Spring 2008

A Weighted Modular Principal Component Analysis Approach for Face Authentication

Chandrika Tummala
Old Dominion University

A WEIGHTED MODULAR PRINCIPAL COMPONENT ANALYSIS APPROACH FOR FACE AUTHENTICATION

By

Chandrika Tummala

B.Tech (Electronics and Communication Engineering) April, 2005

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

May 2008

Approved by:

Dr. K. Vijayan Asari (Director)

Dr. Zia-ur Rahman (Member)

Dr. Jiang Li (Member)

ABSTRACT

A WEIGHTED MODULAR PRINCIPAL COMPONENT ANALYSIS APPROACH FOR FACE AUTHENTICATION

Chandrika Tummala
Old Dominion University, May 2008
Director: Dr. K. Vijayan Asari

A weighted modular approach for face authentication based on the priorities of different facial regions that change with varying poses, expressions and occlusions is presented in this thesis. This helps in verifying the identity of an individual who claims to be a subject in the database and is unaware of the presence of the face authentication system. A sequence of face images is selected from a video in a particular predefined interval and is used for verification. The face images are divided into different horizontal modules based on the regions representing facial features. A principal component analysis on these modules produces low dimensional representations of the sub images representing the facial feature regions. A weighted comparison of feature regions of the test images with the respective regions of the training images provides the authentication outcome. The performance of the proposed face authentication system is evaluated with several individuals belonging to different ethnicities. It is observed that the weighted modular approach outperforms the state-of-the-art face authentication techniques for input images with varying poses and expressions and occlusions. Research studies are progressing to compute the weights adaptively based on the magnitude of the feature variations due to poses and expressions and occlusions. In addition, application of the techniques based on multiple modalities for face authentication is also being investigated.

Dedicated to Mom, Dad, Sisters and dear Husband.

For all of your love and support.

ACKNOWLEDGMENTS

I would like to thank Dr. K. Vijayan Asari for his constant guidance and the opportunity to work with him in the Vision Lab. Without his encouragement and support, none of this would be possible.

I would also like to express my sincere gratitude to Dr. Zia-ur Rahman and Dr. Jiang Li, members of the thesis advisory committee, for their time and assistance.

In addition, I would like to thank my family and friends for their unconditional love and support through the entire time.

TABLE OF CONTENTS

ABSTRACT..... ii

ACKNOWLEDGEMENTS..... v

TABLE OF CONTENTS..... vi

LIST OF FIGURES..... viii

1. INTRODUCTION..... 1

2. LITERATURE SURVEY..... 7

2.1 Face Authentication Based on Eigen Flow..... 9

2.2 Face Authentication Based on Independent Component Analysis..... 11

2.3 Face Authentication Using Support Vector Machines (SVM)..... 12

2.4 Face Authentication Based on Scale Invariant Feature Transform (SIFT)..... 13

2.5 Face Authentication Based on One-Time Templates..... 15

2.6 Face Authentication Based on 3D Methods..... 16

2.7 Summary..... 17

3. FACE AUTHENTICATION USING PRINCIPAL COMPONENT ANALYSIS..... 19

3.1 Principal Component Analysis (PCA)..... 19

3.2 Face Authentication using Principal Component Analysis..... 25

3.3 Experimental Set up, Specifications and Results..... 28

3.3.1 Camera specifications..... 29

3.3.2 Image specifications.....29

3.3.3 System specifications and processing speeds..... 30

3.3.4 Experimental results..... 30

3.4 Summary..... 36

4. FACE AUTHENTICATION USING MODULAR PCA..... 38

4.1 Face Authentication using Modular PCA (MPCA)..... 38

4.2 Experimental Results..... 44

4.3 Summary..... 48

5. FACE AUTHENTICATION USING WEIGHTED MODULAR PCA..... 50

5.1 Face Authentication using Weighted Modular PCA..... 50

5.2 Experimental Results..... 57

5.3 Summary..... 60

6. CONCLUSIONS AND FUTURE WORK..... 61

REFERENCES..... 64

LIST OF FIGURES

Figure 1.1: Illustration of face authentication system..... 4

Figure 2.1: Illustration of the concept of optical flow..... 10

Figure 3.1: Matrix representation of face image of size N×N..... 20

Figure 3.2: Representation of the image space..... 20

Figure 3.3: Illustration of the training phase for PCA based face recognition..... 23

Figure 3.4: Illustration of the testing phase for PCA based face recognition..... 24

Figure 3.5: Representation of the image space for an individual..... 25

Figure 3.6: Sample training images of 20 individuals..... 31

Figure 3.7: Sample testing images of two subjects with PINs 9 and 5..... 32

Figure 3.8: Error characteristics with respect to number of eigenvectors using universal PCA (for images that are similar to training set)..... 34

Figure 3.9: Sample test images of two subjects with occlusions 34

Figure 3.10: Error characteristics with respect to number of eigenvectors using universal PCA (for images with occlusions)..... 35

Figure 3.11: ROC curve of universal PCA..... 36

Figure 4.1: Representation of splitting the face image in modular PCA..... 40

Figure 4.2: Illustration of the training phase for modular PCA based face authentication..... 41

Figure 4.3: Illustration of the testing phase for modular PCA based face authentication..... 43

Figure 4.4: Sample testing images with varying expressions, poses and occlusions used for modular PCA 45

Figure 4.5: Error characteristics with respect to number of eigenvectors using modular PCA (for images with expressions, poses and occlusions)..... 46

Figure 4.6: Error characteristics with respect to number of modules using modular PCA (for images with expressions, poses and occlusions)..... 47

Figure 4.7: ROC curves of universal PCA and modular PCA..... 48

Figure 5.1: Representation of splitting the face image in weighted modular PCA..... 52

Figure 5.2: Illustration of the training phase for weighted modular PCA based face authentication..... 54

Figure 5.3: Illustration of the testing phase for weighted modular PCA based face authentication..... 56

Figure 5.4: Error characteristics with respect to number of eigenvectors using weighted PCA (for images with varying expressions, poses and occlusions)..... 58

Figure 5.5: ROC curves of universal PCA, modular PCA and weighted modular PCA..... 59

Chapter 1

INTRODUCTION

In a real life scenario, given a series of still or video images, authenticating an individual using a stored database of images is the main goal of face authentication. Out of many biometric authentication methods, face authentication is one of the best and most flexible methods in which the individual could be unaware of being authenticated. Face authentication is a field that is related to face recognition. The main difference between the two is that a face authentication system has to verify the claimed individual using the personal identification number that is assigned to him during the enrollment process. In face recognition, the system has to identify the individual from the database. In simple terms, face recognition is a one-to-many match, whereas face authentication is a one-to-one match; that is, face authentication is a subset of face recognition. Face authentication has many applications, which include secure entry, money transaction authentication and various forms of verification in order to prevent identity theft. The test images grabbed from the video may contain variations in expressions, poses and occlusions such as sun glasses, mask, skull cap, makeup, etc. Dealing with these kinds of variations in face images is the main task of this research. Therefore, different algorithms that are based on Principal Component Analysis (PCA) are investigated. Various functional steps in face authentication include detection of the face region, extraction of face features and classification.

The journal model for this work is the *Physical Review*.

There are many methods that can be used for face authentication, such as using subspace methods like PCA, independent component analysis and linear discriminant analysis. These are statistical, structural techniques that use the width of the head, the distance between the eyes, the distance between the eyes to the tip of the nose, [1] or the angles between the eye edges, mouth corners, etc. [2] Hidden Markov Models (HMM) and Gaussian Mixture Models (GMM) [3] based methods use a band of pixels that cover the forehead, eyes, nose, mouth and chin without finding the exact locations of the face. There are many other methods, such as Bayesian methods that use a probabilistic distance metric [4] and the classifier techniques, such as support vector machines and minimum distance pair that are used for face authentication. Frontal face authentication under controlled conditions would give near to 100 percent accuracy for large databases, but problems arise when the face authentication is under uncontrolled conditions such as expressions, poses and occlusions.

The objective of this research is to develop a face authentication system that could be useful for applications related to access control in a distributed environment such as in ATMs, banks or secured buildings. In this case, each subject is given a personal identification number (PIN) during the enrollment. When the subject types his or her respective PIN, video is recorded and the frames are grabbed from this video and are stored for further processing. The facial authentication process extracts the features from these test images, which are compared to the stored features in the database to verify whether the PIN belongs to the claimed identity.

Face authentication system can be explained with an example in brief. When an individual wants to perform a transaction at the ATM, he or she should have a debit card

and a PIN (which is assigned to the individual when he or she opens an account in the bank). While the individual swipes the card and enters the PIN, the video of the individual's face is recorded and is matched by the face authentication process with the images in the database (which are collected during the enrollment process). If the PIN does not match the account number, the individual is asked to enter the PIN again. This happens three times, and if he does not enter the correct PIN in three trials, he is not allowed to perform the transaction. When the individual enters the correct PIN, 10 test images of the individual from the video are grabbed and compared to the information in the database. If at least two matches are correct among the 10 test images, then the individual is allowed to perform a transaction or else the transaction is not allowed to take place. That is, a transaction can be performed only if the account belongs to the individual who is present at the ATM machine. The block diagram illustrating the process of the proposed face authentication system is shown in figure 1.1. The proposed face authentication system can also be used in other access control areas.

In this thesis, the solution to the problems due to variations in expressions, poses and occlusions are addressed, and inspired by the methods that are proposed for face recognition in [5] and [6], a robust algorithm to authenticate an individual is presented. The proposed technique consists of many phases, starting with the face template extraction using face detection based on the Viola and Jones algorithm, feature extraction and then authentication. The occlusions, expressions and poses are handled by the weighted modular PCA approach. The algorithm is computationally simple, and by setting the thresholds and by varying the number of eigenvectors, the authentication system provides better results.

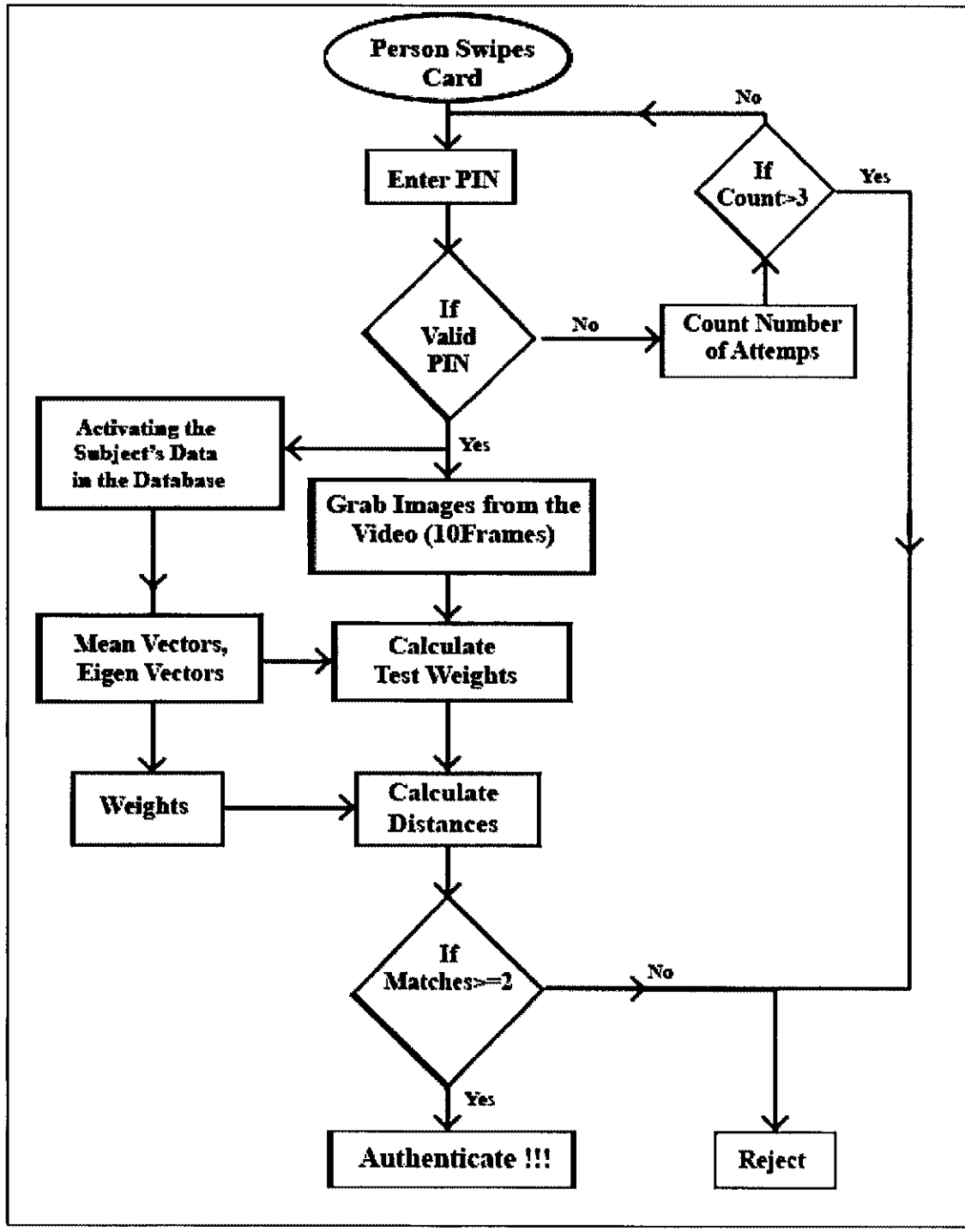

Figure 1.1: Illustration of face authentication system

The main objective of this thesis research is to develop a face authentication system satisfying the following requirements:

- The individuals are not asked to pose in front of the camera; therefore the individuals are to be authenticated even with varying poses, expressions and occlusions.
- The training of the system is done off line.
- The testing is done on the videos that are recorded while an individual enters the PIN in order to perform a transaction or entry.
- The overall system should be simple, fast, accurate and economical.

The specific objectives in the proposed research work are as follows:

- To detect the face region in a video frame using the Viola-Jones face detector,
- To extract face features by weighted modular PCA approach,
- To compare the feature vector with the information in the database using the Mahalanobis distance measure,
- To repeat the process for 10 images captured at a particular rate from the video sequence,
- To generate authentication results based on a predefined number of matches with the information in the database.

The thesis is organized as follows:

A detailed survey of the state of art methods that are used for face authentication and the pros and cons involved in these methods are presented in chapter 2. This also discusses the various problems that are involved in face authentication systems and explains the limitations that are involved in each of the techniques. It can be seen that not all problems are overcome in many of the techniques developed so far.

Chapter 3 presents the basic concept of PCA and how it is used in the real time face authentication systems. The experimental set up, specifications, data acquisition and the performance evaluation of face authentication using PCA is described in this chapter. It also clearly explains why the face authentication using PCA fails and why there is a need for a better technique.

Chapter 4 introduces the modular PCA based technique for face authentication. The performance evaluation of this face authentication system is presented. It also explains how this method outperforms the universal PCA in terms of expressions, poses and occlusions.

Chapter 5 presents the algorithm of weighted modular PCA approach for face authentication and explains the performance evaluations using this technique. This chapter also conveys how this technique outperforms the universal PCA and the modular PCA in terms of varying expressions, poses and occlusions.

Chapter 6 includes the conclusions and directions for the future work.

Chapter 2

LITERATURE SURVEY

Face authentication is a flexible biometric authentication method in which the individual is unaware of being authenticated and performs best in features such as accuracy, cost, and ease of sensing. [7] Face authentication has received significant attention during the past few years, as it is one of the most successful applications of image analysis and understanding. Many efforts have been made and many methods have been proposed within the past few years on video based face authentication. The literature on face authentication is vast and diverse because it is a challenging and interesting field that attracts many researchers of different fields such as psychology, pattern recognition, neural networks and computer vision.

Researchers have proposed many methods for face authentication that are hybrid and it is difficult to categorize these systems based purely on what types of methods they used for representing the training set and for classification purposes. Talking in terms of the general categorization based on the methods used so far, the methods that use whole face region as raw input to the face authentication system are categorized as holistic matching methods. Another widely used method represents face images as Eigenfaces, [8], [9], [10] a concept based on principal component analysis. In feature based matching methods, the local features such as eyes, nose and mouth are extracted first, and the locations of these features and the geometrical statistics are fed into a feature based classifier. The usage of both the local features and the face as the global region, just as the human perception system uses to authenticate an individual, are classified as the

hybrid methods. Within the mentioned categories there can be many other classifications as there are many methods explaining about how to authenticate a person.

A different type of categorization based on three subspace techniques, Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discriminant Analysis (LDA), are all termed as appearance-based statistical methods. All three techniques, PCA, LDA and ICA, belong to the same family of methods called the subspace methods. [11] There are many face authentication systems based on appearance based statistical methods that use purely one technique or a combination of these techniques explained later in this chapter. Researchers [12] reported that performance for combined PCA-LDA method is improved over the pure LDA or pure PCA. There is a controversy between two groups of researchers; one says that LDA performs worse than traditional PCA [13] and the other group says that PCA is better compared to LDA for small training sets, but that does not hold true for large data sets [14]. There is another controversy between two research groups on the performance comparison between ICA and PCA. [15] One report support ICA, [16] and the other supports PCA. [17] All the three subspace techniques have their own advantages and disadvantages. But many researchers say that PCA and some of its modifications perform far better than both ICA [7] and LDA. [18]

As the above survey gives an overview of the statistical techniques, now we look into the structural matching techniques. These matching techniques use the width of the head, the distance between the eyes or the distance between the eyes to the tip of the nose, [1] or the angles between the eye edges, mouth corners, etc. [2] Hidden Markov Models (HMM) [3] based methods use a band of pixels that cover the forehead, eyes,

nose, mouth and chin without finding the exact locations of the face. There are many other methods, such as Bayesian methods that use a probabilistic distance metric [4] and methods that make use of support vector machine (SVM) as the classifier. [19]

2.1 Face Authentication Based on Eigen Flow

In the Eigen flow method, [20] first the optical flow residue is computed between one of the images in the training set and the test image of that particular individual. Second, the principal component analysis is performed on the optical flow images to get the eigen flow. Finally, using linear dicriminant analysis, the eigen flow residue is combined with the optical flow residue, which determines the authenticity of an individual.

Optical flow [21] is used to capture the motion between the face appearances when there is variation in facial expression or pose. Optical flow is generated using four steps given any two training images. 1) The background is removed and zeros are added to the removed background region, 2) The optical flow is determined using the Lucas-Kanade algorithm, [22] 3) In order to speed up the PCA training process, the optical flow is down sampled and 4) Finally, the background and four side boundaries are removed within the small sized optical flow image because the boundaries do not result in accurate motion estimation.

Optical flow determines the velocity fields in x and y directions. It is generally used for motion analysis to estimate the displacement of pixels of one image to the other. The concept of optical flow method is illustrated in figure 2.1. Figures 2.1 (a)

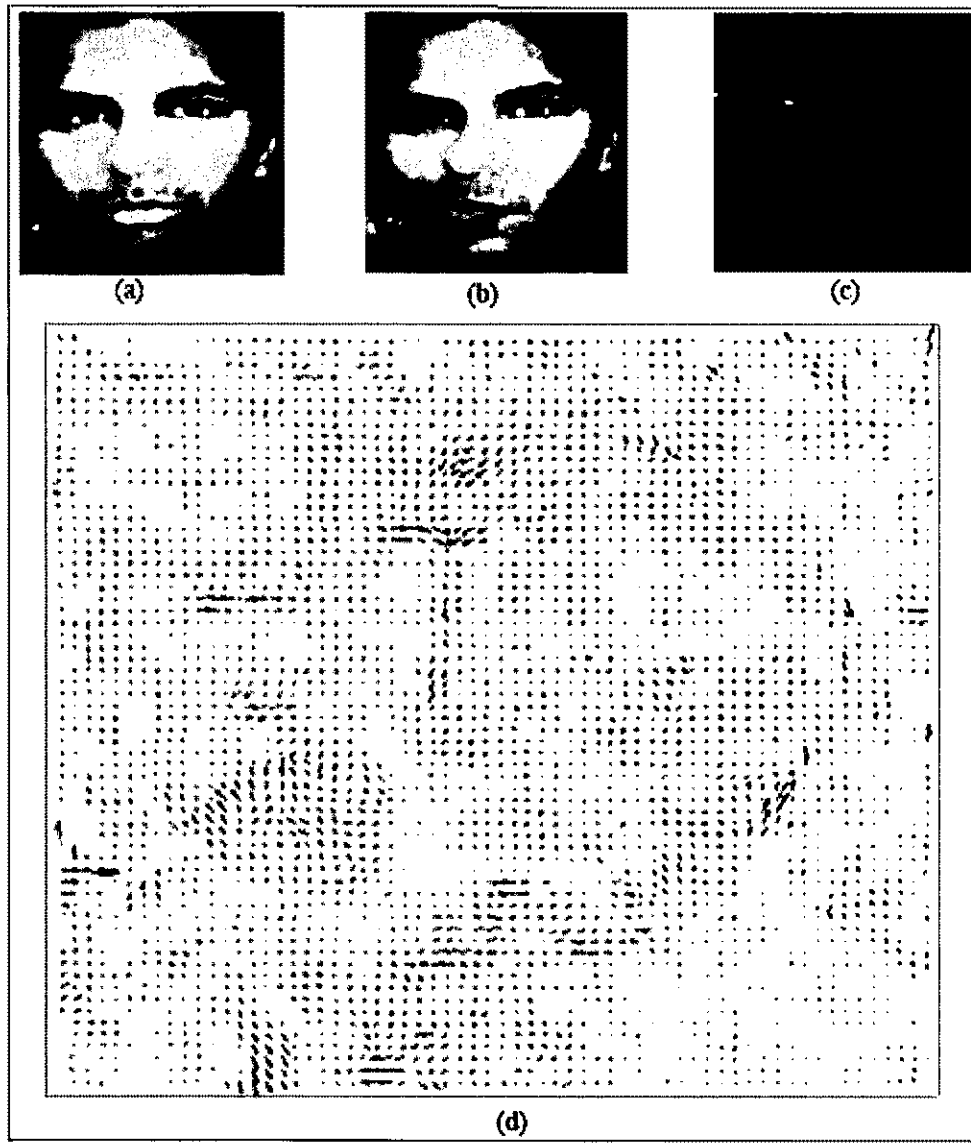

Figure 2.1: Illustration of the concept of optical flow. (a) image 1, (b) image 2, (c) optical flow residue, (d) optical flow between image 1 and image 2

and (b) show two images, image 1 and image 2 of a person with different expressions. The optical flow is determined using the Lucas-Kanade algorithm [22] and the optical flow image is shown in figure 2.1 (d) below the original images. Now using image 1 and the derived optical flow image (d), the second image is to be reconstructed, which is

called the predicted image. The difference between the second image and the predicted image is the optical flow residue image, which is shown in figure 2.1 (c).

With the training set of four images of the same subject with different expressions, we can obtain 12 optical flow images. When different subjects are considered to compute the optical flows, PCA is applied to these optical flow images, which results in a larger residue called the eigen flow residue. The above mentioned residues have the ability to discriminate between the self class and the imposter class; by using the LDA we can combine the two residues. Now LDA [23] is used to find the optimal 1-dimensional feature space that separates the two classes.

The self class is obtained using the projections of all the optical flow images from subject 1, the imposter class projections are obtained from the projections of the optical flow between subject 1 and other remaining subjects in the training set. Now the mean and the covariance matrices of these two classes are determined and then the within-class and between-class scatter matrices are computed. The weight vector that maximizes the ratio between the determinants of the between class scatter matrix to the within class scatter matrix will authenticate the system.

2.2 Face Authentication Based on Independent Component Analysis

The alternative feature extraction algorithm other than principal component analysis is independent component analysis, [24] which is another widely used face recognition/ face authentication technique. As in many other image processing techniques, first the relevant features are extracted. Recovering the original image from the known images is the main aim of source separation where each known image is a

mixture of original images. There are two different architectures for independent component analysis, namely architecture I and architecture II.

In architecture I, [7] the data matrix D is formed in such a way that each row vector represents an image. Images are considered as random variables and the pixels as the trials. Moving along the pixels, two images *a* and *b* are independent because it is difficult to predict the value of the pixel of image b based on the pixel of image a; this is a good example of the independence of images. In architecture II, the data matrix used is a transpose of that used in architecture I. So moving along the images, two pixels *a* and *b* will be independent. In this approach ICA finds a matrix such that the rows of this matrix are statistically independent in order to reconstruct the original image in a least square manner. As the vectors are independent they will be much closer to the natural features of the training set and the ability to represent differences between the face images is high. But researchers report that ICA will not replace PCA as it does not perform well with large datasets. So first PCA is applied to reduce the dimensionality of the data and then ICA is performed. ICA algorithms are a recursive process and sometimes do not converge. [25]

2.3 Face Authentication Using Support Vector Machines (SVM)

A support vector machine [26], [44] is the main decision making tool for many of the image processing techniques that are used for face authentication. It is mainly a classifier technique that is considered to outperform Euclidean distance, normalized correlation, etc. SVMs use many parameters and different kernels, which make the optimization space extensive and do not guarantee to make the best solution. As it is a

classification technique, in training process the faces can be represented as either the eigen-faces (using principal component analysis) or as fisher-faces (using linear discriminant analysis).

Now let us proceed further with the training process using the concept of fisher-faces for feature extraction. SVM is trained to classify between the within-class and between-class images. [27] Researchers say that SVM classifier is one of the best techniques in extracting the features from the training data set. [28] Explaining in brief the theory behind SVMs, they are based on minimizing the expectation or actual test error for the trained machine, which is termed as the principal or structural risk. Considering the linearly separable case, SVM provides the optimal hyper-plane that maximizes the margin of the distances to the closest positive and negative training patterns. Coming to the non-linear case, using the kernel functions (polynomials, exponential and sigmoid functions) the training patterns are mapped onto a high-dimensional space where the decision boundary is linear in this high-dimensional space. [29] The disadvantage of SVMs is that when the data is modified by the feature extraction or normalization, they get highly trained, and that affects generalization.

2.4 Face Authentication Based on Scale Invariant Feature Transform (SIFT)

SIFT technique is somewhat similar to local binary pattern methods [30] [31] but is for producing a view invariant representation of feature extracted 2D patterns. SIFT features presented by David Lowe [32] are invariant to the scaling and rotation of an object and have promising matching for objects with distortions, noise varying

illumination, etc. This approach identifies images even with occlusions and is capable of achieving good results in a real world scenario.

The main steps involved for computing and generating the features for an image are: (a) scale-space extrema detection: It is the first stage and is implemented by using a Gaussian function to derive the key points that are invariant to scale and rotation; (b) key point localization: At each pixel location, a detailed model is fit to compute the location and scale, key points are selected based on the stability; (c) orientation assignment: Based on local image gradient directions, each key point is assigned by one or more orientations and all the operations are on image data that have been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations; and finally, (d) key point descriptor: The local image gradient that are measured at selected scale in the region around each key point are transformed in such a way that it allows for significant levels of local shape distortion and variation in illumination. Therefore each SIFT feature is composed of four parts, namely the location, the scale, the orientation and the descriptor. Using these features there are many classifiers [43] such as (i) minimum pair distance: computes the minimum distance between all pairs of all key point descriptors in two images; (ii) matching the eyes and the mouth, as most of the information of the features lie in these two regions; (iii) matching on a regular grid: image is divided into sub regions or modules using regular grid with overlapping and then matching is done.

2.5 Face Authentication Based on One-Time Templates

This is an interesting technique that is based on applying a transform on the original data that cannot be retrieved later. This technique uses two types of transformations that are similar; one transform is applied to match with Euclidean distance, and the other transform is used to match with cosine functions. Let us use the terminology, a for data and b for the test image and x and y as their corresponding transformations.

Transformation using Euclidean distance: Using an orthogonal matrix M and a vector n that is generated randomly and independently, a transformed template is created. This orthogonal matrix has the property such that $M^T M = MM^T = I$. Using these randomly generated matrices the transformations are performed; at the registration stage the transformation for the data set x is created by $x = Ma + n$, the values of x are stored discarding the values of a, as the values of M and n are randomly generated and it is impossible to retrieve the original data. These values of M and n are stored in the user's smart cards, and for security purposes these values should be different for different users. At the testing stage for face authentication, the test image b is also transformed in a similar manner $y = Mb + n$ with values of M and n that are provided by the user of the smart card. Now the matching is performed using the Euclidean distance between the transformed values of x and y, provided the user gives the correct M and n values.

Transformation using cosine function: The transformation is done in a similar manner; the data is transformed using the equation $x = Ma$, the test image is transformed using $y = Mb$, and the matching is performed using the cosine distance between the x and y values using equation (2-1),

$$\cos(x, y) = \frac{x.y}{\|x\|\|y\|} \tag{2-1}$$

where, $x.y = \sum_{i=1}^{N} x_i y_i$

And coming to the one-time template [33], a new template is created every time a person comes in for authentication by changing the values of M and n. At i^{th} authentication, a person enters his i^{th} M and n values and the i^{th} y values are created. Now at $(i+1)^{th}$ authentication, the new data, which is assigned during the i^{th} authentication, is used to transform the data test image, and the values M, n and x are always updated and stored as soon as the person is authenticated. After every authentication that is processed for an individual, he or she gets new M and n values, which are kept on his smart card. These new values are calculated using the equations

$$M_{i+1} = M_i^T M_i \tag{2-2}$$

$$n_{i+1} = M_i^T n_i + n_i^T . \tag{2-3}$$

The approach used here is a simplification process of semi-randomized access control. [34] This method of transformation is applicable not only to face template matching systems but also to other biometric template matching systems.

2.6 Face Authentication Based on 3D Methods

The above mentioned techniques are some of the 2D face authentication techniques. We now briefly explain how 3D face authentication is carried out. The 3D face authentication systems are divided into 3 parts: data acquisition, feature extraction and finally face authentication. The most common way to represent the 3D data through

scanning is the triangle mesh representation. After the data acquisition is done, the features are extracted; this is done during the training stage and even during the testing stage. There are three steps involved in the feature extraction process: surface point classification, approximating nose tip extraction and finding symmetry plane, and finally critical points, nose and sub face extraction. To compute surface curvature at each vertex of the triangle mesh of the face, the bi-quadratic Bezier patch method [35] is used.

First the specific features such as the nose tip are computed. Spin image method [36] is used as the initial guess to find the nose tip. Second, the face symmetry plane is determined using the modified mirror plane method [37] and then, using the intersection of symmetry plane with the face mesh, the symmetry profile curve is determined. Finally the critical points and the face mask are extracted. The critical points are the eye corner points, mid cheek points, bottom nose point, eye corner points, etc. Face mask is the sub region of the face surface. When the features are extracted, only the face feature information is stored and is matched with the feature information at the client location. Two types of comparisons are used; first is comparison using profile curves [37] and second, using the nose and face masks. The nose and face mask comparison is done using the Iterative Closest Point (ICP) algorithm. [38]

2.7 Summary

In this chapter, various methods and approaches used for face authentication and the pros and cons of each technique has been discussed. This chapter focused mainly on the various problems that are involved in face authentication systems and also discussed the limitations of the technique used so far. It also conveyed that not all problems are

overcome in many of the techniques developed so far, and the necessity of a novel approach to solve the issues faced by varying poses, expressions and occlusions in face regions.

Chapter 3

FACE AUTHENTICATION USING PRINCIPAL COMPONENT ANALYSIS

Principal component analysis is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables. [39] The main idea is to reduce the dimensionality of the original data set that is to express the large single dimensional vector of pixels constructed from a two dimensional facial image into the compact principal components of the feature space, which is called the eigenspace projection. The eigenspace is calculated by finding out the eigenvectors of the covariance matrix derived from a set of facial images. This eigenspace is called the universal eigenspace as it represents the variations in expression and pose among all the subjects.

This chapter deals with the theory of the universal principal component analysis and how it is used to implement the face authentication, and the performance of the face authentication using universal PCA technique with a different number of eigenvectors is assessed.

3.1 Principal Component Analysis (PCA)

For face recognition using PCA, we assume that there are P subjects, and each subject has K training face images. Each face image $I(i, j)$ is a two dimensional matrix of size N by N, as shown in figure 3.1, which is transformed into a single dimensional vector of size $N^2 \times 1$. The image space is created by all the $N^2 \times 1$ vectors of P subjects with K

images of each subject as shown in figure 3.2. The main idea of PCA for face recognition is to express a large one dimensional vector of pixels constructed from a two dimensional facial image into a compact set of principal components of the feature space, which is called the eigenspace projection.

$$I(i, j) = \begin{bmatrix} I_{1,1} & I_{1,2} & I_{1,3} & \cdots & I_{1,N} \\ I_{2,1} & I_{2,2} & I_{2,3} & \cdots & I_{2,N} \\ \vdots & \vdots & \vdots & & \vdots \\ I_{N,1} & I_{N,2} & I_{N,3} & \cdots & I_{N,N} \end{bmatrix}_{N \times N}$$

Figure 3.1: Matrix representation of face image of size $N \times N$

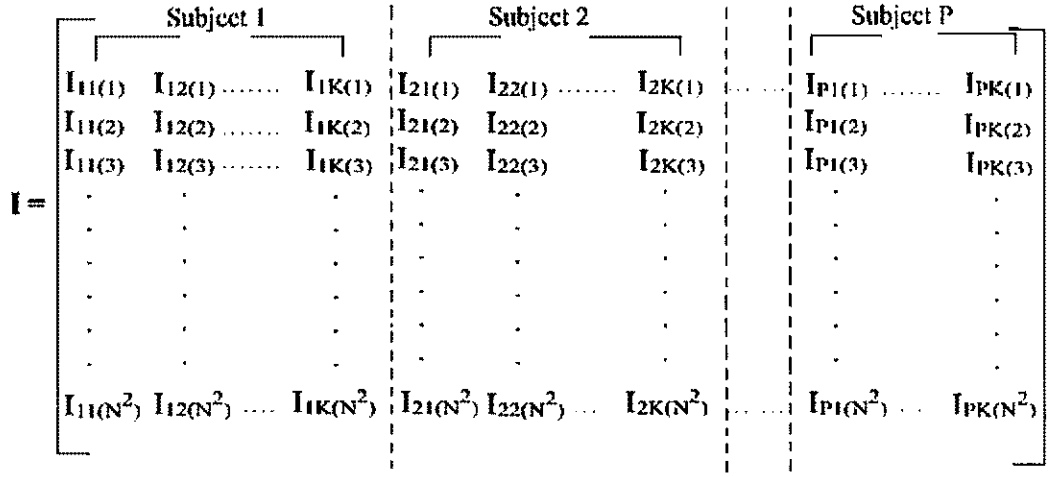

Figure 3.2: Representation of the image space

Each column vector in the image space, which is an image in the training set, is denoted as I_{lm}, which is a $N^2 \times 1$ column vector, where l denotes the index of the subject

and m denotes the index of the face image of an individual, where $1 \le l \le P$ and $1 \le m \le K$. The mean face is calculated from the image space which is defined by,

$$M = \frac{1}{PK} \sum_{l=1}^{P} \sum_{m=1}^{K} I_{lm} \tag{3-1}$$

Each training face differs from the mean by the vector $D_{lm} = I_{lm} - M$. Therefore the difference matrix A is of size $N^2 \times PK$ can be represented as $A=[D_1, D_2, D_3, ..., D_{PK}]$. PCA is applied to the difference matrix by finding a set of R orthonormal eigenvectors V_r, corresponding to the largest eigenvalues of the matrix AA^T, that is

$$AA^T V_r = \lambda_r V_r \qquad \text{for } r = 1,2,.....,R \tag{3-2}$$

where $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_R$ are in the decreasing order of non-negative eigenvalues and R is the number of significant eigenvectors that are chosen with the largest corresponding eigenvalues. The matrix AA^T is of size $N^2 \times N^2$ and determining the N^2 eigenvectors is computationally complex, but the number of training face images, PK is much smaller than N^2, therefore first the eigenvectors V_r' are determined by $A^T A$ which is of size $PK \times PK$, that is

$$A^T A \; V_r' = \lambda_r V_r' \tag{3-3}$$

By solving the equations (3-2) and (3-3), we get $V_r = A V_r' \lambda_r^{-\frac{1}{2}}$. The set of orthonormal basis of a new feature space is formed by these eigenvectors, which is termed as the eigenspaces. [40] Therefore the face images are represented by a set of eigenvectors developed from a covariance matrix formed by the training face images. The idea behind eigenspace is to find a lower dimensional space in which these shorter vectors will describe face images, and essentially it is the subspace representation of all the face

images. Thus the transformation of each face image I_{lm} from the image space to the eigenspace is given by,

$$\omega_{lm_r} = V_r^T (I_{lm} - M) \quad \forall l, m, r \tag{3-4}$$

Using these weights, a comparison is made during the testing phase and the distance is found out in order to recognize an individual. The training phase and the testing phase of face recognition using PCA is illustrated in figure 3.3.

In the testing phase, classification is based on a test face image I_{test} where the projection into the eigenspaces is obtained by,

$$\omega_{test_r} = V_r^T (I_{test} - M) \quad \forall r \tag{3-5}$$

The weights of the training set ω_{lm_r} form a vector $T_{lm} = [\omega_{lm1}, \omega_{lm2}, \omega_{lm3}, \ldots\ldots \omega_{lmR}]^T$, which describes the projection of each input face image in the eigenspace. The weights of the test image ω_{test_r} form a vector $T_{test} = [\omega_{test1}, \omega_{test2}, \ldots\ldots \omega_{testR}]^T$. This vector is used to fit the test face image in the predefined face class. To classify the test image, a simple Euclidean distance measurement technique is used where the distance between T_{test} and T_{lm} is measured using the equation,

$$Dist_{lm} = \| T_{test} - T_{lm} \| \quad \forall l, m \tag{3-6}$$

Finally, when comparing $Dist_{lm}$ with a pre-selected threshold τ, the test image can be classified to be recognized as the l^{th} subject if minimum of $Dist_{lm}$ is less than or equal to τ.

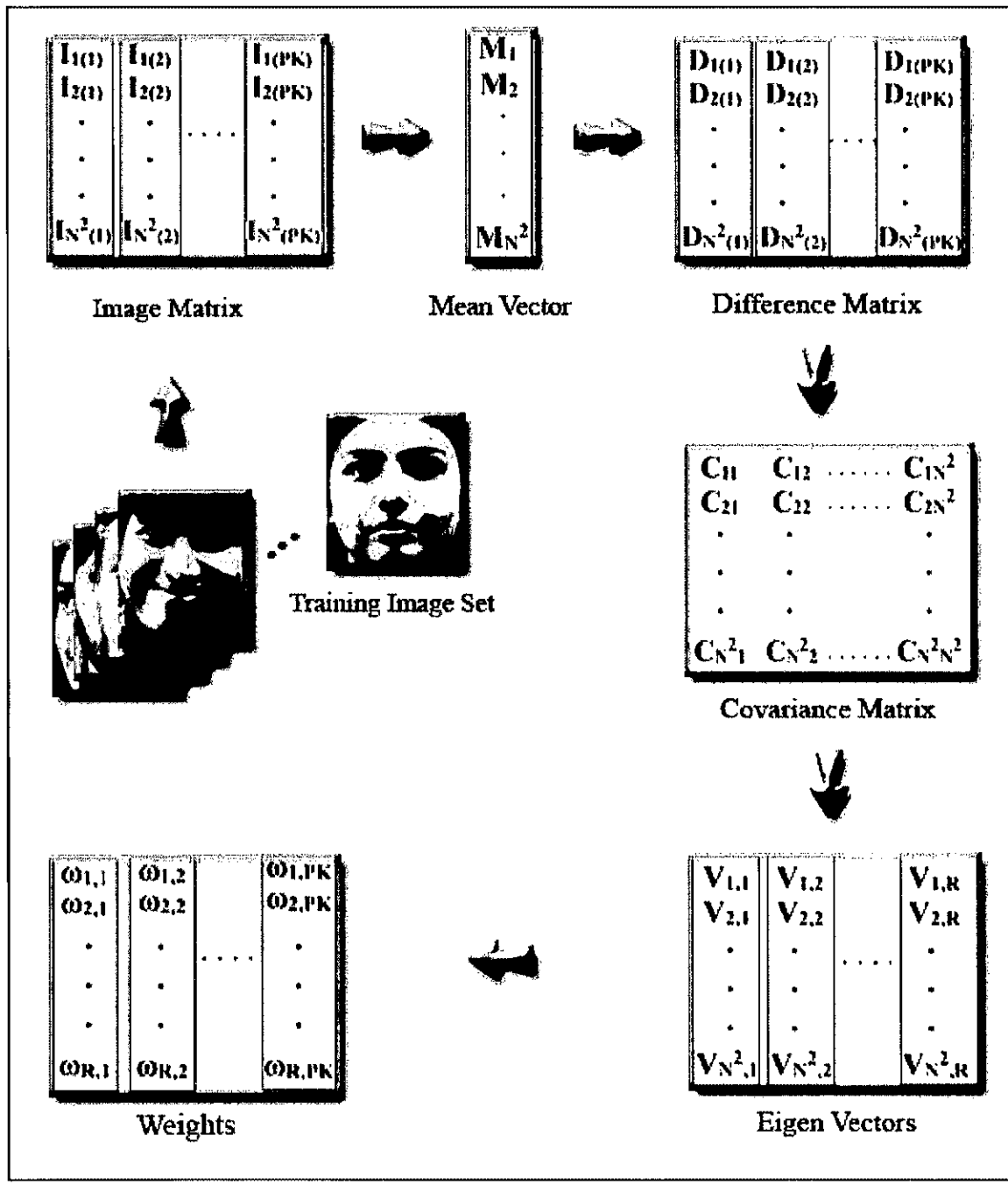

Figure 3.3: Illustration of the training phase for PCA based face recognition

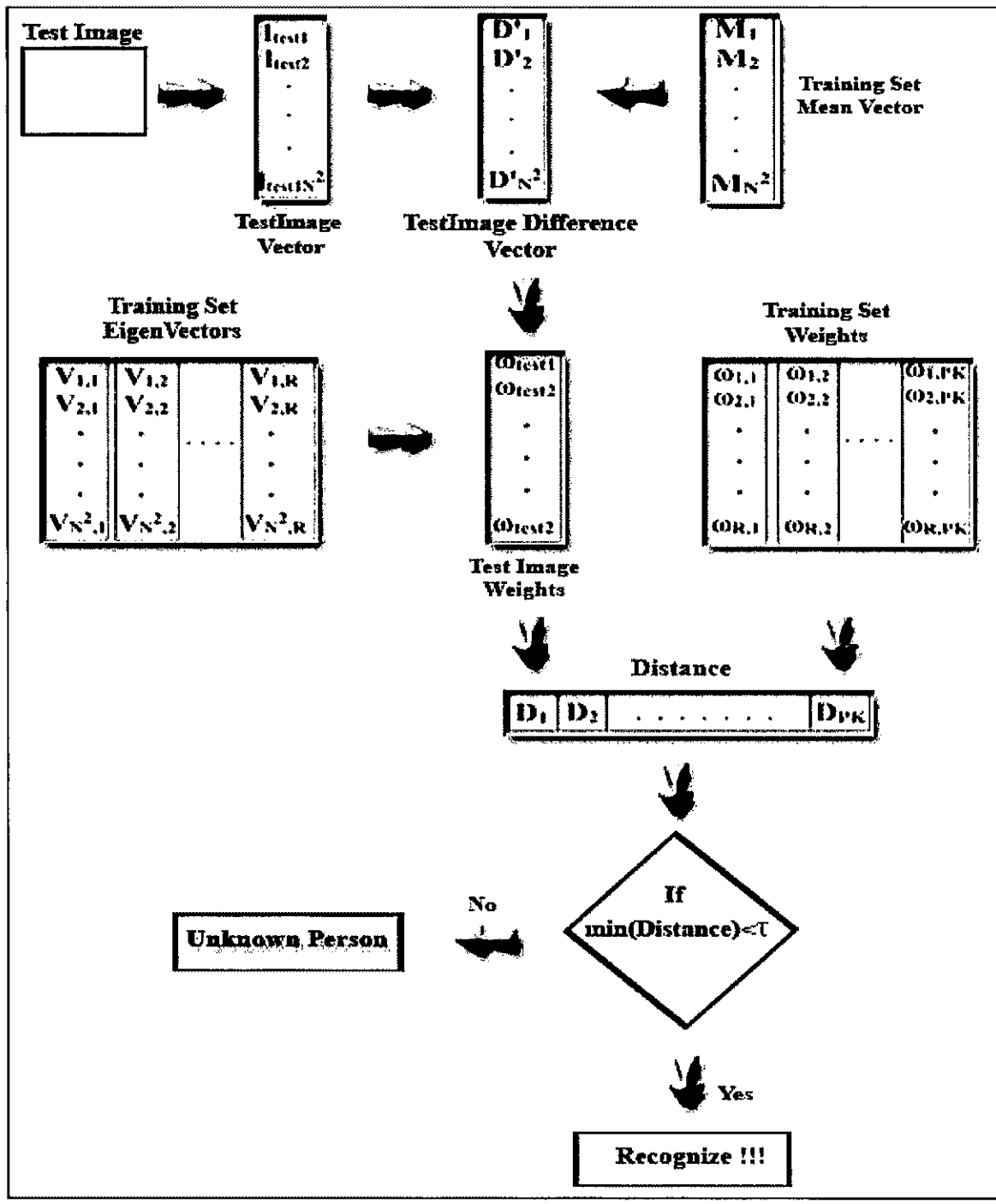

Figure 3.4: Illustration of the testing phase for PCA based face recognition

3.2 Face Authentication using Principal Component Analysis

The above mentioned universal eigenspace represents not only the inter variations between the different training individuals but also the intra variations of each individual with different expressions, different poses, ages, etc., which is mainly applicable for face recognition. Face authentication is different from face recognition in that face recognition is a one-to-many match whereas face authentication is a one-to-one match.

The algorithm for face authentication should be robust to each individual that is the squared norm of the difference between a weight of the test vector and its representation in the eigenspace is the measurement for authenticating an individual. From now onwards we will mainly focus on an individual subject. Therefore, for an image space of K images of each individual, the same method can be repeated P times, as P is the number of subjects.

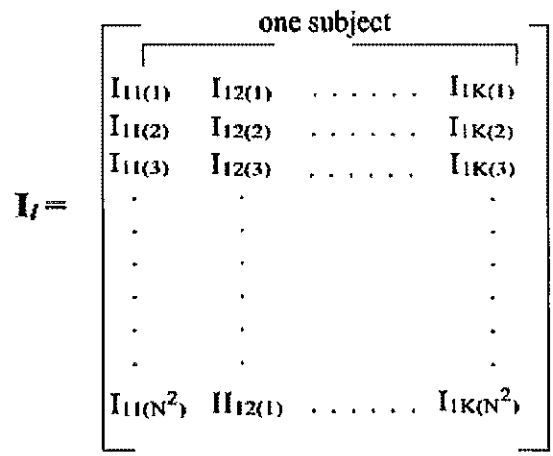

Figure 3.5: Representation of the image space for an individual

The image space of the l^{th} individual is represented by I_l as shown in figure 3.5, where l ranges from 1 to P, and the size of the image space is $N^2 \times K$. One eigenspace is

constructed for each training subject for the individual PCA technique. Now the average face image M_l for each subject can be obtained as,

$$M_l = \frac{1}{K}\sum_{m=1}^{K} I_{lm} \quad \forall l \tag{3-7}$$

Each training face differs from the mean by the vector $D_{lm} = I_{lm} - M_l$, $l=1,2,...,P$. The difference image matrix for the l^{th} subject can be represented as $A_l = [D_{l1}, D_{l2}, D_{l3}, ..., D_{lK}]$. In the individual PCA the eigenvector is denoted as $V_{l,r}$ and each face is projected on its own eigenspace by the equation,

$$\omega_{lm_r} = V_{l,r}^T (I_{lm} - M_l), \forall l, m, r \tag{3-8}$$

where R is the number of eigenvectors chosen.

The above algorithm is repeated for all P subjects. The weights for all the subjects are determined during the training phase. During the test phase of the authentication system, when a face image is captured and the individual claims to be subject l, first the face image I_{ltest} is projected to the l^{th} subject eigenspace to obtain the weights of the test face image by the equation,

$$\omega_{ltest_r} = V_{l,r}^T (I_{ltest} - M_l) \tag{3-9}$$

As mentioned in the universal PCA approach, the weights ω_{lm_r} form a vector $T_{lm} = [\omega_{lm1}, \omega_{lm2}, \omega_{lm3}, ...\omega_{lmR}]^T$ that describes the projection of each input face image of the training subjects in the eigenspace, where this vector is used to fit the test face image in the predefined face class. Similarly the weights ω_{ltest_r} of the test image form a vector T_{ltest} as $T_{ltest} = [\omega_{ltest1}, \omega_{ltest2}, \omega_{ltest3}.....\omega_{ltestR}]^T$.

To verify or authenticate the test image, two different techniques have been used in this thesis research. First is the simple Euclidean distance measurement technique, and the second is Mahalanobis distance [23]. In the Euclidean distance measurement, the Euclidean distance between T_{ltest} and T_{lm} is measured using the equation given by,

$$Dist_m = \left\| T_{ltest} - T_{lm} \right\| \ \forall m \tag{3-10}$$

If minimum of $Dist_m$ is less than or equal to a pre-selected threshold τ_l, then the test image is said to be authenticated as subject l. Setting the threshold is a difficult task, because smaller thresholds lead to a problem of false rejections and higher thresholds lead to false acceptance. Therefore a different threshold is set to each individual.

The second method of distance measurement is the Mahalanobis distance. When all the training face images of each subject I_{lm} are projected to the eigenspace, the mean face vector and the covariance matrix corresponding to each individual are determined using T_{lm} by the following equations,

$$T_l^{Mean} = \frac{1}{K} \sum_{m=1}^{K} T_{lm} \quad \forall l \tag{3-11}$$

$$T_l^{Cov} = \frac{1}{K} \sum_{m=1}^{K} (T_{lm} - T_l^{Mean})(T_{lm} - T_l^{Mean})^T \quad \forall l \tag{3-12}$$

Where T_l^{Mean} and T_l^{Cov} are the mean and the covariance matrices of weight vectors of subject l, which are calculated during the training phase. Coming to the testing phase of the face authentication system, when the individual claims to be subject l, the test face image I_{ltest} is captured and this image is projected to the eigenspace of subject l, and the

projected vector T_{ltest} is obtained. Then the Mahalanobis distance is measured using the mean and the covariance matrices of class l as,

$$Dist = (T_{ltest} - T_l^{Mean})^T T_l^{Cov^{-1}} (T_{ltest} - T_l^{Mean}) \quad (3\text{-}13)$$

The Mahalanobis distance is compared with the pre-selected threshold τ_l similar to that in the case of Euclidean distance measurement and the test face image can be rejected or accepted accordingly. Setting the threshold is a difficult task because a smaller threshold leads to a problem of False Rejection Rate (FRR), while a higher threshold leads to a problem of False Acceptance Rate (FAR). Therefore individual thresholds are set for each subject in the database in order to get the best results.

3.3 Experimental Set-up, Specifications and Results

The experiment for face authentication is done and the performance of this algorithm is tested on videos of different individuals captured with varying expressions, poses and occlusions. For the detection of the face images that are grabbed from the frames, a detection algorithm [42] developed by Viola and Jones is used. In ideal situations, the training and the test images should be frontal face images, properly cropped and resized. Since we are dealing with a real time application, this may not be possible. This is because the individual is not asked to pose before the camera. Instead the camera is positioned in such a way that a full frontal image of the customer's face is automatically captured, and the individual is unaware of being authenticated. This may result in non-frontal images getting captured, which in turn will affect the facial authentication process. To overcome this difficulty, multiple face images are captured over a certain interval of time in order to authenticate efficiently.

Most of the videos were grabbed using a Sony camcorder with 2.11 mega pixel resolution. For the training set of images, the images are grabbed at every twelfth frame from a 10-second video; the images are cropped and resized manually as the training procedure is not a real time process. During the testing phase, 10 face images of an individual are grabbed and stored at every seventh frame from a 3-second video. These 10 testing images of an individual are stored in a folder with their PIN which was entered to perform a transaction; the testing phase is carried out using this folder, and the system may authenticate or reject the individual accordingly. The folder used here is a temporary folder; it is not used later, because it is just to authenticate a person and not for the training purpose.

3.3.1 Camera specifications

The camcorder used for face authentication was a Sony DCR HC85 model and the specifications are:

- Resolution: 2.11 mega pixel
- Live video capture: Up to 640×480 pixels
- Frame rate: Up to 30 frames per second

3.3.2 Image specifications

The specifications of the images that are captured and processed are as follows:

- Input frame rate from the camera: 24 frames per second
- Frame of the image size captured by the camera: 640×480 pixels.
- Minimum size of detected face within each frame: 64×64 pixels
- Size of the face image that is processed for authentication: 64×64 pixels

- The images that are processed are Portable Gray Map (pgm) images.

3.3.3 System specifications and processing speeds

The system specifications and the approximate times of processing various modules are as follows:

System specifications:

- System processor: Intel Xeon(TM) 2.40 GHz
- System memory: 1 GB RAM

Approximate times of operations:

- Time taken to train the subjects into the authentication system: 2.2 sec
- Time taken to authenticate a person: approximately 1.2 sec

3.3.4 Experimental results

For the experiments, images of 20 individuals are used in the training phase. Most of the individuals involved belong to the ODU Vision Lab, and they belong to different ethnicities. The number of images considered for training the algorithm is 20 for each individual, and the size of each image used is 64×64 pixels. All the images used are portable gray map images. A total of 400 images are used for the training algorithm of 20 individuals. These images are grabbed from the video during the enrollment process and are processed to get the weights of each individual. During the testing phase, while the individual enters his or her personal identification number the video is recorded for about 3 seconds, and 10 frames are grabbed and used for authentication. The experiment was

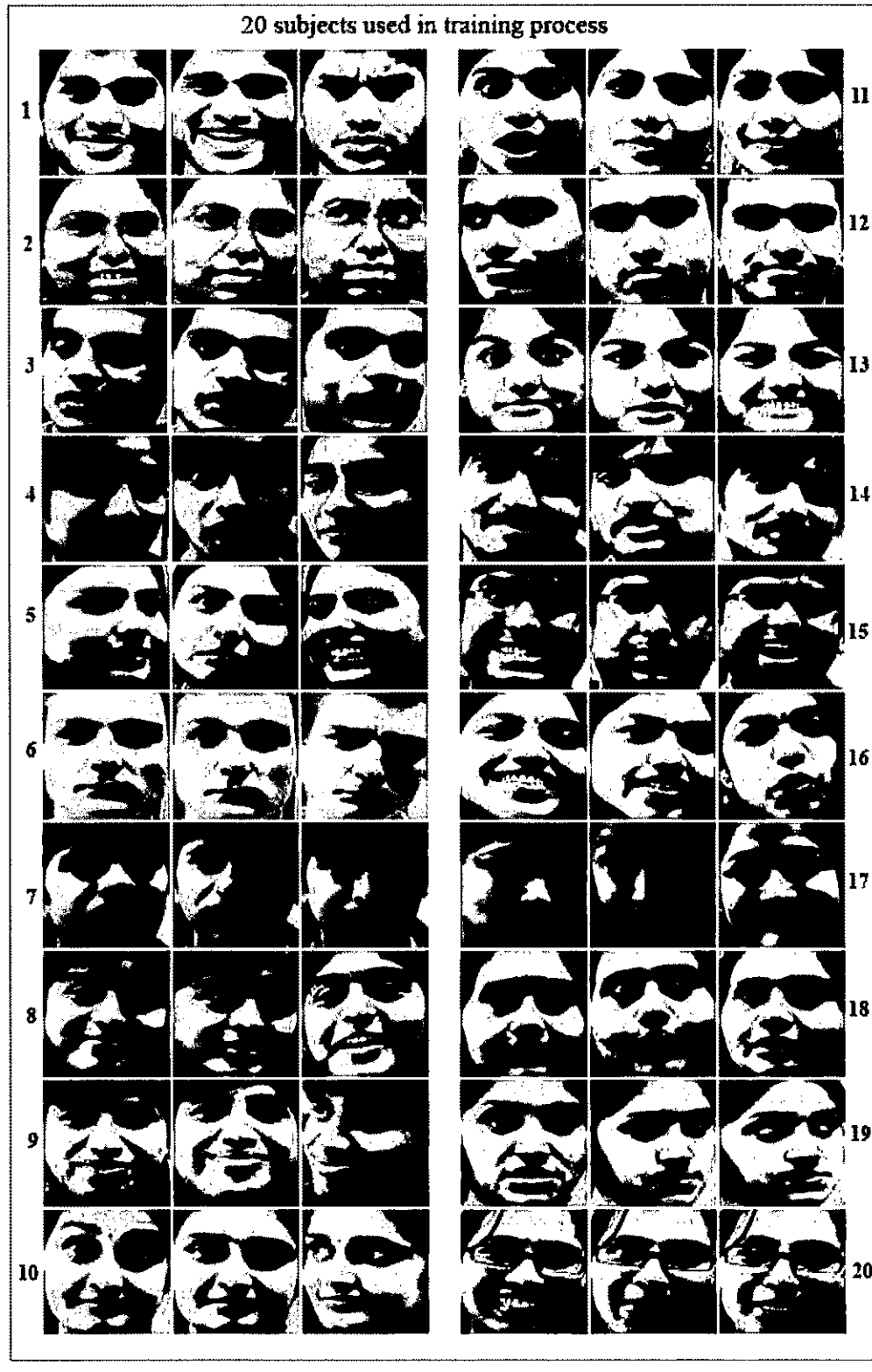

Figure 3.6: Sample training images of 20 individuals

conducted on different individuals that are in the database and also on unknown individuals. Tests were conducted on the face images, which contain occlusions such as sun glasses, masks, facial hair and makeup, along with varying expressions and poses. Sometimes people who come for a transaction may be talking on a phone with different expressions and poses. Therefore, the experiment was conducted in order to get perfect simulations even in such different operational scenarios.

Figure 3.6 shows some of the sample training images of 20 individuals and the corresponding PINs are shown next to the images. It can be observed that all the training face images are fairly neutral with few expression and pose variations. The testing database consists of different videos of different individuals taken at different times; some of the individuals belong to the training database and some of them are unknown individuals.

Figure 3.7: Sample test images of two subjects with PINs 9 and 5

The videos of different individuals are taken at different times with different occlusions and different expressions to test the system. The sample test images shown in

figure 3.7, which are similar to the training images, were grabbed from a video captured at an interval of 10 seconds on different days.

Since the verification or authentication is carried out as a one-to-one matching process, the algorithm is executed and tested for the 10 images of the individual that were grabbed from the video that is captured for around 3 seconds when the person enters his or her personal identification number to perform a transaction. Out of these 10 grabbed face images, if at least two of the test images are matched with the information present in the database, then the person can perform the transaction, else he is not allowed to continue.

The performance of PCA for face authentication was tested by varying the number of eigenvectors. More eigenvectors resulted in an increased authentication rate, thereby decreasing the false acceptance rate. In order to analyze the system, the algorithm was executed several times on videos of known individuals and also of unknown individuals that were captured at different days. The tests were performed for test images that were similar to the training images and also on images with occlusions.

For the above mentioned test, in which face images are similar to the training images, we can observe from the plot shown in figure 3.8 that the authentication rate increases with the number of eigenvectors used for authentication. The number of eigenvectors is varied from 5, 10, 15, 20, 30 and so on up to 80. It was observed that when the number of eigenvectors is greater than 20, the authentication rate is around 91 percent, and there is not much improvement for the number of eigenvectors greater than 20.

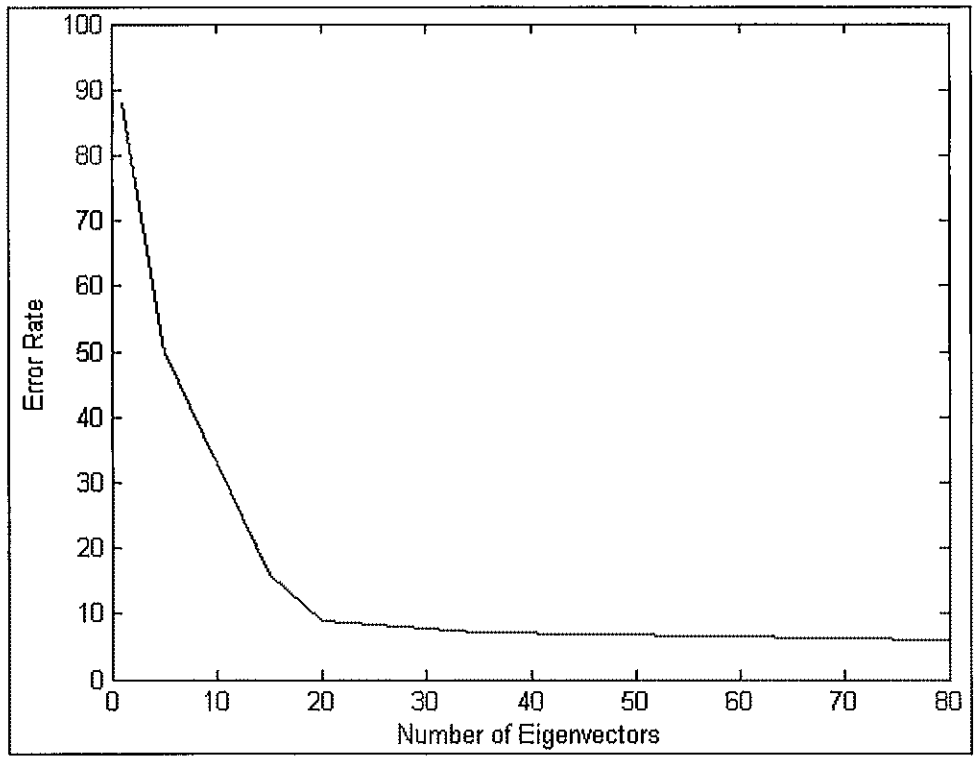

Figure 3.8: Error characteristics with respect to number of eigenvectors using universal PCA (for images that are similar to training set)

Figure 3.9: Sample test images of two subjects with occlusions

Setting the thresholds and getting the best results for videos that have faces with occlusions is the main objective of this process. When the algorithm is tested with test images of individuals with occlusions such as sun glasses, talking on a phone, a face mask and so on as shown in figure 3.9, it can be observed that the authentication rate is not satisfactory. Figure 3.10 shows that the authentication rate increases with respect to the number of eigenvectors for test images with occlusions.

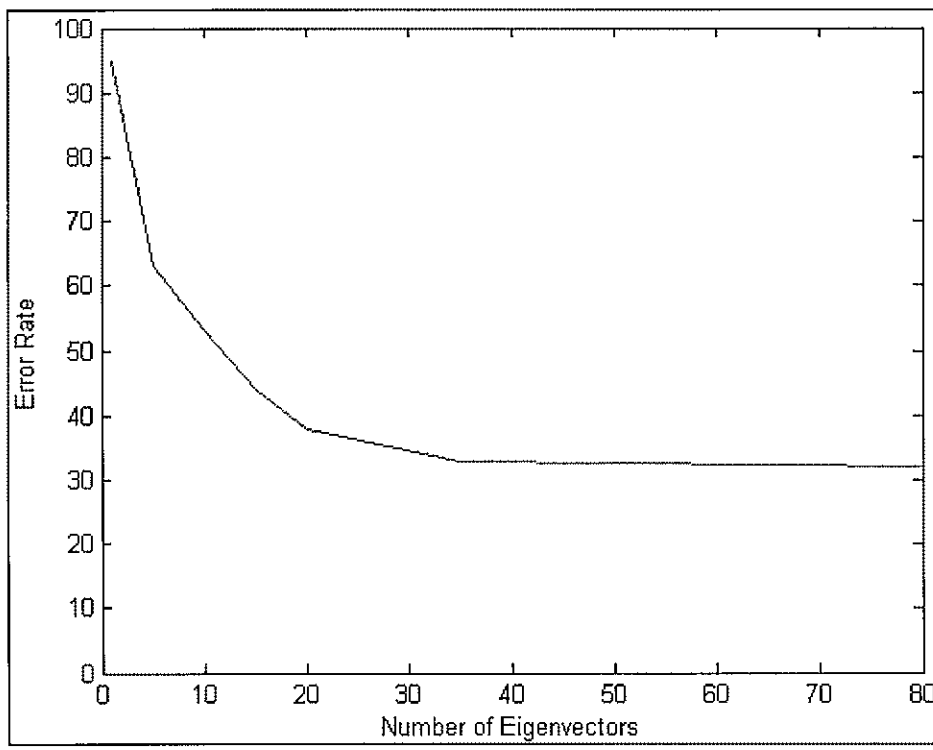

Figure 3.10: Error characteristics with respect to number of eigenvectors using universal PCA (for images with occlusions)

The number of eigenvectors is varied in the same manner as before from 5, 10, and so on up to 80. In this case it is observed that with a number of eigenvectors greater

than 20 the authentication rate is around 64 percent only, and there is not much improvement for a number of eigenvectors greater than 20 for the face images with occlusions. It can be observed from the Receiver Operating Characteristics (ROC) curve as shown in figure 3.11 that for a particular false acceptance rate of 36 percent the authentication rate achieved is 100 percent.

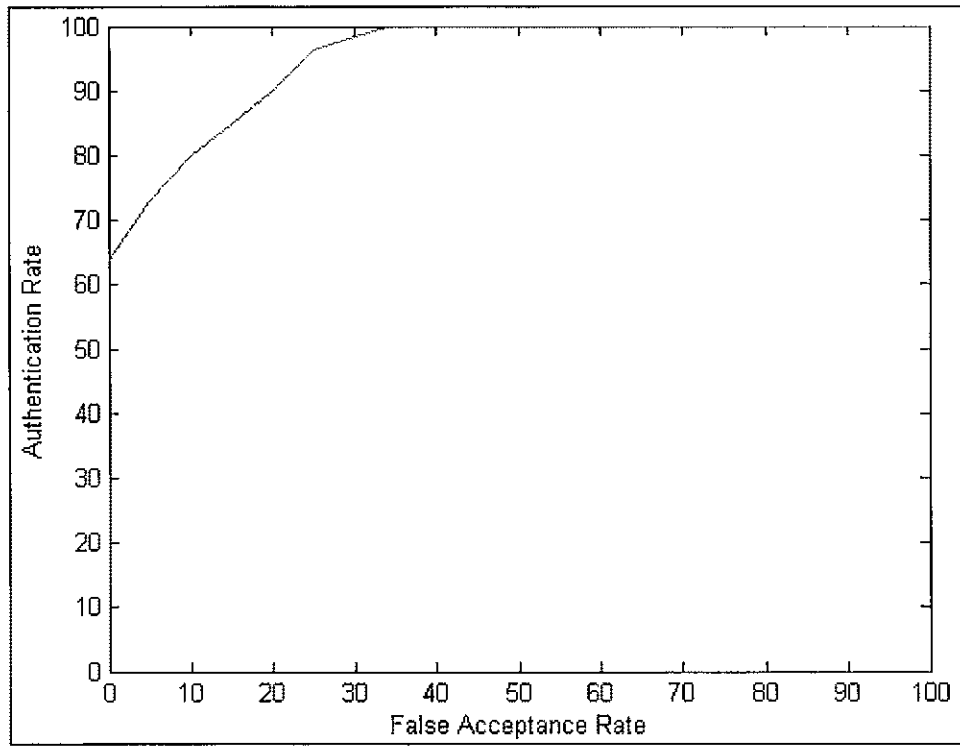

Figure 3.11: ROC curve of universal PCA

3.4 Summary

In this chapter, the basic concept behind the PCA and how it is used in a real time face authentication systems has been presented. The experimental set up and the performance evaluation has been explained. The primary goal of face authentication is to

authenticate a person even with variations in facial expression and pose and with occlusions. Based on the experimental results it can be observed that with the PCA based technique for face authentication the algorithm performed well with the test images that are similar to the training images, but it did not perform well for the test images that were with varying expressions, poses and occlusions. Therefore, there is a need for a technique that can achieve the actual goal for face authentication. A modified PCA method is investigated and presented in chapter 4.

Chapter 4

FACE AUTHENTICATION USING MODULAR PCA

In order to improve the performance of the face authentication algorithm a modified PCA method called modular PCA [5] is investigated. In this technique, the face images are divided into sub images or modules, and then PCA is applied to each of the modules. In the universal PCA approach, the algorithm is trained considering the entire face image as a single unit. If the face region has large variations in pose and expression it may affect the authentication rate profoundly. The modular PCA approach helps in authenticating a person even if some of the face images are affected due to varying poses and expressions and also with occlusions. This increases the performance of the authentication system when compared to the universal PCA approach.

This chapter deals with the theory of modular PCA and its application to face authentication. The increased performance of the face authentication over the universal PCA approach is presented.

4.1 Face Authentication using Modular Principal Component Analysis (MPCA)

The experimental results in chapter 3 showed that face authentication using universal PCA is not efficient under conditions of varying poses and expressions, as the method considers each face image as global information in order to obtain the projections. Under these conditions the projections of the image will vary considerably with the projections of the images with normal pose and expressions, hence making it difficult to authenticate an individual. In modular PCA the face images are divided into

modules and the weights are determined for each of these modules. This results in an effective representation of the local information of the face. In this sub images approach, only some of the face regions vary with varying poses and expressions and the remaining regions are not affected. Therefore only some of these sub images are affected. The projections of test face images that are not affected by varying poses and expressions will closely match the projections of the training sub images of the same subject under normal conditions, thereby making the authentication process more efficient.

In the modular PCA approach, each image in the training set of all the individuals in the database is divided into S non overlapping smaller images. That is, the size of each sub image vector is N^2/S. The mathematical expression for obtaining these sub images is represented by,

$$I_{lmg}(a,b) = I_{lm}\left(\frac{N}{\sqrt{S}}(i-1)+a, \frac{N}{\sqrt{S}}(j-1)+b\right) \text{for}, 1 \le i,j \le \sqrt{S}\ ,\ 1 \le a,b \le (N/\sqrt{S}) \quad (4\text{-}1)$$

where m denotes the index of the face image of an individual, l denotes the index of the subject and g denotes the module number; (a, b) represents the pixel location. The size of the original image is $N \times N$. The relation between g, i and j is $g = (i-1)\sqrt{S} + j$.

Figure 4.1 shows an example of how the face image is divided into 16 sub images; that is, S = 16. To get a sub image S_{10} for all the images of an individual, the value of i is equal to 3, and the value of j is equal to 2. After, all the face images are divided into modules to get the sub images. Each of the corresponding sub images of all the face images of an individual is loaded into a matrix to form the sub image space; for example, if the face images in the training set are divided into 16 sub images, say S_1 to S_{16}, the sub images S_1 of all the face images of an individual are loaded into one sub

image matrix, which is of size $(N^2/S)\times K$, where K is the number of face images of each subject. Similarly we have 15 other sub image spaces, and now the universal PCA is applied to all these 16 sub image spaces separately, and the weights are obtained.

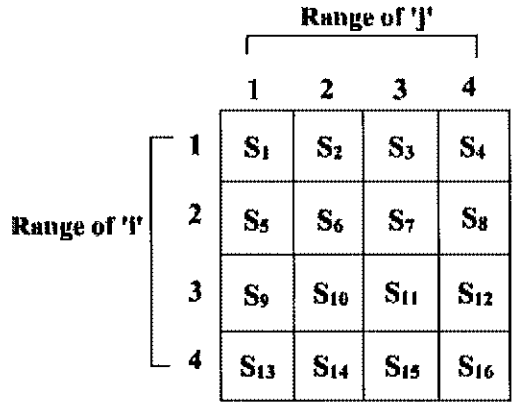

Figure 4.1: Representation of splitting the face image in modular PCA

Figure 4.2 illustrates the training phase of modular PCA based face authentication. The average face image M_{lg} will be different for each subject and for each sub image, and it is obtained as,

$$M_{lg} = \frac{1}{K}\sum_{m=1}^{K} I_{lmg} \quad \forall l, g \tag{4-2}$$

Each training face sub image differs from the mean sub image by the vector $D_{lmg} = I_{lmg} - M_{lg}$. The difference matrix A_{lg} is of size $(N^2/S)\times K$, which can be represented as $A_{lg} = [D_{l1g}, D_{l2g}, D_{l3g}, ..., D_{lKg}]$. As explained in chapter 3, using the covariance matrix $A_{lg}A_{lg}^T$, the eigenvectors V_{lg_r} are obtained. Each face sub image is projected on its own eigenspace by the equation,

$$\omega_{lmg_r} = V_{\lg_r}^T (I_{lmg} - M_{\lg}) \tag{4-3}$$

where, $1 \le r \le R$ and R is the number of eigenvectors chosen.

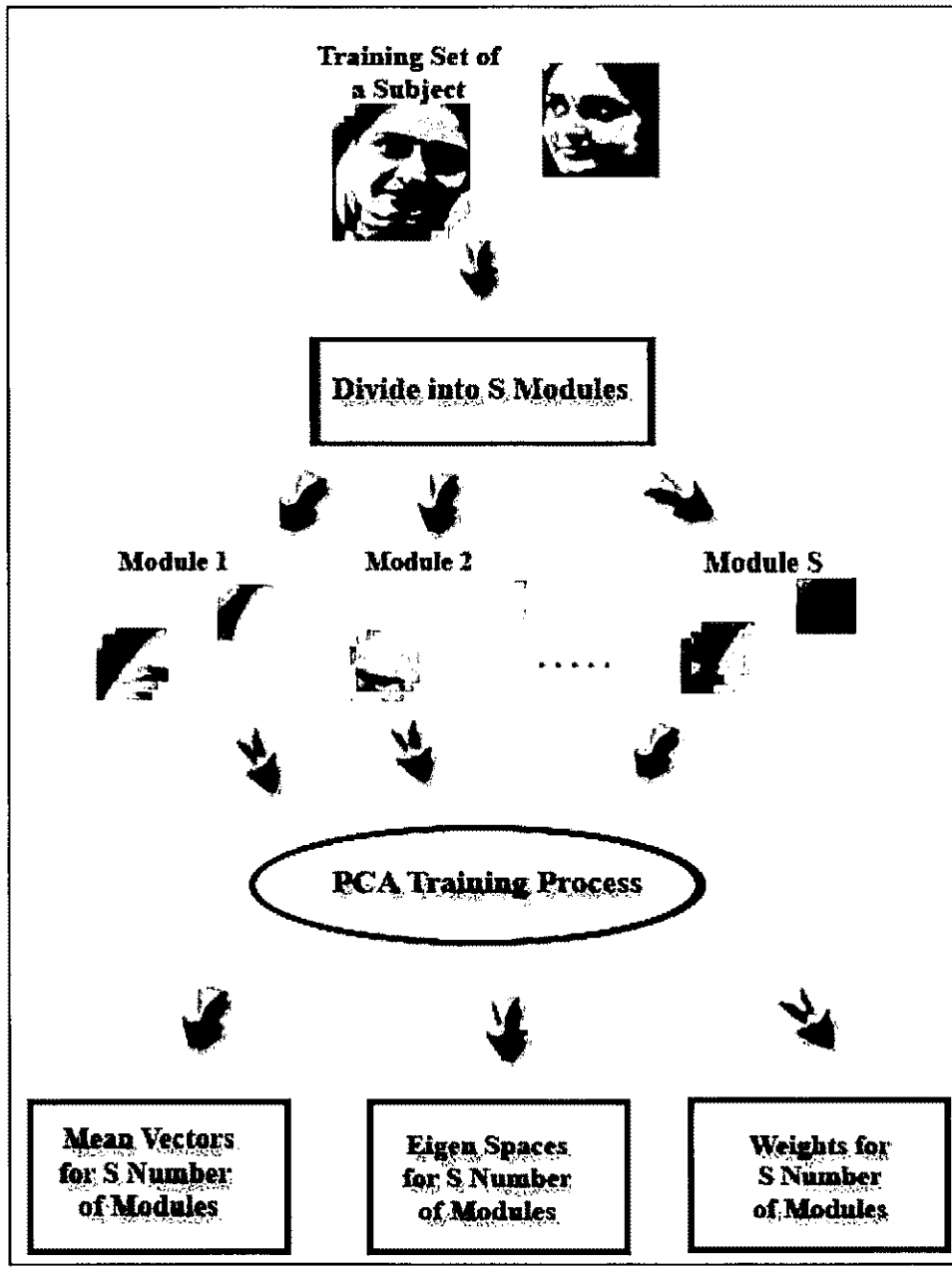

Figure 4.2: Illustration of the training phase for modular PCA based face authentication

The above algorithm is repeated PS times, for S sub images and for P subjects. The weights of all the sub images of all the subjects are determined during the training phase. The weights of the training set ω_{lmg_r} form a matrix T_{lmg} by the equation, $T_{lmg} = [\omega_{lmg_1}, \omega_{lmg_2}, \omega_{lmg_3}, \ldots\ldots \omega_{lmg_R}]^T$. It describes the projection of each sub image of the training subject in the eigenspace, where this eigenspace is used to fit the corresponding test sub image in the predefined face class.

Figure 4.3 illustrates the testing phase for modular PCA based face authentication. During the test phase of the authentication system, when a test face image I_{ltest} is captured and the individual claims to be subject l, the image is divided into modules according to the equation (4-1), and then the sub image I_{ltestg} of the test face image is projected to the l^{th} subject eigenspace to obtain the weights of the test face image by the equation,

$$\omega_{ltestg_r} = V_{lg_r}^T (I_{ltestg} - M_{lg}) \ \forall g, r \tag{4-4}$$

The weights ω_{ltestg_r} of the test image form a vector T_{ltestg} as $T_{ltestg} = [\omega_{ltestg_1}, \omega_{ltestg_2}, \omega_{ltestg_3}, \ldots\ldots \omega_{ltestg_R}]^T$. The same algorithm is used to obtain the eigenspaces for the rest of the sub images, and the corresponding training sub image eigenspaces are used to get the error measurement.

Again, two methods are used for distance computation, the Euclidean distance and the Mahalanobis distance measures. The Euclidean distance measure is determined between T_{ltestg} and T_{lmg} using the equation,

$$Dist_{mg} = \left\| T_{ltestg} - T_{lmg} \right\| \ \forall g, m \tag{4-5}$$

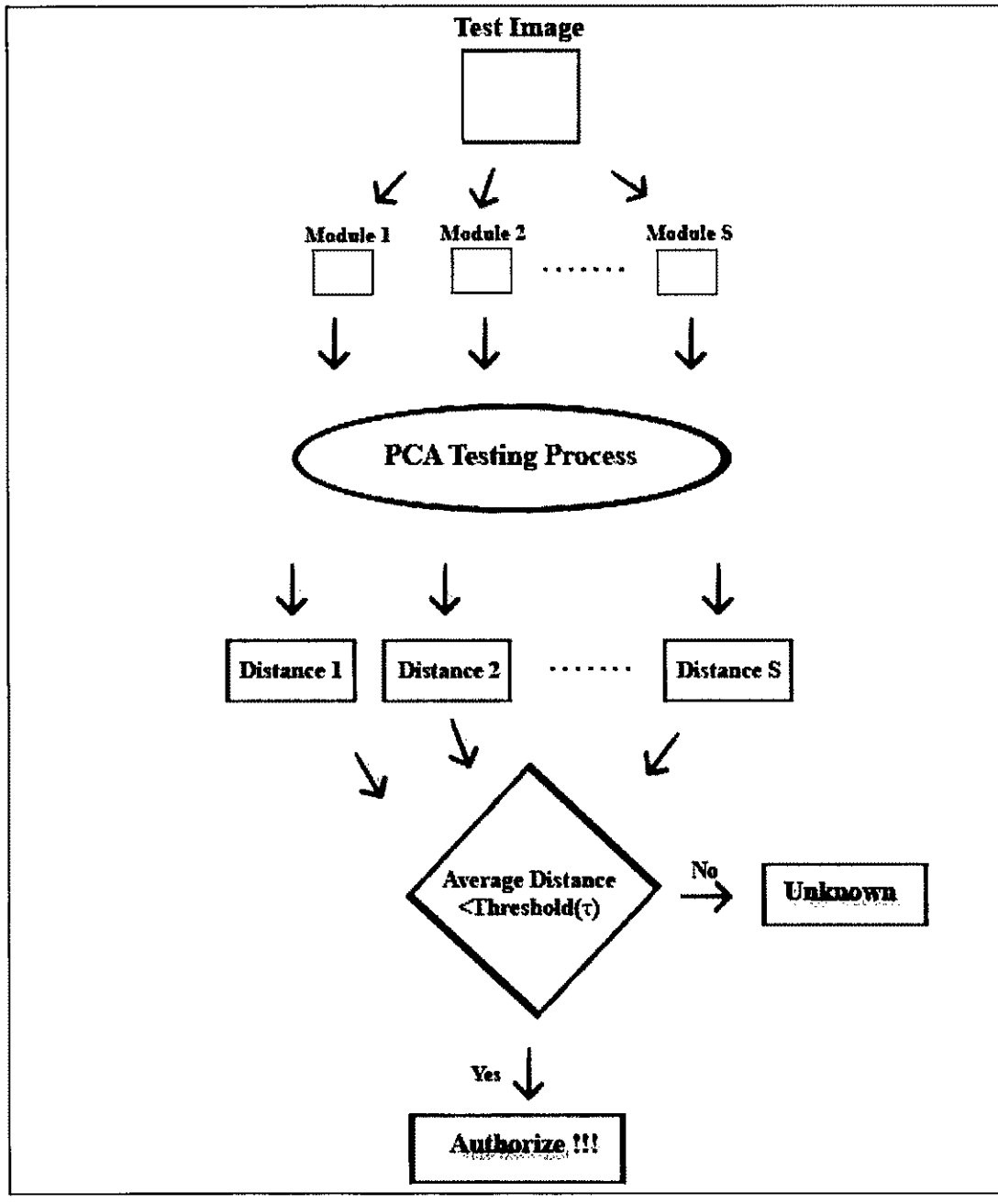

Figure 4.3: Illustration of the testing phase for modular PCA based face authentication

A minimum of $Dist_{mg}$ is determined for each module that is $Dist_g = \min_m(Dist_{mg})$. If the average distance $Dist_g$ for all modules is less than or equal to a pre-selected threshold τ_l, the test image is said to be authenticated as subject l, where τ_l is different for every individual.

The Mahalanobis distance measurement is obtained using the equation,

$$Dist_g = (T_{ltestg} - T_{lg}^{Mean})^T T_{lg}^{Cov^{-1}} (T_{ltestg} - T_{lg}^{Mean}) \quad \forall g \tag{4-6}$$

T_{lg}^{Mean} and T_{lg}^{Cov} are the mean and the covariance matrices of weight vectors of all the corresponding sub images of subject l that are determined during the training phase using the weight vector T_{lmg} as,

$$T_{lg}^{Mean} = \frac{1}{K}\sum_{m=1}^{K} T_{lmg} \quad \forall g,l \tag{4-7}$$

$$T_{lg}^{Cov} = \frac{1}{K}\sum_{m=1}^{K} (T_{lmg} - T_{lg}^{Mean})(T_{lmg} - T_{lg}^{Mean})^T \quad \forall g,l \tag{4-8}$$

If the average of the Mahalanobis distance $Dist_g$ is less than or equal to a pre-selected threshold τ_l, then the test face image is said to be authenticated.

4.2 Experimental Results

The performance of face authentication using modular PCA is evaluated with the same database as discussed in chapter 3. The performance of this algorithm was tested by varying the number of eigenvectors. Greater numbers of eigenvectors resulted in an increased authentication rate and thus decreased false acceptance rate. Tests were

conducted to analyze the algorithm with different videos of individuals with varying poses and expressions and also with occlusions and on different days. Tests were also conducted by varying the number of eigenvectors and by varying the number of modules.

Figure 4.4: Sample testing images with varying expressions, poses and occlusions used for modular PCA

The algorithm was executed and tested for 10 images of an individual that were grabbed from the video that was captured for around 3 seconds while the person entered his or her PIN to perform a transaction. Out of these 10 test face images, if at least two resulted in a match in the face authentication process, then the person was allowed to

perform the transaction, otherwise the person was not allowed to continue. Figure 4.4 shows sample test images of 5 subjects.

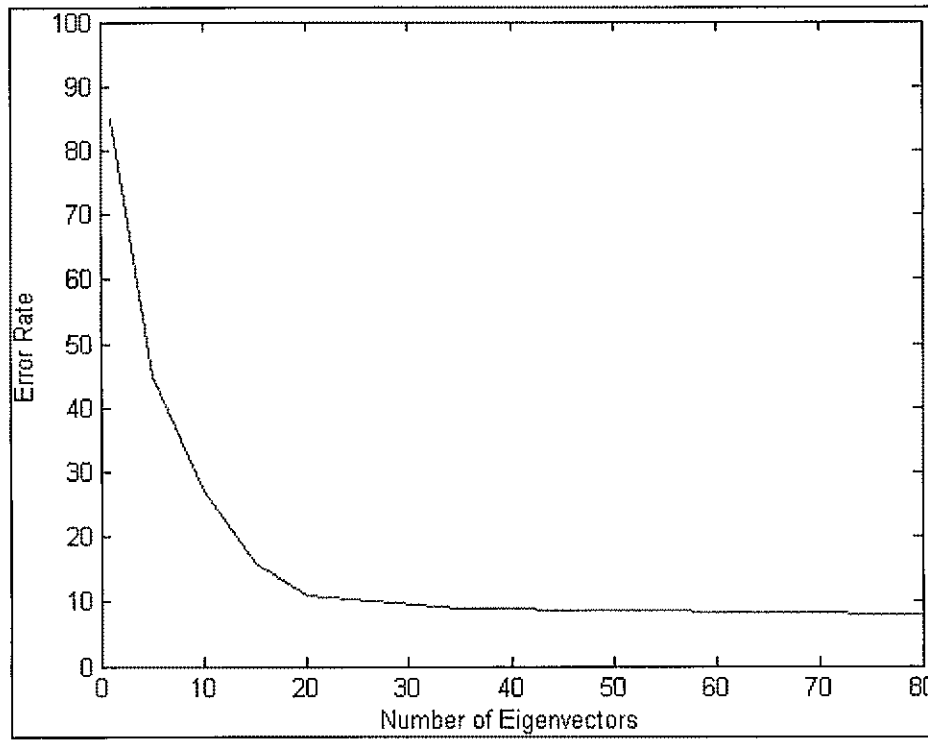

Figure 4.5: Error characteristics with respect to number of eigenvectors using modular PCA (for images with expressions, poses and occlusions)

It can be observed from figure 4.5 that the authentication rate is increasing with the increase in the number of eigenvectors, as in the case of face authentication using PCA. The number of eigenvectors is varied from 5, 10, 15, and so on up to 80, and we can observe that for the number of eigenvectors greater than 20, the authentication rate is around 89 percent, and there is not much improvement for the number of eigenvectors greater than 20. Similarly, the experiment was performed by varying the number of

modules by changing the value of S. As the image size for the training set of face images is 64×64, the S values used were 4, 16, 32 and 64. Figure 4.6 shows the error characteristics for the different number of modules.

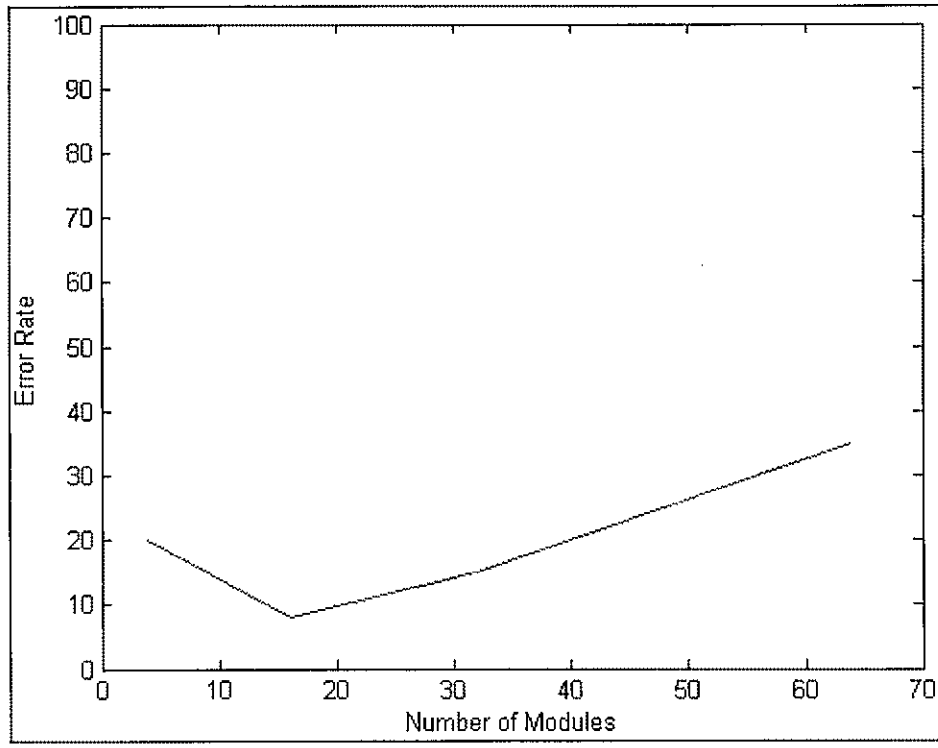

Figure 4.6: Error characteristics with respect to number of modules using modular PCA (for images with expressions, poses and occlusions)

From figure 4.6, we observe that the authentication rate was better for the number of modules equal to 16 and degraded later as the number of modules was increased. For face authentication with number of modules equal to 16 and the number of eigenvectors equal to 25, the authentication rate is 92 percent for test videos with varying expressions and poses and also with occlusions.

It can be observed from the receiver operating characteristics curve as shown in figure 4.7 that for a particular false acceptance rate of 8 percent the authentication rate achieved is 100 percent. We can observe that the modular PCA helped in improving the authentication accuracy over the universal PCA.

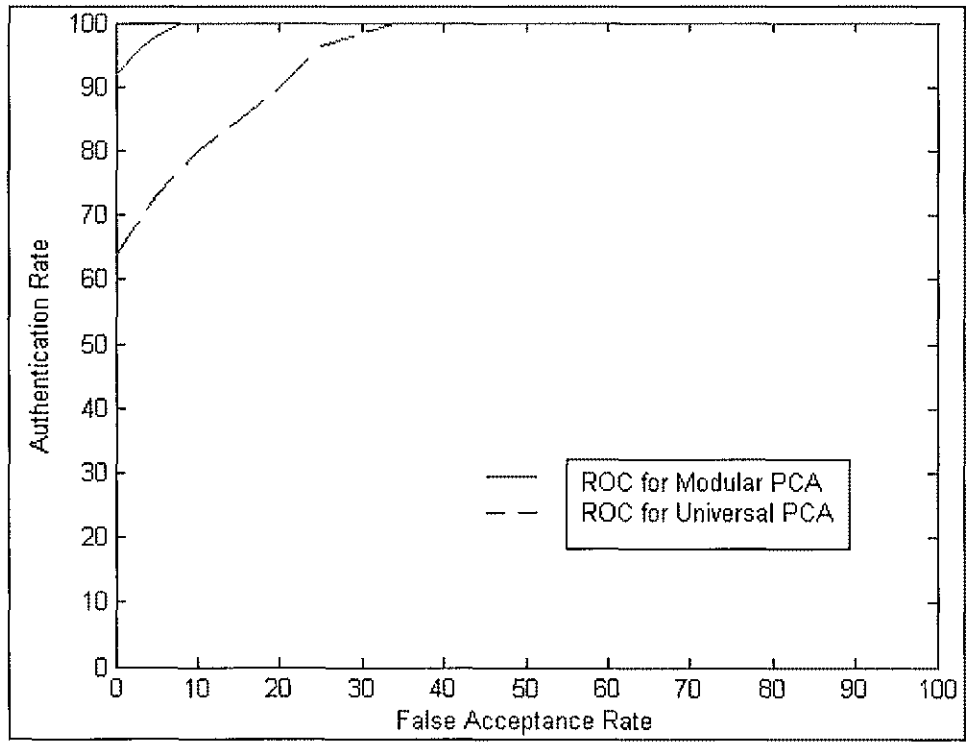

Figure 4.7: ROC curves of universal PCA and modular PCA

4.3 Summary

In this chapter, the application of modular PCA technique in real time face authentication systems has been presented. The performance evaluation for the face authentication is discussed. Based on the experimental results it can be observed that with the modular PCA based algorithm, face authentication performs well with the test images

not only similar to that of the training images but also with varying expressions, poses and occlusions. It can be observed that the dimensionality reduction technique on individual modules of the face images improved the accuracy of face authentication when compared to that of applying on the whole image.

Chapter 5

FACE AUTHENTICATION USING WEIGHTED MODULAR PCA

In order to get better authentication performance, a modified PCA approach named weighted modular principal component analysis (WMPCA) [6] is employed and presented in this chapter. In this method, the face images are divided into four horizontal sub regions such as forehead, eyes, nose and mouth with chin, and PCA is performed on each sub image; in modular PCA the face images are divided into a large number of square modules. The weighted modular PCA approach is based on a comparison between the main features of the individual with the respective information in the database. It is expected to yield better performance for face authentication when compared to the other two methods described in the previous chapters.

This chapter presents the description of the weighted modular principal component analysis approach for face authentication and shows the increased performance of the face authentication over the other two approaches.

5.1 Face Authentication using Weighted Modular Principal Component Analysis

It was shown in chapter 4 that dimensionality reduction techniques on individual modules of the face images improved the accuracy rate of face authentication when compared to the universal PCA where the whole image is considered as a single unit. Therefore the concept of modules is used again in this research, but in a different way. In this method the face images are divided into horizontal sub regions such as forehead, eyes, nose and mouth with chin regions. Each sub region is analyzed separately using the

principal component analysis. This technique is mainly a comparison of the corresponding features of the face images. The sum of the weighted distances between each sub region of test image to the respective sub images in the training set is calculated. The final decision for accepting or rejecting an individual is based on the sum of the weighted distances obtained from each sub region. [6]

In the weighted modular PCA approach, each image in the training set of all the individuals in the database is divided into four horizontal rectangular sub regions. That is the size of each sub image vector is $N^2/4$, the size of the original face images being $N \times N$. The mathematical expression for dividing into sub images is given by,

$$I_{lmg}(a,b) = I_{lm}\left(\frac{N}{4}(g-1)+a,\ b\right) \tag{5-1}$$

where g is the module number used (in this case 4 modules are used so, $1 \le g \le 4$), l is the index of the subject, and m is the index of the face images of each subject; $1 \le a \le (N/4)$ and $1 \le b \le N$. PCA training process is carried out to get the weights of all the sub regions of each individual as mentioned in chapter 4.

For each sub region I_{lmg} of each individual, the mean vector, the difference matrix, the covariance matrix, eigenvectors and the projections are calculated in the same manner as described in chapter 4. All the K images of an individual are divided into sub regions using the equation (5-1). Figure 5.1 shows the division of the image into horizontal sub regions. As the face images in the training set are divided into 4 sub-regions, say S_1 to S_4, the sub region S_1 of all the face images of an individual is loaded

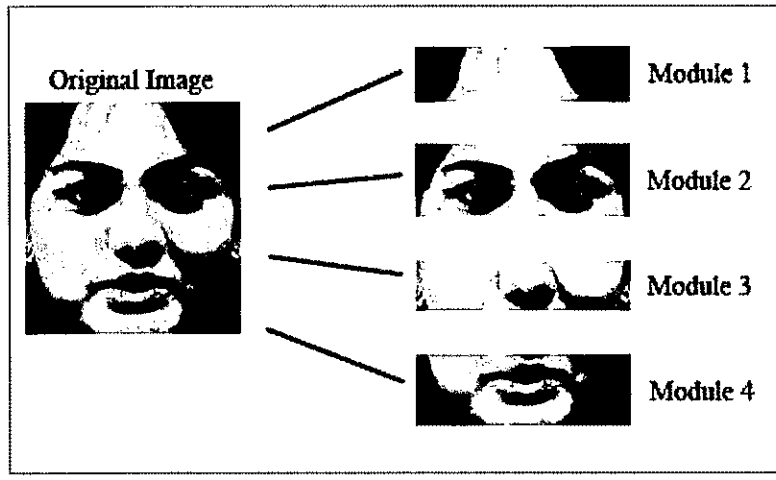

Figure 5.1: Representation of splitting the face image in weighted modular PCA

into one sub region space that is of size $(N^2/4) \times K$. Similarly the other 3 sub region spaces are formed; the universal PCA is applied to all these 4 sub region spaces separately, and the weights are obtained for each sub region.

Figure 5.2 illustrates the training phase of weighted modular PCA based face authentication. The algorithm is explained for only one sub region and the same procedure is applied to all the sub regions of each individual. Now the average face image M_{lg} is calculated as,

$$M_{lg} = \frac{1}{K} \sum_{m=1}^{K} I_{lmg} \tag{5-2}$$

where $1 \le l \le P$ and $1 \le g \le 4$. Each training sub image of an individual differs from the mean sub image by the vector $D_{lmg} = I_{lmg} - M_{lg}$. The difference matrix A_{lg} can be represented as $A_{lg} = [D_{l1g}, D_{l2g}, D_{l3g}, \ldots\ldots, D_{lKg}]$. As explained in chapter 4, using the

covariance matrix $A_{lg}A_{lg}^{T}$, the eigenvectors V_{lg_r} are obtained. Then the weights are calculated using the equation,

$$\omega_{lmg_r} = V_{lg_r}^{T}(I_{lmg} - M_{lg}),\ 1 \le g \le 4 \text{ and } 1 \le r \le R \tag{5-3}$$

where R is the number of eigenvectors chosen. The weights ω_{lmg_r} of the training set sub regions form a matrix T_{lmg} by the equation, $T_{lmg} = [\omega_{lmg_1}, \omega_{lmg_2}, \omega_{lmg_3}, \ldots\ldots\omega_{lmg_R}]^{T}$. It describes the projection of each sub image of the training subject in the eigenspace, where this eigenspace is used to fit the corresponding test sub image in the predefined face class. The above algorithm is repeated *4P* times (that is for *4* sub images and for *P* subjects). The weights of all the sub regions of all the subjects are determined during the training phase.

Intra-subject variance ψ_{lg} for each of the sub regions is calculated and is used as the weight for classification. It mainly assigns the priorities for each sub region. For each sub region g of an individual *l* the variance is computed as,

$$\psi_{lg} = \frac{1}{K}\sum_{m=1}^{K}[T_{lmg} - T_{lg}^{Mean}]^2 \qquad \text{for } 1 \le l \le P \text{ and } 1 \le g \le 4 \tag{5-4}$$

where,

$$T_{lg}^{Mean} = \frac{1}{K}\sum_{m=1}^{K} T_{lmg} \quad \forall l, g \tag{5-5}$$

Figure 5.3 illustrates the testing phase for weighted modular PCA based face authentication. During this phase of the face authentication system, when a test face image I_{ltest} is captured and if the individual claims to be the subject *l*, first the image is

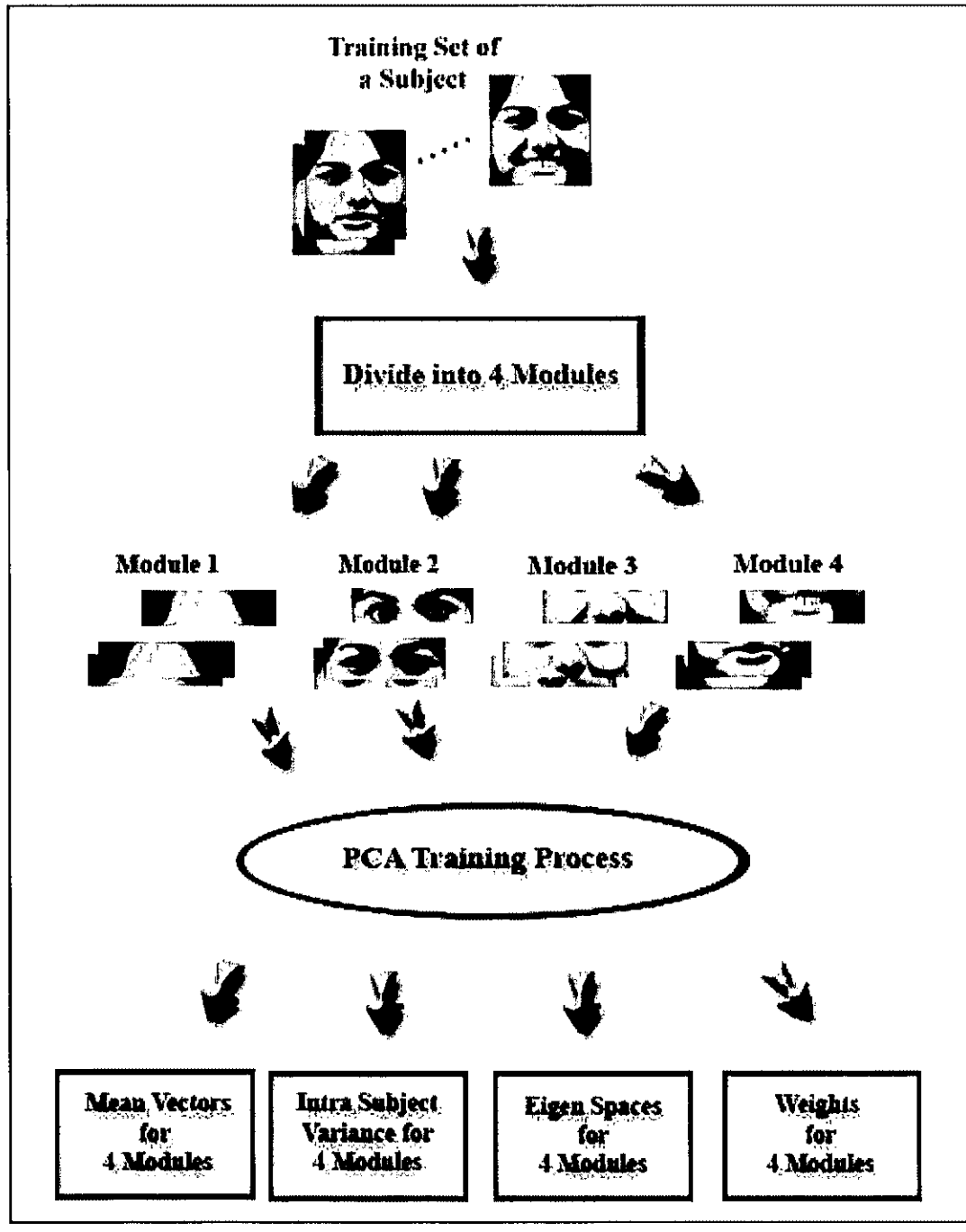

Figure 5.2: Illustration of the training phase for weighted modular PCA based face authentication

divided into four modules using the equation (5.1). Then each sub region I_{ltestg} of the test face image is projected to its respective eigenspaces of subject l to obtain the weights of the test face image by the equation,

$$\omega_{ltestg_r} = V_{\lg_r}^T (I_{ltestg} - M_{\lg}) \quad \text{for, } 1 \le r \le R \tag{5-6}$$

The weights ω_{ltestg_r} of the test image form a vector T_{ltestg} as $T_{ltestg} = [\omega_{ltestg_1}, \omega_{ltestg_2}, \omega_{ltestg_3}, \ldots\ldots\omega_{ltestg_R}]^T$.

Again, two methods are used for distance computation, the Euclidean distance and the Mahalanobis distance. The Euclidean distance measure is determined using the equation,

$$Dist_{mg} = \left\| T_{ltestg} - T_{lmg} \right\| \qquad \text{for } 1 \le g \le 4,\ 1 \le l \le P \text{ and } 1 \le m \le K \tag{5-7}$$

The sub region that is more invariant to the expressions and poses has the highest priority. Similarly, the sub regions that are less invariant to the expressions and poses have less priority. The minimum distance $Dist_g$ for each sub region is found by,

$$Dist_g = \min_m (Dist_{mg}) \quad \forall g \tag{5-8}$$

This final measure for deciding the authentication is the weighted sum of the distances of individual sub regions that is given by the equation,

$$Dist_{test} = \sum_{r=1}^{4} [\psi_{\lg} Dist_g^2] \tag{5-9}$$

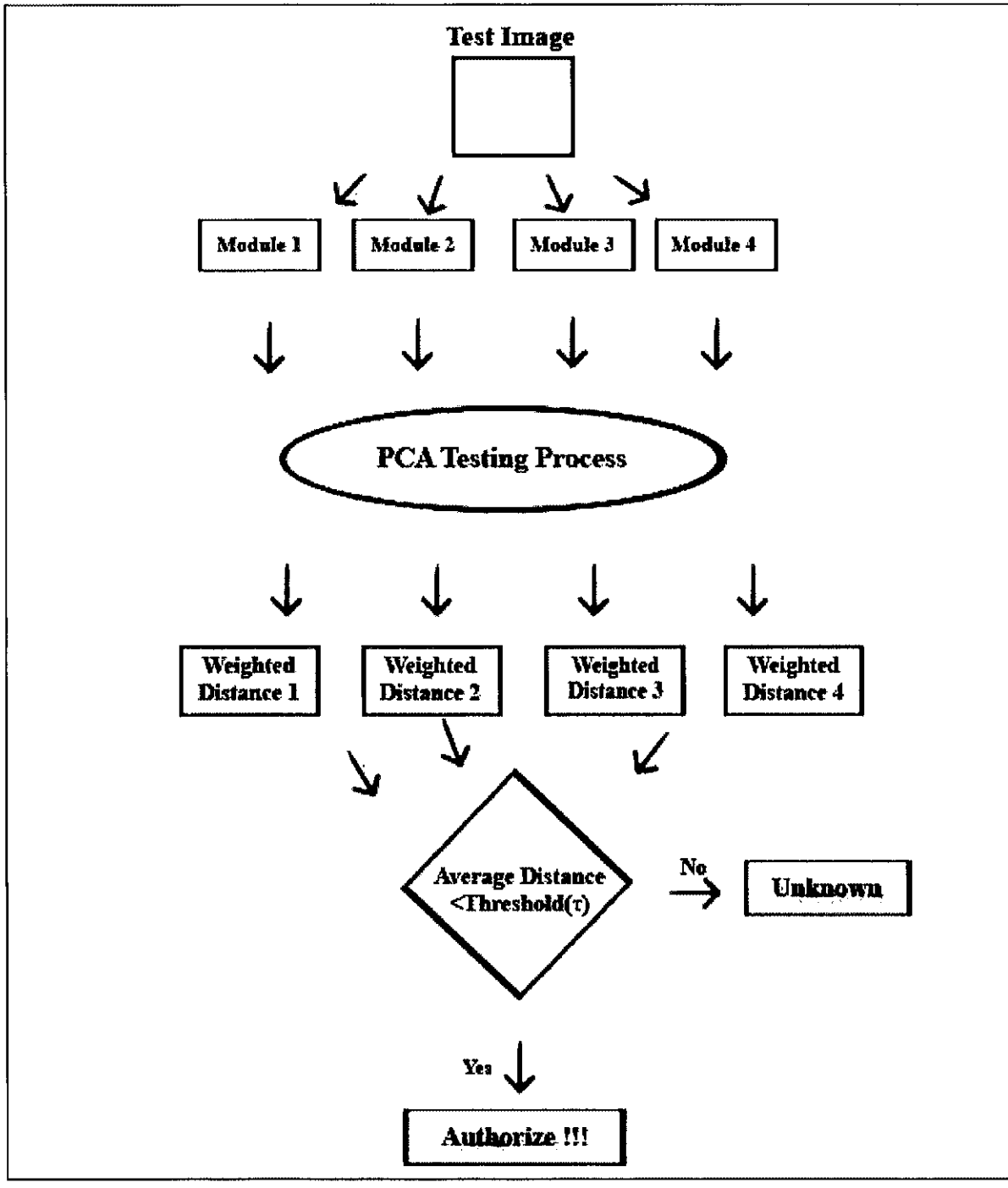

Figure 5.3: Illustration of the testing phase for weighted modular PCA-based face authentication

This final distance should be less than or equal to the pre-selected threshold τ_l so that the subject is authenticated, where τ_l is different for every individual.

The Mahalanobis distance measurement is obtained by,

$$Dist_g = (T_{ltestg} - T_{lg}^{Mean})^T T_{lg}^{Cov^{-1}} (T_{ltestg} - T_{lg}^{Mean}) \tag{5-10}$$

where T_{lg}^{Mean} and T_{lg}^{Cov} are the mean and the covariance matrices of all the corresponding sub images of subject l that are determined using the weight vector T_{lmg}, which is calculated during the training phase.

$$T_{lg}^{Mean} = \frac{1}{K}\sum_{m=1}^{K} T_{lmg} \quad \forall g, l \tag{5-11}$$

$$T_{lg}^{Cov} = \frac{1}{K}\sum_{m=1}^{K} (T_{lmg} - T_{lg}^{Mean})(T_{lmg} - T_{lg}^{Mean})^T \quad \forall g, l \tag{5-12}$$

The classification is done based on the distance equation (5-8) to calculate the sum of weighted Mahalanobis distances. If the distance $Dist_{test}$ is less than or equal to a predefined threshold τ_l, then the subject is said to be authenticated.

5.2 Experimental Results

The performance of face authentication using weighted modular PCA is evaluated with the same training database as discussed in chapter 3. The performance of weighted modular PCA for face authentication was tested by varying the number of eigenvectors. A greater number of eigenvectors resulted in an increased authentication rate, thereby decreasing the false acceptance rate. Tests were conducted to analyze the algorithm with

different videos of the individuals at different times by varying expressions and poses and occlusions. Sample test images are shown in figure 4.4 of chapter 4. The algorithm is executed and tested for 10 images of an individual that were grabbed from the video that is captured for around 3 seconds while the person enters his or her personal identification number to perform a transaction. Out of these 10 face images, if at least two of the test images match, then the person is allowed to perform the transaction; otherwise the person is not allowed to continue. It can be observed from figure 5.4 that the authentication rate is increased with the increase in the number of eigenvectors. We can observe that for the number of eigenvectors equal to 20, the authentication rate is 96 percent.

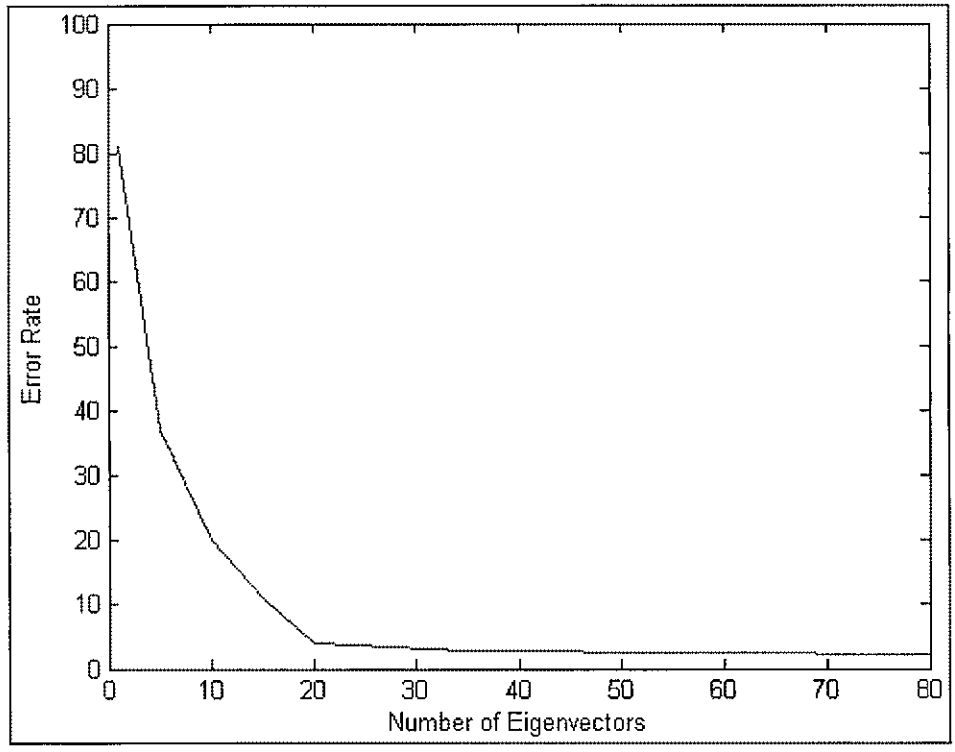

Figure 5.4: Error characteristics with respect to number of eigenvectors using weighted modular PCA (for images with varying expressions, poses and occlusions)

The algorithm performs well for test images with varying expressions and poses and with occlusions; the authentication rate is around 98 with 25 eigenvectors and there is not much improvement for the number of eigenvectors greater than 25.

From figure 5.5, we can observe that for a particular false acceptance rate of 2 percent the authentication rate achieved is 100 percent. Therefore the weighted modular PCA-based face authentication gives accurate results by varying expressions and poses and also with occlusions.

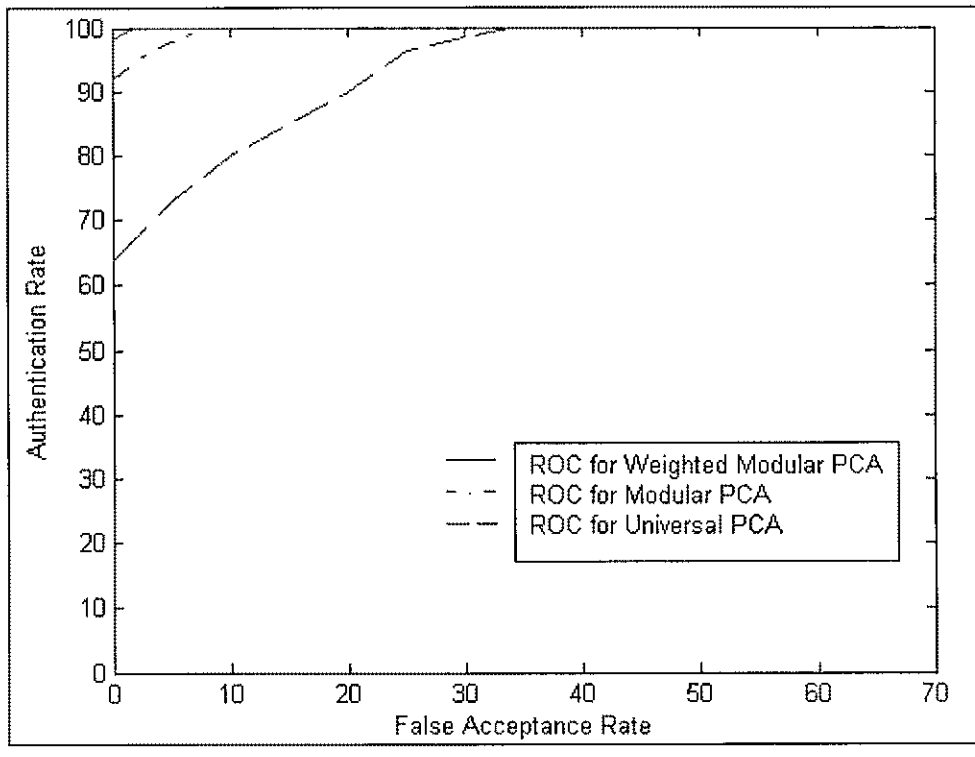

Figure 5.5: ROC curves of universal PCA, modular PCA and weighted modular PCA

5.4 Summary

In this chapter, the application of the weighted modular PCA technique for real-time face authentication has been explained. The performance evaluation for face authentication is discussed. Based on the experimental results it can be observed that with the weighted modular PCA-based approach for face authentication, the algorithm performs well with the problems of varying expressions and poses and with occlusions. The weight estimation based on the variances of individual modules and its impact on distance computations has improved face authentication accuracy. Therefore as this face authentication system performed well, it can be used in many applications such as in financial institutions, ATMs where money transactions are crucial and other secured areas.

Chapter 6

CONCLUSIONS AND FUTURE WORK

A weighted modular approach for face authentication has been presented in this thesis. The module weights were provided based on the priorities of different facial regions that change with the variations in pose and expressions and also with occlusions. This technique helped in authenticating the identity of an individual who claims to be one of the subjects in the database. A sequence of face images were selected from video frames at a particular interval and were used for the authentication purpose. These images will have wide variations in pose, expressions and occlusions when the individual is unaware of being authenticated.

Three different techniques were investigated, namely Principal Component Analysis (PCA), Modular PCA (MPCA) and Weighted Modular PCA (WMPCA). In the PCA technique, the main idea was to express all the two dimensional facial images in the training set in a compact set of principal components of the feature space called the eigenspace projections, where the whole face image was considered to be a single unit. In the MPCA technique, the face images were divided into square modules, and then the PCA approach was applied to each of the modules. Therefore only some of the sub images varied with the variations in facial features as it resulted in an effective representation of the local information of the face. In the WMPCA technique, the face images were divided into four horizontal sub regions, such as forehead, eyes, nose and mouth with chin regions, and then PCA was applied to these sub regions, which resulted in a comparison between the corresponding features of an individual. Weighted

comparison of respective feature regions of the test images with the training images provided a good authentication outcome.

The performance of the proposed face authentication system was evaluated with several individuals belonging to different ethnicities. It was observed that the weighted modular approach outperformed the state of the art face authentication techniques for the input images with varying poses and expressions and also with occlusions. The time taken for the computations was significantly reduced as the eigenspace was computed using the lower dimensions of the covariance matrix and later transformed into the higher dimensions. Two distance measures were used, the Euclidean distance and the Mahalanobis distance. From the results obtained for the two distance measures, it was observed that the Mahalanobis distance measure performed better than the Euclidean distance. The concept of intra subject variance was presented, which assigned high priorities to the facial features with low variance and less priority to the facial features with high variance. The final measure for the classification was based on the sum of the weighted distances using these variances. The final measure was compared to a pre-selected threshold, and the individual was authenticated accordingly. Different thresholds were set to different individuals in order to handle the problems faced by false authentications and false rejections. Comparison of the three techniques employed in this thesis was performed by the Receiver Operating Characteristics (ROC) curves. It can be observed that for a two percent false acceptance rate, the authentication rate achieved was 100 percent by using the WMPCA approach for face authentication.

It is observed that the changes of the representational components of different face feature regions are inconsistent with the variations in pose, expressions and with

occlusions. Hence it is envisaged that an adaptive computation of weights based on the magnitudes of specific feature variations could help in an improved authentication rate. Research work is also progressing to incorporate additional face feature components estimated from multiple modalities to provide better authentication performances.

REFERENCES

[1] M. D. Kelly, "Visual Identification of People by Computer," Technical Report AI-130, Stanford AI Project, Stanford, CA, 1970.

[2] T. Kanade, "Computer Recognition of Human Faces," Birkhauser, Basel, Switzerland, and Stuttgart, Germany, 1973.

[3] A. V. Nefian and M. H. Hayes III, "Hidden Markov Models for Face Recognition," In Proceedings of International Conference on Acoustics, Speech and Signal Processing, pp. 2721–2724, 1998.

[4] P. J. Philips, "Support Vector Machines applied to Face Recognition," Advanced Neural Information Processing System, vol. 11, pp. 803–809, 1998.

[5] 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.

[6] A. P. Kumar, S. Das and V. Kamakoti, "Face Recognition Using Weighted Modular PCA," In Proceedings of International Conference on Neuro-Information Processing, Lecture Notes in Computer Science, vol. 3316, pp. 362–36, 2004.

[7] International Biometric Group http://www.biometricgroup.com/, February 2008.

[8] M. Turk and A. Pentland, "Face Recognition using Eigenfaces," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591, June 1991.

[9] M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[10] M. Turk, "A Random Walk through Eigenspace," IEICE Transaction on Information & Systems, vol. E84-D, no. 12, pp. 1586-1595, December 2001.

[11] E. Oja, "Subspace Methods of Pattern Recognition," Electronic & Electrical Engineering Research Studies, 1983.

[12] W. Zhao, R. Chellappa and A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," In Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, pp. 336-341, April 1998.

[13] J.R. Beveridge, K. She, B. Draper and G.H. Givens, "A Nonparametric Statistical Comparison of Principal Component and Linear Discriminant Subspaces for Face Recognition," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Kaui, HI, USA, pp. 535- 542, December 2001.

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

[15] B. Draper, K. Baek, M.S. Bartlett and J.R. Beveridge, "Recognizing Faces with PCA and ICA," Computer Vision and Image Understanding (Special Issue on Face Recognition), vol. 91, no. 1, pp. 115-137, July 2003.

[16] M.S. Bartlett, J.R. Movellan and T.J. Sejnowski, "Face Recognition by Independent Component Analysis," IEEE Transactions on Neural Networks, vol. 13, no. 6, pp. 1450-1464, November 2002.

[17] K. Baek, B. Draper, J. R. Beveridge and K. She, "PCA vs. ICA: A Comparison on the FERET Data Set," In Proceedings of the Fourth International Conference on Computer Vision, Pattern Recognition and Image Processing, Durham, NC, USA, pp. 824-827, March 2002.

[18] P. Belhumeur, J. Hespanha and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection," In Proceedings of the Fourth European Conference on Computer Vision, Cambridge, UK, vol. 1, pp. 45-58, April 1996.

[19] B. Moghaddam, and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Transactions on Pattern Analysis on Machine Intelligence, vol. 19, pp. 696–710, 1997.

[20] X. Liu, T. Chen and B.V.K.V. Kumar, "Face Authentication for Multiple Subjects using Eigenflow," In Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 1-20, May 2002.

[21] C.L. Fennema and W.B. Thompson, "Velocity determination in Scenes containing several moving objects," Computer Graphics and Image Processing, vol. 9, pp. 301-315, 1979.

[22] B.D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," In Proceedings DARPA IU Workshop, pp. 121-130.

[23] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, Second edition, John Wiley & Sons. Inc., New York, 2001.

[24] A. Hyvärinen, J. Karhunen and E. Oja, Independent Component Analysis, Wiley, New York, 2001.

[25] C. Havran, L. Hupet, J. Czyz, J. Lee, L. Vandendorpe, and M. Verleysen, "Independent component analysis for face authentication," In KES 2002 proceedings of Knowledge-Based Intelligent Information and Engineering Systems, Crema (Italy), pp. 1207-1211, 2002.

[26] I. Guyon and D. Stork, "Linear discriminant and support vector classifiers," A.J. Smola, P. Bartlett, B. Scholkopf, and C. Schuurmans, editors, Advances in Large Margin Classifiers, MIT Press, Cambridge, MA, pp. 147-169, 1999.

[27] B. Moghaddam, W. Wahid, and A. Pentland, "Beyond eigenfaces: Probabilistic matching for face recognition," Automatic Face and Gesture Recognition, Nara, Japan, pp. 30–35, April 1998.

[28] K. Jonsson, J. Kittler, Y. P. Li, and J. Matas, "Support vector machines for face authentication," In Proceedings of British Machine Vision Conference, Nottingham, UK, pp. 543–553, 1999.

[29] O. Deniz, M.Castrillon and M. Hernandez, "Face recognition using independent component analysis and support vector machines," Pattern Recognition Letters, vol. 24, pp. 2153-2157, 2004.

[30] G. Heusch, Y. Rodriguez, and S. Marcel, "Local binary patterns as an image preprocessing for face authentication," IDIAP-RR 76, IDIAP, pp. 6, 2005.

[31] G. Zhang, X. Huang, S. Li, Y. Wang, and X. Wu, "Boosting local binary pattern (lbp)-based face recognition," Advances in Biometric Person Authentication, Springer Verlag, L. 3338, editor, pp. 179–186, 2004.

[32] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[33] Y. Lee, Y. Lee, Y. Chung and K. Moon, "One-Time Templates for Face Authentication," In Proceedings of IEEE international Conference on Convergence Information Technology, pp. 1818-1823, November 2007.

[34] Y. Lee and I. Verbauwhede, "Secure and low-cost RFID authentication protocols," In Proceedings of the 2nd IEEE International Workshop on Adaptive Wireless Networks (AWiN), pp. 1-5, November 2005.

[35] A. Razdan and M. Bae, "Curvature Estimation Scheme for Triangle Meshes Using Biquadratic Bezier Patches," Computer Aided Design, vol. 37, no.14, pp. 1481-1491, December 2005.

[36] A. E. Johnson and M. Hebert, "Using spin-images for efficient multiple model recognition in cluttered 3-D scenes," IEEE PAMI, vol. 21, no. 5, pp. 433-449, 1999.

[37] L. Zhang, A. Razdan, G. Farin, J. Femiani, M. Bae and C. Lockwood, "3D Face Authentication and Recognition Based on Bilateral Symmetry Analysis," The Visual Computer, vol. 22, no. 1, pp. 43-55, 2006.

[38] P. Besl and R. Jain, "Segmentation through variable order surface fitting," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 2, pp. 167-192, 1988.

[39] George H. Dunteman, "Principal Component Analysis," Sage University Paper #69, Newbury Park, CA: Sage Publications, Inc., 1989.

[40] K. Rao, N. Ahmed, "Orthogonal transforms digital signal processing", IEEE transaction on ICASSP Acoustics, Speech and Signal Processing, vol. 1, pp. 136-140, April 1976.

[41] P. Viola and M. Jones, "Robust real time face detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2004.

[42] A. Grosso, E. Tistarelli, M. Bicego, and M. Lagorio, "On the use of sift features for face authentication," Computer Vision and Pattern Recognition Workshop, New York, pp. 35, 2006.

[43] K. Fukunaga, Introduction to Statistical Pattern Recognition, second edition, Academic Press, 1990.

Chandrika Tummala
Department of Electrical and Computer Engineering,
Old Dominion University,
571-230-8615.
ctumm001@odu.edu

Education

Old Dominion University, Norfolk, Virginia
Master of Science in Electrical Engineering, **May 2008**
Current GPA: 3.57/4.0
Jawaharlal Nehru Technological University, India
Bachelor of Technology in Electronics and Communications Engineering, **April 2005**
GPA: 3.43/4.0

Computer Skills

Languages	C, C++, Matlab, HTML, SQL, Oracle 9i, Java
Electronic Simulations	PSpice, VHDL, MultiSim
Web Technologies	JSP, HTML, UNIX, Adobe Photoshop, Micro-media Flash
Documentation	MS Office

Relevant Coursework

Digital Signal Processing, Probability and Stochastic Processes, Engineering Systems and Modeling, Linear Systems, Digital Image Processing

Presentations

Presented a project titled "Smart SMS using Microcontroller 8051" National Level Symposium at Osmania University, Andhra Pradesh, India

February 12, 2005

Awards

Won **2nd Prize** at National Level Symposium, Osmania University, India

March 2005