AUGUST 2016 M.Sc. THESIS IN ELECTRONICS AND COMPUTER ENGINEERING MİTHAT ÇAĞRI YILDIZ

HASAN KALYONCU UNIVERSITY
GRADUATE SCHOOL OF
NATURAL AND APPLIED SCIENCES

FACE FEATURE SELECTION USING GENETIC ALGORITHM UNDER DIFFERENT BIOMETRIC VARIATIONS

M.Sc. THESIS
IN
ELECTRONICS AND COMPUTER ENGINEERING

BY
MİTHAT ÇAĞRI YILDIZ
AUGUST 2016

FACE FEATURE SELECTION USING GENETIC ALGORITHM UNDER DIFFERENT BIOMETRIC VARIATIONS

M.Sc. Thesis
In
Electronics and Computer Engineering
Hasan Kalyoncu University

Supervisor:

Prof. Dr. Celal KORAŞLI

By
Mithat Çağrı YILDIZ
August 2016

© 2016 [Mithat Çağrı YILDIZ]

REPUBLIC OF TURKEY
HASAN KALYONCU UNIVERSITY
GRADUATE SCHOOL OF NATURAL & APPLIED SCIENCES
ELECTRONICS AND COMPUTER ENGINEERING DEPARTMENT

Name of the thesis: Face Feature Selection using Genetic Algorithm under Different Biometric Variations

Name of the student: Mithat Çağrı YILDIZ
Exam date: 05 August 2016
Approval of the Graduate School of Natural and Applied Sciences.

Prof. Dr. Mehmet KARPUZCU
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Celal KORAŞLI
Head of Department

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

Prof. Dr. Celal KORAŞLI
Supervisor

Examining Committee Members:	Signature
Prof. Dr. Celal KORAŞLI	……………….
Doç. Dr. SEMA KOÇ KAYHAN	……………….
Doç. Dr. Muhammet Fatih HASOĞLU	……………….

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by this rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Mithat Çağrı YILDIZ

ABSTRACT

FACE FEATURE SELECTION USING GENETIC ALGORITHM UNDER DIFFERENT BIOMETRIC VARIATIONS

YILDIZ, Mithat Çağrı
M.Sc. in Electronics and Computer Engineering
Supervisor: Prof. Dr. Celal KORAŞLI
August 2016, 73 pages

In the current study face recognition under different biometric variations is investigated applying Principal Components Analysis (PCA). In order to improve the recognition performance Genetic Algorithm (GA) is selected. The algorithm follows optimized selection of PCA features based on GA operations on the datasets ORL, FERET and BANCA.

The maximum recognition rate (MRR) results obtained with ORL and FERET databases are found to be close to the results of computed with WAVELET-PCA-GA-SVM method. Further the MRR results obtained for BANCA database is 100% as that of the computed with WAVELET-PCA-GA-SVM method for YALE and YALE-B databases.

Generally PCA on GA is found to be effective in removing irrelevant data groups and therefore it improves the performance.

Keywords: Biometrics, Face Recognition, Genetic Algorithm, Optimization, Principal Components Analysis.

ÖZET

FARKLI BİYOMETRİK DEĞİŞİMLER ALTINDA GENETİK ALGORİTMA KULLANARAK YÜZ TANIMA ÖZELLİĞİ

YILDIZ, Mithat Çağrı
Yüksek Lisans Tezi, Elektronik ve Bilgisayar Mühendisliği
Danışman: Prof. Dr. Celal KORAŞLI
Ağustos 2016, 73 sayfa

Bu çalışmada Temel Bileşenler Analizi (PCA) uygulanarak farklı biyometrik değişimler altında yüz tanıma performansını artırmak için Genetik Algoritma (GA) uygulanması yapıldı. Tasarımı yapılan algoritmik yöntem ile PCA'nın performansının GA yönteminin ORL, FERET ve BANCA veri tabanlarına uygulanarak optimize edilmiştir.

Geliştirilen bu algoritmik yöntemin ORL ve FERET veri tabanlarına uygulanması ile elde edilmiş MRR (Maksimum Tanıma Oranı) sonuçlarının, WAVELET-PCA-GA-SVM yöntemi ile bulunan sonuçlara yakın olduğu tespit edilmiştir. Ayrıca, yöntemin BANCA veri tabanı uygulanmasından elde edilen 100% Maksimum Tanıma Oranı sonucu, YALE VE YALE-B veri tabanlarına uygulanan WAVELET-PCA-GA-SVM yönetimden elde edilen sonuçla aynı olduğu bulunmuştur.

Genel olarak bu çalışmanın önemli bir neticesi PCA üzerinde yapılan GA işlevi PCA 'nın önemsiz sayılabilecek veri gruplarını kaldırmada etkili olduğu ve performansı arttırıcı özellikleri olduğudur.

Anahtar Kelimeler: Biyometri, Yüz tanıma, Genetik algoritma, Optimizasyon, Temel Bileşenler Analizi.

To My beloved Parents,

My Brothers and Sisters

And all my family

ACKNOWLEDGEMENTS

First and foremost, I would like to express my supervisor Prof. Dr. Celal KORAŞLI for his guidance, advice, criticism, encouragements and many creative insights offered during this thesis.

Special thanks are reserved Doç. Dr. Muhammet Fatih HASOĞLU, and Prof. Dr. Celal KORAŞLI for serving on the committee and their contributions and suggestions to improve the quality of the thesis.

Last but not least, I will be grateful to my parents, my family and all my friends for the encouragement, endless support and assistance

I would like to express my thanks to my family for the supported they provided me through my entire life, without their encouragement and help, I would not have finished this M.Sc. study.

LIST OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

LIST OF ABBREVIATIONS

HE	Histogram equalization
MVN	Mean-Variance Normalization
PCA	Principal Component Analysis
GA	Genetic Algorithm
MRR	Maximum recognition rate
MD	Manhattan Distance

CHAPTER 1

INTRODUCTION

1.1 Biometrics

Biometric recognition systems are broadly used and aim to recognize human beings based on their physical and/or behavioral characteristics [1]. Basically, “biometrics” originally comes from the Greek words “bio” (life) and “metrics” (to measure) [2]. In fact, in order to identify physical and/or behavioral characteristics of individuals based on the statistical measurements biometric is used. The idea of identifying human beings using different parts of the body comes from ancient times. In ancient Babylon about two thousand years ago, merchants recorded the trading transactions sealed deals with fingerprints on clay tablets [3]. Chinese in the 3rd century B.C applied thumbprints and fingerprints on clay tablets as signatures to seal the official documents. Several official document papers dated in Persia bore fingerprint impressions in the 14th century A.D. [4, 5]. In the 14th century, a Portuguese writer used stamped children’s palm print and footprints on paper against Chinese merchants [6] for identification purposes.

The French police, in the 19th century, established an anthropometric system, known as Bertillonage [7], in order to fix the problem of identification of convicted criminals. The identification system was built on the assumption that the bodies of people do not change in the basic characteristics and consist five primary measurements of body parts such as head length; head breadth; length of the middle finger and the length from elbow to end of middle finger. The system recorded the length of the little finger and the eye color. Recently, new methods of biometrics are applied in many applications for authorized person to enter a restricted place and to identify or verify a person.

Generally, biometrics are considered as a method to authenticate human being's identity by measuring the individual’s unique characteristics. By increasing the security needs in places such as airports, ATM, etc., traditional identification/verification systems that use user names, passwords and identification cards that can be forgotten or stolen easily, are being replaced by biometric systems. In fact, extracted information from biometric traits is more stable and reliable during human lifetime and cannot be forgotten or stolen. DNA, Ear, Face, Facial Thermogram, Finger Geometry, Fingerprint, Gait, Hand Geometry, Hand Vein, Iris, Keystroke, Odor, Palmprint, Retina, Signature and Voice are some instances of biometric modalities [4].

Figure 1.1 shows some biometric samples. Face recognition is considered as one of the interested and secure biometric traits [8]. Generally, biometric traits are categorized as, anatomical and behavioral traits in different applications [9, 10]. Iris, face, ear, hand, retinal scan, DNA, palmprint or fingerprint is some instances of Anatomical trait. While speech, handwriting, signature, gait or keystroking are considered as behavioral traits. On the other hand, some kinds of biometric traits; as an example voice; is considered as a combination of both anatomical and behavioral traits [9, 10].

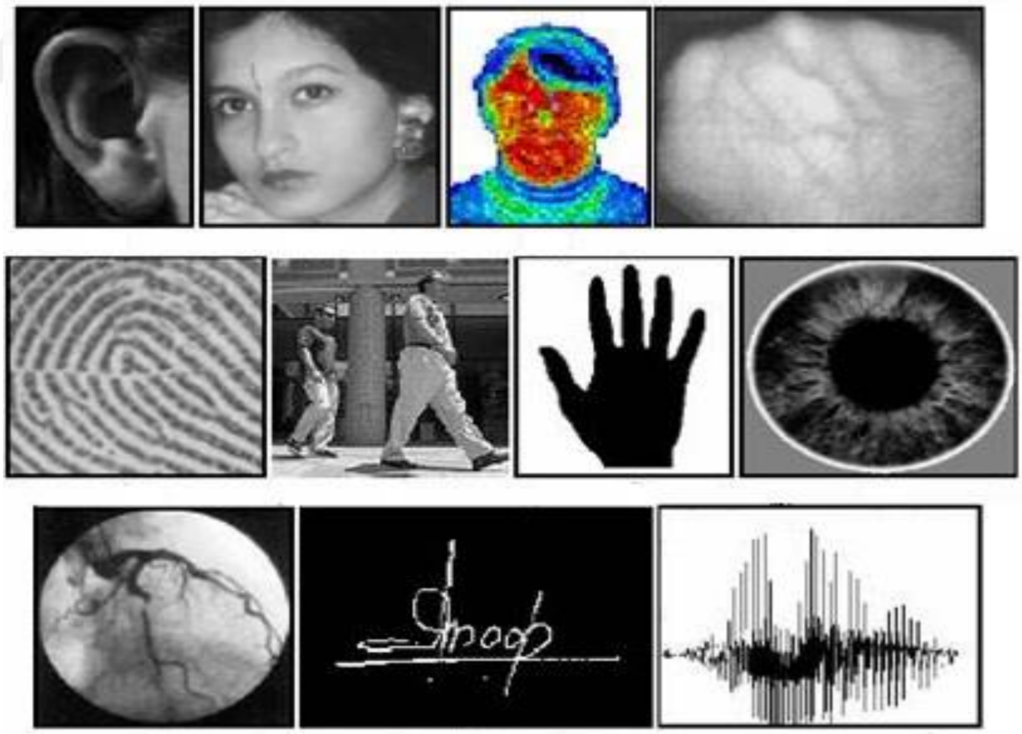


Figure 1.1 Different Biometric Characteristics. [8]

Generally, a biometric recognition system is categorized as verification mode or an identification mode. In the verification mode, an individual identity is validated by comparing the captured biometric trait with his/her own biometric templates that is available in the database. White verification systems usually validate a claim identity that is submitted via a personal identification number (PIN), a user name, or a smart card. Subsequently, the system performs a one-to-one comparison to test the claim validity. On the other hand, the identification mode, a person is recognized by searching the templates of all the individuals in the database for a match. The identification system performs a one-to-many comparison to set up an individual's identity [11].

In order to design a biometric system, consideration of following four main modules is needed [9].

- Sensor module, to acquire the biometric data of a user. The camera is an example of sensor module that captures the face images of an individual.
- Feature extraction module, to extract a set of significant information from the raw biometric traits. For example, extracting the global and/or local information of a face image using feature extraction methods in the feature extraction module of a face-based biometric system.
- Matcher module, to produce matching scores for comparing the features extracted during recognition against the stored templates. For example, reporting the number of matching scores between the input and the template face images. Usually each matcher module involves the encapsulation of a decision making module for verifying a user's claimed identity (verification) or establishing a user's identity (identification) based on the matching score.
- System database module, to store the biometric templates of the registered users. During the registration phase, the biometric trait of a user is scanned by a biometric reader. Generally, multiple templates of a user are stored in the database due to variations observed in the biometric trait. The database templates may be updated over the time.

1.2 Face recognition

Face recognition has been comprehensively studied in recent years. However, finding robust techniques under varying conditions is still an unsolved problem. In fact, face recognition is considered as one of the most relevant applications of image analysis. Plenty of research focuses on face recognition [12-15] with the aim of developing robust methods to recognize faces under different illumination variations, pose, and partial occlusions. In fact, illumination variations such as shadows, underexposure, and overexposure are considered as fundamental problems in a practical recognition system. In this respect, researchers introduce various techniques to deal with illumination changes in the past decades [16].

Illumination variations, pose and occlusions have been studied separately in the literature to improve the robustness of face recognition methods. In [17], a face recognition system based on dynamic similarity features in the presence of occlusion has been proposed, which randomly crops different size patches on the test images. They have validated their scheme on FERET database. The patches were small squares which were placed on the cropped face randomly. The proposed scheme is considered as a generalization of Kernel Discriminant Analysis algorithm and it is robust to random occlusions to some extent.

The authors of [13] have performed some experiments on AR database to recognize faces under different conditions such as using partially occluded faces with sunglasses and scarf, and using synthetic occlusion patterns. The experimental results have demonstrated that their proposed recognition algorithm is better than the other algorithms on real occlusions such as sunglass and scarf occlusions.

In reference [14], a face recognition system based on Lophoscopic PCA has been proposed. The system uses six subsets of images in the training set which correspond to the whole face, and the masking of the left eye, right eye, both eyes, nose and mouth respectively. All the subjects are then fused by using different combination techniques. The authors [14] emphasized that Lophoscopic PCA outperforms PCA in terms of experiments. They have performed their face recognition system on UPC database. The authors of [15] have proposed another face recognition scheme for under expression and/or illumination variation. The scheme is based on wavelet transform and improved the recognition accuracy compared to PCA on JAFFE and ORL databases.

In the field of face recognition there are many studies aim to investigate the robustness of different face recognition techniques in the presence of various difficulties. However, illumination variations, pose, and occlusions on the faces were not considered altogether in the same study. This study aim is to investigate the effect of the aforementioned difficulties on face recognition and its attempts to reduce the effect illumination variations, pose, and occlusions by selecting the optimized subset of features.

1.3 Genetic algorithm

Genetic Algorithms (GAs) are considered as heuristic search algorithms to model evolutionary ideas of natural selection and genetics. Genetic algorithms typically are considered for population sizes larger than one. Each individual of the population corresponds to a single solution, and therefore genetic algorithms are considered as an appropriate measurement to generate the optimal solution set in optimization problems [18].

Generally, genetic algorithm considers an initial set of solutions, called population. Each population is represented by chromosomes. In this respect, generation of new solutions is performed using genetic operators such as recombination, crossover and mutation via different generation. The aim of this procedure is to find better solutions by improving earlier ones. In each iteration (generation), a fitness value F(x) is calculated for each individual representing the goodness of the solution. The algorithm stops when an end condition is satisfied. We can briefly represent the basic concepts of a standard genetic algorithm as of coding, a fitness function and reproduction [18].

1.4 Literature review

To improve face recognition systems, searchers apply face image processing techniques based on faces under different illumination variations, pose, and partial occlusions. As a matter of fact, illumination variations such as shadows, underexposure, and overexposure are considered as fundamental problems in a practical recognition system. In this respect, researchers introduce various techniques to deal with illumination changes in the past decades

Satone, M. and Kharate, G [19] proposed a novel algorithm for face recognition in which a low frequency component of the wavelet is used for PCA representation. Best features of PCA are selected using the genetic algorithm (GA). Support vector machine (SVM) and nearest neighbor classifier (ND) are used for classification. Classification accuracy is examined by changing number of training images, number of features and kernel function. Results are evaluated on ORL, FERET, YALE and YALE-B databases. Experiments showed that proposed method gives a better recognition rate than other popular methods.

In the paper of Fernandes, S.nd Bala, J. [20] performance analysis of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for face recognition is made. This analysis is carried out on various current PCA and LDA based face recognition algorithms using standard public databases. Among various PCA algorithms analyzed, Manual face localization used on ORL and SHEFFIELD database consisting of 100 components gives the best face recognition rate of 100%, the next best was 99.70% face recognition rate using PCA based Immune Networks (PCA-IN) on ORL database. Among various LDA algorithms analyzed, Illumination Adaptive Linear Discriminant Analysis (IALDA) gives the best face recognition rate of 98.9% on the CMU PIE database, the next best was 98.125% using Fuzzy Fisherface through a genetic algorithm on ORL database.

Abdullah, M. et all [21] conducted a study to optimize the time complexity of PCA (eigenfaces) that does not affect the recognition performance. The authors minimize the participated eigenvectors which consequently decreases the computational time. A comparison is done to compare the differences between the recognition time in the original algorithm and in the enhanced algorithm. The performance of the original and the enhanced proposed algorithm is tested on face94 face database. Experimental results show that the recognition time is reduced by 35% by applying our proposed enhanced algorithm.

Bruce A. et all [22] PCA/ICA comparisons by systematically testing two ICA algorithms and two ICA architectures against PCA with four different distance measures on two tasks (facial identity and facial expression). In the process, this paper verifies the results of many of the previous comparisons in the literature, and relates them to each other and to this work. The tests show that the FastICA algorithm configured according to ICA architecture II yields the highest performance for identifying faces, while the InfoMax algorithm configured according to ICA architecture II is better for recognizing facial actions. In both cases, PCA performs well, but not as well as the ICA.

1.5 The present study

This study concentrates on a face recognition system under different variations and occlusions based on the use of Principal Component Analysis (PCA). We also investigate the effect of Genetic Algorithms (GA) on PCA to select the optimized subset of features in order to improve the recognition performance of the system. After performing genetic modification on all images of a database by performing: (i) Histogram Equalization (HE) and Mean Variance Normalization (MVN), (ii) Extracting face image features using a global feature extractor motor, (iii) Matching trained and test face images using Manhattan distance measurement (see Appendix B1) and finally (iv) Appling Genetic Algorithm to choose suitable features, MRR results are compared to those of PCA.

In addressing the variation problems, we explore the suitability of using feature selection for recognizing face images. We have also studied the effect of occlusion on face images with and without optimization of feature sets. We have performed different computations on FERET [23], ORL [24] and BANCA [25] database in order to report the robustness of the scheme.

The organization of the thesis is as follows; Chapter 2 describes face image preprocessing; in Chapter 3 the feature extraction method applied is investigated, the details of Genetic Algorithm as a feature selection strategy are represented in Chapter 4 and in Chapter 5 presents application of GA on PCA for several databases. Elaboration on the results is performed in Chapter 6 and finally Chapter 7 concludes the thesis.

1.6 Proposed work

The method is based on several implementations of PCA-GA method on ORL, FERET and BANCA are performed for different poses of traıned and test datasets. The next attempt would be computed the MRR with PCA- GA method for YALE and YALE-B databases for different biometrics variations. Further, it would be interesting to compare the present MRR findings to those obtained with WAVELET-PCA-GA-ND and WAVELET-PCA-GA-SVM method for BANCA database obtained

CHAPTER 2

FUNDAMENTALS OF FACE IMAGE PROCESSING

2.1 Face image preprocessing

In the field of face recognition, image preprocessing and feature extraction have a significant role especially for 2-D images. The individual images acquired by the camera may usually involve pose of head, illumination and occlusion. One of the most important steps in face recognition is preprocessing. In fact, after detecting the faces, the system should specify the head's position, size and pose. All these are necessary when preparing an image for feature extraction and feature representation. Since in this study, we have been working on the databases with 2-D images, therefore the concentration of preprocessing steps involves two-dimensional images. In this chapter a brief summary of the image preprocessing used in the present study is given. Figure 2.1 depicts the steps to prepare a 2-D image before applying features extraction. Usually 2-D based processing systems use image registration, cropping, histogram equalization and mean variance normalization.

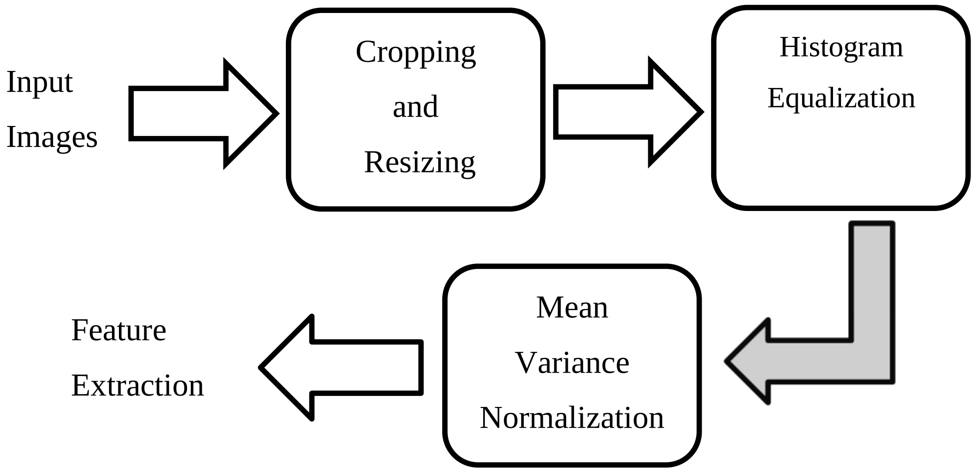
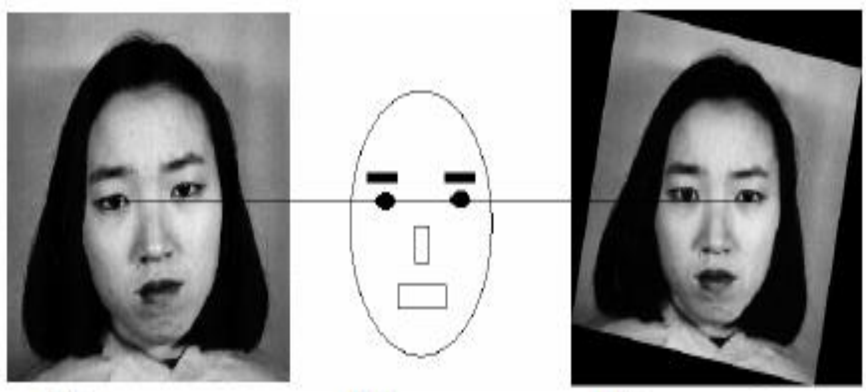


Figure 2.1: Face image preprocessing steps [26]

2.2 Image registration

Generally, image registration step is not considered as the main stage of preprocessing steps. In this stage of processing all the input images are aligned. This stage in general is the preamble to the other preprocessing stages. Different individuals images are captured by the camera are expected to be out of phase with each other due to differences in perspective, terrain or other causes. In fact, this step brings the input image into alignment with the base image by employing a spatial transformation to the input image. This step of image preprocessing is not considered as a necessary stage if a system is able to automatically detect and place position of the facial features in an image for feature extraction. However, in the case of automatic detecting and locating absence, then the image alignment is considered as an important tool in order to extract facial features of two dimensional images successfully. Figure 2.2 (a) depicts the original face image of an individual that is slightly rolled to an angle, it is needed to align this image and other images in the database, therefore different facial features such as both eyes are in the same position with each other. Figure 2.2 (c) shows an aligned face image. In [27] the authors mentioned that different frames of video sequences might contain closed eyes and then it might be unsuitable in some cases to align the eye centers. In case of such problems they suggest to use another position of the eyebrows in facial images instead.



(a) (b) (c)

Figure 2.2: Registration of an original face image (a) using the eye centers of face image (b) to generate an aligned image (c) [27].

2.3 Cropping images and Resizing

Image cropping is considered as another step of image preprocessing. In fact, the face includes the most significant information in an image. The background, the neck region and the hair of a subject in an image can be considered as unnecessary elements for feature extraction. Indeed, including these regions of the images lead to more memory requirements, increase in computation time and subsequently a decrease in recognition performance. Cropping is a kind of preprocessing techniques that is applied to deal with these problems. In general, cropping extracts a portion of the image which includes the face region, but eliminates the neck, hair and background. In addition, we have to resize each face image prior to giving the image to a feature extractor (algorithm) as an input to extract the facial features. Different images may have different size and therefore after cropping the images it is necessary to unify size of all images. Different subjects do not have the same size of the face, therefore due to cropping a different size of the new image is achieved for different subjects. In this case all the images in the database should be resized to a common dimension. Usually, in order to obtain the best performance, all the images in the database resize to the size of the image that has the smallest dimension [28].

2.4 Histogram equalization

In a database, different images may have different number of intensity levels. In addition, congestion of intensities in different levels might be different and these differences among images may lead to decrease the efficiency of face recognition. Histogram equalization (HE) is used as a technique in image processing for contrast adjustment using the image's histogram. Generally, with a low contrast image in a database the effect of applying HE is to spread out the most frequent intensity values. This technique allows the areas of low contrast to obtain a higher contrast without affecting the overall global contrast. In fact, the shape of image histogram is modified after applying histogram equalization technique. HE aims to increase the range of intensity and spreads the intensity distributions. The distributed intensities are therefore better in terms of a histogram since they provide flattened peaks and valleys for an image [28]. This process spreads the contrast of the low contrast parts of an image without affecting the overall contrast of the image. Figure 2.3 demonstrates an image histogram before and after applying HE.

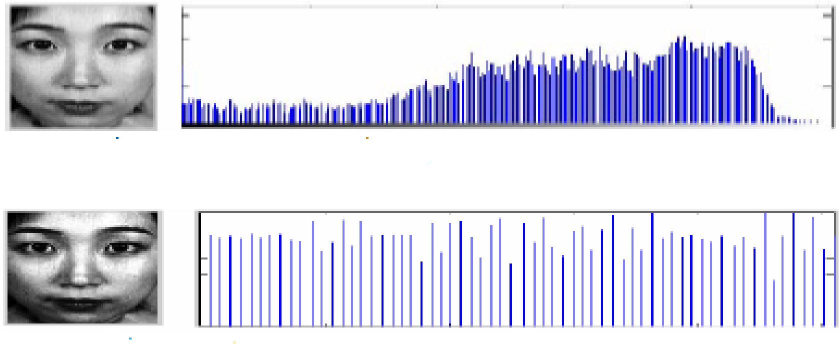


Figure 2.3: Image histogram before and after equalization [28]

For HE the following mathematical formula is used

$$S_K = (L-1)\sum_{j=0}^{k} P_{r(rj)} \tag{2.1}$$

Where L is the total number of possible intensity levels, $k = 0, 1, 2, \ldots, L-1$ and P_r is the estimate of the probability of occurrence of intensity level in an image. Histogram equalization is a function of rounded S_k and N_k where N_k is number of pixels that have intensity value rj, where rj is the intensity level of input image

2.5 Mean-Variance Normalization

Mean and variance normalization (MVN) is considered as a feature normalization technique that is applied in many fields such as` speech recognition, face recognition etc. in [26] employed MVN for increasing the robustness of speech features to noise. In addition, in [29] the authors deal with the problem of illumination variation and pose in face authentication. In general, MVN is typically used to increase the robustness of recognition features [30]. Feature normalization techniques usually try to reduce the actual mismatch between training and testing conditions. MVN is applied to process feature vectors in order to reduce the effect of noise and illumination and then improve the recognition performance. MVN is obtained as follows:

$$R = X - M_x \, , MVN = R/std(R) \qquad (2.2)$$

where X is a matrix consisting of the intensity values of a grayscale image, M_x is mean value of X and *std (R)* is standard deviation of R.

2.6 Preprocessing performed

In this study, prior to applying feature extraction of the image preprocessing steps on all face images in order to reduce the effect of noise, illumination, and this is expected to increase the recognition performance of the system. All the images were cropped in order to remove the noise factors such as background, neck. Then they resized to a common dimension for each database, based on the size of smallest image in the corresponding database. In addition, image registration step was done using Torch3 vision library [31] for BANCA face database to detect and align the face images from video sequences. Histogram equalization and mean variance normalization techniques were employed for increasing the intensity level of images and reducing the effect of illumination and actual mismatch images.

CHAPTER 3

FEATURE EXTRACTOR

3.1 Feature extraction method

After performing preprocessing steps summarized in the previous chapter, an extraction method is needed to extract the face features. Typically, there are two methods, namely holistic and component based methods as feature extractor in the facial expression analysis. The authors in [32, 33] applied holistic methods to extract the facial features. In holistic methods usually the features are extracted from the whole face image. These kinds of algorithms extract plenty of feature information, however, they need large computation time and therefore inappropriate for real-time applications. In addition, producing large dimension of features may lead to redundant information. On the other hand, component based methods are considered as another technique that extract feature information from face images. In fact, component based methods save computation time and therefore, more practical for real-time application, especially in the field of face recognition. However, for these kinds of techniques a prior knowledge of the areas that contain the information for facial expressions is needed. Indeed, without this knowledge some important features are ignored and then the recognition performance of face recognition system may be affected.

In this study, the holistic approach was used to extract the face features; however, in order to remove the redundant information from the extracted features and making the method suitable for real application a Genetic Algorithm based feature selection was applied to keep the optimized set of features. Therefore, the general structure of the feature extraction section contains a dimensionality reduction, feature extraction for feature selection as shown in Figure 3.1.

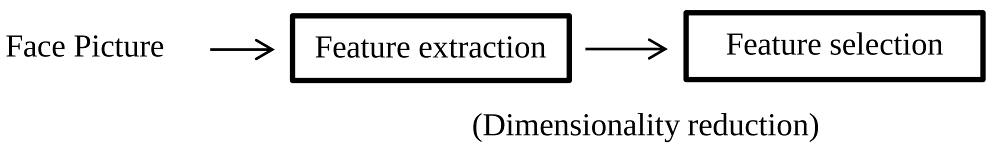


Figure 3.1: Structure of face feature extraction [32, 33]

3.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is considered as a linear non-parametric analysis transform method in the field of pattern recognition. The PCA is known as an efficient technique to extract features successfully in pattern recognition area such as face recognition [34, 35]. This algorithm is used as a simple projection tool in order to reduce a complex data set from a high dimension to a lower dimension.

Generally, the PCA role is to operate directly on whole patterns known as features to explore the global information for consequent recognition using a set of previously found global projectors from a given training pattern set [36]. PCA aims to keep the maximum original pattern information after extracting features, and subsequently it reduces dimensionality [36]. The common steps of PCA algorithm as is explained in [37].

Follows the steps:

I. Collecting I_i images (I_i= [I_1, I_2…, I_M]), where each image is stored in a vector of size L.

II. Mean centering ($\boldsymbol{Y}_i$), the images should be mean centered by subtracting the image mean from each image vector using equation (3.2), where A is the image mean and can be computed using equation (3.1).

$$\boldsymbol{A} = \frac{1}{M} \sum_{i=1}^{M} I_i \tag{3.1}$$

$$\boldsymbol{Y}_i = I_i - \boldsymbol{A} \tag{3.2}$$

Computing the covariance matrix according to equation (3.3).

$$\boldsymbol{C} = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{Y}_i \boldsymbol{Y}_i^T \tag{3.3}$$

III. Specifying the eigenvalues of the covariance matrix using equation (3.4), where E is the set of eigenvectors related to the eigenvalues Λ.

$$\boldsymbol{C}E = \Lambda E \tag{3.4}$$

IV. Sorting the eigenvalues and corresponding eigenvectors in descending order.

V. Projecting each of the centered training images into the created eigenspace based on a new ordered orthogonal basis with the first eigenvector having the direction of the largest variance of the data using equation (3.5), where

E_k's are the eigenvectors corresponding to the Λ significant eigenvectors which are chosen as those with the largest corresponding eigenvalues of $\boldsymbol{C}$ and k varies from 1 to Λ.

$$W_{ik} = (\boldsymbol{E}_k)^T.(\boldsymbol{Y}_i) \qquad \forall i,k \tag{3.5}$$

VI. Recognizing images by projecting each test image I_{test} into the same eigenspace using equation (3.6).

$$W_{(test)k} = (\boldsymbol{E}_k)^T.(I_{test} - \boldsymbol{A}) \quad \forall k \tag{3.6}$$

In fact, dimensionality reduction is considered as a necessary task in any pattern recognition system. In general, we can say the performance of PCA based face recognition system depends on the amount of sample images, number of features and some other factors such as reducing noise, illumination, pose and etc.

3.3 Feature selection method

In order to select the significant facial feature sets in this study, we applied Genetic Algorithm as a feature selection method. [See Section 3.4] In general, based on the relation with the target classifier, feature selection techniques are categorized into two groups; filter and wrapper methods. Feature selection concentration is on areas of application for which datasets with tens or hundreds of thousands of variables (features) are available. Feature selection is different from feature extraction and it is needed to make a distinction between feature extraction and feature selection. A feature extraction algorithm aims to extract the features from the data such as PCA that extract facial features from face in this thesis. A feature extractor usually attempts to build new features from the original data by transformations or combination techniques. In other words, a feature extraction method transforms or combines the data to create an appropriate subspace in the original feature space. [38]

On the other hand, a feature selection algorithm aims to choose the best subset of the input feature set. It attempts to remove irrelevant features and it is often performed after feature extraction. Therefore, features are extracted from the face images, and then an optimal subset of these features is selected by a feature selection method. In this work, after extracting the face features using PCA, the method of Genetic Algorithm was used for the best face from the optimized subset of face images.

3.4 Genetic Algorithm (GA)

Genetic algorithm (GA) has been introduced by John Holland at University of Michigan (1970's) [39] with the aim of presenting effective optimization techniques and machine learning applications in computer science. In this technique directed search algorithms, based on the biological evolution procedures such as inheritance, mutation, natural selection, and recombination (or crossover) are applied. Generally, in GA a populatıon of individuals represented by chromosomes is considered. The individuals in the population then undergo a process of evaluation. The evaluation process denotes an intelligent exploitation of a random search in order to solve optimization problems.

The genetic structure and behavior of chromosomes in a population of individuals is described in [40]. All the individuals in a population attempt to take resources and mates. In this sense, the successful individuals generate more offspring compare to other poorly performed individuals. Propagation of good individual's gens in the population leads to generating the off springs from two good parents and therefore the consecutive generations become more suitable to their environment. In brief, GA starts with a population of size N along with randomly produced chromosomes. It is needed to apply a fitness function on each chromosome of the population. New chromosomes via crossings of selected chromosomes of this population are made and therefore in these chromosomes application of recombination and mutation is needed. In order to keep enough space for insertion the generated new chromosomes and keeping the population with the same N chromosomes, elimination of old population member is needed. The next step is to apply fitness function on the chromosomes and insert them in the population and finally, if the best solution is found or, if the time (or generation number) is finished the chromosome with best fitness result should be returned, otherwise the algorithm continues to find the optimized solution.

The basic computational steps of the Genetic Algorithm based computation process follows the flow chart shown in Figure 3.2.

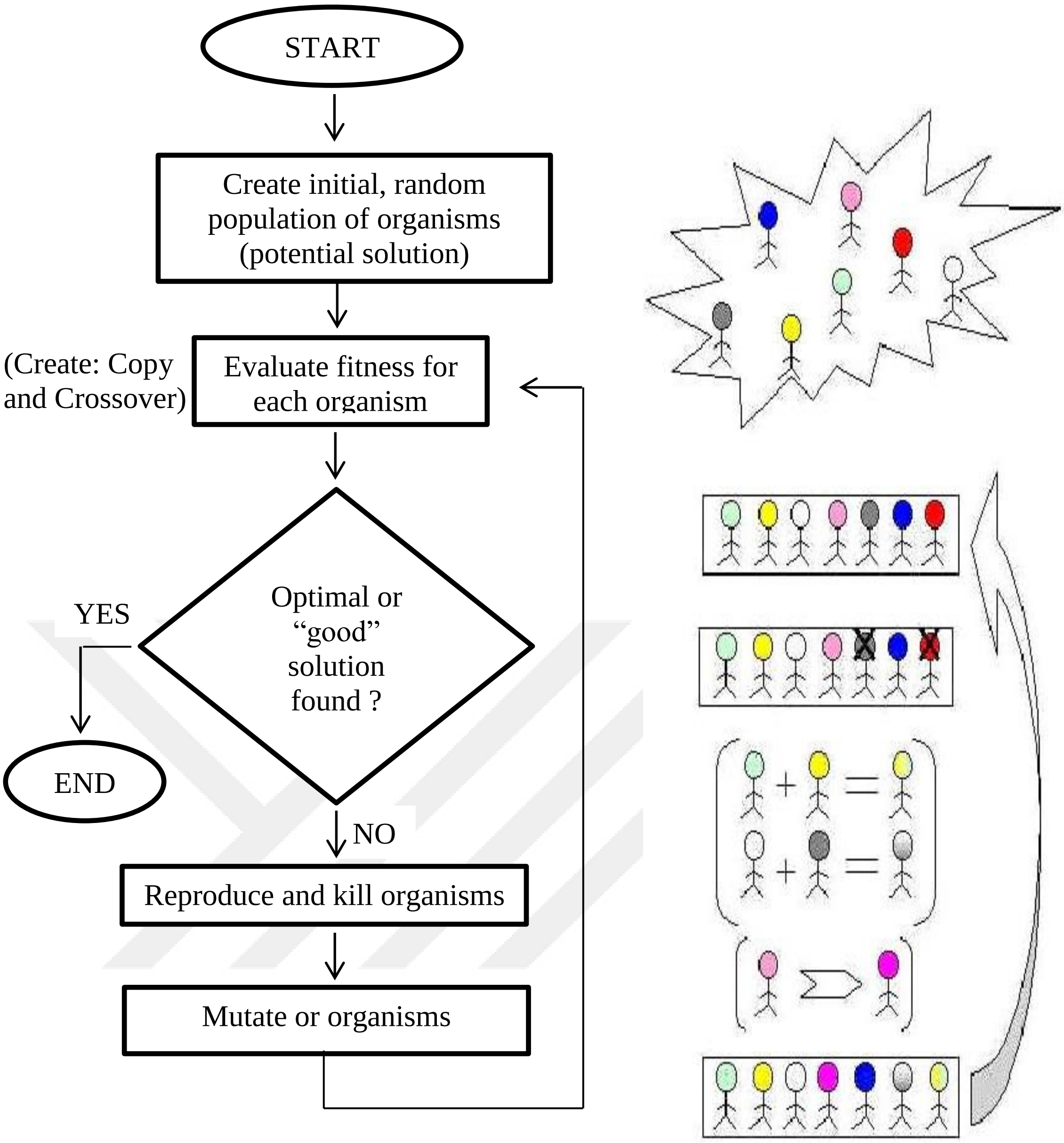


Figure 3.2 The Flow chart of the genetic algorithm [41]

Briefly, GA contains four main stages: evaluation, selection, crossover and mutation. The evaluation process measures the fitness of each individual solution in the population and a relative value corresponding to the definition of optimization criteria is assigned.

i. In the selection process, individuals of the current population are selected randomly for next generation development. Selection of population size is considered probably as the most significant parameter, reflecting the size and complexity of the problem. Usually, the trade-off between additional computational efforts regarding increased population size is a problem specific decision to be determined by the programmers.

ii. In the crossover process selection of two individuals and combination of them about a crossover point is considered into account and as a result two new individuals are created.

Usually, genetic algorithms use 1-point crossover (Figure 3.3) and 2-point crossover operators as the standard procedures in recombination process. Recombination step plays an important role for designing and implementation of robust systems.

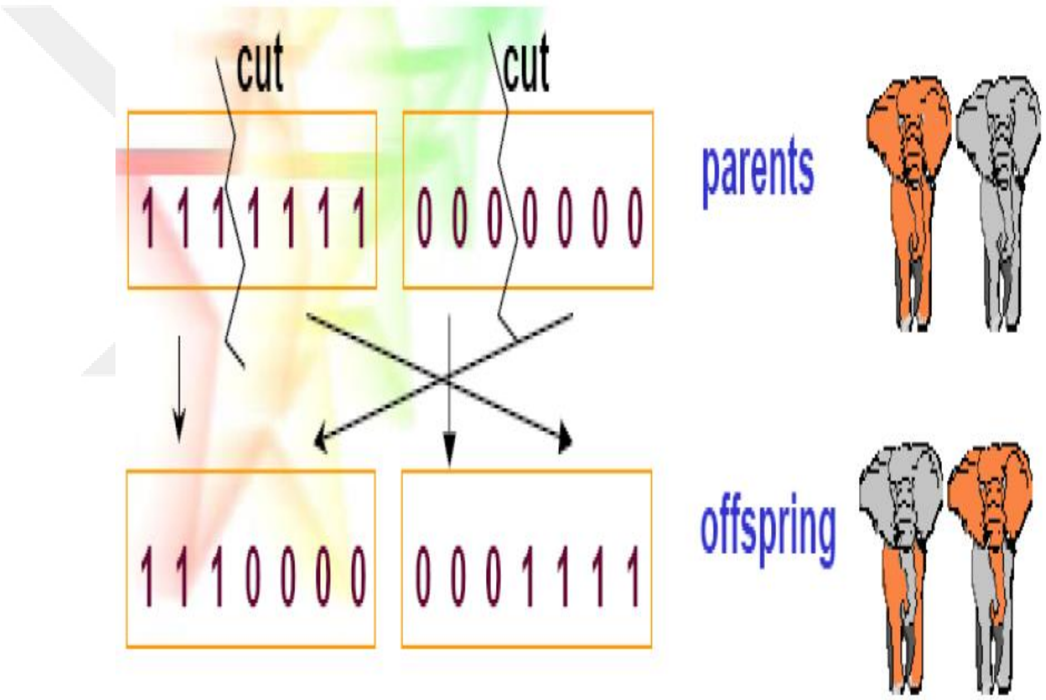


Figure 3.3 Crossover processes [42]

iii. The role of mutation procedure is to randomly modify the genes of an individual subject to a small mutation factor, in order to introduce further randomness into the population. In genetic algorithms, mutation is a genetic operator used to keep genetic variety from one generation of a population of chromosomes to the next. It is considered to be similar to biological mutation [43]. The typical example of a mutation operator includes a probability that an arbitrary bit in a genetic structure is modified from its original state. A typical technique of implementation of the mutation procedure is producing a random variable for each bit in a sequence. This random variable determines whether or not a specific bit will be modified. In fact, the goal of mutation in GAs is to prevent local minima by avoiding the population of chromosomes to be so similar to each other, accordingly slowing or even stopping evolution. Usually mutation is employed to the intermediate population to produce the next population. For instance, Figure 3.4 demonstrates the procedure of mutation operator to obtain a new solution.

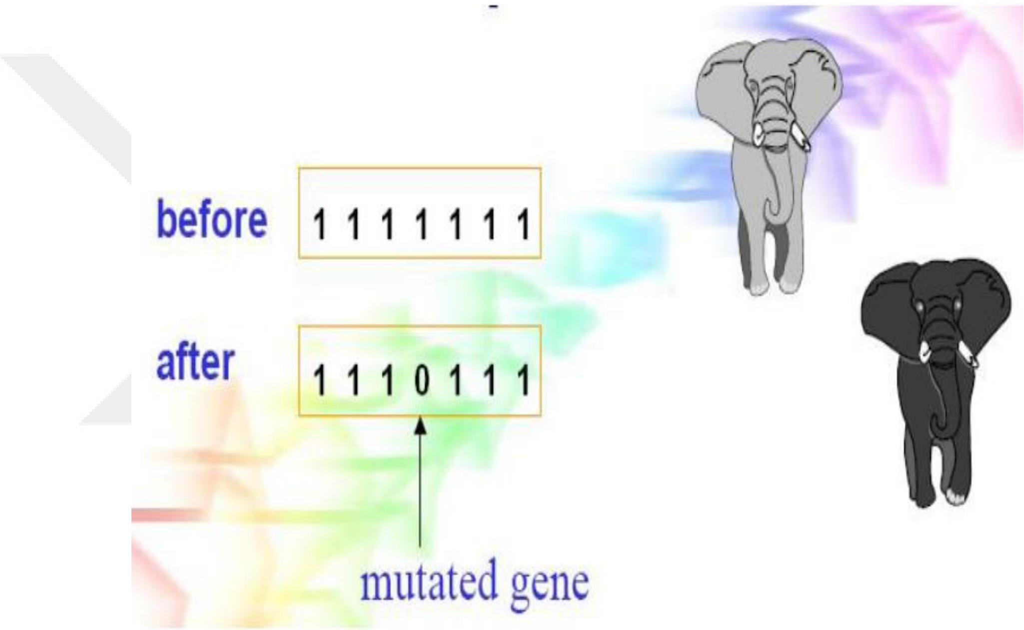


Figure 3.4 Mutation processes [42]

iv. After obtaining one of the possible termination criteria the iterative process of GAs terminates. The termination criteria can be considered as if a known optimal or acceptable solution level is achieved; or if a maximum number of generations have been done; or if a given number of generations without fitness improvement occur.

In the area of GAs consideration of some other parameters including the maximum number of generations, a crossover probability, a mutation probability, a selection method and probably the strategy to be followed is an important issue [42].

3.5 Manhattan Distance (MD)

Manhattan distance computes the distance that one data point to the other. If a grid-like path is followed. the Manhattan distance between two points is the sum of the differences of their corresponding components. The distance between a point $x=(x_1,x_2\ldots x_n)$ and a point $y=(y_1,y_2,\ldots y_n)$ is:

$$MD_{(x,y)} = \sum_{i=1}^{n} | x_i - y_i | \qquad (3.7)$$

Where n is the number of variables, and x_i and y_i are the values of the i^{th} variable, at points x and y respectively [50] .

In the present work, the Manhattan distance is applied to compute the distance score between each train face feature eigenvalue (x_i) and test face feature eigenvalue (y_i) of the corresponding eigenvectors After performing MD operation (Equation 3.7) on the corresponding eigenvalues, the minimum distance (value) among the elements of this MD matrix is used for the identification rate (maximum recognition rate). Thi distance represents the most resemble image for each input data. See Appendix B1 for more details of the Manhattan Distance computations.

CHAPTER 4

DESCRIPTION OF FACE DATABASES

4.1 Face database

In order to test the validity of the implementing face recognition system, we performed several experiments on different subsets of face databases. Face databases employed in this work are FERET [23], ORL [24] and BANCA [25] to test the validity of the face recognition system under different, biometrics conditions in terms of illumination, pose and occlusion in face images. Subsequent subsections have a brief overview on each face database separately.

4.2 AT and T (ORL) database

The computational databases AT and T known as ORL [25] is a standard face database that contains face images of 40 distinct subjects. Each subject has ten different frontal images and they are captured at different times and with a dark homogeneous background. The size of each face image in ORL database is 112×92 pixels. Different variations in facial expression such as open/closed eyes, smiling/non-smiling and scale variations exist in this database images. In this study, all 40 subjects of this database are considered to test the face recognition system. Some sample face images of ORL database are depicted in Figure 4.1.

Figure 4.1 Sample Images of ORL Dataset [25]

4.3 FERET database

The Face Recognition Technology (FERET) database is a standard dataset prepared from 1993 through 1997 in 15 sessions for facial recognition system evaluation by the Defense Advanced Research Projects Agency (DARPA) and the National Institute of Standards and Technology (NIST) [44]. Face images in the FERET database have been captured under semi-controlled conditions. The FERET dataset contains a total number of 14126 face images from 1564 sets of images with involving 1199 individuals and 365 duplicate sets of images [23]. Duplicate sets cover the second image sets of the same individuals captured in different days; it was a gap over two years for taking the images of the same individual in duplicate sets. The dimension of each image is considered as 256×384 pixels.

The features of FERET database under different illumination and pose conditions are: The images in category Fa (1196 pictures) develop the gallery images for four probe sets, Fb, Fc, Dup1 and Dup2. Fb contains 1195 images having variations in expressions, Fc contains 194 images with variations in cameras used and illumination conditions, Dup1 includes 722 pictures which are recorded at different times compared to Fa and Dup2. Dup2 is a subset of Dup1 (234 pictures) with images taken at least 18 months later after the gallery image was taken. Table 4.1 denotes the naming convention based on different categories of FERET database including pose angle, description and number of images and individuals.

Table 4.1 Naming Convention of FERET Database [23].

Two letter code	Pose Angle (degrees)	Description	Number in Database	Number of Subjects
Fa	0 = frontal	Regular facial expression	1762	1010
Fb	0	Alternative facial expression	1518	1009
ba	0	Frontal "b" series	200	200
bj	0	Alternative expression to ba	200	200
bk	0	Different illumination to ba	200	200
bb	+60	Subject faces to his left which is the photographer's right	200	200
bc	+40		200	200
bd	+25		200	200
be	+15		200	200
bf	-15	Subject faces to his right which is the photographer's left	200	200
bg	-25		200	200
Bh	-40		200	200
bi	-60		200	200
ql	-22.5	Quarter left and right	763	508
qr	+22.5		763	508
hl	-67.5	Half left and right	1246	904
hr	+67.5		1298	939
pl	-90	Profile left and right	1318	974
pr	+90		1342	980
Ra	+45	Random images. See note below. Positive angles indicate subject faces to photographer's right	322	264
Rb	+10		322	264
Rc	-10		613	429
Rd	-45		292	238
Re	-80		292	238

Note: See Appendix A.3 for more details of the FERET database codes.(Fa: regular frontal image, Fb: alternative frontal image taken after Fa, ba: the same as Fa, bj: corresponding image of ba and the some as ba, bk: corresponding image of ba taken under different lighting condition)

In this work, we used randomly 100 frontal face images with 4 samples to test our algorithms. Some sample images of FERET database are presented in Figure 4.2.

Figure 4.2 Sample Images of FERET Dataset [23]

4.4 BANCA database

BANCA database is an European project and its aim is to develop a secure system and improve identification, authentication and access control schemes in four different languages (English, French, Italian and Spanish) [24]. In fact, BANCA is a multimodal database with two modalities namely face and voice. In this study, we only used face images to test our face recognition system. The face images in this database were taken under 3 different realistic and challenging operating scenarios. The BANCA database contains 52 subjects, half men and half women. In this database, 12 recording sessions are used for each subject under different conditions and cameras. The data in the first 1-4 trials is captured under controlled conditions and in the trials 5-8 and 9-12 concentrate on Degraded and Adverse scenarios respectively.

Generally, in the face image database, 5 frontal face images are extracted from each recorded video. In order to test the validity of our system, the face images extracted from trials 1 is taken under Controlled conditions are used. Forty subjects of BANCA database with 10 samples are selected randomly to test the algorithms. Figure 4.3 represents a few samples of BANCA database (session 1) face images.

Figure 4.3 Sample Images of BANCA Dataset [24]

In the present study the availability of different variations in facial images of FERET, ORL and BANCA databases applied is summarized in Table 4.2

Table 4.2 Availability of different variations in face databases

Availability Modifications	FACE DATABASES		
	FERET	ORL	BANCA
Pose variations	✓	✓	✓
Illumination	✓	✓	✓
Variations	✓	✓	✓
Facial Expressions	✓	✓	✓
Occlusion (glasses)	✓	✓	✓
Occlusion (mustache)	✓	✓	
Occlusion (beard)		✓	

CHAPTER 5

APPLICATION OF GA FOR FACE RECOGNATION

In this chapter application of both PCA and PCA - GA methods on ORL, FERET and BANCA databases are explained. The databases include 199 eigenvectors (out of 200 eigenvectors) of both trained and test datasets during implementation of PCA.

5.1 Determining number of selected eigenvectors (Features) used in PCA and GA

In the present work, 199 eigenvectors (out of 200 eigenvectors) of each image in the databases were included during implementation of PCA. In order to select the eigenvectors (features) using GA, binary encoding scheme is applied to represent the presence or absence of a specific eigenvector where 1 means presence of the eigenvector and 0 means absence of the eigenvector. Therefore, if the first eigenvector was not selected since the corresponding cell contains 0 and the second eigenvector was selected since its corresponding cell contains 1. A sample row of this new PCA-GA generated eigenmatrix, as selected features (set of eigenvectors), is given in Figure 4.5. The eigenvector at row number 1 was not selected since the corresponding eigenvector is '0', but eigenvector '1' at row number 2 was selected since the corresponding eigenvector is '1. Similarly, the eigenvectors at row numbers 3, 4, 5, 6 were not selected but eigenvector 7 was selected and so on.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	....	200
0	1	0				1	0	0	1	0	1	0	0	0		0

A row of PCA-GA generated eigenmatrix

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	....	200
	1					1			1		1					

Eigenvectors after GA operation

Figure 5.1 An example of eigenvectors selected for the suggested GA

5.2 Structure of the present face recognition system

In the present work PCA is used as a feature extraction method to extract the facial features. Generally, PCA involves consideration of the global information of images by transforming a high dimension feature set to a lower dimension (restraining). In other words, the D-dimensional feature space should be transformed into a K-dimensional subspace where K < D with the purpose of minimizing the loss of relevant information. PCA projects feature vectors on a transformed subspace described by relevant directions and constructed by eigenvectors of the covariance matrix of the data (see Section 3.2). Since of the PCA feature vectors are not relevant for face recognition, depending on the application, in the database of the subspace feature vectors some of eigenvectors should be selected. Usually the number of eigenvectors is specified by sorting all eigenvectors in order of decreasing eigenvalue and then selecting the first L eigenvectors which give an accumulative energy ν above a defined threshold, such as 90% (see Section for L). On the other hand it is known that selecting larger eigenvalues do not essentially guarantee that the resultant eigenvectors involve the "most significant" features in terms of classification [45-49]. Therefore, in the present study we aimed at to find a best set of eigenvectors so that the resultant features have the best discrimination ability.

Genetic Algorithm (GA) was applied to select an optimal subset of extracted eigenvectors as a result of the PCA analysis to enhance the performance of face recognition; each training face image was converted into a vector by row concatenation, and covariance matrix was constructed from a set of training feature image the structure of the recognition is shown in Figure 5.1.

Feature selection is considered as one of the most important steps in pattern recognition to choose eigenvectors. Typically, all the eigenvectors in the eigenspace are not equally informative. The importance of eigenvectors usually is that they are specified as based on the order of the eigenvalues, but this order is not always appropriate to define the data. For example, the first few eigenvectors may contain poses under lighting while some other eigenvectors may contain features such as glasses or mustaches. GA selects an optimal subset of eigenvectors to reduce computation time and to improve the recognition rate. Figure 5.2 displays the basic steps of GA followed in the present work applied for the recognition process.

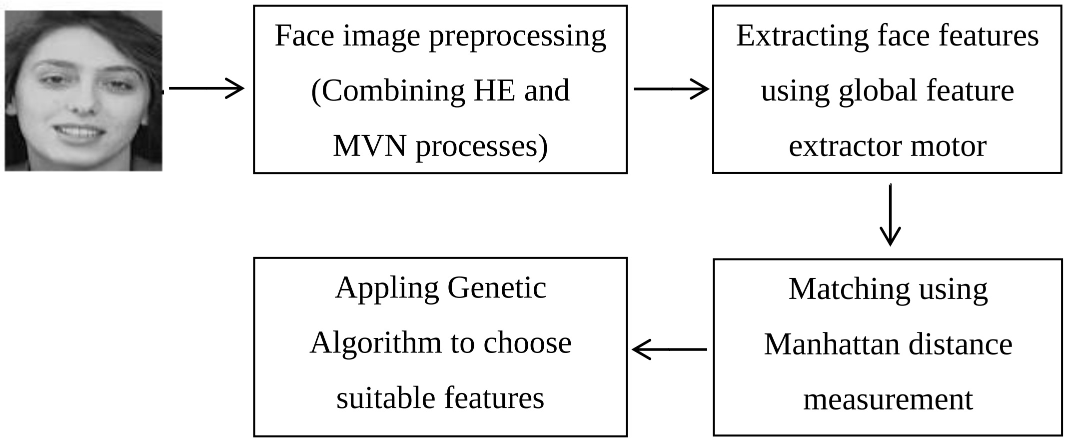


Figure 5.2 Structure of the present face recognition system

In this work, binary encoding scheme is applied to represent the presence or absence of a specific eigenvector where 1 means presence of the eigenvector and 0 means absence of the eigenvector.

The performance of implemented face recognition system was tested on several subsets of three face databases ORL, FERET and BANCA. In ORL face dataset, 10 different frontal face images for 40 different subjects are available. In the present work all 40 subjects of this database were considered; and randomly 5 images were assigned per subject of for training and testing the rest of the images. The parameters of GA set are shown in Table 5.1. During the tests different population between 10 to 30 were selected, and for GA iterations for the mutation and crossover rates were set as 0.05 and 0.95 respectively. To test the optimality of the eigenvalues, the fitness values (Appendix A1) were employed in selection the potential solutions.

Table 5.1 Genetic algorithm parameters obtained with PCA on the subjects of the databases

Parameter	***Value***
Population Sizes	10-15-20-25-30
Number of Iterations	65
Crossover Rate	0.95
Mutation Probability	0.05

CHAPTER 6

RESULTS AND DISCUSSIONS

The procedure to obtain degree of closeness of each train image to its corresponding test image for every individual image of ORL, FERET and BANCA databases are based on computing the maximum recognition rates (MRR). In the following sections the PCA and PCA-GA methods for determining recognition rates for each image in the database were explained and degree of recognition rates obtained together with the number of selected features and the selected population size were illustrated in table format. For all the tests different population sizes ranging from 10 to 30 selected. For both PCA and PCA-GA the number of iterations is limited to 65, and the mutation and crossover rates set to 0.05 and 0.95 respectively. To test the optimality of the eigenvalues, the fitness values (Appendix A1) employed in selection the potential solutions.

6.1 Recognition performances of ORL database determined with PCA and PCA- GA methods

Implementation of PCA and PCA-GA methods on ORL, FERET and BANCA databases to determine recognition performance were performed for different poses of the trained and the test datasets.

To compute recognition performances of ORL databases containing 40 subjects only 10 samples for each individual were selected; 5 samples from the training set and 5 samples from the testing set were randomly selected. The 200 features selected from this database is shown in Table 6.1. In order to compare the training and testing samples each image was resized to 100*200 pixels. The computational steps to perform PCA is given as:

I. For each image collecting 200 images to form the image matrix I_i (I_i= [I_1, I_2…, $I_{M=200}$]), where each image is stored in a vector of size L=20000 has forming 200*20000 pixel I_i matrix

II. Determining the mean image value of 200 subject using Equation (6.1) and subtracting I_i image matrix from the mean image vector $\boldsymbol{A}$ using Equation (6.2) results in the matrix $\boldsymbol{Y}_{i}$, i=200x20000

$$\boldsymbol{A} = \frac{1}{M}\sum_{i=1}^{M} I_i \tag{6.1}$$

$$\boldsymbol{Y}_i = I_i - \boldsymbol{A} \tag{6.2}$$

III. Performing the following covariance matrix operation according to Equation (6.3)

$$\boldsymbol{C} = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{Y}_i \boldsymbol{Y}_i^T \tag{6.3}$$

produces 200x200 order covariance matrix $\boldsymbol{C}$.

IV. *Sorting the* eigenvalues of the covariance matrix $\boldsymbol{C}$ in descending order to form corresponding eigenvectors, then removing the last ‘parasitic’ column produces a sorted covariance matrix $\boldsymbol{C}_s$ for the training ımages.

V. The same procedure described in III is also performed for the test images. maximum similarity between trained and test images, which is listed in 5th column of Table 6.1.

The procedure explained for PCA (Section 3.2) and PCA constrained with GA (Section 5.1) on ORL, FERET and BANCA databases are explained in the following sections.

CASE 1. The first set of computational tests were carried out on all subjects of ORL face database after operating with PCA and GA as given in Table 6.1: The selected features (eigenvectors) achıeved after operating on ORL database are listed on column 3rd and the corresponding recognition rates (accuracies) are given on column 4th. The accuracies achieved with the algorithms in Table 6.1 are based on the selection of maximum number of nonzero eigenvectors and GA-based eigenvector.

In Table 6.1 'Number of Selected Features' is 199 (out of 200 eigenvectors) that is the maximum number of non-zero eigenvectors (features) and therefore, 'Selected Features' column contain all 199 eigenvectors from 1 to 199. This means eigenvectors (features) 1, 2, 3, 4, 5, ….., 199 are used for image recognition. If GA is not applied on PCA 'Selected Population Size' always produces N/A

The recognition rates (accuracy) achieved by PCA algorithm with selection of maximum number of nonzero 199 eigenvectors is 84.50.

CASE 2. To compare MRR for different population sizes PCA-GA method was run for different population sizes and the results are illustrated in Table 6.2: Since the iteration number 65 when population size equals to 10 after 650 (10×65) runs, the best result was achieved for 97 eigenvectors producing MRR of 85.50. When the population size was changed to 15 after 975 (15×65) runs, the best result was obtained for 101 eigenvectors producing MRR of 86.50 and so on. It may be noted that with PCA-GA method, a better MRR reaches to 87.00 after applying PCA-GA feature selection method for 98 eigenvectors for the population size of 20.

Remarkably GA on PCA leads to lower MRR compared to PCA alone which it does not include all maximum non-zero eigenvectors (features).

Table 6.1 Comparison of recognition performances of ORL Dataset -1 in the all obtained with PCA & PCA-GA methods (Case 1)

Method	Number of Selected Features	Selected Features	Selected Population Size	Maximum Recognition Rate
PCA [34-37]	199	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35, 36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67 ,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98, 99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,116,117,118,119,120,121,122, 123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144, 145,146147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167 ,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189, 190,191,192,193,194,195,196,197,198,199	N/A	84.50
PCA-GA	98	2,7,9,16,19,20,21,23,24,25,27,37,38,41,44,45,46,47,49,53,56,59,64,67,68,69,72,73,78,79,80,84,85, 87,89,90,93,96,97,98,101,104,105,106,107,108,109,111,113,115,116,117,122,123,126,127,128, 130,134,139,140,143,145,146,147,148,149,150,152,153,154,155,156,159,161,162,163,164,165, 167,168,169,170,175,178,179,181,183,184,185,188,189,190,191,193,194,195,197	20	87.00

Table 6.2 Recognition performance of ORL Dataset-1 computed for different population sizes using PCA-GA methods (Case 2)

Population Size	Number of Selected Features	Selected Features	Maximum Recognition Rate
10	97	1,2,3,4,5,6,8,11,14,17,18,19,20,21,28,29,30,31,33,36,39,44,45,49,54,55,58,59,62,63,64,66,68,71,72,73,76,78,80,81,82,89,92,95,96,101,102,105,109,111,112,114,116,117,118,119,120,121,125,127,129,130,131,132,135,137,143,144,146,147,148,150,151,153,154,157,159,161,162,163,167,169,170,172,173,177,178,179,181,183,185,187,188,191,193,196,198	85.50
15	101	2,3,6,8,9,10,13,15,16,17,19,21,23,25,26,27,28,29,32,42,44,45,49,50,52,53,56,57,61,62,63,64,65,68,71,74,75,76,79,80,81,84,88,90,91,93,94,96,97,99,101,102,106,107,108,110,112,113,115,116,118,119,120,121,123,126,127,128,132,134,136,138,139,141,144,146,147,148,149,150,151,152,156,158,162,165,166,167,168,169,171,176,177,178,179,180,186,187,191,193,199	86.50
20	98	2,7,9,16,19,20,21,23,24,25,27,37,38,41,44,45,46,47,49,53,56,59,64,67,68,69,72,73,78,79,80,84,85,87,89,90,93,96,97,98,101,104,105,106,107,108,109,111,113,115,116,117,122,123,126,127,128,130,134,139,140,143,145,146,147,148,149,150,152,153,154,155,156,159,161,162,163,164,165,167,168,169,170,175,178,179,181,183,184,185,188,189,190,191,193,194,195,197	87.00
25	115	1,4,5,7,8,12,13,14,16,17,18,22,23,29,32,33,36,38,44,45,46,49,53,54,55,56,58,60,67,70,71,73,74,76,77,81,82,86,88,90,92,93,94,95,97,98,101,102,103,107,108,109,110,112,115,117,118,119,120,121,122,123,125,127,130,131,132,133,134,135,137,138,139,141,142,143,144,145,146,148,149,151,153,155,156,157,159,161,163,164,165,166,169,170,171,172,174,175,176,177,178,179,181,182,183,185,187,189,190,193,194,195,196,197,199	85.50
30	103	1,2,4,7,8,10,11,12,15,18,19,20,22,23,25,28,34,35,39,40,43,44,48,50,51,52,56,58,59,60,62,63,64,65,67,69,71,72,73,74,75,76,78,80,82,87,89,91,98,99,100,103,104,105,106,109,112,115,116,117,118,121,122,125,126,128,129,135,136,137,138,139,144,147,148,149,150,155,157,161,162,163,164,165,167,173,175,176,179,180,182,184,185,186,189,190,193,194,195,196,197,198,199	86.00

CASE 3. The same PCA and PCA-GA methods are applied performed on ORL Dataset-2 having only facial expressions and occlusions in order to investigate the effect of feature selections MRR results. The results are summarized in Table 6.3. For 25 individuals havıng occluded and facial expression produce (25x5)-1=124 features, PCA method results in MRR of 88.00 and surprisingly PCA-GA method does the same for 68 eigenvectors.

Table 6.3 Comparison of recognition performances of ORL Dataset-2 in the presence of facial expression and occlusion obtained with PCA and PCA - GA methods (Case 3).

Method	Number of Selected Features	Selected Features	Selected Population Size	Maximum Recognition Rate
PCA	124	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17, 18,19,20,21,22,23,24,25,26,27,28,29,30, 31,32,33,34,35,36,37,38,39,40,41,42,43, 44,45,46,47,48,49,50,51,52,53,54,55,56, 57,58,59,60,61,62,63,64,65,66,67,68,69, 70,71,72,73,74,75,76,77,78,79,80,81,82, 83,84,85,86,87,88,89,90,91,92,93,94,95, 96,97,98,99,100,101,102,103,104,105, 106,107,108,109,110,111,112,113,114, 116,117,118,119,120,121,122,123,124	N/A	88.00
PCA – GA	68	3,4,7,10,11,12,15,18,20,22,23,26,27, 29,30,31,32,33,35,38,40,41,44,47,48, 50,53,54,55,57,58,60,62,64,65,67,68, 70,71,72,73,75,76,79,82,83,87,88,90, 93,95,96,97,100,101,106,109,110,111, 112,113,114,117,120,121,122,123,124	20	88.00

Here it is clear that the best recognition performance was achieved with the recognition rate of 88.00 in both cases with and without feature selection. Also, when facial expressions exist there is not any improvement in recognition performance. However, we can claim that the same recognition performance is obtained with a small number of eigenvectors as small as 68 which causes lesser memory requirement.

CASE 4. The ORL Dataset-2 containing 25 individuals with 10 samples were selected to generate different population sizes from 10 to 30. MRR becomes 88.00 for every random selection of samples for different population sizes.

Table 6.4 Recognition performance of ORL Dataset-2 for different population size in the presence of facial expression and occlusion (Case 4)

Population Size	Number of Selected Feature	Selected Features	Maximum Recognition Rate
10	62	2,5,6,12,13,14,17,18,20,22,23,24,28,31,33,34,35,36, 37,38,40,42,45,46,47,49,53,54,55,57,60,61,64,66,69, 72,73,74,75,76,78,80,83,84,86,87,92,96,98,100,101, 104,105,107,109,110,111,116,120,123,124,125	88.00
15	61	2,3,6,7,8,9,10,11,12,15,17,19,20,25,28,33,35,36,37, 38,40,43,49,51,52,55,58,59,60,62,64,65,67,68,69,73, 75,78,79,81,83,86,87,91,92,93,96,99,101,102,103, 108,109,112,114,116,117,118,121,123,125	87.20
20	68	3,4,7,10,11,12,15,18,20,22,23,26,27,29,30,31,32,33, 35,38,40,41,44,47,48,50,53,54,55,57,58,60,62,64,65, 67,68,70,71,72,73,75,76,79,82,83,87,88,90,93,95,96, 97,100,101,106,109,110,111,112,113,114,117,120, 121,122,123,124	88.00
25	75	1,3,4,5,9,10,14,15,19,23,26,27,29,30,31,33,36,37,38, 39,40,41,42,44,45,46,47,49,53,54,55,56,57,59,60,63, 65,66,69,70,71,72,73,76,77,78,81,82,84,88,89,91,96, 97,98,99,100,102,103,104,105,106,108,111,113,114, 115,117,118,119,120,121,122,124,125	87.20
30	64	2,4,5,6,7,10,11,13,14,16,17,18,27,28,30,31,34,37,38, 39,40,42,43,46,50,51,52,54,55,59,61,62,63,66,70,71, 73,74,75,78,79,80,82,84,88,89,90,91,92,93,94,96,99, 101,102,103,104,105,108,110,119,120,121,125	88.00

CASE 5. In the next set of tests on ORL face database, we focused on investigation of feature selection using GA for variations in face poses. For this purpose, another subset of ORL face database containing 15 individuals were selected. The results of MRR performance with only PCA and PCA constrained with GA to select the optimized eigenvectors are given in Table 6.5 and MRR performances for different population sizes and for different poses are displayed in Table 6.6.

Table 6.5: Comparison of recognition performances of ORL Dataset -3 in the presence pose variations obtained with PCA & PCA-GA methods (Case 5)

Method	Number of Selected Features	Selected Features	Selected Population Size	Maximum Recognitio n Rate
PCA	74	1,2,3,4,5,6,7,8,9,10,11,12,13,14, 15,16,17,18,19,20,21,22,23,24, 25,26,27,28,29,30,31,32,33,34, 35,36,37,38,39,40,41,42,43,44, 45,46,47,48,49,50,51,52,53,54, 55,56,57,58,59,60,61,62,63,64, 65,66,67,68,69,70,71,72,73,74	N/A	89.333
PCA - GA	29	2,4,6,7,8,11,12,13,19,20,22,23, 32,33,34,37,39,41,44,48,53,62, 65,66,69,71,72,74,75	20	90.667

CASE 6. MRR performance of ORL dataset (/ Dataset-3) for different population sizes in the presence variations in poses. The results of MRR without feature selection was found to be 89.333 while with selecting the optimized subset of the features it was improved to 90.667. In Table 6.6 it can easily be observed the same MRR was also achieved for population sizes 15, 20 and 30. Note that for a population size of 20 that the memory requirements gets smaller.

Table 6.6 Recognition performance of ORL Dataset-3 for different population size in the presence pose variation (Case 6)

Population Size	Number of Selected Features	Selected Features	Maximum Recognition Rate
10	35	4,5,6,7,8,10,13,15,17,18,20,21,24,26,28,29,30, 32,33,35,37,41,42,43,44,45,48,54,55,58,60,62, 63,67,74	88.00
15	38	3,9,11,12,14,15,18,20,23,24,28,29,31,32,33,34, 35,38,39,41,46,47,50,53,55,56,59,60,61,63,64, 65,66,67,68,72,73,74	90.667
20	29	2,4,6,7,8,11,12,13,19,20,22,23,32,33,34,37,39, 41,44,48,53,62,65,66,69,71,72,74,75	90.667
25	37	1,2,4,5,7,8,9,10,13,18,19,20,21,22,31,32,35,36, 37,38,39,42,48,49,50,51,52,54,59,63,65,68,71, 72,73,74,75	89.333
30	44	1,2,3,4,7,10,11,12,14,15,17,18,19,20,21,22,23, 24,26,29,34,37,38,40,41,43,44,47,48,49,50,51, 52,53,56,57,58,60,64,66,70,71,72,75	90.667

6.2 Recognition performances of FERET database determined with PCA and PCA- GA methods

A subset of FERET database for 100 individuals, 4 samples per subject were selected to perform the tests: 2 out of 4 samples were considered as the train set and the rest 2 for the test set. These datasets include all types of face variations and occlusions. The method of analysis of this database is similar to Cases 1 and 2 of ORL databases explained in Section 6.1. Tables 6.7 and 6.8 summarizes MRR performances and corresponding population sizes, and the number of selected features of the FERET database.

In Table 6.8 it is obvious that MRR performance is improved for all population sizes compared to PCA feature extraction method. Further, as the population sizes increases MRR performance is observed to increase. This is an important advantage of implementing face recognition system for large population sizes in terms of reducing computation memory and time.

Table 6.7: Comparison of recognition performances of FERET Dataset -1 obtained with PCA & PCA-GA methods (Case 1)

Method	Number of Selected Features	Selected Features	Selected Population Size	Maximum Recognition Rate
PCA [34-37]	199	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29, 30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54, 55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79, 80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103, 104,105,106,107,108,109,110,111,112,113,114,116,117,118,119,120,121,122, 123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140, 141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158, 159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176, 177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194, 195,196,197,198,199	N/A	88.00
PCA - GA	95	2,3,4,5,7,11,15,17,18,19,26,27,28,29,32,33,34,37,39,40,42,43,45,47,49,51,53, 59,63,64,65,66,68,70,71,72,73,74,76,82,85,91,93,97,99,102,104,105,107,109, 111,112,114,119,123,124,127,129,130,131,139,140,141,143,145,147,152,153, 154,158,159,160,161,162,163,166,167,168,170,172,174,176,178,179,181,182, 183,188,189,193,194,195,196,197,199	30	91.25

Table 6.8: Recognition performance of FERET Dataset-1 computed for different population sizes using PCA-GA methods

Population Size	Number of Selected Features	Selected Features	Maximum Recognition Rates
10	117	1,2,3,4,6,7,8,9,10,12,14,15,18,19,20,21,22,25,26,33,34,35,37,38,40,46,47,51,53,54,55,57,58,59,60,61,64,65,66,67,68,69,70,74,76,78,79,80,81,83,84,88,89,92,93,95,96,97,98,99,100,102,103,104,105,107,109,111,112,113,115,117,118,121,123,124,127,129,130,131,132,135,137,138,139,140,141,145,148,150,151,152,154,158,159,160,161,162,163,167,170,172,175,177,178,179,180,181,182,183,186,188,192,194,196,198,200	88.75
15	106	2,3,6,9,10,11,12,13,15,16,17,18,19,23,24,25,29,33,34,35,36,38,39,40,41,42,45,46,47,48,52,53,56,57,59,61,64,66,67,70,72,73,75,77,79,81,82,83,84,85,86,88,91,93,94,96,97,100,101,102,105,106,109,112,114,116,117,118,119,120,121,130,131,133,135,139,140,141,144,147,148,150,153,157,158,160,162,163,164,165,166,167,168,170,172,174,175,178,179,180,183,186,188,190,195,200	90.00
20	106	1,3,5,6,7,10,11,12,13,14,15,17,21,22,23,24,26,29,31,32,36,37,41,42,43,44,45,46,48,49,50,51,52,55,59,62,64,72,73,75,76,78,79,80,82,85,90,92,93,95,97,98,100,101,102,107,112,113,118,122,123,124,125,128,129,132,133,135,139,140,141,142,143,144,145,146,148,149,151,152,153,156,157,158,159,160,161,163,164,165,168,170,173,175,176,178,179,182,183,184,185,186,191,195,197,199	88.75
25	116	3,4,5,6,7,8,10,12,13,15,18,20,21,22,25,27,28,29,30,31,32,33,34,35,36,38,40,46,50,51,54,55,58,61,62,63,65,66,67,69,70,71,72,75,76,78,80,81,82,85,90,92,94,95,98,100,101,104,105,106,107,110,111,112,113,117,118,119,121,122,123,124,125,126,127,135,138,139,140,141,143,144,145,150,151,153,155,156,158,160,161,164,165,166,167,168,169,170,171,172,175,178,179,181,183,184,185,187,189,190,191,193,194,195,196,198	90.00
30	95	2,3,4,5,7,11,15,17,18,19,26,27,28,29,32,33,34,37,39,40,42,43,45,47,49,51,53,59,63,64,65,66,68,70,71,72,73,74,76,82,85,91,93,97,99,102,104,105,107,109,111,112,114,119,123,124,127,129,130,131,139,140,141,143,145,147,152,153,154,158,159,160,161,162,163,166,167,168,170,172,174,176,178,179,181,182,183,188,189,193,194,195,196,197,199	91.25

6.3 Recognition performances of BANCA database determined with PCA and PCA- GA methods

The last set of tests deals with the extracted face images from BANCA video sequences. In the population size of 40 subjects, 10 samples per subject were selected to perform the tests: 5 out of 10 samples were taken as the train and the rest 5 for the test set. These dataset include all types of face variations and occlusions. The method of analysis of this database is similar to Cases 1 and Case 2 of ORL database as explained in Section 6.1. Tables 6.9 and 6.10 summarize MRR performances for the selected features of BANCA with the corresponding population sizes.

For this dataset PCA-GA method improves MRR performance to 100 for the population size of 25 out of 30 as displayed in Table 6.10.

If one compares MRR results given in Tables 6.9 and 6.10, MRR achieved with PCA method is 97.00 while the feature selection obtained with PCA-GA method improves MRR to 100.

The test results show that the feature selection strategy when PCA is constrained with GA improves MRR performance even up to 100. This is because of the result of removing the eigenvectors containing noise.

Table 6.9: Comparison of recognition performances of BANCA Dataset -1 obtained with PCA & PCA-GA methods

Method	Number of Selected Features	Selected Features	Selected Population Size	Maximum Recognition Rate
PCA [34-37]	199	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32, 33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61, 62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90, 91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113, 114,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134, 135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154, 155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174, 175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194, 195,196,197,198,199	N/A	97.00
PCA-GA	109	2,3,4,13,14,16,18,19,20,21,23,25,26,27,29,30,31,32,34,36,39,40,41,42,44,45,48,49,50, 51,53,54,55,56,55,64,65,67,69,71,74,75,77,79,81,82,83,84,86,89,92,93,97,100,102,105, 107,109,112,116,117,118,122,123,124,125,126,128,129,132,135,137,140,141,142,143, 145,146,147,150,152,154,155,157,160,162,163,165,166,167,168,169,170,171,173,174, 175,177,178,179,180,183,184,188,190,191,192,194,200	25	100.00

Table 6.10: Recognition performance of BANCA Dataset-1 computed for different population sizes using PCA-GA methods

Population Size	Number of Selected Features	Selected Features	Maximum Recognition Rates
10	116	2,4,6,7,11,13,15,16,17,20,21,24,25,27,28,30,32,33,34,36,37,39,40,44,45,46,47,48,49,53,54,56,58,62,64,67,68,69,70,73,74,78,79,81,84,87,90,92,93,94,95,96,98,100,102,105,107,108,110,111,114,116,117,118,121,123,124,125,126,127,128,129,131,132,133,137,139,140,141,142,143,145,146,147,148,151,153,154,155,160,161,163,164,165,166,167,168,170,172,173,174,175,176,179,182,183,184,185,187,191,192,193,194,197,199,200	100.00
15	112	1,3,4,6,7,8,9,12,13,16,17,18,19,20,21,22,23,24,25,26,27,28,30,32,33,34,37,38,41,42,43,44,45,53,54,57,58,61,70,71,72,74,76,77,81,82,83,86,87,90,91,92,94,95,99,102,103,104,105,107,108,110,111,112,114,115,116,117,118,120,121,122,124,125,128,130,131,133,135,137,141,144,145,148,149,150,151,155,156,158,160,166,168,169,170,173,174,177,178,181,185,186,187,188,189,192,193,194,195,196,197,200	100.00
20	114	3,4,7,10,14,15,16,18,19,20,21,23,25,29,30,31,32,33,35,36,40,41,45,46,48,50,51,53,55,58,59,60,61,62,63,65,66,69,72,76,80,83,84,86,89,90,91,92,95,98,99,100,101,104,109,110,111,114,115,116,117,118,119,120,121,122,123,124,125,127,130,131,132,133,137,139,140,141,143,144,146,148,149,152,153,154,155,157,158,159,160,161,163,167,168,169,172,174,176,177,179,180,182,183,184,186,187,188,189,190,193,195,196,200	100.00
25	109	2,3,4,13,14,16,18,19,20,21,23,25,26,27,29,30,31,32,34,36,39,40,41,42,44,45,48,49,50,51,53,54,55,56,57,64,65,67,69,71,74,75,77,79,81,82,83,84,86,89,92,93,97,100,102,105,107,109,112,116,117,118,122,123,124,125,126,128,129,132,135,137,140,141,142,143,145,146,147,150,152,154,155,157,160,162,163,165,166,167,168,169,170,171,173,174,175,177,178,179,180,183,184,188,190,191,192,194,200	100.00
30	110	1,4,5,7,8,9,10,12,14,15,17,18,19,20,21,22,23,24,27,30,31,33,36,41,43,46,48,49,50,52,53,57,58,61,62,63,64,65,68,70,71,74,76,78,79,81,84,87,89,91,92,96,97,99,102,105,106,107,109,110,111,112,114,115,119,122,123,124,125,129,130,131,132,139,141,143,144,146,149,151,153,154,156,160,161,164,165,167,168,169,171,173,175,176,178,179,180,181,182,184,185,186,188,190,191,192,197,198,199,200	100.00

6.4 Comparision of recognition performances of the present study to those computed in the literature

In the current work, the performance of PCA constrained with GA applied on ORL, FERET and BANCA databases in terms of MRR were compared to the other works in the literature. The present results and the results obtained with GA-constrained algorithms, WAVELET-PCA-GA-ND and WAVELET-PCA-GA-SVM methods in the literature, are illustrated in Table 6.11.

One of the interesting findings of the present work is that when PCA is constrained with GA for ORL and FERET databases, MRR results are close to the results computed with WAVELET-PCA-GA-ND and WAVELET-PCA-GA-SVM methods. Because of the simplicity of the present methodology it is expected to achieve considerable saving in computer memory and time. Further, MRR results is also observed to be comparable when WAVELET-PCA-GA-ND method is applied for YALE and YALE-B databases.

The contributive results of the present computation is that for BANCA database, MRR is improved as good as to 100 percent as is computed with WAVELET-PCA-GA-SVM method [19] for YALE and YALE-B databases.

Table 6.11 Comparison of recognition performances (MRR) of the present study to those computed in the literature

ALGORITHMS	**DATABASES**				
	ORL	**FERET**	**BANCA**	**YALE**	**YALE-B**
PCA	84.50 [34-37]	88.00 [34-37]	97.00 [34-37]	-	-
PCA-GA	***87.00***	***91.25***	***100.00***	-	-
WAVELET-PCA-GA-ND	94.3 [19]	94.1 [19]	-	93.3 [19]	97.5 [19]
WAVELET-PCA-GA-SVM	97.3 [19]	98 [19]	-	100 [19]	100 [19]

CHAPTER 7

CONCLUSION

Face recognition is one of the important challenges in appearance-based pattern recognition field. This technology has emerged as an attractive solution to address many new needs for identification and verification of identity. Many factors affect the face recognition performance like facial expression, illumination, pose or occlusion.

In this study, as a feature selection method (to select the optimized subset of features) the effect of GA on PCA is worked out on ORL, FERET and BANCA databases which contain face biometrics under different illumination variations, poses and partial occlusions. During image preprocessing, the histogram equalization and the mean-variance and normalization methods are implemented to reduce the effects of biometric variations.

Analysis of the results obtained for both the trained and the test databases shows that the feature selection strategy is found to improve MRR by constraining PCA with GA for ORL, FERET and BANCA databases.

The interesting findings of the present work is that when PCA is constrained with GA for ORL and FERET databases, MRR results are close to the results computed with WAVELET-PCA-GA-ND and WAVELET-PCA-GA-SVM methods. The contributive results of the present computations are that for BANCA database MRR is improved to be as good as to 100 percent, which is the same as that computed in references [19] with WAVELET-PCA-GA-SVM method for YALE and YALE-B databases.

REFRENCES

[1] A.K., R. Bolle, S. Pankanti, eds.: Biometrics: Personal Identification in Networked Security. Kluwer Academic Publishers, 1999.

[2] G. Koltzsch, "Biometrics-Market Segments and Applications", Journal of Business Economics and Management, Vol. 8, no. 2, pp. 119-122, 2007.

[3] Biometrics History. NSTC. Home page, http://www.biometrics.gov/ Documents /BioHistory.pdf.

[4] History of Fingerprinting. FINGERPRINTING. Home page, http://www. Fingerprinting.com/history-of-finger printing.

[5] J. Meaney., History of Fingerprints Timelines, Home page, http://www. Fingerprintamerica.com/fingerprinthistory.asp.

[6] Wikipedia, the free encyclopedia, Home Page, http://en.wikipedia.org/wiki/ Biometric.

[7] Division of Criminal Justice Services. Taking bertillon measures, 2006, Home Page, http://criminaljustice.state.ny.us/ojis/history/ph-measr.htm.

[8] X. Zhang and Y. Gao, "Face recognition across pose: A review", Pattern Recognition. Vol. 42, No. 11, pp. 2876-2896, 2009

[9] A. K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition", IEE Transaction on Circuits and Systems for Video Technology, Vol. 14, no. 1, pp. 4-20, 2004.

[10] A. K. Jain, A. Ross, and S. Prabhakar, "Biometrics: A Tool for Information Security", IEE Transaction on Information Forensis and Security, Vol. 1, pp. 125-143, 2006.

[11] J. L. Wayman, “Fundamentals of biometric authentication technologies,” Int. J. Image Graphics, vol. 1, no. 1, pp. 93–113, 2001

[12] A.M. Martinez, “Recognizing Imprecisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class”. IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24. pp. 748-763, 2002.

[13] H. J. Oh, K.M Lee, S. U Lee, “Occlusion invariant face recognition using selective LNMF basis images”. ACCV (Asian Conference on Computer Vision) 2006, Hyderabad, India, pp.120-129, 2006.

[14] F. Tarrés, A. Rama, L. Torres, “A Novel Method for Face Recognition under Partial Occlusion or Facial Expression Variations”. 47th International Symposium ELMAR-2005, Multimedia Systems and Applications, Zadar, Croatia, 2005.

[15] S. Harin, J. Sabah, “Face Recognition in the Presence of Expression and/or Illumination Variation”. Fourth IEEE Workshop on Automatic Identification Advanced Technologies, pp. 144-148, 2005.

[16] H. Han, S. Shan, X. Chen, and W. Gao, “A comparative study on illumination preprocessing in face recognition,” Pattern Recognit., vol. 46, no. 6, pp. 1691–1699, 2013.

[17] L. Qingshan, Y. Wang, L. Hanqing, M. Songde, “Occlusion Robust Face Recognition with Dynamic Similarity Features”. 18th International Conference on Pattern Recognition, 2006. ICPR 2006 Volume 3, pp. 544 – 547, 2006.

[18] P. Corné van der, “Evolutionary Multi-Objective Optimization and Preference Modeling in Green Logistics”, Erasmus University Rotterdam, Economics & Informatics Erasmus School of Economics, 2012.

[19] M. Satone and G. Kharate, “Feature Selection Using Genetic Algorithm for Face Recognition Based on PCA, Wavelet and SVM” Inernational Journal on Electrical Engineering and Informatics - Volume 6, Number 1,2014.

[20] S. Fernandes and J. Bala, "Performance Analysis of PCA-based and LDA based Algorithms for Face Recognition" International Journal of Signal Processing Systems Vol. 1, 2013.

[21] M. Abdullah, M. Wazzan and S. Bo-saeed "OPTIMIZING FACE RECOGNITION USING PCA" International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, 2012.

[22] A. Bruce, K. Baek, M. Stewart and J. Beveridgea "Recognizing Faces with PCA and ICA" Computer Vision and Image Understanding 91, pp. 115–137, 2003.

[23] P.J. Philips, H. Wechsler, J. Huang, P. Rauss "The FERET database and evaluation procedure for face recognition algorithms, Image and Vision Computing, Vol.16, pp.295–306, 1998.

[24] E. Bailly-B, S. Bengio, F. Bimbot, M. Hamouz, J., Kittler, J., Mariethoz, J., Matas, K., Messer, V., Popovici, F. Poree, B. Ruiz and J.P. Thiran, "The BANCA database and evaluation protocol", In Audio and Video-Based Biometric Person Authentication: Proc. 4th Int. Conf., AVBPA2003, LNCS 2688, pp. 625–638, 2003.

[25] AT&T Laboratories Cambridge, The ORL Database of Faces: http://www.cam-orl.co.uk/facedatabase.html.

[26] P. Pujol, D. Macho, C. Nadeu, "On Real-Time Mean-and-Variance Normalization of Speech Recognition Features". IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. Volume 1, 2006.

[27] Z. Gengtao, Z. Yongzhao, Z. Jianming, "Facial Expression Recognition Based on Selective Feature Extraction". Sixth International Conference on Intelligent Systems Design and Applications, 2006. ISDA '06. Volume 2, pp. 412-417, 2006.

[28] Y. Fu and T.S. Huang, 'Human Age Estimation with Regression on Discriminative Aging Manifold', IEEE Transactions on Multimedia, 10(4), pp. 578-584, 2008.

[29] S. Malassiotis, and M.G Strintzis, “Combining 2D and 3D Face Images for Relaible Identity Verification”. Proceedings of European Signal Processing Conference (EUSIPCO), 2005.

[30] M. G. Rhodes, ‘Age Estimation of Faces: A Review’, Appl. Cognitive Psychology. 23, pp. 1–12,2009.

[31] Torch3vision, Retrieved, from http://torch3vision.idiap.ch

[32] W. Fellenz, J. Taylor, N. Tsapatsoulis, and S., Kollias, “Comparing template-based, feature-based and supervised classification of facial expressions from static images”. Proceeding of Circuits, System, Communications and Computers, WSES Press, Nugate, Japan, pp. 5331-5336, 1999.

[33] C. Lisetti, and D. Rumelhart, “Facial expression recognition using a neural network”. Proceedings of the 11th International Flairs Conference, AAAI Press, New York, pp. 328-332, 1998.

[34] T. Young, K.-S. Fu, “Handbook of Pattern Recognition and Image Processing”, Academic Press, 1986.

[35] M. Turk and A. Pentland, “Eigenfaces for Recognition”, Journal of Cognitive Neuroscience, Vol.3, No.1, pp. 71-96, 1991.

[36] S. Chen and Y. Zhu, “Subpattern-based principle component analysis”, Pattern Recognition, Vol.3, pp. 1081-1083, 2004.

[37] W. S. Yambor, Analysis of PCA-based and Fisher Discriminant-based Image Recognition Algorithms, Master thesis, Colorado State University,2000.

[38] L. Talavera, “An evaluation of filter and wrapper methods for feature selection in categorical clustering” Lecture Notes in Computer Science, pp. 440-451, 2005.

[39] M. Mitchell, “Genetic Algorithms: An Overview ”, pp. 31–39, 1995.

[40] http://www.cs.ucc.ie/~dgb/courses/tai/notes/handout12.pdf

[41] O. Derya, “Feature Selection for Face Recognition Using a Genetic Algorithm”, “,Bilkent University, Department of Computer Engineering, 2006.

[42] S. Omid, Ü. Ahmet and A. Adnan, "Evolutionary Multi-Objective Optimization for Nurse Scheduling Problem", Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control (ICSCCW), 2-4 September 2009.

[43] Wikipedia Encyclopedia (n.d). Mutation Operator (Genetic algotithms). http://en.wikipedia.org/wiki/Mutation_(genetic_algorithm).

[44] http://www.nist.gov/mml/mmsd/20151215head.cfm

[45] S. Manisha and G.K. Kharate, "Selection of Eigenvectors for Face Recognition", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.3, 2013.

[46] W. Yambor, B. Draper, and R. Beveridge, "Analyzing PCA-based Face Recognition algorithms: Eigenvector Selection and Distance Measures", 2nd Workshop on Empirical Evaluation in Computer Vision,2000.

[47] D. Valentin et al, "Principal Component and Neural Network Analyses of Face Images: What can be Generalized in Gender Classification?", Journal of Mathematical Psychology, vol. 41, pp. 398-413, 1997.

[48] T. Xiang and S. Gong, "Spectral clustering with eigenvector selection," Pattern Recognition, vol. 41, pp. 1012-1029, 2008.

[49] F. Zhao, et al., "Spectral clustering with eigenvector selection based on entropy ranking," Neurocomputing, vol. 73, pp. 1704-1717, 2010.

[50] H. Abul, et al., " Comparative Study Of Distance Metrics For Finding Skin Color Similarity Of Two Color Facial Images", National Conference on Advancement of Computing in Engineering Research, 2013.

APPENDIX A

A.1 Fitness Function and Fitness Value in GA:

Generally, a "fitness function" in GA is responsible to evaluate the solutions by returning a "fitness value", which reflects the optimality of solutions. Therefore, the fitness values are employed in selection process to select the potential solutions for next generation.

In this work we used the recognition rate as a fitness function F(X) as follow:

$$F(X) = Recognition\ Rate = \frac{N_{accepted}}{N_{test}}$$

where $N_{accepted}$ is number of successful recognition and N_{test} is the number of all testing images in the database. The aim of this fitness function is to select the features with maximum recognition rate. Then in each generation, the features with maximum fitness values are selected to be submitted as local optimum solutions to the next generation. Finally, the best fitness value of all generations is selected as the optimal solution based on the recognition performance or rate.

A.2 Genetic Algorithm pseudo code:

The genetic structure and behavior of chromosomes in a population of individuals is described as [40]. All the individuals in a population attempt to take resources and mates. In this sense, the successful individuals generate more offspring compare to other poorly performed individuals. Propagation of good individual's gens in the population leads to generating the off springs from two good parents and therefore the consecutive generations become more suitable to their environment. Figure 9.1 show the GA pseudo code and process, where n is the number of individuals in the population; χ is the fraction of the population to be replaced by crossover in each iteration; and μ is the mutation rate. Generally, n, χ and μ is given as input to GA for generation 0, in order to initial the population of n random individuals and evaluate the fitness function. Then the best fineness value along with the selected gens are passed to next generations via a do loop. Crossover and mutation are applied in each generation on the selected gens to propagate the good individuals. Finally, the best fitness value and gens are returned.

Algorithm: GA(n, χ, μ)

// Initialise generation 0:
$k := 0$;
$P_k :=$ a population of n randomly-generated individuals;
// Evaluate P_k:
Compute *fitness*(i) for each $i \in P_k$;
do
{ **// Create generation $k+1$:**
// 1. Copy:
Select $(1-\chi) \times n$ members of P_k and insert into P_{k+1};
// 2. Crossover:
Select $\chi \times n$ members of P_k; pair them up; produce offspring; insert the offspring into P_{k+1};
// 3. Mutate:
Select $\mu \times n$ members of P_{k+1}; invert a randomly-selected bit in each;
// Evaluate P_{k+1}:
Compute *fitness*(i) for each $i \in P_k$;
// Increment:
$k := k + 1$;
}
while fitness of fittest individual in P_k is not high enough;
return the fittest individual from P_k;

Figure 9.1 Genetic Algorithm pseudo code

A.3 Explanation of shorthand letters for FERET database:

Two lowercase character strings (shorthand letters) in FERET database indicates the kind of imagery [23].

1. fa indicates a regular frontal image
2. fb indicates an alternative frontal image, taken seconds after the corresponding fa
3. ba is a frontal image which is entirely analogous to the fa series

4. bj is an alternative frontal image, corresponding to a ba image, and analogous to the fb image
5. bk is also a frontal image corresponding to ba, but taken under different lighting
6. bb through bi is a series of images taken with the express intention of investigating pose angle effects (see below). Specifically, bf - bi are symmetric analogues of bb - be.
7. ra through re are “random” orientations. Their precise angle is unknown. It appears that the pose angles are random but consistent. The pose angles in the table were derived by manual measurement of inter-eye distances in the image, and in their corresponding frontal image.

APPENDIX B

B.1 Manhattan Distance :

The Manhattan distance computes the sum of difference in each dimension of two vectors in n dimensional vector space. It is the sum of the absolute differences of their corresponding components. Manhattan distance is also called the L_1 distance. If $u = (x_1, x_2 \ldots x_n)$ and $v = (y_1, y_2 \ldots y_n)$ are two vectors in n dimensional hyper plane, then the Manhattan Distance $MD(u, v)$ between two vectors u, v is given by the **Eq. 1.**

$$\boldsymbol{MD(u,v)} = |\boldsymbol{x_1} - \boldsymbol{y_1}| + |\boldsymbol{x_2} - \boldsymbol{y_2}| + \cdots + |\boldsymbol{x_n} - \boldsymbol{y_n}| \qquad (1)$$

$$\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} |\boldsymbol{x_i} - \boldsymbol{y_i}|$$

Now for two RGB scale images of size $p \times q$, $I_1(a,b,c)$ and $I_2(a,b,c)$ where $a = 1,2...,p$, $b = 1,2..,q$ and $c = 1,2,3$ where c represents color intensity values Red, Green, Blue respectively. Manhattan Distance is measured using **Eq. 2.**

$$\boldsymbol{MD(I_1, I_2)} = \sum_{a=1}^{p}\sum_{b=1}^{q}\sum_{c=1}^{3} |I_1(a,b,c) - I_2(a,b,c)| \qquad (2)$$

As the number of pixels, n which falls in skin region varies with varying size of the image,so rather than taking the absolute distance further the distance is being normalized using **Eq. 3.**

$$\boldsymbol{MD_1(I_1, I_2)} = \frac{\boldsymbol{MD(I_1, I_2)}}{n} \qquad (3)$$

where n= number of pixels considered.

Manhattan distance between skin regions of the images shown in Fig 9.2(a) and Fig. 9.2(b) is 66.2244.

Figure 9.2. Manhattan distance between images=66.2244. (a) Facial image 1 (b)Facial Image 2 [50]