FACE RECOGNITION UNDER PARTIAL OCCLUSION AND SMALL DENSE NOISE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ELECTRONIC SYSTEMS AND COMMUNICATIONS

BY

ROHIT KUMAR

ROLL NO. -212EE1210

Department of Electrical Engineering

National Institute of Technology, Rourkela-769008

2014

FACE RECOGNITION UNDER PARTIAL OCCLUSION AND SMALL DENSE NOISE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ELECTRONIC SYSTEMS AND COMMUNICATIONS

BY

ROHIT KUMAR

ROLL NO- 212EE1210

SUPERVISED BY

DR.DIPTI PATRA

Department of Electrical Engineering

National Institute of Technology, Rourkela-769008

2014

National Institute Of Technology, Rourkela

CERTIFICATE

This is to certify that the thesis entitled, "**Face Recognition under Partial Occlusion and Small Dense Noise**" submitted by **Mr. Rohit Kumar** in partial fulfillment of the requirements for the award of Master of Technology in **Electrical Engineering** with specialization in **"Electronic Systems & Communication"** at National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University / Institute for the award of any Degree or Diploma.

Date:

Prof. Dipti Patra

Department of Electrical Engineering

National Institute of Technology, Rourkela

ACKNOWLEDGEMENT

This project is by far the most significant accomplishment in my life and it would have been impossible without people who provided me with an excellent professional environment and supported throughout in accomplishing my project.

I would like to extend my respect and my sincere thanks to my honorable, esteemed supervisor ***Prof.*** **Dipti Patra,** *Department of Electrical Engineering. She is not only a great lecturer, but most importantly a generous person. I sincerely thank her for the kind of exemplary guidance and encouragement as a guide. Her continuous reliance and support inspired me in taking the right decisions and I am very much elated to work under her guidance.*

I am very much grateful to our Head of the Department, **Prof. Anup Kumar Panda,** *for providing us with best facilities in the department and his timely suggestions and all my teachers* **Prof. Susmita Das, Prof. K. Ratna Subhashini, Prof. Prasanna Kumar Sahu, Prof. Supratim Gupta** *for providing a solid background for my studies. They truly have been great sources of inspiration to me and I thank all of them from the bottom of my heart.*

I would like to thank all my friends and particularly my classmates for all the innovative and mind boggling discussions we had, which made us think to a different level all together. I've rejoiced their companionship very much during my stay at NIT, Rourkela.

I would like to thank my parents, who taught me the values and ethics of hard work by setting their own example. They have been an enormous source of inspiration instead of being apart during my stay in NIT Rourkela.

In the end, I would like to thank all those who made my stay in Rourkela an unforgettable and memorable experience.

Rohit Kumar

ABSTRACT

Problem of automatic recognition of human faces from front views with varying expression, illumination, occlusion as well as disguise is considered. Here the problem of recognition is cast as one of the several classifying linear regression models and argued that in handling such problems a new theory using sparse representation of signals is the key. A face recognition algorithm is also introduced which uses ‘L1-minimization’ theory of optimization. This proposed concept handles two crucial problems of face recognition, which are, feature extraction and robust occlusion handling. For extraction of features, PCA is used, but later in the thesis it is shown that if sparsity is properly calculated in the face representation, selection of features doesn’t remain crucial.

However, the number of extracting features is crucial here. Another crucial factor is the authenticity of calculating sparse coefficients. Unconventional feature extraction techniques such as down-sampled images and random projections give results comparable to common features like Eigenfaces, as long as the dimension of the feature space exceeds a particular threshold, predicted by the sparse representation theory. This can handle errors because of occlusion and consistently by using the fact that these errors are frequently sparse with respect to the standard basis. The sparse representation theory helps in predicting that how much of occlusion can be handled using this recognition algorithm and how can the training images be selected so that robustness to occlusion can be maximized. A Number of experiments on freely accessible facial databases are performed to justify the efficiency of the proposed algorithm and the above claims.

Contents

List of figures

Abbreviations

PCA-Principal Component Analysis

SVD- Singular Value Decomposition

SVM- Support Vector Machine

SPCA- Sparse PCA

CCTV- Closed-circuit television

LDA- Latent Dirichlet allocation

PCs- Principal Components

COV- Covariance Matrix

List of Tables

CHAPTER 1

INTRODUCTION

1.1 : Face Recognition

Human Face is a complex, multidimensional structure which requires efficient computing techniques for the recognition process. The face has been our priority and focus of attention in playing an important role in identifying an individual's face. We recognize a large number of faces learned throughout the life and recognize those faces at first glance even after several years. There may be some variations in faces because of aging and factors like beard, hair-style or even change of glasses. Face recognition is also vital in biometrics. In biometrics basic properties of human faces are matched to the existing data and depending on the result, the identification of an individual is confirmed. Features of the facial databases are extracted and implemented through different efficient algorithms and required valuable changes are done to improve the existing algorithms**.** Computers which distinguish and perceive confronts could be connected to a large number of real world applications like criminal distinguishing proof, security frameworks, character check and so on. All the face recognition systems basically find the identity of any given face image after comparing it to a database or memory. Such memory is formed using a training set of face images. Face detection and recognition is utilized as a part of numerous places these days, in sites facilitating pictures and person to person communication destinations. Face recognition and recognition might be attained utilizing innovations identified with software engineering. Characteristics separated from a face are transformed and contrasted and compared with faces present in the dictionary. On the off chance that a face is remembered, it is known or the framework may demonstrate a comparable image already existing in the formed else it is obscure. In a surveillance system if an obscure face seems more than once, then it is put away in a database for further recognition. These steps are exceptionally helpful in identification of mischievous persons. When all is said in done, face recognition strategies might be isolated into two parts focused around the face representation they utilize appearance-based, which utilizes all-encompassing composition characteristics and is connected to either entire face or particular locales in a face picture and characteristic based, which utilizes geometric facial characteristics (mouth, eyes, eyebrows, cheeks and so forth.) and geometric connections between them.

1.2: Motivation to the work

With the advancements of machine innovation, individuals abuse these advancements for negative purposes. They mask themselves to beguile the security framework, subsequently influencing the execution. Acknowledging face recognition framework, individuals change their looks to trap the security framework by blanketing the face with a scarf or hand**.** Literature studies uncover that faces could be perceived in a limited environment with high exactness. Yet in true situations it is as of now testing, for example, enlightenment, posture variety and impediments need to be overcome in which impediments, for example, sunglasses, scarf and so forth is more essential. Consequently, diverse procedures must be received to tackle the issues**.** Numerous authors have handled the issue of partial occlusion, which is portrayed in the later segments. Dealing with face recognition under controlled situations has been on scene for the past numerous years, however recognition under uncontrolled conditions like enlightenment, interpretation and fractional occlusion are a late issue. An extraordinary measure of work has been carried out to handle recognition under changing outflows and lighting conditions**.**

The fractional occlusion influences the local characteristics, yet the recognition methods might be made powerful if these local characteristics are united together intelligently. Martinez utilized robust recognition with fractional occlusion by blending neighborhood characteristics based on similarities. SVM can provide robustness if neighborhood characteristics are treated with it. Face recognition methods, whether linear or nonlinear are characterized into three groups handling occlusion in face images, i.e., characteristic based methods that deal with characteristics like eyes, mouth, nose and build a geometrical correspondence between them. The second classification is appearance-based methods that focus on the holistic features of face images by acknowledging the entire face region and third class deals with the hybrid local and global features of face images to be utilized for recognition purpose. In view of these classes an overview is directed to analyze every individual system in taking care of the partial occlusion and the improvements made by different creators to handle the issue. Additionally recorded in the content are the databases on which tests were directed and results were concentrated in the wake of performing the tests.

While there are several general available techniques for recognizing frontal face images, which perform very well under controlled environments. But these tend to fail under uncontrolled environments like sharp illumination changes, Partial occlusion and large variation in poses. The whole work in this thesis has been concentrated upon this issue only, more specifically on the problem of partial occlusion. This can be very helpful in the applications like terrorist recognition, surveillance where the subject intends to disguise the technical systems and CCTV cameras.

1.3: Challenges in face recognition

Face recognition is sensitive under the conditions written below

- Large variation in pose
- Drastic change in illumination
- Face under Partial occlusion

1.3.1: Large variation in pose

Fig 1.1: Examples of images with large changes in pose

1.3.2: Drastic change in illumination

Fig 1.2**:** Examples of images with drastic illumination changes

1.3.3: Partial Occlusion

Hindrance in the view of an image refers to Occlusion. It may be natural, as well as synthetic. Natural hindrance refers to hindrance of perspectives between the two picture objects without any intension while manufactured hindrances refer to a fake barricade of purposefully blanket the picture's perspective with a white/dark solid rectangular piece. Fractional occlusion has been found in numerous areas of picture handling. It is seen in iris recognition where the eyelashes impede the iris; distinguishing proof through ear can likewise be impeded by the ornaments. Indeed continuously requisition face picture gets blocked by means of extra accessories (sunglasses/scarf/ hair or even by hand). Other than biometric picture processing, it is additionally experienced in the medicinal field where the supply routes may be blocked because of elevated cholesterol level.

When there is drastic change in the environment or target face is under partial occlusion, the recognition of the faces becomes a tedious task. Previously proposed methods and algorithm fail to make an impact under such challenging conditions. To make the recognition process robust, there is need of algorithm which can tackle these challenges well.

Fig 1.3**:** Examples of Partially Occluded images

1.4: Literature survey

Automatically recognizing faces has been and still is a challenging research field of computer vision, machine learning and biometrics applications. Capturing images for database preparation depends on the specific application, for example, in surveillance purpose a video camera is used to capture the facial images. Hence, depending upon the acquisition of facial data, techniques of face recognition are mainly divided into three categories: method dealing with intensity images, methods dealing with images from a video sequence and methods which deals with images from any sensor like infrared images or 3D images.

Enlarged surveys give some light on these methods falling under above categories and try to give an idea about some of their advantages and also their drawbacks in general [1,2].

Methods for face recognition, which deals with intensity images mainly come under two categories as follows: feature-based and holistic [3-5]. Approaches based on features first process the input images and extract distinct features from them such as eyes, nose, mouth etc. Features from other facial marks come handy. Later computers try to find the geometric relation among these features and thus the image is reduced to a feature vector. Using these measurements, standard recognition schemes try to find the exact match of the images. Earlier works related to automated face recognition were also based upon these techniques. Kanade[6] made one of such attempts, he employed general methods of image processing for extraction of feature vector of 16 facial parameters. Those parameters were dependent upon areas, distances and angles so that varying size of images could be compensated. He then used Euclidean distance for purpose of matching.

More advanced feature extraction techniques these days are involving deformable templates [7], [8], [9], Hough transform methods [10], Reisfeld's symmetry operator [11] and Graf's filtering and morphological operations [12]. However, all of such techniques heavily rely on factors such as restricting the search subspace with geometrical constraints [13]. In addition to that, a must tolerance must be introduced into the models since they seldom fit perfectly the structures in the image. The main drawback of such methods is the difficulty of automatic feature detection (as discussed above) and the fact that the implementer of any of these techniques has to make arbitrary decisions about which features are important [14].

Holistic approaches use a global representation of images and then try to identify images. i.e they use entire image specifications for extracting features rather than only using local features of the images: statistical and AI approaches. In a simple version from the holistic approaches, the image is made to be a 2D array of intensity values and recognition is further done by comparing direct correlation values between the test face and all the other training faces in the dictionary. Even though this method performs under some limited environments [15] (i.e., equal illumination, scale, pose, etc.), it is computationally this is not efficient and lacks a straightforward correlation-based approach, such as sensitivity to face orientation, size, illumination changes, background and noise etc.[16]. The main shortcoming of the direct matching methods' efficiency is that they try to perform classifications in very high dimensions [17].

To manage this issue of dimensionality, many other schemes have been proposed which incorporates dimensionality reducing techniques to get and have the most dominant feature dimensions ahead of face matching. Sirovich and Kirby [18] were the first to use Principal Components Analysis (PCA) [19, 20] to face images representation in lower dimension. They showed that any specific facial image can be easily represented along a different coordinate space known as Eigen space, and any face can be easily reconstructed by using mere a small number of eigenvectors and the corresponding projecting along each Eigen picture. Turk and Pentland [21, 22] later felt, based on Sirovich and Kirby's findings, that projections of images along Eigen pictures could be used as a tool to extract features and recognizing faces. They managed to make a face recognition system that forms Eigen faces, which are the eigenvectors of the associated covariance matrix formed by known face (patterns), and then tried to recognize specific faces by comparing each of their projections along the respective Eigen faces of the known faces of many individuals. The Eigen faces forms a feature space that drastically diminishes the dimension of the original space, and further process of face recognition is done in the lower dimension sub-space.

A measure of work has been carried out to handle recognition under changing outflows and illumination conditions, using methods like PCA, LDA, neural networks and several variations of them are used but each has its limitations. Although these are successful in many applications, but they don't give good results when the face image gets partially occluded. Local features in the image gets affected by the partial occlusion but the recognition can be made efficient if those local features are managed prudently. In order to overcome this problem sparse PCA [24] has been used in this work. Will,Todd[23] later introduced SVD theory where they introduced about Principal components calculation in a different way. They showed that any matrix, square or not, of any dimension can be represented as products of three matrices. Here they showed that principal components can be directly calculated using this theory that too in sorted format. J. Wright, A. Yang, A. Ganesh, S. S. Sastry[24] introduced Sparse representation of images so that subject can be represented by less number of elements. Using this theory they showed that the occlusion and noises could be easily compensated if the sparsity is properly harnessed. They further used this theory for facial recognition application.

1.5: Objective of the Thesis

- To extract the features of the face image using Principal component analysis.
- To reduce the dimension of image using PCA for fast computation and memory conservation.
- Further reduction of the dimension of test images in sparse domain.
- Compensation for the occlusion and illumination changes in test image using L1-norm minimization for robust face recognition.

1.6: Thesis contribution

- Dictionary of facial images was created and its features using PCA were extracted successfully.
- The dimension of the images was reduced after projecting images over a feature domain.
- Test image was represented in sparse domain and occlusion was handled using a trivial template.
- Sparse coefficients were calculated using L1-minimization and occlusions were compensated efficiently for robust recognition.

CHAPTER 2

Background theory

2.1: Eigenvalues and Eigenvectors

- Eigenvalues measure the amount of variation (information) explained by each principal component and will be larger for the first PC and smaller for the subsequent PCs.
- An eigenvalue greater than 1 indicates that principal component accounts for more variance than accounted by one of the original variables in standardized data. This could be used in thresholding of data, which is later used to decide the required number of eigenvectors.
- Eigenvectors provide the weights to compute the uncorrelated principal components, which are the linear combination of the original variables.

2.2 Principal component analysis (PCA)

Principal component analysis (PCA) is very general data-processing and dimension reduction technique, with so many applications in engineering, biology, and social science. Briefly, PCA is used to find important contributors of any data. Thus, instead of taking all possible contributors to a result, only few important ones are used. Eventually a lot of calculation is reduced for the remaining important contributors. A large amount of memory used is also saved while performing analysis.

It is a standard mathematical tool for analyzing data and extracting relevant information from the data sets. It helps in converting the complex data into a low-dimensional data set and helps in revealing the underlying information in it. It was used as a feature extraction algorithm in this project. PCA has a lot of applications in data analysis as follows:

2.2.1: Applications

- Used for compression and classification of data.
- The motive is to reduce the number of variables, at the same time, retain most of the information.
- New variables known as principal components (PCs) are mutually uncorrelated, and are ordered according to the amount of information they contain.
- PC's are a series of linear least squares fits to a sample, each orthogonal to all previous.
- Reduces the dimension of the dataset.
- Decreases the redundancy of the data.
- Filters part of the noise from the data.
- Used to prepare the data for further analysis under several techniques.

2.2.2: Disadvantages of PCA

- The components are not independent but uncorrelated. It would be even better if we have a representation in which components are independent to each other.
- PCA seeks for linear combinations of the original variables. The nonlinear combination may even yield better representation. PCA has an extension for doing this type of analysis, Nonlinear PCA.
- Instead of L2 norm, it may be advantageous to use L1 norm. Especially, if the signal that we want to represent is sparse or has a sparse representation in some

other space. PCA is extended for this specific problem as well, which is called Sparse PCA.

2.3: Object representation based on PCA (principal component analysis)

According to PCA theory, the data are analyzed to find the most important elements and structure to remove noise and redundant part, and thus dimension of data is reduced.

2.3.1: Steps to calculate Eigen faces

- Get data matrix [A]
- Subtract mean from each column of data matrix to get [B].
- Find out covariance matrix [COV].

  $[COV] = (1/n\text{-}1)\ B*B^T$

- Calculate eigenvalues and eigenvectors of [COV].
- Arrange the eigenvectors in order of priority basis (vector corresponding to highest eigenvalue is placed along 1^{st} column and so on) along columns of a matrix [U]; this matrix is called Eigen-basis.

But calculating Eigen basis matrix for face dictionaries can be a hard task because of dimensionality problem. For example if there are 100 faces in training set of dimension 200×200,then after reshaping each image into a column vector and concatenating those to form a matrix will result into a matrix of dimension [40000×100].Thus dimension of covariance matrix[COV] will be 40000×40000,which will result into 40000 Eigen-basis columns. Handling such a large dimension matrix is still beyond the capacity of normal computers available.

Hence, for making the computation easy, we calculate eigenvector of $[B^T*B]$ instead of$[B*B^T]$, because the dimension of prior one reduces to just 100×100 which makes the computation faster.

As per the theory of PCA, by using these 100 eigenvectors we can calculate the real Eigen basis vectors of Eigenbasis matrix;

$$U_i = B*V_i$$

Thus the dimension of Eigen-basis matrix [U] becomes [40000×100], where each column if reshaped back into the dimension of the original image will result in a distorted face kind of image, hence these columns are also called as Eigenfaces.In this way reduction in the dimension is achieved without losing much information about the facial images.

These so called 100 vectors are predominant vectors out of those 40000 vectors and carry most of the information about faces in the training set. This way, we save memory and enhance our calculation speed without much compromising with the object information.

2.4: Image compression using PCA

Suppose we have a mean adjusted image matrix $[A]_{m\times n}$ .Covariance matrix of the adjusted image matrix is $[cov(A)]_{m\times m}$. Eigen vector of covariance matrix is calculated and its sorted columns are known as “Principal Components of image matrix”.

Now, we take r columns out of total m columns of principal components, such that $r<m$.

Data is projected in lower dimension as;

$$[\mathrm{Y}]_{r*n} = [V]^{T}_{r*m} * [X]_{m*n} \tag{2.1}$$

Now original data is obtained as;

$$[X]_{m*n} = [V]_{m*r} * [Y]_{r*n} \tag{2.2}$$

In this way we obtain a data of the same size as the original one, but later have less information. These both take the same amount of memory on a computer, but computational speed gets enhanced in a later case because this is the result of multiplication of two low dimensional matrices.

Amount of information in reconstructing an image depends on the number of Principal Components included in the projection. It has been seen that most amount of image information lies with only first few eigenvectors or principal components of Eigen basis Matrix.

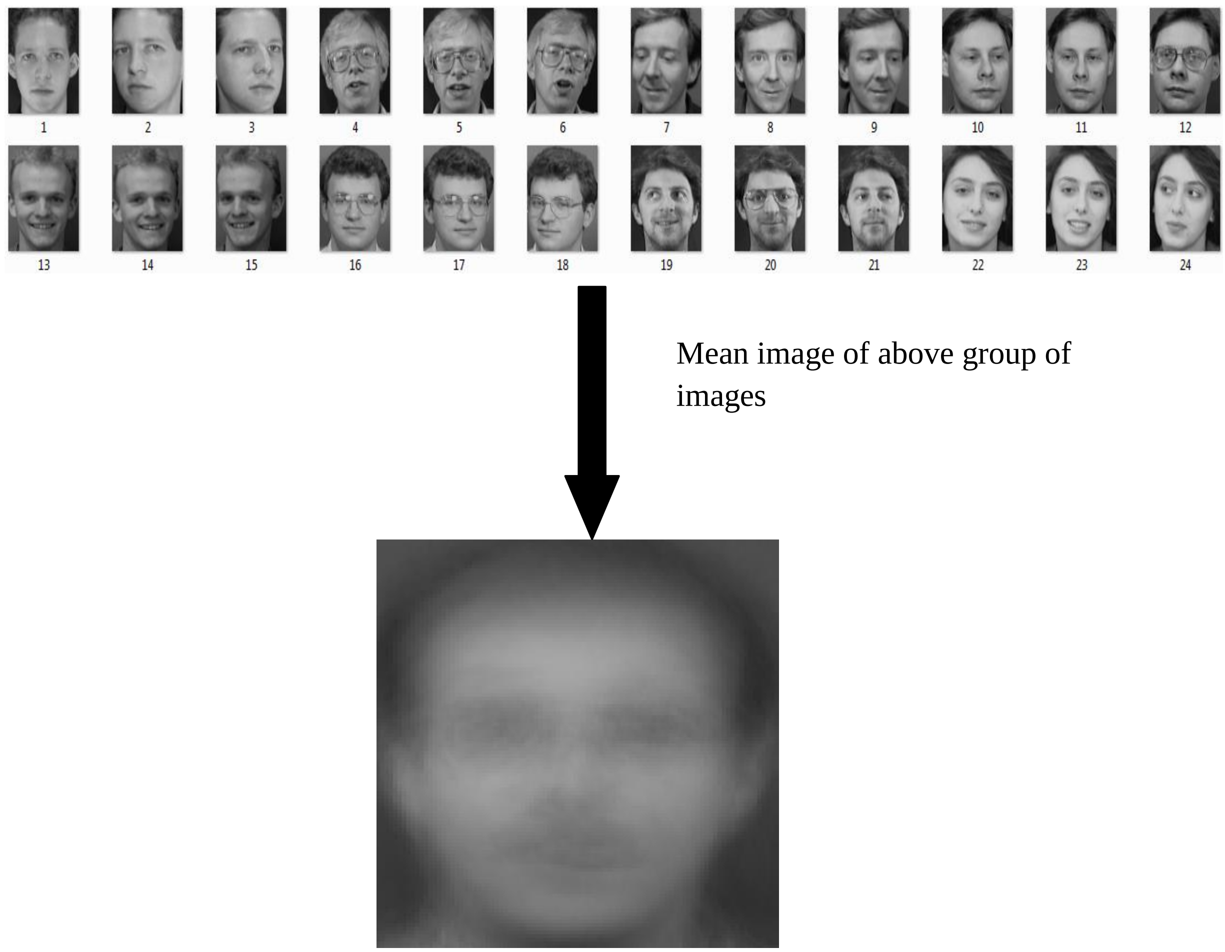

Fig 2.1: Set of training images and their mean image

2.5: Singular Value Decomposition (SVD)

SVD is one of common way to perform Principal Component Analysis (PCA). It is a matrix diagonalization procedure that allows to "diagonalizable" any matrix– square or not square, invertible or not invertible.

$$[U\ \Sigma\ V] = SVD\ (B) \tag{2.3}$$

Any matrix $[B]$ of dimension m*n can be decomposed and written in a form of

$$B = U * \Sigma * V^T \tag{2.4}$$

$U \in R_{m*n}$, $V \in R_{m*n}$ are both Orthogonal to each other and normalized matrices, whereas $\sum$ is diagonal matrix. $\|U\| = 1$ and $\|V\| = 1$, i.e both matrices are unitary.

- The $[U]$ and $[V]$ contain Eigen-vectors in order of highest to lowest variance.
- Highest variance vector suggests that the vector incorporates those features that change the most.
- Eigen-vectors are orthogonal to each other. That means two Eigen-vectors don't share the same features.
- Eigen-values of $[\Sigma]$ explain the importance of each respective Eigen-vector.

  $\Sigma = Diag(\sigma_1, \sigma_2, \sigma_3 \dots \dots \dots \sigma_n)$; Such that, $\sigma_1 \geq \sigma_2 \geq \dots . \geq \sigma_n$. These are called singular values of $[B]$ and are calculated by performing square root over the Eigen values ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$) of matrix $[B]^T[B]$.

Columns of V are eigenvectors of $[B]^T[B]$ and and columns of U are calculated as;

$$U_i = (\frac{1}{\sigma_i}) B * V_i \qquad (2.5)$$

In short, SVD tells us that which specific vector will give the most "diverse" result from the dataset, and comparatively less important vectors in descending orders. Which means that the first column of the matrix $[U]$ is the most important; the second column is the second most important, and so on.

CHAPTER 3

Feature Extraction

3.1: Introduction

Human facial characteristics end up being critical for face recognition and examination. By study it has been resolved that eyes, mouth, and nose is amongst the most prevailing characteristics for facial recognition. Perceiving an individual from facial characteristics makes the methodology of recognition more mechanized. It is worth noting that in light of the fact that few different frameworks use the spatial geometry of recognizing facial characteristics, they don't utilize hair styles, facial hair, or other comparable components. Facial recognition is by and large utilized for police work. Case in point, open security, terrorist suspects and missing persons.

Facial characteristic extraction has a few issues which must be thought and be understood. A few issues of facial characteristic extraction are given as follows: Small varieties of face size and introduction might be affecting the result. As the information picture originates from the webcam in the room condition the caught picture has distinctive splendor, shadows and clearness which could fail the procedure. Frequently facial characteristics may be covered by different things, for example, a cap, a glass, hand or hairs. Human faces have an amount of feelings by numerous diverse expressions, however, this framework can distinguish the corner of the characteristics on account of neutral, tragic, joyful and shock. Most facial characteristic extraction techniques are touchy to different non-ideals, for example, variation in light, corruption, introduction, time intensive and color space utilized. A great characteristic extraction will additionally expand the execution of face recognition framework.

There are many accessible procedures for the extraction of characteristics from a facial picture, for example,

- Color features
- Shape features
- Texture features

Here we have used Eigen basis as our feature extraction technique, which uses Principal component analysis techniques for the purpose. PCs can be calculated using eigenvectors of covariance matrix or there is another way of calculating it, which is Singular value decomposition. The main difference between general computation and using SVD is that SVD sorts the eigenvectors automatically while we need to sort these vectors manually later in the prior case.

3.2: Eigen basis extraction

According to the SVD theorem of matrix; any data matrix $[A] \in R_{m*n}$, can be decomposed as,

$$[A] = U * \Sigma * V^T \tag{3.1}$$

$U \in R_{m*n}, V \in R_{m*n}$ are Orthogonal and normalized unitary matrices & Σ is a diagonal matrix. $\Sigma_1 = Diag(\sigma_1, \sigma_2, \sigma_3 \dots \dots \dots \sigma_r)$;, such that, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$, these are called the singular values of $[A]$ and are calculated by performing square root over the Eigen values $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r)$ of matrix $[A]^T A$.Our matrix of interest is $[U]$, which is used to represent the object. Suppose we have an image matrix X of size $N_1 \times N_2$, we concatenate the window matrix to form a column vector I of size $d \times 1$ ($d = N_1 \times N_2$).

For ‘n’ images from different classes, we construct a n-column vector, A= [I1, I2, I3……In] (d<<n).

$$[A] \in R_{d*n}$$

$[A]$ is our data matrix.

We can get Eigen basis vector $[U]$ by computing the SVD of the centered data matrix, the columns of the data matrix are equal to respective training images minus their mean ($\bar{I}$).

3.3: Eigenfaces

PCA has a very good application which is in the computer vision domain, called Eigen faces. Eigen face is a name for eigenvectors which are the components of the face itself. It has been used for face recognition where the most variations considered as important. It has been quite successful in face recognition application for a couple of decades. When any eigenvector of the Eigen basis matrix is reshaped back into the dimension of original images, a distorted face like image is obtained, that is why such reshaped vectors are known as Eigen faces. These look like similar to that of a witch like face.

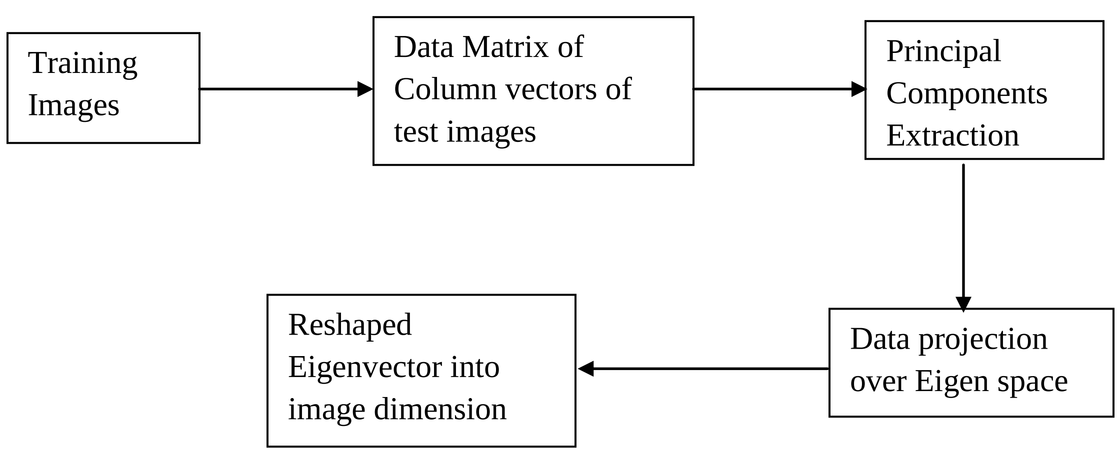

Fig 3.1: Flow chart for Eigen faces extraction

CHAPTER 4

Sparse Modelling of Test Image

4.1: Sparse Representation of a Test Image

4.1.1: SPARSE REPRESENTATION

A signal or image is said to be sparse if it contains a very small number of non-zero values. A signal could be sparse in all domains some may be sparser in frequency domain, some more in time domain or vice versa. Sparsity solutions are a recently emerging technique and are used for many applications. By compressing sensing technique only 30% of the signal samples are needed to reconstruct the complete image. It is a recent method for signal recovery in applied mathematics, which uses L1-optimization technique.

The appearance of one image under different environmental conditions is supposed to lie approximately in a low dimensional subspace. Given a target template set T = [t_1 $t_{2 \ldots\ldots}$ t_n] $\in R_{m*n}$ (d>>n), containing 'n' column vectors.

A valid test result y ∈ R_d approximately lies in the linear span of T,

$$y = Ta = a_1 t_1 + a_2 t_2 + \cdots + a_n t_n \qquad (4.1)$$

Where $a = (a_1, a_2, a_3 \ldots a_n) \in R$ is called a Target coefficient vector.

4.1.2: Robustness to occlusion

In several visual identification scenarios, objects are frequently corrupted by noise or partially occluded. Such occlusion generates unpredictable errors. Any part of the image can get affected and can appear at any place on the image.

Eigenbasis extraction is done using compressive sampling and Eigen basis matrix is used for representing the object. A valid test image 'y' approximately lies in the linear span of $[U]$.

$$Y \simeq \mathrm{Ua} = a_1 U_1 + a_2 U_2 + \cdots + a_n U_n \tag{4.2}$$

The effect of occlusion and noise can be incorporated as,

$$Y = \mathrm{Ua} + \mathrm{Z} \tag{4.3}$$

A trivial template, which is nothing but a identity matrix **I**= [**I**1, **I**2, ..., **I**d], $R_{m*n} \in$ to capture the occlusion as

$$Y = \mathrm{Ua} + \mathrm{Ie} \tag{4.4}$$

$$Y = [U \;\; I] \begin{bmatrix} a \\ e \end{bmatrix} = TC \tag{4.5}$$

Where, $[I]_{m*n}$ is identity matrix. 'e' is an error coefficient vector.

The locations of noise are unknown to the system. The above equation is known to be under-determined, which means there will not be any particular solution of the system of equations rather it will have an infinite number of solutions.

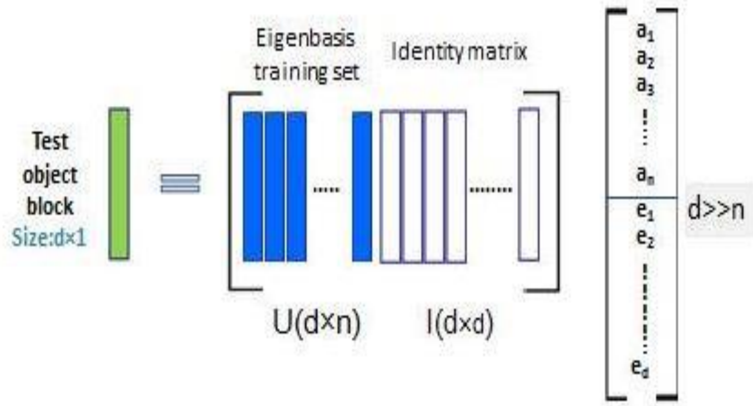

Fig 4.1: Dimension of underdetermined equation

4.2 Achieving Sparseness through L1-Minimization

If the number of atoms is greater than that of rows then the system is an under-determined. It then becomes an optimization problem to determine the non-zero entries. In order to get sparsest solution, we go for an L1 optimization method. If a signal x is reconstructed on the base of this equation, y = TC, here 'T' is the dictionary matrix, C is the original signal and 'y' is the 'n' measurements. In order to reconstruct the image we use the relationship

$$C = y^T T \tag{4.6}$$

The above system of linear equation is under-determined and hence does not yield a unique solution for C. The failure brought about by occlusion and noise ordinarily damages a small amount of the picture pixels. Therefore, for a better recognition result, there are just a few nonzero coefficients in 'e' that record for the noise and fractional occlusion. Thus, we need to have a sparse result. We exploit the compressibility in the transform domain by tackling the issue as an l1-regularized least squares problem, which commonly yields sparse Solution of the equation.

4.3: Compensation of Partial Occlusion/Noise in an Image

When the most sparse solution $[C^{*}] = [a^{*}, err^{*}]^{T}$ of the above under-determined system of linear equation is obtained through minimizing the L1-norm, using an optimizing algorithm, the occlusion compensated clean image of the subject is obtained after subtracting the error coefficients from the test image vector. This gives a clean image of the test image free from occlusion or noises. This can be represented as,

$$\text{Compensated image vector} = \text{Test image vector} - \text{err}^{*}$$

The test image vector is then reshaped back to the dimension of test image, which will result in a visibly clean test image.

4.4: Face Recognition Algorithm

1. Input: Matrix T

$$T = [U\ I]$$

2. Normalize columns of matrix [T] to have unit L2 norm.

3. Using L1-minimization for the equation sparest solution of under-determined system is

$$C^{*} = argmin_{c} \|C\|_{1} \ \text{ Subject to } TC = y$$

4. Compute the residuals, $r_i(y) = \|y - err^{*} - T\sigma_i(C^{*})\|_2$ for $i = 1,2,3 \ldots\ldots k$.

5. Output: identity $y = argmin_i r_i(y)$.

Pseudo code

- Accessed face images from the database in MATLAB.
- Images were resized to a lower dimension.
- Pixel values of every image was extracted and then reshaped to form a column vector.
- Data matrix was formed after horizontally concatenating all column vectors calculated in the previous step.
- Then, the columns of data matrix were mean adjusted.
- Eigen basis vectors were calculated using SVD. Thus sorted eigenvector matrix was obtained.
- Eigen vectors corresponding to a particular threshold value close to zero were eliminated. Thus some reduction in dimension is obtained.
- Centered images were projected onto feature subspace.
- Test image was obtained.
- Sparse coding is done using a large trivial template (identity matrix).
- Sparse solution for the underdetermined system of equation was calculated using L1-norm minimization.
- Residual vector was calculated corresponding to each coefficient in sparse vector.

  $r_i(y) = \|y - err^* - T\sigma_i(C^*)\|_2$ for $i = 1,2,3 \ldots\ldots k$.
- Index corresponding to the minimum residual value gives the index recognized image in the training database.

Chapter 5

Results Analysis

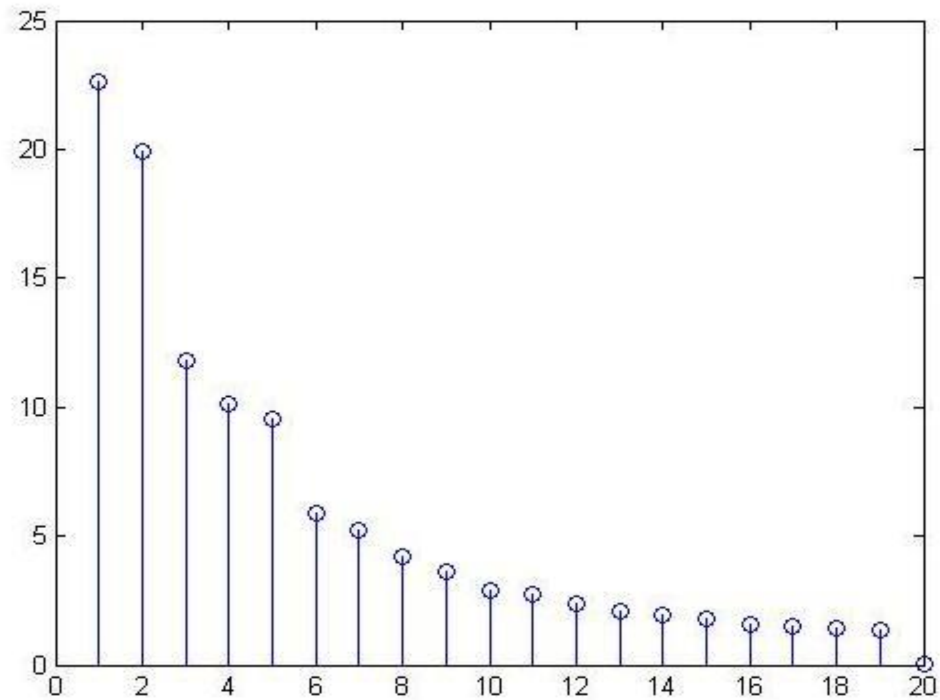

Fig 5.1: weight value of corresponding 20 eigenvectors of covariance matrix

Above graphical figure demonstrates the weight value of eigenvectors of Eigen-basis matrix. Weight of first vector is highest; weight of second vector is less than the first one and so on. Weight values also show the importance of each vector in the matrix. Thus first vector is of most importance and importance keeps on decreasing as we move towards the other columns. It also shows that the first vector contains the most amount of information about the image database and rest other contain less information respective of each other.

Fig 5.2(a): Group of image demonstrating image compression using PCA

Fig 5.2(b): Group of image demonstrating image compression using PCA

In above two figures 5.2(a) and 5.2(b), Image compression using PCA is demonstrated. In Fig 5.2(a) on a standard image of size 512×512 PCA is applied and Eigen basis matrix is calculated. This matrix has 512 eigenvectors. A group of 4 figures is kept together for quality comparison purpose. It is quite evident that projected image with mere 50 eigenvectors is comparable in quality with the one with 512 eigenvectors. i.e. 512-50=462 eigenvectors can be supposed to be redundant and is removed to save memory without losing much image information. In Fig 5.2(b) original image is of dimension 256×256. PCA is applied and group of 4 compressed images with different number of principal components is shown and those explain how number of principal components affect the quality of compressed image.

Fig 5.3(a): Eigen faces for Cropped YALE Face Database

Fig 5.3(b): Eigen faces from a dataset of three images per class

In above two figures the Eigen faces results are shown. Both figures contain 16 Eigen faces. In Fig 5.3(a) Eigen faces for cropped YALE database is shown. A fig 5.3 (b) Eigen faces for a face database with three faces per class is presented.

(I) (II) (III)

Fig 5.4(a): (I) compensated image for occlusion (II) Occluded test image (III) Estimated error after sparse coefficients calculation

Fig 5.4(b): (I) compensated image for occlusion (II) Occluded test image (III) Estimated error after sparse coefficients calculation

Fig 5.5(a): partially occluded image

Fig 5.5(b): occlusion compensated image

Fig 5.6(a): partially occluded image

Fig 5.6(b): occlusion compensated image

In above two set of figures, comparative view of occlusion compensation is demonstrated. First figure of the batch contains the image with fractional occlusion while second one is the occlusion compensated image. It is quite evident from the comparison that occlusion is successfully removed and clean image is reconstructed using Sparse PCA and L1- Norm minimization.

.

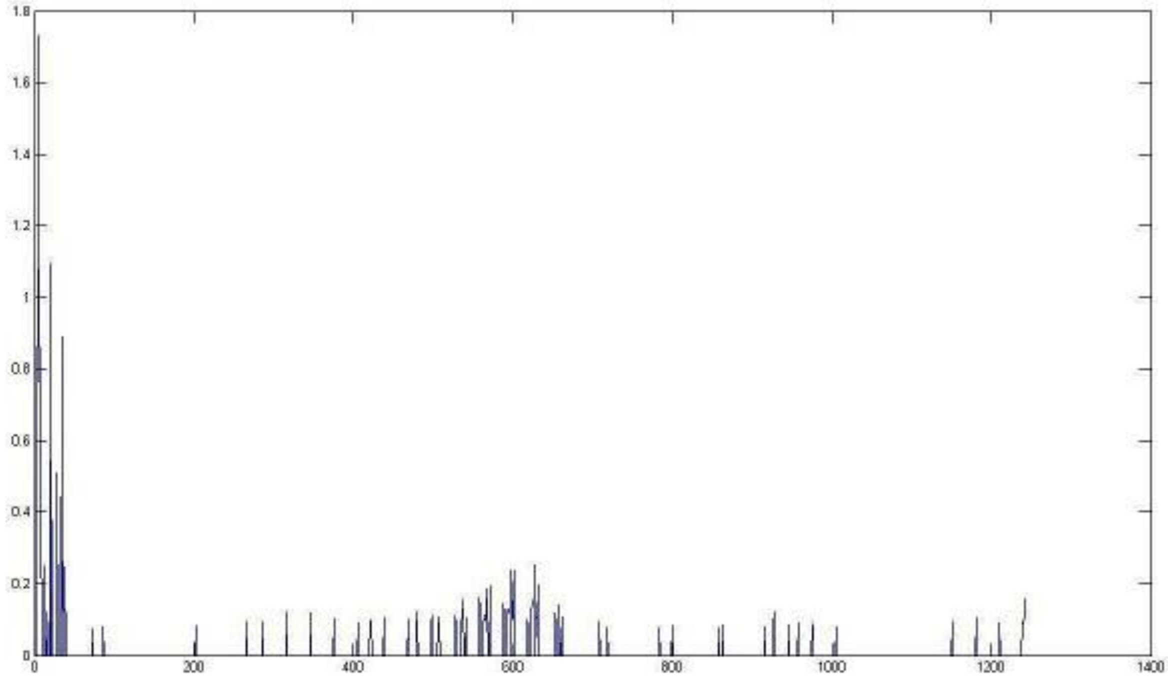

Fig 5.7(a): Sparse coefficients for a valid test image

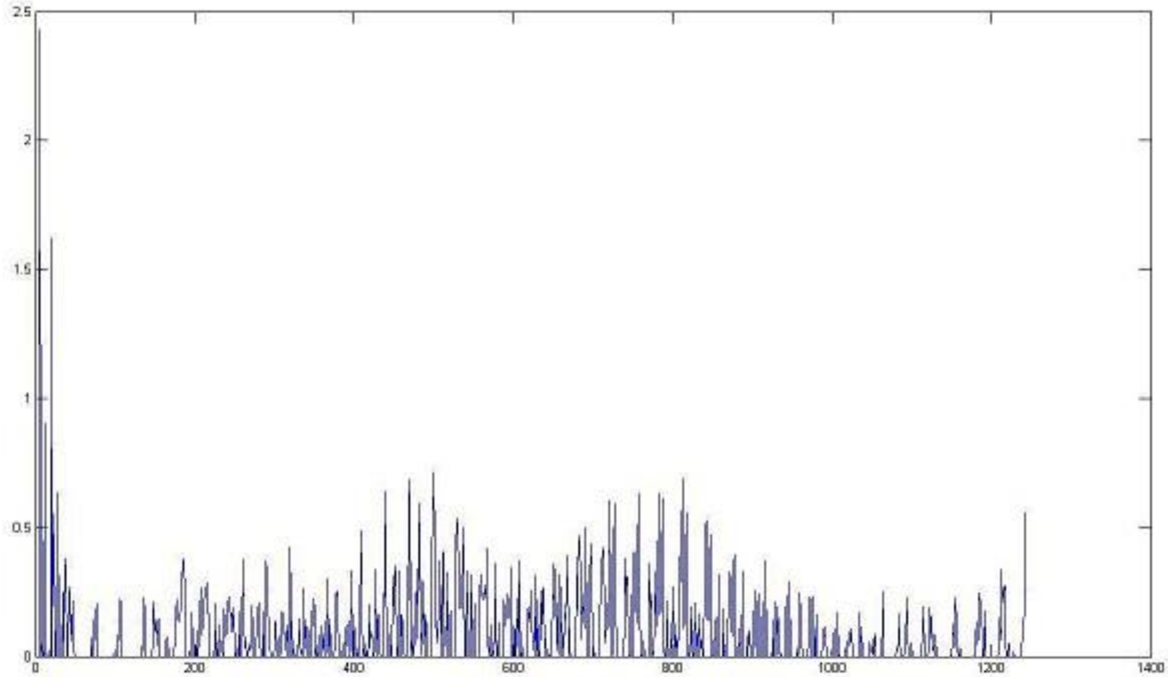

Fig 5.7(b): Sparse coefficients for an invalid test image

In the above two plots, calculated sparse coefficients for a valid test image and an invalid test image is shown. Valid test image refers that the test image is from the same class of database, while invalid test image means that the image is not close to the images of database. Through above two plots it has been tried to show that the sparse coefficients for a valid test image is Sparse while it is scattered in case of an invalid test image.

Recognition results :

Input image   Identified image

Input image   Identified image

Input image

Identified image

Input image

Identified image

Table 1: Dimensions of calculated matrix in the process:

Original image	Down sampled image	Reshaped image column vector	Number of training images	Data matrix	Eigen basis matrix(U0)
168×192	30×40	1200×1	100	1200×100	100×100

Table 2: Performance Comparison Table under different amount of Occlusion:

Percentage(%) of corruption	10%	20%	30%	40%	50%	60%	70%	80%	90%
Recognition rate for PCA	92.3%	75.1%	61%	11.9%	-	-	-	-	-
Recognition rate for SPCA	100%	100%	100%	100%	95.7%	87.2%	57.4%	12.8%	0%

CHAPTER 6

Conclusion and scope of future work

Conclusion:

Theory of sparse representation and its application onto face recognition is introduced. We verify that the feature extraction is no longer critical to recognition once the sparsity of the problem is properly harnessed. Common PCA technique has been in uses for over a decade now, which is able to give results under ideal environments but fails to perform under inconsistent environments like image under partial occlusion or under drastic changes in illumination. Sparse representation is quite a new concept in the field of compressive sampling and computer vision analysis. Here this concept is incorporated with traditional PCA to represent the test image and model the occlusion/noise present in test image. The algorithm was tested under noisy and occluded images. The experiment results show that the proposed algorithm outperforms other techniques under all circumstances under comparison. This incorporation of sparse representation is surely going to help in further research works of image processing and face recognition.

Although there are some limitations of proposed method too, result is not satisfactory when the occlusion portion of the image becomes more than 40%.

- It was also observed that as the amount of occlusion in the image increases the reconstructed image is not very clean. Hence percentage of occlusion affects the recognition process directly.

- It was also observed during experiments that the occlusion compensation becomes more accurate when number of training images increases in the formation of database.

Limitations:

- Memory is an issue in face recognition systems, so it is observed here also. Memory limitations of traditional computers restrict us to deal with the images of high resolution. Thus it affects the recognition process ultimately. Memory also affects the estimation of sparse coefficients, hence ultimately affects the error estimated.
- One more limitation is that it ceases to perform above an occlusion/noise limit, which is about 40% as per the experimental observations.

Future scope:

The future scope of the lies on the fact that whether the algorithm can be applied to real time video object tracking.

References:

[1] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys*, Vol.35, pp.399-458, 2003.

[2] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognition Letters*, Vol.28, pp.1885-1906, 2007.

[3] R. Brunelli and T. Poggio, "Face recognition: features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.15, pp.1042-1052, 1993.

[4] M. A. Grudin, "On internal representations in face recognition systems," *Pattern Recognition*, Vol.33, pp.1161-1177, 2000.

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

[6] T. Kanade, "Picture Processing System by Computer Complex and Recognition of Human Faces," Kyoto University, Japan, PhD. Thesis 1973.

[7] A. Yuille, D. Cohen, and P. Hallinan, "Feature extraction from faces using deformable templates," in *IEEE Computer Society Conference on Computer Vision and Templates*. San Diego, CA, USA, pp.104-109, 1989.

[8] N. Roeder and X. Li, "Experiments in analyzing the accuracy of facial feature detection," *Vision Interface '95*, pp.8-16, 1995.

[9] C. Colombo, A. D. Bimbo, and S. D. Magistris, "Human-computer interaction based on eye movement tracking," *Computer Architectures for Machine Perception*, pp.258-263, 1995.

[10] M. Nixon, "Eye spacing measurement for facial recognition," in *SPIE Proceedings*, pp.279-285, 1985.

[11] D. Reisfeld, "Generalized symmetry transforms: attentional mechanisms and face recognition," Tel-Aviv University, PhD. Thesis, 1994.

[12] H. P. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating faces and facial parts," in *International Workshop on Automatic Face- and Gesture-Recognition*, pp.41-46, 1995.

[13] I. Craw, D. Tock, and A. Bennett, "Finding face features," in *Second European Conference on Computer Vision*, pp.92-96, 1992.

[14] R. Cendrillon and B. C. Lowell, "Real-Time Face Recognition using Eigenfaces," in *Proceedings of the SPIE International Conference on Visual Communications and Image Processing*, Vol.4067, pp.269-276, 2000.

[15] R. J. Baron, "Mechanisms of Human Facial Recognition," *International Journal of Man-Machine Studies*, Vol.15, pp.137-178, 1981.

[16] R.-J. J. Huang, "Detection Strategies for face recognition using learning and evolution," George Mason University, Fairfax, Virginia, Ph. D. Dissertation 1998.

[17] L. Sirovich and M. Kirby, "Low-dimensional Procedure for the Characterization of Human Faces," *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, Vol.4, pp.519-524, 1987.

[18] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. New Jersey: Prentice-Hall, 1988.

[19] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, second ed. Boston, MA: Academic Press, 1990.

[20] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.586-591, 1991.

[21] M. Turk and A. Pentland, "Eigenfaces For Recognition," *Journal Of Cognitive Neuroscience*, Vol.3, pp.71-86, 1991.

[23] Will,Todd, “Introduction to the Singular Value Decomposition,” *Davidson College*, *www.davidson.edu/academic/math/will/svd/index.html,*1999.

[24] J. Wright, A. Yang, A. Ganesh, S. S. Sastry, Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No.2, pp.210-227, 2009.