COMPARATIVE STUDY BETWEEN FEATURE EXTRACTION METHODS FOR FACE RECOGNITION

SITI FAIRUZ BINTI ABDULLAH

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS*

JUDUL: COMPARATIVE STUDY BETWEEN FEATURE EXTRACTION METHODS FOR FACE RECOGNITION

SESI PENGAJIAN: 2013/2014

Saya SITI FAIRUZ BINTI ABDULLAH

mengaku membenarkan tesis Projek Sarjana Muda ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenark membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
	TIDAK TERHAD	

(TANDATANGAN PENULIS)

Alamat tetap:

Tarikh:

(TANDATANGAN PENYELIA)

Nama Penyelia

Tarikh:

CATATAN: * Tesis dimaksudkan sebagai Laporan Projek Sarjana Muda(PSM).
** Jika tesis ini SULIT atau atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

© Universiti Teknikal Malaysia Melaka

COMPARATIVE STUDY BETWEEN FEATURE EXTRACTION
METHODS FOR FACE RECOGNITION

SITI FAIRUZ BINTI ABDULLAH

This report is submitted in partial fulfillment of the requirements for the
Bachelor of Computer Science (Artificial Intelligence)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2014

CHAPTER I

INTRODUCTION

1.1 Project Background

Image recognition becomes one of the most popular topics for researcher nowadays. As we can see, image recognition has been implementing in Facebook, android, camera digital, personal laptop etc. Facebook is the one of the social network that implements image recognition for their user to ease their user in photo tagging process.

There are many different identification technologies available. The most common is an identification method today is a Personal Identification Number system. Those systems always have trouble with forgery, theft, and lapses in users' memory. Therefore pattern recognition technique has been chosen because these techniques identify people using their physiological characteristics like face and fingerprint.

Biometrics has been used to identify people as the identification using an individual's unique physical characteristic. By using the individual's physical characteristic, it can never be forgotten, lost or copied, like a card or password. Biometrics has been widely accepted as a fast, accurate and dependable way to identify a person.

Face recognition is the one of biometrics example that becoming the most reliable way to verify a person. It has even been developed into an application for smart phone users. This proves that it is moving into our everyday lives.

Face recognition has gained attention among researcher. This is evidenced by the emergence of a specific face recognition conference such as AFGR and AVBA and FRT (Javidi B., 2002). Face recognition can be used without anyone noticing since facial images easily accessible remotely from the camera (Rabia Jafri at el, 2009). This system gives beneficial for security and surveillance purposes (Rabia Jafri at el, 2009). Government agencies are more interested in increasing the use of biometric security system based on behavioural characteristics (Andrea F.Abate at el, 2007).

Face recognition is become the most challenging problems in biometric systems because there are many image variations in real-life such as lighting, illumination, facial expression, partial occlusion and imprecise localization of face area. (Martínez, A. M., 2002)

The need of face recognition is for security issue where to cope up with the present day crime. Face recognition system is proving to be very efficient in the present day market, which is also meeting the general criteria of our respectable clients through various perspectives. Face recognition able to provide the maximum security better than any other security solution.

1.2 Problem Statement

Previous system, in order to recognize authorized user is by inserting username and password to access into the system. Those systems always have trouble with forgery, theft, and lapses in users' memory. To overcome this weakness in this study, face recognition is proposed as a solution.

Biometrics is usually associated with the applicability of unique physiological characteristics to identify an individual. It used to authenticate the person's identity by using special characters like fingerprint, face, iris and etc. A biometric system can be used in two ways, whether in validation or identification. For identification, biometrics can be apply in determine a person's identity without knowledge, for example, scanning a crowd using face recognition technology. For validation, biometrics can be applied to verify person's identity, for example, retinal scan and fingerprint scans.

Refer to previous analyses, face has been choosing because it has a high level of accuracy (Rabia Jafri at el, 2009). But there is a problem with human face where human face will change over time. For example wrinkles, beard, mustache, glasses and shape of the head that can affect the performance must be considered. Besides that, environment also can affect the performance, such as lighting etc.

Nowadays, the existence of the face recognition system is not very good in ability to recognize the real picture (Rouhi, R., Amiri, M., & Irannejad, B., 2012). As a feature extraction is an important step in face recognition operation, researchers keeps find the best features extraction method for system in performing under different condition. Illumination, variation of pose and different expression is a challenging problem in the feature extraction of the face recognition (Rouhi, R., Amiri, M., & Irannejad, B., 2012).

The most popular group of feature extraction methods for appearance based are Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) (Turk, M., & Pentland, A., 1991). Linear Discriminant Analysis (LDA) usually gives better recognition accuracy because LDA finds linear transformation when applied to a set of image (Yu, H., & Yang, J., 2001). Therefore, study and investigation of Linear Discriminant Analysis (LDA) as a feature extraction method become one of research branches.

1.3 Objective

The objectives of the project

1. To explore Linear Discriminant Analysis (LDA) as feature extraction technique for face recognition.
2. To analyse the capability of Linear Discriminant Analysis Algorithm (LDA) in face recognition.

1.4 Scope

The scope of this project consists of feature extraction in face recognition algorithm. There are different methods in face recognition which are LDA, PCA, Sobel, Prewitt, Robert, Canny involve a series of steps which is capturing, analyzing and comparing a face to a face database.

In analyzing part, these methods or techniques will be approached to do the feature extraction part. This research will include another two stages which are image processing part and classification part.

1.5 Project Significance

This project intends to help researcher, public and industries to understand the concept of image or face recognition and to choose the best method for the accuracy of the system.

The wide use of this system or application needs us to improve the current system in order get better application of this system for the future. The significance of this research is to investigate the capability of Linear Discriminant Analysis (LDA) as a feature extraction method in face recognition domain.

1.6 Expected Output

At the end this research, Linear Discriminant Analysis (LDA) and other feature extraction techniques will be applied to the images and LDA techniques will give the best accuracy in for face recognition.

1.7 Conclusion

Face recognition technology is gives convenient for the human authentication. It gives beneficial for security and surveillance purposes. The existence of the face recognition system is not very good in ability to recognize the real picture.

Nowadays, government agencies are more interested in increasing the service of biometric security system. There are three stage in face recognition which is processing, feature extraction and classification part. As a feature extraction is an important part in face recognition, the study of feature extraction method is needed. The study and investigate of feature extraction method become one of research branches.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction to Face Recognition

Face recognition and authentication is most widely applied in several applications such as authentication within access control devices and etc (Karima Ouji at el, 2009). Face recognition can be used without anyone noticing since facial images easily accessible remotely from the camera (Rabia Jafri at el, 2009). This system gives beneficial for security and surveillance purposes (Rabia Jafri at el, 2009). Government agencies are more interested in increasing the use of biometric security system based on behavioural characteristics (Andrea F.Abate at el, 2007). Since a few years ago, face recognition gained significant attention as the one of successful applications of image analysis (Bahram Javidi, 2002).

Nowadays, system developers come up with the face recognition application that are capable of extracting faces and picking up faces from the crowd and compared to an image source or database. Some face recognition software application was designed to recognize similarity using pattern recognition. This outline carries the characteristics of a typical pattern recognition system (Phil Brimblecombe, 2005).
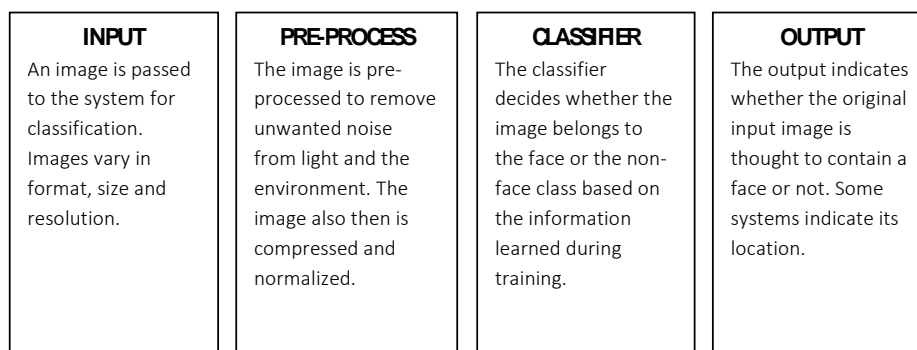
INPUT	PRE-PROCESS	CLASSIFIER	OUTPUT
An image is passed to the system for classification. Images vary in format, size and resolution.	The image is pre-processed to remove unwanted noise from light and the environment. The image also then is compressed and normalized.	The classifier decides whether the image belongs to the face or the non-face class based on the information learned during training.	The output indicates whether the original input image is thought to contain a face or not. Some systems indicate its location.

Figure 2.1: Pattern recognition system

Face recognition or identification based on physiological characteristics is included in biometrics based technologies. Recent years, biometrics-based techniques are the successful techniques for recognizing individuals (Rabia Jafri at el, 2009). The basic idea behind biometrics is that our bodies contain unique properties that can be used to distinguish us from others. Firstly, there is no voluntary action for face recognition. Other biometrics need to do some voluntary action by the user. For example, of fingerprint identification, the user needs to place his fingerprint at the hand geometry detection and for the retina identification, the user need to stand in a fixed position in front of the camera.

Face recognition can be used without anyone noticing since facial images easily accessible remotely from the camera (Rabia Jafri at el, 2009). Facial image can be easily obtained from the inexpensive camera. Face recognition and identification is totally non-intrusive and does not carry such health risks compared to other physiological characteristics (Rabia Jafri at el, 2009).

2.1.1 Face Recognition System

Face recognition or identification based on physiological characteristics and behavioral traits is included in biometrics based technologies. Recent years, biometrics-based techniques are the successful techniques for recognizing individuals (Rabia Jafri at el, 2009).

The basic idea behind biometrics is that our bodies contain unique properties that can be used to distinguish us from others. Face recognition has their own benefit compared to other physiological characteristics. Biometrics involves two categories which are physiological biometrics and behavioral biometrics (Paul Reid, 2004).
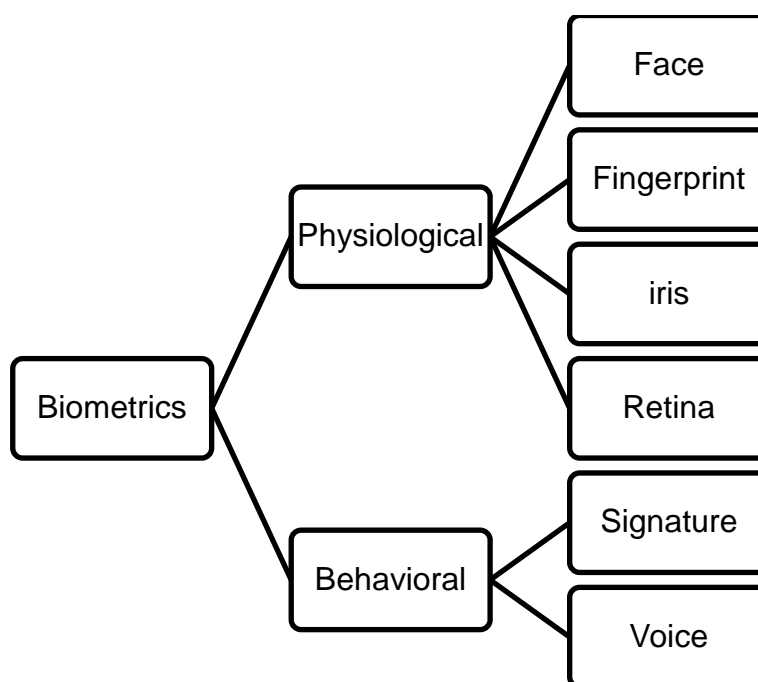

Figure 2.1.1.1 Biometrics characteristics

Table 2.1.1.2 Comparison of Various Biometric Technologies

	Univers-ality	Distinctive-ness	Permane-nce	Collecta-ble	Performa-nce	Acceptabi-lity
Face	High	High	Medium	High	Low	High
Fingerprint	Medium	High	High	Medium	High	Medium
Iris/Retina	High	High	High	Medium	High	Low
Signature	Low	Low	Low	High	Low	High
Voice	Medium	Low	Low	Medium	Low	High

The table shows the comparison of various biometric technologies. There is universality which these biometric technologies measured by the amount of people use these technologies.

For distinctiveness, it has been measured by how people can distinguish all the technologies based on an identifier. For permanence it has been measured on how permanent is the identifier. For collectable, it has been measured by how well that identifier can be captured and quantified. For performance, it is measured by matching speed and the accuracy of that biometric technology and for acceptability it has been measured by the willingness of people to accept that technology (Manivannan at el, 2000).

There are three tasks in face recognition process which is pre-processing, feature extraction and classification part. The step of feature extraction involves obtaining the relevant facial features from the data. These features could be face region, variations, angle or measures and etc. The last step is identification step which the process compare, classify and accuracy measure. This phase will do classification process using data mining or etc.
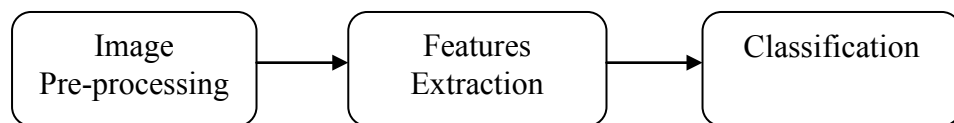
Image Pre-processing → Features Extraction → Classification

Figure 2.1.1.3 Flow Chart of Face Recognition Process

2.1.2 Basic Algorithms of Face Recognition

Face recognition problem can be divided into two main categories which is face verification and   face identification. The process begins with the input of the face. The recognition stage includes feature extraction where important for discriminant is saved, and the matching where the recognition results is given with the aid of face database.

Figure 2.1.2.1 Basic step of Face Algorithms

2.2 Feature Extraction

Feature extraction is a form of dimensionality reduction in pattern recognition and image processing. Feature extraction is used when the input data is too large to be processed by transforming the input data into a set of features. The set of features extracted are choosing based on relevant information in order to perform the desired task.

2.3 Linear Discriminant Analysis (LDA)

Linear discriminant analyses (LDA) is used in pattern recognition to find a linear combination of features which characterizes or separates two or more classes of objects or events. LDA attempts to express one dependent variable as a linear combination of other features. It is discriminant function analyses which are linear transformation. LDA also known as a classical statistical approach for supervised dimensionality reduction and classification in the feature extraction algorithm (Adnan N. F., 2012).

Data sets can be transformed into vectors and can be classified in the transformed space by two different approaches.

(i)   *Class-dependent transformation*: This type of approach involves maximizing the ratio of between class variance to within class variance. The main objective is to maximize this ratio, so that adequate class separability is obtained. The class-specific type approach involves using two optimizing criteria for transforming the data sets independently.

(ii)   *Classification-independent transformation*: This approach involves maximizing the ratio of the overall variance to within class variance. This approach uses only one optimizing criterion to transform the data sets and hence all data points irrespective of their class identity are transformed using this transform. In this type of LDA, each class is considered as a separate class against all other classes.

The goal of the Linear Discriminant Analysis (LDA) is to find an efficient way to represent the face vector space. LDA uses class specific information which best discriminates among classes.

2.3.1 Step of LDA

1. Get images from MIT-CBCL face database.

2. Compute the average of all faces.

3. Compute the average face of each person.

4. Subtract them from the training faces.

5. The within-class scatter matrix SW

6. From this scatter matrix we calculate the Fisher face vectors.

2.3.1.1 LDA Equation

Scattering matrix is defined as

$$S_i = \sum_{x \in Gi}(x - mi)(x - mi)^t - (1)$$

Where,

x = a data in a group Gi,

m_i = means the centroid of the group

Gi = between-class scatter matrix (S_B) and within-class scatter matrix (S_W) are defined respectively as follows.

$$S_B = (m_1\text{-}m_2)(m_1\text{-}m_2)^t - (2)$$

$$S_W = S_1 + S_2 - (3)$$

The solution to maximize the ratio between S_B and S_W

$$w = S^{-1}_W (m_1 - m_2) - (4)$$

S_W should be non singular.

2.3.1.2 LDA in Face Recognition

Linear Discriminant Analysis is the method of feature extraction which is the faces of the same subjects are grouped into separate classes and the variations between the images of different classes are discriminated using the eigenvectors at the same time minimizing the covariance within the same class (Woodward at el, 2001).

SW is the within-class scatter matrix and SB is the between-class scatter matrix. These scatter matrices are defined as:

$$SB = \Sigma_C N_i (ClasAvg_i - AvgFace) (ClasAvg_i - AvgFace)^T - (5)$$

$$SW = \Sigma_C \Sigma X_i (x_k - ClasAvg_i) (x_k - ClasAvg_i)^T - (6)$$

Where,

C = the number of distinct classes,

N_i = the number of images for each class i,

$ClasAvg_i$ = the average face image of faces in class i,

X_i = represents the face images that are in class i,

AvgFace = the average face image for all images in the dataset.

The steps involved in implementing the algorithm are (P.N. Belhumuer at el, 2000):
Training phase:

1. Represent the faces in the database in terms of the vector X as shown in Equation (2).

2. Compute the average face AvgFace and subtract the AvgFace from the vector X as in Equation (4).

3. Classify the images based on the number of unique subjects involved. So the number of classes, C, will be the number of subjects who have been imaged.

3. Compute the scatter matrix using Equation (5) and (6).

4. Use PCA to reduce the dimension of the feature space to N − C. Let the eigenvectors obtained be WPCA.

5. Project the scatter matrices onto this basis to obtain non-singular scatter matrices SWN and SBN.

6. Compute the generalized eigenvectors of the non-singular scatter matrices SWN and SBN so as to satisfy the equation SB*WLDA = SW*WLDA*D, where D is the eigenvalue. Retain only the C-1 eigenvectors corresponding to the C-1 largest eigenvalues. This gives the basis vector WLDA.

7. Then the image vector X is projected onto this basis vector and the weights of the image are computed. Recognition of a new face is done similar to that of the PCA method. Here the given face is moralized by subtracting the AvgFace and then the weights are calculated by projecting the image onto the basis vectors. Then the Euclidean measure is used as the similarity measure to determine the closest match for the test image with the face in the trained database.

2.4 Edge Detection

Edge detection plays an important role in computer vision and image analysis because edge is the basic characteristic of the image. Edge detection can be the main approach to image analysis and recognition because of the useful and identical information that's contained in edge of the sub - image. Edge detection is a fundamental tool in image processing which are in the areas of feature detection and feature extraction. By applying an edge detection algorithm to an image, the amount of data to be processed may be reducing. Edge detection algorithms are filtering out information that may be regarded as less relevant, but it still preserving the important structural properties of an image. If the edge detection step is successful, the subsequent task of interpreting the information contents in the original image may therefore be substantially simplified.

2.4.1   Sobel Edge Detector

Sobel operator is based on convolving the image with a small, separable based on integer valued filter in horizontal and vertical direction and relatively easier in terms of computations. The result of the Sobel operator for each point image is either the corresponding to gradient vector or the norm of the vector. The gradient approximation that produces is relatively crude for high frequency variations in the image.

2.4.1.1 Step of Sobel Edge Detector

1) Consider the arrangement of pixels about the pixel (i, j):

$$
\begin{matrix}
a0 & a1 & a2 \\
a7 & [i,j] & a3 \\
a6 & a5 & a4
\end{matrix}
$$

2) The partial derivatives can be computed by:

$$M_x = (a2 + ca3 + a4) - (a0 + ca7 + a6) \; -(7)$$

$$M_y = (a6 + ca5 + a4) - (a0 + ca1 + a2) \; -(8)$$

3) The constant c implies the emphasis given in pixels closer to the center of the mask.

4) Setting c = 2, we get the Sobel operator:

$$
M_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}
\qquad
M_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}
$$

2.4.2   Prewitt Edge Detector

Prewitt operator is a discrete differentiation operator that computes an approximation of the gradient of the image intensity function. The result of the Prewitt operator at point of the image is either corresponding to gradient vector or the norm of the vector. The gradient approximation that produces is relatively crude for high frequency variations in the image.

2.4.2.1 Step of Prewitt

1) Consider the arrangement of pixels about the pixel (i, j):

$$
\begin{array}{ccc}
a0 & a1 & a2 \\
a7 & [i,j] & a3 \\
a6 & a5 & a4
\end{array}
$$

2) The partial derivatives can be computed by:

$$M_x = (a2 + ca3 + a4) - (a0 + ca7 + a6) - (9)$$

$$M_y = (a6 + ca5 + a4) - (a0 + ca1 + a2) - (10)$$

3) The constant c implies the emphasis given in pixels closer to the center of the mask.

4) Setting c = 1, we get the Prewitt operator:

$$
M_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}
\qquad
M_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}
$$

2.4.3   Canny Edge Detector

The canny edge detector is an operator that uses a multi-stage algorithm to detect a wide range of edges in images. The aim was to discover the optimal edge detection where the optimal edge detector means to have good detection in edges, good localization for edges and minimal response to the edges. Good detection is how the algorithm should mark as many real edges in the image as possible. For the good localization, the edged marked should be as close as possible to the edge in the real image while minimal response by giving an edge in the image should only be marked once and image noise should not create false edges.

2.4.3.1 Step of Canny

1) Consider the arrangement of pixels about the pixel (i, j):

$$
\begin{matrix}
a0 & a1 & a2 \\
a7 & [i,j] & a3 \\
a6 & a5 & a4
\end{matrix}
$$

2) The partial derivatives can be computed by:

$$M_x = (a2 + ca3 + a4) - (a0 + ca7 + a6) - (11)$$

$$M_y = (a6 + ca5 + a4) - (a0 + ca1 + a2) - (12)$$

3) The constant c implies the emphasis given in pixels closer to the center of the mask.

4) Setting c = 1, we get the Canny operator:

$$
M_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}
\qquad
M_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}
$$

2.4.4 Roberts

The Roberts cross operator is to approximate the gradient of an image through discrete differentiation which is achieved by computing the sum of the squares of the differences between diagonally adjacent pixels.

2.4.4.1 Step of Roberts

If 2x2 window is used,

$$\begin{vmatrix} p1 & p2 \\ p3 & p4 \end{vmatrix}$$

Where the filter is center on p1 with p2 being pixel[x+1] [y] and p3 being pixel[x] [y+1], etc. then the formula to calculate the resulting new p1 pixel is

$$\text{Pixel} = \text{abs}\ (p1-p4) + \text{abs}\ (p2-p3) - (13)$$

$$\text{Pixel} = \text{SQRT}\ ((X*X) + (Y*Y)) - (14)$$

Where X = abs (p1-p4) and Y = abs (p2-p3)

The two filters are basic convolution filters of the form:

$$M_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad M_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

2.5 Principle Component Analysis (PCA)

Principle Component Analysis (PCA) is a type of dimensional reduction or ordination analysis. Ordination analysis attempts to embed objects distributed in high dimensional space into lower dimensional space. Dimensionality reduction is achieved by projection to lower dimensional space using a linear transformation. PCA can effectively reduce redundant information. PCA converts a set of observations correlated variables into a set of value of linear uncorrelated variables. The number of principle components will be lesser than the original variable.

PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. PCA is a powerful tool for analyzing data. The main advantage of PCA is that to find these patterns in the data, and to compress the data, for example by reducing the number of dimensions, without much loss of information.

2.5.1 Step of PCA

1. Get images from MIT-CBCL face database.

2. Subtract the mean. The mean subtracted is the average across each dimension. All the x values have $\bar{x}$(the mean of the x values of all the data points) subtracted, and all the y values have you subtracted from them. This produces a data set whose mean is zero.

3. Calculate the covariance matrix.

4. Calculate the eigenvectors and Eigenvalues of the covariance matrix.

5. Choosing components and forming a feature vector.

6. Deriving the new data set