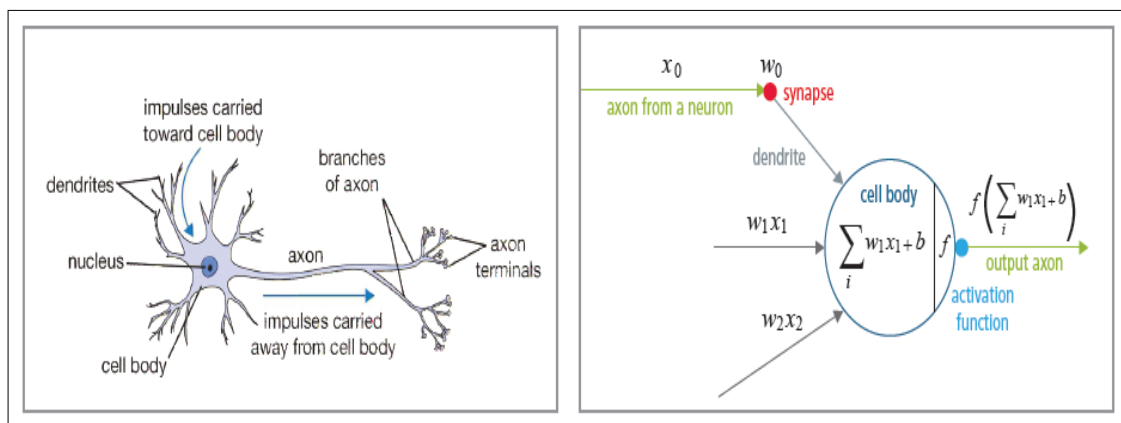
$$\begin{aligned} W_{OPT}^\Phi &= \arg \max_{W^\Phi} \frac{|(W^\Phi)^T S_B^\Phi W^\Phi|}{|(W^\Phi)^T S_W^\Phi W^\Phi|} \\ &= \arg \max_{W^\Phi} \frac{\alpha K Z K \alpha}{\alpha K K \alpha} \\ &= [\mathbf{w}_1^\Phi \dots \mathbf{w}_m^\Phi] \end{aligned} \quad (3.27)$$

Assigning the vectors $\Phi(x)$ in R^f , in a way that is similar to KPCA, to a location that is surrounded by the eigenvectors w^Φ with low dimensions (see section 3.3). By applying the same technique in the Fisherface method (in order to avoid singularity problems when calculating W_{OPT}^Φ) in face recognition (see section 3.2), the extracted eigenvectors in (21) is called Kernel Fisherfaces or KFA [42].

3.5 NEURAL NETWORKS

In order to process the information that a human brains receives, it uses a set of inter connected processing elements called neurons. These neurons are self-governing and work with a non-parallel pattern. When faced with a problem like recognizing faces in the crowd the human brain uses those neurons to identify them. The biological neural processing system is what inspired a neural network study [43]. Figure (3.3) shows the biological neuron and its mathematical model.

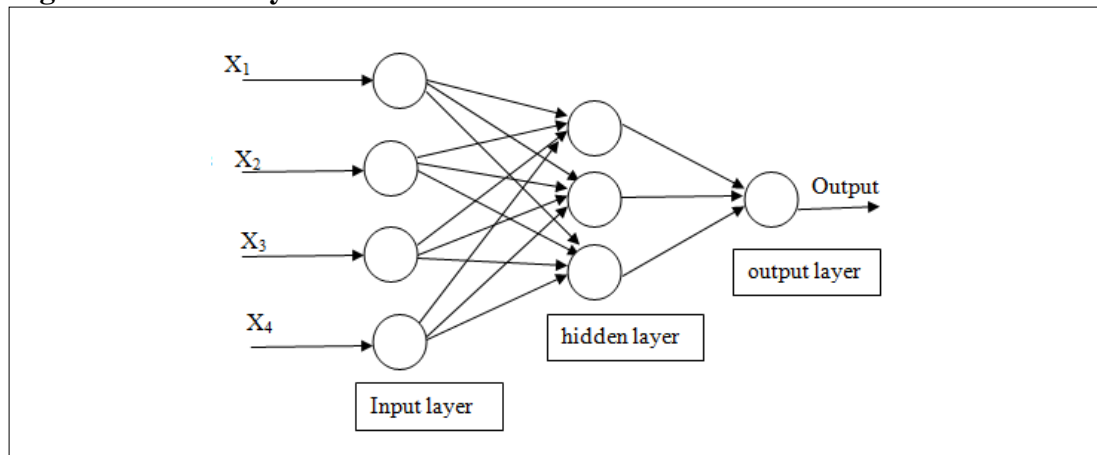
Figure 3.3: An illustration of a biological neuron and its mathematical model



Source: N. Revathy, T.Guhan, "Face Recognition System Using Back Propagation Artificial Neural Networks".

Neural network is adds the learning ability into the network component, acting as an adaptive retrieval system, where the network weights represent flexibility. Many types of neural networks exist, some of them are fairly simple and some are very difficult and complicated [43]. In this thesis, we are focusing more on the back propagation neural network. Figure (3.4) shows a general multilayer neural network.

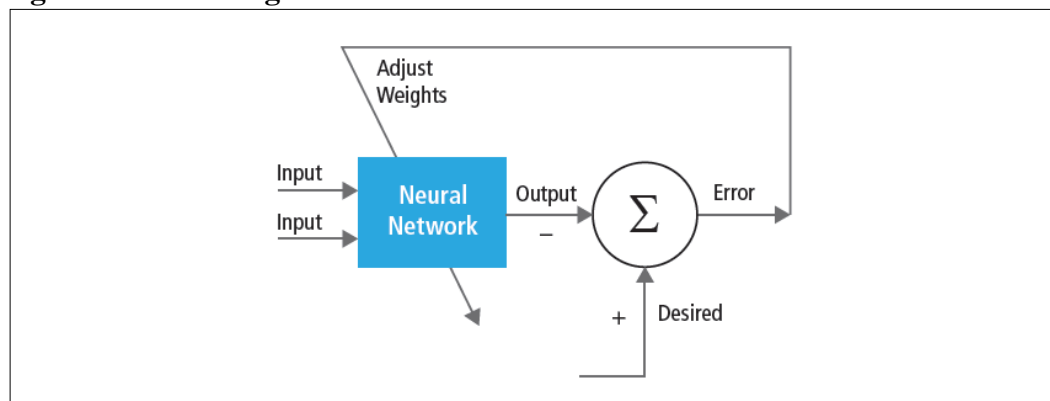
Figure 3.4: Multilayered neural network



Source: S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition", IP Group, Cadence.

Whereas figure (3.5) shows the general training of a neural networks.

Figure 3.5: Training of neural networks

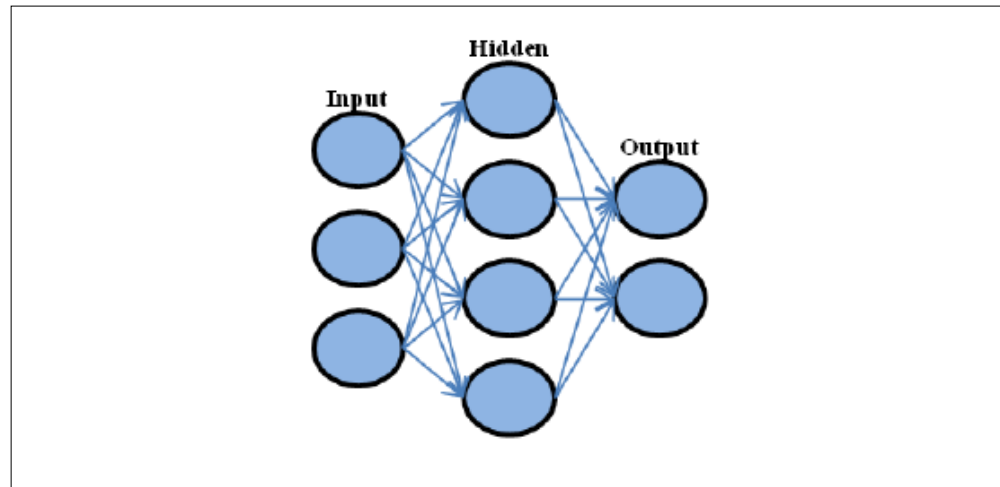


Source: V. Balamurugan, M. Srinivasan, Vijayanarayanan.A , "A New Face Recognition Technique using Gabor Wavelet Transform with Back Propagation Neural Network".

3.5.1 BPN

The Back-Propagation network is a learning algorithm that commonly used for the training of a multilayer perceptrons (MLP). The input, hidden and output layers are formed by the sensory units of the MLP network. The way the information move in a back-propagation algorithm is in two directions, a forward sweep and a backward sweep. The forward sweep is the description of the network with a fixed weight from the input to the output layer. On the other hand, the backward sweep is the description of the network from the output to the input layer [47].

Figure 3.6: Three layered BPN



Source: S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition", IP Group, Cadence.

3.5.1.1 BPN Advantages

- 1) A computationally efficient method.
- 2) Minimization of the output's total squared error
- 3) Face images that are non-linear can be recognized very easily when the BPN methods is combined with PCA method.

3.5.1.2 BPN Disadvantages

- 1) Very time consuming.
- 2) Less accurate than other NN techniques.

3.5.1.3 Mathematical Representation of BPN

A controlled learning approach with an objective output vectors that are classified in the system is used by the BPN algorithm. BPN starts the learning process with arbitrary input variables of a model. The standard summation of products is found by using the net total input which is defined as:

$$net_j = \sum_{i=0}^i x_i \cdot w_{ij} \quad (3.28)$$

Where x equals the input vectors, and w equals the weight vectors.

Mostly, an activation function rule is used for measuring the output value that will be forwarded to the other parts, and the output value of it is called the activation value. Sigmoid function is usually used by the Back-propagation algorithm as the start function, and it is defined by the equation (23), where the hidden layer of the activation function is represented as $f(net_j)$.

$$f(net_j) = 1 / [1 + \exp(-net_j)] \quad (3.29)$$

Below the equation is used to calculate the net total output:

$$net_k = w_k \cdot f(net_j) \quad (3.30)$$

After that, these two equations are used to calculate the yields of both hidden and output layer:

$$O_j = f(net_j) \quad (3.31)$$

$$O_k = f(net_k) \quad (3.32)$$

Until the pattern of the output is generated, the input pattern is in fact spread out through the entire network. The generalized delta rule is used by the BPN mostly in order to find out the error. The hidden layer error units are represented by δ_j and are described as the equation below and the error throughout all the output layer units is represented by δ_k and is described in equation below:

$$\delta_j = f'(net_j) \cdot \sum_{k=1}^k \delta_k \cdot w_{kj} \quad (3.33)$$

$$\delta_k = (t_k - o_k) f'(net)_k \quad (3.34)$$

In order to decrease the error signals each unit changes the direction of its input weights to some extent which continues for the next patterns as well. The unit's weight change in the hidden layer after applying a learning rate η , is determined by:

$$\Delta w_{ij} = \eta \delta_j x_i \quad (3.35)$$

While the weight change for the output layer can be determined from:

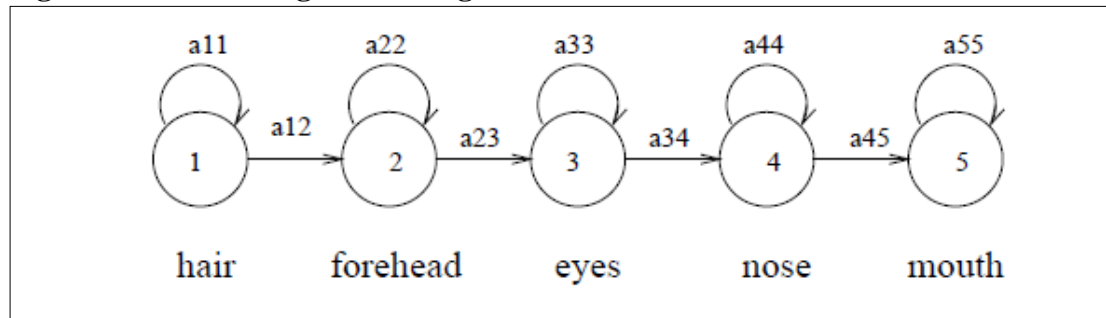
$$\Delta w_{kj} = \eta \delta_k f(\text{net}_j) \quad (3.36)$$

Updating the weight values is the final step in back-propagation and is determined by the equation below [48].

$$\Delta w_{ij}(n + 1) = \eta \delta_j o_i + \alpha \Delta w_{ij}(n) \quad (3.37)$$

3.6. HMM

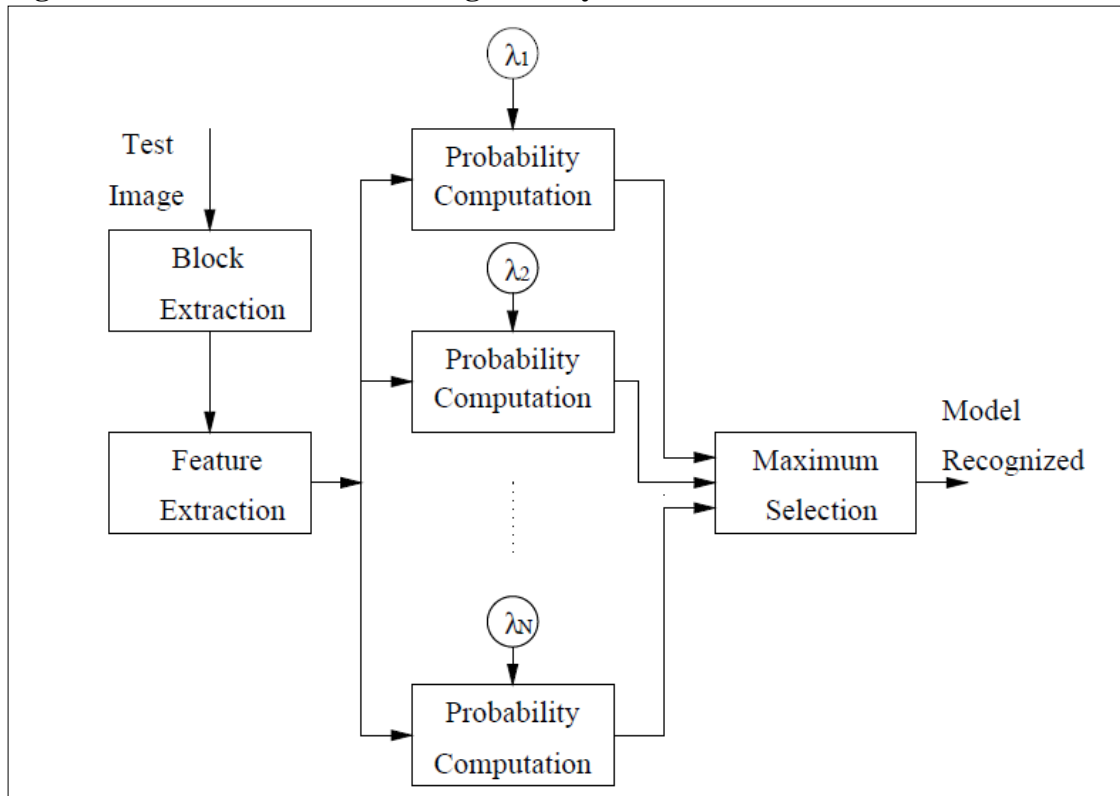
Figure 3.7: Face recognition using HMM



Source: A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition".

Hidden Markov Models (HMMs) have proven to be a successful tool to use when the used data is one-dimensional just as in speech recognition. It is very fast and efficient method especially when used for face recognition. However extended it to two-dimensional data has shown to be computationally very difficult. Many successful works have been achieved using HMM in the field of face recognition later. Since HMMs are a group of statistical models that are used to describe the statistical properties of a signal [49]. Figure (3.8) shows the structure of a face model in a HMM based face recognition system.

Figure 3.8: HMM based face recognition system



Source: A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition,".

3.6.1 Advantages of HMM

- 1) Very flexible.
- 2) Provides better compression.

3.6.2 Disadvantages of HMM

- 1) Does not take important feature parts of the face into account.
- 2) Scanning methodology is fixed.

3.6.3 Mathematical Representation of HMM

HMM has two corresponding steps, the first is hidden Markov string that has a fixed number of states. The second is a number of probability density functions that are connected with every state. HMM which can be represented as λ , is defined by the following components:

- a) The number of states in the model is defined as N . The set of states is defined as S where $S = \{S_1, S_2, \dots, S_N\}$. T is the number of frames, and $q_t \in S, 1 \leq t \leq T$ defines the state of the model at a time t .
- b) The number of different observation symbols is represented as M . The set of all possible observation symbols is represented as V , where $V = \{v_1, v_2, \dots, v_M\}$.
- c) The transition matrix of the state probability is represented as \mathbf{A} , where $\mathbf{A} = \{a_{ij}\}$ and

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i] \quad 1 \leq i, j \leq N \quad (3.38)$$

including these two constraints,

$$0 \leq a_{ij} \leq 1, \\ \sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N;$$

- d) The probability matrix of the observation symbol is represented as \mathbf{B} , where $\mathbf{B} = \{b_j(k)\}$ and

$$b_j(k) = P[\mathbf{O}_t = v_k | q_t = S_j] \quad (3.39)$$

where

$$1 \leq j \leq N, 1 \leq k \leq M$$

Also the observation symbol at time t is represented as \mathbf{O}_t .

- e) The probability distribution of the initial state is represented as $\mathbf{\Pi}$, where $\mathbf{\Pi} = \{\pi_i\}$ and

$$\pi_i = P[q_1 = S_i] \quad 1 \leq i \leq N \quad (3.40)$$

For convenience, an HMM is represented by the triplets below:

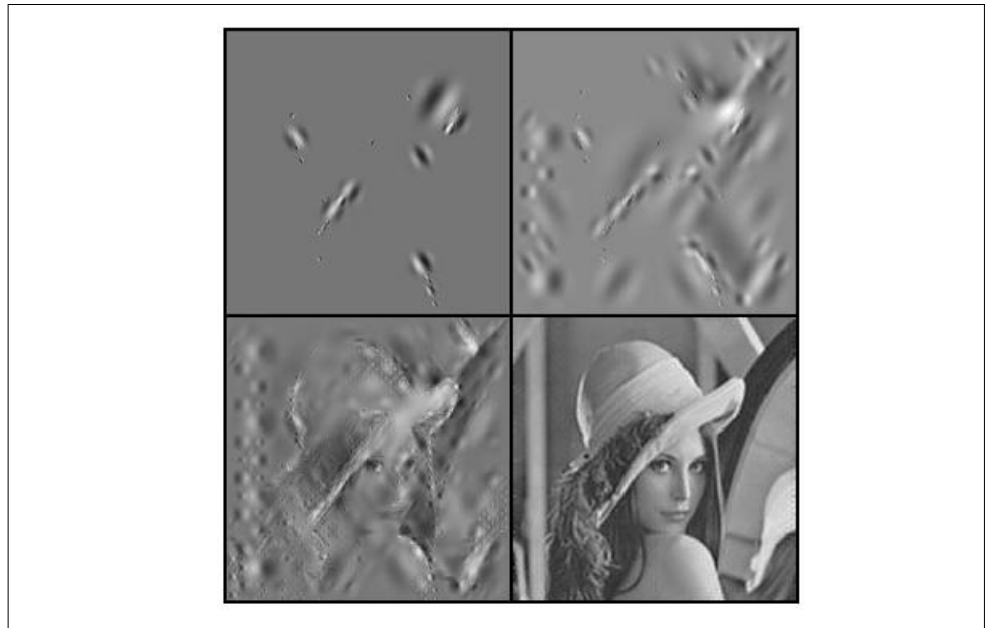
$$\lambda = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi}) \tag{3.41}$$

The description above is a representation of HMM, in which the interpretations are chosen from an alphabet set of distinct symbols as $V = \{v_1, v_2, \dots, v_M\}$ [49].

3.7 GABOR WAVELET TRANSFORM

GWT is a powerful tool that has been used in many applications, like image and video processing, statistical analysis, and telecommunication. In the last few years, lots of attention was directed towards GW because of its ability to represent faces with few two-dimensional GWs. A combination of locally or globally selected points in a specific image is used in Gabor wavelet. The wavelet transform can perform multi-resolution time-frequency analysis [50]. The characteristics of spatial localization, orientation, spatial frequency and face relationship are detained by GWT. Among different wavelets, GWT offers the best resolution in terms of time and frequency which is why it is set to be the best base for extracting local features for a number of reasons, to achieve multi-resolution and multi-orientation [51]. Figure (3.9) shows the GW representation of an image.

Figure 3.9: The GW representation of an image



Source: http://www.anilaggrawal.com/br/vol_002_no_002/reviews/ts/page003.html.

3.7.1 GWT Advantages

- 1) Recognizes face images with different illumination, pose and expression.
- 2) Is not effected by change, rotation and size enlargement.
- 3) Generalization and abstraction of training data is possible.
- 4) Keeps the relationships between the pixels that are close to each other.
- 5) Fast recognition and low computational cost.

3.7.2 GWT Disadvantages

- 1) Require high computational efforts.

3.7.3 Mathematical representation of Gabor Wavelet

The GWT is usually defined as:

$$\phi_{(x,y)}(z) = \frac{\|k_{x,y}\|^2}{\sigma^2} e^{-\frac{\|k_{x,y}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{x,y}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (3.42)$$

The orientation and scale of GW is defined by x and y , where $z = (a,b)$, $\| \cdot \|$ represents the standard operator and $k_{x,y}$ which is the wave vector, is described by [36]

$$k_{x,y} = k_y e^{i\phi_x} \quad (3.43)$$

$k_y = k_{max}/f^y$ and $\phi_x = \pi x/8$. Where the maximum frequency is identified as K_{max} and the frequency domain that contains the spacing element between kernels in is defined by f .

The kernels in the GW represented equation (32) are all the same since they are all produced by the same source GW, and that is through the wave vector $K_{x,y}$. A Gaussian packet is what produces the kernels and the complex plain wave. However the wave part of the kernel is determined by the part between the square brackets in equation (32), and the DC value is determined by the second term. Keeping in mind that when the percentage of the width of a Gaussian window to that of wavelength, determined by the parameter that is σ , that is used the effect of the DC value becomes irrelevant unless they have very large values. Gabor wavelets are valuable when they are used in five

different sizes, $y \in \{0, \dots, 4\}$ and eight different orientations $x \in \{0, \dots, 7\}$ along with these constraint $\sigma = 2\pi$, $K_{max} = \pi/2$ and $f = \sqrt{2}$ [51].

The complication of the image is represented by the GW and it is just like defined by (32). By considering $I(x,y)$ as an image with a grey-level scatter, I represents the convolution of an image and is described by:

$$O_n(z) = I(z) * \varphi_{n,b} z \quad (3.44)$$

In which the convolution operator is represented as, $z = x, y$, and the convolution result that corresponds to the Gabor kernel by orientation n and scale is represented as b . $O_n(z)$. This produces:

$$S = \{O_n, : n \in \{0, \dots, 7\}, b \in \{0, \dots, 4\}\} \quad (3.45)$$

$I(z)$ is the GW representation of the image and is formed by the equation (36) [52].

4. RESULTS AND DISCUSSION

In this section, experiments are done and performances of each method are evaluated using MATLAB. The ORL database was used for evaluation purposes. First, a set of test results were obtained by running the methods without the use of Gabor filters. Then, another set of test results were obtained using Gabor filters. The assessment was done by comparing the test results obtained from each test group.

4.1 THE ORL DATABASE

In the experiments the Olivetti Oracle Research Lab (ORL) was used which is also known as the AT&T database. The database includes 10 images with different poses, facial expressions, lighting and glasses, for each of the 40 different people. The taken images are all in a frontal pose with some side movement and dark background. The dimensions of all the images are 92×112 . Figure (4.1) shows the ORL database.

Figure 4.1: ORL Database



Source: Y. Jin, Q. Ruan, "Face Recognition Using Gabor-Based Improved Supervised Locality Preserving Projections".

4.2. RESULTS

4.2.1 Matlab

Experiments were performed using MATLAB R2013a. MATLAB is a programming language that was developed by MathWorks. The MATLAB platform is optimized for solving engineering and scientific problems.

4.2.2 Results of face recognition methods without Gabor wavelet

In this section, each method was performed and tested independently and results are presented. The used methods are executed without the use of Gabor filters.

In order to assess the performance of PCA, KPCA, LDA, and KFA, the codes acquired and used were from the PhD (Pretty helpful Development functions for) face recognition toolbox [53, 54]. Results of each method were obtained by running each code. The recognition rate of PCA was 66.07 percent. Whereas the recognition rate of LDA was 86.07 percent which is a lot better than PCA. After that, face recognition was performed using kernel methods along with PCA and LDA. KPCA did not provide a better recognition performance. On the contrary, the recognition rate had dropped to 49.29 percent. The same thing applied to KFA, but the performance did not drop as much as the KPCA. The recognition rate with KFA was 85.71 percent which is not a lot than the recognition rate with LDA.

The performance of BPN was tested by acquiring the code from the MathWorks website. The results have shown a better performance than the previously used methods. The recognition rate using BPN methods was 95 percent which is very high.

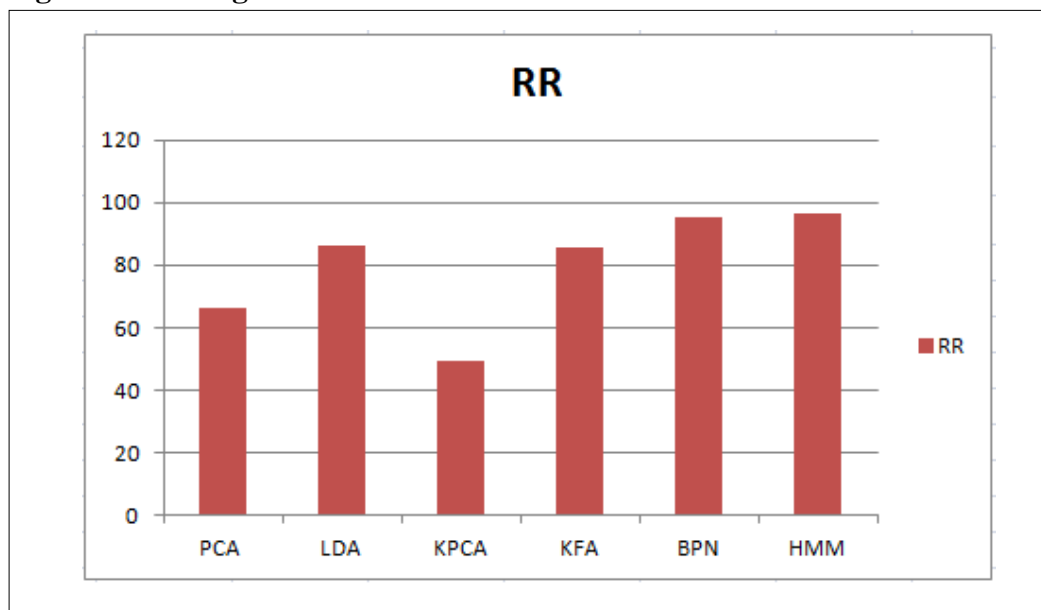
The last method that has been experimented is HMM. The code was also acquired from the Mathworks website. Face recognition using HMM method showed the best performance result amongst the tested methods. The recognition rate was 96.5 percent which is the highest. Table (4.1) illustrates the recognition rates of the tested methods. Recognition using HMM based system showed the best performance and using KPCA the worst.

Table 4.1: Test results of face recognition methods

Method	Recognition Rate (in %)
PCA	66.07
LDA	86.07
KPCA	49.29
KFA	85.71
BPN	95
HMM	96.5

Figure (4.2) demonstrates the recognition rate of the used methods without Gabor wavelet.

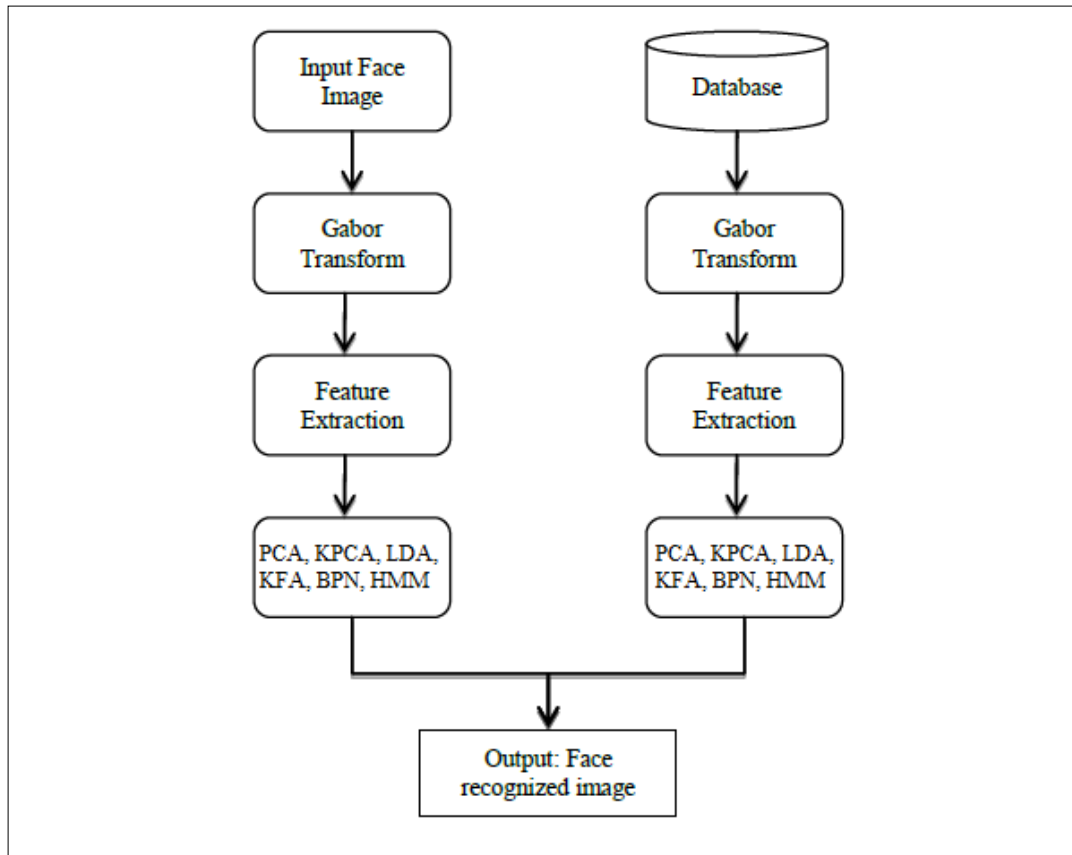
Figure 4.2: Recognition rate without Gabor wavelet



4.2.3 Results of face recognition methods using Gabor wavelet

In this section, Gabor wavelet is used with the same methods that were tested in the previous section and the results obtained are presented. Gabor wavelet is a known and powerful tool for face recognition as stated previously and it has been used in many applications because it increases the performance dramatically. Figure (4.3) illustrates the face recognition system using Gabor wavelet with common methods.

Figure 4.3: Face recognition system with Gabor wavelet



Source: V. Balamurugan, M. Srinivasan, Vijayanarayanan.A , " A New Face Recognition Technique using Gabor Wavelet Transform with Back Propagation Neural Network".

The codes for the tested Gabor wavelet based methods were acquired from the PhD tool as well [53, 54]. The first tested method was Gabor based PCA, and the experiment results have shown a great improvement. The recognition rate was 74.17 percent which is a lot higher than the results acquired by running the PCA algorithm alone. As for LDA, Gabor based LDA has also shown a better result with a recognition rate of 93.33 percent. Gabor based KPCA and KFA have also shown a great improvement with recognition rates of 80 percent and 93.33 percent repeatedly. As noticed, Gabor based LDA and KFA have done similar performances with recognition rate of 93.33 percent.

The code for testing the performance of Gabor based BPN was acquired from the MathWorks website. The test results came with a recognition rate of 97.90 percent which is also high compared to the test result of BPN algorithm without Gabor wavelet.

Gabor based HMM algorithm has also showed an improvement in performance with a recognition rate of 96.8 percent [56] which is slightly higher than the HMM based test result, yet considered as an improvement.

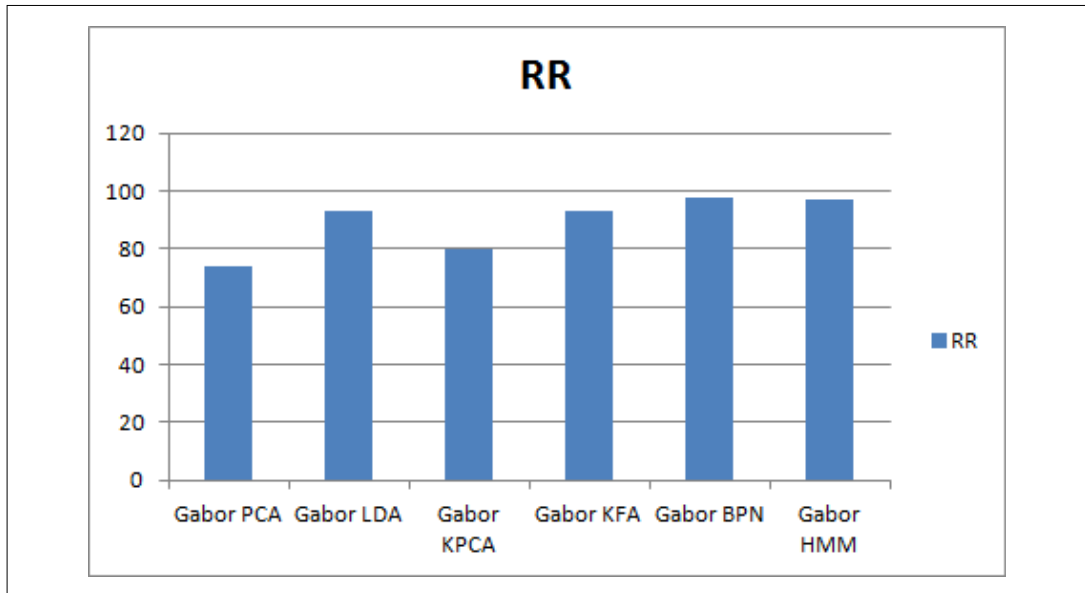
Table (4.2) shows the recognition rates using Gabor filters.

Table 4.2: Test results of face recognition methods with Gabor wavelet

Method	Recognition Rate (in %)
Gabor-based PCA	74.17
Gabor-based LDA	93.33
Gabor-based KPCA	80
Gabor-based KFA	93.33
Gabor-based BPN	97.9
Gabor-based HMM	96.8 [56]

Figure (4.4) illustrates the recognition rate of the used methods using Gabor wavelet.

Figure 4.4: Recognition rate using Gabor wavelet



4.3 DISCUSSION

Experimental results obtained from the previous two sections have showed the effectiveness of using Gabor filters with the most commonly used face recognition methods. As seen, the performances of the tested methods have increased dramatically in some case more than the others. Yet one thing was obvious in all of the test results obtained and that is Gabor filters increases the recognition rate when used. For Gabor-based PCA the increase was by 8.1 percent more than the conventional PCA method. Gabor-based LDA method showed an increase of 6.63 percent than the usual LDA methods. KPCA probably showed the best increase result with a difference of 30.71 percent than the original KPCA methods. KFA increased by 7.62 percent. BPN increased by 2.5 percent, and HMM increased by 0.03 percent which is the least change amongst all the used methods. Table (4.3) shows a comparison between the test results obtained without the use of Gabor filters and the test results obtained using the Gabor filters.

Table 4.3: Comparing test results of face recognition methods with and without Gabor wavelet filters

Methods	Recognition rate (in %) without use of Gabor filters	Recognition rate (in %) using of Gabor filters
PCA	66.07	74.17
LDA	86.07	93.33
KPCA	49.29	80
KFA	85.71	93.33
BPN	95	97.9
HMM	96.5	96.8 [56]

Figure (4.5) shows a comparison between recognition rate without the use of GW and with the use of GW.

Figure 4.5: A comparison between recognition rates



In the first test set, without Gabor filter, HMM achieved the best performance with a recognition rate of 96.5 percent. Yet in the second test set, with Gabor filter, BPN performed better than the rest of the methods with a recognition rate of 97.6 percent.

5. CONCLUSION

Face recognition has been the most challenging yet most important application in computer vision. Various researches have been done in face recognition area in the past 30 years and yet, still no one universal system has managed to solve all the encountered obstacles. Researches will most likely continue in the coming 30 years to achieve the best system.

In this thesis, the problem with face recognition has been addressed. Common methods of face recognition were explained in detail along with their advantages and disadvantages.

The aim of this thesis was to show the effectiveness of Gabor wavelet filters when used with the most commonly known methods of face recognition. The experimental results have shown a great increase in the performance of the methods when used with Gabor filters.

REFERENCES

Books

R. Szeliski, "Computer Vision: Algorithms and Applications", September 3, 2010.

Periodicals

- [1] M. Nandini, P. Bhargavi, G.R. Sekhar, "Face Recognition using Neural Networks", International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013.
- [2] Y.B. Wang, T.Y. Zhou, and B.J. Hu, "A Step-wise Refinement Algorithm for Face Recognition Based on Blocking Wavelet Transforms", Journal of Information Hiding and Multimedia Signal Processing, Volume 6, Number 3, May 2015.
- [3] J.-K. Fung, "Face Recognition by Eigenface and Elastic Bunch Graph Matching", Department Of Computer Science and Engineering, The Chinese University of Hong Kong.
- [4] Z. Riaz, M. S. Sarfraz and M. Beetz, "Towards Unconstrained Face Recognition Using 3D Face Model, Biometric Systems", Design and Applications, 2011.
- [5] Shang-Hung Lin, "An Introduction to Face Recognition Technology", Informing Science Special Issue on Multimedia Informing Technologies-Part 2, volume 3 No.2, 2000.
- [6] R. Bhatia, "Biometrics and Face Recognition Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 5, May 2013.
- [7] O'Toole, A. J., Jiang, F., Roark, D., and Abdi, H., "Predicting human face recognition", Face Processing: Advanced Methods and Models, Elsevier, 2006.
- [8] O'Toole, A. J., Phillips, P. J., Jiang, F., Ayyad, J., P'enard, N., and Abdi, H., "Face recognition algorithms surpass humans matching faces over changes in illumination". IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.
- [9] R. Szeliski, "Computer Vision: Algorithms and Applications", September 3, 2010.
- [10] Sinha, P., Balas, B., Ostrovsky, Y., and Russell, R. , "Face recognition by humans: Nineteen results all computer vision researchers should know about", Proceedings of the IEEE, 2006.
- [11] Kanade, T., Cohn, J., Tian, Y.-L. "Comprehensive database for facial expression analysis", Proceedings of International Conference on Face and Gesture Recognition, 2000.

- [12] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, 1997.
- [13] J. Yang D. Zhang, A. F. Frangi, and J. Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, 2004.
- [14] Luis Torres, "Is There Any Hope For Face Recognition?", Technical University of Catalonia, Barcelona, Spain.
- [15] R.S. Malpass and J. Kravitz, "Recognition for faces of Own and other Race", *Journal of Personality and Social Psychology*, Vol. 13, 1969.
- [16] M. D. Hegde, Sayeesh, "Face expression recognition using score level fusion method", *International Journal of Multidisciplinary Research and Development* 2015.
- [17] W.-L. Chao, "Face Recognition", GICE, National Taiwan University.
- [18] V.V. Starovoitov, D.I Samal1, D.V. Briiliuk, " Three Approaches for Face Recognition", The 6-th International Conference on Pattern Recognition and Image Analysis, Velikiy Novgorod, Russia, October 21-26, 2002.
- [19] I. Marques, "Face Recognition Algorithms", June 16, 2010.
- [20] Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. , "Face recognition: A literature survey", *ACM Computing Surveys (CSUR)*, 2003.
- [21] B. S. Bargavi, C. Santhi, " Global and Local Facial Feature Extraction using Gabor Filters", *International Journal of Science, Engineering and Technology Research (IJSETR)*, Volume 3, Issue 4, April 2014.
- [22] M. Ballantyne, R. Boyer, L. Hines, "Woody Bledsoe: His Life and Legacy," *AI Magazine* 17, no. 1:7-20, Spring 1996.
- [23] A. J. Goldstein, L. D. Harmon, and A. B. Lesk, "Identification of Human Faces," *Proc. IEEE*, Vol. 59, May 1971.
- [24] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Pattern Analysis and Machine Intelligence*, vol. 12, 1990.

- [25] M. A. Turk , A. P. Pentland, "Face Recognition Using Eigenfaces," Proc. IEEE, 586-591, 1991.
- [26] U. Bakshi, R. Singhal," A New Approach of Face Recognition Using DCT, PCA, and Neural Network in MATLAB", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 3, May-June 2014.
- [27] Jayanthi T and Aji, "Face Recognition Using Kernel PrincipalComponent Analysis", Advances in Vision Computing: An International Journal (AVC) Vol.1, No.1, March 2014.
- [28] A. Nichal1, H. Jagtap, S. Ingale, N.Patil, "Face Recognition using PCA and Eigen Face Approach", International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, Vol. 2, Issue 5, May 2014.
- [29] L. Chen, H. M. Liao, M. Ko, J. Lin, G. Yu, "A new LDA-based face recognition system which can solve the small sample size problem", The Journal of the Pattern Recognition society",1713-1726, 2000.
- [30] J. A. Dargham, A. Chekima, E. G. Mounq, " Fusion of PCA and LDA Based Face Recognition System", International Conference on Software and Computer Applications (ICSCA), vol. 41, 2012.
- [31] P. Dhoke , M.P. Parsai, " A MATLAB based Face Recognition using PCA with Back Propagation Neural network", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2014.
- [32] O. M. Parkhi, A. Vedaldi, A. Zisserman, "Deep Face Recognition", 2015.
- [33] S. K Jameel, "Face Recognition System Using PCA and DCT in HMM", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 1, January 2015.
- [34] S. Raut and S. H. Patil," Face Recognition Using Maximum Confidence Hidden Markov Model", International Journal of Advances in Engineering & Technology, May 2012.
- [35] H. Cho, R. Roberts, B. Jung, O. Choi, and S. Moon, "An Efficient Hybrid Face Recognition Algorithm using PCA and GABOR Wavelets", 2014.
- [36] Y. Jin, Q. Ruan, "Face Recognition Using Gabor-Based Improved Supervised Locality Preserving Projections", Computing and Informatics, Vol. 28, 2009.

- [37] S. Suganya and D. Menaka, "Performance Evaluation of Face Recognition Algorithms", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 1, January 2014.
- [38] Y. Bakhshi, S. Kaur, P. Verma, "A Study based on Various Face Recognition Algorithms", International Journal of Computer Applications (0975 – 8887), Volume 129 – No.13, November 2015.
- [39] A. Nichal, H. Jagtap, S. Ingale, N. Patil, "Face Recognition using PCA and Eigen Face Approach", International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, Vol. 2, Issue 5, May 2014.
- [40] H. Kadiya, "Comparative Study on Face Recognition using HGPP, PCA, LDA, ICA and SVM", International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064, Volume 2 Issue 1, January 2013.
- [41] V. Solunke, P. Kudle, A. Bhise, A. Naik, Prof. J. R. Prasad, "A Comparison between Feature Extraction Techniques for Face Recognition", International Journal of Emerging Research in Management & Technology, February 2014.
- [42] B. Schölkopf A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem", Technical Report No.44, December 1996.
- [43] M.-H. Yang, "Face Recognition Using Kernel Methods", Advances in neural information processing Systems, Vol. 14, 2002.
- [44] N. Revathy, T. Guhan, "Face Recognition System Using Back Propagation Artificial Neural Networks", International Journal of Advanced Engineering Technology, Vol. III, Issue I, January-March, 2012.
- [45] S. Hijazi, R. Kumar, and C. Rowen, "Using Convolutional Neural Networks for Image Recognition", IP Group, Cadence.
- [46] V. Balamurugan, M. Srinivasan, Vijayanarayanan.A, "A New Face Recognition Technique using Gabor Wavelet Transform with Back Propagation Neural Network", International Journal of Computer Applications (0975 – 8887) Volume 49– No.3, July 2012.
- [47] U. Chaudhary, C. Mubarak, A. Rehman, A. Riyaz, S. Mazhar, "Face Recognition Using PCA-BPNN Algorithm", International Journal of Modern Engineering Research (IJMER), Vol.2, Issue.3, May-June 2012.
- [48] P. Dhoke, M.P. Parsai, "A MATLAB based Face Recognition using PCA with Back Propagation Neural network", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 8, August 2014.

- [49] A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '98), pp. 2721–2724, Seattle, Wash, USA, May 1998.
- [50] B. Sharma, R. Rana, R. Mehra, "Face Recognition Using Gabor Wavelet for Image Processing Applications", Association of Computer Electronics and Electrical Engineers, 2013.
- [51] Thomas, L. L. , Gopakumar, C. and Thomas, A. A."Face recognition based on Gabor Wavelet and Backpropagation Neural Network", 2013.
- [52] Liu, C.and Wechsler, H. "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition", 2002.
- [53] V. Štruc and N. Pavešić, "The Complete Gabor-Fisher Classifier for Robust Face Recognition", EURASIP Advances in Signal Processing, pp. 26, 2010, doi=10.1155/2010/847680.
- [54] V. Štruc and N. Pavešić, "Gabor-Based Kernel Partial-Least-Squares Discrimination Features for Face Recognition", Informatica (Vilnius), vol. 20, no. 1, pp. 115–138, 2009.
- [55] V. Balamurugan, M. Srinivasan, Vijayanarayanan.A , " A New Face Recognition Technique using Gabor Wavelet Transform with Back Propagation Neural Network", International Journal of Computer Applications (0975 – 8887) Volume 49– No.3, July 2012.
- [56] P. Nicholl, A. Amira, D. Bouchaffra, and R. H. Perrott, "A Statistical Multiresolution Approach for Face Recognition Using Structural Hidden Markov Models", EURASIP Journal on Advances in Signal Processing, Volume 2008, Article ID 675787, 13 pages
- [57] N. Aronszajn, "Theory of Reproducing Kernels", Transaction of the American Mathematical Society, Volume 68, Issue 3, May 1950, 337-404.

Other sources

J G Daugman, "Computer Vision", Notes.

https://github.com/edwardtoday/PolyU_MScST/tree/master/COMP5422/Project/RefProj/PhD_tool

<https://www.mathworks.com/matlabcentral/fileexchange/28679-bpn-based-face-recognition>

<https://www.mathworks.com/matlabcentral/fileexchange/37351-face-recognition-software>

<https://www.mathworks.com/matlabcentral/fileexchange/48591-face-recognition-biometric-with-wavelet-and-neural-network-matlab-code>