University of Management and Technology

FINAL YEAR PROJECT REPORT

Face Recognition Using

Principle Component Analysis

Project Advisor

Mr.Mirza Mubasher Baig

Submitted by

Ilker Binal- 101620165

Department of Computer Science

School of Science and Technology

University of Management and Technology

Face Recognition Using Principle Component Analysis

Project Report submittedto the
Department of Computer Science, University of Management and Technology
In partial fulfilment of the requirements for the degree of
Bachelor of Science
In
Computer Science

Ilker Binal-101620165

Advisor

Head of Department

---

10 January 2013

Abstract

All transactions made in everyday life, time, speed and reliability are important. The entrance of a large company, employees, or similar methods taught at each entrance and exit of the card use is inevitable. However, these methods are used in the safety chain must be broken. Every day in this new and high performance methods are proposed. Automatic human face recognition is one of them. Because it is the human face can be distinguished features and easy to change is seen as a promising method.

Based on work done previously, taking into account the needs of today's automated human face recognition using Principal Component Analysis has been made and residues resulting from high performance.

Dedication

To My Family

Acknowledgements

I express sincere appreciation to my supervisor Mr.Mirza Mubasher Baig for his guidance and insight throughout the research and people who participate in my project experiments are gratefully acknowledged.

Table of contents

List of figures

List of tables

List of Results

Chapter 1

Introduction

Biometrics

Biometrics is the automated recognition of individuals based on their behavioural or physiological characteristics The physiological characteristics are related to the shape of the body. The most common example is fingerprint. Other examples include face recognition, hand geometry and iris recognition. The behavioural characteristics are related to the behaviour of a person. Signature is one example of these characteristics which is still widely used today. Modern approaches are the study of keystroke dynamics and voice.

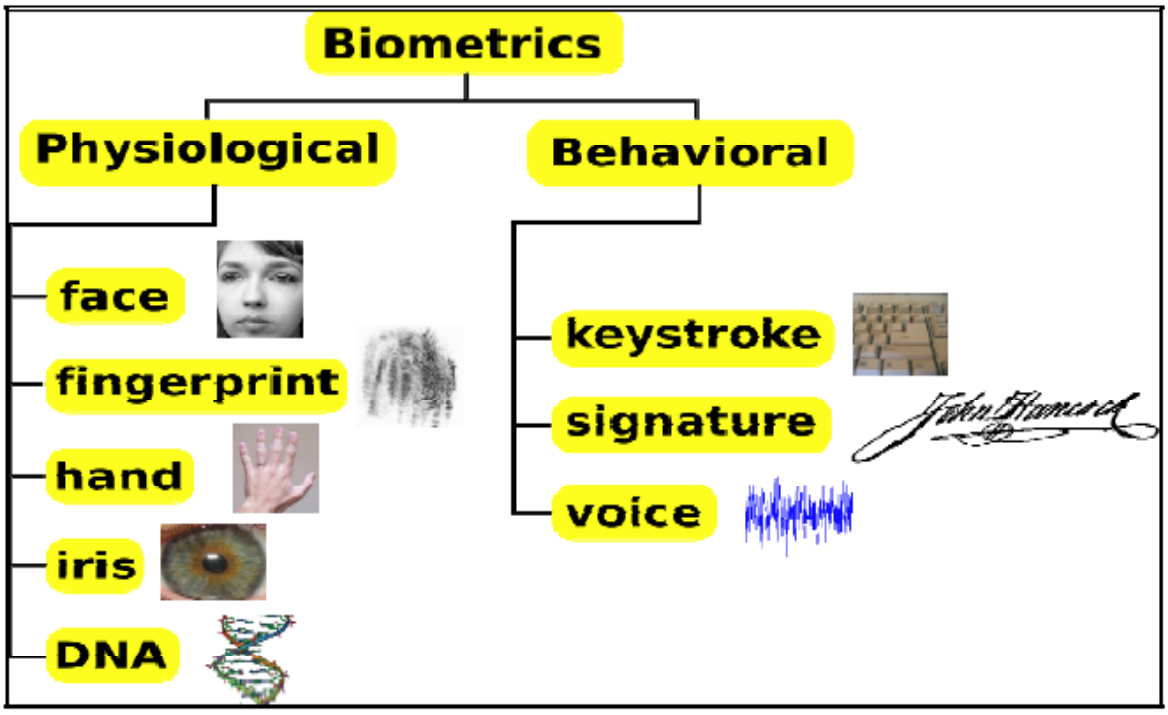

Figure 1

Face recognition

Humans often use faces to recognize individuals and advancements in computing capability over the past few decades now enable similar recognitions automatically. Early face recognition algorithms used simple geometric models, but the recognition process has now matured into a science of sophisticated mathematical representations and matchingprocesses. Major advancements and initiatives in the past ten tofifteen years have propelled face recognition technology into thespotlight. Face recognition can be used for both verification andidentification (open-set and closed-set)

Adatabasecreatedfrom thesamplefacesto be made overthe recognition ofthe personas a result ofautomated processes, and thisrecognition processcan be donewitha variety ofmethods.

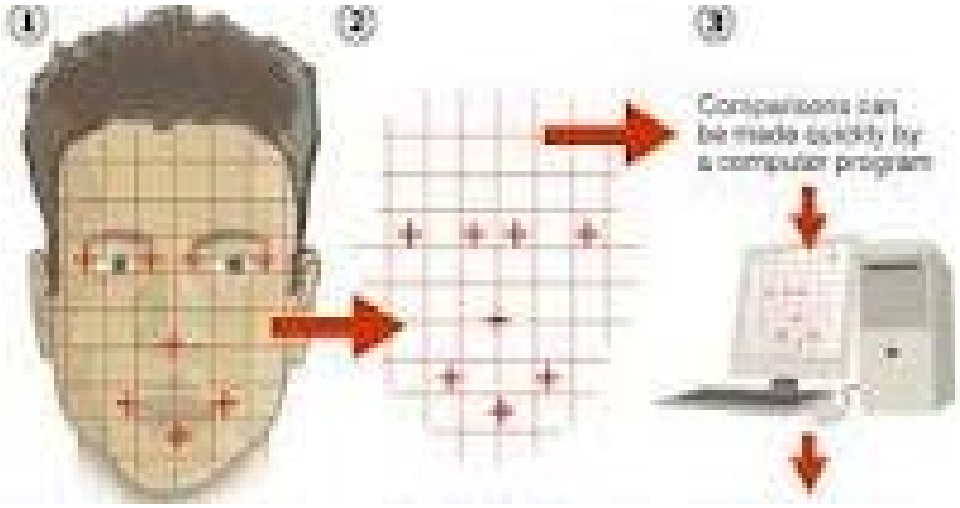

Figure 2

Face recognition problems can be analysed mainly in three groups. The first is the identification and removal of an image on the face. These various approaches are used to fix the problem. The second person is a problem of verification; any two of these systems are required to decide whether the image is the same person. The third and most difficult problem concerns the determination of identity. These types of sample taken from a system image, compared with faces in the database are asked to determine what is started

Every human being has a different and dissimilar creativity which are the foundations of Face recognition system. If it is expanded, there are dissimilar structures in the face physique as it is present in finger print. But the dissimilar face structure is not known by people. Everybody has a face structure and the operation of face recognition is made by contours known as symmetry of faces is matched through computer. Staff can pass from the gate if their face structure is matched with computerized face recognition. Face recognition system analyses characteristic features of faces then compares with face database.

Face recognition is biometric identification by scanning a person’s face and matching it against a library of known faces. Face recognition is defined as the identification of a person from an image of their face. Face Verification: Given a face image that might not belong to the database, verify whether it is from the person it is claimed to be in the database.

As one of the most successful applications of image analysis and understanding, face recognition has recently received significant attention, especially during the past few years.Given the requirement for determining people's identity, the obvious question is what technology is best suited to supply this information? There are many different identification technologies available, many of which have been in wide-spread commercial use for years.

The most common person verification and identification methods today are Password/PIN (Personal Identification Number) systems, and Token systems (such as your driver's license).Because such systems have trouble with forgery, theft, and lapses in users' memory, there has developed considerable interest in biometric identification systems, which use pattern recognition techniques to identify people using their physiological characteristics. Fingerprints are a classic example of a biometric; newer technologies include retina and iris recognition. In the last couple of years we have seen an enormous growth in electronically available services, such as banking through ATMs, the internet and voice services (phone). Humans are integrated closer to computers every day, and computers are taking over many services that used to be based on face to face contact between humans.

This has prompted an active development in the field of biometric systems. Face recognition is more advantageous than the other biometrics used. Whereas manybiometrics require the subjects co-operation and awareness in order to perform an identification or verification, such as looking

into an eye scanner or placing their hand on a fingerprint reader, face recognition could be performed even without the subject's knowledge.

History of Face Recognition

Automated face recognition is a relatively new concept. Developed in the 1960s, the first semi-automated system for facerecognition required the administrator to locate features (such aseyes, ears, nose, and mouth) on the photographs before it calculated distances and ratios to a common reference point, which were then compared to reference data. In the 1970s,Goldstein, Harmon, and Lesk used 21 specific subjective markers such as hair colour and lip thickness to automate the recognition. The problem with both of these early solutions was that the measurements and locations were manually computed. In 1988,Kirby and Sirovich applied principle component analysis, a standard linear algebra technique, to the face recognitionproblem. This was considered somewhat of a milestone as itshowed that less than one hundred values were required to accurately code a suitably aligned and normalized face image.

In1991, Turk and Pentland discovered that while using theeigenfaces techniques, the residual error could be used to detectfaces in images– a discovery that enabled reliable real-timeautomated face recognition systems. Although the approach was somewhat constrained by environmental factors, it nonetheless created significant interest in furthering development of automated face recognition technologies. The technology first captured the public's attention from the media reaction to a trial implementation at the January 2001 Super Bowl, which captured surveillance images and compared them to a database of digital mug shots. This demonstration initiated much-needed analysis onhow to use the technology to support national needs while being considerate of the public's social and privacy concerns. Today, face recognition technology is being used to combat passport fraud, support law enforcement, identify missing children, and minimize benefit/identity fraud.

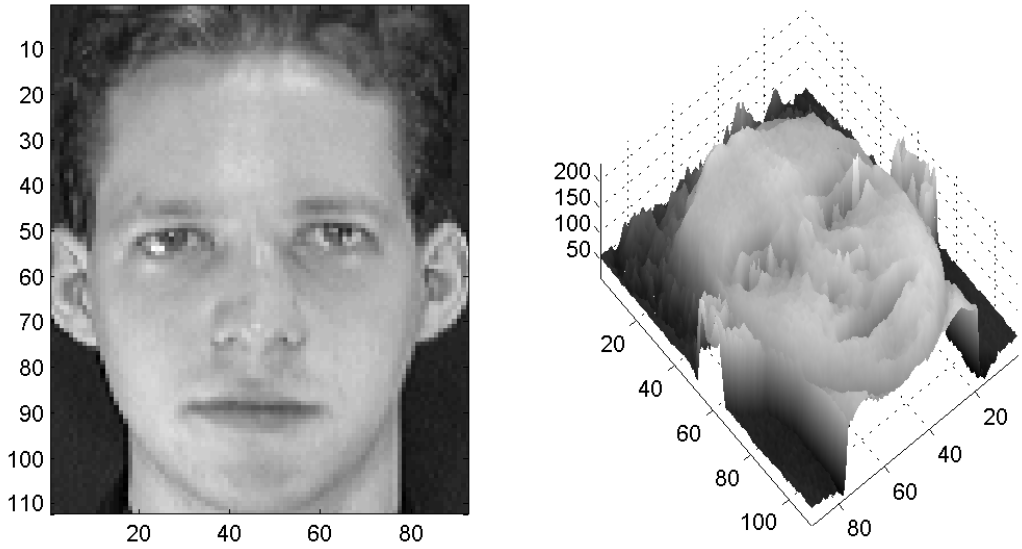

Figure 3

Face recognition refers to an automated or semi-automated process of matching facial images. Many techniques are available to apply face recognition of which is Principle Component Analysis (PCA). PCA is a way of identifying patterns in data and expressing the data in such a way to highlight their similarities and differences. Before applying this method to face recognition, a brief introduction is given for PCA from mathematical point of view.

There are many other types of identification such as password, PIN (personal identification number) or token systems. Moreover, it is nowadays very instilled the usage of fingerprints and iris as a physiological identification system. They are very useful when we need an active identification system; the fact that a person has to expose their body to some device makes people feel being scanned and identified. The pause-and-declare interaction is the best method for bank transactions and security areas; people feel conscious of it, as well as comfortable and safe with it. However, we do not want to interact with people that way in many other applications that required identification. For example, a store that wishes to recognize some customers or a house that has to identify people that live in there. For those application, face as

well as voice verification are very desirable. It is also important that an identification technique is closer to the way human beings recognize each other.

Face viewed as a vector

A face, which is an image, can be viewed as a vector. If the image's width and height are **w** and **h** pixels respectively, the number of components of this vector will be **w*h**. Each pixel is coded by one vector component. The construction of this vector from an image is performed by a simple concatenation - the rows of the image are placed each beside one another, as shown on the figure 5

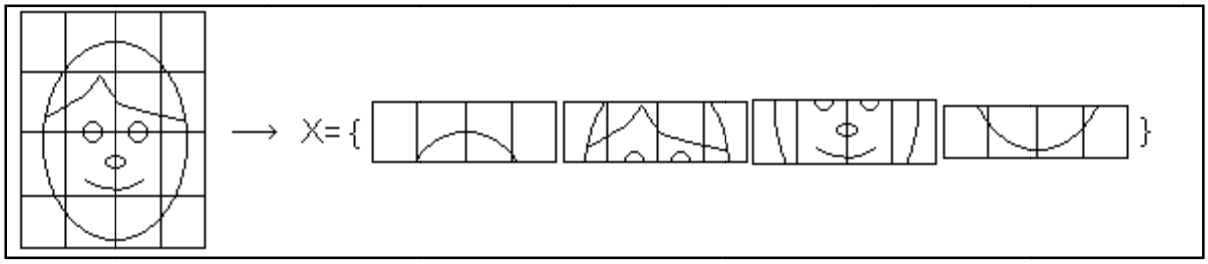

Figure 4: formation of the face's vector from the face's image

Image Space

The face vector described in the previous section belongs to a space. This space is the image space, the space of all images whose dimension is **w** by **h** pixels. The basis of the image space is composed of the following vectors:

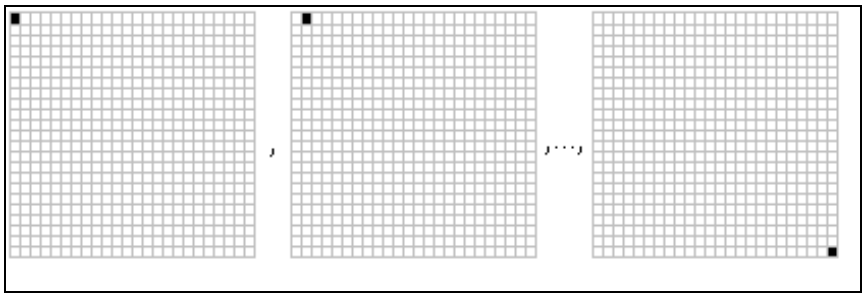

Figure 5: Image space's basis

All the faces look like each other. They all have two eyes, a mouth, a nose, etc. located at the same place. Therefore, all the face vectors are located in a very narrow cluster in the image space, as shown in the figure 6

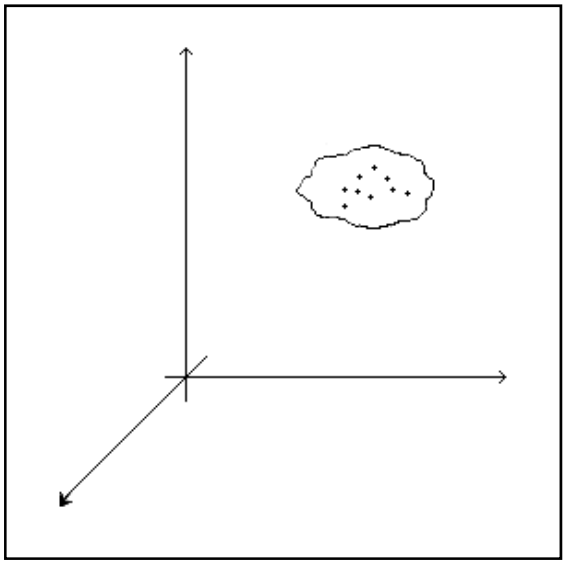

Figure 6: Image Space and face cluster

The dimension of the image space is **w*h**. Of course, all the pixels of a face are not relevant, and each pixel depends on its neighbours. So, the dimension of the face space is less than the dimension of the image space. The dimension of the face space cannot be determined, but it is sure that it is to be far less than that of the image space.

Principal Components Analysis aims to catch the total variation in the set of the training faces, and to explain this variation by a few variables. The fact that it reduces the dimension is important. In fact, an observation described by a few variables is easier to handle and easier to understand than if it was defined by a huge amount of variables. And when many observations, or faces, have to be processed the dimensionality reduction is of first importance.

The goal of the method presented here, Principal Components Analysis, is to reduce the dimension of a set or space so that the new basis better describes the typical ‘models’ of the set. In our case the ‘models’ are a set of training faces. The new basis will be constructed by linear combination. Components in this face space basis will be uncorrelated and will maximise the variance accounted for in the original variables.

Face identification

(”Who am I?”) is a one-to-many matching process that compares a query face image against all the template images in a face database to determine the identity of the query face (see Fig. 3). The identification of the test image is done by locating the image in the database who has the highest similarity with the test image. The identification process is a ”closed” test, which means the sensor takes an observation of an individual that is known to be in the database. The test subject’s (normalized) features are compared to the other features in the system’s database and a similarity score is found for each comparison. These similarity scores are then numerically ranked in a descending order. The percentage of times that the highest similarity score is the correct match for all individuals is referred to as the ”top match score.” If any of the top *r* similarity scores corresponds to the test subject, it is considered as a correct match in terms of the cumulative match. The percentage of times one of those *r* similarity scores is the correct match for all individuals is referred to as the ”Cumulative Match Score”,. The ”Cumulative Match Score” curve is the rank *n* versus percentage of correct identification, where rank *n* is the number of top similarity scores reported.

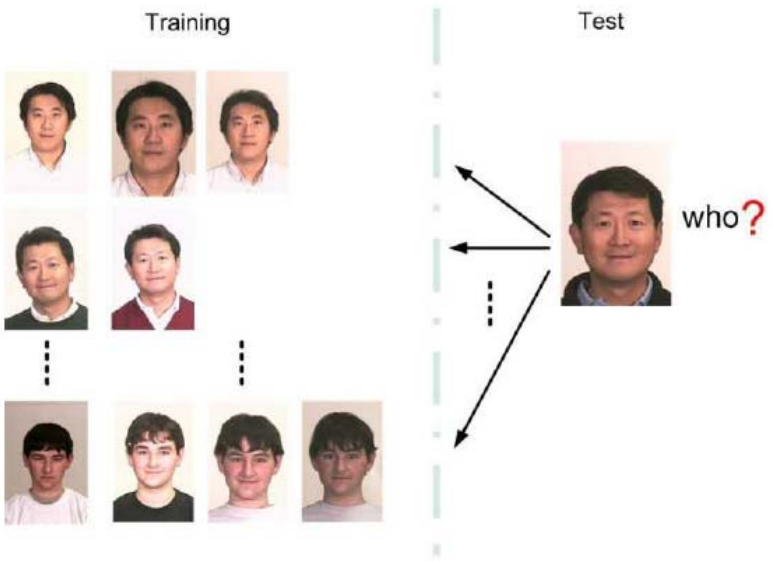

Figure 7

Face Recognition System Methods

Facerecognition is required,whichfaceour databaseisbased onmore thana match. Recognition offaceinformationrequestedfrom the otherfaces thatare normalizedandis classifiedby comparing

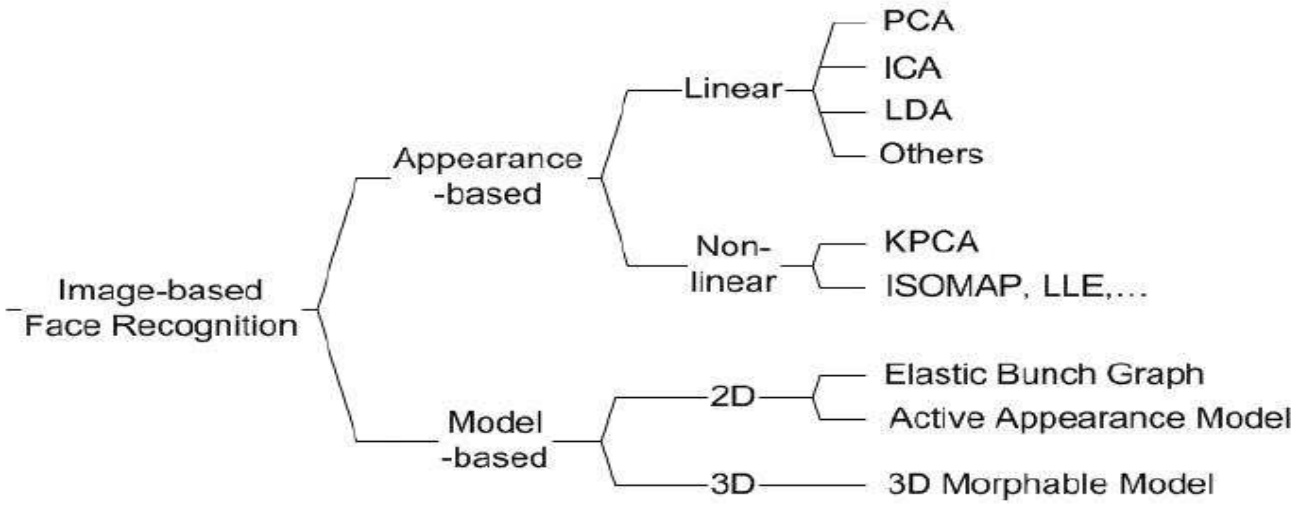

Figure 8

As mentionedin the introductionthere aremany methodsused inface recognitionsystem. When comparingtheratesofsomemethods. These methods are;

1. ICA (Independent Component Analysis)
2. Artificial Neural Networks
3. Template Matching
4. PCA(Principle Component Analysis)

1. ICA (Independent Component Analysis)

Independentcomponent analysismethod,a multivariatestatistical method thatdetermines thestructure ofthe data setthat is hidden. ICAmethod,the storedfactorsare defined asindependent components,the unknowndata forobservationby mixing withstirringto formthe matrix. The observationvector x, the source vector,s, andmixingwiththe matrixAis shown inthe followingformulais obtained asthe mainimpuritymodel. ICAmethodusingobjectiveobservation vector, is toestimatethe unknownvalues ofA ands.

$$X=A.s$$

Linear ICA in a square matrix of mixing, the resources normal (Gaussian) distribution is considered to have a non-ICA high degree of information for the calculation of the individual components must be known. Assuming normal distribution of data is appropriate, this information is zero. Therefore, ICA does not result in the data with normal distribution. Acceptance of the mixing matrix is square, with the number of resources means that the same number of independent components. This assumption facilitates estimation of the matrix decomposition. Because parsing, such as the mixing matrix in the formula below is based on inverse matrix

$$S=W.x,\ W=A^{-1}$$

WhereW isthe matrixdecomposition. ICA is statistical dependencebetween sourcesto minimizetheobjectivepresence oftheWmatrix.

ICA is alsowidely usedtwodifferent architecturalstructures. Displays thevariables inthe firststructure, the pixel valuesrepresenttheexperimentalvalues. In thisapproach,the imagesare sensitive tothe independenceorthe independence ofthe functions ofthe image. Accordingto this approach,the valueof a pixelin an imagepixelvalue ofthe otherimage,the samecannot be predicted, these two imagesisindependent of each other. The secondstructure,the firstexposureconditionistrans. Pixel values of thevariablesin this case; the images arein the experimentation. In thisapproach, the pixelsare sensitive tothe independenceorthe independence ofthe functions ofpixels. According to this approachis similarto each otherin the same imagepixels arepixels thatcannot be predictedusingthe value of oneand the otherisindependent of each other.

2. Artificial Neural Networks

In general,network, learnthenecessaryinput to outputmap. Changethe weights ofthe learnedassociationandreciprocal linkmapis a method usedto correctthis error. Each inputresulting fromthe requiredoutputare compared andmodifiedweightsusedin order to obtaincorrectoutput. Neuralnetworks,according to the functionnodetype layoutresults. Renewal ofneurons inrulestructureis completely dynamic.

The simpleststructurefunction ofthe thresholdresultsfrom scratch is the individualneurons. Here,itapplies andsome internalthresholdlevels ofneuroncomparethe totalentries. If youare satisfied withlevel, ittranslates, if not suppliedis removed.

The neurons areconnected to each otherin anetworkstructure is created.

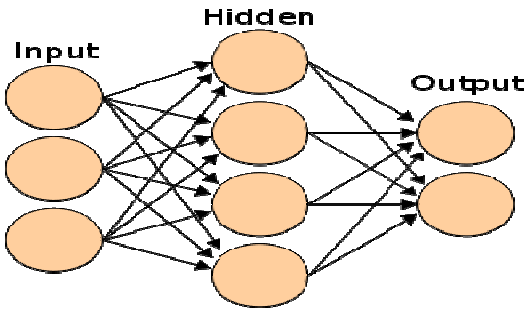

Figure 9

3. Template Matching

Templatematching,with the aid ofa computeras an approach toface recognitionprocess, based onthe entireimageusingthe grey-level patternrecognitionis a very simplemethod.The simplestversion of thetemplatematchingmethodcomprising the values ofpictureintensityrepresented asamatrixanda singlemouldthatis compared tothe whole face. There aresmarterwaysof coursethat you can applytemplatematchingmethod. For example,a matrix comprisinga differentviewdirection ofthe greylevelsused to makedetection. Another importantmethodisto usemultiple diestoa singledirection ofview.

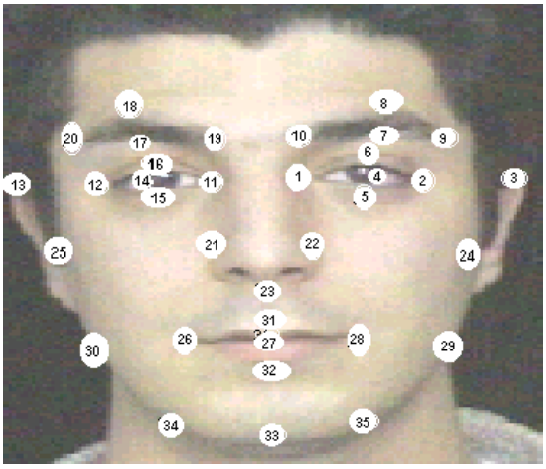

Figure 10

feature	Distance	feature	Distance
1	0.5*((1,2)+(11,12))	16	0.5*((25,30)+(24,29))
2	0.5*((5,6)+(15,16))	17	0.5*((30,34)+(29,35))
3	(3,13)	18	0.5*((1,22)+(11,21))
4	(24,25)	19	((10,19)
5	(29,30)	20	0.5*((2,9)+(12,20))
6	(34,35)	21	0.5*((9,10)+(19,20))
7	(26,34)	22	0.5*((11,19)+(1,10))
8	(28,35)	23	0.5*((6,7)+(16,17))
9	(26,28)	24	0.5*((7,8)+(17,18))
10	(27,31)	25	0.5*((18,19)+(8,10))
11	(27,32)	26	0.5*((18,20)+(8,9))
12	(32,33)	27	(11,23)
13	(23,31)	28	(1,23)
14	(21,22)	29	0.5*((1,28)+(11,26))
15	0.5*((13,25)+(3,24))	30	0.5*((12,13)+(2,3))

Table 1

4. PCA(Principle Component Analysis)

Turk and Pentland [1991] developed a Principal Component Analysis approach not only to recognize faces (see Section 2.5.1), but also to track and detect faces in an image or sequence of images. By constructing a face-space of pictures that are known to be faces, sections of a larger image can be compared to the face-space, and if the distance between the section of the image and the face space is smaller than some threshold, we can conclude that that area of the larger picture contains a face. PCA is currently one of the most promising areas in face recognition and detection [GottumukkalandAsari 2004]. While it has some drawbacks, it is fast and can produce excellent results assuming thetraining data is in the correct form.PCA also known as KarhunenLoeve projection. PCA calculates the Eigen vectors of the covariance matrix, and projects the original data onto a lower dimensional feature space, which is defined by Eigen

vectors with large Eigen values. PCA has been used in face representation and recognition where the Eigen vectors calculated are referred to as Eigen faces. In gel images, even more than in human faces, the dimensionality of the original data is vast compared to the size of the dataset, suggesting PCA as a useful first step in analysis. There are many approaches to face recognition ranging from the Principal Component Analysis (PCA) approach (also known as Eigen faces). Prediction through feature matching. The idea of feature selection and point matching has been used to track human motion. Eigenfaces have been used to track human faces. They use a principal component analysis approach to store a set of known patterns in a compact subspace representation of the image space, where the subspace is spanned by the Eigen vectors of the training image set. PCA is a useful statistical technique that has found application in fields such as face recognition and imagecompression, and is a common technique for finding patterns in data of high dimension. The basic goal is to implement a simple face recognition system, based on well-studied and well-understood methods. One can choose to go into depth of one and only one of those methods. The method to be implemented is the PCA(Principle Component Analysis). It is one of the more successful techniques of face recognition and easy to understand and describe using mathematics. This method involves using Eigen faces.