Face Recognition Based on Local Zernike Moments

Mostafa Malekan

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the Degree of

Master of Science
in
Electrical and Electronic Engineering

Eastern Mediterranean University
June, 2015
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Serhan Çiftçioğlu
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Chair, Department of Electrical and
Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Erhan A. İnce
Supervisor

Examining Committee

1. Prof. Dr. Hasan Demirel

2. Prof. Dr. Hüseyin Özkaramanlı

3. Assoc. Prof. Dr. Erhan A. İnce

ABSTRACT

The thesis combines Principal Component Analysis (PCA) (a dimensionality reduction tool), Local Zernike Moments (LZM) (a filtering method), and some image processing techniques to design a face recognition system based on minimum distance criteria. During simulations, training images were first partitioned and then passed through an LZM transformation to compute moments at each pixel by considering local neighborhood of the pixels in each sub-image. Repeating this calculation for different moment components a set of complex moment images were obtained. Then for each moment image phase and magnitude histograms (PMHs) were extracted. To cut down on processing time the extracted histograms were then concatenated and PCA was applied to extract a reduced dimensionality set that represents the data well.

The thesis presents simulation results that are two-fold. The first set of results are achieved by computing moment images that are obtained by partitioning the input intensity image, normalizing the sub-regions, passing them through an LZM filter and then extracting and concatenating the phase and magnitude histograms (PMH) of each smaller moment-image to create a final feature vector that can be used for recognition purposes. The second set of results are obtained by performing almost the same steps but after the LZM filtering the reduced size moment-images are passed through a 2D Gaussian weighting operation. The Gaussian weighting kernel used was the same size as the sub-region and the standard deviation was taken as 8. The histograms were created by multiplying the magnitude of each pixel with the corresponding kernel weight before histogram binning.

In both simulations since the final feature vectors obtained after concatenating the PMHs of all the sub-moment-images are long (high computational complexity) we have used PCA to reduce the size of the feature vectors. Finally our system was tested using some probe sets from the 'The Face Recognition Technology' (FERET) image database and rank1 to rank5 matches were found based on minimum distance calculations.

The recognition accuracy for our system was tested using four different probe sets (FaFb, FaFc, Dup-I and Dup-II) from FERET image database. While using reduced dimensionality feature vector based on PMHs the accuracy of recognition for the probe sets FaFb, FaFc, Dup-I and Dup-II were respectively 94%, 86%, 73%, and 70%. When a 2D Gaussian kernel is applied on the magnitude of the moment images obtained from the LZM filter the corresponding accuracies for the same probe sets became 97%, 95%, 78%, and 76%. This indicates that filtering the moment-images using a 2D weighting function helps improve the results on average by 5.75 %.

Finally we have compared the accuracy of our proposed system with face recognition systems using Local Binary Patterns (LBP) and Histograms Local Zernike Moments (H-LZM). Results indicate that our H-LZM-PCA methods would provide higher recognition rates for all probe sets when compared with the LBP and while using reduced dimensionality feature vector by PCA algorithm, features vectors were limited to facial features by removing other additional information and since the dimensions are reduced, computation time reduced when compared with H-LZM.

Keywords: Face Recognition; Local Zernike Moments; Principle Component Analysis; FERET Database; Local Binary Patterns;

ÖZ

Bu tez çalışmasında, en küçük mesafe kriterine dayalı bir yüz tanıma sistemi tasarlanmak üzere, Temel Bileşen Analizi (TBA) (bir boyutluluk indirgeme aracı), Yerel Zernike Momentleri (YZM) (bir filtreleme yöntemi) ve bazı görüntü işleme teknikleri birleştirilmiştir. Simülasyonlar boyunca, ilk olarak eğitme görüntüleri bölümlenmiş olup daha sonra ise her bir alt-görüntüdeki piksellerin yerel komşuluğu dikkate alınarak her bir pikseldeki momentler hesaplanmak üzere bir YZM dönüşümünden geçirilmiştir. Bu hesaplama işlemi farklı moment bileşenleri için tekrarlanarak bir karmaşık moment görüntü seti elde edilmiştir. Daha sonra her bir moment görüntüsü için faz ve büyüklük histogramları (FBH) çıkarılmıştır. Daha sonra işlem süresinin azaltılması amacı ile çıkarılan histogramlar birleştirilmiş olup verileri uygun bir şekilde temsil eden boyutluluğu indirgenmiş bir setin çıkarılması amacı ile TBA uygulanmıştır.

Bu tez çalışması iki-kat olan simülasyon sonuçlarını sunmaktadır. Sonuçların ilk seti, giriş yoğunluk görüntüsünün bölümlenmesinden elde edilen moment görüntülerinin hesaplanması, alt-bölgelerin normalize edilerek bir YZM dönüşümünden geçirilmesi ve daha sonra ise tanıma amaçları için yararlanılabilecek nihai bir özellik vektörünün oluşturulması için daha küçük olan her bir moment-görüntü için faz ve büyüklük histogramlarının (FBH) çıkarılması ve birleştirilmesinden elde edilmiştir. Sonuçların ikinci seti ise neredeyse aynı adımların uygulanması sonucunda elde edilmiş olup ancak YZM filtreleme işleminden sonra boyutları indirgenmiş olan moment-görüntüleri iki boyutlu bir Gauss filtresinden geçirilmiştir. Kullanılan Gauss ağırlıklandırma çekirdeği alt-bölge ile aynı boyuta sahip olup standart sapma ise 8

olarak dikkate alınmıştır. Histogramlar, her bir pikselin büyüklüğünün histogram gruplandırmasından önceki ilgili çekirdek ağırlığı ile çarpması sonucunda oluşturulmuştur.

Her iki simülasyonda da tüm alt-moment-görüntülerin FBH'larının birleştirilmesi sonucunda elde edilen nihai özellik vektörlerinin uzun olması nedeniyle (yüksek hesaplama karmaşıklığı) özellik vektörlerinin boyutlarının indirgenmesi amacı ile TBA kullanılmıştır. Son olarak sistemimiz "Yüz Tanıma Teknolojisi" (YTT) görüntü veri tabanından bazı sonda setleri kullanılarak test edilmiş olup 1.Derece ile 5.Derece arası dereceler, en küçük mesafe hesaplamalarına dayanılarak elde edilmiştir.

Sistemimiz için tanıma doğruluğu, YTT görüntü veri tabanından dört farklı sonda seti (FAFB, FAFC, Dup-I ve Dup-II) kullanılarak test edilmiştir. FBH'a dayalı boyutluluğu indirgenmiş özellik vektörü kullanıldığında FaFb, FaFc, Dup-I ve Dup-II sonda setleri için tanıma doğruluğu sırasıyla %94, %86, %73 ve %70 olarak elde edilmiştir. YZM filtresinden elde edilen moment görüntülerinin büyüklüğü üzerinde iki boyutlu bir Gauss çekirdeği uygulandığında ise aynı sonda seti için ilgili doğruluk oranları sırasıyla %97, %95, %78 ve %76 olmuştur. Bu sonuçlar iki boyutlu ağırlıklandırma fonksiyonu kullanılarak moment-görüntülerin filtrelenmesinin sonuçların ortalama %5.75 oranında iyileştirilmesine yardımcı olduğunu göstermektedir.

Son olarak Yerel İkili Örnekler (YİÖ) ve Yerel Zernike Moment Histogramları (H-YZM) kullanılarak önerilen sistemin doğruluğu, yüz tanıma sistemleri ile karşılaştırılmıştır. Elde edilen sonuçlar, önerilen H-YZM-TBA yöntemlerinin YİÖ ile karşılaştıtrıldığında tüm sonda setleri için daha yüksek tanıma oranları sunacağını

göstermektedir ve TBA algoritması tarafından azaltılmış boyut özellik vektörü kullanırken, özellik vektörlari diğer ek bilgiler kaldırarak yüz özellikleri sınırlandırıdı ve boyutları azaltıldi bu yana H-YZM ile karşılaştırıldığında hesaplama süresi azaltılmış.

Anahtar Kelimeler: Yüz Tanıma; Yerel Zernike Momentleri; Temel Bileşen Analizi; YTT Veri Tabanı; Yerel İkili Örnekler;

ACKNOWLEDGMENT

First, I would like to thank my supervisor Assoc. Prof. Dr. Erhan A. İnce for guiding and helping me during my master's study. He has always been patient and has kindly shared his knowledge with me.

I also wish to thank all the faculty members at the department of Electrical and Electronic Engineering.

My sincere thanks go to my best friends Dr. Kiyan Parham, Mr Sina Ghasempour, who have always been supportive of me.

Last but not least, I would like to express my appreciation to my parents who have always given me support and guided me in the right direction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

μ	Mean Value
σ	Overall Variance
v_{nm}^{k}	Moment Based Operators
R_{mn}	Radial Polynomials
G	Kernel Matrix
σ	Standard Deviation Of The Distribution
P	Ratio Of The Polar Coordinate
θ	Angle Of The Polar Coordinate
EBGM	Elastic Bunch Graph Matching
FERET	Face Recognition Technology
GZM	Global Zernike Moments
GPZMA	Geodesic Pseudo Zernike Moment Array
H-P-LZM	Histogram-PCA-LZM
LZM	Local Zernike Moments
LDA	Linear Discriminant Analysis
LTP	Local Ternary Pattern
LBP	Local Binary Pattern
PCA	Principal Component Analysis
UDWT	Un-decimated Discrete Wavelet Transform
ZM	Zernike Moment

Chapter 1

INTRODUCTION

Face recognition is a task that humans perform routinely and effortlessly in their daily lives. The subject of face recognition or more generally pattern recognition is a widely studied subject within digital image processing both because of the practical importance of the topic and the theoretical interest from cognitive scientists. Despite the fact that other methods of biometric identification such as fingerprint or iris recognition can be more accurate, face recognition has always remained a major focus of research because of its *non-invasive* nature. The performance of face recognition systems has improved significantly since the introduction of the first automatic face recognition system that was developed by T. Kanade [1]. Moreover, face detection, facial feature extraction, and recognition can now be performed in 'real-time' for images captured under favorable situations. Although progress in face recognition has been encouraging, the task has also turned out to be a difficult endeavor, especially for unconstrained tasks where viewpoint, illumination, expression and occlusions are considered.

Research in face recognition is motivated not only by the fundamental challenges that the recognition problem poses but also by numerous practical application where human identification is needed. Face recognition, as one of the primary biometric technologies, became more and more important owing to rapid advances in technologies such as digital cameras, the internet and mobile devices, and increased

demands on security. Face identification can be classified in three main groups. Identification of individual faces based on facial features is investigated in the first and leading group [2] [3] [4]. In these techniques, the estimation of dynamic facial muscle contractions of human faces are considered. Feature vectors taken out from profile silhouettes forms the basis for the second group [5] [6]. In the last group, feature vectors extracted from frontal views of people are employed [7]. While extracting feature vectors it is possible to follow two different approaches; 1) make use of the entire picture (global image processing) and 2) work on fundamental elements of the face [8], (local region(s) processing) and combine feature points extracted therein. Widely accepted fundamental elements of a face image include: chin, eyes, mouth and nose.

1.1 Applications of Face Recognition

In [9], the main applications of face recognition were categorized by Jafri and Arabnia as follows:

- Verification: One-to-one verification which covers the case where the identity of a single individual has to be verified.
- Identification: One-to-many identification which is applied when an image is given, and the identity of the individual is to be determined from an available database by comparing to with all available image.

Face recognition techniques can be employed for various practical purposes. Some of these practical applications are as listed below:

- Identification in order to grant access to buildings, ATM machines and for airport security [10],
- Application in smart cards,

- Indexing of videos [11],
- Searching images within a database (for example to locate a missing child),
- Justice systems' application in criminal cases,
- Gender recognition [12] [13],
- Determination of facial expressions [14] [15].

1.2 Proposed Work and Objectives

Generally, identification through face recognition is used to grant access to building and restricted zones. Therefore, accuracy of recognition is an important issue. To improve face recognition rates various studies in the literature have proposed alternative solutions. With the same goal in mind, in this thesis we propose to use local Zernike Moments based filtering on sub-blocks of the image and create feature vectors by extracting and concatenating phase-magnitude histograms for each sub-block from each moment image. To reduce the computational complexity the thesis proposes to make use of Karhunen-Loeve transformation. Our system will be tested using some probe sets from the 'The Face Recognition Technology (FERET)' image database and matches will be found based on minimum distance calculations. A block diagram, that depicts the general processing steps for the proposed method, is provided in Figure 1.1.

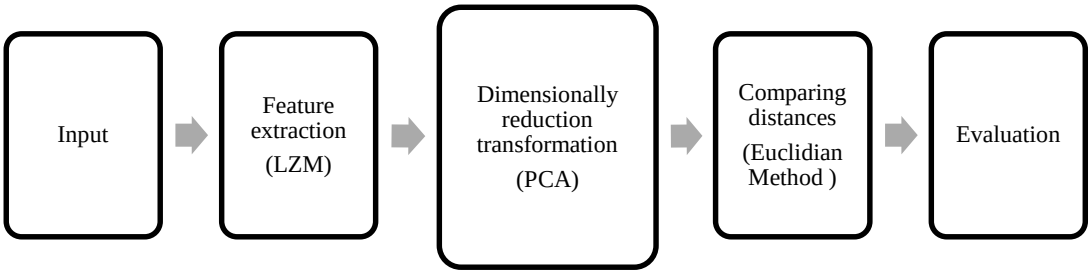


Figure 1.1: Block diagram of our proposed system

1.3 Literature Survey

Over the last two decades, face recognition has become an active area of research in computer vision, neuroscience, and psychology. Progress has advanced to the point that face recognition systems will soon be demonstrated in real-world settings. The rapid development of face recognition is due to a combination of factors: active development of algorithms, the availability of a large database of facial images and the availability of procedure for calculating [1]. A significant amount of research have been presented in the literature to take further the ongoing research in this field aims to improve the performance of this system against these factors.

In [16] Farokhi et al. proposed an innovative algorithm for face recognition that makes use of global and local characteristics. The method proposed by Farokhi combines un-decimated discrete wavelet transform (UDWT) and Zernike moments for segregating images in local and global features. It was proven that the investigated method had some advantages in comparison to the other common methods. One of the major issues in face recognition subjects deals with the fact that a single set of features is not enough for producing satisfactory results. In [17] and [18], Singh utilized the combination of two advantageous methods namely Zernike moments (ZMs) and local binary pattern (LBP) /local ternary pattern (LTP) in face recognition. The results obtained showed that the recognition rates were 10-30% improved with this new approach. In [19], Hajati proposed a novel method based on the Geodesic Pseudo Zernike Moment Array (GPZMA) for 3D face recognition. A similar study that was conducted by Sariyanidi was also presented in [20]. In [21], Matthew and Pentland introduce a face recognition method that encodes only the most relevant information in a set of face images. In the literature this method is

known as Principal Component Analysis (PCA) where the set of face images are transformed into a set of characteristic feature images called “eigenfaces”. Here the recognition would be performed by first projecting each query image into the sub-space spanned by the eigenfaces and then classifying the face by comparing its position in the face space with positions of known individuals.

1.4 Thesis Organization

Chapter 1 provides a general overview for face recognition systems and also summarizes the state of the art. Chapter 2 provides descriptions for feature vector generation and dimensionally reduction. In Chapter 3, details of the proposed system are provided. A summary of widely known FERET database is given in Chapter 4. Simulation results and discussions are provided in Chapter 5. Finally, conclusions are made and some suggestions for future work are provided in Chapter 6.

Chapter 2

FEATURE EXTRACTION AND DIMENTIONALLY REDUCTION METHODS

The objective of a face recognition system is to have the ability to either verify or identify facial images. Since faces in general have common structures, face recognition can be a challenging task. One approach for face recognition is through the use of facial features. In the literature various methods have been proposed for the detection of facial features. In this chapter, we give an overview of geometric methods for feature extraction and also briefly describe PCA and Linear Discriminant Analysis (LDA) as two alternative methods that can be used for dimensionally reduction. To extract facial features approaches such as Local Binary Pattern (LBP), Elastic Bunch Graph Matching (EBGM), Gabor filter and Global Zernike Moment (GZM) have been proposed. Regardless of which method is used once the facial features are extracted, reduced dimensionality feature vectors would be formed via the use of PCA or LDA. A block diagram of a general face recognition system has been provided in Figure 2.1. The diagram suggests that for feature extraction LBP, EBGM, GABOR and GZM can be chosen and for dimensionally reduction either PCA or LDA can be used.

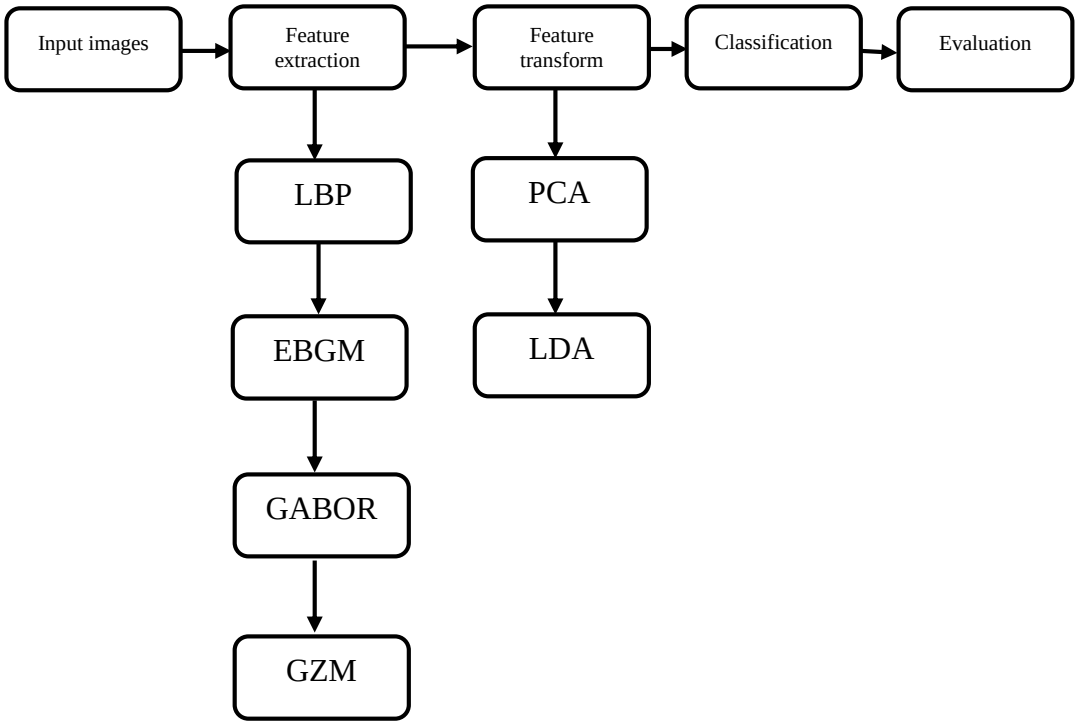


Figure 2.1: A block diagram of a general face recognition system with possible feature extraction methods and transformation methods.

2.1 Feature Extraction Methods

Feature extraction is a specific procedure of dimensionality reduction and denotes the acquirement of the image features from the image such as visual features, statistical pixel features, transform coefficient features, and algebraic features, with emphasis on the algebraic features, which represent the intrinsic attributes of an image. Face recognition represent to perform the classification to the above image features in terms of a certain criterion. Some used methods are described as follows.

2.1.1 Local Binary Pattern (LBP)

Local binary patterns (LBP) operator is one of the best performing texture descriptors and it has been widely used in various applications in computer vision. LBP is highly discriminative and has the advantage that it is invariant to monotonic gray-level changes. LBP is the particular case of the texture spectrum model proposed in 1990 [22]. LBP operator labels the pixels of an image by thresholding

the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models [22].

2.1.2 Elastic Bunch Graph Matching (EBGM)

Likewise to the previous method, EBGM is a data-extraction method wherein facial features are extracted from input facial graphs. The local features of a facial landmark are represented by a jet, where a jet is a set of Gabor wavelet features. A face bunch graph is created as a generalized representation of faces of various individuals, thus, it consists of a 'bunch' of jets corresponding to a landmark. To obtain the optimal face graph to represent a new face, a two-step approach is adopted [23].

2.1.3 Gabor Filter

Dennis Gabor [24] in 1946 presented a filter which is employed as a feature extraction tool. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave [25].

2.1.4 Global Zernike Moment (GZM)

In this method, the microstructure around any pixel is obtained by separating the image into moment components. The results obtained from by this method through utilizing FERET database did not perform adequately. But on the other hand it produced satisfactory results for recognizing alphabet letters and fingerprints [20].

2.2 Dimensionally Reduction Transformation Methods

When using a holistic approach for face recognition, an image of size ($n \times m$) is represented in an ($n \times m$) dimensional space. In practice this many spaces is too large to allow robust and fast face recognition. A common way to attempt to resolve this problem is to use dimensionality reduction techniques. Dimensionality reduction transformation helps reduce the number of random variables under consideration. It helps to only select the ones which best characterize the subject. The most widely used transformation methods are PCA and LDA. Following subsections provide a brief description for each method.

2.2.1 Principal Component Analysis

Principal Component Analysis is a linear transformation technique that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA can predict, remove redundancy, extract features, compress data etc. The main idea of using PCA for face recognition is to express the large 1-D vector of pixels constructed from 2-D facial image into the eigenspace projection. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The principal components are orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric [26] [27].

2.2.2 Linear Discriminant Analysis

Linear discriminant analysis is a generalization of Fisher's linear discriminant, a method used to find a linear combination of features that characterizes or separates

two or more classes of objects or events. LDA is also closely related to PCA and they both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data. PCA on the other hand does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables must be made [28].

Chapter 3

GENERATION OF REDUCED DIMENSIONALITY FEATURE VECTORS

In this chapter, we combine PCA (a dimensionality reduction tool), LZM (a filtering method), and some image processing techniques to design a face recognition system. During simulations, training images are first partitioned and then transformed with the help of an LZM filter. The LZM filter will help to compute moments at each pixel by considering local neighborhood of each pixel in each sub-block. Repeating this calculation for different moment components a set of complex moment images can then be obtained. Afterwards, for each moment image phase and magnitude histograms are extracted and a general feature vector is constructed via the concatenation of the extracted histograms. To cut down on processing time the proposed method would then utilize PCA to extract a reduced dimensionality set that best represents the data. Throughout the experiments, it was observed that when reduced size moment images were processed with a 2D Gaussian Kernel recognition rates will be improved by 3-5 percent. The Gaussian weighting function used was the same size as the sub-region. The histograms were created by multiplying the magnitude of each pixel with the corresponding kernel weight before histogram binning. Finally, Euclidian distance metrics were calculated to select five feature vectors (rank1-rank5 results in chapter 5) which are the closest to the feature vector of the query image. Figure 3.1 provides details of the processing steps for the proposed face recognition system using LZM-PCA.

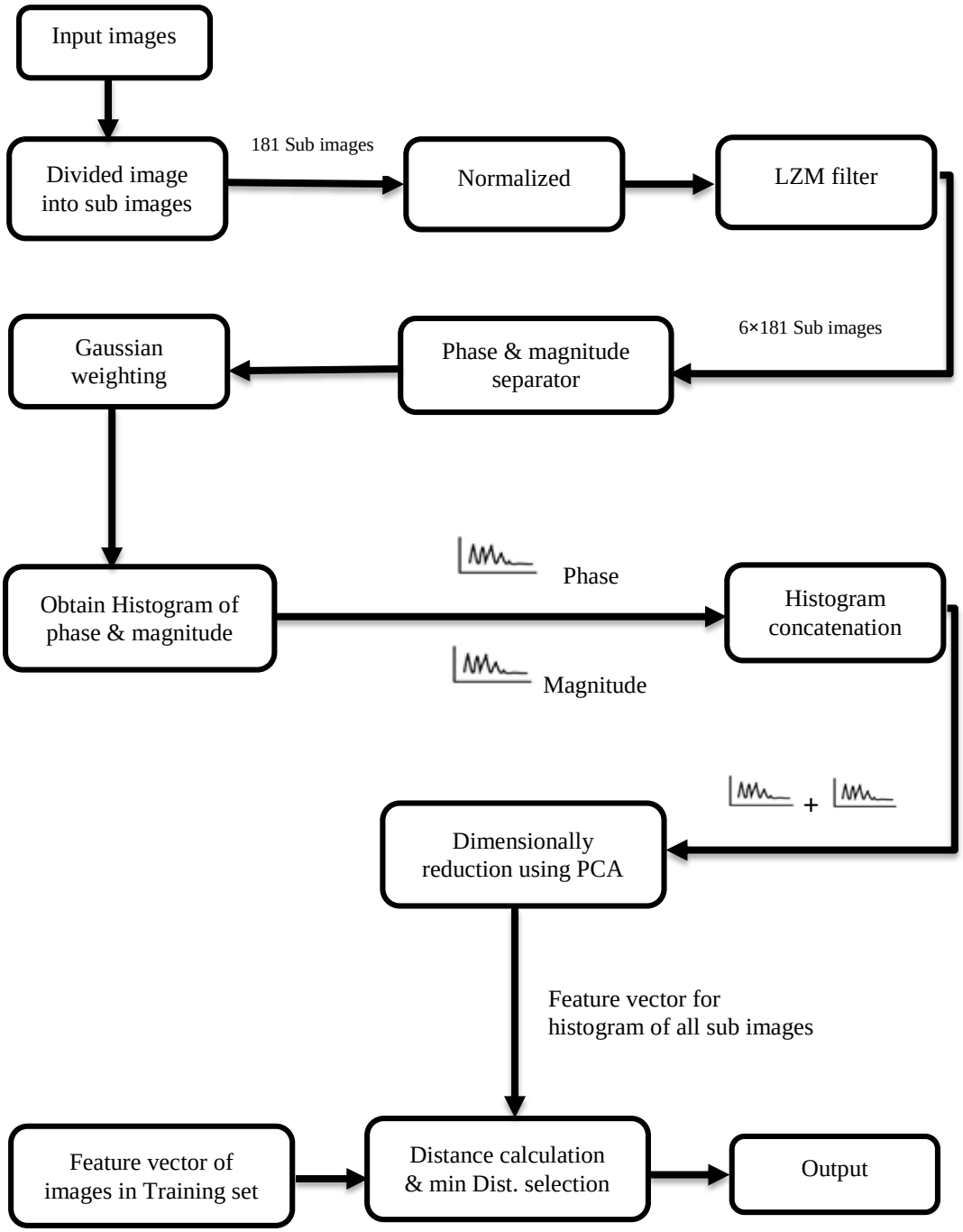


Figure 3.1: Block diagram for H-LZM-PCA processing

3.1 Breaking Input Images into Sub-Images

Since the face recognition methods must succeed in low resolution images with pose and illumination variations, most of the existing methods, try to solve the low-resolution face problem by considering these conditions. FERET database is separately divided into color and grayscale images consisting of low resolution images with pose and illumination variations. Since experiments previously carried out using FERET database had revealed the superior performance of LZM over other methods [29], in this study, we propose a method that uses LZM and Gaussian weighting that makes use of the FERET database. Since the authors of [30] had demonstrated that performance purely based on gray scale input images was quite satisfactory, in this study the color images in the target and probe sets were first transformed to grayscale using the RGB-to-Gray color conversion model.

As previously explained in Chapter 1, face recognition techniques can be broken down into two rough categories, [31]: (i) Global techniques and (ii) Local or component-based techniques. Global techniques use a single feature vector that represents the whole face image. Component-based recognition techniques are designed to compensate for pose changes by allowing a flexible geometrical relation between the components in the classification stage. Generally, a skin region analysis is carried out first, and then features like eyes, nose and mouth are located and cropped our as local regions. Afterwards template matching can be administered either to the local regions or to the corner points extracted from them. Figure 3.2 shows an eye, a nose and a mouth image extracted from the image on the left as described above.

Figure 3.2: Facial components

Note from Figure 3.3 that, since all facial components in the input image are not always at a fixed location and also the distance from the camera, as well as the photo angle play effective role, to attain an acceptable recognition rate scale and rotation invariant techniques become important. Also to deal with aging, the recognition system has to be well trained using sample images taken at different time intervals.

(a) (b) (c) (d)

Figure 3.3: Facial components with different distance and angel
(a) Frontal view, (b) Frontal view (with distance from the camera), (c) Profile view from quarter angle, (d) Profile view from half angle

In this thesis, instead of locating features in the face region, we subdivide the input image into sub-regions and extract and use PMHs to create our feature vector. Due to the fact that some of these sub-regions contain facial components when we concatenate their PMHs into our feature vector we are in essence doing the same

thing as component-based techniques. However, a problem with local histograms is that in the presence of small geometric variations, features along the borders may fall out of the local histogram. To deal with this, we down-weighted the features along the borders by applying a Gaussian window peaked at the center of each sub-region similar strategies are employed in a number of histogram-based representations. To account for the down-weighted features, we apply a second (inner) partitioning, where a higher emphasis is placed on features down-weighted at the first (outer) partitioning. Single histogram suffers from losing structure information and the special structure is important in face recognition. Hence images are decomposed into non-overlapping sub-region from which local features are extracted. The proposed technique breaks down the input moment images into sub-regions in two stages.

First we divide the image to ($N \times N$) equal-sized blocks (outer partitioning) beginning from the top-left of the image. Then, we divide the image to ((N-1) × (N-1)) blocks of the same size as the previous ones with a grid shifted half a block size from top left (inner partitioning). In the first stage, the available pixels along the length is divided by N and denoted (*i)*. Similarly, the number of pixels along the width is divided by N and denoted (*j)*. Subsequently, intervals indexed by (*i)* and (*j)* are determined along the length and the width. In Figure 3.4, input image is broken into ($N \times N$) sub-images. Note that the dimensions of all the obtained sub-images are the same and after extensive experiments, the researchers found good performance when setting N=10 [20].

(a) (b)

Figure 3.4: Image outer Partitioning
(a) Original input image, (b) ($N \times N$) sub-images

In the second stage, we provide a ((*N-1)* × *(N-1*)) sub-image by eliminating $\left\lfloor \frac{i}{2} \right\rfloor$ from the first column, $\left\lceil \frac{i}{2} \right\rceil$ from the last column, $\left\lfloor \frac{j}{2} \right\rfloor$ from the first row and $\left\lceil \frac{j}{2} \right\rceil$ from the last row. Therefore, the remaining pixels along the length is divided by *N*-1 and denoted *(i').* Similarly, the remaining pixels along the width is divided by *N*-1 and denoted *(j').* Subsequently, intervals indexed by *(i')* and *(j')* are determined along the length and the width. Figure 3.5 (a) shows input image after eliminating $\left\lfloor \frac{i}{2} \right\rfloor$ from the columns and $\left\lfloor \frac{j}{2} \right\rfloor$ from the rows. In Figure 3.5 (b), image is broken into ((*N*-1) × (*N*-1)) sub-images.

(a) (b)

Figure 3.5: Image inner portioning
(a) Input image, (b) ((*N*-1)×(*N*-1)) sub-images

Since, we have $N^2 + (N-1)^2$sub-regions for each moment component (181 sub-region), in order to increase the robustness against illumination variations further, each sub-region is then normalized.

3.2 Normalizing Sub-Images

Based on the researches from literature [32], the Zernike moments (the special case of which is the LZM filter) produces desirable results only if its input is a normalized image. In other hand, to increase the robustness of recognition against illumination variations, each sub-image is normalized [20]. The scenario of normalization is performed as follows, firstly the mean value (μ) and overall variance (σ) of the sub-images are obtained (Note that n is dimensional of sub-images). Thus, the sub-images are normalized to have zero mean and unit variance after they are cropped, and their sizes are fixed to 130×150. Equation (3.1) depicts mean value and overall variance formula.

$$\mu = \frac{\sum_i \sum_j x_{ij}}{n} \quad , \sigma^2 = \frac{\sum_i \sum_j (x_{ij} - \mu)^2}{n-1} \tag{3.1}$$

Any pixel of the sub-image is substituted considering the equation (3.2), afterwards, the normalized sub-images are obtained.

$$a_{ij} = \frac{x_{ij} - \mu}{\sqrt{\sigma^2}} \tag{3.2}$$

3.3 LZM Filtering

Zernike moments are based on the calculation of the complex moment coefficients and are successful in character recognition of images that contain distinctive shape information like characters. However these holistic moments were seen inadequate for the face images and for this reason a novel face representation method called Local Zernike Moments (LZM) was proposed and is shown to be successful in face recognition [29]. The LZM method localizes the calculation of the moments around each pixel.

LZM transformation is not sensitive to rotation which can be proved by the equations in [32] and this transformation provided a rich image representation by successfully exposing the intensity variations around each pixel. The components of the representation (i.e. complex images corresponding to different moment orders) could be very robust to illumination variations [20]. A brief description of the LZM method is provided below in section 3.3.1.

3.3.1 LZM Filter Design

In this stage, first the V_{nm}^{k} (Moment based operator) is defined for being convolved on the normalized sub-images. Where V_{nm}^{k} express the $k \times k$ sized filter kernels of Zernike polynomial, n is the order of polynomial and m is number of iteration which are specific value within the Zernike polynomial satisfying the following conditions:

1. n-$|m|$ is even.
2. $n \geq |m|$.

When the value of *n* and *m* are increased, the observed frequencies are increased accordingly, but the filter is gotten more damage by noises. Researchers [20] proved that this method produced the best results when sized filter kernels of Zernike polynomial, *k* is 7, for the reason of the fact that, the corresponding value requires a relatively small memory space for execution and provides acceptable accuracy.

Earlier than calculating moment based operator, initially $V_{nm}^{k}(i,j)$ must be converted into its polar coordinate using $V_{nm}^{k}(i,j) = V_{nm}(\rho,\theta)$. This is because, the radial polynomial ($R_{nm}(\rho)$) is expressed in polar coordinates as denoted by equation (3.3) and (3.4).

$$V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm\theta} \tag{3.3}$$

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s \rho^{n-2s}(n-s)!}{s!\,(\frac{n+|m|}{2}-s)!\,(\frac{n-|m|}{2}-s)!} \tag{3.4}$$

The procedure starts by setting the origin at the center of the filter kernel Zernike polynomial which has (7×7) grid size as shown in Figure 3.6, therefore, the distances and angles of any component of the filter kernel Zernike polynomial from the origin are calculated for the aim of obtaining the corresponding ρ and θ values.

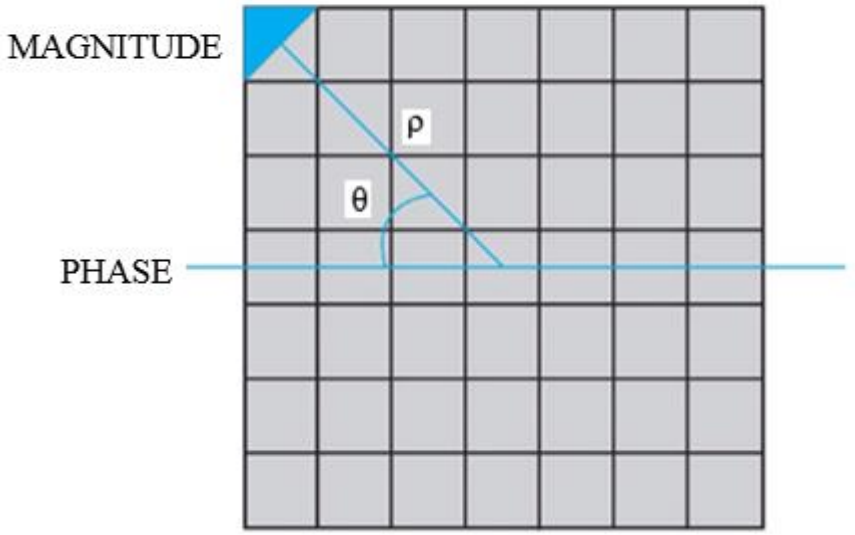


Figure 3.6: The (7×7) filter kernel Zernike polynomial

Referring to Figure 3.6 and equation (3.5), the system find out the polar coordinate of each filter kernel Zernike polynomial.

$$\rho = \sqrt{\Delta x^2 + \Delta y^2}\,, \quad \theta = tan^{-1}\frac{y}{x} \tag{3.5}$$

Where ρ is magnitude and θ is phase corresponding the polar coordinate and x and y represent the Cartesian coordinate. Thus, the system calculate $V_{nm}(\rho, \theta)$ for each filter kernel Zernike polynomial by using the equation (3.3).

In [20] it was shown via simulation that acceptable performance could be obtained when the moment order, *n*, is set to 4. The number of active moments, *K*, could then be calculated through equation (3.6) and would be equal to six. In this thesis the 6 moment images that were computed were $Z_{11}, Z_{22}, Z_{31}, Z_{33}, Z_{42} and\ Z_{44}$.

$$K(n) = \begin{cases} \dfrac{n(n+2)}{4} & \text{if } n \text{ is even} \\ \dfrac{(n+1)^2}{4} & \text{if } n \text{ is odd} \end{cases} \tag{3.6}$$

3.3.2 2D Convolution between LZM Filters and Sub-Images

For LZM processing we derive moment-based operators V_{nm}^{k}, which are (k×k) kernels calculated using the equation $V_{nm}^{k}(i,j) = V_{nm}(\rho,\theta)$. These operators are similar to 2D convolution kernels used for image filtering as depicted in Figure 3.7 (b). The LZM transformation can be defined by these kernels as equation (3.7).

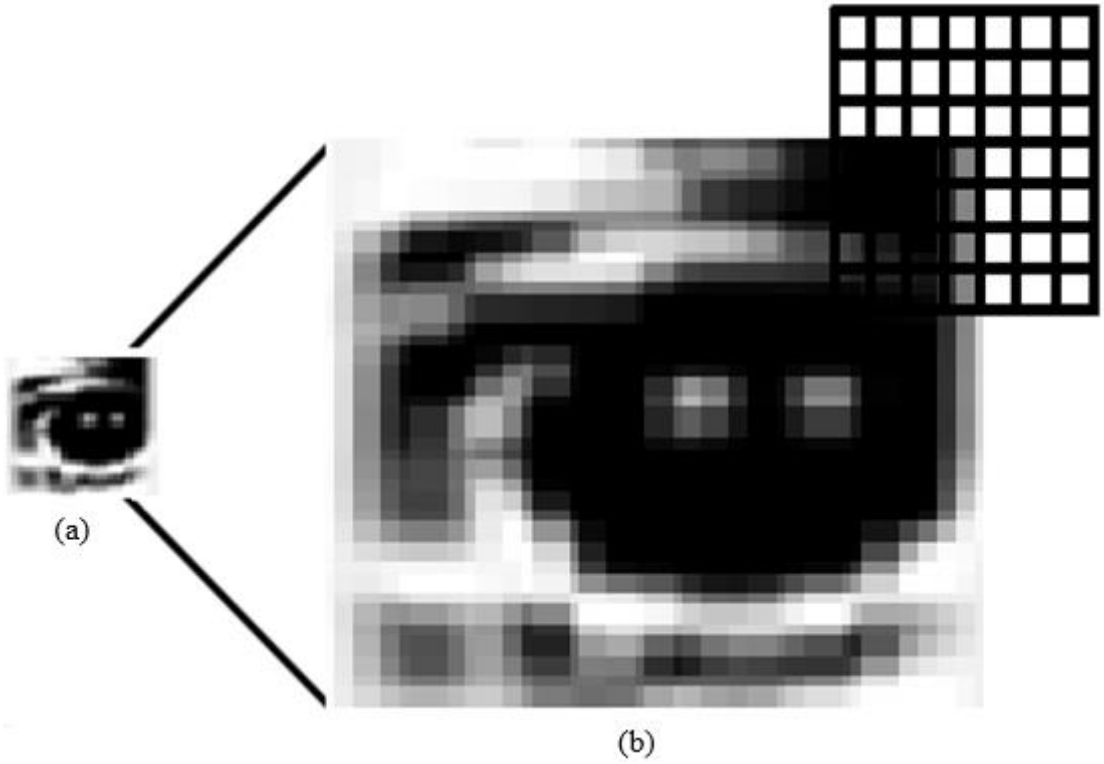


Figure 3.7: 2D convolution between LZM filter and sub-Image
(a) One sub-image, (b) LZM filter convolution

$$z_{mn}^{k}(i,j) = \sum_{p,q=-\frac{k-1}{2}}^{\frac{k-1}{2}} f(i-\rho, j-q) V_{nm}^{k}(\rho,q) \tag{3.7}$$

Finally, the filters would transform every 181 sub-images into 1086 filtered sub-images preparing for the next stage. The total number of filtered sub-images for the above stated conditions is calculated from the equation (3.8).

$$(N^2 + (N-1)^2) \times K = 1086 \quad (3.8)$$

3.4 Phase and Magnitude Separator

Since the six filter kernel convolved though sub-images involves Zernike polynomial expansions and these expansions included complex matrix. Thus, each pixel of the filtered sub-images is determined via its corresponding phase and magnitude. Herein, phase and magnitude of filtered sub-images are separated based on LZM transformation as shown in Figure 3.8.

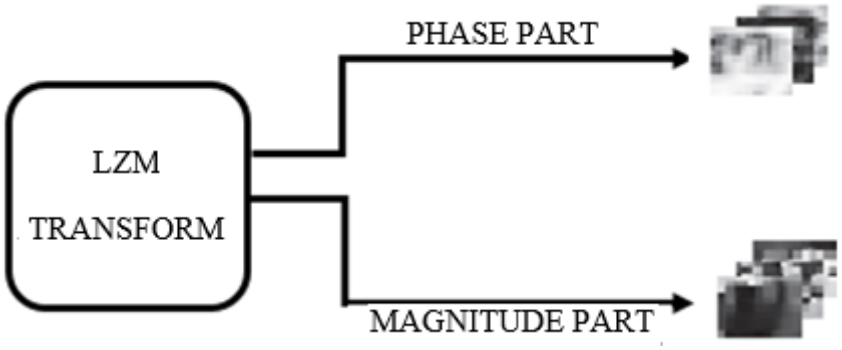


Figure 3.8: Phase and magnitude of moment sub-images

Figure 3.9 (a) represents one sub-image convolving six LZM filters so that their phase and magnitude are separated as depicted in Figure 3.9 (b) and (c).

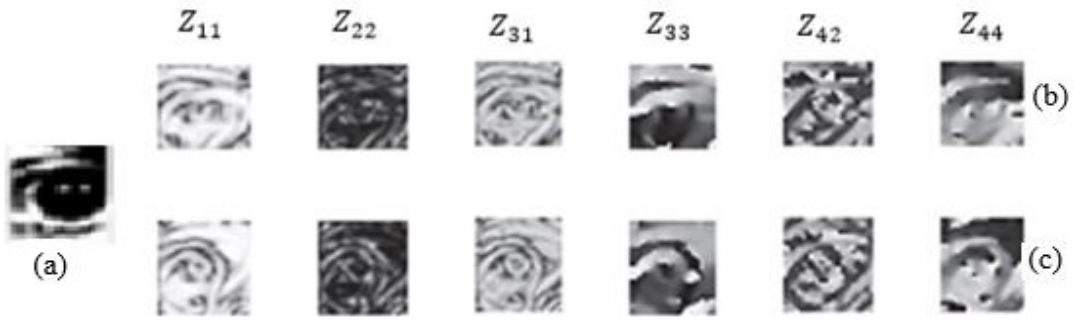


Figure 3.9: Phase and magnitude separator
(a) Sub-image, (b) Magnitude part, (c) Phase part

3.5 Gaussian Weighting

In this work Gaussian weighting implies multiplying the magnitude of each pixel of the sub-image by a corresponding weight as depicted in Figure 3.10. This step is executed before adding each pixel's magnitude to a relevant histogram bin [20].

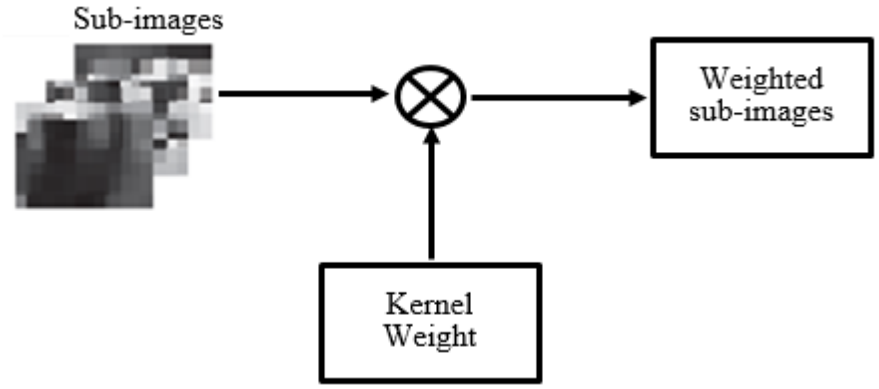


Figure 3.10: 2D Gaussian kernel

The elements of the (7×7) Gaussian mask are defined using (3.9) where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis and σ is the standard deviation of the distribution. The weights give higher significance to pixels near the edge and the main goal is to reduce edge blurring for each sub-image. The amount of smoothing is controlled by σ (larger σ for more intensive smoothing). Gaussian weighting multiplies the magnitude of the, pixels in the sub-image by weights taken from the kernel prior to binning the amplitudes. Here, weighting kernel has the same size as a sub-image. Standard deviation was taken as σ = 8 as suggested in [20]. Finally, in one clock cycle, the filter executes one multiplication operation and generates one smoothed pixel. Gaussian distribution curve with μ= (0, 0) and σ = 1 is depicted in Figure 3.11.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^{2+}y^2}{2\sigma^2}} \tag{3.9}$$

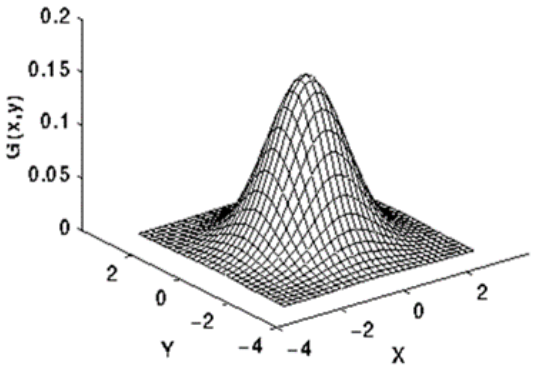


Figure 3.11: Gaussian distribution curve with μ=(0, 0) and σ = 1

3.6 Phase and Magnitude Histograms

A histogram is a graphical representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (quantitative variable) and was first introduced by Karl Pearson. To construct a histogram, the first step is to "bin" the range of values. After extensive experiments, researchers found that good performance is obtained when setting b = 24. Several types of histograms can be employed to utilize the output of LZM transformation, including magnitude histograms (MHs) and phase-magnitude histograms (PMHs). Experiments have led to conclude that PMHs perform substantially better than MHs [20]. By employing this method, 2127 histograms are constructed for each image which acquire as following equation (3.10):

$$(N^2 + (N-1)^2) \times K \times 2 = 2127 \tag{3.10}$$

Figure 3.12 depicts six phase and six magnitude histograms, one for each moment image obtained from one sub-block of the image.

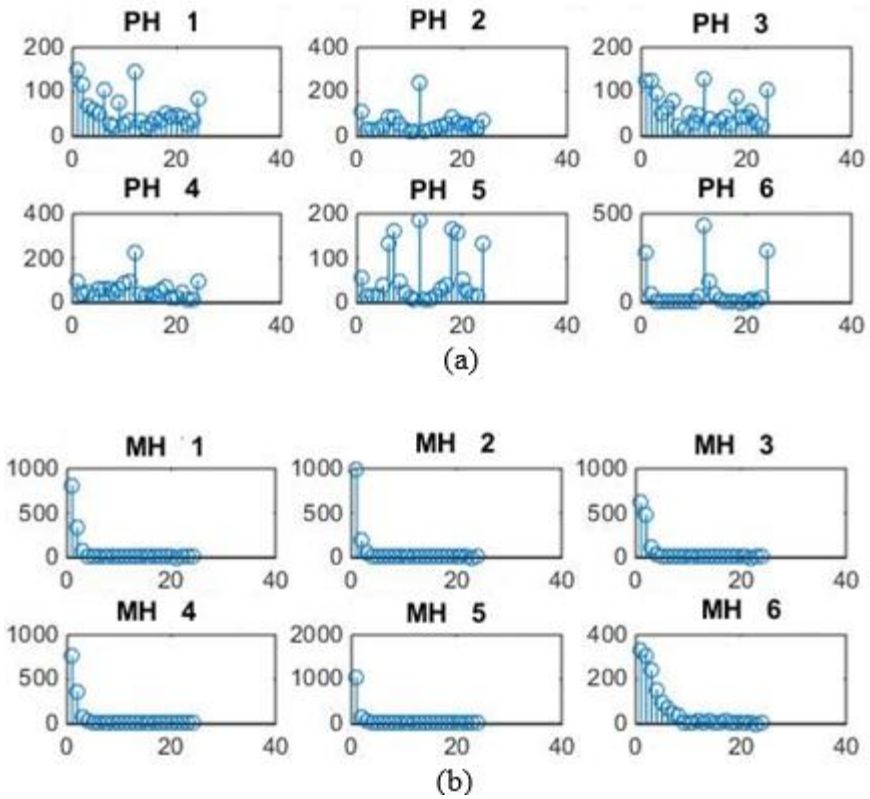


Figure 3.12: Phase and Magnitude Histograms for the six moment images obtained from one sub-block
(a) Phase histograms, (b) Magnitude histograms

3.7 Histogram Concatenation

Final feature vector is obtained by concatenating the extracted phase-magnitude histograms at each sub-image that is formed by dividing each moment image to non-overlapping sub-images. As shown in Figure 3.13, after phase-magnitude histograms are added together for each sub-image by this way, the histograms are concatenated to obtain the final feature vector of the input image.

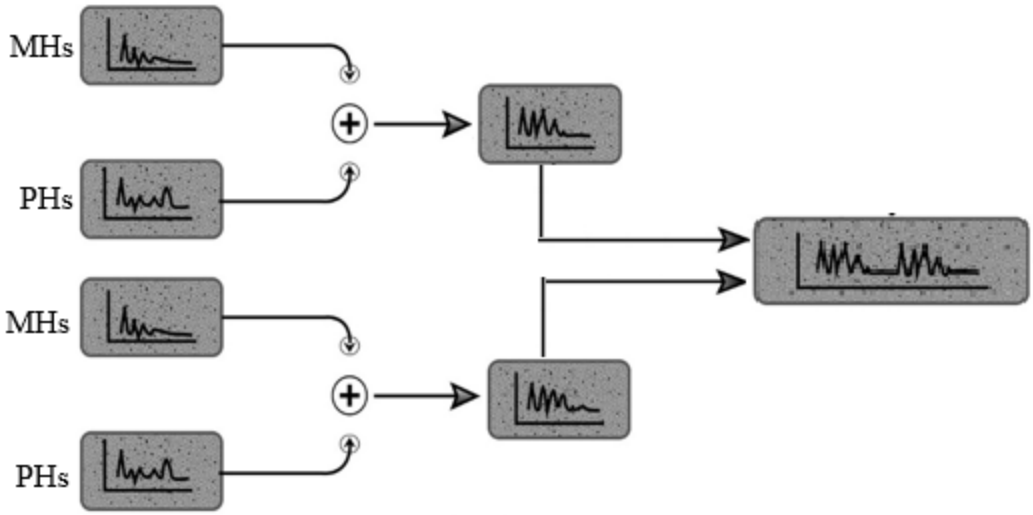


Figure 3.13: Concatenation of phase and magnitude histograms

Figure 3.14 shows the final feature vector of input image after concatenation of phase-magnitude histograms for its sub-images.

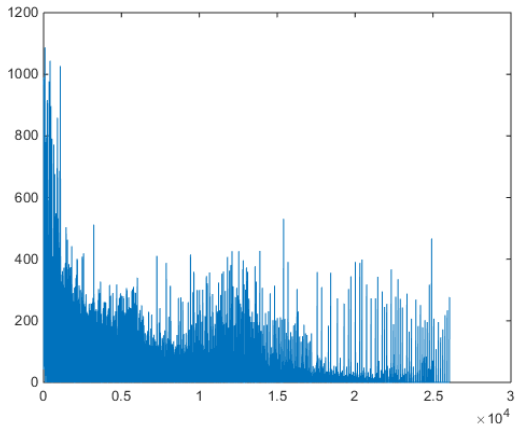


Figure 3.14: Final feature vector of the input image

Since the interval of 0-23 is used for the range of values and the number of active moments is six, the length of the final feature vector is 26064 which is acquire as following equation (3.11).

$$(N^2 + (N-1)^2) \times K \times b = 26064 \tag{3.11}$$

3.8 Dimensionally Reduction Using PCA

PCA is a useful statistical technique used for facial recognition, image processing, and image compression. This method determines new data set for the feature vector with respect to original data set. In the new data set, the first axis is oriented in such a way that data variance is maximized (i.e. along the direction with the greatest data dispersion). The second axis must be set perpendicular to the first axes which is expected to lead to the maximizing data. Afterward, the principal amount of data dispersion would be achievable in any data set. Not that, the new features are linear function of the old features. The main advantage of this method is that reduce the number of dimensions, without much loss of information. The redundancy is measured by correlations between data elements and using only the correlation has the advantage that analysis can be based on second order statistics only [27]. This approach has two advantages:

1. The final feature vector can be limited to facial features just by eliminating redundancy.
2. Since the dimensions are reduced, computation time can be reduced.

As previously studied, the achieved final feature vector is a 26064 dimensional vector. There are some redundancy in this dimensional vector which are not very effective in recognition and computation of this large vector take much more time. Thus, PCA should be employed for reducing the dimensions of the obtained feature vector. In this section, the proposed algorithm is explained as follow stages.

3.8.1 Calculate the Covariance Matrix

The input data of PCA can be in the n-dimensional vector. In this study, extracted feature vector is a vector in 2 dimensional with the length of its distinct numbers (x_i),

(y_i). The covariance matrix are first calculated from the respective equations as presented in equations (3.12) and (3.13). Where ($\overline{x}$) is mean value, *V*(*x*) is variance of values, the standard deviation of values is shown by *(σ)*.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\,, \qquad V(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \quad , \qquad \sigma = \sqrt{V(x)} \tag{3.12}$$

$$COV(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) \tag{3.13}$$

3.8.2 Calculate the Eigen Values and Eigen Vectors of Covariance Matrix

In this stage, the Eigen values and Eigen vectors for cov (*x, y*) are calculated as stated by equation (3.14). This is important because it means that you can express the data in terms of these perpendicular eigenvectors, instead of expressing them in terms of the *x* and *y* axes. According to linear algebra theorems, *(n×n)* matrix has (*n)* independent Eigen vectors and (*n)* independent Eigen values. According to Kramer theorem for homogeneous system of linear equations, only the answer is nontrivial when determinate coefficient is zero.

$$A_{n\cdot n}V_{n\cdot 1}=\lambda V_{n\cdot 1} \quad \text{or} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}\begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}=\lambda \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}, \ |A-\lambda I|=0 \tag{3.14}$$

The direction of Eigen values and Eigen vectors are shown in Figure 3.15 and it is evident that one of the two vectors is oriented such that the data contained within them have the maximum dispersion (*u*) and the other vector is perpendicular to it (*v*). These Eigen values actually represent the degree of data dispersion along the corresponding Eigen vectors.

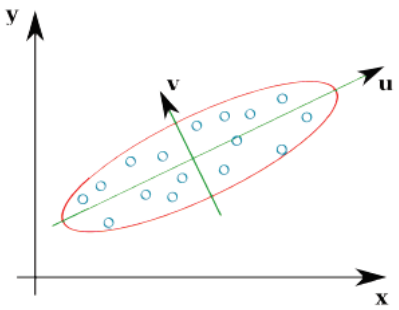


Figure 3.15: Direction of Eigen value and Eigen vector

3.8.3 Selection of Components and Constructing Feature Vectors

Herein, the Eigen vectors obtained through the previous stage are sorted in a descending order in terms of their Eigen values as shown in Figure 3.16 (note that all Eigen values obtained for the covariance matrix are zero or greater than zero). If you look at the eigenvectors and eigenvalues from the previous section, you will notice that the eigenvalues are quite different values. In fact, it can be concluded that the Eigen vector with the highest Eigen value is the principal component of the existing data.

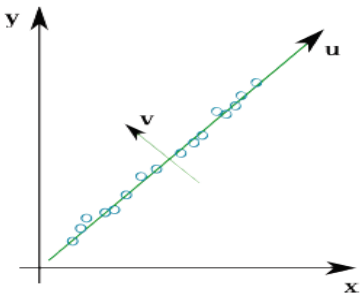


Figure 3.16: Dimension reduction of feature vector

Therefore, the data components are normalized according to their significance from more important to less important. Here, the less important components are also left out for decreasing the dimensions of the data so that this leads to a negligible data loss.

There are three different ways in the literature for eliminating some of the eigenvectors that come at the end of a sorted list. The first approach which is proposed by [33] is to remove the last 40% of the eigenvectors. This is a heuristic approach and the threshold has been obtained experimentally. The second approach tries to select minimum number of eigenvectors that will guarantee that energy *G* is greater than a typical threshold of 0.9.

In a system with *p* eigenvectors, the cumulative energy *G* up to and including the j^{th} eigenvector is the sum of the energy content across all of the eigenvalues from 1 through *j*:

$$G = \sum_{k=1}^{j} g_k \tag{3.15}$$

The goal is to choose a value of *j* as small as possible while achieving a reasonably high value for *G* on a percentage basis.

$$\frac{\sum_{k=1}^{j} g_k}{\sum_{k=1}^{p} g_k} \geq 0.9 \tag{3.16}$$

The third and last approach is to retain eigenvectors which has corresponding stretch values, s_i, greater than 0.01 [34]. The stretch value s_i is defined as the ratio of the i^{th} eigenvalue divided by the largest eigenvalue (λ_L).

$$s_i \geq 0.01 \quad \text{where} \quad s_i = \frac{(\lambda i)}{(\lambda L)} \tag{3.17}$$

Figure 3.17 below shows energy and stretching dimensions for the FERET data.

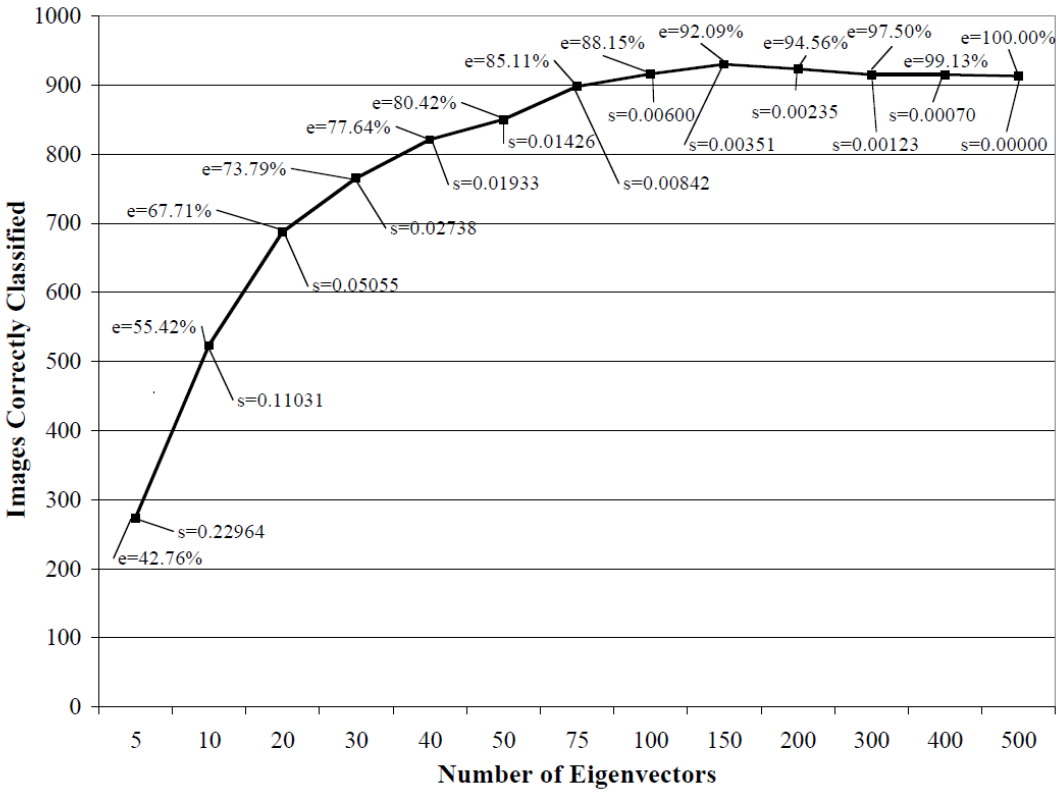


Figure 3.17: Energy and stretching dimensions for the FERET data

In this work to simplify the computational complexity, 25% of the trailing eigenvectors were eliminated. This is a bit lower than the 40 % limit suggested in [33]. With a training set of 1400 images this means that only the largest 1050 eigenvectors were kept and the remaining were eliminated.

3.8.4 Obtaining New Data Set

In the final stage of the PCA, the transpose of feature vector is multiplied by the transpose matrix of the normalized data. Once we have chosen the components (eigenvectors) that we wish to keep in our data and formed a feature vector, we simply take the transpose of the vector and multiply it on the left of the original data set, transposed.

Recall that the final transform is this:

$$\text{FinalData} = \text{RowFeatureVector} \times \text{RowDataAdjust} \tag{3.18}$$

Which can be turned around so that, to get the original data back,

$$\text{RowDataAdjust} = \text{RowFeatureVector}^{-1} \times \text{FinalData} \quad (3.19)$$

Where RowFeatureVector is the matrix with the eigenvectors in the columns transposed so that the eigenvectors are now in the rows, with the most significant eigenvector at the top and RowDataAdjust is the mean-adjusted data transposed.

Herein, it turns out that the inverse of our feature vector is actually equal to the transpose of our feature vector. This is only true because the elements of the matrix are all the unit eigenvectors of our data set. This makes the return trip to our data easier, because the equation becomes:

$$RowDataAdjust = RowFeatureVector^{T} \times FinalData \quad (3.20)$$

In special case, to get the actual original data back, we need to add on the mean of that original data because of all Eigen values obtained for the covariance matrix are zero or greater than zero. So, for completeness,

$$\text{RowOriginalData} = (\text{RowFeatureVector}^{\text{T}} \times \text{FinalData}) + \text{OriginalMean} \quad (3.21)$$

Not that, this formula also applies when there are not all the eigenvectors in the feature vector. So even when you leave out some eigenvectors, the above equation still makes the correct transform.

3.9 Distance Calculations and Selection of Minimum Distance

The final part of this project entails using the best available methods to compare feature vectors for images. There are various comparison methods as stated in [35]. In section 3.9.1, we summarize some of the most widely accepted ones.

3.9.1 Distance Calculations Methods

This section provides brief details on how Euclidean distance, Cosine distance and Mahalanobis distance metrics can be computed.

3.9.1.1 Distance Calculation via the Euclidean Distance Method

In this method, the difference between each feature vector of testing image (Q) and its corresponding recorded feature vector of training image (T) is obtained. Then, the square root is the sum of the squares of these differences as shown in equations (3.22).

$$d_{euc} = \sqrt{(Q - T)^2} \tag{3.22}$$

3.9.1.2 Distance Calculation via the Cosine Distance Method

In this method, the cosine distance computes the difference in direction, irrespective of vector lengths. The distance is given by the angle between the two vectors as shown in equation (3.23). By the rule of dot product.

$$Q \cdot T = Q^t T = |Q| \cdot |T| \cos\theta$$

$$d_{cos}(Q,T) = 1 - \cos\theta = 1 - \frac{QtT}{|Q| \cdot |T|} \tag{3.23}$$

3.9.1.3 Distance Calculation via Mahalanobis Distance Method

The Mahalanobis distance is a special case of the quadratic-form distance metric in which the transform matrix is given by the covariance matrix obtained from a training set of feature vectors, that is $A = \Sigma^{-1}$. In order to apply the Mahalanobis distance, the feature vectors are treated as random variables $X=[x_0, x_1 \ldots x_{n-1}]$, where x_i is the random variable of i_{th} dimension of the feature vector. Then, the correlation matrix is given by R where $R=[r_{ij}]$ and $r_{ij} = E\{x_i x_j\}$. $E\{x\}$ gives the mean of the random variable x. Then, the covariance matrix is given by Σ, where $\Sigma=[\ \sigma_{ij}^2\ ]$ and the Mahalanobis distance between two feature vectors Q and T is obtained by letting $X_Q = Q$ and $X_T = T$, which gives in equation (3.24).

$$\sigma_{ij}^2 = r_{ij} - E\{x_i\}E\{x_j\}$$

$$d_{mah} = [(XQ - XT)\Sigma - 1\ (XQ - XT)]^{\frac{1}{2}} \tag{3.24}$$

3.9.2 Justification of our Choice for Distance Calculation

In the literature there are many distance computation metrics. Some of the most widely accepted ones have been explained in Section 3.9. In this work we choose to adapt the Euclidean distance metric. This decision was based on the precision vs recall graph presented in [35], which is also shown below in Figure 3.18.

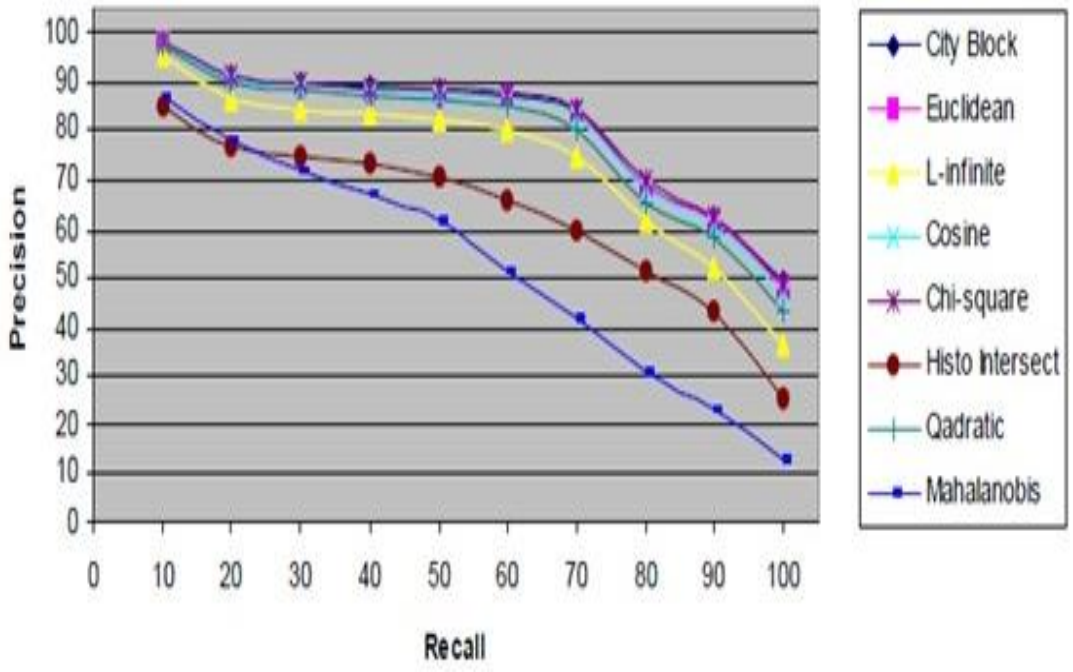


Figure 3.18: Performance of different distance methods

Also, in [35] and [36] it was stated that the Euclidian distance metric can be computed faster. Hence, in this work we adapted the Euclidean metric for distance calculations.

Recall which is also known as the true positive rate or sensitivity expresses how many relevant objects were recognized. Precision on the other hand expresses how relevant the returned set is.

Equations 3.25 and 3.26 can be used to compute the Recall and Precision values and both will vary in the range 0-1.

$$r = \frac{TP}{TP + FN} \tag{3.25}$$

$$p = \frac{TP}{TP + FP} \tag{3.26}$$

Where TP stands for the true positions, *FP* and *FN* denote respectively the false 9positives and false negatives.

Chapter 4

DATABASE WITH GALLERY, PROBE AND TARGET SETS

Face Recognition Technology (FERET) had emerged in Media Laboratory of MIT University as an image recognition data base [37]. The FERET was set out to establish a large database of facial images that was collected independently from the algorithm developers. The database collection was a collaborative effort between Dr. Wechsler and Dr. Phillips. The aim of the FERET is to develop automatic face recognition abilities that can be employed to assist security, intelligence, and law enforcement personnel in the performance of their duties.

The available images in FERET are classified in two separate categories. In the first set, target images and the second set covers probe images including person's face under different conditions comprising illumination and pose variation. Additionally, facial images taken several years after the original image are accessible in probe set.

As mentioned earlier, the images in FERET database are divided into two set as known probe images and target images which are mainly used for facial recognition tests. Structure of FERET database are shown in Figure 4.1.

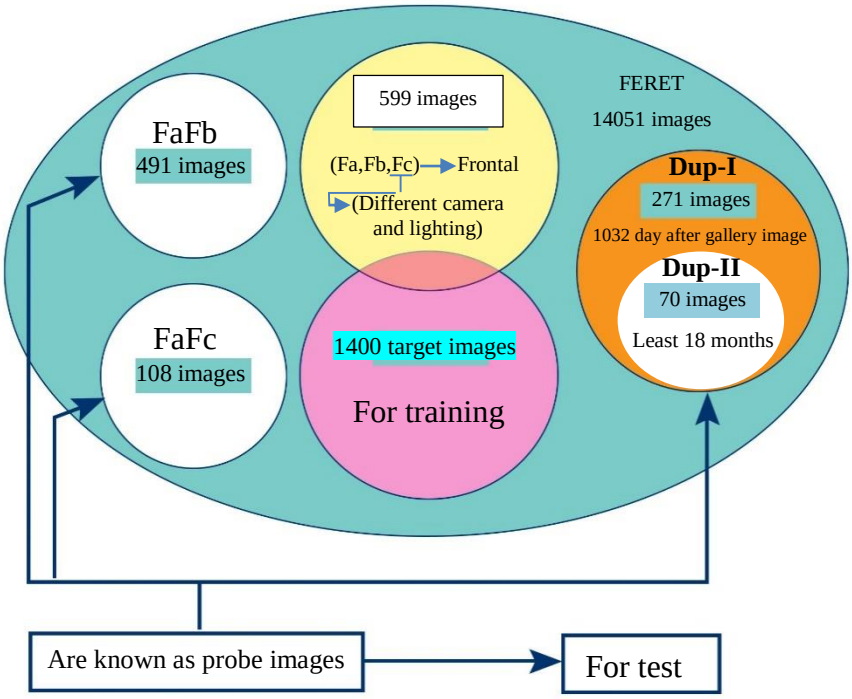


Figure 4.1: FERET database

There are 1400 images for entrance images for training. As shown in Figure 4.1, FERET introduces some of target and probe images from data base. Firstly, the target images become train, therefore, the system has been tested according the probe images by FaFb, FaFc, Dup-I, and Dup-II sets. The images of FaFb and FaFc sets have taken photo from frontal view, moreover, the images of FaFc set have taken photo from different cameras with illumination variation. Dup-I set have taken 1031 days after respective gallery matches. Furthermore, the images of Dup-II set is the subset of Dup-I set images which have taken 18 months after respective gallery matches.

The FERET database includes 14051 facial images with 256×384 image resolution. Some sample of images from gallery set are shown in Figure 4.2.

Figure 4.2: Sample images from gallery set

Some sample images of FaFb set are shown in Figure 4.3, they have been taken photo in frontal view including 491 facial images.

Figure 4.3: Sample images from FaFb set

Some sample images of FaFc set are shown in Figure 4.4. These photos have been taken in frontal view from different cameras with illumination variation including 108 facial images.

Figure 4.4: Sample images from FaFc set

Some sample images of Dup-I set are shown in Figure 4.5. Dup-I set includes 271 facial images which have been taken 1031 days after respective gallery matches.

Figure 4.5: Sample images from Dup-I set

Some sample images of Dup-II set are shown in Figure 4.6. The images of Dup-II set is subset of Dup-I set which have been taken after 18 months from respective gallery matches including 70 facial images.

Figure 4.6: Sample images from Dup-II set

Finally, the images of FaFb, FaFc, Dup-I and Dup-II sets are known as the probe images which can be entered in the system as images of test. Note that, target images are trained in this stage, therefore, their results are compared with the results of probe images based on minimum distance.

Chapter 5

SIMULATION RESULTS

This section provides a performance analysis for the LZM based proposed method and compares the results obtained with those obtained from other state of the art face recognition methods. For comparisons the Local Binary Patterns (LBP) and Histograms Zernike Moments (H-LZM) methods were simulated and compared with our H-LZM-PCA method. During the experiments the Face Recognition Technology (FERET) database introduced in Chapter 4 was used. The images available in the FERET database are divided into two groups: the probe sets and target set. The target set is the name given to the set of known facial images and query set refers to the set of unknown facial images that needs to be identified. In some papers in the literature target and probe sets are referred to as training and testing sets.

First, the proposed method obtains feature vectors for all images in the target set using the block diagram provided in Figure 3.1 and stores them in a database. This part of the processing is referred to as the training stage. Then individual image selected from the probe set is processed to obtain the reduced dimensionality feature vector. The algorithm computes the minimum Euclidean distance between all feature vectors stored in the database and the feature vector of the selected probe image. Therefore, the system compares the name of candidates and if the first five letters of the candidate's names turn out to be equal to selected probe image name, the

candidate is specified and introduce in rank list (A rank ordered candidate list of the percent most likely matches for any given probe image).

In this work the system compares the probe image against a whole target images in order to establish a match in rank list. In comparing the probe image with the target images, a similarity score is normally generated. These similarity scores are then sorted from the highest to the lowest (where the lowest is the similarity that is equal to the operating threshold). This means that a higher threshold would generate a shorter rank list and a lower threshold would generate a longer list. The operator is presented with a ranked list of possible matches in descending order. A probe image is correctly identified if the correct match has the highest similarity score (i.e., is placed as "rank 1" in the list of possible matches). The percentage of times that the highest similarity score is the correct match for all individuals submitted is referred to as the top match score. It is unlikely that the top match score will be 100% (i.e., that the match with the highest similarity score is indeed the correct match). Thus, one would more often look at the percentage of times that the correct match will be in the n^{th} rank (i.e., in the top n matches) [38].

The closed-set identification outcome was a rank 50 list of the best matches per identification run against the total database. A match was considered to have been made if the person appeared in the list of fifty best matches. The performance of a closed-set identification system will typically be described as having an identification rate at rank n. For example, a system that has a 94.5% identification rate in FaFb set at rank 3 would mean that the system will be 94.5% sure that the person in the probe image is in either position 1, 2, or 3 in the ranked list presented to the operator.

While a person was deemed to have been successfully identified when that person's image was included in the rank 5 list of best matches, we provide the closest 5 matches as rank1 to rank5 where rank1 indicates the highest recognition percentage. Respectively, other images of probe set are entered for comparing with all target images. In Table 5.1, the closet five ranks are shown for images in probe set (Dup-II) when compared with images in the target set.

Table 5.1: Rank1-Rank5 matches for images in probe set Dup-II when compared with images in the target set.

Closest 5 matches	Rank5	Rank4	Rank3	Rank2	Rank1
Similarity	0.720	0.720	0.731	0.755	0.760

The thesis presents simulation results in two-fold, the first part of the results are achieved by combining PCA, LZM and some image processing techniques to design a face recognition system based on minimum distance criteria. Tables 5.2-5.5 respectively present simulation results for FaFb, FaFc, Dup-I and Dup-II query sets.

The second part of the results are obtained by performing almost the same steps but after the LZM filtering the reduced size moment-images are processed by a 2D Gaussian weighting function. Tables 5.6-5.9 respectively present simulation results for FaFb, FaFc, Dup-I and Dup-II probe sets.

Table 5.2: Face recognition results using probe set FaFb (no Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.939	0.942	0.945	0.947	0.947

Table 5.3: Face recognition results using probe set FaFc (no Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.858	0.858	0.863	0.863	0.863

Table 5.4: Face recognition results using probe set Dup-I (no Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.720	0.722	0.724	0.729	0.731

Table 5.5: Face recognition results using probe set Dup-II (no Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.667	0.667	0.671	0.690	0.700

Table 5.6: Face recognition results using probe set FaFb (Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.972	0.972	0.972	0.972	0.972

Table 5.7: Face recognition results using probe set FaFc (Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.911	0.930	0.930	0.930	0.950

Table 5.8: Face recognition results using probe set Dup-I (Gaussian weighting)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.751	0.751	0.765	0.765	0.780

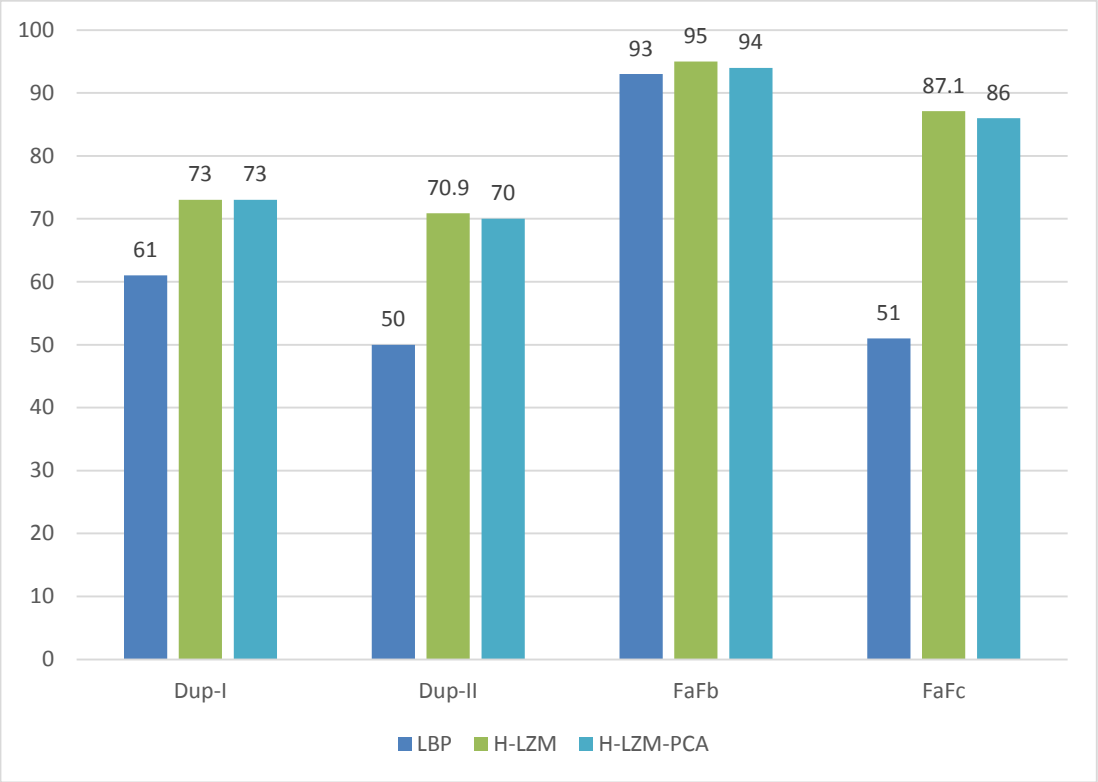
Table 5.9: Face recognition results using probe set Dup-II (Gaussian weightining)

Rank 5	Rank 4	Rank 3	Rank 2	Rank 1
0.720	0.720	0.731	0.755	0.760

According to tables results, while using reduced dimensionality feature vector based on PMHs the accuracy of recognition for the probe set FaFb, FaFc, Dup-I and Dup-II were respectively 94%, 86%, 73%, and 70%. When a 2D Gaussian kernel is applied on the magnitude of the moment images obtained from the LZM filter, the accuracy of recognition for the same probe sets became 97%, 95%, 78%, and 76%. This indicates that filtering the moment-images using a 2D weighting function helps improve the results on average by 5.75 %.

Finally we compared the accuracy of our proposed system with face recognition systems using Local Binary Patterns (LBP) and Histograms Local Zernike Moments (H-LZM). As can be seen in Table 5.10, our H-LZM-PCA (no weighting) are compared with the LBP and H-LZM (no weighting).

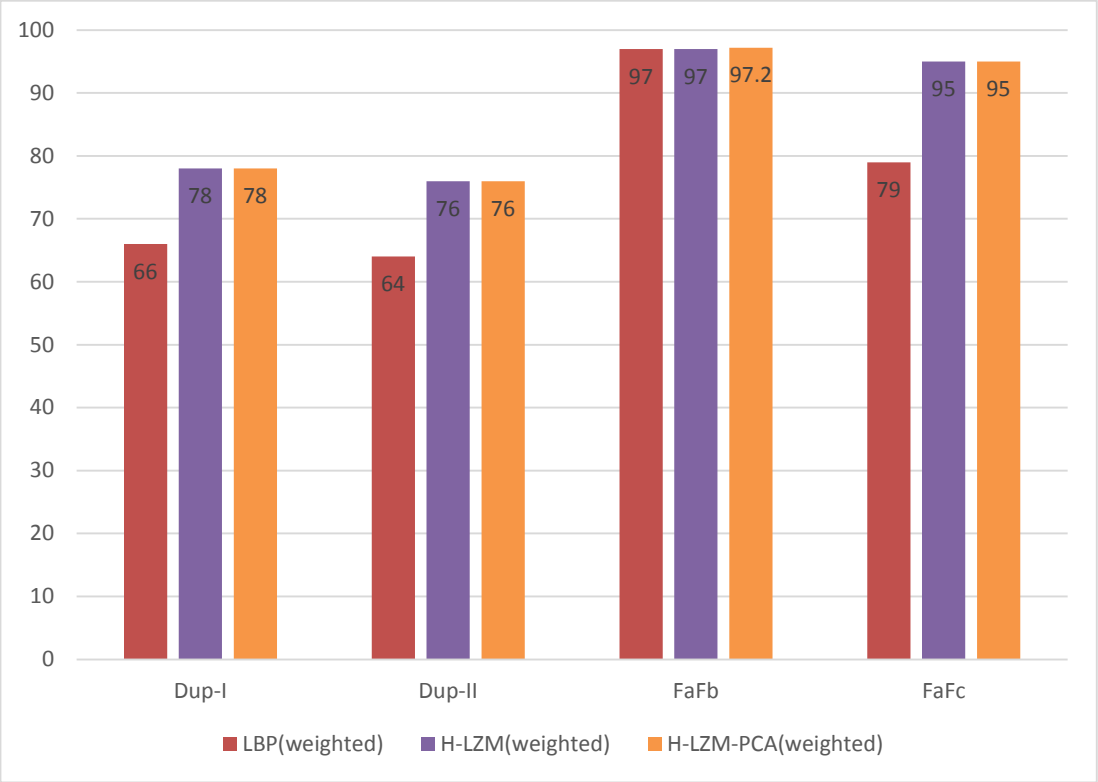
Table 5.10: Comparison of recognition rates for the proposed method, LBP and H-LZM



We have confirmed that using PCA, features vectors were limited to facial features by removing other additional information and since the dimensions are reduced, computation time reduced [39] [40]. Also results obtained from applying the proposed algorithm to the FaFb and FaFc sets were satisfactory where facial expression variations are included with brightness variations. This is due to the fact that, in this method, first normalized sub-images are obtained so the change in the person's expression in one image component and illumination variations produced satisfied results.

As can be seen in Table 5.11, our H-LZM-PCA (weighting) are compared with the LBP and H-LZM (weighting).

Table 5.11: Comparison of recognition rates for the proposed method, LBP and H-LZM with 2D Gaussian weighting.



Results indicate that proposed algorithm using Gaussian weighting to the FaFb and FaFc sets were satisfactory. The results obtained for Dup-I and Dup-II sets were less than FaFb and FaFc sets results since the aging variations changed the tested feature vector. This solution is discussed in the Future work section.

Compression of Table 5.10 and Table 5.11 is indicated, the results of FaFb, FaFc, Dup-I and Dup-II data sets using a 2D weighting function helps improve the results on average by 5.75 % and it is because the weights give higher significance to pixels near the edge and reduce edge blurring for each sub-image. Our H-LZM-PCA methods would provide higher recognition rates for all probe sets when compared with the LBP and while using reduced dimensionality feature vector by PCA algorithm, computation time reduced when compared with H-LZM.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

The thesis combined PCA, LZM and some image processing techniques to design a face recognition system based on minimum distance criteria. Firstly, the system computes moment images that were obtained by partitioning the input intensity image, normalizing the sub-images, passing them through an LZM filter and then extracting and concatenating the PMHs of each smaller moment-image to create a final feature vector that could be used for recognition purposes. Secondly, after the LZM filtering the reduced size moment-images are passed through a 2D Gaussian weighting.

Our system was tested using some probe sets from the FERET image database. While using reduced dimensionality feature vector by PCA algorithm based on PMHs, the accuracy of recognition for the probe sets were respectively 94%, 86%, 73%, and 70%. When a 2D Gaussian kernel is applied on the magnitude of the moment images obtained from the LZM filter the corresponding accuracies for the same probe sets became 97%, 95%, 78%, and 76%. This indicates that filtering the moment-images using a 2D weighting function helps improve the results on average by 5.75 % and it is because the weights give higher significance to pixels near the edge and reduce edge blurring for each sub-image.

Finally we had compared the accuracy of our proposed system with face recognition systems using LBP and H-LZM. Results indicate that our H-LZM-PCA methods could provide higher recognition rates for all probe sets when compared with the LBP and while using reduced dimensionality feature vector by PCA algorithm, features vectors were limited to facial features by removing other additional information and since the dimensions are reduced, computation time reduced when compared with H-LZM.

6.2 Future Work

In the future, our algorithm can be modified such that it takes the phase and magnitude of the moment images obtained from the first LZM transform and simultaneously feeds them into two different second layer LZM filter blocks. We expect that applying the LZM transformation twice will help increase the systems accuracy in recognizing people with a higher degree of efficiency.

The results obtained for Dup-I and Dup-II data sets were less than FaFb and FaFc sets results since the aging variations changed the feature vectors. For eliminating the effects of time-based dimensions on the facial recognition system, a “live” system can be devised for updating image’s feature vector based on his/her latest encounter with the system. Such a measure shall reduce the changing effects in the image’s feature vector due to aging.

REFERENCES

[1] Li, S.Z., & Jain, A., "Handbook of face recognition", *Springer-Verlag London*, 2th ed., ch.1, pp.1-8, 2011.

[2] Kirby, M., & Sirovich, L., "Application of the Karhunen-Loeve procedure for the characterization of human faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.1, pp.103-108, 1990.

[3] Sirovich, L., & Kirby, M., "Low-dimensional procedure for the characterization of human faces", *Journal of the Optical Society of America A (JOSA A), Optics and Image Science*, vol.4, no.3, pp.519-524, 1987.

[4] Terzopoulos, D., & Waters, K., "Analysis of facial images using physical and anatomical models", *IEEE Third International Conference on Computer Vision*, Osaka, Japon, pp.727-732, 1990.

[5] Manjunath, B.S., Chellappa, R., & von der Malsburg, C., "A feature based approach to face recognition", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Champaign, Urbana, pp.373-378, 1992.

[6] Kaufman Jr., G.J., & Breeding, K.J., "The automatic recognition of human faces from profile silhouettes", *IEEE Transactions on Systems, Man and Cybernetics*, vol.6, no.2, pp.113-121, 1976.

[7] Kerin, M.A., & Stonham, T.J., "Face recognition using a digital neural network with self-organising capabilities", *IEEE International Conference on Pattern Recognition*, New Jersy, USA, vol.1, pp.738-741, 1990.

[8] Yuille, A.L., Cohen, D.S., & Hallinan, P.W., "Feature extraction from faces using deformable templates", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Harvard Univ., Cambridge, MA, USA, pp.104-109, 1989.

[9] Jafri, R., & Arabnia, H.R., "A survey of face recognition techniques", *Journal of Information Processing Systems* (*JIPS)*, vol.5, no.2, pp.41-68, 2009.

[10] Liu, J.N.K., Meng, W., & Bo, F., "An internet-based intelligent robot security system using invariant face recognition against intruder", *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol.35, no.1, pp.97-105, 2005.

[11] Acosta, E., et al., "An automatic face detection and recognition system for video indexing applications", *IEEE International Conference on Acoustics, Speech and Signal Processing*, Florida, USA, vol.4, pp.3644-3647, 2002.

[12] Balci, K., & Atalay,V., "PCA for gender estimation: which eigenvectors contribute? ", *IEEE 16th International Conference on Patern Recognition*, Qebec, Canada, vol.3, pp.363-366, 2002.

[13] Moghaddam, B., & Ming-Hsuan, Y., "Learning gender with support faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, no.5, pp.707-711, 2002.

[14] Colmenarez, A., Frey, B., & Huang,T.S., "A probabilistic framework for embedded face and facial expression recognition", *IEEE Computer Society Conference in Computer Vision and Pattern Recognition*, NY, USA, vol.1, pp.597, 1999.

[15] Shinohara, Y., & Otsu, N., "Facial expression recognition using Fisher weight maps", *Sixth IEEE International Conference in Automatic Face and Gesture Recognition*, Seoul, Korea, pp.499-504, 2004.

[16] Farokhi, S., et al., "Near infrared face recognition by combining Zernike moments and undecimated discrete wavelet transform", *Elsevier-Digital Signal Processing: A Review Journal*, vol.31, pp. 13-27, 2014.

[17] Singh, C., Mittal, N., & Walia, E., "Complementary feature sets for optimal face recognition", *Springer-EURASIP Journal on Image and Video Processing*, vol.2014, no.1, pp.1-18, 2014.

[18] Singh, C., Walia, E., & Mittal, N., "Robust two-stage face recognition approach using global and local features", *Springer-Verlag in Visual Computer*, vol.28, no.11, pp.1085-1098, 2012.

[19] Hajati, F., Raie, A.A., & Gao, Y., "3D face recognition using geodesic PZM array from a single model per person", *Ieice Transactions on Information and Systems*, vol.E94D, no.7, pp.1488-1496, 2011.

[20] Sariyanidi, E., et al., "Local Zernike Moments: A new representation for face recognition", *IEEE International Conference on Image Processing*, Istanbul, Turkey, pp.585-588, 2012.

[21] Turk, M.A, & Pentland, A.P., "Face recognition using eigenfaces", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

[22] Ahonen, T., Hadid, A., & Hadid, P., M., "Face description with local binary patterns: Application to face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Media lab., MIT, Cambridge, MA, USA, vol.28, no.12, pp.2037-2041, 2006.

[23] Wiskott, L., et al., "Face recognition by elastic bunch graph matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.775-779, 1997.

[24] Gabor, D., "Theory of communication. Part 1: The analysis of information", *Part III: Radio and Communication Engineering, Journal of the Institution of Electrical Engineers*, vol.93, no.26, pp.429-441, November 1946.

[25] Yi, J., & Ruan, Q.Q., "Face recognition using Gabor-based improved supervised locality preserving projections", *Computing and Informatics*, vol.28, no.1, pp.81-95, 2009.

[26] Ince, E.A., & Ali, S.A., "Rule based segmentation and subject identification using fiducial features and subspace projection methods", *Journal of Computers (Finland)*, vol.2, no.4, pp.68-75, 2007.

[27] Jolliffe, I.T., "Principal component analysis", *Springer-Verlang*, 1986.

[28] Han, P., Wu, J., & Wu, R., "SAR target feature extraction and recognition based on 2D-DLPP", *Elsvier-Physics Procedia*, Vol.24, no.B, pp.1431-1436, 2012.

[29] Alasag, T., & Gokmen, M., "Face recognition in low resolution images by using local Zernike moments", *International Conference on Machine Vision and Machine Learning*, Czech Republic, no.125, pp.1-7, August 2014.

[30] Gonzalez, R.C., & Woods, R.E., "Digital Image Processing", *Prentice Hall, Upper Suddle River*, 2th Ed., 2002.

[31] Brunelli, R., & Poggio, T., "Face recognition: features versus templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.15, no.10, pp.1042-1052, Oct 1993.

[32] Khotanzad, A., & Hua Hong, Y., "Invariant image recognition by Zernike moments", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.5, pp.489-497, May 1990.

[33] Yambor, W.S., Draper, B.A., & Beveridge, J.R., "Analyzing pca-based face recognition algorithms: Eigenvector selection and distance measures", *2nd Workshop on Empirical Evaluation in Computer Vision*, Dublin, Ireland, 2000.

[34] Kirby, M., "Dimensionality Reduction and Pattern Analysis: An Empirical Approach", 2000.

[35] Zhang, D., & Lu, G., "Evaluation of similarity measurment for image retrieval", *Neural Network and Signal Processing, Proceeding of the 2003 International Conference on*, vol.2, pp.928-931, Dec 2003.

[36] Draper, B.A., Baek, K., Bartlett, M.S., & Beveridge, J.R., "Recognizing faces with PCA and ICA", *Elsevier-Computer Vision and Image Understanding*, vol.91, no.1, pp.115-137, July–August 2003.

[37] Rizvi, S.A., Phillips, P.J., & Moon, H., "The FERET verification testing protocol for face recognition algorithms", *Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japon, pp.48-53, April 1998.

[38] Introna, L., & Nissenbaum, H., "Facial Recognition Technology A Survey of Policy and Implementation Issues", 2010.

[39] Panahi, N., Shayesteh, M.G., Mihandoost, S., & Varghahan, B.Z., "Recognition of different datasets using PCA, LDA, and various classifiers", *5th IEEE International Conference on Application of Information and Communication Technologies (AICT)*, Baku, Azerbaijan, pp.1-5, Oct 2011.

[40] Kekre, H.B., & Shah, K. "Energy efficient face recognition using row, column feature vectors of Slant Transform and performance comparison with PCA", *IEEE Symposium on Industrial Electronics & Applications (ISIEA),* vol.1, pp.150-155, Oct 2009.