Fall 2007

A Modular Kernel Linear Discriminant Analysis of Gabor Features for Improved Face Recognition

Neeharika Gudur
Old Dominion University

A MODULAR KERNEL LINEAR DISCRIMINANT ANALYSIS OF GABOR FEATURES FOR IMPROVED FACE RECOGNITION

By

Neeharika Gudur
B. Tech (E.C.E) May 2005, MGIT, Jawaharlal Nehru Technological University, India.

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

ELECTRICAL ENGINEERING

OLD DOMINION UNIVERSITY
DECEMBER 2007

Approved by:

Vijayan K. Asari (Director)

Zia-ur Rahman (Member)

Jiang Li (Member)

ABSTRACT

A MODULAR KERNEL LINEAR DISCRIMINANT ANALYSIS OF GABOR FEATURES FOR IMPROVED FACE RECOGNITION

Neeharika Gudur
Old Dominion University
December 2007
Director: Dr. Vijayan K. Asari

Automatic face recognition is one of the major challenges in computer vision and pattern analysis. This thesis presents an efficient face recognition system that is robust with regards to changes in illumination, facial expressions and partial occlusions. Modular Kernel Linear Discriminant Analysis performed on Gabor Features obtained from the face images is employed for improving face recognition accuracy. A face image is pre-processed using the 2D Gabor wavelet transform to achieve invariance to illumination in images. Modular approaches that divide the pre-processed images into smaller sub-images provide improved accuracy, as the facial variations in an image are confined to local regions. Kernel methods are applied to these images in order to extract the most discriminating features, thus improving the classification accuracy of the system. Dimensionality reduction of these generated higher dimensional features is obtained by applying Linear Discriminant Analysis, thus improving the computational speed.

Performance of the proposed technique is tested and evaluated for face recognition accuracy including factors like changes in illumination conditions, facial expressions and partial occlusions. The AR and FERET databases are used for training and testing processes. Results indicate that the proposed technique has better face recognition accuracy when compared to state of the art techniques like Principal Component Analysis (PCA), Modular Principal Component Analysis (MPCA), Linear Discriminant Analysis (LDA), and Kernel Principal Component Analysis (KPCA). Research is continuing for pose invariant face recognition by considering multiple face recognition modules, trained for different facial views.

Dedicated to my parents Sudersanam and Shobha, and my brother Naveen.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Dr. Vijayan K. Asari for his support, advice and motivation for not only finishing this thesis but also guiding me throughout my masters program.

I would also like to thank Dr. Zia-ur Rahman and Dr. Jiang Li for agreeing to be on my committee and for contributing their valuable time.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Face recognition has received significant attention during the past three decades as one of the most successful applications of image analysis and understanding. There are several commercial, security and forensic applications that require face recognition technologies. These applications include automated crowd surveillance, access control, mug-shot identification, face reconstruction, design of human computer interface, multimedia communications and content-based image database management. There are several commercial face recognition systems available including, 2D systems from Cognitec Systems [1], Visage [2] and 3D systems from A4Vision [3], Geometrix [4], and Genex technologies [5]. However, these systems work under controlled lighting and environmental conditions.

A biometric is a representation of a unique characteristic or feature of an individual that has the potential capability to distinguish between an authorized person and an impostor. Since biometric characteristics are distinctive, they cannot be duplicated. The person to be authenticated needs to be physically present at the point of identification. Biometrics are inherently more reliable and more capable than traditional knowledge based (password or Personal Identification Number (PIN)) and token-based techniques (passport, driver's license, and ID card). Currently there are many biometric technologies used for personal authentication: face, fingerprint, hand geometry, iris, retina, signature, voice, etc. Although other methods of identification (such as fingerprint, or iris scans) are accurate, face recognition has always remained a major focus of

Reference model for this work is IEEE Pattern Analysis and Machine Intelligence research because of its non-invasive nature and because that is how human being primarily identify each other. Since the late eighties there has been an explosive growth in face recognition research because of the practical importance of the topic and theoretical interest of both cognitive scientists and computer vision and pattern recognition researchers.

1.1. Challenges in Face Recognition

Although a great deal of effort has been made to improve face recognition, it still remains a challenge. Frontal face recognition methods under controlled environments such as controlled lighting and controlled pose angle of the face etc have produced nearly 100% accuracy for large databases. Successful 2D face recognition systems have been deployed only under constrained situations like slight variations in pose, expressions and lighting. However, major challenges are yet to be addressed to make face recognition an authentic phenomenon. One major factor limiting the applications of 2D face recognition systems is that the human facial-image potentially has very large intra-subject variations due to:

1. Illumination,
2. Facial expressions,
3. Occlusions due to facial hair, face accessories and other objects,
4. Aging, and
5. Pose.

Various techniques and methodologies have been developed over the past three decades targeting the above challenges.

1.2. Motivation

A large number of face recognition algorithms have been presented in the literature. With a number of different databases available, it is very difficult to compare different face recognition algorithms. Even when the same database is used, researchers may use different protocols for testing. While many of the algorithms perform well on a particular database, they do not achieve similarly good results on other databases.

The basic idea behind this research is to develop a face recognition algorithm that is not only robust enough to handle variations in illumination and facial expressions but also efficient and applicable to real-time applications. Further, after comprehensive testing of the recognition algorithm against a number of different databases and maximizing its performance, it will be implemented as a component of a fully automatic face recognition system, complete with face detection module. In summary, a face recognition system should not only be able to cope with variations in illumination, expression and pose but also recognize a face in real-time.

1.3. Focus of Research and Specific Objectives

The main focus of this research is to make the face recognition process robust with regards to the first three challenges mentioned in section 1.1, i.e., to improve the recognition accuracy on face images that are affected due to illumination and/or facial expressions and/or partial occlusions. It is also possible to extend the recognition technology developed for a frontal face image in this research to other views of a face. However, this would need the face recognition system to be trained for multiple views.

The specific objectives of this thesis can be summarized as follows:

1. Extraction of facial features using Gabor wavelet transform;
2. Development of a modularization technique that improves face recognition for local variations;
3. Enhancement of classification accuracy by applying kernel subspace methods along with Linear Discriminant Analysis (LDA);
4. Integration of the above processes to develop a face recognition system that is robust to varying illumination, facial expressions and occlusions (wearing sunglasses, a scarf, etc);
5. Testing and evaluation of the proposed face recognition system's performance using different face image databases under the conditions mentioned in step 4.

1.4. Organization of Thesis

A general survey of various technologies and methods for face recognition and a thorough survey of the methods relevant to this thesis are presented in chapter 2. Chapter 3 describes the different mathematical expressions and concepts used in this thesis. A detailed explanation of 2D Gabor wavelet used for feature extraction and the Kernel Linear Discriminant Analysis (KLDA) along with centering of the data is also provided in chapter 3.

The implementation of the Gabor wavelet feature extraction combined with the kernel method for improved face recognition is presented in chapter 4. The details of the development and analysis of the modularization technique is presented. Additionally, discussion regarding projection of data into modular linear subspaces as well as projections into high dimensional spaces is presented. A step-by-step description of the algorithm is also presented in chapter 4.

Chapter 5 presents the experimental results and performance comparisons of the proposed face recognition technique with existing methods. An analysis of the effects of changes in lighting conditions, facial expressions and occlusions is also provided in that chapter. Chapter 6 presents conclusions and future works.

CHAPTER 2

LITERATURE SURVEY

Automatic face recognition is one of the fundamental problems with computer vision and pattern analysis, and much research has been occurred over the last two decades [1-5]. Many face recognition technologies based on different methodologies have been developed and documented in the literature. A tremendous amount of work is still in progress to make automatic face recognition technology a reality. A general survey of various technologies and methods for face recognition is presented in this chapter.

The first approach in recognition of faces was the correlation method [6]. Such methods were computationally expensive requiring dimensionality reduction schemes for fast computation. Developments in this field have been categorized broadly into appearance-based and model-based algorithms. The algorithm proposed in this thesis falls into the former category. The classification in appearance-based techniques [7-12] is performed by considering the intensity image as the input. Features are extracted from the given intensity images and the classification is performed on the features with or without an intermediate step of dimensionality reduction in the feature based techniques [13-16].

2.1. Appearance-Based Face Recognition

Appearance-based techniques depend on a representation of images that induces a vector space structure (i.e., the image represented as a 1D/2D vector). These techniques represent an object in terms of several object views. Many view-based approaches use statistical techniques to analyze the distribution of the object image in the vector space and derive an efficient and effective representation (feature space) according to different applications. Given a test image, the comparison between the stored prototypes and the test view is then carried out in the feature space.

Image data can be represented as vectors and hence interpreted as points in a multi-dimensional vector space. One approach to cope with high dimensionality is to reduce the dimensionality by combining features. Linear combinations are particularly attractive because they are simple to compute and analytically tractable. In effect, linear methods project the high-dimensional data onto a lower-dimensional space. Two classical approaches to finding effective linear transformations are explained here. These are:

1. Principal Component Analysis (PCA) – Seeks a projection that best represents the data in the least square sense.
2. Linear Discriminant Analysis (LDA) – Seeks a projection that best separates the data in the least squares sense.

2.1.1. Principal Component Analysis (PCA)

PCA [8] [9] has been one widely used approach for face recognition. Face recognition is treated as a two-dimensional recognition problem, taking advantage of the fact that faces are normally upright and thus may be described by a small set of 2D characteristic views. The face images are represented by a set of eigenvectors obtained from the covariance matrix of the training set. The aim is to find a set of M orthogonal vectors in data space that account for the variance of data as much as possible. Projecting the data from their original N-dimensional space onto an M-dimensional subspace spanned by these vectors performs a dimensionality reduction that often retains most of the intrinsic information in the data. The first principal component is taken to be along the direction with the maximum variance. The second principal component is constrained to lie in the subspace perpendicular to the first. Within that subspace, it points in the direction of the maximum variance. Then, the third principal component (if any) is taken in the maximum variance direction in the subspace perpendicular to the first two, and so on.

Although face recognition is a high level visual problem, there is quite a bit of structure imposed on the task. The key is to take advantage of some of this structure by proposing a recognition scheme [8] based on an information theory approach that seeks to encode the most relevant information in a group of faces. This information will best distinguish the faces from one another. The approach transforms face images into a small set of characteristic feature images, called "eigenfaces", that are the principal components of the initial training set of face images. Recognition is performed by projecting a new image into the sub-space spanned by the eigenfaces ("face space") and then classifying the face by comparing its position in face space with the positions of known individuals.

Hence, a lower dimensional space is found using PCA where each face is described by a shorter vector. An example of PCA is shown below in figure 2.1. The data is initially randomly distributed, as shown in figure 2.1 (a); the data is then correlated to be grouped, and belongs to a certain location in the given coordinate space as shown in figure 2.1 (b).

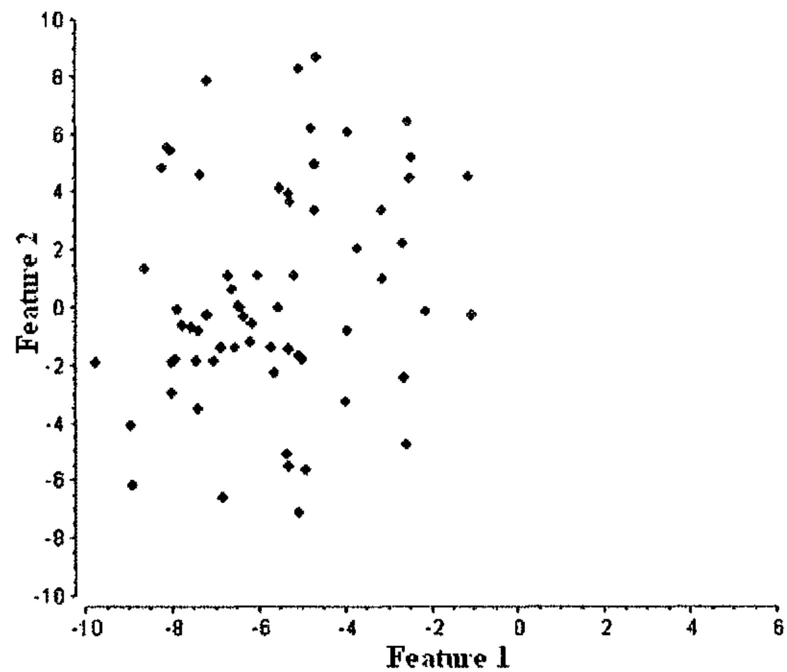

Figure 2.1 (a). An example of PCA - Original data.

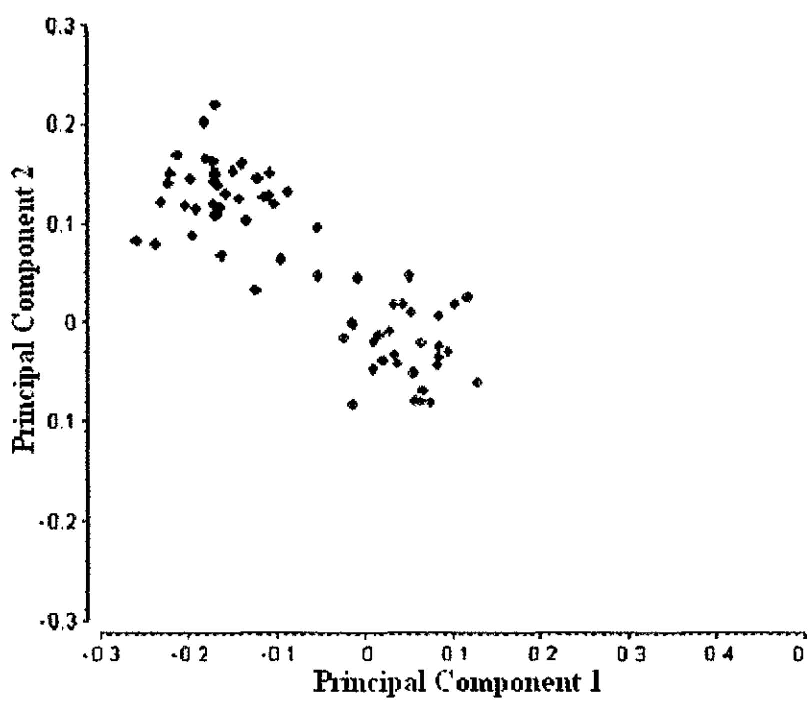

Figure 2.1(b). An example of PCA - Correlated data.

2.1.2. Linear Discriminant Analysis (LDA)

Although PCA finds components that are useful for representing data, there is no reason to assume that these components are also useful for discriminating between data in different classes. Grouping the samples together suggests that the components discarded by PCA might represent distinguishing features that are needed to distinguish between classes. PCA seeks components that are efficient for representation; Linear Discriminant Analysis (LDA), on the other hand, seeks components that are efficient for discrimination.

LDA [10-12] easily handles the case where the Within-Class frequencies are unequal and the performance of LDA has been examined on randomly generated test data. This method maximizes the ratio of the Between-Class variance to the Within-Class variance in any particular set data, thereby guaranteeing maximum separability. The differences between PCA and LDA are discussed in table 2.1.

Table 2.1. Differences between PCA and LDA.

PCA	**LDA**
Does more feature classification	More data classification
The shape and location of the original data sets change when transformed to a different space	Does not change the location. Only tries to provide more class-separability, and draws a decision region between the given classes

Thus, LDA provides a better understanding of the distribution of feature data. An example of LDA is shown in figure 2.2 (a), where a set of randomly distributed data is shown. The discriminating components are formed, and the test vectors are classified accordingly, as shown in figure 2.2 (b).

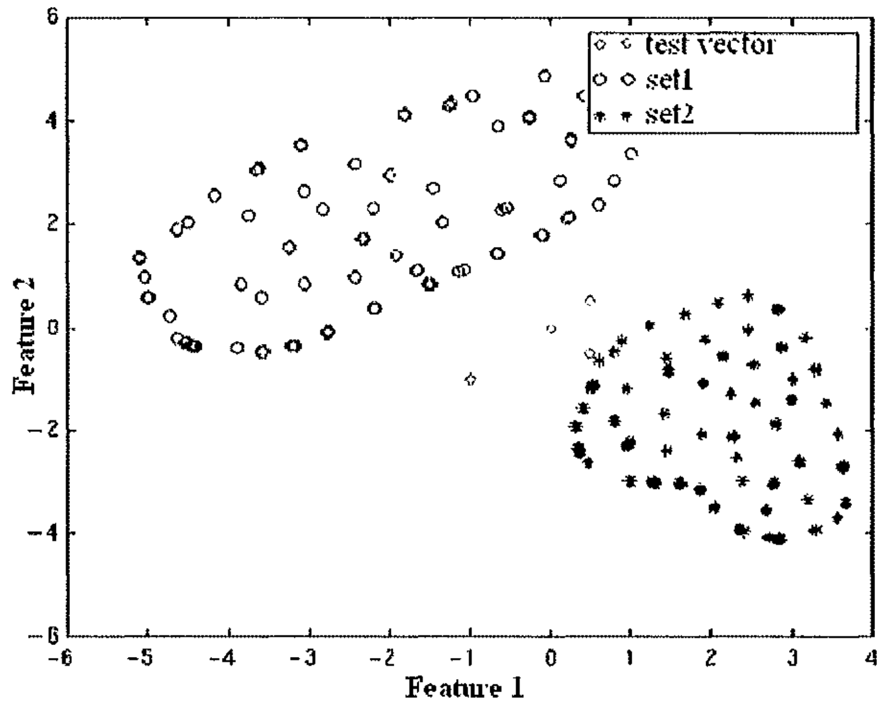

Figure 2.2 (a). LDA - Original data sets and the test vectors.

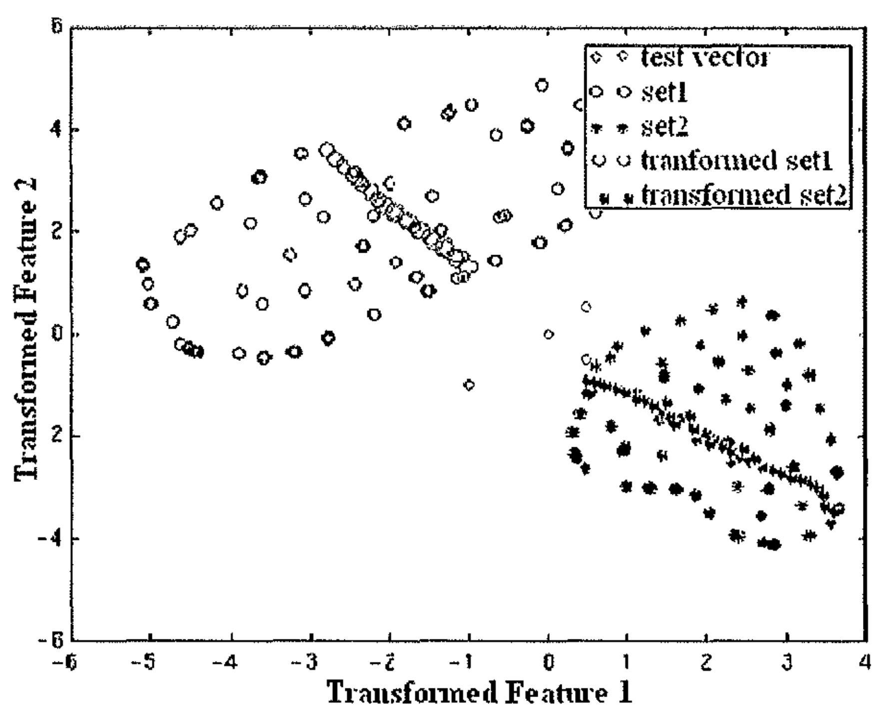

Figure 2.2 (b). LDA - Data sets in original and transformed space along with the transformation axis.

2.1.3. Higher Dimensional Spaces

PCA encodes the pattern information based on second order dependencies (i.e. pixel wise covariance among the pixels) and are insensitive to the dependencies of multiple (more than two) pixels in the patterns. Since the eigenvectors in PCA are the orthonormal bases, the principal components are uncorrelated. In other words, the coefficients for one of the axes cannot be linearly represented from the coefficients of the other axes. Higher order dependencies in an image include nonlinear relations among the pixel intensity values, such as the relationships among three or more pixels in an edge or a curve, which can capture important information for recognition. Explicitly mapping the vectors in input space into higher dimensional space is computationally intensive. Kernel methods [17] [18] have therefore come in handy. Using the kernel trick (section 3.3.1), one can compute the higher order statistics using only dot products of the input patterns. Kernel PCA (KPCA) [17] has been applied to face recognition applications and is observed to be able to extract nonlinear features. The process of obtaining the weights for the input patterns in the KPCA transformed space is described below. The general idea of obtaining the transformed space from the higher dimensional space is described below.

1. The data in vector space X (the input space) is mapped to a vector space H (the feature space) via a nonlinear mapping $\Phi(\cdot): X \rightarrow H$.

2. The algorithm executed on the vector representation $\Phi(x)$ of the data is made.

In other words, non-linear analysis of the data is performed using linear methods like PCA, LDA, etc. The purpose of the map $\Phi(\cdot)$ is to translate non-linear structures of the data into linear ones in H.

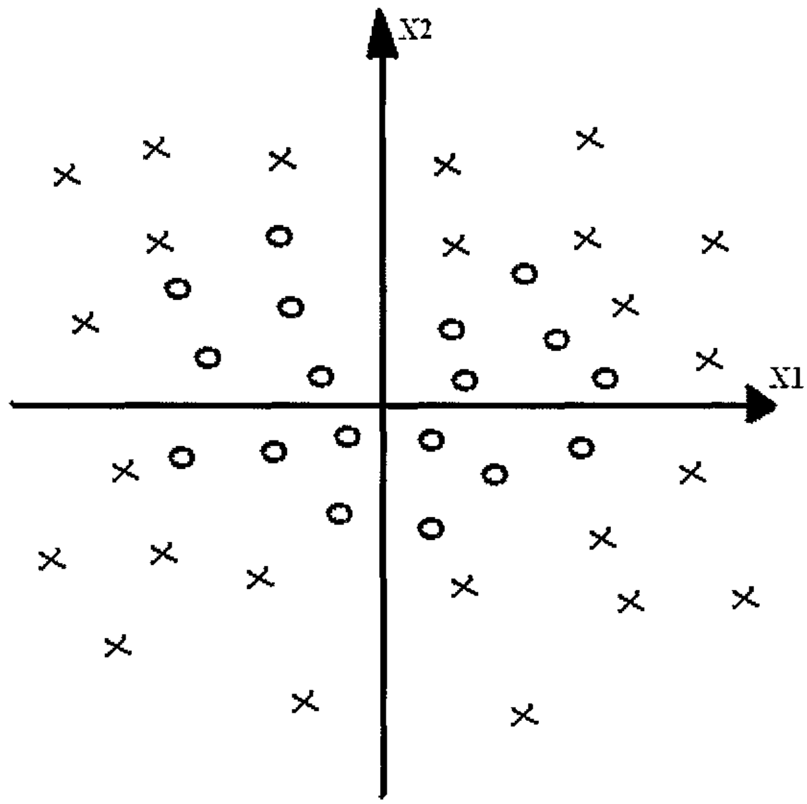

Figure 2.3 (a). 2D representation of the original input data X.

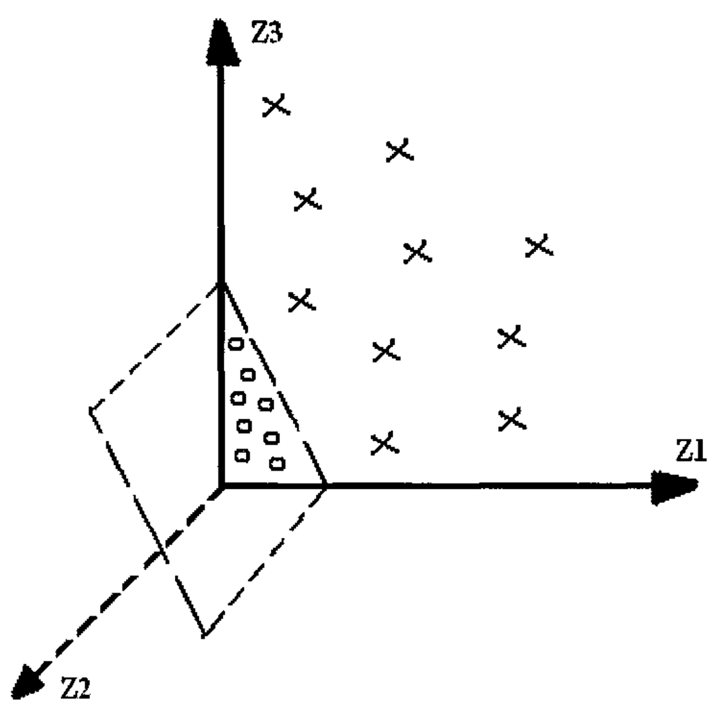

Figure 2.3 (b). Transformed data (3D representation).

Consider the following discrimination problem, shown in figure 2.3 (a), where the goal is to separate two sets of points. In the input space, the problem is non-linear, but after applying the transformation Φ $(\cdot)$ that maps each vector to the three monomials of degree two formed by its coordinates, the separation boundary becomes linear. The transformed data is shown in figure 2.3 (b).

2.2. Feature Based Methods

The importance of facial features for face recognition cannot be overstated. Many face recognition systems need facial features in addition to the holistic face, as suggested by studies in psychology [13]. Feature based techniques are targeted particularly towards achieving illumination invariance in pattern classification. Many feature based face recognition techniques have been developed during recent years [13-16].Three types of feature extraction methods can be distinguished: (1) generic methods based on edges, lines, and curves; (2) feature-template-based methods; (3) structural matching methods that take into consideration geometrical constraints on the features.

2.2.1. Gabor Wavelet Transform

Despite remarkable progress so far, the general task of face recognition remains a challenging problem due to complex distortions caused by various variations in illumination, facial expression and pose. It is widely believed that local features in face images are more robust to such distortions and a space-frequency analysis is often desirable to extract such features [19].With good characteristics of space-frequency localization, wavelet analysis seems to be the right choice for this purpose [20] [21]. In particular, among various wavelet bases Gabor functions provide the optimized resolution in both the space and frequency domains [22] [23]. The Gabor wavelet was originally developed by Gabor (1946) when he proposed to represent signals as a combination of elementary functions. The 2D counterpart of the Gabor elementary function was then introduced by Granlund in [23]. Daugman (1985) [24] reviewed the 2D Gabor wavelet family and showed that this family can well model the 2D receptive-field profiles of simple cells in the mammalian visual cortex. Thus such visual neurons could optimize the general uncertainty relations for resolution in space, spatial frequency and orientation. From an information theoretic viewpoint, Okajima [25] derived the Gabor function as solutions for a certain mutual-information maximization problem. The work showed that the Gabor-type receptive field can extract the maximum information from local image regions. Due to the useful characteristics of Gabor functions, they have been widely and successfully applied for texture segmentation [26], handwritten numerals recognition [27], fingerprint recognition [28] and face recognition [29]. The wide application of Gabor functions has also resulted in different terminologies that may be quite confusing for researchers. Some examples are the Gabor wavelet, Gabor filter, Gabor expansion, Gabor transform and Gabor function. Since this study starts from joint time-frequency analysis of signals, the terminology of the Gabor wavelet is used in this thesis. Gabor Features are used to represent the features extracted by a set of Gabor wavelets when the wavelet family is applied at a certain facial feature point. A detailed survey of Gabor wavelet based face recognition methods, both analytic and holistic, will follow later.

While analytic methods utilize the Gabor Features extracted from prominent feature points for recognition, holistic methods normally extract features from the whole face image. An augmented Gabor Feature vector [29] can be derived by concatenating the Gabor Features at all pixel locations. Since the feature vector consists of all useful information extracted from different frequencies, orientations and locations, this representation can produce discriminant features for recognition. Similar to typical holistic face recognition methods, faces need to be detected and normalized in size and orientation prior to recognition. Various works have shown that such Gabor Features are much more robust than gray-level intensity values with regards to misalignment caused by the normalization procedure [30]. A number of researchers have developed different recognition systems based on this feature vector. In Liu's early work [29], he applied the Enhanced Fisher linear discriminate Model (EFM) on the Gabor Feature vector for face recognition. Results show that the novel Gabor-Fisher Classifier outperformed both PCA and LDA. Since the 40 Gabor filtered images are concatenated together to form a feature vector, the dimension is huge, e.g., 163,840 for 64×64 pixel images. As a result, down-sampling is first used to reduce the dimension to a manageable size. Liu also applied Independent Component Analysis (ICA) [31] to the augmented feature vector and developed a so-called Independent Gabor Feature (IGF) for recognition. The results show that ICA performs significantly better than eigenfaces. One of Liu's recent works [32] utilized KPCA with fractional power polynomial kernel to reduce the dimension of the extracted Gabor Feature vector and enhance the discriminative power at the same time. However, no direct comparison among those proposed approaches is presented. Shen and Bai [33] [34] mapped the augmented Gabor Features to kernel space, i.e., the extracted

Gabor Feature is analyzed by Generalized Discriminant Analysis (GDA), or Kernel Direct Discriminant Analysis (KDDA) for further feature enhancement. Experimental results show that kernel methods achieve much better results than linear methods such as PCA and LDA. The works of both Liu [29] and Shen [33] have shown that Gabor feature based methods can achieve significant improvement over those using raw pixels, which proved the discrimination ability of Gabor Feature. Similar work can also be found in [34], which applies Null LDA (NLDA) to the augmented Gabor Feature vector for recognition. Once the dimension of the extracted feature vector has been reduced and the discrimination ability enhanced by a certain subspace analysis, simple nearest neighbor classifier and Euclidean distance measure can be applied for classification. When the simple Euclidean distance measure seems to be enough, research results do suggest that different distance measures may affect the performance of the system, and an appropriate distance measure has to be chosen for different subspace analysis approaches [29], [33].

2.3. Local Region Based Techniques

The Modular PCA (MPCA) method [35] is one of the appearance-based methods that try to overcome facial variations by exploring the face's local structure. In this method, a face image is first partitioned into several smaller sub-images, and a single conventional PCA is then applied to each of them. There are recent publications [36-39] that use the concept of local region analysis for expression, occlusion and lighting invariant face recognition. In [37] and [38] a weighted distance measure is used that reduces the impact of pixels in the test image that have undergone a significant movement from the corresponding positions in the training images. This technique was mainly targeted towards achieving expression invariance [40, 41].

A number of other interesting approaches have been explored from different perspectives, such as local feature analysis statistical model based [42] and component-based face recognition methods [43, 44]. Examples of the statistical model based scheme are the Hidden Markov Model (HMM) [45] and the Gaussian Mixture Model [46]. Instead of considering face images from a global view, component-based schemes analyze each facial component separately.

2.4. Summary

A detailed literature survey of various 2D face recognition algorithms has been provided in this chapter. The algorithms are broadly classified into appearance-based and feature based techniques. An overview of classical statistical methods, PCA, LDA and kernel methods has been presented. Also, a brief survey of modularization has been presented for exploring the local variations in the face image. Due to robustness with regards to complex distortions caused by various changes in illumination, facial expressions and poses, Gabor wavelets seem to show promise in extracting the local features that are useful for face recognition.

CHAPTER 3

MATHEMATICAL TECHNIQUES AND CONCEPTS

In this chapter, theoretical and mathematical explanations of the various concepts used in this thesis are presented. The concept and importance of the Gabor wavelet transform is provided first; then the different concepts of dimensionality reduction and data analysis are discussed. Finally, the kernel method is presented.

3.1. Gabor Wavelet Transform

The 'quantum principle' of information states that the conjoint time-frequency domain for 1D signals must necessarily be quantized such that no signal or filter occupies less than a certain minimal area (threshold) in it [23]. This minimal area reflects the inevitable trade-off between time resolution and frequency resolution and has a lower bound on their product, analogous to Heisenberg's uncertainty principle in physics.

The 2D counterpart of a Gabor elementary function was first introduced by Granlund in 1978 [23]. In 1985 Daugman [24] showed a surprising equivalence between the 2D Gabor function and the organization and characteristics of the mammalian visual system. By generalizing the time-frequency resolution uncertainty to the 2D domain, Daugman showed that the joint 2D resolution of Gabor wavelets actually achieves the theoretical limit independent of the values of any of the parameters. From an information theoretic viewpoint, in 1998 Okajima [25] derived the Gabor functions as solutions for a certain mutual-information maximization problem. The work shows that the Gabor-type receptive field can extract maximum information from local image regions.

Daugman [24] generalized the Gabor function to the following 2D form:

$$\Psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2}\left[\exp(\frac{-\|k_{u,v}\|^2\|x\|^2}{2\sigma^2})\right]\left[\exp(ik_{u,v}.x) - \exp(-\frac{\sigma^2}{2})\right] \tag{3.1}$$

where σ is the standard deviation of the elliptical Gaussian along u and v. The 2D Gabor function is thus a product of an elliptical Gaussian and a complex plane wave. Mathematically, the 2D Gabor function achieves the resolution limit in the conjoint space only in its complex form. Pollen and Ronner [47] found that simple cells of virtual cortex exist in quadrature-phase pairs.

Daugman [24] and others have proposed that an ensemble of simple cells of virtual cortex is best modeled as a family of 2D Gabor wavelets that sample the frequency domain in a log-polar manner. The decomposition of an image f into these states is called the Wavelet transform of the image. A particular Gabor elementary function can be used as the mother wavelet to generate a whole family of Gabor wavelets.

3.2. Linear Subspace Analysis

3.2.1. Principal Component Analysis (PCA)

The aim of PCA is to identify a subspace spanned by the training images $\{x_1, x_2, x_3, x_4, \ldots\ldots x_M\}$ that could de-correlate the variance of pixel values. This can be achieved by eigen analysis of the covariance matrix $\Sigma = \frac{1}{M-1}\sum_{i=1}^{M}(x_i - \bar{x})(x_i - \bar{x})^T$:

$$\Sigma E = \Lambda E \tag{3.2}$$

where E and Λ are the resultant eigenvectors (also referred to as eigenfaces) and eigenvalues respectively. The representation of a face image in the PCA subspace is then obtained by projecting it onto the coordinate system defined by the eigenfaces [8].

3.2.2. Linear Discriminant Analysis

While the projection of face images into PCA subspace achieves de-correlation and dimensionality reduction, LDA aims to find a projection matrix W which maximizes the quotient of the determinants of S_b and S_w [48].

$$W = \arg_w \left[\max \frac{\left|W^T S_b W\right|}{\left|W^T S_w W\right|} \right] \tag{3.3}$$

where S_b and S_w are the Between-Class Scatter and Within-Class Scatter respectively.

Consider a C-class problem and let N_c be the number of samples in class C. A set of M training patterns from the C class can be defined as {X_{ck}, c=1, 2, ….., C; k=1, 2, ….., N_c }, $M = \sum_{c=1}^{C} N_c$. The S_b and S_w of a training set can be computed as:

$$S_w = \frac{1}{C} \sum_{c=1}^{C} \frac{1}{N_c} \sum_{k=1}^{N_c} (X_{ck} - \mu_c)(X_{ck} - \mu_c)^T \tag{3.4}$$

$$S_b = \frac{1}{C} \sum_{c=1}^{C} (\mu_c - \mu)(\mu_c - \mu)^T \tag{3.5}$$

where μ is the mean of the entire training set, and μ_c is the mean of the class C. It was shown in [49] that the projection matrix W can be computed from the eigenvectors of $S_w^{-1}S_b$. However, due to the high dimensionality of the feature vector in face recognition applications, S_w is usually singular, i.e., the inverse of S_w does not exist. As a result, the original face vectors are first projected to a lower dimensional space by PCA, which is then subjected to LDA analysis. Let W_{pca} be the projection matrix from the original image space to the PCA subspace. The LDA projection matrix W_{lda} is thus composed of the eigenvectors of $(W_{pca}^T S_w W_{pca})^{-1}(W_{pca}^T S_b W_{pca})$. The final projection matrix W_{mfd} can thus be obtained by:

$$W_{mfd} = W_{pca} W_{lda} \tag{3.6}$$

Note that the rank of $S_b \leq C-1$, while the rank of $S_w \leq M-C$. As a result, it is suggested that the dimension of the PCA subspace should be M-C [50].

3.3. Non-Linear Kernel Subspace Analysis

As per section 3.2, both PCA and LDA are linear methods. Since facial variations are mostly nonlinear, PCA and LDA projections can only provide suboptimal solutions for face recognition tasks [51]. Recently, kernel methods have been successfully applied to solve pattern recognition problems because of their capacity in handling nonlinear data. By mapping sample data to a higher dimensional feature space, a nonlinear problem defined in the original image space is effectively turned into a linear problem in the feature space. PCA or LDA can subsequently be performed in the feature space. Kernel

Principal Component Analysis (KPCA) [17] can also be performed in the feature space. Experiments show that KPCA is able to extract nonlinear features and thus provide better recognition rates in applications such as character recognition [52] and face recognition [53].

Algorithms in feature spaces make use of the following idea, via a nonlinear mapping.

$$\Phi: R^N \rightarrow F \text{ (A mapping from input space, } R^N \text{ to feature space, F)}$$

$$\text{and } X \rightarrow \Phi(x) \text{ (Mapping of each input data X)} \quad (3.7)$$

The data $\{X_k \in R^N, k=1, \ldots.., M\}$ is mapped into a potentially much higher dimensional feature space, F. Higher order dependencies in an image include nonlinear relations among the pixel intensity values (such as the relationship among three or more pixels in an edge or a curve) that can capture important information for recognition. Explicitly mapping the vectors in input space into higher dimensional space is computationally intensive. Kernel trick has therefore come in handy. Using the kernel trick one can compute the higher order statistics using only dot products of the input patterns (explained in section 3.3.1).

The most widely used kernels functions are the Polynomial kernel given in equation (3.8) and the Gaussian kernel given in equation (3.9).

$$K_{i,j} = \{\phi(x_i)\phi(x_j)^T + 1\}^d \quad (3.8)$$

$$K_{i,j} = \exp\{\frac{-\left\|\phi(x_i) - \phi(x_j)\right\|^2}{\sigma}\} \quad (3.9)$$

where d is the degree of the Polynomial kernel, σ is the standard deviation of the Gaussian kernel and $\Phi(x)$ is the mapping from input space to higher dimensional feature space, as explained above.

3.3.1. Kernel Principal Component Analysis

Kernel Principal Component Analysis (KPCA) has been applied to face recognition applications and is observed to be able to extract nonlinear features. The process of the KPCA transformed space is described below.

The sample covariance matrix C, of the data set X_i, where i=1,2,3......N, is given by:

$$C = \frac{1}{N}\sum_i X_i X_i^T \qquad (3.10)$$

The eigenvalues of this matrix represent the variance in the eigen-directions of data space. The eigenvector corresponding to the largest eigenvalue is the direction in which the data is most stretched out. The second direction is orthogonal to the first and picks the direction of largest variance in that orthogonal space and so on. Thus, to reduce the dimensionality of the data, we project the data onto the retained eigen-directions of the largest variance:

$$C = \sum_a \lambda_a (u_a u_a^T) \qquad (3.11)$$

where 'λ_a' are the eigen values and 'u' is the eigenvector.

The projection is given by,

$$y_i = U_k^T X_i \text{ , for all } i \tag{3.12}$$

where U_k is the $d \times k$ sub-matrix containing the first k eigenvectors as columns.

Now, consider that there are more dimensions than data-cases, i.e. some dimensions remain unoccupied by data. The eigenvectors that span the projection space lie in the subspace spanned by the data-cases as follows:

$$\begin{aligned}
\lambda_a u_a &= C u_a \\
&= \frac{1}{N} \sum_i X_i X_i^T u_a \\
&= \frac{1}{N} \sum_i (X_i^T u_a) X_i \\
\Rightarrow u_a &= \sum_i \frac{(X_i^T u_a)}{N\lambda_a} X_i \\
&= \sum_i \alpha_i^a X_i
\end{aligned} \tag{3.13}$$

where u_a is some arbitrary eigenvector of C. It can therefore be assumed that every eigenvector can be expressed exactly as the linear combination of the data-vectors and, hence, lie in its span. This also implies that instead of the eigenvalue equation $Cu=\lambda u$, we may consider the N projected equations,

$$X_i^T C u = \lambda X_i^T u \text{ for all i.} \tag{3.14}$$

From this equation the coefficients α_i^a can be efficiently computed from a space of dimension N as follows,

$$X_i^T C u_a = \lambda X_i^T u_a$$

$$X_i^T \frac{1}{N}\sum_k X_k X_k^T \sum_j \alpha_j^a X_j = \lambda_a X_i^T \sum_j \alpha_j^a X_j$$

$$\frac{1}{N}\sum_{j,k} \alpha_j^a [X_i^T X_k][X_k^T X_j] = \lambda_a \sum_j \alpha_j^a [X_i^T X_j] \tag{3.15}$$

Renaming $[X_i^T X_j]$ as K_{ij} , so that,

$$K^2 \alpha^a = N\lambda_a K \alpha^a$$

$$K\alpha^a = (\lambda_a)\alpha^a \text{ with } \lambda_a = N\lambda_a \tag{3.16}$$

Therefore, an eigenvalue equation for α, which completely determines the eigenvectors u had been derived.

Normalizing u,

$$u_a^T u_a = 1$$

$$\sum_{i,j} \alpha_i^a \alpha_j^a [X_i^T X_j] = 1$$

$$\alpha_a^T K \alpha_a = 1$$

$$N\lambda_a \alpha_a^T \alpha_a = 1$$

$$\|\alpha_a\| = \frac{1}{\sqrt{N\lambda_a}} \tag{3.17}$$

Finally, whenever a new test data vector t comes up, its projections onto the reduced space are computed as,

$$\begin{aligned} u_a^T t &= \sum_i \alpha_i^a X_i^T t \\ &= \sum_i \alpha_i^a K(X_i, t) \\ &= \frac{1}{\sqrt{N\lambda_a}} \sum_i K(X_i, t) \end{aligned} \tag{3.18}$$

where $K(X_i,t)$ is the kernel matrix (i.e., $\Phi(x)$, as described above in section 3.3). The above equation is central to most of the kernel methods. It is called the Dot product kernel.

Therefore, the final equation to kernelize any given data set is given by,

$$K_{i,j} = \phi(x_i)\phi(x_j)^T \tag{3.19}$$

3.3.2. Centering Data in Feature Space

After kernelizing the data, the efficiency of the algorithm only depends on the kernel matrix, so the kernel matrix is centered as,

$$K_{i,j} = \phi_i \phi_j^T \tag{3.20}$$

The features are centered using the equation,

$$\phi_i = \phi_i - \frac{1}{N}\sum_k \phi_k \tag{3.21}$$

Hence, in terms of the new features the kernel is given by,

$$\begin{aligned} K_{i,j}^C &= (\phi_i - \frac{1}{N}\sum_k \phi_k)(\phi_j - \frac{1}{N}\sum_l \phi_l)^T \\ &= \phi_i \phi_j^T - [\frac{1}{N}\sum_k \phi_k]\phi_j^T - \phi_i[\frac{1}{N}\sum_l \phi_l^T] + [\frac{1}{N}\sum_k \phi_k][\frac{1}{N}\sum_l \phi_l^T] \\ &= K_{i,j} - k_i 1_j^T - 1_i k_j^T + k 1_i 1_j^T \end{aligned} \tag{3.22}$$

with $k_i = \frac{1}{N}\sum_k K_{ik}$, $k = \frac{1}{N^2}\sum_{i,j} K_{i,j}$ and 1_i or $1_j = (ones)_i$ or $(ones)_j$

Therefore, the equation to center the data is,

$$K^C(t_i, x_j) = K(t_i, x_j) - k(t_i)1_j^T - 1_i k(x_j) + k 1_i 1_j^T \tag{3.23}$$

Hence, the centered kernel can be computed in terms of non-centered kernel alone without accessing any features.

3.4. Summary

The Gabor wavelet transform is one of the efficient pre-processing techniques to achieve better invariance to changes in illumination and also to extract prominent features. LDA is one the best techniques for dimensionality reduction of the data and also to attain better classification by providing maximum class-separability. Kernel subspace analysis is one of the efficient techniques for dealing with nonlinear data. These techniques form the basis of the algorithm that is presented in chapter 4.

CHAPTER 4

MODULAR KERNEL LINEAR DISCRIMINANT ANALYSIS ON GABOR FEATURES

This chapter deals with the proposed algorithm for performing Modular Kernel Linear Discriminant Analysis (MKLDA) on the Gabor Features. The main steps of this algorithm are:

1. Pre-process each of the input face images in the database by applying the Gabor wavelet transform;
2. Divide each of the Gabor wavelet transformed images into smaller regions called modules;
3. Perform the Kernel Linear Discriminant Analysis (KLDA) on the modules (sub-images) obtained from the previous step;
4. Apply the Minimum Distance Classifier to classify the data, and finally determine the face recognition accuracy;

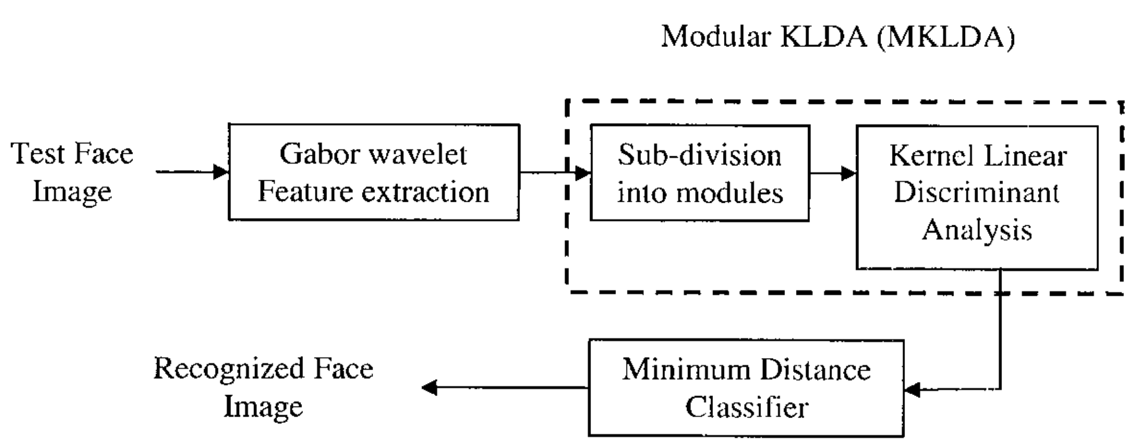

Figure 4.1. Block diagram of the proposed technique.

Figure 4.1 shows a block diagram of the proposed algorithm. This is described in detail by explaining the mathematical interpretation involved in every algorithmic step in the next sections.

4.1. Gabor Wavelet Transform

Use of the 2D Gabor wavelet transform in computer vision was pioneered by Daugman in the 1980s [24]. Also, vonder Malsburg's group has developed face recognition by representing images in terms of Gabor wavelets [54], [55].

Gabor wavelets (filters) characteristics for frequency and orientation representations are quite similar to those of the human visual system. These have been found appropriate for texture representation and discrimination. This Gabor wavelet based extraction of features directly from the gray-level images is successful and has been widely applied to texture segmentation [27] and fingerprint recognition [28].

A complex-valued 2D Gabor wavelet function is a plane wave restricted by a Gaussian envelope:

$$\Psi_{u,v}(z) = \frac{\left\|k_{u,v}\right\|^2}{\sigma^2}\left[\exp\left(\frac{-\left\|k_{u,v}\right\|^2\left\|x\right\|^2}{2\sigma^2}\right)\right]\left[\exp(ik_{u,v}.x)-\exp\left(-\frac{\sigma^2}{2}\right)\right] \qquad (4.1)$$

Here, u and v define the orientation and scale of the Gabor filters, and the wave vector K $_{u,v}$ is defined as: $k_{u,v} = k_v \exp(i\phi_u)$. The multiplicative factor $\|k_{u,v}\|^2$ ensures that filters tuned to different spatial frequency bands have approximately equal energy. Here, $k_v = \frac{k_{\max}}{f^v}$ and $\phi_u = \frac{u\pi}{8}$. K_{max} is the maximum frequency, and f is the spacing factor between the kernels in the frequency domain. These Gabor kernels in equation (4.1) are self similar, as they can be generated from one kernel (a mother wavelet). Each kernel is the product of a Gaussian envelope and a complex plane wave. The first term $exp(ik_{u,v} \cdot x)$ in equation (4.1) controls the oscillatory part of the kernel and the second term $exp(-\sigma^2/2)$ adjusts the mean so that the filters are insensitive to the overall level of illumination, making the kernel independent of the DC component. Hence, when the parameter σ is large enough, it is not necessary to consider the DC effect.

In most applications, the following values are used: $\sigma=2\pi$, $K_{max=}\ \pi/2$, $f=\sqrt{2}$, v ϵ(0,1,2,3,4) and u ϵ (0,1,..7).

The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor wavelet filters as defined by equation (4.1). The convolution of the image with a Gabor wavelet filter is defined as:

$$\Gamma_{k,x}(z) = \Gamma(z) * \Psi_{k,x}(z) \tag{4.2}$$

where $\Gamma_{k,x}(z)$ denotes the result of convolution of the Gabor filter at different orientations and scales, u ϵ (0,1,..7) and v ϵ (0,1,2,3,4) respectively.

In order that this Gabor representation encompasses the information of different frequencies and orientations, a discriminative feature vector Γ^r is derived to represent the image Γ by concatenating the vectors $\Gamma_{k,x}$ that are the results of convolution.

$$\Gamma^r = \left((\Gamma_{0,0})^T (\Gamma_{0,1})^T \ldots\ldots (\Gamma_{4,7})^T \right)^T \tag{4.3}$$

Five different scales and 8 different orientations are used so that there are 40 different Gabor wavelet filters, as shown in figure 4.2. These Gabor wavelet filters are convolved with each image to extract the features. Figure 4.3 shows an image after applying Gabor wavelet transform for one particular scale and orientation. From this Gabor wavelet transformed image, it is seen that variation in illumination is reduced and also the features are more prominent (due to the orientations).

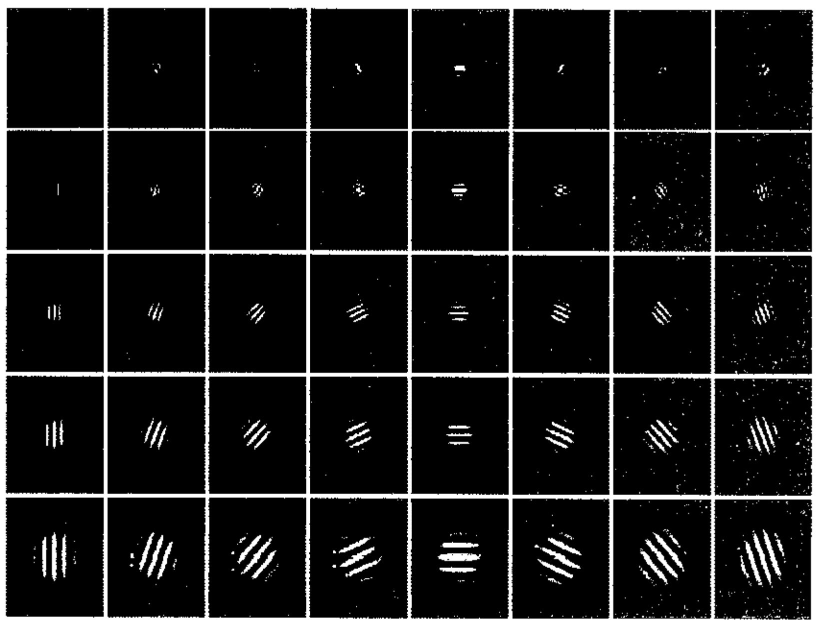

Figure 4.2. The Gabor wavelet filters used; Total of 40 filters i.e 5 × 8 (Scales=5 & Orientations=8).

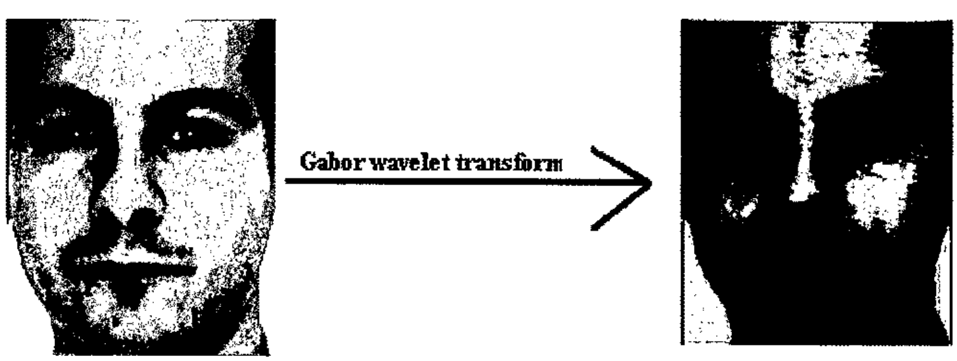

Figure 4.3. Gabor wavelet transformed image.

For an image of size M × N, the number of Gabor wavelet representations will be of the order of M × N × 40. The feature space obtained is hence 40 times larger than the original space. The Gabor wavelet representations reside in a higher dimensional space, so it is important to reduce the feature space to a lower dimensional representation. In this thesis, Kernel Linear Discriminant Analysis (KLDA) is used to reduce the dimensions and also to select the most discriminating features among these Gabor Features.

4.2. Modular Approach

Most of the appearance-based techniques, namely PCA [8] and LDA [10] [12], are not very effective under the conditions of varying illumination, facial expression, pose, etc, as they consider the global information of an image and represent it with a set of weights. Under these conditions the weight vectors will vary considerably from the weight vectors of the images with normal pose and illumination. Hence, it is difficult to identify the test images correctly. If, however, the face images were divided into smaller regions and the weight vectors computed for each of these regions, then the weights would be more representative of the local information of the face. When there is a variation in the pose or illumination only some of the face regions will vary while the rest of the regions will remain the same. Hence, weights of the face regions not affected by varying pose and illumination will closely match with the weights of the same individual's face regions under normal conditions. It is therefore expected that improved recognition rates can be obtained by following a modular approach.

A standard face database consists of face images, I_1, I_2, I_3, ...I_Z, each image of size M × N. Each of these face images is represented as a vector I_v (a column or a row matrix). The face image is segmented into modules using one of the two division strategies: non-overlapping modules and overlapping modules. These face images are resized to a size of L × L pixels. In non-overlapping module strategy there is no overlap between one module and another module, but with overlapping modules certain features that might have been overlooked in a particular module will be included by another neighboring overlapping module.

A non-overlapping module strategy is used in this thesis. Each image can be divided into modules (sub-images) of size 32 × 32 pixels, so a total of 4 modules are generated for an image of size 64 × 64 pixels. Hence 4 (32 × 32), 16 (16 × 16), 64 (8 × 8) modules can be generated using the non-overlapping strategy. With the increase in the number of modules, more detailed local features are obtained, but the performance will degrade if the face image is divided into very small regions as the global information of the face is lost.

Every image in the database is divided into N_M number of smaller sub-images. The size of each sub-image is L^2/N (size of each image being L × L).The sub-images are represented as:

$$\mathrm{I}_{ij}(m,n) = \mathrm{I}_i\left(\frac{L}{\sqrt{N_M}}(j-1)+m, \frac{L}{\sqrt{N_M}}(j-1)+n\right) \quad \forall i,j \tag{4.4}$$

where i varies from 1 to Z (Z being the number of images in the database), j varies from 1 to N_M (N_M being the number of sub-images), and m and n vary from 1 to $L/\sqrt{N_M}$. Figure

4.4 shows the result of dividing a Gabor wavelet transformed image into 16 smaller images using equation (4.4), with N_M=16.

Here, the size of the module is an important aspect. If the face images are divided into very small modules, the face information may be lost and the recognition accuracy may deteriorate. On the other hand, if the module size is large (comparable to the size of the original face image), the effects of variations (illumination, pose, etc.) would be prominent. Thus, an optimum module size is essential for the success of the method. Chapter 5 presents details about the module size used in this work.

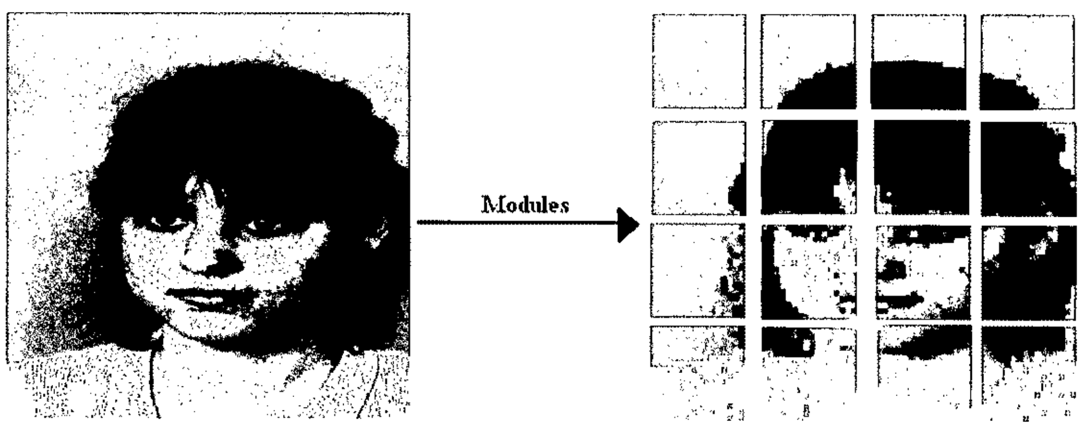

Figure 4.4. Original image of size 64 × 64 is divided into sub-images of size 16 × 16 i.e. N_M=16.

4.3. Kernel Linear Discriminant Analysis

Selection of discriminating features and reduction in dimensions from the Gabor wavelet feature space is described in this section. The LDA combined with the kernel approach, called Kernel Linear Discriminant Analysis (KLDA) is used for this purpose. The algorithmic steps for performing KLDA are described below.

Consider a standard face database consisting of images of size M × N. Let the training set consist of N_T images. Equation (4.5) is applied in order to kernelize the data set.

$$K_{i,j} = I_i \cdot I_j^T \text{ (Dot product kernel) or } K_{i,j} = (I_i \cdot I_j^T + 1)^d \text{ (Polynomial kernel)} \quad (4.5)$$

The kernel matrix $K_{i,j}$ is centered as shown in equation (4.6).

$$K_{i,j}^C = K_{i,j} - K_i 1_j^T - 1_i K_j + 1_i K_j 1_j^T \quad (4.6)$$

Here, $1_i \; or \; 1_j = (ones)_i \; or \; (ones)_j$. The eigenvalues and the corresponding eigenvectors of $K_{i,j}^C$ are calculated, and are normalized as,

$$e_a = \frac{1}{\sqrt{e_a}}, \text{ where } e_a \text{ are the eigenvalues} \quad (4.7)$$

$$E_a^1 = \frac{1}{\sqrt{e_a}} \cdot E_a \text{ , where } E_a \text{ are the eigenvectors} \quad (4.8)$$

Eigenvectors E_a^1 that are associated with the largest M^l eigenvalues are selected so that the images are projected onto a lower dimensional space given by,

$$K_p = [E_a^1]^T . K_{i,j}^C \quad (4.9)$$

LDA of this feature space is performed to retain the most distinguishing features to achieve maximum separability between different classes (individuals). The Within-Class Scatter matrix is computed as:

$$S_w = \sum_{i=1}^{D} S_i \tag{4.10}$$

$$\text{where } S_i = \sum_{k_p \varepsilon D} (K_p - m_i)(K_p - m_i)^T \tag{4.11}$$

Here $m_i = \frac{1}{J}\sum_{k_p \varepsilon d} K_p$, where J is the number of samples in class d, d being the number of classes (individuals). Between-class Scatter matrix is computed as:

$$S_b = \sum_{i=1}^{D} (m_i - M)(m_i - M)^T \tag{4.12}$$

where M is the total mean vector given by $M = \frac{1}{N_T}\sum_{Kp} K_P$. Here N_T where is the total number of images in the training set.

Eigenvectors of the product matrix $(S_w^{-1} S_b)$ are computed, which form the discriminant vectors onto which the data is projected. The weights are computed from the eigenvectors as,

$$W_K = E_P^T . K_P \tag{4.13}$$

Therefore, the final feature space consisting of the most discriminating features with reduced dimensionality is obtained.

4.4. Minimum Distance Classifier

Whenever a test image is presented to the system, KLDA is performed to compute its weights W_{test}. The mean of the weights in the training set for each class d is given by,

$$\Gamma_d = \frac{1}{n_k} \sum_{i=1}^{k} W_{dk} \tag{4.14}$$

where n_k is the number of training images for a particular subject. Next, the minimum distance is computed as,

$$D_d = \left| W_{test} - \Gamma_d \right| \text{ for all d} \tag{4.15}$$

If min (D_d) < θi,, for a particular threshold value for d, the corresponding face class in the training set is the closest one to the test image. Hence the test image is recognized as belonging to that dth face class.

4.5. Summary

An algorithm independent of variations in illumination and facial expressions is presented in this chapter. Face images are pre-processed by Gabor wavelet transform to make them illumination independent and also to extract prominent features (obtained by using different orientations). Further, these images are divided into small modules in order to remove the effect of local variations. Non-linear KLDA is applied to the modularized Gabor feature space to achieve maximum separability and dimensionality reduction. This kernel trick ensures improved efficiency under complex situations such as the presence of occlusions. The efficiency and accuracy of the proposed algorithm is tested against various face databases and presented in the next chapter.

CHAPTER 5

EXPERIMENTAL RESULTS

This chapter describes the various experiments conducted with the proposed algorithm to test its robustness with regards to changes in illumination, facial expressions and occlusions. A detailed analysis of the results is also provided. Two face image databases, AR [56] and FERET [57] are used for the experiments.

5.1. Results for AR Database

The AR database consisting of 40 individuals with 13 images for each individual is considered here. These face images have changes in illumination and facial expressions and have partial occlusions such as the wearing of sunglasses and scarves.

Figure 5.1 (a). AR - Training images.

From the 13 images available for each individual, 5 images that have slight changes in illumination and facial expressions are used for training the algorithm, and the remaining 8 images that have partial occlusions are used for testing. These images are shown in figures 5.1 (a) and 5.1 (b). Therefore, the training set has 200 images (40 × 5), and the testing set has 320 images (40 × 8). Each image is resized to 64 × 64 pixels for ease of computation in the experiments.

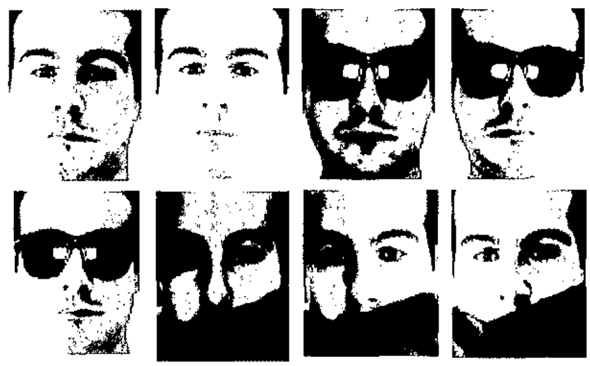

Figure 5.1 (b). Test images.

5.1.1. Results – PCA on Gabor Features

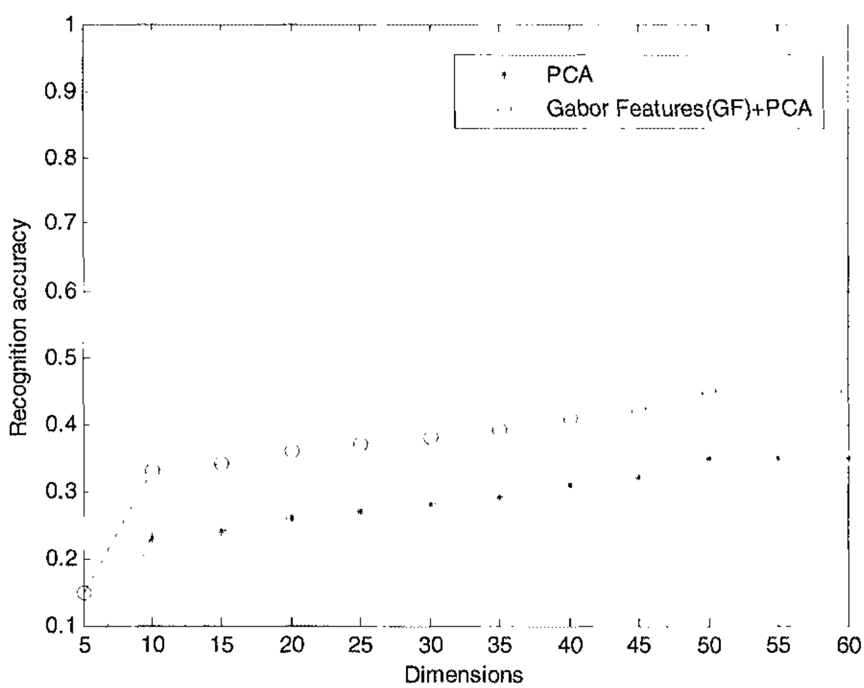

Figure 5.2. Comparison between PCA on original intensity images with PCA on Gabor Features (GF).

The experiments are conducted on the training set from the AR database after Gabor feature extraction. The procedure is tested with the test set from the AR database after performing Gabor feature extraction combined with PCA to determine the recognition accuracy. The graph in figure 5.2 shows a comparison of recognition accuracies obtained for PCA on Gabor Features and for PCA alone (i.e. on intensity images). It is observed that GF+PCA performs better than PCA alone since the Gabor wavelet makes the image illumination independent and also extracts the prominent features by considering different scales and orientations.

5.1.2. Modular PCA versus PCA

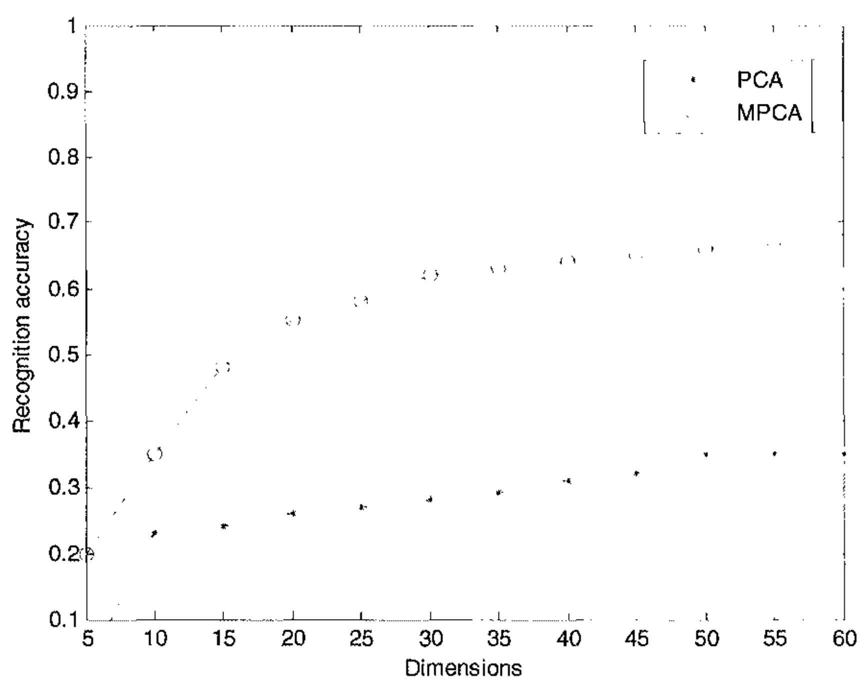

Figure 5.3. Comparison between PCA and Modular PCA (MPCA).

Face images 64 × 64 pixels in size are divided into non-overlapping modules of size 32 × 32 pixels, 16 × 16 pixels and 8 × 8 pixels. Figure 5.3 shows a comparison between Modular PCA and the conventional PCA. Here, it is seen that the accuracy of Modular PCA (MPCA) is significantly better than that of PCA, as the effect due to local variations is reduced.

Table 5.1. Recognition accuracies for different module sizes.

Module Size	Recognition accuracy
32×32	42.90
16×16	58.56
8×8	62.31

Table 5.1 shows the recognition accuracies obtained for different module sizes with 30 eigenvectors. From the results obtained, a module size of 16×16 pixels is found to be optimum, thus maintaining a balance between recognition accuracy and computational complexity.

5.1.3. Results – Modular PCA on Gabor Features

The performance of MPCA on Gabor Features [58] is compared with both PCA and MPCA on original intensity images. From the graph in figure 5.4, it observed that the approach of MPCA on Gabor Features gives a substantial increase of almost 40% when compared to conventional PCA and 20% when compared with MPCA. This increase is due to illumination independence and also reduction in the effect of local variations.

Figure 5.5 shows the graphs obtained for all of the above discussed approaches. It is seen that the MPCA on Gabor Features provides better recognition accuracy than the other methods.

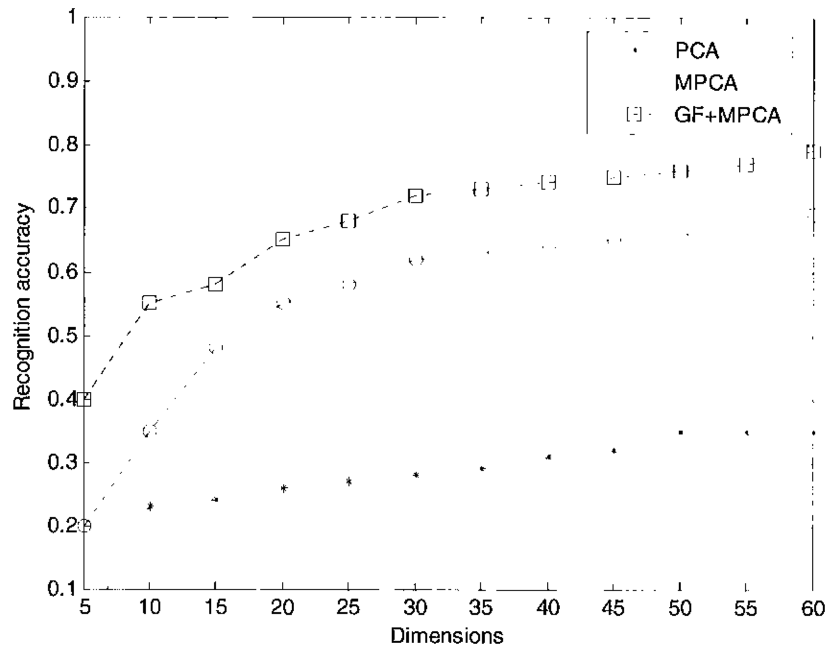

Figure 5.4. Comparison between PCA on original intensity images with Modular PCA on Gabor Features.

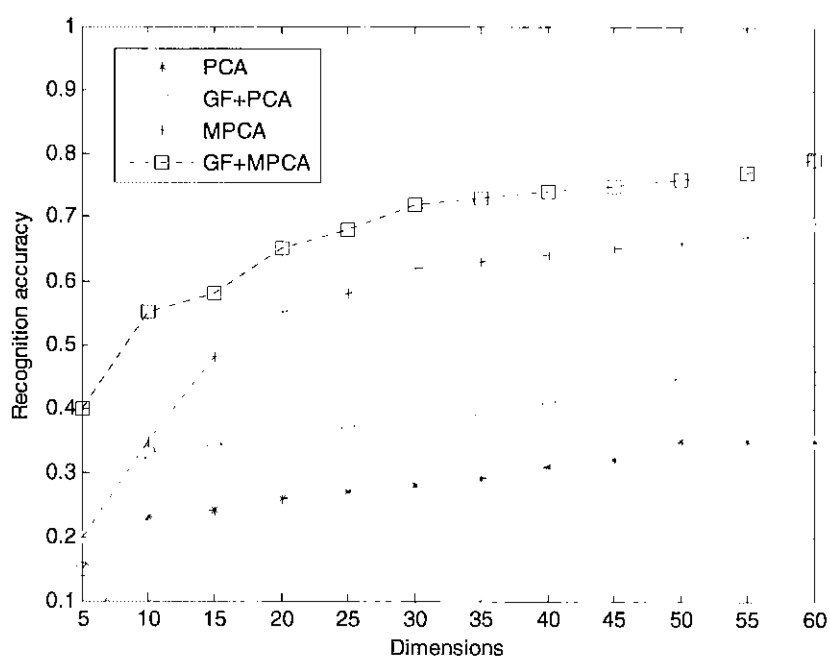

Figure 5.5. Comparison of all PCA with Modular PCA on original intensity images, PCA+Gabor Features and Modular PCA+Gabor Features.

5.1.4. Comparison of Different Dimensionality Reduction Techniques

In this section, different dimensionality reduction techniques are compared with each other. Figure 5.6 shows a comparison between PCA and LDA. It is seen that LDA performs better than PCA as it provides better class-separability that, in turn, increases accuracy.

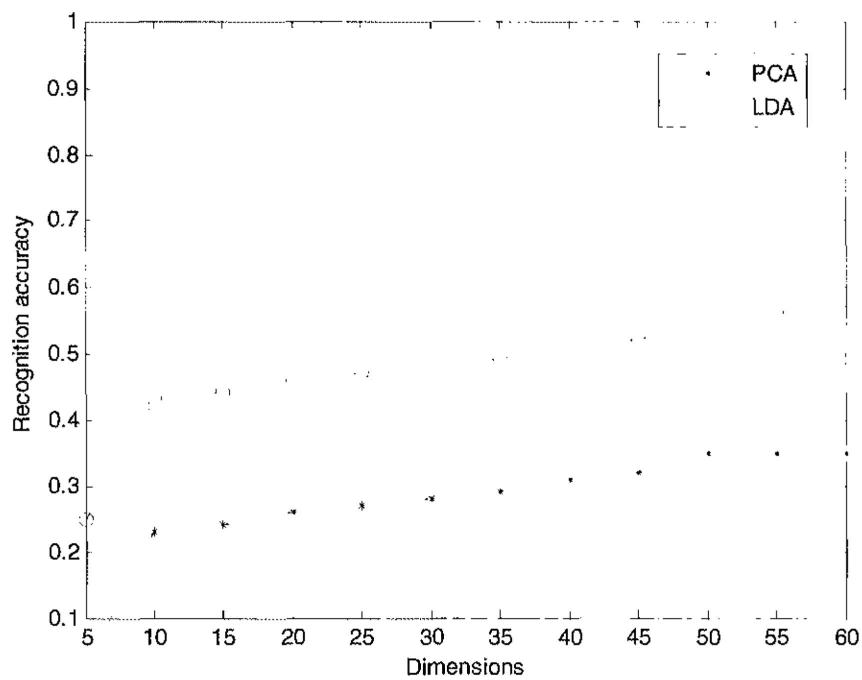

Figure 5.6. Comparison between PCA with LDA.

Figure 5.7 shows a comparison between Kernel PCA and PCA. KPCA works better than PCA for the case of nonlinearity, providing better classification.

Figure 5.8 shows a comparison between PCA and Kernel LDA (KPCA + LDA). It performs much better than PCA, as it works well for non-linear cases along with providing better class-separability. Figure 5.9 shows the comparisons of all of the above discussed methods.

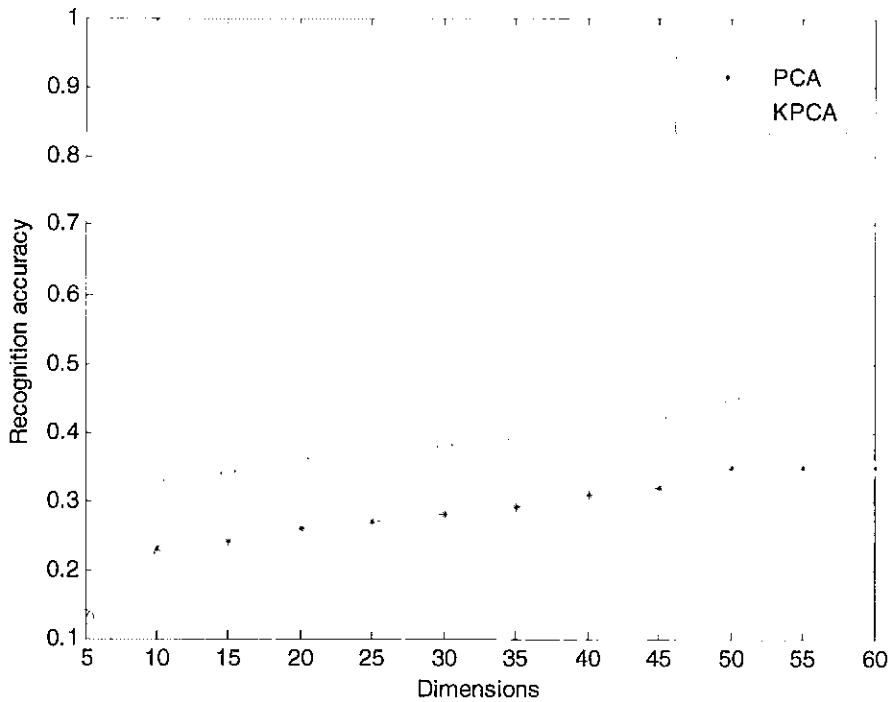

Figure 5.7. Comparison between PCA with Kernel PCA.

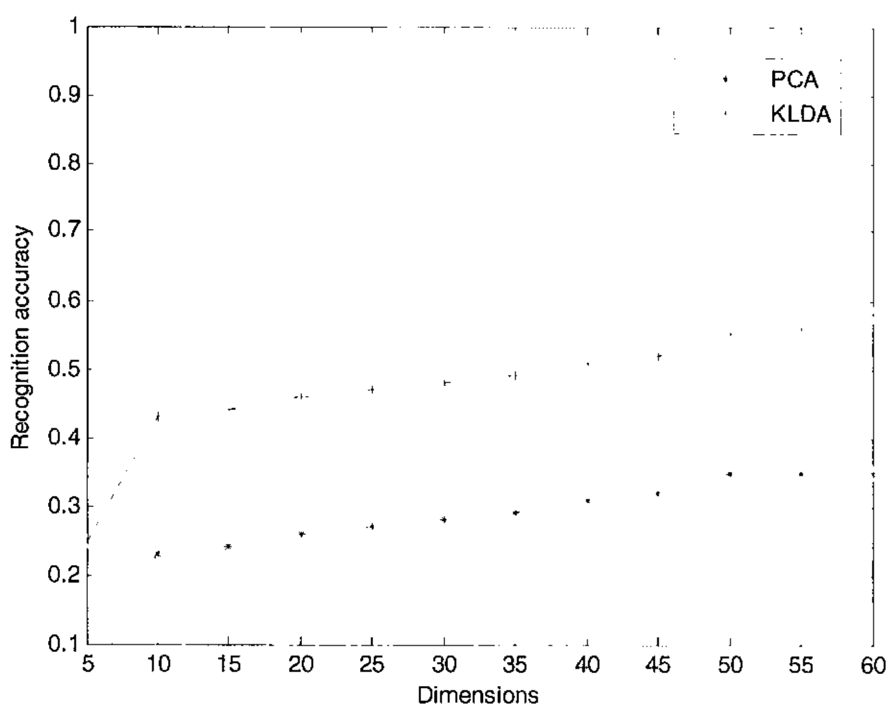

Figure 5.8. Comparison between PCA with Kernel LDA.

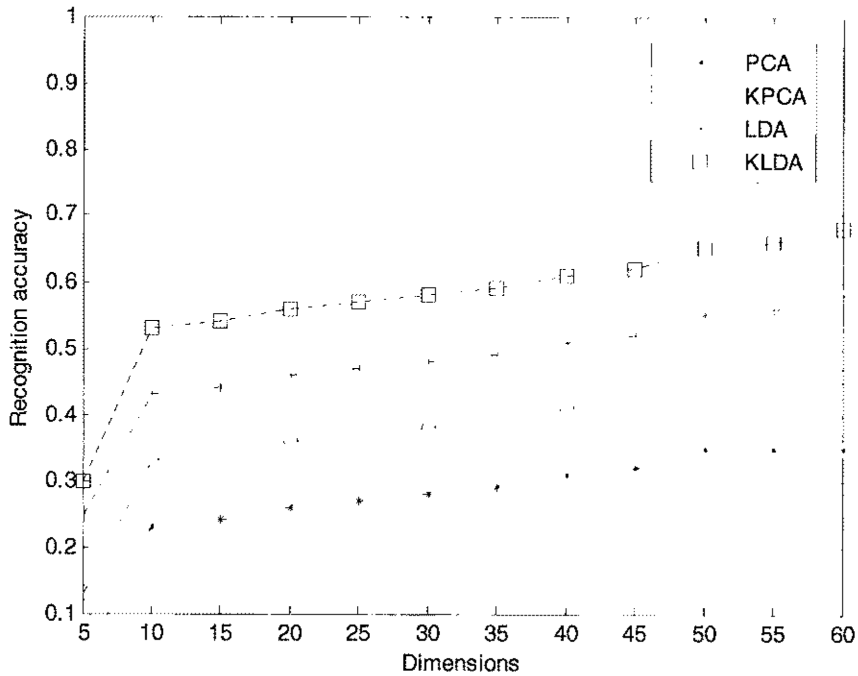

Figure 5.9. Comparison of PCA, LDA, Kernel PCA and Kernel LDA.

5.1.5. Results – MKLDA on Gabor Features

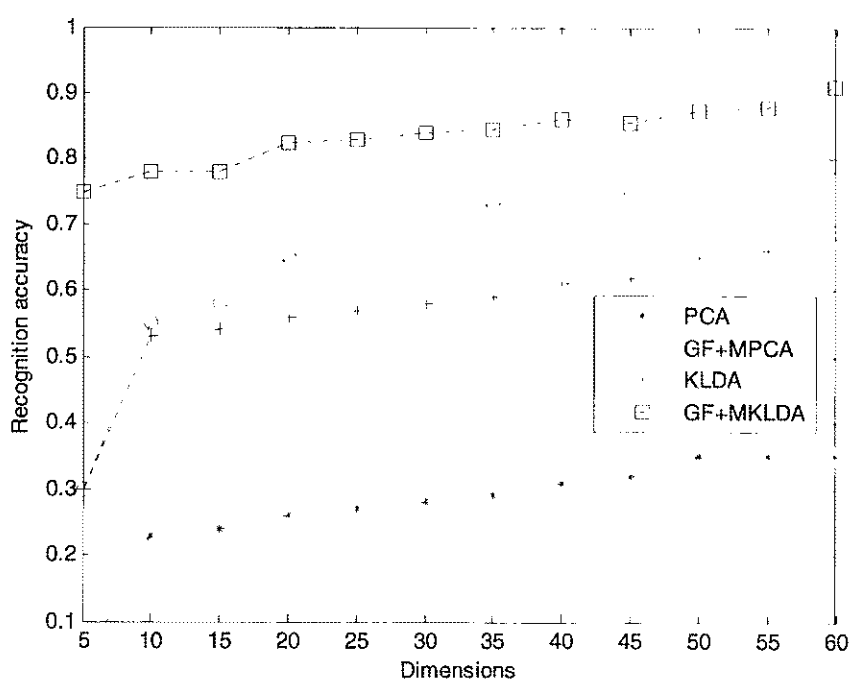

Figure 5.10. A comparison of the proposed method with the above discussed techniques.

Here, the modular approach combined with KLDA and Modular KLDA (MKLDA) on Gabor Features (proposed technique) is performed on the AR database. From the graph in figure 5.10, it is observed that the proposed technique performs much better when compared to the previously discussed techniques. A recognition accuracy of 92.3% is obtained for the proposed MKLDA on Gabor Features for the AR database.

5.2. Results for FERET Database

The FERET [57] face image database is a result of the FERET program, which was sponsored by the U.S. Department of Defense through DARPA. It has become a standard database for testing and evaluating state-of-the-art face recognition algorithms. The proposed algorithm was tested on a subset of the FERET database, consisting mainly of frontal views with variations in illumination, facial expressions and occlusions such as closing the eyes and having a beard. This subset includes 400 images of 50 individuals (each individual has 8 images). Three images of each subject (individual) are randomly chosen for training, while the remaining 5 images are used for testing. Thus, the training sample set consists of 150 images and the testing sample set consists of 250 images. A sample of the training and testing images from this subset is shown in figures 5.11 (a) and 5.11 (b).

Figure 5.11 (a). FERET - Sample Training images.

Figure 5.11 (b). FERET - Sample Test images.

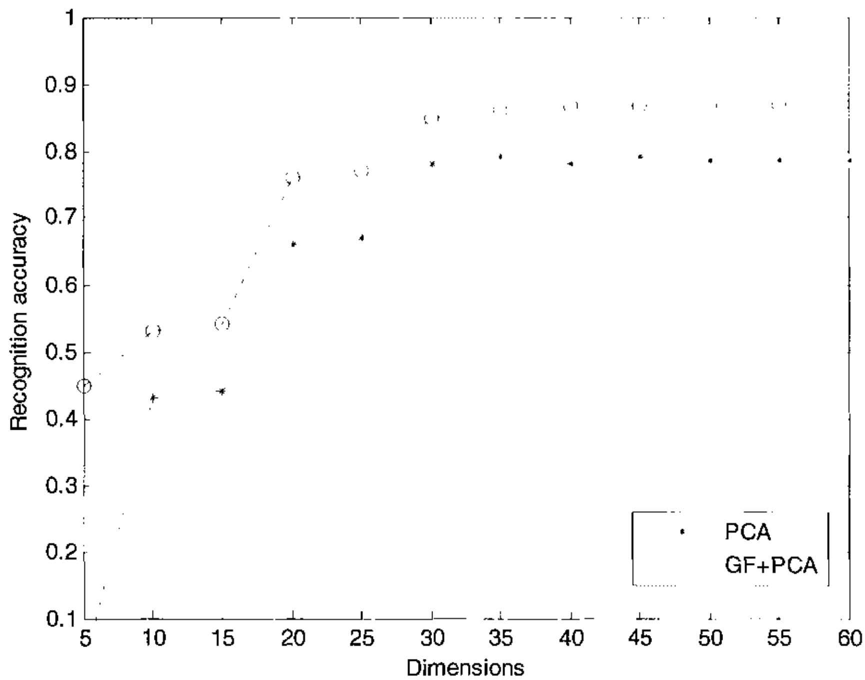

Figure 5.12. Comparison between PCA on Gabor Features with PCA on original intensity images.

The Gabor wavelet filter is applied to each of the face images in the defined FERET subset to extract the features in order to and make the images, illumination invariant and to make features more prominent. The performance of PCA on Gabor Features is compared to PCA on original intensity images, which is shown in figure 5.12. An improvement of recognition accuracy is observed.

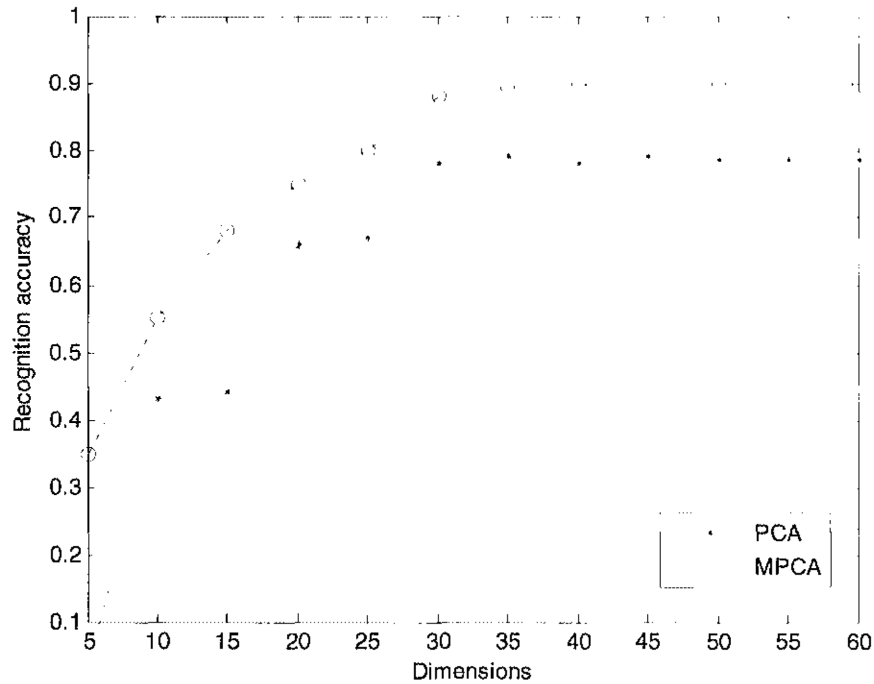

Figure 5.13. Comparison between Modular PCA with PCA.

Figure 5.13 shows the comparison of MPCA with the conventional PCA. Each of the input face images is resized to 64 × 64 pixels and sub divided into modules of size 16 × 16 pixels. The recognition accuracy increases as the module size is smaller, but the computational complexity increases. Hence, there is a trade-off between the module size and the computational complexity from the experiments conducted in this thesis work. A module size of 16 × 16 pixels is found to be optimum. From this graph, it is observed that the Modular PCA provides about 20% more accuracy when compared to PCA.

Experiments for Modular PCA approach on Gabor Features [58] are conducted. The recognition accuracy of this method increases by a considerable amount, i.e the accuracy is nearly 95% for 30 selected eigenvectors as shown in figure 5.14.

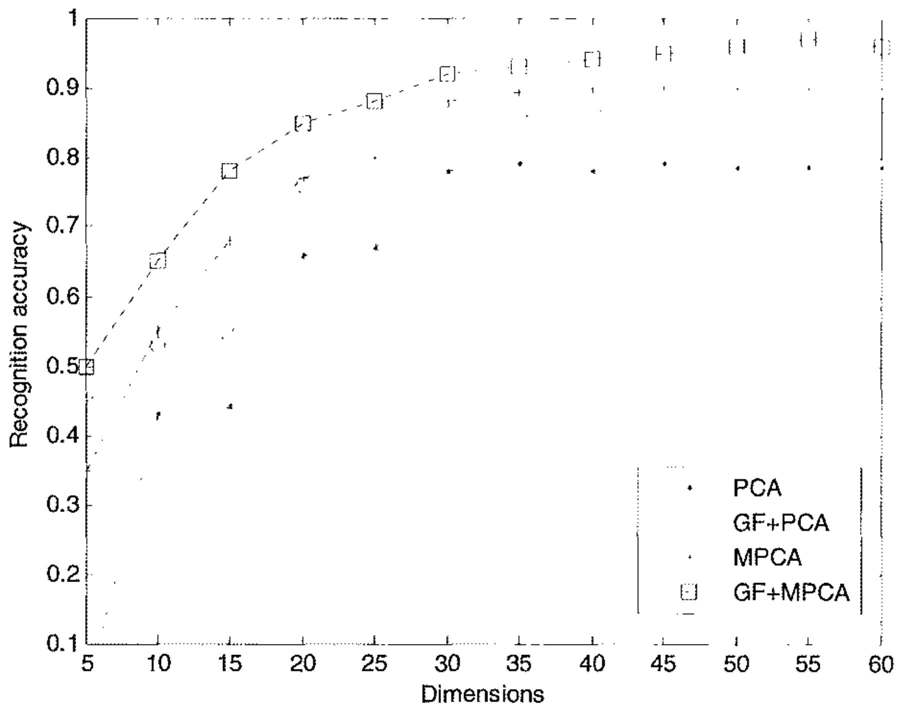

Figure 5.14. Comparison between Modular PCA on Gabor Features with other conventional techniques.

Next, different dimensionality reduction techniques are compared in order to determine the most efficient way to reduce the number of dimensions and also achieve maximum possible classification accuracy. Figure 5.15 shows this comparison, from which it is understood that non-linear analysis (kernel methods) perform best in complex situations like non-uniform lighting conditions and in the presence of occlusions because they extract the most discriminating features that achieve maximum separability. Therefore, KLDA is the best possible technique in terms of achieving maximum efficient separability under complex situations.

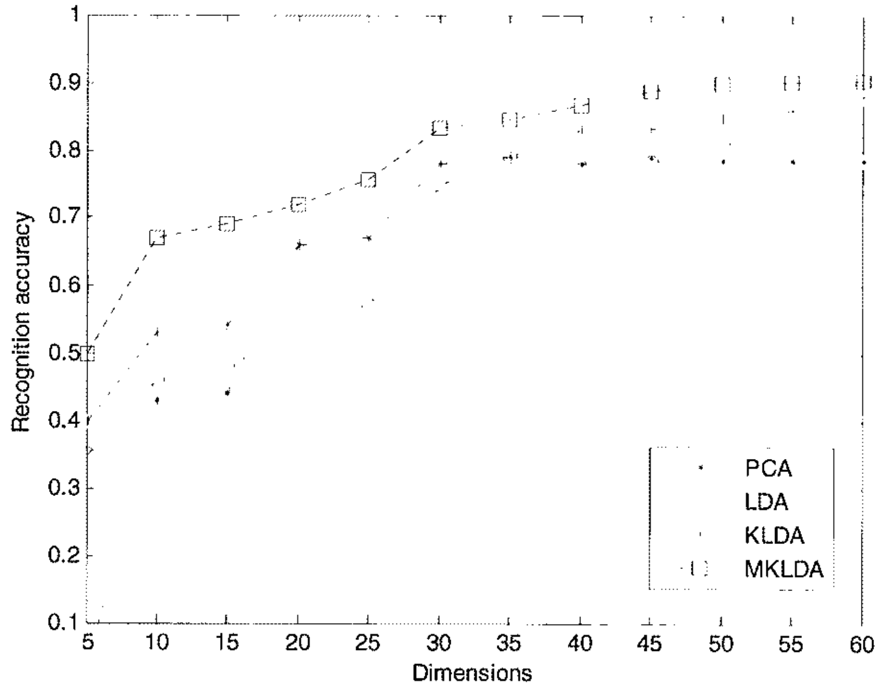

Figure 5.15. Comparison between Modular KLDA approach with - PCA, LDA and KLDA.

Experiments using different kernels, Dot product kernel, Polynomial kernel and Gaussian kernel were performed. The recognition accuracies for these different kernels are tabulated in table 5.2. The Polynomial kernel (d=2), where d is the degree of the polynomial, is used in this thesis.

Table 5.2. Recognition accuracies for different kernels

Kernel	**Recognition accuracy**
Dot product	88.6%
Polynomial (d=2)	90.1%
Gaussian (σ=3)	89.2%

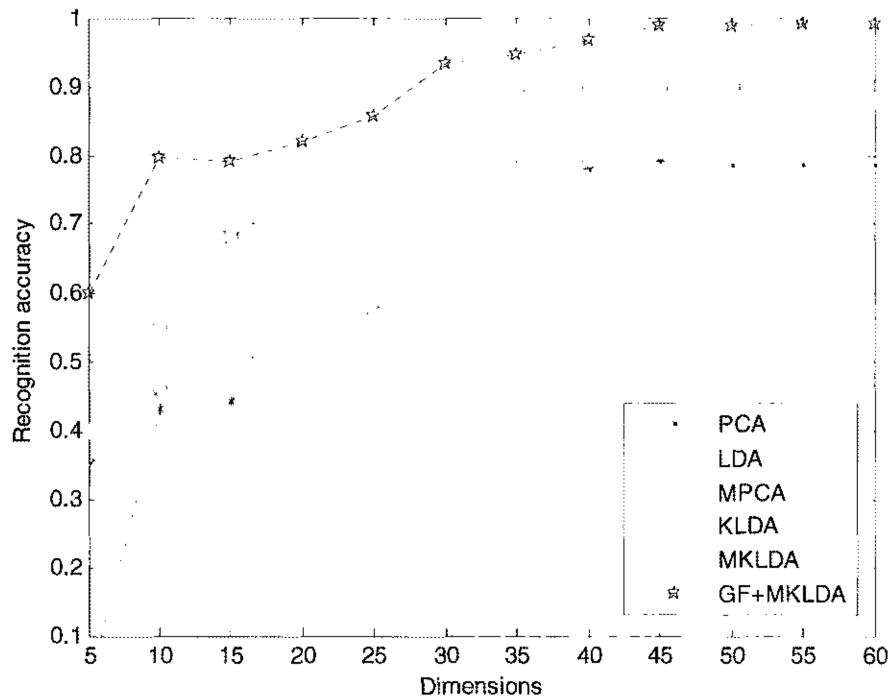

Figure 5.16. Comparison of the proposed Modular KLDA on Gabor Features with all other classical techniques.

Finally, the Gabor wavelet feature extraction technique combined with the MKLDA is compared with various previously discussed techniques, as shown in figure 5.16. It is seen that the proposed technique outperforms all of these techniques, achieving a recognition accuracy of 98.9%.

5.3. Receiver Operating Characteristics (ROC) Curves

Experiments were also conducted to evaluate the performance of the algorithm by comparing the recognition rate and false recognition rate of two methods – MKLDA on Gabor Features and KLDA on Gabor Features – on both AR and FERET databases. A module size of 16 x 16 pixels is used here, so a total of 16 modules are generated, i.e.

N_M=16. A threshold θ_{N_M}, the number of modules that would be assigned to each individual (subject), is varied from 2 to 16 to observe the effect of N_M on face recognition. 30 eigenvectors are used in these experiments. In the first iteration of the algorithm threshold θ_{N_M} = 16, and it is reduced by one for each iteration until a threshold of $\theta_{N_M} = 2$.

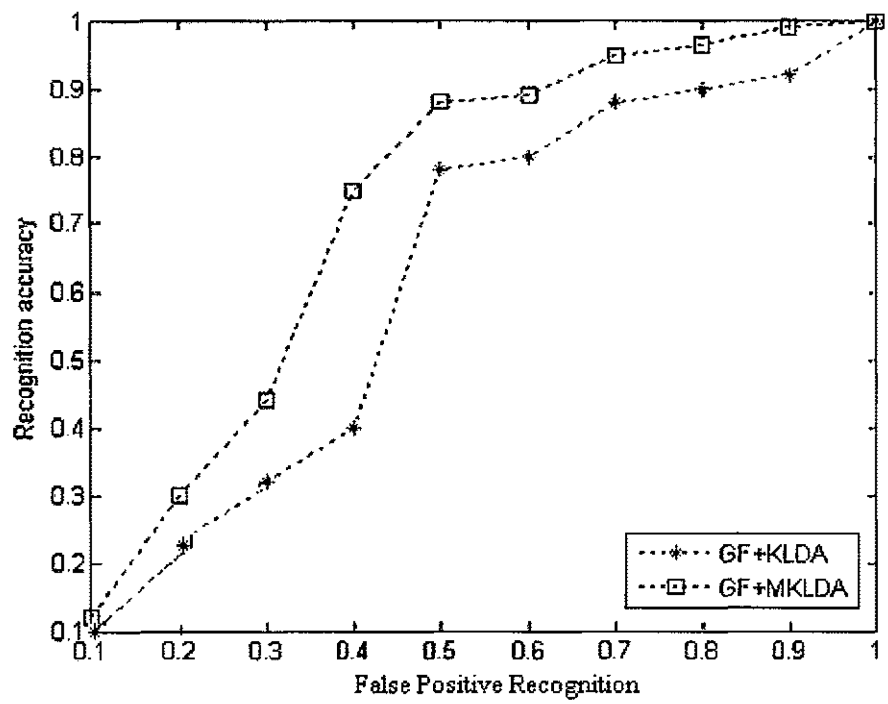

Figure 5.17. Comparison between ROC curves for MKLDA on Gabor Features and KLDA on Gabor Features for FERET database for AR database.

The ROC results for the AR database are shown in figure 5.17. From figure 5.17 it is observed that for a false positive rate of 0.1 (threshold θ_{N_M} = 16), the recognition rate for GF+KLDA is zero and 0.12 for GF+MKLDA. For a false positive recognition of 0.4

(threshold $\theta_{N_M} = 9$), the recognition rate is 0.38 for GF+KLDA and 0.74 for GF+MKLDA, i.e., a substantial increase of 36% in the recognition rate.

The ROC results for FERET subset database are shown in figure 5.18. For a false positive rate of 0.6 (threshold $\theta_{N_M} = 9$), the recognition rate is 0.88 for GF+MKLDA and 0.64 for GF+KLDA, i.e. an increase of 24% in recognition rate.

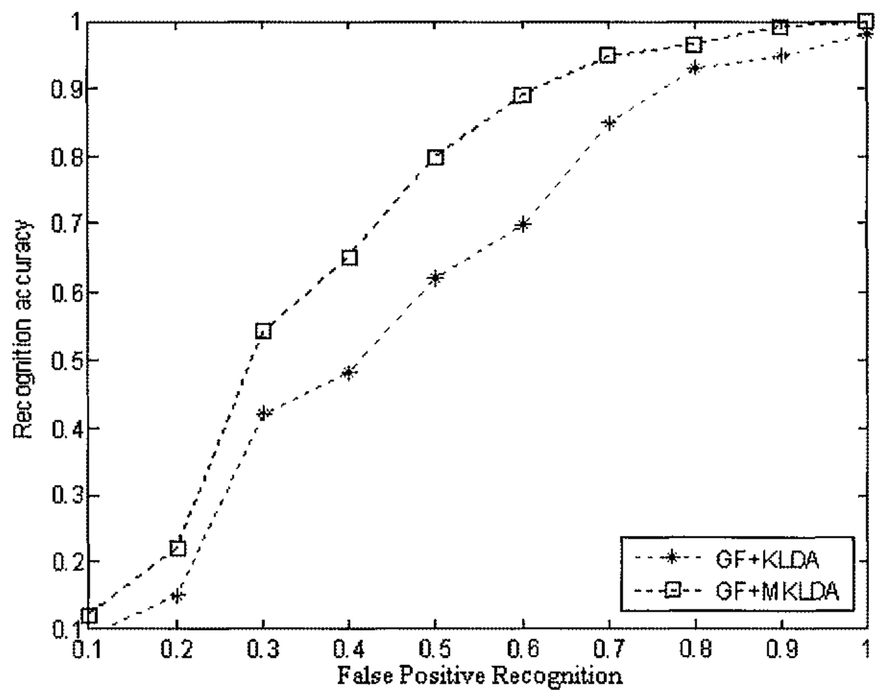

Figure 5.18. Comparison between ROC curves for MKLDA on Gabor Features and KLDA on Gabor Features for FERET subset database

It is seen that the recognition rate and false positive rate keep increasing with reduction in threshold θ_{N_M}, in the case of both the AR and FERET subset databases. Better recognition accuracy is obtained when the number of false positives is less for the proposed MKLDA on Gabor Features for both the AR and FERET subset databases.

Therefore, it can be said that the proposed technique performs better than KLDA on Gabor Features.

5.4. Summary

Different experiments were conducted to test the efficiency and robustness of the proposed technique on the AR and FERET databases. It has been observed that the Gabor wavelet feature extraction increases the recognition accuracy by making the illumination uniform throughout the image and also by extracting the more prominent features (obtained from different orientations). Also, the recognition accuracy is increased by reducing the effect of local variations as seen by the modular approach. Further, KLDA enhances the recognition accuracy by selecting the most discriminating features. Therefore, the MKLDA on Gabor Features extracts the most discriminating and local features of an image providing a substantial increase in recognition accuracy. Thus, the proposed method has proven to be robust in complex situations (occlusions, illumination, expressions, etc.).

CHAPTER 6

CONCLUSION AND FUTURE WORK

An efficient and robust Modular Kernel LDA (MKLDA) on Gabor Features has been proposed for face recognition in this thesis, and the method has been comprehensively tested using the AR and FERET face databases.

6.1. Summary

6.1.1. An Overview of Gabor Wavelets

A detailed review of the background and mathematical interpretation of Gabor wavelets is presented in Chapter 2 and Chapter 3. As a member of the wavelet family, mathematical analysis shows that the Gabor wavelet achieves optimal resolution in both time and frequency domains. Motivated by the mathematical background and biological evidence, 2D Gabor wavelets have been widely applied in different computer vision and pattern recognition applications including face recognition.

6.1.2. An overview of Modular Approach

Most of the appearance-based techniques (PCA, LDA, etc.) consider the global information of the image and thus are not very effective under conditions of varying illuminations, expressions, poses, etc. Under these conditions a modular approach is found to be more efficient. In the modular approach the face images are divided into smaller images, that represent the local information of a face. When variations occur in pose or illumination only some of the local face regions will vary. The rest of the regions will remain the same as the face regions of the normal image. Hence weights of the face regions not affected by varying pose, and illumination will closely match with the weights of the same individual's face regions under normal conditions. A detailed description and literature review of the modular approach has been presented in this thesis.

6.1.3. Gabor Wavelets and Modular Kernel Linear Discriminant Analysis

Although face recognition has been an active research area for many years, it is still a challenge due to the complex distortions caused by expression, pose and illumination variations. However, with the aid of complex perceptual systems such as the visual cortex it is very easy for a human to recognize hundreds of people, even in the presence of dynamic variations of face shape, pose, expression and appearance. Gabor wavelet's characteristics for frequency and orientation representations are quite similar to those of the human visual system. Hence Gabor wavelet applications are adopted in this thesis as a way to extract robust features for face recognition. Once the features are extracted, nonlinear kernel subspace analysis combined with the modular subdivision approach is applied for dimensionality reduction and class-separability enhancement. The combination of Gabor wavelets and modular kernel methods has been successfully applied to face recognition and tested using face databases, e.g. AR and FERET. The proposed method has achieved better performance than other state of the art recognition algorithms on the AR and FERET databases.

6.2. Future Work

A Gabor Feature is simply extracted by applying a wavelet at a certain image location. While Gabor wavelets with varied frequency and orientations are applied at different locations, the approach reflects the fact that different image regions display varied texture features. However, the candidate features in this work are extracted using a pre-defined set of 40 Gabor wavelets. While the most appropriate wavelet in the candidate set is chosen for a certain image location, the optimal wavelet for the position might not be included in the defined set. The search space of the wavelets can be extended to include all possible parameter spaces for better performance in the algorithm.

A non-overlapping module strategy was used in this work. Further, considering the relationship between neighboring modules might improve the performance of the algorithm since neighboring modules may contain critical information for recognition. The optimum solution would be to consider a combination of both the overlapping and non-overlapping module strategies to ensure that critical information is not lost.

6.2.1. Pose Invariant Face Recognition

The experiments conducted in this thesis mostly consider frontal views of the face. The proposed technique has been proven successful for frontal views under conditions of changes in illuminations, facial expressions and occlusions. Since face images (captured from the dynamic real world) have complex patterns, a pose estimation module could be used. This would aid in comparison of test images with images having similar pose. Thus, a face recognition system consisting of multiple face recognition modules could be developed that is robust to pose variance.

REFERENCES

[1] Cognitec systems, < http://www.cognitec.com/>

[2] Visage, Inc., <http://www.viisage.com/>

[3] A4 Vision, Inc., <http://www.a4vision.com>

[4] Gemometrix, Inc., <http://www.geometrix.com/>

[5] Genex Technologies, < http://www.genextech.com/>

[6] L. Capodiferro, R. Cusani, G. Jacovitti and M. Vascotto, "Correlation based technique for shift, scale, and rotation independent object classification", IEEE International Conference on Acoustics, Speech and Signal Processing, April 1987.

[7] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces", Journal of Optical Society of America, vol. 4, no. 3, pp. 519-524, March 1987.

[8] M. Turk and A. Pentland, "Eigenfaces for recognition", Journal on Cognitive Neuroscience, pp. 71-86, March 1991.

[9] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for object representation", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.

[10] P. Belhumer, J. Hespanha and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", Proceedings of the 4th European Conference on Computer Vision, vol. 1, pp. 45-5814-18, April 1996.

[11] A. Martez and A. Kak, "PCA versus LDA", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, February 2001.

[12] R. Fisher, "The use of multiple measurements in taxonomic problems", Annals of Eugenics, vol. 7, pp. 179-188, 1936.

[13] J. H. Yuen, P. C. Feng, G. Can and Lia, "Face recognition using holistic Fourier invariant features", Pattern Recognition, vol. 34, no. 1, pp. 95-109, 2001.

[14] M. Zhou and H. Wei, "Face Verification Using Gabor Wavelets and Adaboost", 18th International Conference on Pattern Recognition, vol. 1, pp. 404-407, 2006.

[15] L. Shen and L. Bai, "Gabor Feature Based Face Recognition Using Kernel Methods", Proceedings of the IEEE 6th International Conference on Automatic Face and Gesture Recognition, 2004.

[16] S. Chen and B. C. Lovell, "Illumination and expression invariant face recognition with one sample image", International Conference on Pattern Recognition, pp. 300-303, 2004.

[17] M. H. Yang, N. Ahuja and D. Kriegman, "Face recognition using Kernel Eigenfaces", Advances in NIPS, vol. 14, 2002.

[18] M. H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods", IEEE Conference on Automatic Face and Gesture Recognition, 2002.

[19] W. Zhao, R. Chellapa, A. Rosenfield and P. J. Phillips, "Face recognition: A literature survey", ACM Computing Surveys, vol. 35, no. 4, pp. 399–458, December 2003.

[20] S. Qian and D. Chen, "Joint Time-Frequency Analysis: Method and Applications", Prentice Hall PTR, 1996.

[21] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", IEEE Transactions on Information Theory, vol. 36, pp. 961-1005, 1990.

[22] D. Gabor, "Theory of communications", Journal of Institution of Electrical Engineers, vol. 93, pp. 429-457, 1946.

[23] G. H. Granlund, "Search of A General Picture Processing Operator", Computer Graphics and Image Processing, vol. 8, pp. 155-173, 1978.

[24] J. G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial- Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters", Journal of the Optical Society of America A: Optics, Image Science and Vision, vol. 2, pp. 1160- 1169, 1985.

[25] K. Okajima, "Two-dimensional Gabor-type receptive field as derived by mutual information maximization", Neural Networks, vol. 11, pp. 441-447, 1998.

[26] A. K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters", Pattern Recognition, vol. 24, pp. 1167-1186, 1991.

[27] Y. Hamamoto, S. Uchimura, M. Watanabe, T. Yasuda, Y. Mitani and S. Tomita, "A Gabor filter-based method for recognizing handwritten numerals", Pattern Recognition, vol. 31, pp. 395-400, 1998.

[28] C. J. Lee and S. D. Wang, "Fingerprint feature extraction using Gabor filters", Electronics Letters, vol. 35, pp. 288-290, 1999.

[29] C. J. Liu and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition", IEEE Transactions on Image Processing, vol. 11, pp. 467-476, 2002.

[30] S. G. Shan, W. Gao, Y. Z. Chang, B. Cao and P. Yang, "Review the strength of Gabor Features for face recognition from the angle of its robustness to misalignment", Proceedings of the 17th International Conference on Pattern Recognition, vol. 1, pp. 338-341, 2004.

[31] C. J. Liu and H. Wechsler, "Independent component analysis of Gabor feature's for face recognition", IEEE Transactions on Neural Networks, vol. 14, pp. 919-928, 2003.

[32] C. J. Liu, "Gabor-based kernel PCA with fractional power polynomial models for face recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp. 572-581, 2004.

[33] L. Shen and L. Bai, "Face recognition based on Gabor Features using kernel methods", In Proceedings of the 6th IEEE Conference on Face and Gesture Recognition, pp. 170-175, Korea, 2004.

[34] L. Shen and L. Bai, "Gabor wavelets and Kernel direct Discriminant analysis for face recognition", Proceedings of the 17th International Conference on Pattern Recognition, pp. 284-287, Cambridge, UK, 2004.

[35] R. Gottumukkal and K V Asari, "An Improved face recognition technique based on Modular PCA approach", Pattern Recognition Letters, vol. 25, pp. 429-436, 2004.

[36] A. Martinez, "Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 712–717, January 2000.

[37] Y. Zhang and A.M. Martinez, "Recognition of Expression Variant Faces Using Weighted Subspaces", Proceedings of International Conference on Pattern Recognition, 2004.

[38] K. Tan and S. Chen, "Adaptively weighted sub-pattern PCA for face recognition", Neurocomputing, vol. 64, pp. 505-511, 2005.

[39] A.U. Batur, "Segmented linear subspaces for illumination-robust face recognition", International Journal of Computer Vision, vol. 57, no. 1, pp 49-66, 2004.

[40] P. K. Manglik, U. Misra and M. Hima Bindu, "Facial expression recognition", IEEE International Conference on Systems, Man and Cybernetics, pp. 2220-2224, 2004.

[41] Y. Zhangt, E. C. Prakash and E. Sung, "Synthesis of Facial Expressions Using A 3D Anatomical Model", Proceedings of the 7th International Conference on Control, Automation, Robotics and Vision, pp. 704-709, 2002.

[42] P. Penev and J. Atick, "Local feature analysis, A general statistical theory for object representation", Network computation in Neural Systems, vol. 7, pp. 477-500, 1996.

[43] A. Pentland, B. Moghaddam and T. Starner, "View-based and modular eigenspaces for face recognition", Proceedings of the IEEE Computer Society Conference on Computer Vision an Pattern Recognition, pp. 84-91, 1994.

[44] B. Heisele, P. Ho, J. Wu and T. Poggio, "Face recognition: component-based versus global approaches", Computer Vision and Image Understanding, vol. 91, pp. 6–21, 2003.

[45] A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition", IEEE International Conference on Acoustics, Speech, and Signal Processing, vol.5, pp. 2721-2724, May 1998.

[46] F. Cardinaux, C. Sanderson and S. Marcel, "Comparison of MLP and GMM classifiers for face verification on XM2VTS", Proceedings of the 4th International Conference on Audio and Video Based Biometric Person Authentication, pp. 911-920, 2003.

[47] D. A. Pollen and S. F. Ronner, "Phase Relationship Between Adjacent Simple Cells in the Visual Cortex", Science, vol. 212, pp. 1,409- 1411, 1981.

[48] W. Zhao, A. Krishnaswamy, R. Chellapa, D. L. Swets and J. J. Weng, "Discriminant Analysis of Principal Components for Face Recognition", Face Recognition: From Theory to Applications, CAR-TR-914 edition, pp. 73-85, 1998.

[49] K. Fukunnaga, "Introduction to statistical pattern recognition", Academic Press, 2nd edition, 1991.

[50] D. L. Swets and J. J. Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, pp. 831-836, 1996.

[51] H. Gupta and A. K. Agrawal, "An experimental evaluation of linear and kernel-based methods for face recognition", Proceedings of the 6th IEEE Workshop on Applications of Computer Vision (WACV), pp. 13-18, 2002.

[52] B. Scholkopf, A. Smola and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem", Neural Computation, vol. 10, pp. 1299-1319, 1998.

[53] K. I. Kim, K. Jung and H. Kim, "Face recognition using kernel principal component analysis", IEEE Signal Processing Letters, vol. 9, pp. 40-42, 2002.

[54] H. Murase and S. Nayar, "Visual learning and recognition of 3-D objects from appearance", International Journal on Computer Vision, vol. 14, pp. 5-24, 1995.

[55] B. Moghaddam, A. Pentland and T. Starner, "View-based and modular eigen spaces for face recognition", IEEE Conference on Computer Vision and Pattern Recognition, June 21-23, 1994.

[56] AR database, <http://cobweb.ecn.purdue.edu/~aleix/aleix_face_DB.html/>

[57] FERET database, <http://www.itl.nist.gov/iad/humanid/feret/feret_master.html/>

[58] N. Gudur and K. V. Asari, "Gabor Wavelet based Modular PCA approach Expressions and Illumination Invariant Face Recognition", Proceeding on 35th IEEE workshop on Applied Imagery and Pattern Recognition (AIPR), 2006.

CURRICULUM VITAE

Neeharika Gudur

Education:

- Master of Electrical Engineering
  Old Dominion University (ODU), Norfolk
  Graduation: December 2007
  GPA: 3.61/4.0
- Bachelor of Electrical & Communications Engineering
  MGIT, Jawaharlal Nehru Technological University (JNTU), Hyderabad, India
  Graduation: May 2005
  GPA: 3.75/4.0

Work Experience:

- Research Assistant, Computer Vision and Machine Intelligence (CIMV) Lab, ODU, Norfolk, VA. [Spring '06 – Spring '07]

- Teaching Assistant, ODU, Norfolk, VA.
  1. Digital Signal Processing-I [Fall '06]
  2. Digital Image Processing [Spring '07]

Publications:

- "Gabor Wavelet based Modular PCA approach for Expression and Illumination Invariant Face Recognition", presented at AIPR, Oct 2006 (Conference Paper)
- "Long Range Face Recognition", presented at Old Dominion University on Research day, Apr 2007 (Conference/Poster)
- "Face Recognition in Multi-Sensor Images based on a Novel Modular Feature Selection Technique", submitted to Information Fusion Journal, Elsevier, 2007 (Journal Paper)

Address: 1069W, 49th Street, #B, Norfolk, VA-23508

E-mail: ngudu001@odu.edu