A Combination Approach to Face Recognition

by

Andrew Enughwure

A thesis submitted to the
Department of Computer Science
in conformity with the requirements for
the degree of Master of Science

Bishop's University
Sherbrooke, Quebec, Canada
November 2015

Abstract

The purpose of this research project is to develop a face recognition system capable of outperforming existing solutions in terms of accuracy. Particularly, it's implementation focuses on the PCA, LDA, GaborPCA and GaborLDA appearance-based recognition approaches. Methods of combinations are developed in order to combine the outputs of each of these four approaches in hopes that through combination, a face recognition system with a higher level of accuracy will be achieved.

Acknowledgement

I would like to acknowledge my supervisor, Dr. Madjid Allili for his input in this project. His knowledge on the subject matter and his directions helped strongly in shaping this project. Also Dr. Nelly Khouzam and the Bishop's computer science department, who in one way or another assisted me during my time at the University. Finally, the ITS Helpdesk. Thanks for resolving my computer related issues each time I came by. You guys rock!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Face recognition is a subject that garners automatic interest once mentioned. As the reader would agree, there is a degree of excitement induced by the anticipation of learning anything related or closely related to the subject matter. Hollywood and television has for decades fueled our collective curiosity on face recognition, and it is believed that this project will shed some light on the mysteries surrounding the subject. So what is face recognition? Simply described, it is the ability of a computer system to recognize a test face in a given image, video frame or database of images. In face recognition, computer scientists try to imitate the brain's ability to accurately recognize a face which it has previously seen. This can be at various viewing angles as well as in poor lighting. They also try to imitate the brain's ability to recognize a face based on faces it has previously seen. For example, being able to recognize a friend's brother because of their family resemblance. How to accurately recognize patterns and faces however, has been a topic of hot debate and research. Over time, these researches has led to the development of the various approaches and methods of face recognition which will be discussed in this thesis.

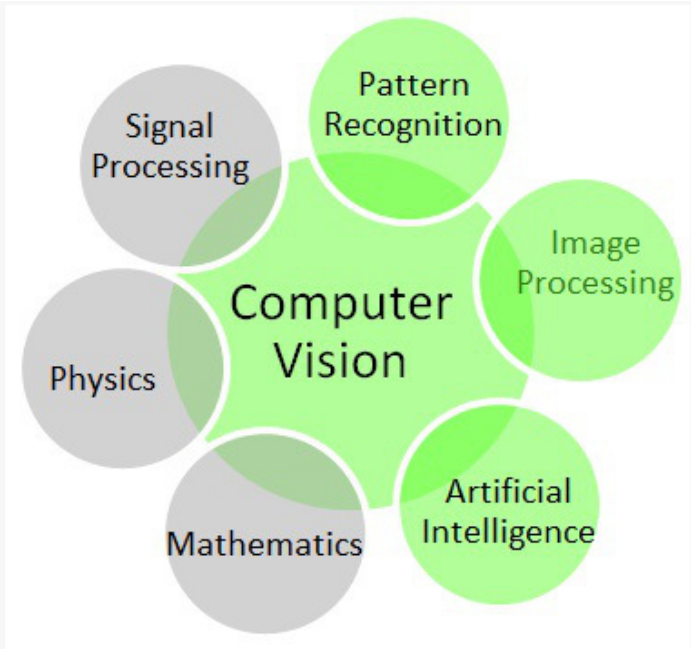


Figure 1.1: Fields of Computer Vision

The image seen in Figure 1.1 displays various fields of computer vision which include pattern recognition and image processing. These fields can be considered as abstractly related because usually, advances in one field could potentially lead to advances in other fields as well. Developing a successful face recognition system requires a cumulative knowledge from all of these fields.

1.1 Overview

This research project aims to understand completely the domain of face recognition with respect to the various face recognition approaches currently existing. Following this research, it hopes to build a face recognition system which utilizes multiple face recognition methods to identify a given test image. The focus of this research will be on how to make decisions on what the identified image should be based on a ranking system generated through the implementation of the multiple methods.

1.1.1 Chapter Summaries

This thesis has been sectioned into five chapters. The paragraphs below give a quick summary of what the reader can expect in each of this chapters.

Chapter one is the introduction. It introduces, face recognition to the reader. It lays the foundation for this thesis by explaining why it is an interesting area of research and also explains what the reader will gain by investing his time in reading through the entirety of this thesis. Finally, it gives a short summary of what the reader is to expect in each of the existing chapters in this thesis.

Chapter two is the background chapter. As the name suggests, it gives a detailed background of face recognition. It's history and origin story. As can be imagined, a face recognition system cannot be built without an understanding of the associated algorithms; As such, a literature review is done which provides information on existing algorithms for face recognition. The review provides information on the most popular approaches to face recognition; stating their advantages and disadvantages. In the latter sections of the chapter, the algorithms which have been selected for the implementation of the face recognition system are mentioned (they are discussed extensively in the methodology chapter), also the face database which was selected for implementing the face system is mentioned and finally, the tools (hardware and software) required to be able to develop the face recognition system are also discussed.

Chapter three is methodology. In this chapter, the face recognition algorithms which were selected for the implementation of the face recognition system are discussed in-depth. Following the in-depth analysis, methods of combination are discussed with the objective of building a face recognition system with higher accuracy. An important concept in image processing is pre-processing. This chapter looks at a technique to correct an image before it is processed. Finally, the techniques for evaluating the outputs of individual face recognition systems are discussed.

Chapter four is for results and analysis. This chapter focuses on the results obtained after implementing and testing the face recognition system. It is split into two sections. The first section looks at the implementation results for the individual face recognition methods, and the second section looks at the implementation results for combined methods. It gives an analysis of the system results and as well as the system performance.

Chapter five is the conclusion. In addition to providing a final summary of the thesis as a whole, the conclusion will also provide the reader with the opinion of the writer as well as discuss future works related to this thesis.

1.2 Chapter Summary

This introductory chapter gave the reader an overview of what to expect in this thesis. It is hoped that by reading it, the reader is eager to find out more about face recognition. It is believed that at the end of this thesis, the reader will have a solid knowledge of face recognition; how it is implemented, the various ways it can be beneficial to society, and most importantly, what the future holds for face recognition.

Chapter 2

Background

In simple terms, computer vision can be described as the process of teaching a computer to see. The first attempt to teach a computer to see was made in MIT in 1966 by Seymour Papert. In his project titled "*The Summer Vision Project*" [1], he attempted to use the summer workers at MIT to participate in the construction of a pattern recognition system. The project was not a success. It exposed the complexities of teaching a computer to see as humans do. This failure however, opened up a whole new field of research for computer scientists as they attempted through algorithms, theories and hypotheses to solve the enigma that is computer vision.

In researching face recognition, other areas of image processing such as face detection have benefited. Modern cameras now come equipped with face detection algorithms which are capable of identifying a face before a picture is taken. This leads to higher accuracy and better pictures for the photographer. Of course, in some face recognition systems, face detection is an integral step in the process of recognition.

2.1 History of face recognition

When the idea of computer vision started spreading following the MIT project in 1966, it attracted lots of brilliant minds, each proposing different theories that achieved various degrees of success. However, it wasn't until 1988 when Kirby and Sirovich [2] developed the

principal components analysis (PCA) that face recognition really took off.

Prior to PCA, face recognition had been done by detecting and analyzing features of the face such as the eyes, nose and mouth. These methods required the extraction of face features from a test image and comparing with similar features from other images in a given database of images. However, the error rates for these approaches were high due mostly to alignment issues. The input image had to be exactly aligned as slight variations in alignment of the face as well as the face pose within the input image could throw off calculations in the face recognition process, resulting in high levels of failure.

What Kirby and Sirovich proved with PCA was that instead of dividing the input image into sections and dealing with local features only, the entire input image could be used without division. They made use of statistical methods whose results enabled them to decide which pixels from the image being processed were irrelevant for the recognition process. They then calculated Euclidean distances [3] of the input image and the images within the image database in order to find a match. Their approach was highly appreciated by the science community and it re-sparked the dwindling interest in face recognition. As a result of their work with PCA, other algorithms such as LDA, MPCA, and ICA were developed,

2.2 Importance of face recognition

We have discussed so far that face recognition has in some way been in existence for quite a while, with brilliant scientists dedicating enormous amounts of their precious time in research in this field. However, the question that could be asked is why? What makes face recognition so important that people from all works of life, governments and companies from various sectors of industry dedicate resources to this area of computing? Below are listed some of the benefits of face recognition which hopefully answers the question.

1. **Security:** We live in an age where security is a major issue for most nations. From local neighborhood crimes to global crimes such as terrorism, the need to have an effective tool to fight against these crimes has never been more needed. As such, face recognition can play an active role in the fight against crime.

2. **Human Computer Interaction (HCI):** This term encompasses a lot of fields in computing including robotics, and gaming. Facial recognition is crucial in robotics where the aim is to develop robots that can recognize objects and people. Also in gaming with the introduction of cameras for motion sensing, the facial recognition could play a massive part in future gaming.

3. **Industrial Sectors:** Imagine a world, where a person can walk into a bank, is identified by a face recognition system and before he gets to the customer service, his information is already loaded. This would ease the transaction process while also ensuring security against fraud in the sense that he is facially verified before he can make a transaction from his bank account. Such scenarios can also be transferred to other sectors like marketing where specific products can be advertised to targeted audiences because a facial recognition system running in background has recognized and matched the profile of the people around the advertising system.

2.3 Approaches for Face Recognition

Over the decades, many algorithms and approaches to face recognition have been developed, some more successful than others. However, the two most popular approaches are the local geometric approach and the global appearance-based face recognition approach. This project will utilize the appearance-based methods for implementation however, an overview of the geometric approach will also be given. Please note that not all existing methods will be listed as that would be an exhaustive process. Instead a select number of methods will be discussed. The following subsections describe them in more details.

2.3.1 Geometric Approach

As stated earlier, the geometric approach extracts common features of the face, such as position of the eyes, nose and mouth in any given image and using the coordinates to determine whether two faces were identical. Some of these attempts involved the use of simple image processing technique such as edge detection, signatures etc. For example in [4], edge maps were extracted first from input images and then matched to big templates, with possible variations in position and size. Faces were then confirmed to be present by searching for edges at approximated positions of some features like the eyes and mouth.

The following methods are some of the techniques which can be used with the geometric approach.

Canny Edge Detection

In a digital image, an edge is a point in the image where the brightness changes sharply. The canny edge detector [4] was developed by John F. Canny in 1986. It is used to detect a wide range of edges in images. Below are some of the attributes of the Canny Edge detector:

- **Good Detection:** In determining a true or false edge, thresholds are required. The Canny edge detector can be fine-tuned with the right threshold to provide good edges on average.
- **Noise sensitivity:** The Canny edge detector eliminates or reduces noise that could corrupt results.
- **Orientation sensitivity:** The Canny edge detector accurately detects not just the edge magnitude, but also the edge orientation. Which can be used in post processing to connect edge segments and in turn suppress non-maximum edge magnitude.
- **Speed and efficiency:** The Canny edge detector allows for recursive implementation which improves efficiency.

Gabor Wavelets

In dealing with problems involving variations in illumination and facial expressions, it is believed by some that extracting the local features of the face first is a great way of resolving these problems. Many believe wavelets to be the best method for feature extraction due to it's space-frequency localization characteristics.

There are several reasons why Gabor wavelets [5] seem to be an ideal choice. These include:

- Biological Motivation: The shapes of the wavelets are shaped liked the simple cells found in the receptive fields of the primary visual cortex.
- Empirical motivation: Wavelets have been used for handwritten numeral recognition, fingerprint recognition and texture segmentation. As such it is believed it can be useful for image pattern recognition also.

In their work on Dynamic Link Architechture(DLA) [6], Lades et al. pioneered the use of Gabor wavelets for face recognition. In their system, features were extracted at deformable nodes from faces which were represented by rectangular graphs. This extraction was done using Gabor wavelets which was referred to as Gabor Jets.

DLA was later extended to Elastic Bunch Graph Matching (EBGM) [7] by Wiskott et al. In EBGM, graph nodes are located at a number of facial landmarks in order to map the distinct features of the face. Since the original, other elastic bunch based methods have been proposed.

Gabor wavelets will be used as part of this project and will be analyzed further in a later chapter.

Advantages of Geometric Approach

The major advantage of geometric based approaches is that the extractions of geometric points or features occur before any matching analysis is done. As a result, the approach is relatively robust to position variation. Also, it can be made invariant to size, orientation and or light. Another advantage of this approach is the speed of matching due to the compactness of the representation.

Disadvantages of Geometric Approach

The major disadvantage with the geometric approach is that it is difficult to automatically detect the features of the face. Whoever is implementing the geometric technique for facial recognition has to make arbitrary decisions on which features of the face is most important. This is necessary because if the feature set lacks discrimination ability, no amount of subsequent processing can compensate for that intrinsic deficiency.

2.3.2 Appearance Based Approaches

Where geometric approaches used local features of the face, appearance based approaches use the entire face instead. Since 1988, when Kirby and Sirovich [8] introduced the principal component analysis method, appearance based approaches became the common basis of research into face recognition. This resulted in multiple variations of PCA which will be discussed in this section.

Principal Component Analysis (PCA)

Processing image data can be computationally expensive. This is due to the large amount of data that the computer has to process in a given image. With PCA, it is possible to reduce the dimensional size of the given image. This is done by extracting a subspace of the input data that best describes the given image. Assuming our goal is to reduce the dimensions of a d-dimensional dataset by projecting it onto a k-dimensional subspace (where $k < d$).

It would then be necessary to know what the size of k should be and also if our subspace represents our data well.

Figure 2.1: PCA Graph

The main idea of PCA is to find vectors that best account for variation of face images in entire image space. These vectors are called eigenvectors. Figure 2.1 shows two graphs in the x and y coordinate system. The first graph from the left shows the data in its original form while the second graph shows the de-correlated version of the original data.

PCA is one of the methods that will be used in this project and will be further analyzed in a later chapter.

Independent Component Analysis (ICA)

ICA can be considered as a generalization of PCA. It is a technique for extracting statistically independent variables. While PCA tries to get a representation of the input data based on uncorrelated variables, ICA tries to get a representation of the input data based on independent variables. The concept of ICA was conceived by Pierre Common in his work [9].

To derive the ICA transformation, Comon developed an algorithm which consists of three different operations: whitening, rotation and normalization. These operations are

sometimes applied to ICA components as pre-processing operations in order to make the data ready for use. The whitening operation transforms a random vector X into another vector U which has unit covariance matrix.

Linear Discriminant Analysis (LDA)

Similar to PCA and ICA, LDA is a dimension reduction technique. However, unlike the aforementioned techniques, it takes into consideration the class information of the given data. LDA was developed by Ronald A. Fisher in 1936 [13] and is also sometimes known as Fisher Discriminant Analysis (FDA). It was originally a two class classifier; however it was later generalized as a multi-class tool by C.R Rao in 1948 [14]. The goal with LDA is to project a dataset onto a lower-dimensional space with good class-seperability in order to avoid over-fitting.

The LDA is another method which will be used in this project and will be analyzed further in Chapter 3 which is on Methodology.

Advantages of Appearance Based Approach

Where geometric methods try to search for features of the face, appearance based methods use the entire image. The major benefit to this is that time is not wasted on algorithms which try to locate exactly where the mouth, nose or eye is. The entire image is used and as such, various statistical methods can be implemented on the pixels of the image in order to determine who the person or persons in the image are.

Disadvantages of Appearance Based Approach

Depending on the conditions of the images in use, it is sometimes necessary to perform pre-processing operations such as normalization, histogram equalization and others in order to improve the accuracy of results when the appearance-based approaches are applied to the images.

2.4 Major Challenges

2.4.1 Geometric

The geometric approach involves analyzing each pixel in a given image in order to extract the features of the face image. After which comparisons can be done in order to find a match. This is computationally expensive as well as time consuming. Also if the image in use is not of a certain quality, it can throw errors into the recognition algorithm, resulting in poor performance of the algorithm.

2.4.2 Appearance-Based

With the appearance-based approach, the reduction in dimension allows for faster processing of data. However, the reduction also leads to loss of information. We assume that the information lost is insignificant. The appearance based approach also takes the entire face image into account, thereby removing the need to search for the individual features of the face. Other major challenges are pose, lighting and angular position of the face.

2.5 Classifiers

For each of the methods mentioned in both the geometric and appearance-based approaches, a critical step in determining the matching image for a given test image is the classification step. Whereby based on the results of the methods the correct image is selected. The subsections below discuss the most popular classifiers in use today.

2.5.1 Neural Networks

Neural Networks are systems that can be trained to recognize patterns. An example can be given of a child who is repeatedly shown different breeds of dogs. After a time, he will come to know many breeds of dogs and even when shown a new breed of dog he has not seen before, based on the characteristics of the previous breeds of dogs he has seen, he can make an educated guess that the new image it is seeing is a dog. Neural Networks operate

in similar fashion. The use of Neural networks for face recognition has been shown by [15] and [16]. In figure 2.2, the general structure of a Neural Network is given.

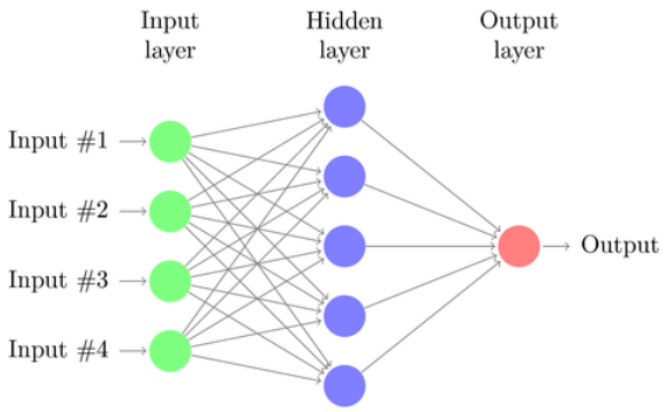


Figure 2.2: Structure of a Neural Network

In geometric face recognition, a Neural Network can be used to determine if the set of features extracted from a given image is a mouth, nose or eyes. Neural Networks is multi-purpose, and can also be used in other areas of image processing and computer vision.

2.5.2 Nearest Neighbor

This classifier is by far the most popular method of classification. It compares test and database images by evaluating their distances in the image subspace. The smaller the distances between images, the more alike they are. The most popular type of distance associated with the nearest neighbor classifier is the Euclidean distance. However, the Mahalanobis distance [17] is also commonly used.

2.6 Face Database

In face recognition and face detection research, it is common among researchers to use a set of standardized and widely accepted face database for their research. This is so that their work can be classed as credible and also so that other members of the research community can replicate their work if they choose to. Below is a list and brief description of existing

face databases.

1. **FERET**

   The FERET program was an initiative by the Department of Defense. Its aim being to develop a face recognition system that could assist officials in their fight against crime. The program ran from 1993 to 1997. Its face database contains 1000 people with male and female faces. There are 14051 images with different views. It can be used for testing time related face algorithms because the time range of each persons image taken was at least a year.

   Link: `http://www.itl.nist.gov/iad/humanid/feret/`

2. **ORL Database of Faces**

   The ORL Database of faces contains ten different images of 40 distinct subjects. This images consist of various lighting, facial expressions and facial details. The size of the images are 92x112 pixels with 256 grey levels per pixel.

   Link: `http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html`

3. **BioID**

   BioID has 1521 gray level images of 23 different test persons. The standard image size is 384x286 pixel. The faces, recorded under natural conditions, i.e. varying illumination and complex background.

   Link: `https://www.bioid.com/About/BioID-Face-Database`

4. **Yale Face DB B**

   The Yale Face DB B concentrates on illumination. It contains 5760 single light source images of 10 subjects. These subjects are seen in 576 views (9 poses and 64 illumination conditions). Also the ambient (background) illumination for every subject in a particular pose was also for every subject in a particular pose was also captured. Bringing the total number of images available in the database to 5760 + 90 = 5850.

Link: http://vision.ucsd.edu/~leekc/ExtYaleDatabase/Yale%20Face%20Database.htm

5. **XM2VTS**

   This is a multi-modal face database. It has four recordings of 295 subjects taken over a four-month period. Each recording contains a speaking head shot and a rotating head shot. The database reflects the project which was to find a means of successful face recognition through audio-visual processes.

   Link:http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/

6. **PIE Database**

   Researchers at the Carnegie Mellon University created the PIE database. It contains 41,368 images of 68 people. Each person was captured under 13 different poses, 43 different illumination conditions, and with 4 different expressions.

   Link: http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261

2.7 Commercial Software

Face recognition is an area of immense potential both security wise, as well as financial rewards. In the last ten years, there have been companies which have released facial recognition software. Below is a list of some:

1. **Cognitec:** The primary product offered by Cognitec is FaceVACS. This product is split into FaceVACS-DBScan, FaceVACS-Entry, FaceVACS-PortraitAcquisition and FaceVACS-VideoScan.

   Link: http://www.cognitec.com/

2. **Sensible Vision:** SV offers a product called FastAccess. They are primarily focused on secured access to hardware devices.

Link: `http://www.sensiblevision.com/en-us/technology/overview.aspx`

3. **Iridian Technologies:** This Company offers 3 solutions for Iris recognition. The Knowho Authentication Server, PrivateID and Proof Positive.

   Link: `http://www.iridiantech.com/`

4. **NextgenID:** The primary product offered is the BioAxs9800-FR. It combines the power of multiple biometric credentials such as face and fingerprint recognition to create a secure access control platform.

   Link: `http://www.nextgenid.com/`

5. **Technest Inc:** A leader in 3D Imaging, 3D facial recognition and intelligence surveillance. Their products include 3D FaceCam, 3D FaceMap and 3D FaceMatch.

   Link: `http://www.genextech.com/`

2.8 Hardware and Software Requirements

In this section, the various hardware and software requirements for the project are reviewed. Matlab R2013a will be used as the development platform for the system. The following requirements are the requirements listed on the official Matlab website [2] in order to run the Matlab R2013a software successfully.

2.8.1 Hardware

The face recognition system will be built on a Windows 7 computer with 4GB of RAM.

Minimum Hardware Requirements for Windows

1. 1GB RAM minimum (2GB Recommended)

2. Processor with at least 1GHz of processing speed.

3. Disk space of at least 1GB.

These minimum hardware requirements are for Windows computers running R2013a version of Matlab. For minimum hardware requirements of other operating systems visit the Matlab website [18].

2.8.2 Software

As stated already, the Matlab R2013a will be used for developement as it contains all the neccessary image processing libraries.

Minimum Requirements for Windows

1. Windows 10
2. Windows 8.1
3. Windows 8
4. Windows 7 Service Pack 1
5. Windows Vista Service Pack 2
6. Windows XP Service Pack 3
7. Windows XP x64 Edition Service Pack 2
8. Windows Server 2012
9. Windows Server 2008 R2 Service Pack 1
10. Windows Server 2008 Service Pack 2
11. Windows Server 2003 R2 Service Pack 2

Due to licensing cost and restrictions, only MATLAB version R2013a was tested for full compliancy. However it is believed that the system should also run on higher versions of the MATLAB IDE. As the MATLAB libraries used are still the same.

2.9 Project Tools

Tools used in this project will be described in this section

2.9.1 MATLAB

MATLAB is a programming language for technical computing. It combines computation, visualization, and programming in a user friendly environment. Some of it's uses include Mathematics and computation, algorithm development, modelling, simulation and prototyping.

MATLAB stands for Matrix laboratory. Over the years, MATLAB has evolved to become the tool of choice for high-productivity research, development, and analysis. MATLAB features application specific solutions called toolboxes which allow users to learn and apply specialized technologies.

The MATLAB working environment, handle graphics, mathematical function library and application programming interface make it one of the best tools for sophisticated research, and the best tool for implementing the face recognition system discussed in this thesis.

2.10 Examples of Face Recognition

SkyBiometry is a company which offers commercial face detection and recognition solutions. Their website offers a demo of face recognition which allows you to upload two images of your choice and see if their algorithm is able to detect and recognize the faces. The screenshots below show the result of such tests.

Figure 2.3 shows a completed face recognition test. The image on the left is the test image while the images on the right are the database images from which to search for a matching image to the test image. As can be seen there is a hundred percent match for the third image on the database. The notes underneath the image on the left simply provide detected/predicted information about the image such as gender, smiling face, glasses etc.

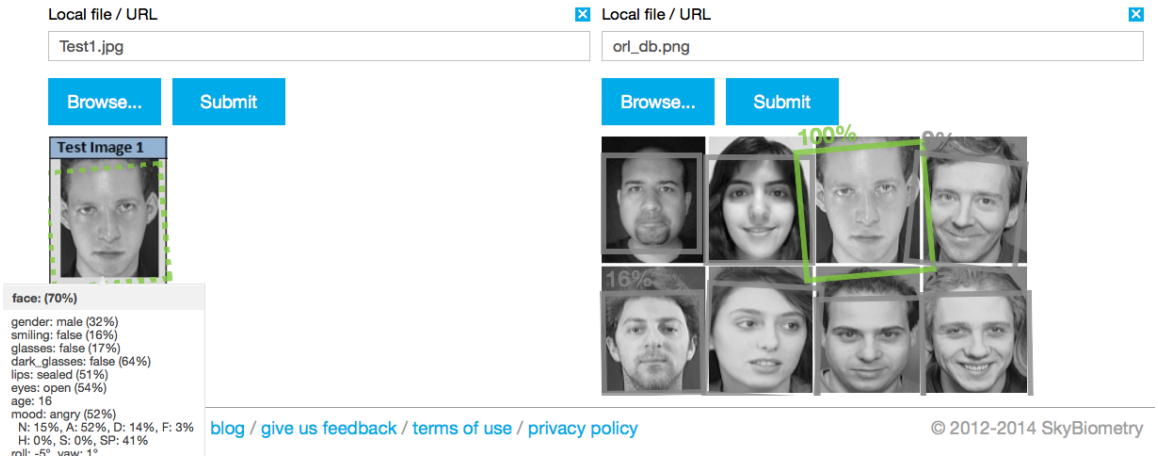


Figure 2.3: Face Recognition Demo

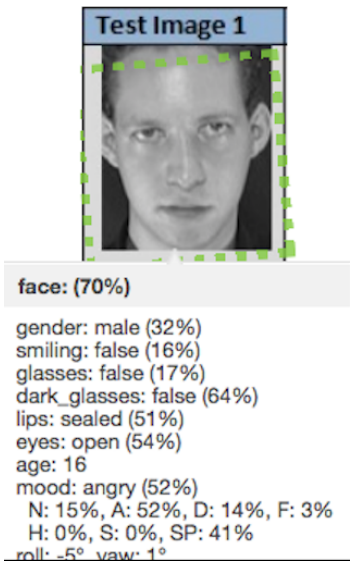


Figure 2.4: Enlarged test image

Figure 2.4 shows an enlarged version of the left side of Figure 2.3.

2.11 Proposed Solution Specifics

In the previous sections, two approaches to face recognition were given. Namely the Geometric local approach and the Appearance-based Global approach. We also looked at the

various methods which fall under each of the approaches.

For implementing our project, we will focus on the following methods.

1. Principal Componenents Analysis (PCA)
2. Linear Discriminant Analysis (LDA)
3. Gabor Principal Components Analysis (GaborPCA)
4. Gabor Linear Discriminant Analysis (GaborLDA)

2.11.1 Proposed Solution

There is a popular expression which says "*Two heads are better than one*". In this project we say "*Four methods are better than one*". Four was the number of methods chosen as it was thought of as a suitable number of methods to use in order to achieve the objective. Each of the methods can be used individually for face recognition. Therefore, it is believed that if a method is found which can be used to combine the outputs of all four methods, the level of accuracy will be improved drastically as each method will complement the other in the cases where one or more of the selected four methods perform badly.

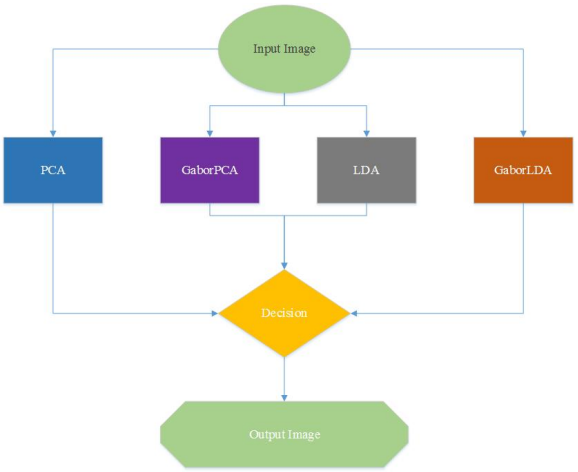


Figure 2.5: Flowchart: Proposed Face Recogntion System

2.11.2 Proposed Face Database

The ORL database mentioned in the literature review will be used for implementing and testing each of the methods. The reason for its selection is that it offers images with different poses, illuminations and multiple expressions. Each image is also small in dimension (92x112)so the proposed system would require less time in computing the input images.

2.11.3 Proposed Classifier

The nearest neighbor classifier is the proposed classifier to be used in this project. It will perform classification by finding the Euclidean distances between the test and database eigenvalues projected to the image subspace. The database image with the smallest distance to the test image will be deemed as its match. In Figure (2.5) above, the nearest neighbor classification occurs at the decision stage.

2.12 Chapter Summary

In this chapter, a literature review was carried out. It looked at the history of face recognition and how it ties in with image processing and computer vision as a whole. As part of the review, the two major approaches to face recognition (geometric and appearance-based) were discussed. The most common methods of each approach were also discussed along with the major challenges associated with each approach. Also classifiers were discussed. Following the section on classifiers, the hardware and software requirements were investigated, image databases which have been standardized for use with face recognition research were discussed and examples of commercial face recognition systems were given. Finally, the proposed solution for the face recognition system was given. It was clarified that four appearance-based methods will be used in conjunction with the Euclidean Nearest Neighbor classifier.

Chapter 3

Methodology

In this chapter, the methods which have been chosen for the implementation of the face recognition system will be looked at in more details. Section 3.1 discusses the pre-processing technique used before implementing the face recognition algorithm. Section 3.2 will provide an in-depth mathematical overview of the PCA, LDA and Gabor methods. Sections 3.3 and 3.4 focus on the combination of the four methods. They look at how the methods can be combined in order to create a face recognition system that is designed to have higher accuracy than each of the methods implemented separately.

3.1 Pre-processing

In order to improve the result of face recognition processes, it is sometimes necessary to pre-process the data before feeding it to the methods. In essence, this processing will help enhance the data in preparation for the processing phase.

3.1.1 Histogram Equalization

Histogram equalization is a technique used to globally increase the contrast of images. This technique will be used to resolve illumination issues prior to the recognition stage.

Figure 3.1 shows the equalization of an image. The graphs show the intensity distributions before and after the distribution. As seen, the illumination of the image changes profoundly. Histogram Equalization calculations are relatively easy to achieve [19].

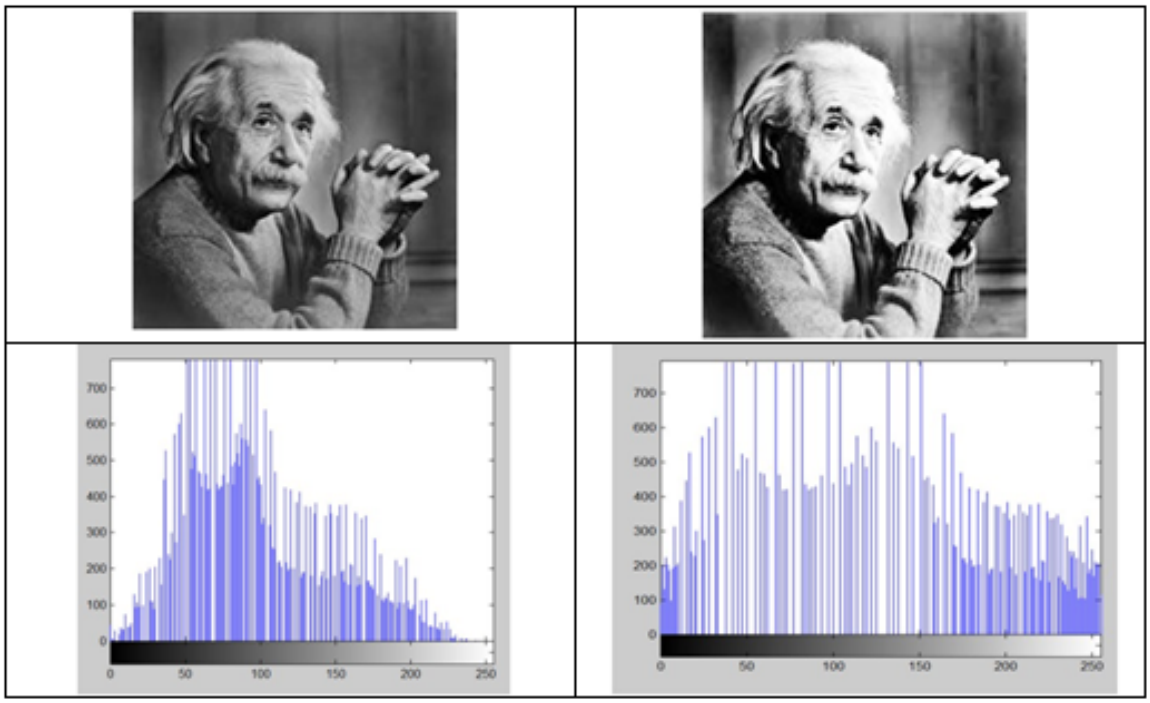


Figure 3.1: Histogram Equalization

3.2 In-depth Methods Analysis

3.2.1 Principal Component Analysis

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of possibly correlated variables into a set of uncorrelated variables called principal components. PCA is defined in such a way that the first principal component has the largest possible variance. Each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components, hence it is uncorrelated with them. The ranking of the components allows to achieve an efficient dimension reduction of the data in many applications such as compression and pattern recognition. PCA was invented in 1901 by Karl Pearson and was later independently developed (and named) by Harold Hotelling in the 1930s.

Kirby and Sirovich were among the first to use PCA to efficiently represent images of

human faces. They showed that face images can be optimally compressed (minimizing the mean squared error between the original images and their reconstructions) by representing them along the eigenvectors coordinate space given by PCA and approximately reconstructing them using a small collection of projections onto the PCA eigen-space [2, 8]. Turk and Pentland popularized the use of PCA for face recognition by introducing the well-known eigenfaces method for face recognition [10]. Since then, PCA has become the most widely used subspace projection technique for face recognition.

Derivation of Principal Components

Let X be a random vector in $\mathbb{R}^n$ consisting of the components $x_1, \ldots, x_n$. We want to find a new random vector Y consisting of the new random variables $y_1, \ldots, y_n$ that are linear combinations of the original variables $x_1, \ldots, x_n$, that is for $i = 1, \ldots, n$

$$y_i = \sum_{j=1}^{n} a_{ij} x_j = \mathbf{a}_i^T X, \quad i = 1 \ldots n. \tag{3.1}$$

Let's denote A the matrix of coefficients a_{ij}. Equation 3.1 can be rewritten as

$$Y = A^T X. \tag{3.2}$$

We seek the orthogonal transformation (matrix) A such that the new variables $y_1, \ldots, y_n$ have maximal variance and they are uncorrelated. The variance of each component y_i is given by

$$\begin{aligned}
\mathrm{Var}\,(y_i) &= E(y_i^2) - E(y_i)^2 \\
&= E(\mathbf{a}_i^T X X^T \mathbf{a}_i) - E(\mathbf{a}_i^T X) E(X^T \mathbf{a}_i) \\
&= \mathbf{a}_i^T \left(E(XX^T) - E(X)E(X^T) \right) \mathbf{a}_i \\
&= \mathbf{a}_i^T \Sigma \mathbf{a}_i,
\end{aligned}$$

where Σ is the covariance matrix of X, that is

$$\Sigma = E(XX^T) - E(X)E(X^T).$$

Maximizing $\mathrm{Var}\,(y_1)$ under the constraint $\mathbf{a}_1^T\mathbf{a}_1 = 1$ is equivalent to finding the stationary value of the unconditional form

$$F(\mathbf{a}_1) = \mathbf{a}_1^T\Sigma\mathbf{a}_1 - \lambda(\mathbf{a}_1^T\mathbf{a}_1 - 1)$$

where λ is a Lagrange multiplier. By differentiating with respect to each component of $\mathbf{a}_1$ and equating to zero, we get the equation

$$\Sigma\mathbf{a}_1 - \lambda\mathbf{a}_1 = 0.$$

It follows that $\mathbf{a}_1$ must be an eigenvector of Σ associated with the eigenvalue λ and

$$\mathrm{Var}\,(y_1) = \mathbf{a}_1^T\Sigma\mathbf{a}_1 = \lambda.$$

Let's denote by $\lambda_1, \lambda_2, \ldots, \lambda_n$ the eigenvalues of Σ (not all necessarily distinct and not all zero), and let's assume that without loss of generality, they are ordered so that $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$. It is clear that in order to maximize the variance of y_1, it is enough to choose $\lambda = \lambda_1$. The variable y_1 is called the first principal component of X and has the largest variance of any linear combination of the original variables $x_1, \ldots, x_n$. Following the same reasoning, we can show that each subsequent principal component $y_i = \mathbf{a}_i^T X$ is obtained by choosing $\mathbf{a}_i$ as the eigenvector of the i-th largest eigenvalue of Σ. Note that imposing that the new variables y_i are uncorrelated is equivalent to having the eigenvectors $\mathbf{a}_i$ orthogonal to each other. Indeed, y_i and y_j are uncorrelated if and only if

$$E(y_i y_j) - E(y_i)E(y_j) = 0. \tag{3.3}$$

We have

$$\begin{aligned}
E(y_i y_j) - E(y_i)E(y_j) &= E(\mathbf{a}_i^T X \mathbf{a}_j^T X) - E(\mathbf{a}_i^T X)E(\mathbf{a}_j^T X) \\
&= E(\mathbf{a}_i^T X X^T \mathbf{a}_j) - E(\mathbf{a}_i^T X)E(X^T \mathbf{a}_j) \\
&= \mathbf{a}_i^T \left(E(XX^T) - E(X)E(X^T)\right)\mathbf{a}_j \\
&= \mathbf{a}_i^T \Sigma \mathbf{a}_j,
\end{aligned}$$

and since $\mathbf{a}_j$ is an eigenvector of Σ, it follows that equation 3.3 is equivalent to

$$\mathbf{a}_i^T \mathbf{a}_j = 0,$$

that is, $\mathbf{a}_i$ and $\mathbf{a}_j$ are orthogonal. We now have a method to determine the principal components of the random vector X by performing an eigenvalue decomposition of the covariance matrix Σ of X, which is symmetric and positive definite and use the corresponding eigenvectors as the columns of the transformation matrix A that will yield the principal components.

First, equation (3.2) is updated as follows

$$Y = A^T(X - E(X)). \tag{3.4}$$

Subtracting the mean vector from X allows to center the new vector Y (i.e., $E(Y) = 0$). The eigenvectors $\mathbf{a}_i$ indicate the main directions containing the maximal variability in the original variables. Each new coordinate y_i represents the projection

$$y_i = a_i^T(X - E(X))$$

of the vector $X - E(X)$ onto the direction of the eigenvector $\mathbf{a}_i$. Each eigenvalue λ_i represents the variance along the direction of its corresponding eigenvector, that is

$$\mathrm{Var}\,(y_i) = \lambda_i.$$

The sum of the variances of the principal components is given by

$$\sum_{i=1}^{n} \mathrm{Var}\,(y_i) = \sum_{i=1}^{n} \lambda_i$$

which is the sum of the eigenvalues of the covariance matrix of X. The result of the PCA is a coordinate transform to best represent the variance of the original variables. The vector X can be reconstructed from the projections of $X - E(X)$ as follows

$$X = E(X) + \sum_{i=1}^{n} y_i \mathbf{a}_i.$$

Indeed, since A is orthogonal, we have $A^T = A^{-1}$. Moreover, we have $Y = A^T(X - E(X))$. It follows immediately that

$$X = E(X) + AY = E(X) + \sum_{i=1}^{n} y_i \mathbf{a}_i. \tag{3.5}$$

Important Remarks

1. **Sampling**

   We have assumed so far that the covariance matrix Σ of the random vector X is given. However, in most practical problems, the vector X and its statistical properties can only be estimated from a set of sample vectors. The expected value of X and its covariant matrix will be replaced by the sample mean and the sample covariance matrix of the N data vectors, i.e. $E(X)$ is replaced by the sample mean

   $$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

   and the covariance matrix

   $$\Sigma = E(XX^T) - E(X)E(X^T)$$

   is replaced by the sample covariance matrix given by

   $$S = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)(X_i - \mu)^T = \frac{1}{N}\sum_{i=1}^{N} X_i {X_i}^T - \mu\mu^T.$$

2. **Standardization**

   The method to derive the principal components dependends on the scales used to measure the original variables. In case one of the variables has a range of values that greatly exceeds the others, then we expect the first principal component to lie in the direction of that variable. If the units of each variable differ, then the principal components will most likely depend on which units are used. The practical solution to this problem is to standardize the data so that the variables have equal range.

One way to achieve the standardization is by transforming the data to have zero mean and unit variance, so that we find the principal components from the correlation matrix. This gives equal importance to all the original variables. Other forms of standardization are possible (see [11]).

Dimension Reduction and Approximation

In order to represent the given data in a reduced dimension while keeping most of its variance, we may reduce the number of principal components used in the processing of the data by retaining only the first few principal components with maximal variance. The exact number of components to use is determined by the data and the application. If the $k < n$ first components are retained, then an n-dimensional vector $X \in \mathbb{R}^n$ is projected to a low-dimensional feature space $Y \in \mathbb{R}^k$ using an $n \times k$ matrix W whose columns are the first k eigenvectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$ of the sample covariance matrix S:

$$Y = W^T X.$$

From $Y \in \mathbb{R}^k$, the reconstruction leads to an approximation of X given by the value $\tilde{X} = WY$. The error of the reconstruction is given by the quantity

$$||X - WW^T X||$$

The transformation WW^T is of rank k. It can be shown that the mean square error (MSE) between any vector X and its reconstruction using only k principle eigenvectors is given by the expression

$$MSE = \sum_{i=k+1}^{n} \lambda_i.$$

See Chapter 4.1.1 for the individual implementation results for PCA

3.2.2 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a statistical method that seeks to reduce dimensionality of data by finding a suitable linear combination of variables (features) that best

discriminates among the classes in the data. The main idea consists of projecting the data onto a lower-dimensional space that provides a good class-separability and also reduce computational costs in the classification process.

LDA is a generalization of Fisher's linear discriminant (FLD) although the two terms are often used interchangeably in statistics and pattern recognition. The FLD is formulated by Ronald A. Fisher in 1936 [13]. In the original form, the linear discriminant was described for a 2-class problem. It was later generalized as *multi-class Linear Discriminant Analysis* by C. R. Rao in 1948 [14]. The goal remains to find the projection of the feature space such that the ratio of the **between-class scatter** and the **within-class scatter** is maximized.

Let us consider a set of N samples $x_1, \ldots, x_N$ in $\mathbb{R}^d$. Let us also consider a linear transformation that projects the original d-dimensional space onto an m-dimensional space, where $m < n$. The new feature vectors y_k are defined by the following linear transformation

$$y_k = W^T x_k, \quad k = 1, \ldots, N.$$

For the sake of simplicity, we illustrate the theory in the case $m = 1$, when y is a scalar and the projection space is a line. We also assume that the sample data consists of two classes ω_1, and ω_2. We would like to select the line that maximizes the separability of the scalars. We first attempt to define a measure of separation as the distance between the projected means of the two classes, i.e.

$$J(W) = |\tilde{\mu}_1 - \tilde{\mu}_2| = |W^T(\mu_1 - \mu_2)|,$$

where

$$\mu_i = \frac{1}{N_i} \sum_{x \in \omega_i} x$$

and

$$\tilde{\mu}_i = \frac{1}{N_i} \sum_{x \in \omega_i} y = \frac{1}{N_i} \sum_{x \in \omega_i} W^T x = W^T \mu_i.$$

It turns out that the distance between projected means is not necessarily a good measure,

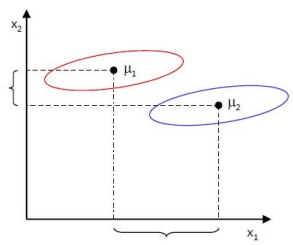


Figure 3.2: Projected means.

since it does not account for the standard deviation within classes. As illustrated in Figure 3.3, sometimes the axis that yields the maximal distance between the means is not the one that provides the maximal separation between the classes.

In order to overcome this problem, Fisher suggested to maximize the difference between the means, normalized by the **within-class scatter**. For each class we define the scatter, an equivalent of the variance, as

$$\tilde{S}_i^2 = \sum_{x \in \omega_i} (y - \tilde{\mu}_i)^2 = \sum_{x \in \omega_i} \left(W^T x - \tilde{\mu}_i\right)^2.$$

The quantity $\tilde{S}_1^2 + \tilde{S}_2^2$ is called the **within-class scatter**. The **Fisher linear discriminant** (FLD) is defined as the linear function $y = W^T x$ that maximizes the ratio

$$J(W) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{S}_1^2 + \tilde{S}_2^2}.$$

Therefore, we would like to find the projection that on one hand maximizes the distance between different classes and on the other hand minimizes the distance between the samples of the same class. However, we need first to express $J(W)$ as a function of W. We define the scatter matrix for each class in the feature space as

$$S_i = \sum_{x \in \omega_i} (x - \mu_i)(x - \mu_i)^T.$$

The **within-class scatter** is defined as $S_w = S_1 + S_2$. The scatter of the projection of the

classes can be expressed as a function of the scatter matrix in the feature space as follows

$$\begin{aligned}
\tilde{S}_i^2 &= \sum_{y \in \tilde{\omega}_i} (y - \tilde{\mu}_i)^2 \\
&= \sum_{x \in \omega_i} (W^T x - W^T \mu_i)^2 \\
&= \sum_{x \in \omega_i} W^T (x - \mu_i)(x - \mu_i)^T W \\
&= W^T S_i W.
\end{aligned}$$

It follows that $\tilde{S}_1^2 + \tilde{S}_2^2 = W^T S_w W$. In a similar way, we can see that

$$(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (W^T \mu_1 - W^T \mu_2)^2 = W^T (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T W = W^T S_b W,$$

where $S_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$ is defined as the **between-class scatter** of the two classes ω_1 and ω_2. Consequently, the Fisher criterion can be expressed as

$$J(W) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{S}_1^2 + \tilde{S}_2^2} = \frac{W^T S_b W}{W^T S_w W}.$$

We maximize $J(W)$ by differentiating it with respect to W and setting the result to zero:

$$\begin{aligned}
\frac{d}{dW} J(W) &= \frac{\frac{d}{dW}(W^T S_b W) W^T S_w W - \frac{d}{dW}(W^T S_w W) W^T S_b W}{(W^T S_w W)^2} \\
&= \frac{(2 S_b W) W^T S_w W - (2 S_w W) W^T S_b W}{(W^T S_w W)^2} \\
&= 0.
\end{aligned}$$

It follows that

$$\begin{aligned}
(S_b W) W^T S_w W - (S_w W) W^T S_b W &= 0 \\
\Rightarrow \frac{W^T S_w W (S_b W)}{W^T S_w W} - \frac{W^T S_b W (S_w W)}{W^T S_w W} &= 0 \\
\Rightarrow S_b W - \frac{W^T S_b W (S_w W)}{W^T S_w W} &= 0 \\
\Rightarrow S_b W - \lambda S_w W &= 0
\end{aligned}$$

which is a generalized eigenvalue problem. If S_w has full rank, the problem can be converted to an eigenvalue problem

$$S_w^{-1} S_b W = \lambda W. \tag{3.6}$$

The vector S_bW points in the same direction as $\mu_1 - \mu_2$ for any vector W. Indeed, we have

$$S_bW = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T W = (\mu_1 - \mu_2)\underbrace{(\mu_1 - \mu_2)^T W}_{\alpha} = \alpha(\mu_1 - \mu_2).$$

By setting $W = S_w^{-1}(\mu_1 - \mu_2)$, we have

$$S_w^{-1} S_b \left[S_w^{-1}(\mu_1 - \mu_2)\right] = S_w^{-1}\left[\alpha(\mu_1 - \mu_2)\right] = \alpha S_W^{-1}(\mu_1 - \mu_2),$$

i.e.

$$S_w^{-1} S_b W = \alpha W.$$

We have therefore solved the eigenvalue problem in equation 3.6. It follows that

$$W^* = \arg\max \frac{W^T S_B W}{W^T S_w W} = S_w^{-1}(\mu_1 - \mu_2).$$

The FLD and LDA can be generalized to a multi-class problem with C classes ω_1, ω_2, $\ldots$, ω_C, each containing N_i samples. We denote by $N = N_1 + \ldots + N_C$ the total number of samples. In this case, we can project on a space of dimension $m = C - 1$. We seek m projection vectors W_1, $\ldots$, W_m which can be arranged in columns to form the projection matrix W which defines the linear transformation

$$y = W^T x$$

that maximizes the ratio of the **between-class scatter** and the **within-class scatter** which generalize to the multi-class case as follows. The **within-class scatter** matrix is defined as

$$S_w = \sum_{i=1}^{C} S_i,$$

where

$$S_i = \sum_{x \in \omega_i} (x - \mu_i)(x - \mu_i)^T \quad \text{and} \quad \mu_i = \frac{1}{N_i} \sum_{x \in \omega_i} x.$$

The **between-class scatter** matrix is defined as

$$S_b = \sum_{i=1}^{C} (\mu_i - \mu)(\mu_i - \mu)^T$$

where

$$\mu = \frac{1}{N}\sum_{x} x = \frac{1}{N}\sum_{i=1}^{C} N_i\mu_i. \tag{3.7}$$

In a similar way as in the two class case, we can derive the scatter matrices for the projected samples as

$$\tilde{S}_w = W^T S_w W \quad \text{and} \quad \tilde{S}_b = W^T S_b W.$$

The ratio of between-class to within-class scatters cannot be defined directly since the projection is no longer a scalar. This is replaced by the determinant of the scatter matrices to obtain a scalar objective function

$$J(W) = \frac{\tilde{S}_b}{\tilde{S}_w} = \frac{|W^T S_b W|}{|W^T S_w W|}. \tag{3.8}$$

It can be shown (see [11] and the references therein) that the projection matrix W^* that maximizes the ratio $J(W)$, i.e.

$$W^* = \arg\max \frac{|W^T S_b W|}{|W^T S_w W|}$$

is the one whose columns W_i^* are the eigenvectors corresponding to the largest eigenvalues of the following generalized eigenvalue problem

$$(S_b - \lambda S_w)W_i^* = 0. \tag{3.9}$$

Note that $S_b = \sum_{i=1}^{C}(\mu_i - \mu)(\mu_i - \mu)^T$ is the sum of C matrices each of rank at most 1 which is constrained by the relation 3.7. It follows that S_b is of rank at most $C - 1$. This means that only $C - 1$ of the eigenvalues will eventually be non-zero. If S_w has full rank, then the projections with maximum class separability information are the eigenvectors corresponding to the largest eigenvalues of $S_w^{-1}S_b$.

See Chapter 4.1.2 for the individual implementation results for LDA

3.2.3 Review of Gabor Filters

Gabor filters, named after Dennis Gabor [5], are bandpass filters which are used in image processing for feature extraction and texture analysis [12, 21]. Mathematically speaking, each Gabor filter is essentially a Gaussian kernel function modulated by a sinusoidal plane wave and can be defined in the spatial domain as follows

$$\psi_{f,\theta}(x,y) = \frac{f^2}{\pi\gamma\eta} \exp\left(-\left(\frac{f^2}{\gamma^2}x_r^2 + \frac{f^2}{\eta^2}y_r^2\right)\right) \exp\left(j2\pi f x_r\right), \tag{3.10}$$

where

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

f is the central frequency of the modulating sinusoidal plane wave, and θ is the orientation of the major axis of the elliptical Gaussian. Gabor filters exhibit optimal localization in both the spatial and frequency domain. Hence as feature extraction tools, they provide multi-resolution local features in space corresponding to confined frequency bands. Frequency and orientation representations of Gabor filters are similar to those of the mammalian visual cortex, which makes them relevant from the biological point of view for feature representation and discrimination. The 2D Fourier transform of the Gabor filters defined in equation 3.10 is given by

$$\Psi_{f,\theta}(u,v) = \exp\left(-\pi^2\left(\frac{\gamma^2}{f^2}(u_r - f)^2 + \frac{\eta^2}{f^2}v_r^2\right)\right), \tag{3.11}$$

where

$$\begin{bmatrix} u_r \\ v_r \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

The Gabor filters with different values of the parameters f, γ, η and θ form the Gabor filter bank. The parameters determine the shape and characteristics of the filters and can be used to define different families of Gabor filters. The parameters values that are frequently used in face recognition are $\gamma = \eta = \sqrt{2}$,

$$\theta = \frac{n\pi}{8}, \quad n = 0, \ldots, 7 \qquad \text{and} \qquad f = \frac{0.25}{2^{\frac{m}{2}}}, \quad m = 0, \ldots, 4.$$

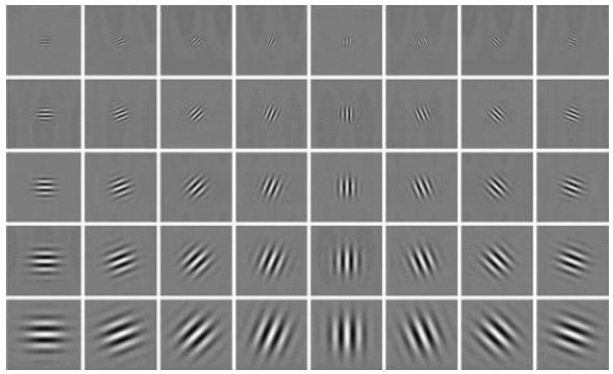


Figure 3.3: The real parts of the Gabor filter bank at 5 scales and 8 orientations.

These values form a filter bank comprising 40 Gabor filters of five scales (frequencies) and eight orientations (see [22, 23, 24]).

In the context of face recognition, each input image $I(x, y)$ is convolved (filtered) with each of the 40 filters from the filter bank (see Figure 3.3) which increases the dimensionality of the original data to 40 times its initial size. The convolution with each Gabor kernel can be carried out directly using the formula

$$C_{f,\theta}(x, y) = I(x, y) * \psi_{f,\theta}(x, y), \tag{3.12}$$

or by using the forward and backward Fast Fourier Transforms (FFT) and the convolution theorem

$$C_{f,\theta}(x, y) = \mathcal{F}^{-1}\{\mathcal{F}\{I(x, y)\}(u, v)\Psi_{f,\theta}(u, v)\}(x, y). \tag{3.13}$$

We obtain the Gabor feature vector representation of the image $I(x, y)$ denoted

$$\mathcal{R} = \{C_{f_m,\theta_n}(x, y) : m = 0, \ldots, 4 \quad \text{and} \quad n = 0, \ldots, 7\}.$$

The augmented feature vectors obtained using the Gabor filters exhibit the same scale, locality and orientation properties as those of the Gabor filters. This allows to build a face representation that is robust to changing facial expressions as well as to illumination

variations (see [24]). In this work, we apply the Gabor filters in conjunction with PCA and LDA in order to reduce the dimensionality of the feature vectors.

See Chapter 4.1.3 and Chapter 4.1.4 respectively for the implementation results for GaborPCA and GaborLDA

3.3 Combined Methods Investigation

In this section, the emphasis will be on describing a method which can be used to collect the outputs of the individual results and through statistical manipulation, be able to decide what the final output image is. In other words, a method will be described which will be used in deciding the resulting output image from the combination of proposed methods. The diagram below gives an overview of the combination process:

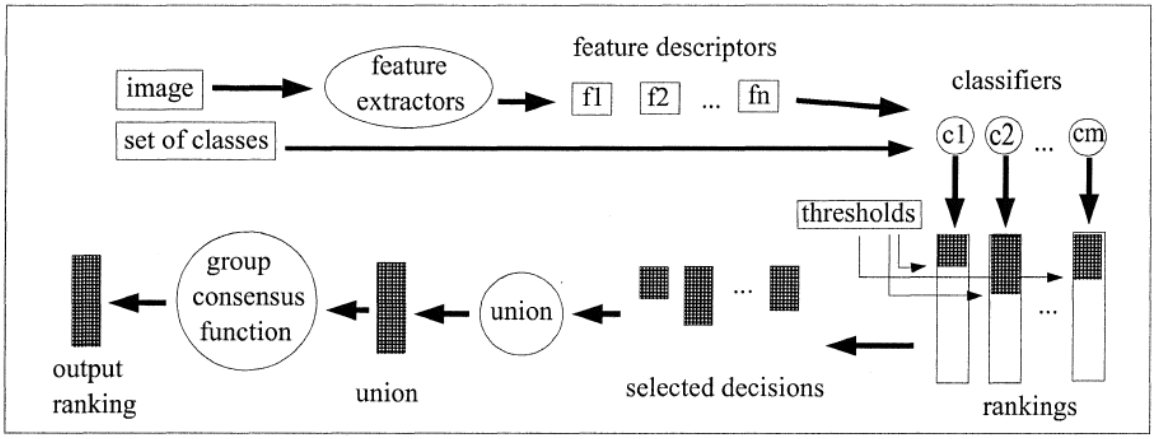


Figure 3.4: Proposed combination method

In the example seen in Figure 3.4, features from the given test images and database images are extracted. In our case the features are the eigenfaces and fisherfaces. For each of the classifiers, (PCA, LDA, GaborPCA and GaborLDA), the outputs are generated. These outputs can then be processed by a statistical method (in this case union). Following that, the final output ranking is gotten from which the best matching image can be derived.

3.4 Ranking Outputs

The four methods which will be used in this project have already been discussed. This section will now look at a way in which these four methods can be combined and through a ranking system, be able to determine an accurate output image.

3.4.1 Understanding Ranking

For each of the approaches in the previous section, subspace analysis is done using the Euclidean distance. As already discussed, the Euclidean distance operates by comparing the features of the test image against the features of the database images in the image subspace. It then outputs the database image with the least difference from the test image as the most likely or equivalent match for the test image. As such, the Euclidean distance can be used to perform a ranking of similarity for all the training images against the test image.

Now if each of the four methods generate Euclidean based rankings, we can then come up with methods on how to compare all the rankings in order to make decision as to what the output image should be. The table below gives sample Euclidean distances.

Sample Euclidean Distances										
	Img1	**Img2**	**Img3**	**Img4**	**Img5**	**Img6**	**Img7**	**Img8**	**Img9**	**Img10**
PCA	1.44	5.71	0.43	0.04	2.34	0.93	1.33	9.32	5.55	3.45
LDA	2.12	6.33	0.49	0.01	2.98	1.12	1.39	8.02	4.58	3.22
GaborPCA	2.11	7.09	0.77	0.06	2.88	0.87	1.52	7.98	5.09	2.55
GaborLDA	1.99	6.32	0.58	0.09	2.79	1.29	1.72	6.98	4.87	2.98

Table 3.1: Sample Distances

As can be seen in Table 3.1, sample distances of a test image against ten database images are given. Each distance was generated by subtracting the test and database images in the image subspace. For each method, the database image with the smallest distance is considered the matching image for the test image. The ten distances for each method can

be sorted from smallest to largest to create rankings. The next section discusses methods which can be used for combination, and they are implemented on these sample rankings for further illustration.

3.4.2 Methods of Combination

In order to get the final output image that is a match for the test image, the Euclidean distances that lead up to the rankings must be combined in ways that ensure that the best match is what is chosen in the end. This section will discuss two methods which will be used for the combination. These are the mean of the distances (averages), and the product of the distances. Finally, these methods will be applied to the sample rankings in Table 3.1 for further illustration.

Mean of Euclidean Distances

The mean of distances is achieved by adding the corresponding distances for each approach and then dividing the total by the number of approaches used. Below is an example of this.

$$d = \left\{ \frac{d_1^{PCA} + d_1^{LDA} + d_1^{GPCA} + d_1^{GLDA}}{4}, ..., \frac{d_n^{PCA} + d_n^{LDA} + d_n^{GPCA} + d_n^{GLDA}}{4} \right\} \quad (3.14)$$

Mean of Distances										
	Img1	**Img2**	**Img3**	**Img4**	**Img5**	**Img6**	**Img7**	**Img8**	**Img9**	**Img10**
PCA	1.44	5.71	0.43	0.04	2.34	0.93	1.33	9.32	5.55	3.45
LDA	2.12	6.33	0.49	0.01	2.98	1.12	1.39	8.02	4.58	3.22
GaborPCA	2.11	7.09	0.77	0.06	2.88	0.87	1.52	7.98	5.09	2.55
GaborLDA	1.99	6.32	0.58	0.09	2.79	1.29	1.72	6.98	4.87	2.98
Mean	**1.92**	**6.36**	**0.57**	**0.05**	**2.75**	**1.05**	**1.49**	**8.08**	**5.02**	**3.05**
Mean Ranking	**5**	**9**	**2**	**1**	**6**	**3**	**4**	**10**	**8**	**7**

Table 3.2: Mean of Distances

Product of Euclidean Distances

The product of the euclidean distances as it suggests combines the distances through multiplication. For each approach, the distances are multiplied together to get a new value which is a combination of all four methods. Below is an example of such.

$$d = \{d_1^{PCA}.d_1^{LDA}.d_1^{GPCA}.d_1^{GLDA}, ..., d_n^{PCA}.d_n^{LDA}.d_n^{GPCA}.d_n^{GLDA}\} \tag{3.15}$$

Product of Distances										
	Img1	**Img2**	**Img3**	**Img4**	**Img5**	**Img6**	**Img7**	**Img8**	**Img9**	**Img10**
PCA	1.44	5.71	0.43	0.04	2.34	0.93	1.33	9.32	5.55	3.45
LDA	2.12	6.33	0.49	0.01	2.98	1.12	1.39	8.02	4.58	3.22
GaborPCA	2.11	7.09	0.77	0.06	2.88	0.87	1.52	7.98	5.09	2.55
GaborLDA	1.99	6.32	0.58	0.09	2.79	1.29	1.72	6.98	4.87	2.98
Product	**12.8**	**1619.6**	**0.1**	**0.0**	**56.0**	**1.2**	**4.8**	**4163.4**	**630.1**	**84.4**
Ranking	**5**	**9**	**2**	**1**	**6**	**3**	**4**	**10**	**8**	**7**

Table 3.3: Product of Distances

Summary of Sample Rankings Implementation

In applying the two methods to the sample distances and then generating the rankings, it can be seen that the equivalent image for all is Image 4. Note that the sample rankings in Table 3.1 have been simplified so as to illustrate the point. Actual tables may not be so accommodating.

3.5 Evaluation Strategy

In order to analyse the results for the individual methods, the following tools will be used:

1. **ROC Curve**

   The receiver operating characteristics (ROC) curve is a graphical plot used for displaying the performance of a classifier system. It captures and displays the change as the discrimination threshold is varied.

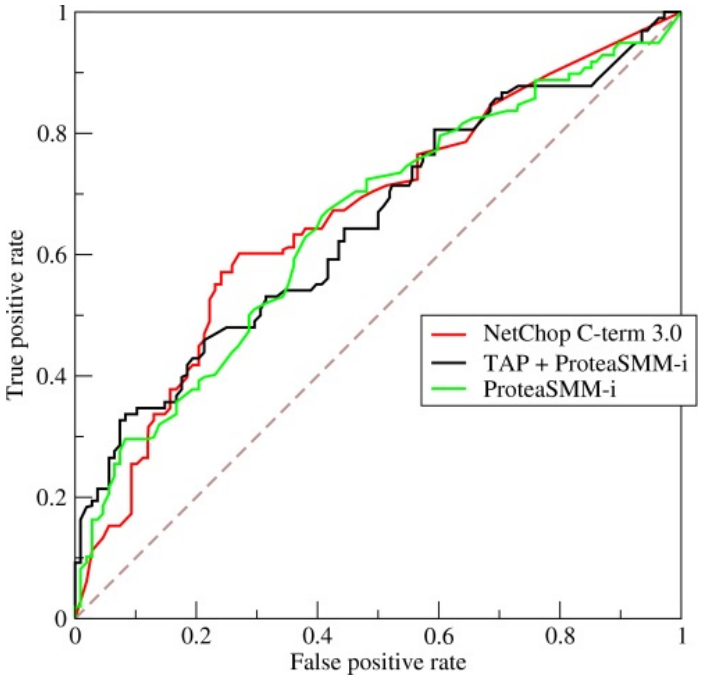


Figure 3.5: An example of a ROC graph

As seen in Figure 3.5, the ROC curve is generated by plotting true positive rates against false positive rates at various thresholds. The true positive rate calculates the proportion of positives that are identified correctly as such (e.g the percentage of database images which are correctly identified as an equivalent match to the test image), while the false positive rate calculates the proportion of negatives that are identified positively (e.g the percentage of database images which are not matches to the test image but are identified as equivalent matches to the test image). The perfect plot would be one in which a hundred percent of the true positives are measured (all database images that should be equivalent to the test image are identified) and a hundred percent of the false negatives are measured (all database images that are not equivalent to the test image are identified as not equivalent). However, this is hard to achieve as errors are prone to occur.

2. **CMC Curve**

   A Cumulative Match Characteristic curve measures how well the identities in a database within an identification system are ranked with respect to unknown test images.

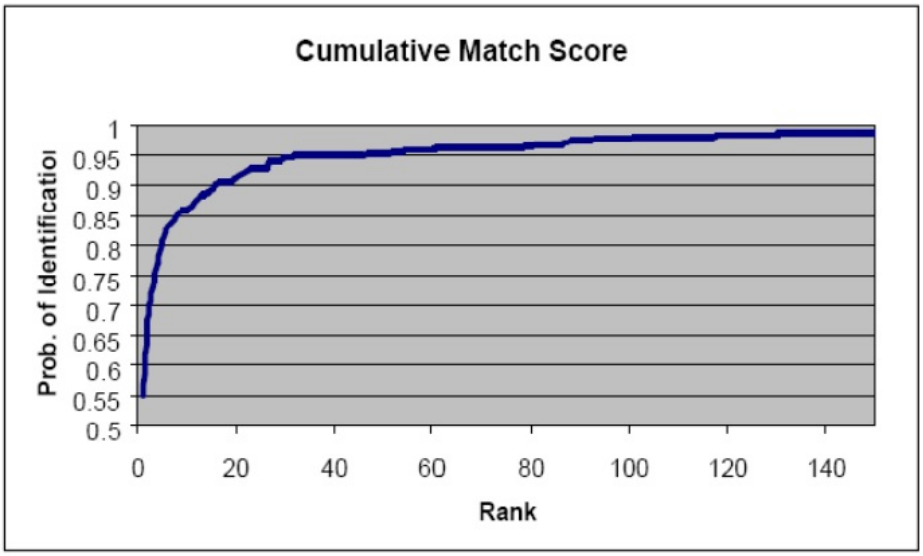


Figure 3.6: An example of a CMC graph

   In other words, the ranking capabilities of an identification system are judged by CMC measures. A key feature of CMC is that in a graph plot that has all the possible ranks, (if the database has 140 images and the CMC goes through rank 140), the probability of identification is 100% at the highest (140 as seen in the chart example) rank. This is because every image is in the image database and the plot is showing the identification rate for the entire image database.

3.6 Chapter Summary

This chapter analyzed in detail the mathematics behind the selected appearance based methods, it then discussed the histogram pre-processing method used in implementation. After that, it then went on to list the steps to implementation for both the individual and combined face recognition methods. Ranking methods were discussed as a way to combine

outputs and finally, the evaluation strategy for the results were also discussed.

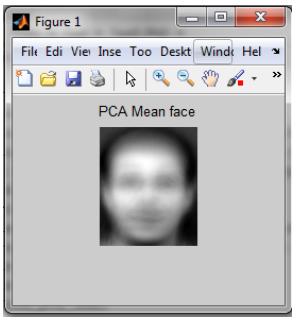
Chapter 4

Results and Analysis

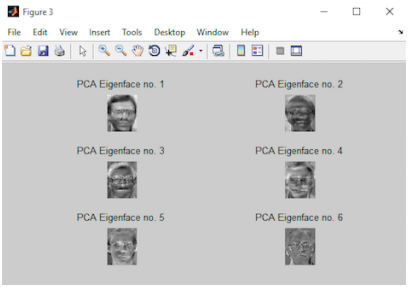
This chapter will analyze the results for first the individual methods, and then the combined methods algorithm using images selected from the ORL image database. The ORL database as already stated contains images of 40 individuals, each with 10 varying image poses; bringing the total of the images in the database to 400. For the purposes of this tests, 5 images of 5 individuals were randomly selected making a total of 25 images used.

4.1 Individual Methods Results

4.1.1 Principal Components Analysis



(a) PCA Mean Face

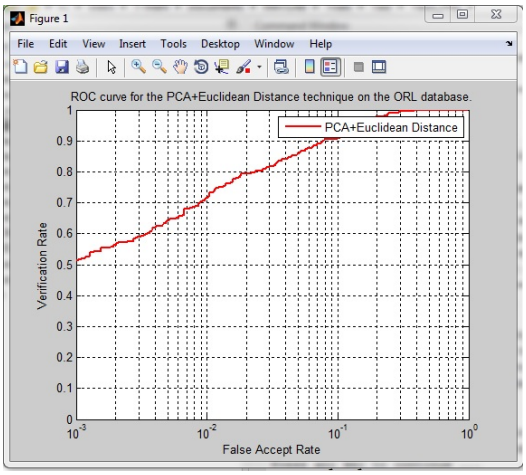


(b) PCA Eigenfaces

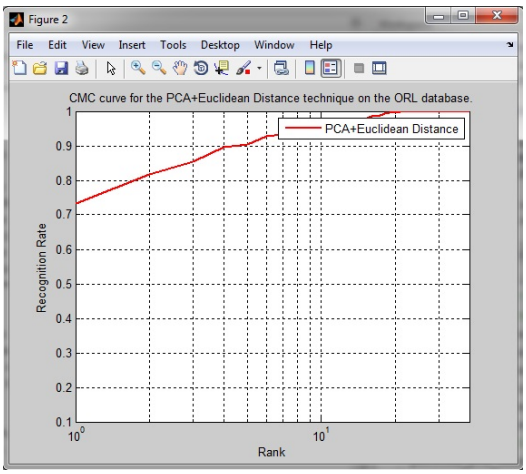
Figure 4.1: PCA Implementation Images

Figure 4.1a shows the Mean face of all the images used for training in the ORL database and Figure 4.1b shows selected Eigenfaces of the images used for training in the ORL

database.



(a) ROC Evaluation Graphs

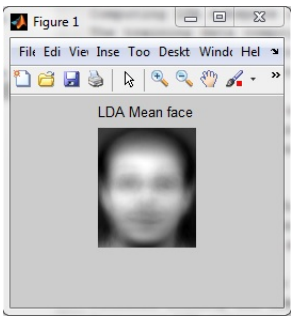


(b) CMC Evaluation Graphs

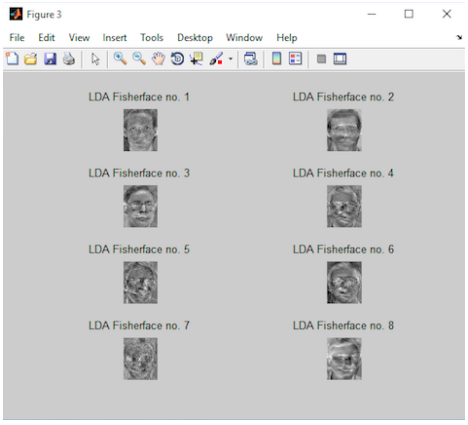
Figure 4.2: PCA Evaluation Graphs

For PCA, the verification rate from the ROC chart is 51% and the recognition rate is 71%.

4.1.2 Linear Discriminant Analysis



(a) LDA Mean Face

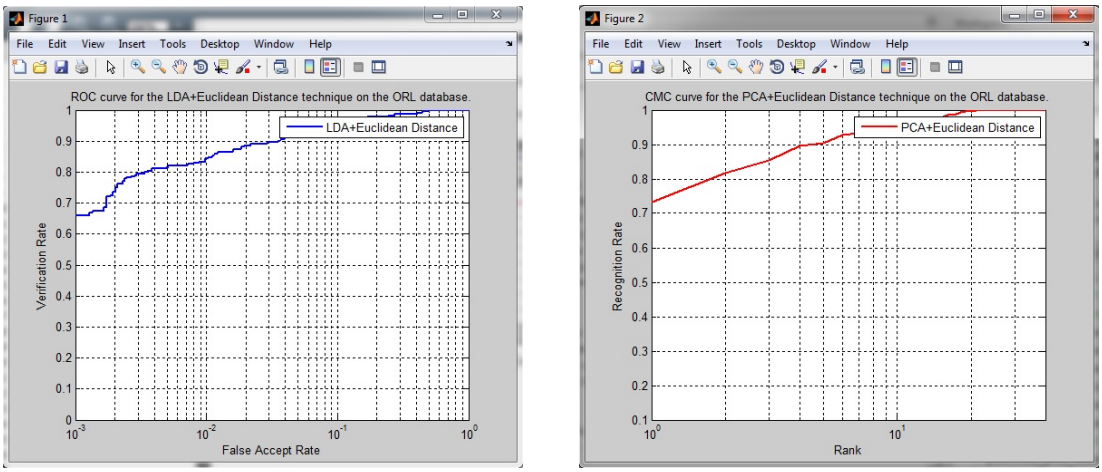


(b) LDA Fisherfaces

Figure 4.3: LDA Implementation Images

Figure 4.3a shows the LDA Mean face of all the images used for training in the ORL

database and Figure 4.3b shows the Fisherfaces of all the images used for training in the ORL database.



(a) ROC Evaluation Graphs **(b)** CMC Evaluation Graphs

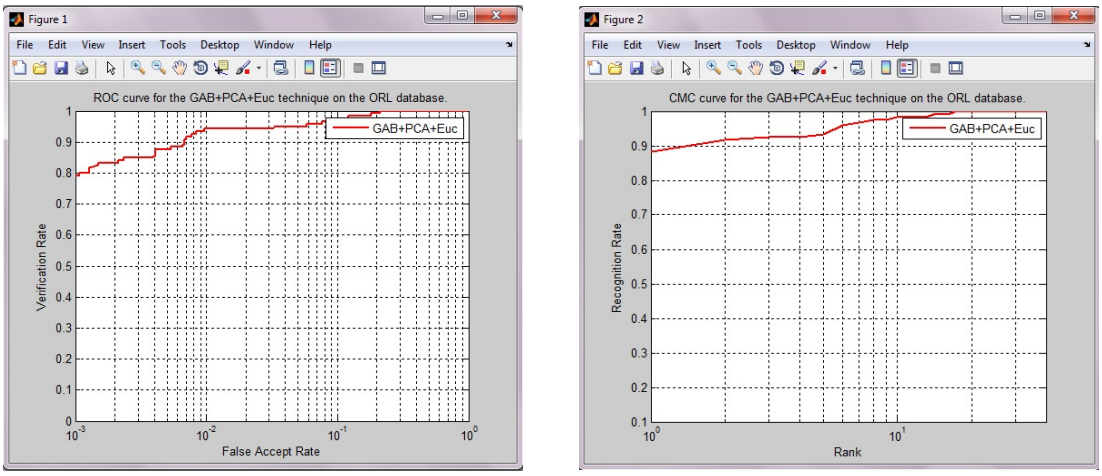
Figure 4.4: LDA Evaluation Graphs

For LDA, the verification rate from the ROC chart is 65% and the recognition rate is 83%. When compared to the PCA, it can be seen that the LDA method functions better on the ORL database than the PCA method.

4.1.3 Implementing Gabor wavelets with PCA

Figure 4.5: Extracting the features from images in the ORL Database

Figure 4.7 shows the Gabor wavelets extraction process for each image in the ORL database.



(a) ROC Evaluation Graph **(b)** CMC Evaluation Graph

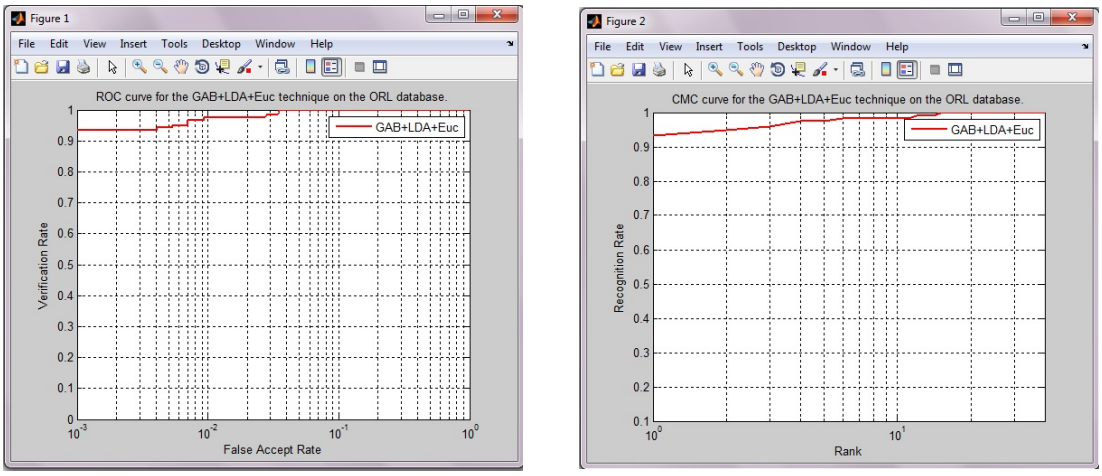
Figure 4.6: Gabor PCA Evaluation Graphs

For GaborPCA, the ROC verification rate is 79% and the recognition rate is 88%. This is higher than both the standalone PCA and LDA methods.

4.1.4 Implementing Gabor wavelets with LDA

Figure 4.7: Extracting the features from images in the ORL Database

Figure 4.7 shows the Gabor wavelets extraction process for each image in the ORL database.



(a) ROC Evaluation Graph **(b)** CMC Evaluation Graph

Figure 4.8: Gabor LDA Evaluation Graphs

For GaborLDA, the verification is 80% and the recognition rate is 89%. Based on this, it can be said that the GaborLDA method outperforms the other three methods when using the ORL database.

4.2 Combined Methods Results

Figure 4.9 below shows the selected images. Each person has a set of five images with an associated test image. In order to test the effectiveness of our combined algorithm, a test was carried out using each of our test images. The Euclidean distances and the output rankings for each of the methods were collated in a table and the rankings processed. The ideal would be that for each table shown in the test cases, the images that match the test image would be ranked between 1 - 5. However as will be seen, this very rarely is the case when dealing with algorithms. Please note that for the rankings go from 1 - 25. An imaged ranked 1 by any of the four methods is an image which has been identified as the closest match to the test image by that method. Accordingly, an image ranked 25, by a method is considered the least closest match to the test image. The combination methods of Product and Averages also have rankings listed alongside for comparison.

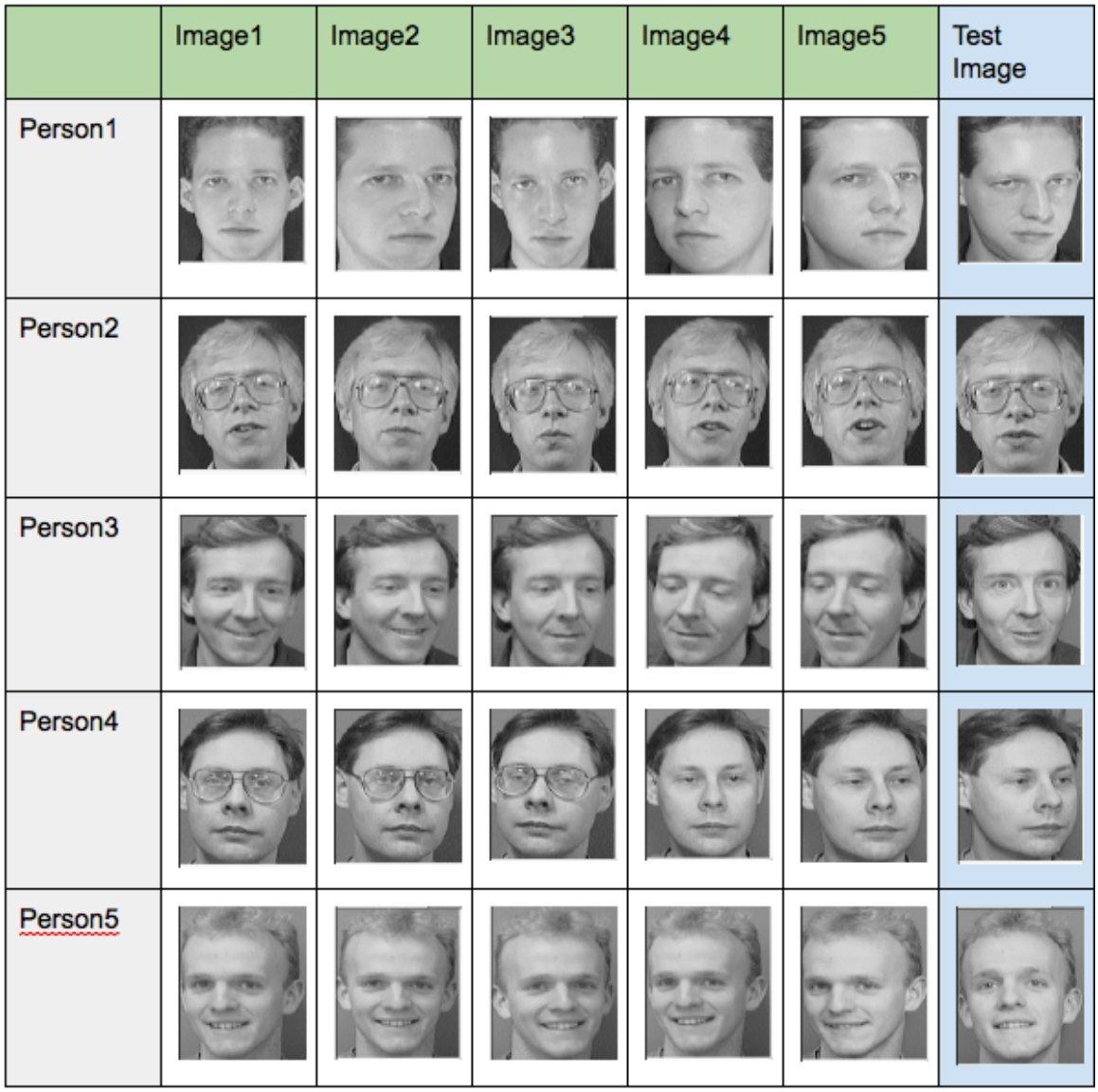


Figure 4.9: DB Images

4.2.1 Test Case 1

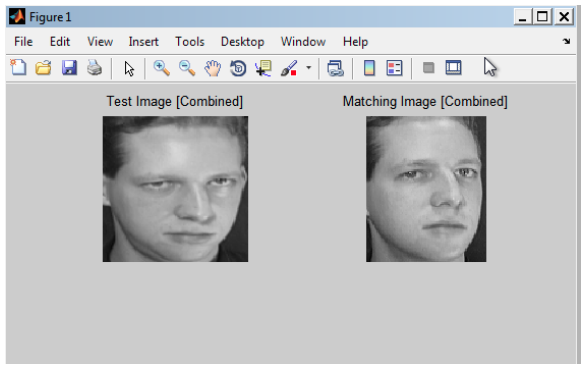


Figure 4.10: TC1 - Recognition Result

Figure 4.10 shows the test and equivalent image for test case 1.

	PCA	PCA-Rank	LDA	LDA-Rank	GaborPCA	GPCA-Rank	GaborLDA	GLDA-Rank	PRODUCT	Prod-Rank	AVERAGE	Avg-Rank
img1	1.2551	2	0.8895	16	0.5344	4	4.3312	8	2.58403844	5	1.75255	7
img2	1.7228	11	0.8525	13	0.4579	2	2.8657	2	1.927217	2	1.474725	2
img3	1.5082	4	0.7669	7	0.5133	3	4.5183	9	2.682526381	6	1.826675	8
img4	1.6805	10	0.9237	17	0.9787	21	7.1806	24	10.90887043	17	2.690875	22
img5	0.6628	1	0.2962	1	0.444	1	2.5818	1	0.2250469443	1	0.9962	1
img6	1.6387	8	0.65	5	0.9302	20	4.9996	13	4.953639582	10	2.054625	11
img7	1.6329	6	0.4809	2	0.7248	11	3.6557	3	2.080669493	3	1.623575	3
img8	1.5789	5	0.6217	4	0.8163	16	4.9201	12	3.942386676	9	1.98425	10
img9	1.6777	9	0.5434	3	0.6881	8	4.0707	7	2.553610137	4	1.744975	6
img10	1.7494	12	0.8605	14	0.822	17	5.4645	15	6.76179881	13	2.2241	13
img11	2.2045	18	1.3618	24	0.7137	10	6.1749	20	13.2302807	21	2.613725	20
img12	1.9496	15	1.2228	20	0.6941	9	3.7088	5	6.13700398	11	1.893825	9
img13	1.8113	14	1.1481	18	0.7589	13	4.5843	10	7.234819281	15	2.07565	12
img14	2.3618	20	1.1523	19	1.0637	25	8.416	25	24.36315713	24	3.24845	25
img15	2.8384	25	1.4009	25	1.0071	22	7.0298	22	28.15116024	25	3.06905	24
img16	2.6118	21	1.2801	22	0.8073	15	4.6928	11	12.66633043	19	2.348	15
img17	2.0837	17	1.2585	21	0.669	7	7.0851	23	12.42969619	18	2.774075	23
img18	2.6396	22	1.2969	23	0.8841	18	5.4227	14	16.41200268	23	2.560825	19
img19	1.6368	7	0.7686	8	0.6061	6	3.7623	6	2.868756607	8	1.69345	5
img20	1.4135	3	0.8786	15	0.5851	5	3.7042	4	2.691606307	7	1.64535	4
img21	2.6795	23	0.8294	10	1.0598	24	5.9929	19	14.11493032	22	2.6404	21
img22	2.2296	19	0.8511	12	0.8007	14	5.5377	17	8.414083145	16	2.354775	16
img23	2.0465	16	0.7972	9	0.7404	12	5.5089	16	6.654421988	12	2.27325	14
img24	2.7178	24	0.8308	11	1.029	23	5.5924	18	12.99354288	20	2.5425	18
img25	1.7629	13	0.732	6	0.9028	19	6.1804	21	7.200238681	14	2.394525	17

Figure 4.11: Person 1 Ranking

In test case one, images 1 - 5 would be considered as good matches for the test image. the table shown in figure 4.11, it can be seen that by using the combined method, the product method was able to recognize image 1 to be part of the five images associated with the test image where the LDA and GaborLDA were off. This can be said also for image 2 where the PCA and LDA methods were very inaccurate. By using the combined

method, the images were successfully recognized as equivalent to the test image. Image 4 was inaccurately recognized by all four methods, and as such the combined method could not in any way improve it.

4.2.2 Test Case 2

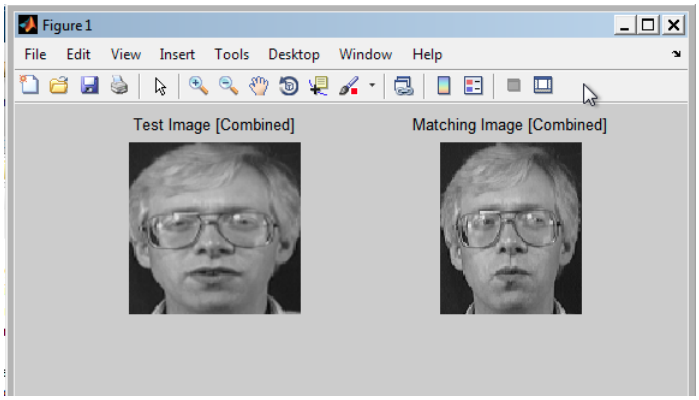


Figure 4.12: Test Case 2

Figure 4.12 shows the test and equivalent image for test case 2.

	PCA	PCA-Rank	LDA	LDA-Rank	GaborPCA	GPCA-Rank	GaborLDA		PRODUCT		AVERAGE	
img1	1.4564	6	0.9102	14	0.9178	9	0.7081	13	0.8615096554	7	0.998125	6
img2	2.6366	24	1.0238	19	1.2065	23	0.5669	6	1.846261257	19	1.35845	24
img3	1.4748	7	0.9633	16	1.125	16	0.7752	18	1.238970528	14	1.084575	11
img4	2.8178	25	0.8548	13	1.1773	21	0.7841	19	2.22348025	20	1.4085	25
img5	2.4757	23	0.8345	12	1.1638	20	0.5619	5	1.351019889	16	1.258975	18
img6	0.6236	3	0.4933	3	0.2726	2	0.2473	2	0.02073801527	2	0.4092	2
img7	0.5831	2	0.4752	2	0.4407	3	0.3976	3	0.04855219845	3	0.47415	3
img8	0.2093	1	0.1995	1	0.0446	1	0.0422	1	0.000078588579	1	0.1239	1
img9	0.6337	4	0.6145	5	0.5362	5	0.5377	4	0.1122722537	4	0.580525	4
img10	0.7138	5	0.7366	10	0.4893	4	0.6509	9	0.1674548557	5	0.64765	5
img11	2.1013	17	1.2181	24	1.1164	15	0.9569	23	2.734370665	25	1.348175	23
img12	2.2237	21	1.1148	23	1.278	25	0.7143	14	2.263000553	22	1.3327	21
img13	2.1029	18	0.9914	17	1.1863	22	0.6489	8	1.604869931	18	1.232375	16
img14	1.7227	11	0.9939	18	1.2659	24	1.1195	25	2.426475117	23	1.2755	19
img15	1.9988	14	1.2599	25	1.1603	19	0.9259	22	2.70545175	24	1.336225	22
img16	2.0751	16	1.0413	20	0.823	6	0.6895	12	1.226165252	13	1.157225	14
img17	2.123	19	1.1126	22	0.9502	10	1.007	24	2.260130658	21	1.2982	20
img18	2.0352	15	1.0875	21	0.8959	8	0.6766	11	1.341614952	15	1.1738	15
img19	2.1503	20	0.8278	11	0.8641	7	0.5807	7	0.8931827113	9	1.105725	12
img20	2.2647	22	0.9362	15	1.0692	11	0.6689	10	1.516350026	17	1.23475	17
img21	1.6897	10	0.5743	4	1.1127	14	0.774	17	0.835732842	6	1.037675	8
img22	1.5776	8	0.7145	9	1.0752	13	0.7225	15	0.8756413016	8	1.02245	7
img23	1.6249	9	0.6946	8	1.075	12	0.8303	20	1.007406897	11	1.0562	9
img24	1.7261	12	0.6502	7	1.1368	18	0.7667	16	0.9781882593	10	1.06995	10
img25	1.8582	13	0.6462	6	1.1365	17	0.8764	21	1.196000107	12	1.129325	13

Figure 4.13: Person 2 Ranking

In test case 2, images 6 - 10 are equivalent matches for the test image. In general, the

four methods were pretty accurate. However, as seen in the table displayed in Figure 4.13, the LDA method was very inaccurate. It did not recognize image 10 as equivalent to the test image. Using the combined method however, both the product and average combination methods were able to recognize image 10 as an equivalent image.

4.2.3 Test Case 3

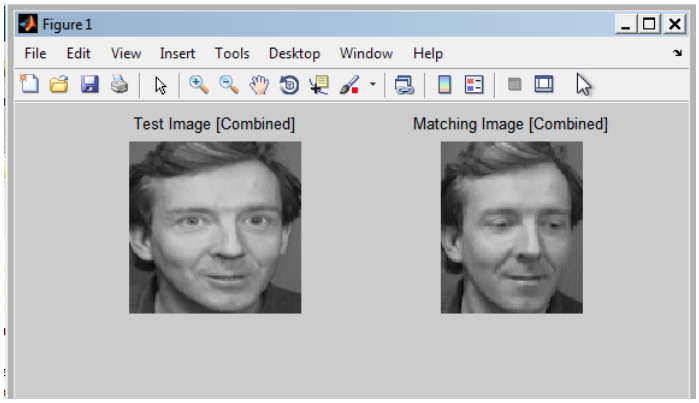


Figure 4.14: Test Case 3

Figure 4.14 shows the test and equivalent image for test case 3.

	PCA	PCA-Rank	LDA	LDA-Rank	GaborPCA	GPCA-Rank	GaborLDA	GLDA-Rank	PRODUCT	Product-Rank	AVERAGE	Avg-Rank
img1	1.9273	13	0.9134	16	0.8558	11	5.7078	15	8.599067498	14	2.351075	13
img2	3.6551	25	1.8169	24	1.0928	23	6.3383	20	45.99851017	24	3.225775	24
img3	2.9751	22	1.6635	23	0.8871	12	6.6512	22	29.20094858	23	3.044225	23
img4	3.0887	24	1.8369	25	1.1501	25	7.4556	24	48.64961922	25	3.382825	25
img5	2.9948	23	1.5236	22	0.962	17	5.1545	7	22.6256156	22	2.658725	21
img6	2.3941	20	1.0397	20	1.0386	21	5.8927	18	15.23396595	20	2.591275	18
img7	2.1509	15	0.9435	18	0.9096	13	5.4143	11	9.994357763	18	2.354575	14
img8	2.1991	16	0.9467	19	0.9474	14	6.5194	21	12.85873849	19	2.65315	20
img9	2.2496	17	0.9205	17	0.8506	10	5.476	12	9.64534828	16	2.374175	16
img10	2.6181	21	1.4504	21	0.9558	16	5.4836	13	19.90246256	21	2.626975	19
img11	0.643	2	0.4434	2	0.3083	2	3.4402	2	0.3023875303	2	1.208725	2
img12	0.6778	3	0.5941	3	0.3983	3	3.6071	4	0.5785349572	3	1.319325	3
img13	0.2633	1	0.2184	1	0.1938	1	1.6445	1	0.01832699003	1	0.58	1
img14	1.4412	7	0.8325	15	1.102	24	7.5149	25	9.936039195	17	2.72265	22
img15	1.706	9	0.707	7	1.0405	22	5.8657	17	7.361399248	11	2.3298	10
img16	1.5491	8	0.7854	14	0.8231	9	4.6894	6	4.696131308	8	1.96175	6
img17	1.0961	6	0.7581	11	0.777	8	6.7263	23	4.342840973	7	2.339375	11
img18	1.7282	10	0.7678	12	0.9498	15	5.7357	16	7.228708329	10	2.295375	9
img19	0.7558	5	0.5968	4	0.5867	4	3.4675	3	0.9176313872	4	1.3517	4
img20	0.6994	4	0.6562	6	0.6723	6	4.5823	5	1.413866759	5	1.65255	5
img21	2.384	18	0.7093	8	0.9903	20	5.3867	10	9.020399644	15	2.367575	15
img22	1.9353	14	0.7258	9	0.7556	7	5.5902	14	5.933139445	9	2.251725	8
img23	1.7336	11	0.7538	10	0.6088	5	5.2188	8	4.151932926	6	2.07875	7
img24	2.3872	19	0.6263	5	0.9867	18	5.3594	9	7.90628595	12	2.3399	12
img25	1.7828	12	0.7724	13	0.9883	19	6.3143	19	8.593278712	13	2.46445	17

Figure 4.15: Person 3 Ranking

In test case 3, images 11 - 15 are equivalent images for the test image. As can be seen in the table displayed above, Images 11 -13 were accurately recognized by the four methods. Images 14 and 15 however were not as accurately predicted. Using the combined method did offer some slight improvements over some of the methods but not for all.

4.2.4 Test Case 4

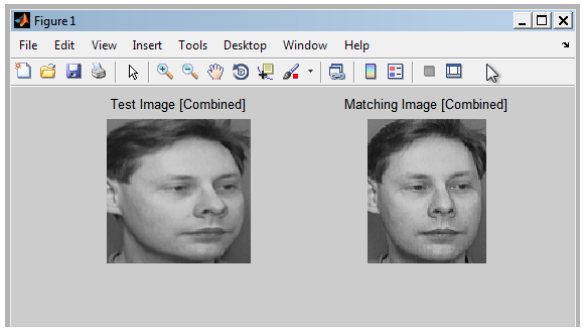


Figure 4.16: Test Case 4

Figure 4.16 shows the test and equivalent image for test case 4.

	PCA	PCA-Rank	LDA	LDA-Rank	GaborPCA	GPCA-Rank	GaborLDA	GLDA-Rank	PRODUCT	Product-Rank	AVERAGE	Avg-Rank
img1	1.9044	11	1.0217	18	0.9183	13	6.1814	13	11.04467646	14	2.50645	14
img2	3.3342	24	1.439	22	0.8269	8	5.3202	9	21.10733446	21	2.730075	18
img3	3.1579	23	1.6986	24	0.8962	11	7.0742	19	34.00726977	24	3.206725	24
img4	3.3564	25	1.8004	25	1.1196	21	7.8376	23	53.02597974	25	3.5285	25
img5	3.0103	22	1.4855	23	0.7035	4	4.2804	4	13.46576069	17	2.369925	9
img6	2.3202	18	1.0335	19	1.2566	25	7.4502	20	22.4492011	22	3.015125	20
img7	2.1221	13	0.8217	9	0.9569	16	5.0509	6	8.427804586	10	2.2379	7
img8	2.2876	17	1.0097	17	1.1384	23	7.6436	21	20.09857575	20	3.019825	21
img9	2.2483	16	0.875	12	0.9662	18	6.0646	11	11.52740384	15	2.538525	15
img10	2.3954	19	1.1565	21	1.1791	24	7.7844	22	25.42725425	23	3.12885	22
img11	1.2839	6	0.8916	13	0.9227	14	7.8435	24	8.284602588	9	2.735425	19
img12	0.9022	4	0.7633	5	0.7113	5	4.0131	3	1.965761729	3	1.597475	3
img13	0.9885	5	0.8327	10	0.7582	6	4.8969	5	3.05611895	5	1.869075	4
img14	1.7588	10	0.9148	14	1.1226	22	8.9253	25	16.12094415	19	3.180375	23
img15	2.1981	15	1.0404	20	1.0061	20	6.6089	17	15.2061097	18	2.713375	17
img16	1.4969	7	0.87	11	0.8411	9	5.3641	10	5.875658411	6	2.143025	5
img17	0.854	3	0.7679	7	0.5783	3	6.4334	15	2.439811563	4	2.1584	6
img18	1.7014	8	1.0075	16	0.9425	15	6.1028	12	9.859660924	12	2.43855	11
img19	0.6606	2	0.6087	2	0.5083	2	3.9571	2	0.8087960215	2	1.433675	2
img20	0.1551	1	0.1539	1	0.1449	1	0.9745	1	0.003370549011	1	0.3571	1
img21	2.4903	20	0.7666	6	0.9868	19	5.3001	8	9.984669364	13	2.38595	10
img22	2.1806	14	0.939	15	0.8548	10	6.6499	18	11.639149	16	2.656075	16
img23	1.9612	12	0.8119	8	0.8173	7	6.3033	14	8.203022493	8	2.473425	12
img24	2.5218	21	0.7231	4	0.9628	17	5.1055	7	8.963618495	11	2.3283	8
img25	1.7291	9	0.695	3	0.9178	12	6.5644	16	7.240157362	7	2.476575	13

Figure 4.17: Person 4 Ranking

In test case 4, images 16 - 20 are equivalent images to the test image. As can be seen in the table displayed in Figure 4.17, image 16 was inaccurately recognized by all four methods.

However, the combined method using averaging was able to associate it as a match. This is also true for image 17. The PCA and GaborPCA methods correctlly recognized it as an equivalent match, but the LDA and GaborLDA did not. The combination via product however, was able to correctly match it.

4.2.5 Test Case 5

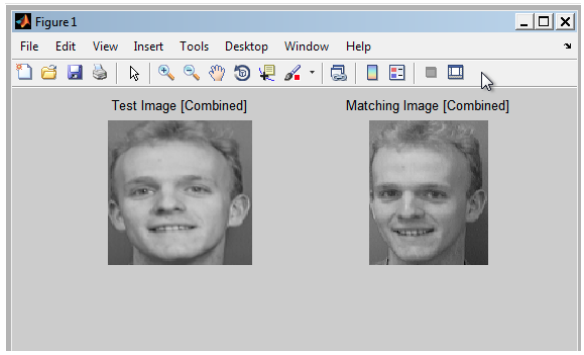


Figure 4.18: Test Case 5

Figure 4.18 shows the test and equivalent image for test case 5.

	PCA	PCA-Rank	LDA	LDA-Rank	GaborPCA	GPCA-Rank	GaborLDA	GLDA-Rank	PRODUCT	Product-Rank	AVERAGE	Avg-Rank
img1	1.551	11	0.8531	16	0.6521	7	5.0886	13	4.390603847	9	2.0362	11
img2	1.7441	14	1.0544	22	1.0161	25	5.5122	18	10.30002307	24	2.3317	22
img3	1.8549	19	0.9909	20	0.7319	14	6.0312	23	8.113454539	23	2.402225	24
img4	1.7703	15	0.829	14	0.8396	20	5.982	22	7.370895236	17	2.355225	23
img5	1.8211	17	0.8254	13	0.9904	24	5.0969	14	7.58778477	19	2.18345	14
img6	2.1634	23	0.6486	8	0.8073	18	5.4291	15	6.150020498	13	2.2621	17
img7	2.0779	22	0.6126	7	0.7632	17	4.8046	12	4.667638724	11	2.064575	13
img8	2.0569	21	0.7279	10	0.7534	16	5.6872	21	6.415182484	15	2.30635	20
img9	2.2049	25	0.6812	9	0.697	11	4.2991	7	4.500635713	10	1.97055	9
img10	2.1695	24	0.876	17	0.744	15	5.4844	17	7.75471459	20	2.318475	21
img11	1.7168	13	1.1045	24	0.7315	13	5.4365	16	7.540829956	18	2.247325	16
img12	1.8424	18	1.0504	21	0.9133	22	4.4357	9	7.839967484	22	2.06045	12
img13	1.8086	16	0.9702	19	0.8944	21	4.3417	8	6.813894403	16	2.003725	10
img14	1.4132	9	0.8798	18	0.7278	12	6.0954	24	5.515715387	12	2.27905	18
img15	1.5456	10	1.0646	23	0.6859	10	5.5833	19	6.301375174	14	2.21985	15
img16	1.399	8	0.7689	11	0.6403	6	4.0597	6	2.796179316	7	1.716975	7
img17	1.9375	20	1.117	25	0.8323	19	7.085	25	12.76187932	25	2.74295	25
img18	1.3517	7	0.8182	12	0.6583	8	4.7855	11	3.484102832	8	1.903425	8
img19	1.3495	6	0.5008	5	0.6703	9	3.9537	4	1.791060026	6	1.618575	5
img20	1.701	12	0.8453	15	0.9653	23	5.648	20	7.839207801	21	2.2899	19
img21	0.8766	4	0.4214	2	0.3407	1	2.7418	1	0.3450673775	2	1.095125	1
img22	0.8423	2	0.4654	4	0.3934	4	2.7961	2	0.431201472	3	1.1243	2
img23	0.8548	3	0.5696	6	0.5322	5	4.5923	10	1.189979872	5	1.637225	6
img24	0.9728	5	0.4515	3	0.3915	3	3.3235	3	0.5714901719	4	1.284825	4
img25	0.4284	1	0.3294	1	0.3763	2	3.9711	5	0.2108716027	1	1.2763	3

Figure 4.19: Person 5 Ranking

For test case 5, images 21-25 are equivalent matches. In general, the four methods did a

good job of identifying most of the images. However, as can be seen in the table displayed in Figure 4.19, the LDA method failed to recognize image 23 as part of the equivalent image being off by 1. The GaborLDA method was even further off. However, via the product combination, this was successfully recognized.

4.3 Analysis of Results

Overall, the algorithms produced very interesting results. On analyzing the ROC and CMC graph plots generated from the implementations of the individual algorithms, we can see that the GaborPCA and GaborLDA methods offered greater verification and recognition rates to PCA and LDA on their own. This is somewhat explainable as GaborPCA and GaborLDA are enhanced variations of the two methods. There was a small percentage of error for each method but that is to be expected considering the number of images (400) which each method processed. For the combined section, selecting only a subset of the entire ORL database allowed for dissertation, processing and display of the results of the combination methods. From our results we can see that in some cases, images that were equivalent to the test images but were not recognized by some or all the methods was correctly identified by the combined methods. Even for cases where the four methods got an image wrong as such leading to poor output for the combination methods, the combination methods were in some way able to improve the rankings for some of the methods. It is worth mentioning that there were some cases where the combination method performed less than the standalone methods. This is to be expected as other of the methods might perform very badly, leading to unfavorable output results for the combination methods. A possible solution to this is mentioned in the future works section of the conclusion chapter. However, based on the overall evidence, it can be said that the combination methods developed in this thesis works and can in fact help in increasing the accuracy of face recognition over some of the individual methods, and in some cases, over all the individual methods.

4.4 Chapter Summary

In this chapter the results based on the implementations which were carried out as part of the methodology were displayed and analyzed. It looked at the results for the implementation of the individual methods before going on to show the results of the methods combined, which was the focus of this thesis. For the combined section, tables of rankings were displayed which helped in looking at the performances of the individual methods against the combined methods.

Chapter 5

Conclusion

This project took a profound look into face recognition. It was seen that currently, there were many methods and approaches to face recognition. In this project, the focus was on four appearance-based face recognition methods which utilized the nearest neighbor (NN) and Euclidean distance approach to face recognition.

In this project, we came up with a method for combining multiple face recognition approaches for the purpose of optimizing and improving the accuracy of face recognition.

Our results showed that for some cases, our approach had significant advantage over existing methods and as such, it can be concluded that the project was a success.

5.1 Future Work

An issue which was encountered in this project was individual methods which performed very badly compared to the other methods affecting the performance of the combination methods. It is believed that if an enhanced method which checks the corresponding rankings of the other methods before deciding whether to use that particular ranking for the combination methods was developed, it would result in greater accuracy for the combined methods as the shortfalls of an individual method would not affect the overall performance of the combination rankings.

Bibliography

[1] Papert, Seymour, *The summer vision project*, http://dspace.mit.edu/handle/1721.1/6125

[2] L. Sirovich, M. Kirby. *A low-dimensional procedure for the characterization of human faces.* Journal of the Optical Society of America 4 (1987) 519–524.

[3] Deza, Elena; Deza, Michel Marie. *Encyclopedia of Distances. Springer.* (2009), p. 94.

[4] Canny, John, *A Computational Approach To Edge Detection*, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679698, 1986.

[5] D. Gabor. *Theory of communication.* In Jounal of the Institution of Electrical Engineers, vol. 93 , pp. 429–457, 1946.

[6] Lades M et al, *Distortion invariant object recognition in the Dynamic Link Architecture*, 1993. IEEE Trans Comput42(3):300311

[7] Wiskott L et al, "Face recognition by elastic bunch graph matching", 1997, IEEE Trans PAMI 19(7):775779

[8] M. Kirby, L. Sirovich. *Application of the Karhunen-Loeve procedure for the characterization of human faces.* IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1990), pp. 103–107.

[9] Pierre Comon, *Independent component analysis, A new concept?.* 1994. Signal Processing, 36(3):287314 (The original paper describing the concept of ICA)

[10] M. Turk, A. Pentland. *Eigenfaces for Recognition.* Journal of Cognitive Neuroscience 3 (1991), pp. 71–86.

[11] A. Webb. *Statistical Pattern Recognition*, Wiley & Sons (2002).

[12] D. A. Forsyth, J. Ponce. *Computer Vision: A Modern Approach*, Prentice Hall 3.

[13] R. A. Fisher. *The Use of Multiple Measurements in Taxonomic Problems.* Annals of Eugenics 7 (2), pp. 179–188, 1936.

[14] C. R. Rao. *The Utilization of Multiple Measurements in Problems of Biological Classification.* Journal of the Royal Statistical Society. Series B (Methodological) 10 (2), pp. 159–203, 1948

[15] Fan X. ; Verma, B. ; *A comparative experimental analysis of separate and combined facial features for GA-ANN based technique*, Sixth International Conference on Computational Intelligence and Multimedia Applications, 2005.

[16] Shaoning Pang; Daijin Kim; Sung Yang Bang; *Face membership authentication using SVM classification tree generated by membership-based LLE data partition*, IEEE Transactions on Neural Networks, Volume: 16 , Issue: 2 , pp. 436 446, 2005

[17] Mahalanobis, Prasanta Chandra (1936). *On the generalised distance in statistics.* Proceedings of the National Institute of Sciences of India 2 (1): 4955. Retrieved 2012-05-03.

[18] Matlab Website, *Official Matlab*, http://www.mathworks.com/

[19] Acharya and Ray, *Image Processing: Principles and Applications*, Wiley-Interscience 2005 ISBN 0-471-71998-6

[20] K. Fukunaga, L. Hostetler. *The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition.* IEEE Transactions on Information Theory, vol 21, pp. 32–40, 1975.

[21] A. Jain, N. Ratha, and S. Lakshmanan. *Object detection using gabor filters.* Pattern Recognition, vol. 30, pp. 295-309, 1997.

[22] V. Štruc, N. Pavešic. *The Complete Gabor-Fisher Classifier for Robust Face Recognition.* EURASIP Journal on Advances in Signal Processing, vol. 2010, article ID 847680, 2010.

[23] L. Shen, L. Bai. *A review on Gabor wavelets for face recognition.* Pattern Anal. Applic. 9, pp. 273–292, 2006.

[24] C. Liu, H. Wechsler. *Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition.* IEEE Transactions on Image Processing, vol. 11, no. 4, pp. 467–476, 2002.