MULTIMODAL BIOMETRIC VERIFICATION AND IDENTIFICATION USING FACE AND HAND

by

Elif Sürer

B.S., Computer Engineering, Boğaziçi University, 2005

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computer Engineering
Boğaziçi University
2007

MULTIMODAL BIOMETRIC VERIFICATION AND IDENTIFICATION USING FACE AND HAND

APPROVED BY:

Prof. Lale Akarun ....
(Thesis Supervisor)

Prof. Sadık Fikret Gürgen ....

Prof. Bülent Sankur ....

DATE OF APPROVAL: 31.07.2007

ACKNOWLEDGEMENTS

I would like to thank to my thesis supervisor Prof. Lale Akarun for her endless support, encouragement and guidance. I would like to thank Prof. Sadık Fikret Gürgen and Prof. Bülent Sankur for participating in my thesis committee and for their revisions on the thesis.

I appreciate the commitment and endless help of Dr. Berk Gökberk and Helin Dutağacı. I want to thank PILab members - Oya Aran, İsmail Arı, Mehmet Gönen and Aydın Ulaş - for their friendship and encouragement. I am grateful to the 40 people who permitted their face and hand data to be acquired. I appreciate their participation in the questionnaire and their patience. Also, I would like to thank Kemal Doğan for encouraging me to pursue my thesis on multimodal biometry.

I am grateful to my family and my beloved one for their endless love and support.

ABSTRACT

MULTIMODAL BIOMETRIC VERIFICATION AND IDENTIFICATION USING FACE AND HAND

The goal of this thesis is to develop user interfaces that identify and verify legitimate users in real-time due to their biometric characteristics. Face and hand geometry were the biometric identifiers that were used in the applications. Application programs using these biometric identifiers were developed in unimodal and multimodal cases.

In this study, six Java applications were developed to capture users' hand and face images in real-time, to enrol, to verify and identify them. All of the applications were designed as wizard applications and were built upon the same scenario.

In the face recognition, Gabor jets were used to extract feature values of the face images. Matching the person was done according to the Gabor based feature vectors. In the hand recognition task, hand shape-based approach was applied on the hands. Hand segmentation, normalization and feature extraction were the important steps of this approach.

In order to test the performance of the applications, data were acquired from 40 people. Recognition performances of the algorithms were analyzed for face, hand and fusion of two modalities. Besides performance, user evaluation was carried out in order to improve the usability of the user interface. A questionnaire regarding the speed, user interface and usability of the applications were collected by the people whose data were acquired. The interface was revised after the recommendations of the users.

ÖZET

YÜZ VE EL KULLANARAK ÇOKLU ŞEKİLDE GERÇEKLEME VE ÖZDEŞLEŞTİRME

Bu tezin amacı, meşru kullanıcıları biyometrik özelliklerine göre özdeşleştiren ve gerçekleyen gerçek zamanlı kullanıcı arayüzleri geliştirmektir. Uygulamalarda yüz ve el geometrisi biyometrik belirleyici kimlikler olarak kullanılmıştır. Bu biyometrik belirleyicileri kullanan uygulamalar, tekli ve çoklu şekillerde geliştirilmiştir.

Bu çalışmada, kullanıcıların el ve yüz imgelerini gerçek zamanlı olarak toplamak, onları kaydetmek, gerçeklemek ve özdeşleştirmek için altı tane Java uygulaması geliştirilmiştir. Bütün uygulamalar "sihirbaz uygulama" olarak tasarlanmış ve aynı senaryo üzerine oturtulmuştur.

Yüz tanımada, yüz imgelerinin özniteliklerini elde etmek için Gabor dalgacıkları kullanılmıştır. Kişi eşleştirme Gabor tabanlı öznitelik vektörlerine göre yapılmıştır. El tanıma kısmında, el şekli tabanlı yaklaşım uygulanmıştır. Bu yaklaşımın önemli adımları el kesimleme, normalleşme ve özniteliklerin çıkarımı olmuştur.

Uygulamaların başarımını sınamak için 40 kişiden veri toplanmıştır. Tanıma algoritmalarının başarımları, el, yüz ve çoklu şekiller için incelenmiştir. Başarımın yanısıra, arayüzün kullanışlılığını artırmak için kullanıcı değerlendirmesi yapılmıştır. Hız, arayüz ve uygulamaların kullanılırlıklarını içeren bir anket sonucu, veri alınan kişilerden toplanmıştır. Kullanıcıların önerilerinden sonra arayüz tekrar gözden geçirilip düzeltilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

μ	Central moment
c	Number of distinct classes
I	Inertial matrix
R	Rotation matrix

AAM	Active Appearance Models
ATM	Asynchronous Transfer Mode
EBGM	Elastic Bunch Graph Matching
EER	Equal Error Rate
FAR	False Acceptance Rate
FBG	Face Bunch Graph
FERET	Face Recognition Technology
FMJ	Freedom for Media in Java
FRR	False Rejection Rate
ICA	Independent Component Analysis
ICP	Iterative Closest Point
ID	Identity
IR	Infrared
LDA	Linear Discriminant Analysis
LoG	Laplacian of Gaussian
PC	Principal Component
PCA	Principal Component Analysis
PIN	Personal Identification Number
ROI	Region of Interest

1. INTRODUCTION

1.1. Motivation

Identification and verification applications have become significantly important in recent years. As the application systems need to corroborate or find out the identity of their users, personal recognition techniques have become highly essential. These techniques focus on accepting the legal users of the system while rejecting non−users and impostors. Building access, computers and cellular phones are the examples of the applications which require the application of personal recognition techniques and block impostors from using their services.

For years, conventional recognition methods have been used while accessing the applications. These methods can be classified as knowledge-based and token-based systems [1]. Knowledge-based systems rely on "what users know" whereas token-based systems rely on "what users have". Examples of knowledge-based systems are personal identification numbers (PINs) and passwords, and examples of token-based systems are identity (ID) cards and badges.

These methods have not been adequate since the entrance to the applications has been easily violated. Computer Emergency Response Team at Carnegie Mellon University reports that bad passwords cause 80 per cent of all network problems [2]. Also, stolen ID cards have been a great flaw on the security of application systems. Since passwords are easily forgotten and ID cards are stolen or borrowed, finding more person-based methods has been unavoidable.

Biometric identifiers - physiological or behavioral characteristics of a human - are more trustworthy and more suitable than knowledge-based and token-based techniques in distinguishing between a confirmed person and a misleader [3]. Since biometric identification allows or disallows a person to use the system by his/her distinctive physiological or behavioral properties, deceiving the system is much more difficult

Biometric recognition systems are based on fingerprints, face, hand, iris, retina, signature, voice and so on [4]. These identifiers are widely used in commercial and non-profit applications.

Biometric systems can either identify or verify the users. In identification, the identity of an unknown user is found between all previously registered users. In verification, a claimed identity is either accepted or rejected [5].

Biometric systems include four fundamental modules. These are:

1. *Sensor module:* This module retrieves the data.
2. *Feature extraction module:* Data that have been acquired by sensor module is processed in this module and feature vectors are extracted.
3. *Matching module:* This module compares the extracted feature vectors against the ones in the template.
4. *Decision-making module:* This final module decides whether to accept or reject the user.

In order for the physiological and behavorial traits to serve as a biometric characteristic, they should be:

1. *Universal:* Everyone should have those traits;
2. *Distinct:* They should be unique to the owner;
3. *Permanent:* They should not change over a period of time;
4. *Collectable:* They should be acquired easily.

Besides, when real-time applications are taken into consideration, they should be fast, acceptable and robust enough to different methods [6].

Although biometric systems have some limits, they are better than traditional security methods since they are unique to the person. Also, they are more convenient since the need to remember passwords is not necessary. Furthermore, biometric identifiers enable negative recognition where the biometric application decided whether the person is who he or she disclaims to be [5].

The aim of this thesis is to implement a verification and authentication system which uses face and hand as biometric identifiers.

1.2. Literature

The most commonly used biometric identifiers are face, fingerprint, hand geometry, retinal patterns, iris, signature and other alternative techniques. All these methods have their strengths and weaknesses. This section gives a brief summary of the techniques and lists their characteristics.

1.2.1. Face

Face is one of the most widely used biometric characteristic of identification. Face identification applications' techniques may employ static or dynamic images with complex backgrounds. Face recognition approaches can be classified as the ones focusing on facial attributes such as the eyes, eyebrows, nose, lips, chin, and the ones focusing on overall face image and a combination of both.

Face verification steps are as follows: extracting the feature set of the user's face image and matching that feature set with the templates that are already stored [6]. In the verification or identification phase, a person's image is acquired with a normal or video camera and a template is calculated using the facial features. Currently formed template is compared with the templates that are already saved in the facial database [7].

Although the identification performances of face applications are generally acceptable, performance deteriorates when identifying faces from different views, complex background and changed illumination. In order for face recognition systems to be widely acceptable, these limitations should be overcome and the applications should detect the face automatically and recognize the face without the background, pose and illumination constraints [1]. Another difficulty arising from the nature of faces is that they change over time and with the expressions [6].

The advantages of face recognition applications are:

1. They are non-intrusive.
2. They operate at low-cost since the price of the cameras has decreased dramatically.
3. They can be used robustly in several places [7].
4. People are accustomed to presenting their faces for identification [8].

Face recognition technology has become popular with its use in criminal studies. Visionics Corp's FaceIt software has been used in Florida, Boston, Dallas and West Palm Beach since 2001 to identify wanted criminals. In January 2002, facial recognition applications have been used to scan the people watching SuperBowl. In spite of its limitations, this technology is highly promising for the future [8].

1.2.2. Fingerprint

A fingerprint is the composition of ridges and furrows on the surface of a fingertip. They are very distinct: Even identical twins have different fingerprints.

The fingerprint identification process is divided into three steps: image acquisition, feature extraction and matching. Fingerprint image of a person is acquired with a digital scanner that records the characteristics of the fingerprint (whorls, arches, loops, patterns of ridges, furrows and minutiae). After the images are acquired by fingerprint sensors, feature values of the image are calculated in the feature extraction module and stored. Feature values of the image are the position and orientation of minutiae points. In the matching process, the stored data are processed with the programs that search for the similar minutiae points and their relative pattern. In order to save time, the image is converted into a character string [6][7].

Since the cost of fingerprint sensors has dropped significantly, fingerprint-based applications will be widely affordable in the near future. Besides, with multiple fingerprints, millions of people can be identified easily. Although these advantages increase the expectancy that fingerprint technology will be a leading one, this technology has also some disadvantages. Since fingerprint technology has been associated with criminal studies, people hesitate to give their fingerprints. Also, fingerprint identification requires huge computational power. Finally, for a small fraction of people fingerprint identification is not suitable because of genetic or environmental issues [3].

Fingerprint technology has been widely used at Asynchronous Transfer Mode (ATM) transactions, over the Internet and criminal issues. Some fingerprint readers also check for blood flow to make sure that a fake finger is not used [7]. Since this technology is widely used, the price of fingerprint readers has decreased dramatically [8].

Also, the hesitation of giving fingerprint has begun to diminish through years although people are still sceptical. In 1997, 75 per cent of 1000 adults questioned responded that they would be comfortable on submitting their fingerprints. 20 per cent of the respondents claimed that fingerprint submission reminds them of crime [8]. Although these statistics are very promising, fingerprint technology has still not been widely accepted.

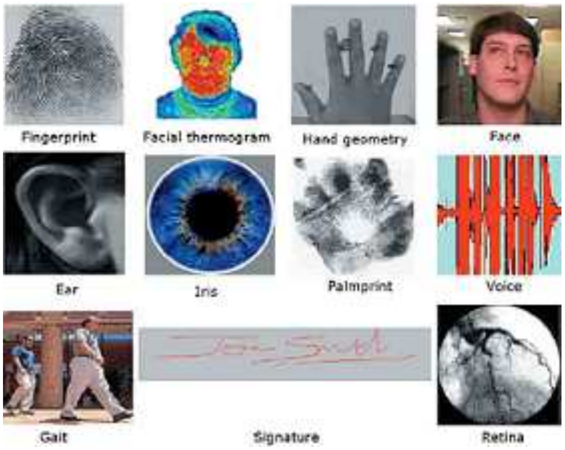


Figure 1.1. Commonly used biometrics [5]

1.2.3. Hand Geometry

Hand shape and size of fingers are also regarded as biometric characteristics. Recognition by hand is a simple, easy to use and cheap technique, so it is used at hundreds of locations throughout the world.

The user places his hand on a metal surface with or without pegs on that guide the user on the correct alignment of the hand. The device reads features of the hand and they form a template in order to be stored in a database or to be compared [7].

The main disadvantage of this technique is its low identification capability. Since hand geometry technique is not very distinctive, using this technique in verification mode rather than identification mode will increase the success of the applications.Also, the size of hand geometry causes this technique to be restricted to certain applications [1][3].

Since hand size causes the technology to be used in certain applications, some verification systems use a few fingers instead of hand [6]. Although these devices are smaller than the ones used for hand geometry, they are still larger than the devices used for fingerprint, face and voice [1].

Although hand geometry is not as distinctive as fingerprint, researchers have developed this method as an alternative to fingerprint. People may harm their fingerprints with certain chemicals. Also, fingerprint enrollment is very difficult in dry weather. Since hand geometry is less likely to change over time, it can be an important alternative to fingerprint.

Companies use hand geometry devices in closed-end applications like access to certain facilities [8].

1.2.4. Retinal Pattern

Retinal recognition can be regarded as “eye signature” of the person. Since it is protected under the eye, it is not easy to change or replicate. Therefore, retinal recognition is one of the most secure biometric identifiers [6].

Veins under the surface of the retina form unique patterns and these patterns can be acquired by projecting light to the eye and capturing the image of the retina with a retinascope. The scanning device uses infrared (IR) light to illuminate the retina and when it is reflected back, the devices uses the reflected light to extract features by using different algorithms [7]. In order to get a correct image, the user should look into an eye-shape structure and focus on a certain spot.

Although retinal pattern is the one of the most secure biometric identifiers, it has several important disadvantages. The degree of user cooperation is very high, retinal scanners are expensive and retinal vasculature can cause some medical conditions such as hypertension. Because of these disadvantages, the public acceptance of this biometric is low and it is not widely used [1][3].

This biometric technology is used in places where high security is compulsory. Retina scans are used for high-security places such as government buildings, military quarters, prisons and banks. Japanese banks have been using this technology in their ATMs since mid 1990s [8].

1.2.5. Iris

Iris is the region of the eye bounded by the pupil and white of the eye. Iris carries such distinctive information on people that even twins have different irises. Arching ligaments, furrows, and ridges are the important features that make the iris unique.

Iris scanning is less intrusive than retina. Although capturing the iris is easier than retina, it still requires high degree of user participitation. In iris scanning, when the user places his head in front of the device, he should see the reflection of his iris in order for a clear image to be obtained [7]. Iris scanners use cameras that analyze the patterns of color and the texture of iris [8].

In order to gain public acceptance, iris scanners' price should be decreased and capturing process should become more user-friendly [1][3].

1.2.6. Signature

Although every person has a unique handwriting style, the identification results of signature is not sufficiently high. This stems from the fact that, no two signatures of a person are identical.

There are two methods on signature-based identification: static and dynamic. Static signature identification uses only the shape features of a signature, while dynamic signature identification uses both shape and dynamic features such as acceleration, velocity and pressure. Dynamic signature method is more distinctive than the static method since dynamic signature is a behavioral biometric and although it is easy to duplicate signatures, it is not easy to duplicate behavioral characteristics.

Signature acquisition is done with a special device. The device consists of a pen (stylus) and a writing tablet that is connected to a computer. After the user signs on the tablet, the system collects all the featural information, forms a template and stores it in a database [8].

In order for the enrolment to be a proper one, the user has to sign the tablet multiple times. The advantage of signature identification is its ease of use and non-invasive nature. The main disadvantage of this method is its dependency on the length of the signature. If the signature is too long, many inconsistent behavorial features are acquired by the system which is difficult to identify. If the signature is too short, enough number of data points may not be collected [7].

1.2.7. Speech

Vocal tracts, mouth, nasal cavities and lips are the reasons of individuality of human speech.

Voice recognition matches a particular voice sample against the ones in the database. Voice systems are trained with the user's voice at the enrolment time and several enrolment sessions are necessary. At the feature extraction level, formants or sound characteristics that are unique to the user are measured. Finally, pattern matching algorithms are used to find the individual [6].

Although speech is distinctive, it is not optimal for large-scale recognition. Speech-based verification can be text dependent or text-independent. In text-dependent verification, the individual is verified by the features of a predetermined phrase. Although text-independent verification is much more difficult, it is more secure.

People are willing to accept speech-based biometric systems, but those systems have some disadvantages such as poor accuracy, sensibility to background noise and dependency on the emotional state of the user [3].

Voice recognition systems are suitable for phone-based applications but recognition performance decreases due to the microphone and communication channel quality [1][7].

Another disadvantage of voice verification systems is that they can be easily tricked by imitators or by someone who records the subject's voice beforehand. Voice recognition techniques are used as a mix of other biometrics and they are widely used in web-based transactions [8][9].

1.2.8. Alternative Techniques

Recently, alternative techniques such as a person's gait, odor and keystrokes have begun to be considered as alternative identification techniques [8].

1.2.8.1. Odor. Each object spreads around an odor which is a chemical composition. By using an array of chemical sensors, each sensitive to a certain group of compounds, the objects may be distinguished. The identification performance of the system diminishes with deodorant smells [1][6].

1.2.8.2. Keystroke. Identification is based on a hypothesis that each person types on a keyboard in a different way. Although this behavioural biometric is not unique, it still helps to distinguish between people. The keystrokes of a person are monitored unobtrusively as the user is typing [3].

1.2.8.3. Gait. This is one of the newer technologies and is yet to be researched in more detail. Basically, gait is the peculiar way one walks and it is a complex spatio-temporal biometrics. It is not supposed to be very distinctive but can be used in some low-security applications. Gait is a behavioral biometric and may not remain the same over a long period of time, due to change in body weight or serious brain damage. Acquisition of gait is similar to acquiring a facial picture and may be an acceptable biometric. Since video-sequence is used to measure several different movements this method is computationally expensive [6].

1.2.9. Classification of Biometric Systems

The biometric systems that use only one type of biometric identifier are considered as unimodal systems whereas systems that use more than one type of biometric identifier as multimodal systems. Since each biometric identifier has its disadvantages, combining them in multimodal systems achieves greater recognition results.

1.2.9.1. Unimodal Biometric Systems. Biometric systems that use only one type of biometric identifier has the following restrictions [1]:

1. *Noise in sensed data:* If the sensed data are noisy, biometric data may be incorrectly matched with the templates in the database thereby incorrectly rejecting or accepting the user.
2. *Intra-class variations:* The biometric data acquired during enrolment sessions may be very different from the verification. This can stem from the user's incorrect interaction with the sensor and modified sensor properties during the verification.
3. *Distinctiveness:* The biometric trait may have inter-class similarities which degrades the recognition performance.
4. *Non-universality:* While each user is expected to have the biometric identifier, a subset of users may not have some certain biometric traits.
5. *Spoof attacks:* An imitator may try to spoof the biometric system in order to enter the system like a legal user.

1.2.9.2. Multimodal Biometric Systems. A multimodal system combines two or more individual biometric identifiers. Multimodal systems improve identification performance, work for a large number of people and increase the robustness against spoof attacks. The combinations can be done in the following levels:

1. **Decision Combination:** Each biometric device makes the decision of accept and reject individually. The decisions are combined with the methods of voting, weighted sum or other methods.

2. **Score Combination:** Each biometric identifier performs its evaluation method and outputs a score. Then, scores of two systems are fused in order to give a single score which is compared with the acceptance threshold. This is the most common technique while combining the identifiers.

3. **Feature Combination:** Feature combination is applied when the feature sets of two individual identifiers are extracted. After the extraction, the feature sets are combined and overall performance of the system is found. This method is not very useful since in order to combine feature sets, good normalization is required, but this is not very likely. Also, most biometric systems do not provide data on feature extraction level [10][4].

1.3. Focus of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, facial recognition algorithms and the approach that is implemented in this thesis are explained in detail. In Chapter 3, an overview of hand recognition algorithms is given and the techniques implemented in the thesis are presented in detail. The developed user interface applications are explained in Chapter 4. In Chapter 5, the performance evaluation of the applications and the results of usability tests are discussed. Finally, conclusions are given in Chapter 6.

2. 2D FACE RECOGNITION

2.1. Literature

As the biometric applications have received great attention, 2D face recognition technology has also developed during the past few years. Several approaches and algorithms have been improved and implemented. Face recognition is the recognition of the identity of a person based on a 2D image of his face. The approaches in 2D face recognition can be classified into two: appearance based and model based. The most important appearance based algorithms are Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Active Appearance Models (AAM) and Elastic Bunch Graph Matching (EBGM) are the strong examples of model based approaches.

2.1.1. Appearance Based Recognition

Appearance-based approaches focus on representing an object in terms of different object views. In this approach, the images are represented as high-dimensional vectors, whose components are pixel intensities, so they can be regarded as points in a high-dimensional vector space.

Statistical techniques are used to analyze the object image vectors and feature space in different contexts. If a test image is acquired, the matching criterion is built upon the similarity in the feature space.

2.1.1.1. Eigenfaces. The "Eigenface" method is synonymous with "Principal Component Analysis" or PCA. Principal Component Analysis is a statistical method that aims to reduce the large dimensionality of the data space [11]. The PCA method achieves this by projecting the input image vector onto a small subset of eigenvectors, as seen in Figure 2.1, and using the projection magnitudes.

The eigenvectors correspond to the maximum magnitude eigenvalues of the covariance matrix. Therefore, they are the directions that explain the data best.

Eigenfaces are commonly used in face representation, recognition and detection. Eigenface representations are very good at reducing the sensitivity to noise. They work well under background changes, blurring and partial occlusions [13]. The most useful applications that use the eigenface approach are surveillance, identification and security applications [11].

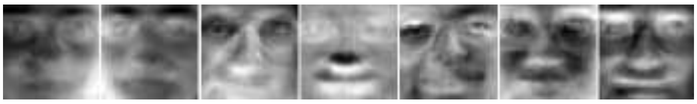


Figure 2.1. Eigenvectors corresponding to the seven largest eigenvalues [14]

2.1.1.2. Fisherfaces. Linear Discriminant Analysis or Fisherfaces approach aims to find an effective face vector space representation. In the Fisherface algorithm, different classes with different statistics are defined and training images are divided into the corresponding classes. Then, techniques that are used in the Eigenface algorithm are applied.

LDA maximizes the ratio of between-class to within-class distribution [15]. The difference between PCA and LDA is that PCA represents the data in orthogonal linear space whereas LDA represents it in a linear separable space. PCA based systems are found to work better in databases of small number of classes and LDA is better for the large number of classes. LDA algorithm is used in speech recognition and in pattern recognition. However, LDA algorithm has problems in reducing the dimensionality of the image data [16].

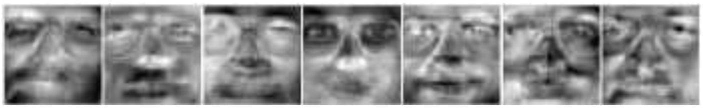


Figure 2.2. First seven basis vectors [14]

2.1.2. Model Based Recognition

Model-based face recognition aims to construct a model of the human face to handle facial variations. This approach is composed of three steps. These are:

1. Constructing the model;
2. Fitting the model to the given face;
3. Finding the similarity between the query face and stored faces by using the parameters of the fitted model.

2.1.2.1. Active Appearance Model. Active Appearance Model uses all the information of the target object, it is not based on near modeled edges. An AAM has grey-level appearance of the target object and a statistical shape model. In order to find a correct match, model parameters that minimize the difference between the image and a synthesized model's projected image are sought. To find the correct parameters, AAM measures the differences and uses the model to correct the current parameters in an iterative approach. A good fit is obtained in a few iterations [17].

In the Active Shape Model approach, the face is found by Active Shape Model search and the shape parameters are found. Then, the face is deformed into an average shape to find the grey-level parameters. Shape parameters and grey-level parameters are the classification parameters of the Active Shape Model.

When a new face is given, the shape and grey-level model parameters are extracted. By using the Mahalanobis distance measure, the nearest shape is found in the training set.

The algorithm used in searching and fitting the model is as follows:

1. Place the appearance model in the image, e.g. at the center of the image.
2. For the current model and the target object:

   (a) Find the distance between the target object and the current model.
   (b) Update the new values for the parameters.
   (c) Find the new difference measure. If the difference has been lowered, save the model.

3. Iterate until convergence has been reached [19].

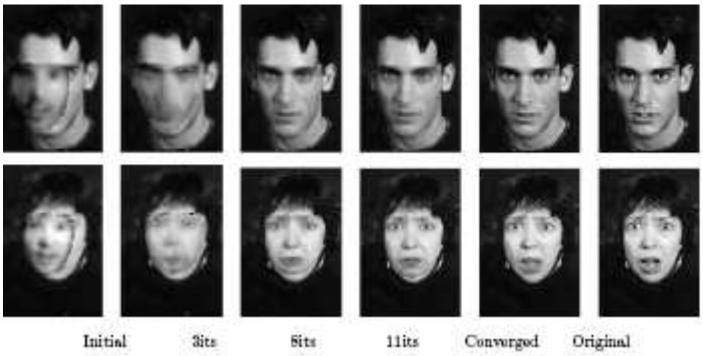


Figure 2.3. Examples of the model fitting iterations [14]

2.1.2.2. Elastic Bunch Graph Matching. Real face images have some characteristics that are not handled by linear analysis. These characteristics are pose, illumination and expression. In order to solve this issue, Gabor wavelet transforms that project the face into a grid are used. A Gabor jet is a node on the elastic grid. Gabor jet's response to a Gabor filter is used to extract the features of the image. Recognition is based on the similarity of the jet's responses. The most difficult problem of Elastic Bunch Graph Matching is the localization of landmark points [20].

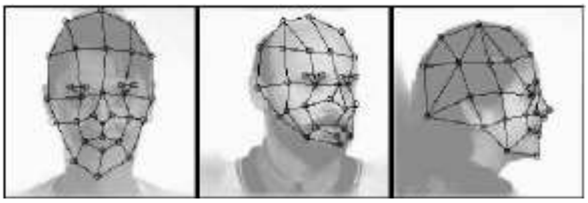


Figure 2.4. Multiview faces overlaid with labeled graphs [22]

Elastic Bunch Graph Matching is based on comparing the images with graphs and creating new graphs. A single labelled graph, which has some nodes that are called jets on it, is matched on an image. Gabor wavelet transform is applied on the jets and the response of the jet after convolution is calculated. Comparing the images is done via comparing the similarities of the jets.

While a labelled graph represents a single object, bunch graphs are used to represent a whole class of objects. When matching is done on a bunch graph, similarity measure becomes more complex since it is possible to choose a different jet for each bunch [21].

The first graph definition of the system is done manually in three steps:

1. A set of fiducial points are marked. These points define the important features of the face such as left and right pupil, the tip of the nose, the top of the chin and so on.
2. Between fiducial points and edge labels, edges are drawn by calculating the differences between nodes.
3. Gabor wavelet transform is calculated and jets are extracted.

Graph similarity is an important measure in Elastic Bunch Graph Matching. Similarity between and image graph and face bunch graph of the same pose is calculated. The similarity is dependent on jet similarities and grids.

The aim of Elastic Bunch Graph Matching is to find fiducial points and to extract a graph that minimizes the similarity between face bunch graph. The matching procedure is composed of four steps:

1. In order to find the approximate face position, an average graph is calculated by averaging the magnitudes of jets in each face bunch graph.
2. Average graph is rescaled and refined due to different positions and sizes.
3. Size and aspect ratio is rescaled by relaxing the x any coordinates of the graph.
4. The position of each node is updated in order to increase the similarity to face bunch graph.

The final graph is defined as the “image graph” and used in face recognition. After extracting image graphs, recognition is done by comparing an image graph to all model graphs and choosing the one with the highest similarity value.

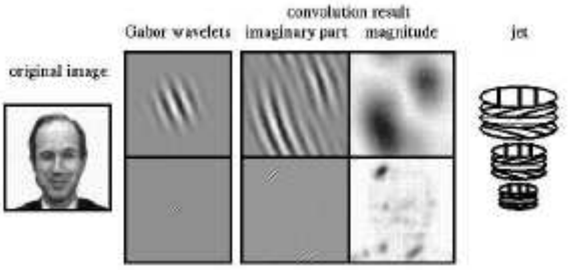


Figure 2.5. Jet [22]

2.2. Gabor Based 2D Face Recognizer

2.2.1. 2D Gabor Wavelets

2D Gabor wavelets are commonly used in feature-based face representation. Since Gabor kernels are similar to cells in the visual cortex and they have properties of multi-orientation and multi-resolution, they are highly preferable. Because the full convolution of Gabor kernels with the image is not a feasible task, sparse sampling is applied.

The face image is convolved with the 2D Gabor kernels at convolution points. The result of the convolution operation for each convolution point is calculated and a local feature vector is calculated.

The image is convolved with Gabor kernels having eight different orientations and five different frequencies. Several Gabor kernel resolutions have been tried in order to find the optimum resolution. The result of Gabor convolution is a vector of size 40 for each convolution point.

In order to represent the face region, a rectangular sampling grid is placed over the face image [23].

In order to ease the problem of landmark location, we have used a square mesh structure to represent the face and the feature extraction has been uniform. 2D Gabor wavelet has been applied at each grid point. The response of the wavelet describes a

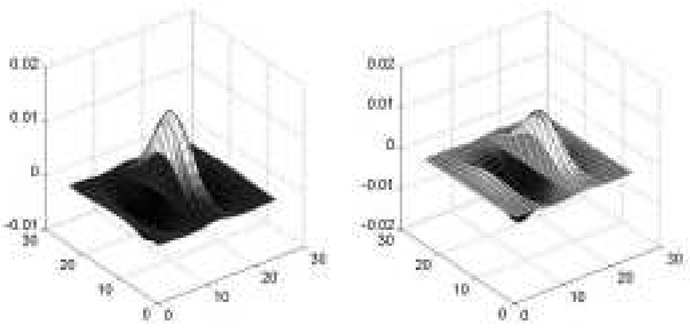


Figure 2.6. Real and imaginary parts of a Gabor kernel

patch of gray values in an image $I(\vec{x})$ around a given pixel $\vec{x} = (x, y)$. The response is a convolution

$$J_j(\vec{x}) = \int I(\vec{x}')\psi_j(\vec{x} - \vec{x}')d^2\vec{x}' \tag{2.1}$$

with Gabor kernels

$$\psi_j(\vec{x}) = \frac{{k_j}^2}{\sigma^2}\exp\left(-\frac{(k_j)^2x^2}{2\sigma^2}\right)\left[\exp(i\vec{k_j}\vec{x}) - \exp\left(-\frac{\sigma^2}{2}\right)\right] \tag{2.2}$$

in the form of plane waves with wave vector $\vec{k}_j$. The wave vector is restricted by a Gaussian. A discrete set of 5 different frequencies, with $v = 0, ..., 4$, and 8 orientations, with $w = 0, ..., 7$

$$\vec{k}_j = (k_{jx}, k_{jy}) = (k_v \cos\varphi_w, \sin\varphi_w), k_v = 2^{-\frac{v+2}{2}}\pi, \varphi_w = w\frac{\pi}{8} \tag{2.3}$$

are applied where $j = w + 8v$. The width $\frac{\sigma}{k}$ of the Gaussian is controlled by the parameter $\sigma = 2\pi$ [24].

After the feature values of the face image is acquired, i.e. the responses to the Gabor kernels, each image is represented as a $40 \times K^2$ size feature vector, where K^2 is the square of the grid size. In the verification phase, L_2 normalization is calculated on the query image and template images. If the calculated distance is less than the threshold, the person is accepted, otherwise rejected.

2.2.2. Face Recognizer's Processing Steps

2.2.2.1. Placing a Grid on the Face. Since convolving Gabor masks with the full image is not a feasible task, a grid is placed on the face. The corners of the grid are the points of convolution with the Gabor mask. Each image is represented with the feature vectors calculated in the convolution.

The grid sizes that have been used in this study have been 7 × 7, 9 × 9 and 15 × 15. The optimum grid size has been found by iteratively testing the whole grid sizes with the Gabor masks. If the grid size is chosen as 7 × 7, 49 convolution points are present, so the size of the feature vector will be gabor mask size-by-49. An example grid is seen below:

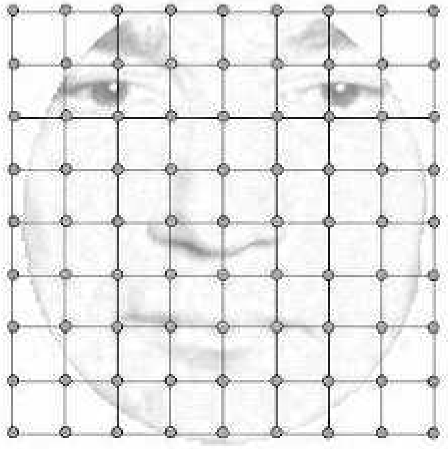


Figure 2.7. Example of a grid

2.2.2.2. Choosing the Gabor Mask. The size and type of the Gabor masks are the important parameters of the face recognition. The sizes of the Gabor masks that have been used in this study have been 3 × 3, 5 × 5, 7 × 7, 9 × 9, 11 × 11, 13 × 13, 15 × 15, 21 × 21, 29 × 29, 33 × 33. Each of the Gabor structures contains masks for five different frequencies and eight different orientations. Therefore, each Gabor mask is a structure having 40 different masks each having specified grid sizes. For example, a Gabor structure with the size of 3 × 3 size contains 40 masks each having the size of 3 × 3. These structures contain both real and imaginary numbers.

2.2.2.3. Convolution with the Gabor Mask. After the grid is placed on the image and the Gabor mask has been chosen, convolution operation starts. The convolution points are the corners of the grid. Convolution is done for both real and imaginary values of the Gabor mask. Then, the magnitude of the real and imaginary convolution is calculated.

In Figure 2.8, the effect of Gabor masks can been. The top row shows Gabor-filtered images for frequencies 1, 2, 3 and 4 and the bottom row shows Gabor-filtered images for orientations 0, 2, 4 and 6.

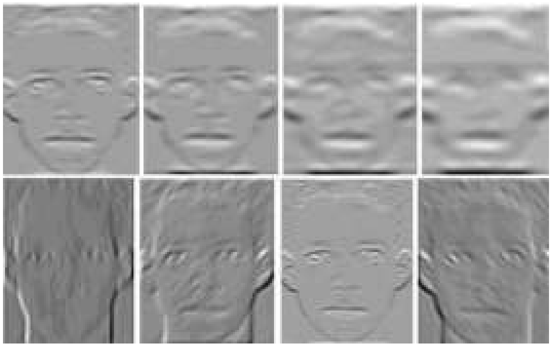


Figure 2.8. Effect of Gabor masks

2.2.2.4. Linear Normalization. Recognition process is done by comparing the feature vectors of the images that are formed with Gabor convolution. Euclidean distances of the feature vectors are calculated and the image with the minimum distance is accepted as the near image if the distance is not greater than a given threshold. The threshold is used in order to have a reject option.

For vectors $x = [x_1\ x_2\ ...\ x_n]^T$ and $y = [y_1\ y_2\ ...\ y_n]^T$, Euclidean distance $d(x, y)$ is calculated with:

$$d(x, y) = \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2} \tag{2.4}$$

3. HAND RECOGNITION

3.1. Literature

Although hand geometry is accepted as a medium security biometric identifier, it has several important properties:

1. Low implementation cost
2. Fast results stemming from low-computational cost algorithms
3. Low size of the template
4. Acceptable by the users
5. Not regarded as an identifier related with police, justice and crime [25].

The recognition with hand geometry is done with several approaches such as hand shape, palm shape, finger size and palmprint.

In the work of Bulatov [26], hand images have been acquired with a document scanner. 30 features are extracted from the hand and for each image, a bounding box is found. The similarity measure of this verification method is the distance to the bounding box. In verification, a previously decided threshold is used as the similarity measure. In the classification, feature vectors are mapped to a higher dimension by Gaussian kernels and Voronoi diagram of the centers are chosen as the metric for classification.

In Chin-Chuan Hana's work [27], a scanner-based authentication system is implemented by using the palm-print features. In the enrolment stage, the images are collected from the people and preprocessing, feature extraction and modelling are applied on the acquired images. In the verification module, an unknown image is preprocessed and its features are extracted in order to compare with the templates stored in the database and decide whether the user is a legitimate one or not.

In the identification mode, preprocessing, feature extraction and template generation are done. Preprocessing step includes thresholding, board tracing, wavelet-based segmentation and region of interest (ROI) location operations. Feature extraction is done via Sobel and morphological operations. In order to extract a square region that possesses the palm-print data, image thresholding, border tracing, wavelet-based segmentation and ROI location are applied. Image thresholding operation is used to binarize the gray images to have hand images. Border tracing lets the boundary of hand image to be found. Wavelet-based segmentation is used to find the fingertips and four fingerroots of the hand.

In Jain et al's method [28], before extracting the features of the hand, hand shape is aligned. Proposed method includes the following operations:

1. Removing the pegs
2. Extracting the contours
3. Extracting and aligning the fingers
4. Computing pairwise distance
5. Verifying the hand

In order to find the correct correspondences between the hands, corresponding fingers are aligned by generating a match matrix and applying the Iterative Closest Point (ICP) algorithm.

Kumar's work [30] attempts to increase the performance of the palm-print based verification systems by adding hand geometry features. From a low-resolution hand image, both palmprint features and hand geometry features can be extracted. Principal lines, wrinkles, minutae and delta points are palm-print features whereas area/size of the palm, length and width of the fingers are the geometric features.

In the image acquisition and alignment phase, first binarized shape of the hand is tried to be found by using an ellipse. Parameters of the best-fitting ellipse, major axis of the ellipse and rotation angles are used in order to approximate the orientation of the hand. After finding the binarized hand, palmprint is extracted by using morphological erosion. Then, palmprints are normalized to reduce the noise on the data and illumination effects. Finally, line features of the palm-print using some directional masks or line detectors.

In Öden et al's work [31], implicit polynomials, which are frequently used in object modelling and recognition, have been applied to the hand recognition task. The most important point in using implicit polynomials is finding the implicit polynomial function or coefficient vector that best represents the object. Modelling the fingers with implicit polynomials have improved the success of the hand recognition.

In preprocessing phase, a LoG edge detector is used. By using the boundary information and signature analysis, fingers are extracted and appropriate points are found. A fourth degree polynomial was fitted to the fingers by using gradient-one fitting and 3L fitting. The results of these operations have been used to calculate the Keren's and Civi's invariants.

3.2. System Implemented

The hand recognition modality of this thesis is based on the "Shape−Based Hand Recognition" study [32] [33] [34]. The algorithm of this study focuses on hand recognition by solely using hand shape. After preprocessing the hand image, segmentation, normalization and feature extraction are applied in order for the system to recognize hands. The only requirements of the system are:

- The background color should be homogenous
- Fingers should not touch each other

3.2.1. Segmentation

Segmentation of the hand is the process of extracting the hand image from the background. This is a hard task since rings, cuffs and wristwatchs cause some artifacts and decrease the accuracy of the hand contour's delineation. Successive application of clustering and morpological operations have been the solutions to the segmentation problem.

Firstly, K-means algorithm is used in order to separate the hand foreground and the darker background. Since the result of this operation ends up with maps having holes and blobs, the largest connected components in the foreground are found and the debris is removed with area-based size filtering.

Similarly, background, which is obtained from the reverse image, is accepted as the largest connected component and area-based filtering is applied to end the segmentation process.

3.2.2. Hand Registration

In hand biometry, the most crucial step is the normalization of the hand. The aim of the hand normalization tasks of this work is to minimize intra-person variances of hand postures, finger orientations and illumination.

Normalization process involves registration of the hand (rotation and translation of hand images) and re-orientation of fingers to standardized directions. By this way, the hand is translated and rotated to a reference frame, illumination correction is applied and the fingers are rotated accordingly.

The operations of translation, rotation and illumination correction are compulsory before feature extraction is applied. If hand normalization is not applied, even the same person's hand contours may not match after different image acquisition sessions.

Registration of the hand process starts with the translation of the hand image. Translation to the centroid of the hand image and rotation toward the larger eigenvector are done with the following procedures:

The moments of the binary image, $m_{i,j}$ for objects x^i and y^j are defined with the following equation:

$$m_{i,j} = \sum \sum_{(x,y) \in object} x^i y^j \tag{3.1}$$

The centroids are calculated with:

$$\overline{x} = \frac{m_{1,0}}{m_{0,0}} \tag{3.2}$$

and

$$\overline{y} = \frac{m_{0,1}}{m_{0,0}} \tag{3.3}$$

The central moments are defined as:

$$\mu_{i,j} = \sum \sum_{(x,,y) \in object} (x - \overline{x})^i (y - \overline{y})^j . \tag{3.4}$$

Therefore, sample covariance matrix of the object (inertial matrix) is:

$$I = \begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix} \tag{3.5}$$

The orientation of the object is defined with the direction of the largest eigenvalue. The value of the rotation angle is defined as:

$$\theta = 0.5 \arctan \left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right) \tag{3.6}$$

By applying these translation and rotation procedures sequentially, the hand image is moved to a reference frame.

3.2.3. Localization of Hand Extremities

Hand extremities are the fingertips of the hands and the valley between the fingertips. Tips of the fingers and bottom of the inter-finger valleys are found by radial plotting. Radial plotting is finding the radial distance with respect to a reference point. In this study, the reference point is chosen as the first intersection point of the larger eigenvector of the inertial matrix with the wrist line. The radial sequence's minima and maxima points give the nine extremum points i.e. points corresponding to the five tips of the fingers and four inter-finger valley points.

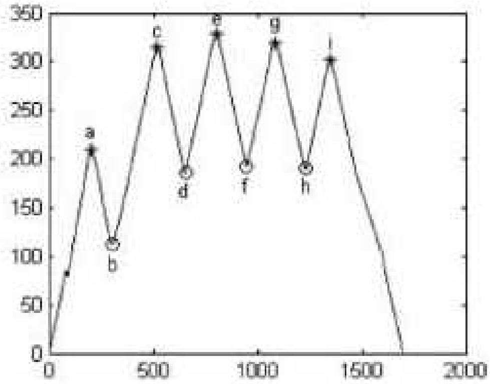


Figure 3.1. Radial plot of hand contours

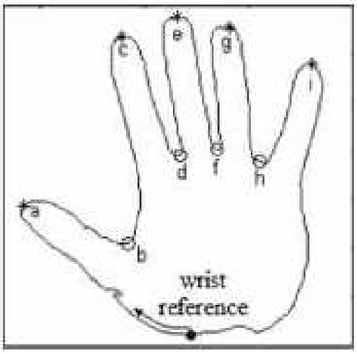


Figure 3.2. Hand extrema

3.2.4. Ring Artifact Removal

Since the rings may cause severance of the finger from the palm or an isthmus on the finger, removing this effects is important. An isolated finger can be found by the size of its connected component and background debris can be removed with morphological operations. A severed finger can also be reconnected to its palm by prolonging its straight lines till it meets the palm.

The presence of the isthmus can be detected by the contour distance to the finger's major axis. Local minimum on the left or right side of the finger is assumed as a cavity caused by the ring when the distance is over a threshold. After detecting the isthmus, it is removed by bridging over the cavities and filling the inside of the cavities.

3.2.5. Finger Registration

After finding the hand extremities and removing the artifacts of the rings, finger registration process starts. This process contains seven steps. These steps are listed as:

1. Finger extraction: Firstly, the tip point's segment is extended along the finger toward the neighbour valley points. After finding the shorter segment, the point is swung like a pendulum towards the other valley point. By this way, finger is extracted. The finger extraction method is different is different for the thumb.

2. Finger pivots: Each finger rotates around a joint, which is approximately located at a distance 20 per cent larger than the finger length. These joints are pivot points that are used to calculate the scale and orientation of the hand.

3. Hand pivotal axis: Pivot line is the line that connects the pivot points of the index finger and little fingers. This line is critical in order to register hand images to a pivot line angle, to rotate fingers and register the thumb.

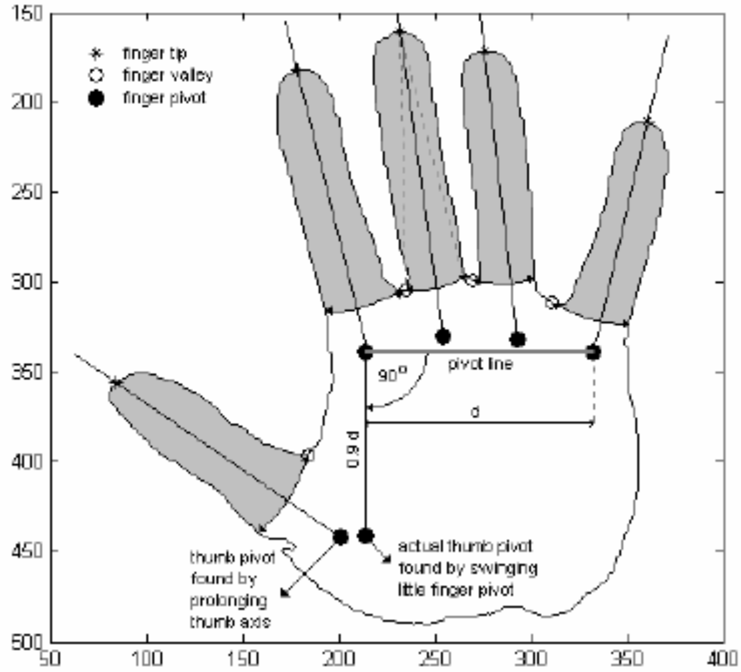


Figure 3.3. Finger extraction

4. Rotation of the fingers: The major axis of each finger is calculated from its own inertial matrix and its orientation angle θ. Each finger i is rotated by the angle $\Delta\theta_i = \theta_i - \Psi_i$, for i = index, middle, ring, little, and where Ψ_i is the goal orientation of that finger. The finger rotations are accomplished by multiplying the position vector of the finger pixels by the rotation matrix:

$$R_{\Delta\theta} = \begin{bmatrix} \cos\Delta\theta & -\sin\Delta\theta \\ \sin\Delta\theta & \cos\Delta\theta \end{bmatrix} \tag{3.7}$$

5. Processing for the thumb: Since thumb rotates with respect to two distinct joints, the motion of the thumb is handled by a rotation followed by a translation. Another difficulty with the thumb was is the stretched skin between the thumb and the index finger.

Since the stretched skin causes difficulties in determining the valley point and thumb, the thumb is assumed to be the same length as the person's little finger. The complex movement of the thumb is handled with the translation and rotation.

6. Centering and rotation of the hand: After the normalization, hands are translated to a fixed reference point in the image plane. Then, rotation is applied on the whole hand to align the hand with the fixed position.

7. Wrist Completion: The hand contours that are obtained after segmnetation have some irregularities that stem from the clothing, the difference in the angle of the forearm and the pressure applied on the imaging device.

   These irregularities cause different wrist segments in different image acquisition sessions and these different wrist segments degrade the recognition performance of the algorithm. In order to solve the irregularities in the wrist regions, a uniform wrist region which is consistent for every hand image is created.

3.2.6. Feature Extraction and Recognition

Selection of features can be done with several different methods in hand biometry. This work concantrates on a scheme that considers the whole scene image containing the normalized hand and background of the image, and applies subspace methods on it. This appearance-based method is Independent Component Analysis (ICA).

ICA is a widely used feature extraction technique in image processing. Although what we obtain is a binary image consisting of the silhouette of the normalized hand before extraction, ICA analysis can easily be enlarged to include texture and palmprint information of the hand as well.

What ICA aims to do is to find a linear tranform W for the inputs that minimizes the dependence between the outputs. Transformation matrix W is also called separating and de-mixing matrix. The formula is as follows:

$$S = Y = WX \tag{3.8}$$

where Y is the vector of outputs, S is the vector of outputs and X is the vector of inputs.

In the hand recognition part of this thesis, this feature extraction method has been used.

4. FACE AND HAND BASED BIOMETRIC VERIFICATION SYSTEM

In this study, six Java desktop applications - Face Recognition, Face Authentication, Hand Recognition, Hand Authentication, Multimodal Recognition and Multimodal Authentication- have been developed. The focus of the study is to implement face and hand recognition algorithms in real-time and to provide easy-to-use interfaces.

4.1. Technical Information

4.1.1. Development Platform

The applications have been developed in Java, using Java 2 Platform Standard Edition Development Kit 5.0 [35]. The modules regarding hand and fusion communicate with MATLAB 7.4.0 (R2007a) [36], use already written M-files of hand recognition in MATLAB and return the results of registration or recognition to the Java user interface.

4.1.2. Packages

Freedom for Media in Java (FMJ) [37] package has been used and modified in order for the application to use two USB cameras interchangeably. The databases of the applications have been designed in MySQL 5.0 [38]. JGoodies [39] package has been used to make Swing user interfaces look more elegant. JMatio [40] package has been used in order for the system to use .mat files (Gabor kernels) which are necessary for the face recognition algorithm.

Performance and recognition tests of the system have been implemented via JUnit package [41]. In order for the system to log the errors and warnings, Log4j [42] package has been used and several information messages have been inserted to the application.

4.2. System Design

All of the applications have two main tasks: One of them is registering a new user and the other one is recognizing an already registered user among the registered users' pictures.

If the application is an authentication application, the announced user is searched within the pictures of the claimed identity. If the application is an identification application, user is searched within the pictures of all registered users' pictures.

4.2.1. New User Scenario

A new user, who is not registered to the verification system chooses this option and starts the registration process.

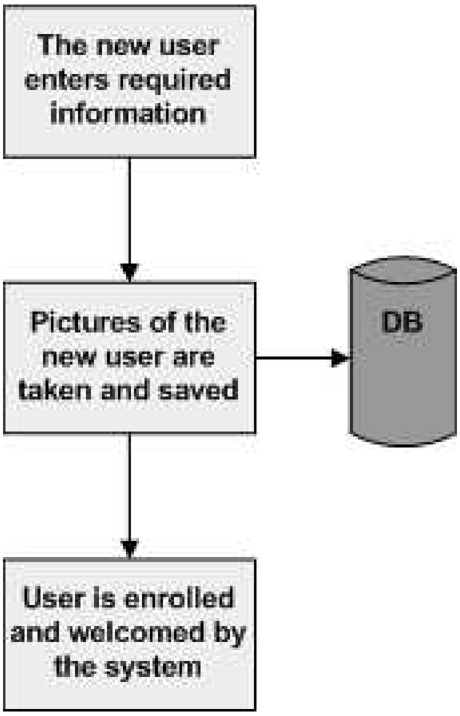

The new user enters required information
Pictures of the new user are taken and saved
DB
User is enrolled and welcomed by the system

Figure 4.1. New user workflow

The user enters his/her name, surname, email and department and starts the picture capturing process. After the picture is saved, features of the picture are extracted and saved in the database. Finally, the new user is welcomed by the system.

4.2.2. Existing User Scenario

An already registered user chooses the “Existing User” option in order for the system to recognize himself/herself. The user saves his/her current picture in order for the current picture to be matched by the verification system.

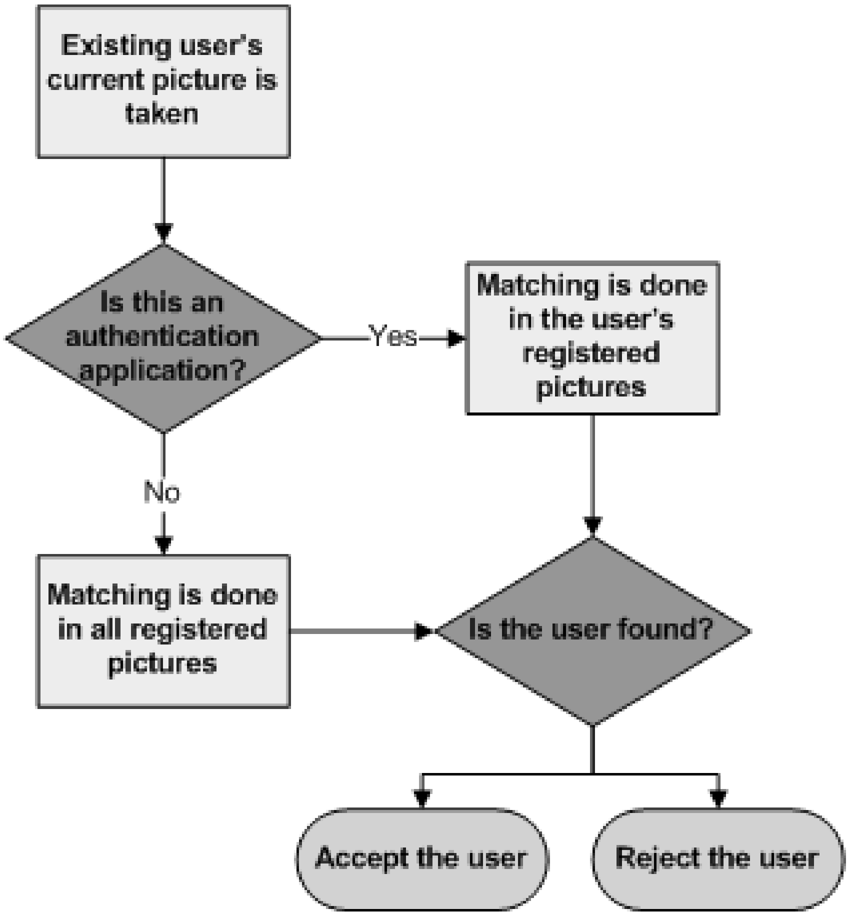

Existing user's current picture is taken
Is this an authentication application?
Yes
Matching is done in the user's registered pictures
No
Matching is done in all registered pictures
Is the user found?
Accept the user
Reject the user

Figure 4.2. Existing user workflow

4.2.3. Matching Scenario for Unimodal and Multimodal Applications

Although the general scenario - i.e New vs Existing User - is the same for both unimodal and multimodal applications, there are some differences in the flow when the recognition is taken into consideration. In unimodal applications, the user is accepted or rejected due to the threshold value. If the user’s current picture’s distance between the nearest picture is less than the threshold value, the user is accepted. Otherwise, the user is rejected and the flow ends. In multimodal applications, the user enters two modalities sequentially if he/she is rejected by the first modality -i.e. face. If the user

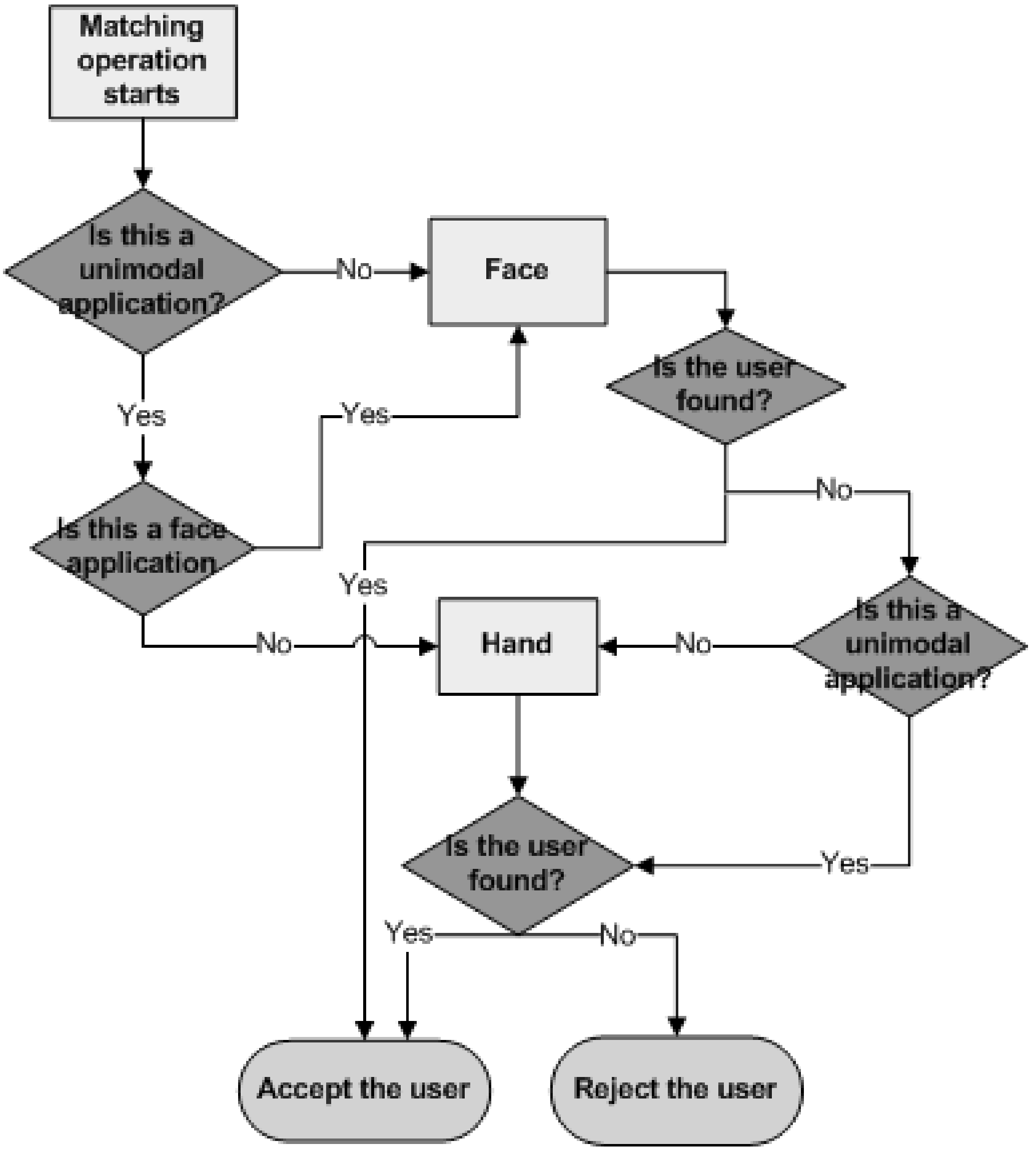

Matching operation starts
Is this a unimodal application?
No
Face
Is the user found?
Yes
Yes
Is this a face application
No
Yes
No
Hand
No
Is this a unimodal application?
Is the user found?
Yes
Yes
No
Accept the user
Reject the user

Figure 4.3. Unimodal and multimodal matching scenario

is accepted after the face verification, the flow ends. However, if the user is rejected by the face modality, the user enters the hand module. If the user is also rejected by the hand modality, user cannot use the system. Otherwise, user is accepted to the system.

4.3. User Interface

4.3.1. Camera Setup Menu

All of the applications include a “Camera Settings” icon in order to change the preferred camera. Although FMJ package enables the user to change the camera for once, trying to return to the previous camera was causing errors. Therefore, the FMJ package was modified to make camera switching more robust.

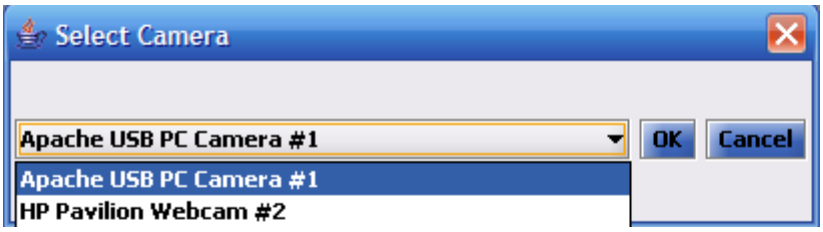


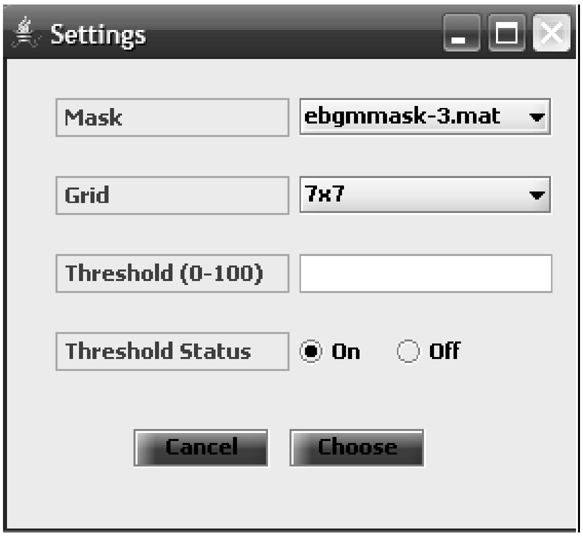
Figure 4.4. Choosing the camera

All of the applications are initialized with the first camera of the detected cameras and this icon enables the user to change the preferred camera. Clicking the Camera Settings icon opens a new dialog that lists the detected cameras of the computer and the user selects the preferred camera from the list and clicks OK.

4.3.2. Settings Menu

“Settings” is an administrative menu option that adjusts the system’s recognition parameters. When the Settings icon is clicked, a new dialog is opened. Hand-based systems’ Settings dialog includes the options of enabling/disabling threshold and adjusting the values of threshold.

Face-based systems’ Settings dialog includes the options of enabling/disabling threshold, adjusting the values of the threshold, changing the mask and changing the



Mask: ebgmmask-3.mat

Grid: 7x7

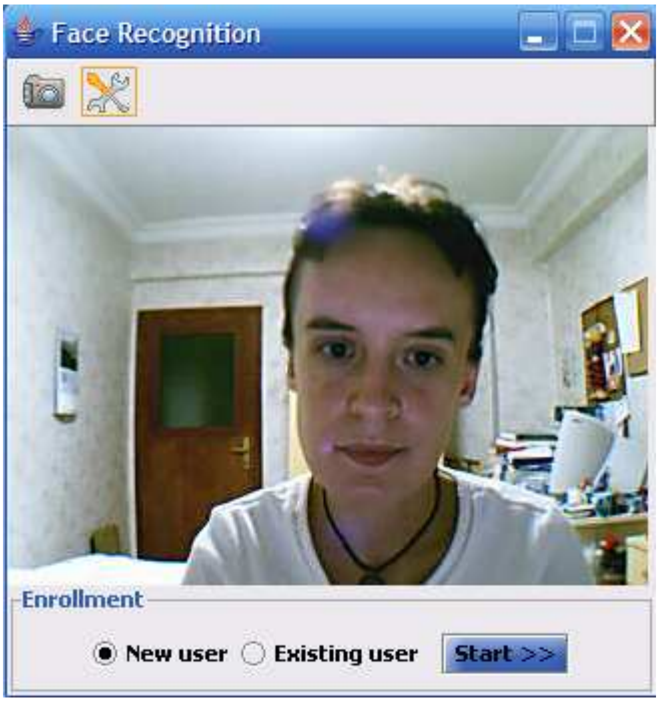
Threshold (0-100):

Threshold Status: ● On ○ Off

Figure 4.5. Adjusting mask, grid and threshold

size of the grid. Since changing the mask and grid size requires all of the registered pictures' feature vectors to be recalculated, a dialog is open to warn the user on this heavy-duty task. It is advised that the updates are done offline by the administrator when time and recognition task's performance are taken into consideration.

4.3.3. Enrollment Options Screen



Enrollment

● New user ○ Existing user

Figure 4.6. Initial screen of the application

The initial screen of all of the applications is Enrollment Options screen. On this screen, the user can adjust the camera and change the settings of the system. Also, the user chooses whether he/she is a new user or an already registered user. After choosing the correct option, enrollment or authentication/identification processes start depending on the scenario.

4.3.4. Enrollment Screen

If the New User option is chosen by the user, enrollment process starts. On this screen, the user enters name, surname, email and department fields in order for the system to register the new user.

Name, surname and department fields are compulsory, so when the user leaves any of these fields blank, an error dialog is displayed and the user is prompted to enter those fields. Also, duplicate name surname combination is not accepted. When the user enters an already registered user's name surname combination, an error message is displayed by the system to prevent illegal updates.

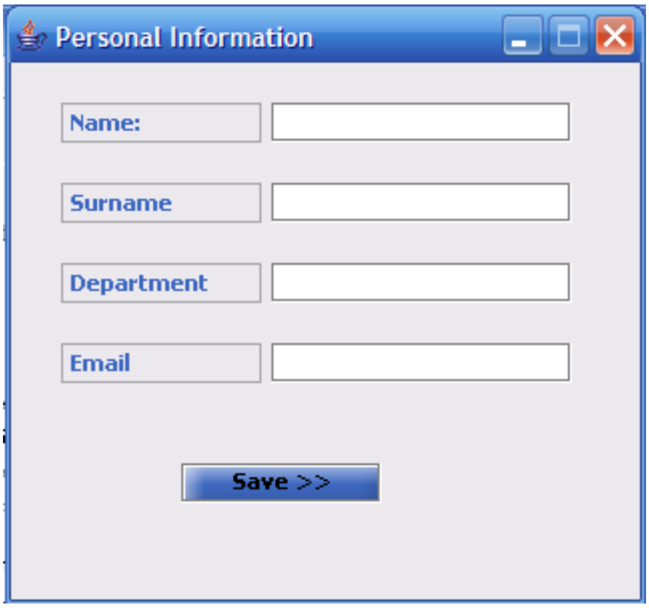


Figure 4.7. Enrollment screen of the application

4.3.5. Picture Snapshot Screens

Face and hand pictures of the user are taken regarding the type of the application. If picture snapshot is on Face mode, the user is expected to put his/her face in an ellipse on the screen so that the face's feature vector is calculated on the cropped face region. Each screen saves two pictures of the user so that recognition process has few pictures of each user to train the recognition process.

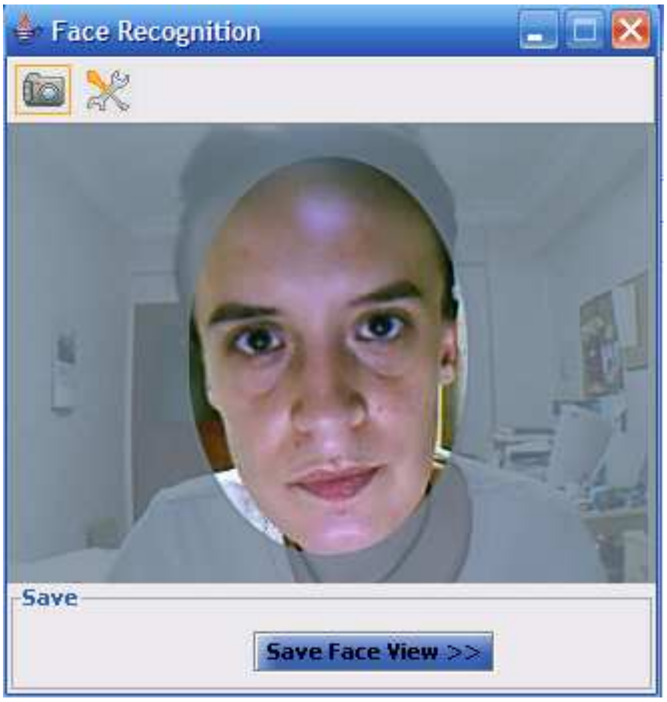


Figure 4.8. The user saves the current snapshot

If picture snapshot is on Hand mode, the user is expected to put his/her hand in a rectangle on the screen and putting his/ her hand in a way parallel to the screen and in a way fingers do not touch each other. If the user does not put his/her hand in the correct position, a warning message is displayed and the user is expected to adjust the position of his/her hand. This is repeated twice: Each screen saves two pictures of the user so that recognition process has four pictures of each user to train the recognition process. After the user saves his/her pictures, he/she completes the process of registration. Heavy-duty tasks such as database operations and training new pictures are done in the completion part of the process.

4.3.6. Welcome Screen

After the user saves his/her pictures, he/she completes the process of registration. By this way, each user registers his/her four pictures and saves name, surname, email and department information on the database. At the end of registration process, a Welcome screen appears where the first snapshot and name surname information of the user are displayed. The user completes the registration process by clicking the Complete button.

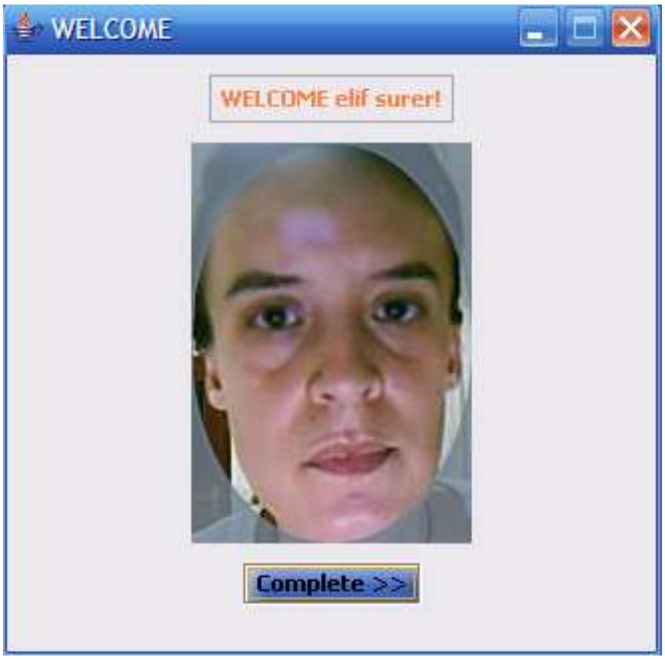


Figure 4.9. The user completes the registration process

4.3.7. Recognition Screen

If the system is in Face mode, the user saves his/her current picture in order the current picture to be recognized/authentized by the verification system. The user places his/her face into the ellipse in order for the current face's feature vector to be calculated. Decision is done by comparing the L2 distances of the current picture's feature vector and the registered pictures's feature vectors. Then, the minimum distance found is compared with the threshold value.

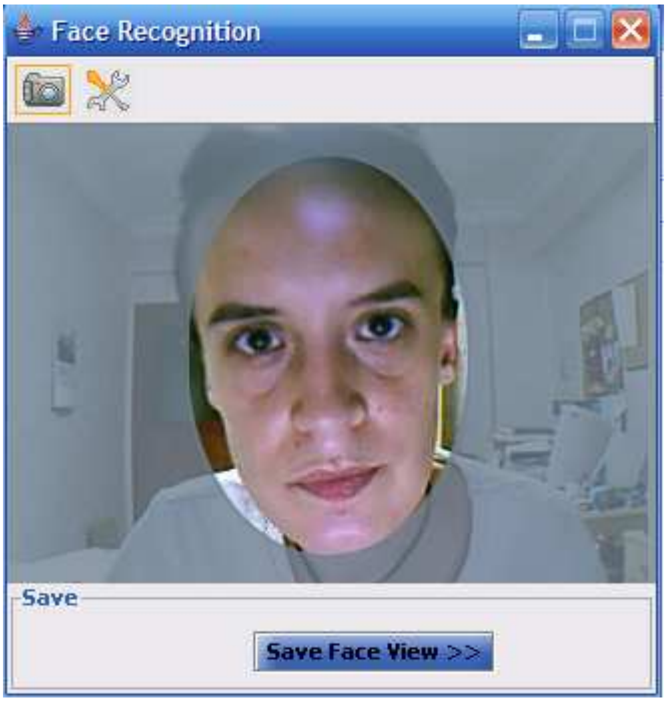


Figure 4.10. The application recognizes the current snapshot

If thresholding option is clicked on the Settings Menu, the distance is checked to see whether it is less than the thresholded value of the minimum distance picture or not. If the minimum distance is not less than the distance calculated by threshold value, "the user is not found" message is displayed and the user is rejected. If thresholding option is not chosen, the minimum-distanced picture's user is the found user and the user is accepted by the system. The reject/acceptance ratio of the system can be set by changing the Threshold value on the Settings menu.

The idea is the same for the hand recognition and authentication as well. The only difference between them is the feature vector calculation, i.e the algorithm and preprocessing operations.

For the multimodal systems, firstly face recognition/authentication is tried. If the user is not found with face modality, hand recognition/authentication screens are displayed. The screens displayed are the same with the screens that are used in unimodal system's picture screenshot screens.

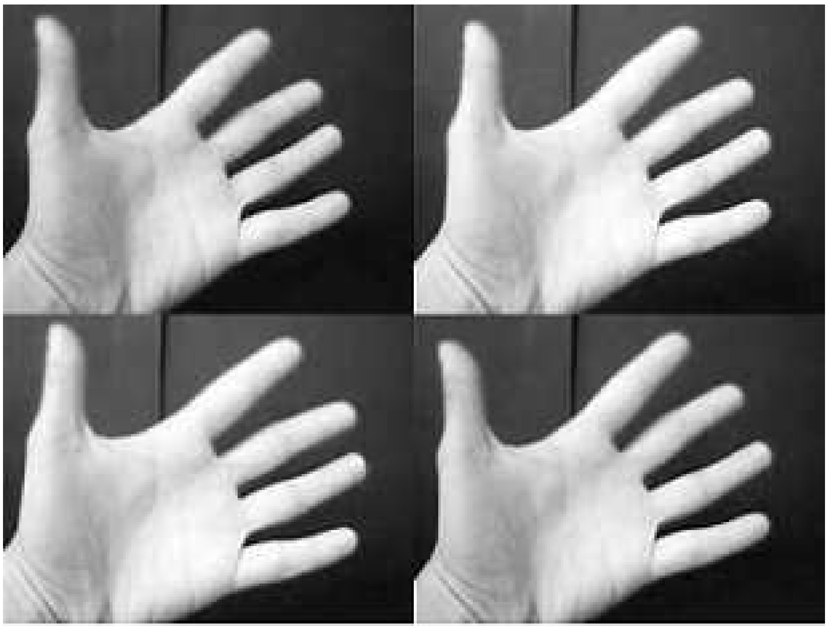


Figure 4.11. Four hand image snapshots acquired from the same person

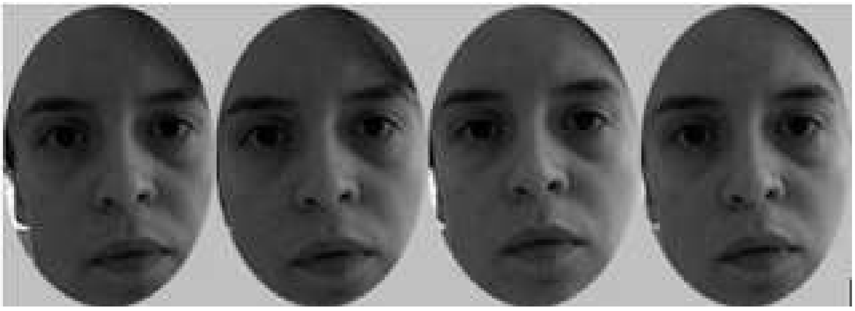


Figure 4.12. Four face image snapshots acquired from the same person

5. RESULTS

5.1. Performance Evaluation

In order to test the recognition and verification performances of the application, several tests have been undertaken with different datasets. Results have been analyzed for unimodal and multimodal cases. For multimodal cases, test results due to different fusion methods have been compared.

5.1.1. Face-based Biometry

Face tests are grouped as identification and verification tests where three different datasets have been used. Dataset-1 is composed of the pictures of 40 people which are taken with the Face Recognition module of the application. Five pictures have been taken from each person and four of them have been used in training set and the fifth image has been taken as the test set.

Dataset-2 is a subset of Dataset-1 where two pictures are used as training set and one picture is used as the test set. Dataset is formed by the data acquired from 40 people.

Dataset-3 is a subset of Face Recognition Technology (FERET) database where pictures of 146 people have been used. Two pictures of each person are included in the training set and one picture is used as the test set.

5.1.1.1. Identification Tests. Before finding the best grid and mask type, two images of each person have been used as the training set and the remaining two images have been used as the validation set.

The sets have been repeated six times with different combinations and their averages have been taken. The optimum parameters have been recorded. 15 × 15

Table 5.1. Average of false matches on 80 pictures

	Mask Width									
Grid Size	**3**	**5**	**7**	**9**	**11**	**13**	**15**	**21**	**29**	**33**
7×7	15.3	14.7	15.7	15.7	16.3	25.7	23.7	20.5	17.5	16.5
9×9	18	17.7	18.8	18	19.3	19.3	18.8	15	13.8	13.8
15×15	14	12.8	12.7	12.2	12	17.7	17.5	13.5	13.7	13.3

grid size yielded the best results. All of the identification tests have used this optimum grid size. In Table 5.1, average of false matches is given.

In Dataset-1, two out of 40 people have been confused. The recognition performance of the system is 95 per cent. In Table 5.2, a subset of the confusion matrix is shown.

Table 5.2. Subset of face confusion matrix (Dataset-1)

	Person 12	**Person 18**	**Person 31**	**Person 27**	**Person 38**
Person 12	**1**	0	0	0	0
Person 18	0	**0**	0	1	0
Person 31	1	0	**0**	0	0
Person 27	0	0	0	**1**	0
Person 38	0	0	0	0	**1**

In Dataset-2, three out of 40 people have been confused. The recognition performance of the system is 92.5 per cent. The performance of Dataset-1 is better since the training set is composed of four images of each person instead of two pictures. Table 5.3 shows a subset of the confusion matrix.

In Dataset-3, 10 out of 146 people have been confused. The recognition performance is 93.15 per cent. In Table 5.4, false results are shown.

Table 5.3. Subset of face confusion matrix (Dataset-2)

	Person 10	**Person 19**	**Person 27**	**Person 38**	**Person 5**
Person 10	**0**	0	0	1	0
Person 19	0	**0**	0	0	1
Person 27	1	0	**0**	0	0
Person 38	0	0	0	**1**	0
Person 5	0	0	0	0	**1**

Table 5.4. False results (Dataset-3)

False Results (1-5)		**False Results (6-10)**	
Queried Person	**Result**	**Queried Person**	**Result**
Person 1	Person 49	Person 82	Person 74
Person 20	Person 137	Person 84	Person 86
Person 36	Person 24	Person 124	Person 136
Person 40	Person 38	Person 127	Person 123
Person 44	Person 38	Person 144	Person 140

5.1.1.2. Verification Tests. In order to have a reject option, a threshold value is needed in the verification system. This threshold value is used to reject the people whose distance to the current picture's feature vector is greater. Tuning the threshold is an important issue since if it is too low, then even authorized people can be rejected by the system which will increase the False Rejection Rate (FRR) of the system. If the threshold value is too high, then most of the people in the database will be accepted by the system, so the False Acceptance Rate (FAR) of the system will increase. Equal Error Rate (EER) is given by the intersection of FAR and FRR curves. Optimum threshold value is accepted as the value that gives the EER. Verification tests have been analyzed with different thresholds and FAR and FRR values have been found and plotted for three datasets.

Table 5.5. Verification results (Dataset-1)

	Threshold									
Rates	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
FAR	0	0	7.1	35.1	66.3	87.7	96.8	99.9	100	100
FRR	100	81.3	21.3	2.5	2.5	0	0	0	0	0

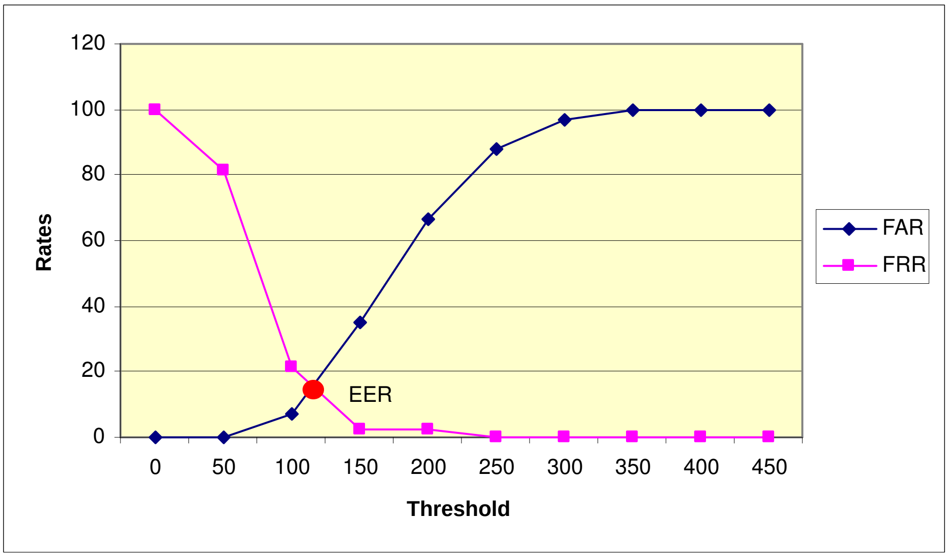


Figure 5.1. Effect of threshold on rates (Dataset-1)

Table 5.6. Verification results (Dataset-2)

	Threshold									
Rates	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
FAR	0	0	7.1	34.8	66.4	86.6	96.3	99.8	100	100
FRR	100	75	20	2.5	2.5	0	0	0	0	0

Table 5.7. Verification results (Dataset-3)

	Threshold									
Rates	**0**	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
FAR	0	0	0	0	1.4	18.5	65	95.4	100	100
FRR	100	100	92.1	66.4	21.9	0.7	0	0	0	0

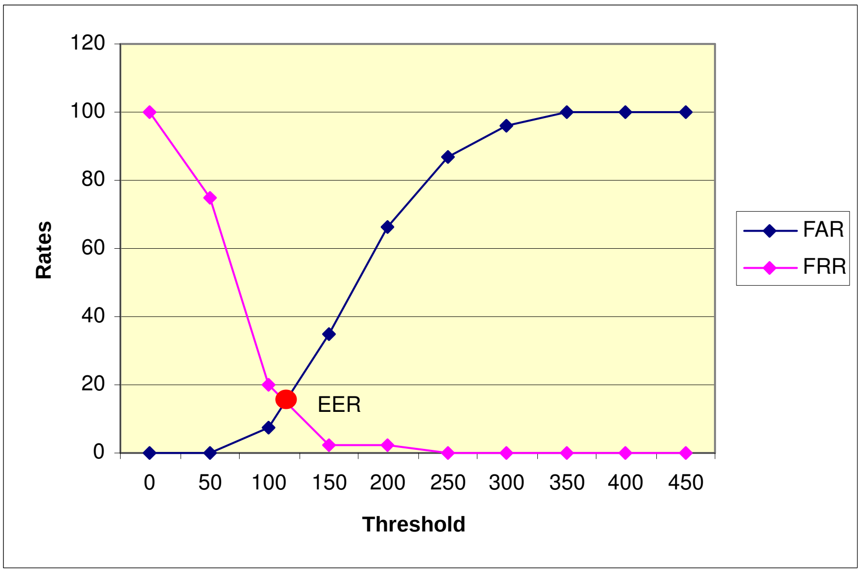


Figure 5.2. Effect of threshold on rates (Dataset-2)

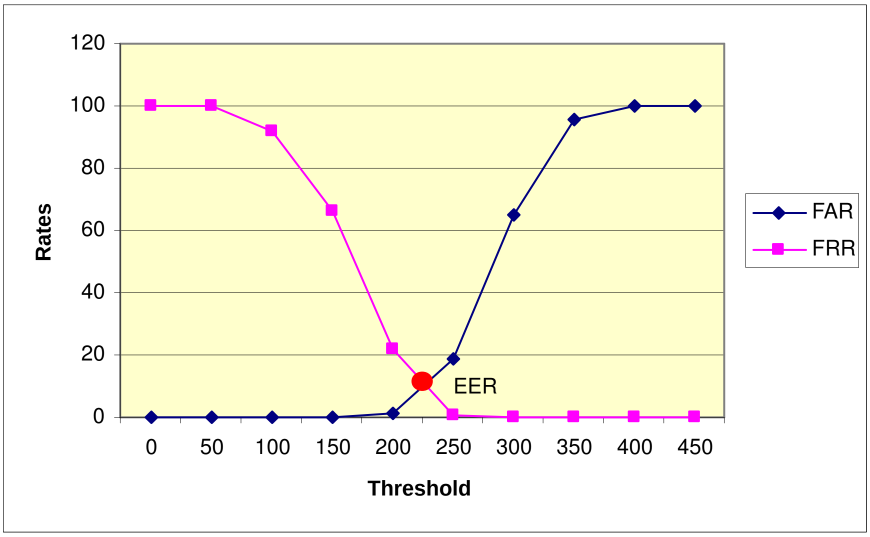


Figure 5.3. Effect of threshold on rates (Dataset-3)

5.1.2. Hand-based Biometry

Hand tests are grouped as identification and verification tests where three different datasets have been used. Dataset-1 is composed of the pictures of 40 people which are taken with the Hand Recognition module of the application. Five pictures have been taken from each person and four of them have been used in training set and the fifth image has been taken as the test set.

Dataset-2 is a subset of DataSet-1 where two pictures are used as training set and one picture is used as the test set. Dataset is formed by the data acquired from 40 people.

Dataset-3 is a dataset where pictures of 146 people which have been acquired during Yöruk et al.'s work [33]. Two pictures of each person are included in the training set and one picture is used as the test set.

5.1.2.1. Identification Tests. In Dataset-1, four mismatches have occurred out of 40 people in Rank-1, so the success of the system is 90 per cent.

In Table 5.8, the subset of hand confusion matrix is given where Per. is the abbrevation of Person.

Table 5.8. Subset of hand confusion matrix (Dataset-1)

	Per.1	**Per.6**	**Per.7**	**Per.8**	**Per.32**	**Per.33**	**Per.40**
Person 1	**1**	0	0	0	0	0	0
Person 6	0	**1**	0	0	0	0	0
Person 7	0	1	**0**	0	0	0	0
Person 8	0	0	1	**0**	0	0	0
Person 32	0	0	0	0	**1**	0	0
Person 33	1	0	0	0	0	**0**	0
Person 40	0	0	0	0	1	0	**0**

In Dataset-2, 7 mismatches out of 40 people have occurred and the recognition performance of the system is 82.5 per cent. The performance is worse than the Dataset-1 since Dataset-2 includes two training images of each person instead of four.

Table 5.9. List of false identification results (Dataset-2)

False Results (1-4)		**False Results (5-7)**	
Queried Person	**Result**	**Queried Person**	**Result**
Person 6	Person 35	Person 24	Person 11
Person 7	Person 13	Person 33	Person 22
Person 12	Person 5	Person 18	Person 6
Person 23	Person 24		

In Dataset-3, recognition result of one out of 146 people is wrong. Recognition performance is 99.32 per cent.

Table 5.10. List of false identification results (Dataset-3)

False Result	
Queried Person	**Result**
Person 24	Person 75

Summary of unimodal identification results are given in Figures 5.4-5.6. Face and hand modalities of the system have been tested with three different datasets.

5.1.2.2. Verification Tests. As in the face modality, a threshold value is needed in order to have a reject option. FAR and FRR values of Dataset-1, Dataset-2 and Dataset-3 have been calculated and optimum thresholds - the ones giving EER - have been found. Verification values and threshold graphics have been plotted for each dataset.

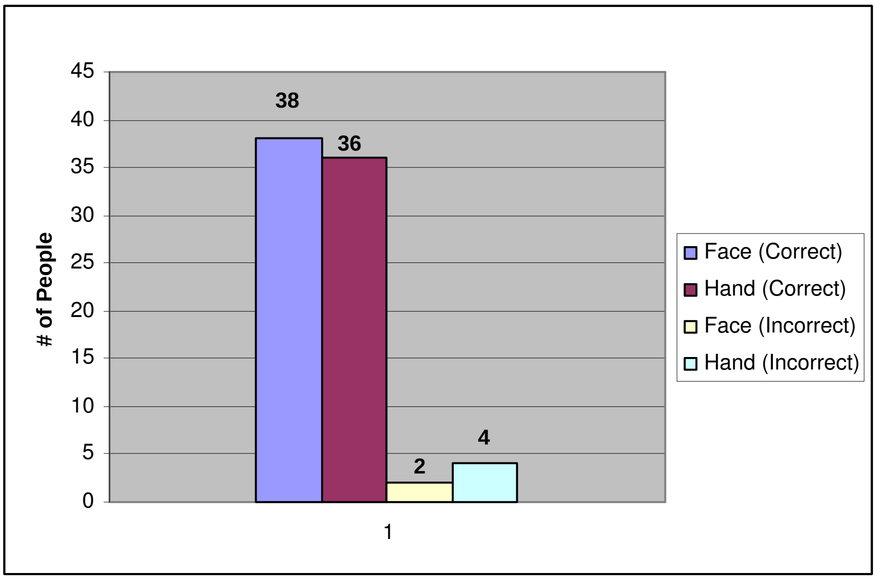


Figure 5.4. Summary of unimodal results (Dataset-1)

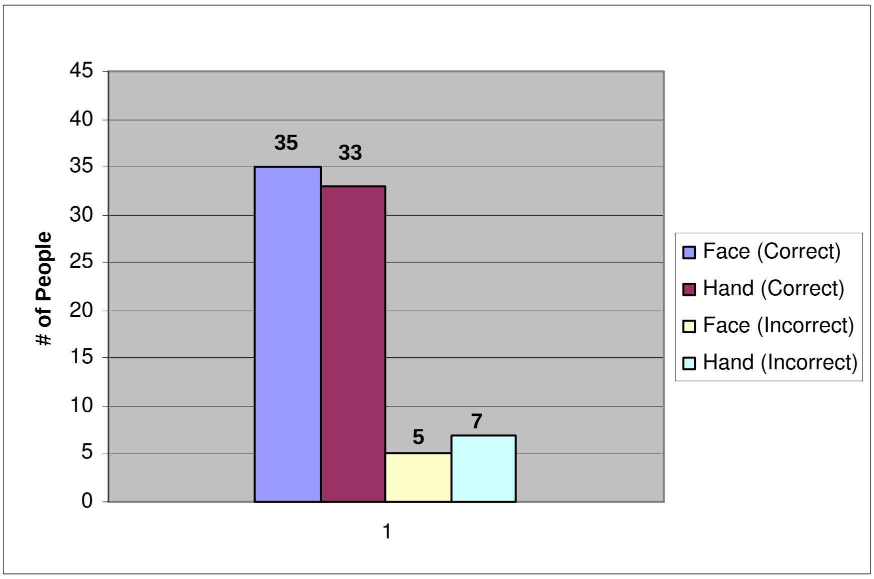


Figure 5.5. Summary of unimodal results (Dataset-2)

Table 5.11. Hand verification results (Dataset-1)

	Threshold						
Rates	**10**	**12**	**14**	**16**	**18**	**20**	**22**
FAR	0	0	0.3	55.1	95.9	98.8	100
FRR	100	94.4	68.8	15	2.5	0	0

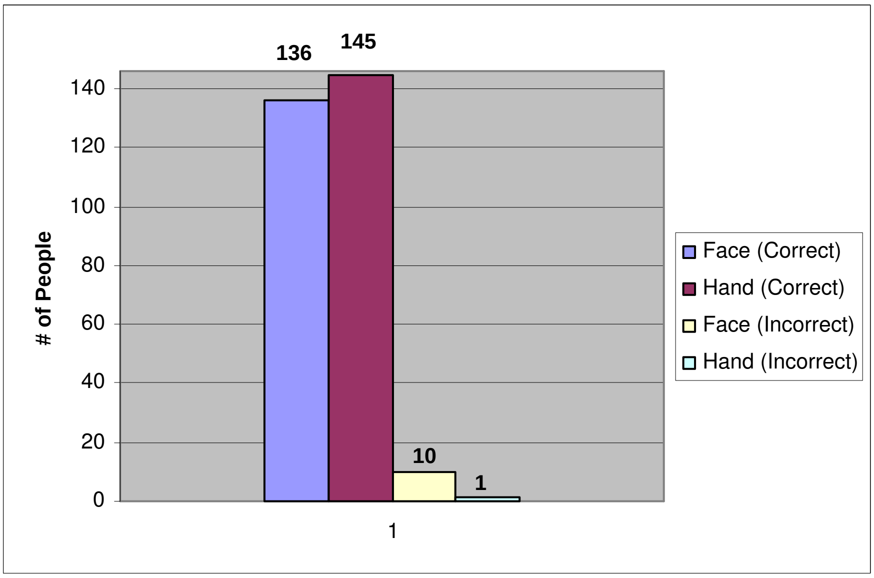


Figure 5.6. Summary of unimodal results (Dataset-3)

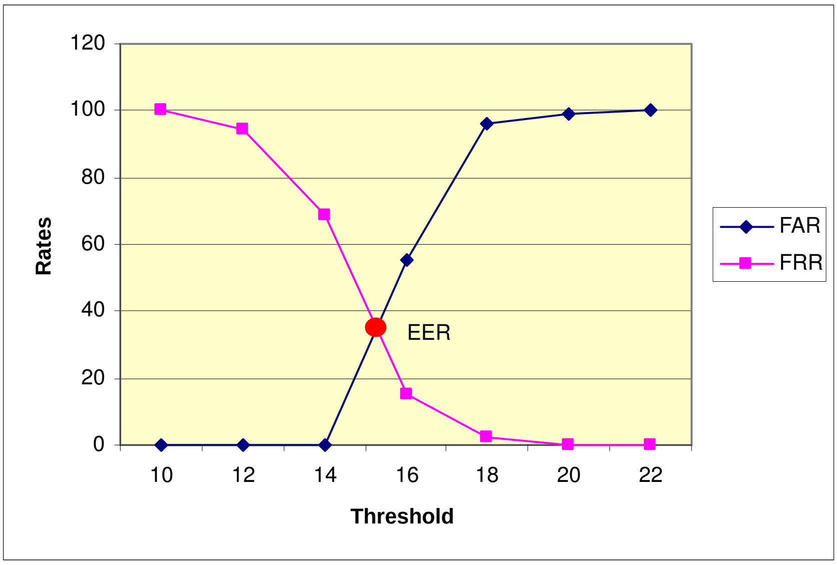


Figure 5.7. Verification results (Dataset-1)

Table 5.12. Hand verification results (Dataset-2)

	Threshold										
Rates	**7**	**7.75**	**8.5**	**9.25**	**10**	**10.8**	**11.5**	**12.3**	**13**	**14**	**14.5**
FAR	0	0	0	0.2	2.7	19.6	70.4	90.5	98.4	100	100
FRR	100	96.3	93.8	82.5	53.8	26.2	6.3	1.3	0	0	0

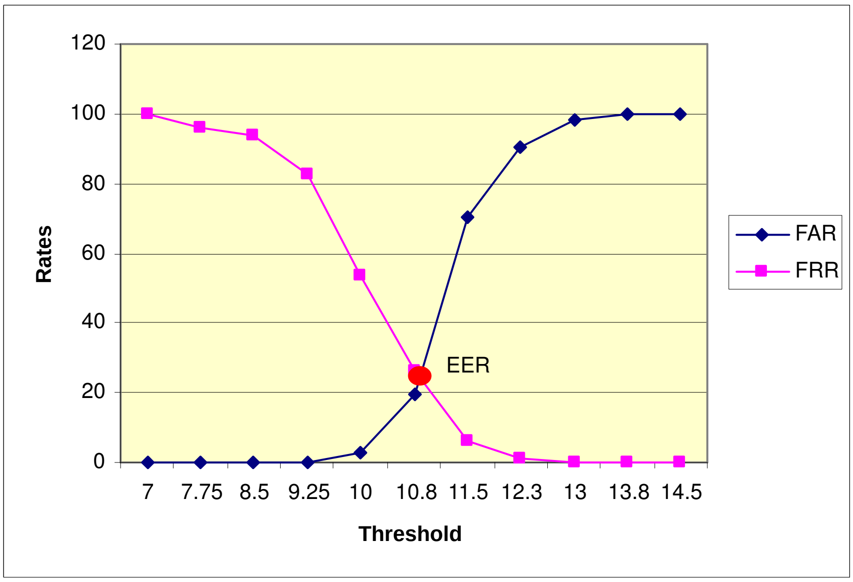


Figure 5.8. Verification results (Dataset-2)

Table 5.13. Hand verification results (Dataset-3)

	Threshold						
Rates	**0**	**4**	**8**	**12**	**16**	**20**	**24**
FAR	0	0	0	0	0.1	75.9	100
FRR	100	94.2	51	26.7	8.9	0	0

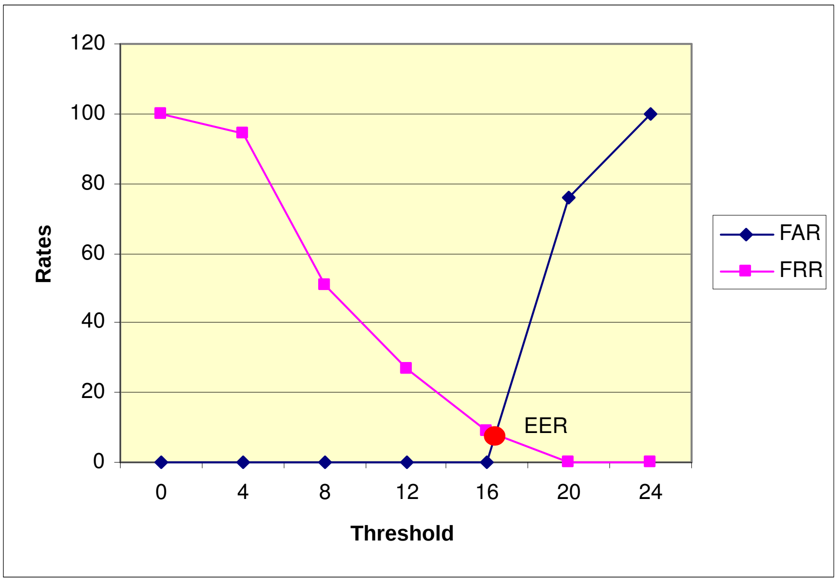


Figure 5.9. Verification results (Dataset-3)

5.1.3. Multimodal Fusion Methods

Different classifiers can be combined in order to increase the accuracy of the classification. Individual classifiers can be combined or fused in the following levels: decision level, score level and feature level. In this thesis, decision level has been the level of fusion. The most commonly used decision-level fusion methods are:

5.1.3.1. Plurality Voting. Plural voting is a commonly used method where the class label that has the greatest vote is the output. The formal expression of the plurality voting is as follows:

$$\arg \max_{j=1}^{c} \sum_{i=1}^{L} t_{i,j}, \tag{5.1}$$

where L is the number of classifiers, and c is the number of classes. $t_{i,j}$ is a measure of vote and $t_{i,j} = 1$ if classifier c_i thinks that the unknown pattern belongs to class ω_j, and $t_{ij} = 0$ otherwise.

5.1.3.2. Borda Count. This method is used when the individual pattern classifiers produce lists of ranks. It is possible to decrease the number of fusion classes in this method.

Also, it is possible to choose the top-k ranked classes in fusion. The classifier is chosen in the following way:

$$\arg \min_{j=1}^{c} \sum_{i=1}^{L} d_{i,j}, \tag{5.2}$$

where $d_{i,j}$ represents distance of classifier and c is the number of classes and L is the number of classifiers.

5.1.3.3. Fixed Arithmetic Combination Rules. Arithmetical rules such as sum, min, max and median can be used as fusion rules if classifiers produce similarity scores. In order for these rules to be applied, score values should be normalized beforehand. The formulas of fixed arithmetic combination are:

1. Sum rule:

$$\arg\max_{j=1}^{c} \sum_{i=1}^{L} d_{i,j} \tag{5.3}$$

2. Product rule:

$$\arg\max_{j=1}^{c} \prod_{i=1}^{L} d_{i,j} \tag{5.4}$$

3. Max Rule:

$$\arg\max_{j=1}^{c} (\max_{i=1}^{L}[d_{i,j}]) \tag{5.5}$$

where $d_{i,j}$ represents distance of classifier and c is the number of classes and L is the number of classifiers.

5.1.3.4. Confidence-added Fusion Rules. In this method, estimated confidences of individual classifiers are used. The confidences can be estimated via the similarity scores of the classifiers. To find the confidence levels of the classifiers, the relative distance between two nearest neighbors of the classifier is taken into consideration.

$$d_i{}' = \frac{(d_1 - d_2)}{Med(d_1, d_2, ...d_c) - d_1}, i = 2...c \tag{5.6}$$

while $Med(d_1, d_2, ...d_c) - d_1$ is the median estimate [43].

5.1.4. Multimodal Biometrics

After face and hand modalities of the system have been developed, several fusion techniques have been used in order to analyze the performance of the multimodal case.

Table 5.14. Multimodal confusion matrix (Dataset-1 and Dataset-2)

	Face Result	**Hand Result**
Person 12	Person 31	Person 7

Table 5.15. Multimodal confusion matrix (Dataset-3)

	Face Result	**Hand Result**
Person 12	Person 31	Person 7

The most commonly used fusion methods and the results of the fusion techniques that have been used in this thesis are summarized.

5.1.4.1. Face-Hand Sequence Results. For Dataset-1 and Dataset-2, when the face and hand modalities of the system have been fused in the order of face and hand, one person out of 40 people has been confused. The confusion result is the same for both datasets. The recognition performance of the system has increased.

Dataset-3 is composed with a Chimeric approach. Although face images are from the FERET database and hand images are from the dataset acquired during Yöruk et al.'s study [33], two arbitrary images which happen to have the same identification code has been accepted as being acquired from the same person. For example, the first face image of the FERET database and the first hand image of hand dataset is accepted as being acquired from the same person. This approach is used in order to enlarge the database tested since Dataset-1 and Dataset-2 which have been acquired with the user interface of this thesis include 40 people.

For Dataset-3, one out of 146 people have been confused. Therefore, the recognition performance has increased.

5.1.4.2. Borda Count Results. Fusion results with Borda Count has been calculated via giving weights to the three ranks. Rank-1 has been multiplied with three, Rank-2 has been multiplied with two and Rank-3 has been multiplied with one. In Table 5.16, the calculation example can be seen.

Table 5.16. Borda Count calculation example

Queried Person	**Ranks**	**Face**	**Hand**	**Borda Result**
Person 2	rank-1 (3)	Person 2	Person 2	Person 2 (7)
	rank-2 (2)	Person 1	Person 12	
	rank-3 (1)	Person 1	Person 2	

In Dataset-1 and Dataset-2, none of 40 people have been confused, so the performance is 100 per cent. In Dataset-3, one out of 146 people has been confused, so the performance is 99.32 per cent. In Table 5.17, false results of Borda Count are shown.

Table 5.17. False results of fusion with Borda Count (Dataset-3)

Wrong Result		
Person	**Score**	**Result**
Person 24	3	Person 75

5.1.4.3. Fixed Arithmetic Combination Results. Fusion results with Fixed Arithmetic Combination Sum Rule has been calculated by using the normalized distances. If a person is not found in first three ranks of a modality, the distance is calculated as 1. In Table 5.18, the calculation example can be seen.

In Dataset-1, two out of 40 people have been confused, so the performance is 90 per cent. In Table 5.19, the calculation results have been given.

In Dataset-2, two out of 40 people have been confused, so the performance is 90 per cent. In Table 5.20, the calculation results have been given.

Table 5.18. Sum Rule calculation example

Person	**Face**	**Value**	**Hand**	**Value**	**Min. Value**	**Result**
Person 2	Person 2	0.15	Person 2	0.1	0.53	Person 2
	Person 1	0.65	Person 12	0.22		
	Person 1	0.7	Person 2	0.28		

Table 5.19. False results of fusion with Score Level (Dataset-1)

Wrong Results (1-2)		
Person	**Distance**	**Result**
Person 7	1.3	Person 6
Person 27	1.13	Person 18

In Dataset-3, seven out of 40 people have been confused, so the performance is 95.21 per cent. In Table 5.21, the calculation results have been given.

5.1.4.4. Confidence Level Results. Fusion results with Confidence Level has been calculated by using the median distance of first six ranks and Rank-1 ve and Rank-2 values. The result having the higher confidence level has been chosen as the correct result. In Table 5.22, the example is given.

In Dataset-1, two out of 40 people have been confused, so the performance is 90 per cent. In Dataset-2, five out of 40 people have been confused, so the performance is 87.5 per cent. In Dataset-3, 5 out of 146 people have been confused, so the performance is 96.6 per cent.

5.1.5. Performance Under Adverse Conditions

In order to test the effect of changes in appearance of the people, additional data have been acquired from 10 people after 2.5 months have passed from the initial session. New data have been acquired with the following properties:

Table 5.20. False results of fusion with Score Level (Dataset-2)

Wrong Results (1-2)		
Person	**Distance**	**Result**
Person 10	1.12	Person 38
Person 19	1.12	Person 5

Table 5.21. False results of fusion with Score Level (Dataset-3)

Wrong Results (1 to 4)			**Wrong Results (5 to 7)**		
Person	**Distance**	**Result**	**Person**	**Distance**	**Result**
Person 20	1.48	Person 137	Person 56	1.65	Person 52
Person 36	1.58	Person 24	Person 84	1.32	Person 86
Person 40	1.58	Person 38	Person 127	1.33	Person 123
Person 44	1.46	Person 38			

1. Normal pose
2. Too near
3. Too far
4. With different lighting condition
5. With expression - smile
6. With glasses

5.1.5.1. Identification Tests. The results of the identification tests have been 68.33 per cent. The performance degradation stemmed mostly from the effect of glasses and changes in appearance (beard, hair). Also, the poses that are too far has caused some trouble since the faces have been very small and background information caused some artifacts.

5.1.5.2. Verification Tests. Verification tests have been undertaken in order to find the optimum threshold value.

Table 5.22. Confidence Level calculation example

Queried Person	**Face**	**Confidence**	**Hand**	**Confidence**	**Result**
Person 1	Person 1	93%	Person 1	43%	Person 1

Table 5.23. False results of fusion with Confidence Level (Dataset-1)

Wrong Results		
Person	**Confidence (%)**	**Result**
Person 12	57	Person 7
Person 34	88	Person 2

5.2. Usability Tests

Evaluation is the task of assessing the functionality and usability of a system. It may be done in laboratory or field and with or without the help of the users. Evaluation can be done for design phase and for the implementation. In this thesis, evaluation was performed for the implementation.

Evaluating the systems aim to assess the functionality of the system, the success of the user interface and to determine the challenges. Evaluating the implementations can be grouped into three types of methods. These are:

1. empirical methods
2. observational methods
3. query techniques

5.2.1. Empirical Methods

Empirical methods (experimental evaluation) use controlled experiments that focus on analyzing the specific features of the user behavior. The evaluator defines a hypothesis to test by means of measuring the interactive behavior. Several experimental

Table 5.24. False results of fusion with Confidence Level (Dataset-2)

Wrong Results		
Person	**Confidence (%)**	**Result**
Person 7	31	Person 13
Person 12	50	Person 5
Person 19	67	Person 5
Person 24	79	Person 11
Person 33	71	Person 21

Table 5.25. False results of fusion with Confidence Level (Dataset-3)

Wrong Results		
Person	**Confidence (%)**	**Result**
Person 1	75	Person 49
Person 36	80	Person 24
Person 44	31	Person 38
Person 84	20	Person 86
Person 124	50	Person 136

conditions are designed that have certain controlled variables. The changes in the behavior are grouped into different conditions.

The success of this method depends on the chosen subjects, tested variables and designed hypotheses. When choosing the subjects, expected user population should be taken into consideration and sample size should be large enough to represent the actual users of the system. Variables of the systems can be divided into two: dependent and independent variables.

Examples of independent variables are interface style, level of help and number of menu items. The dependent variables can be the time taken to complete a task, the number of errors made by the user and user performance.

Table 5.26. True (T) and False (F) results of recognition

	Normal	**Near**	**Far**	**Light**	**Smile**	**Glasses**
Person 1	T	T	T	T	T	F
Person 2	T	T	T	T	T	F
Person 3	T	T	T	T	T	T
Person 4	T	T	T	T	T	T
Person 5	T	T	F	T	F	T
Person 6	T	T	F	T	T	F
Person 7	F	T	F	F	F	F
Person 8	F	F	F	F	F	F
Person 9	T	T	F	T	T	F
Person 10	T	T	T	T	T	T

Table 5.27. Verification results of adverse conditions

	Threshold									
	0	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**
FAR	0	8	10	31.2	64.3	84.2	98	100	100	100
FRR	100	75	25	2	2	2	0	0	0	0

Hypotheses, which are the predictions of the result of an experiment - are formed by taking dependent and independent variables into consideration. The idea is as follows: Hypotheses aim to prove that any change in an independent variable will cause a change in a dependent variable. In other words, they try to disprove null hypothesis -i.e. the claim that states there is no difference in the dependent variable in terms of changes in independent variable.

5.2.2. Observational Techniques

Observational techniques include thinking aloud method, whose recording can be done in several protocols and analyzed in a walkthrough process.

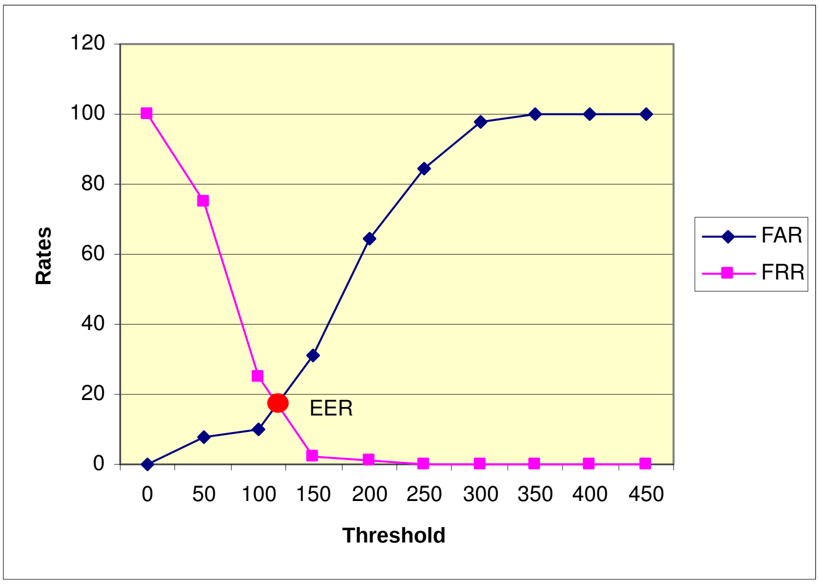


Figure 5.10. Verification results of adverse conditions

Thinking aloud techniques focus on observing the subject when the subject actually uses the system. The evaluator observes the user's behaviors and records them. Evaluator's notes, audio recording, video recording, computer logging and user's notes are the widely used tools in this technique.

After acquiring user's behavioral data, the walkthrough phase starts. In this phase, the weak parts of the implementation are detected and a refinement procedure starts.

5.2.3. Query Techniques

Query techniques are interviews and questionnaires.

5.2.3.1. Interviews. Interviewing the users on their experience with the system is a straightforward approach. An interview usually has a top-down structure starting with general questions and following with more specific ones. In order for an interview to be effective, it should be prepared before the interview and with a set of central questions that form the focus of it [44].

5.2.3.2. Questionnaires. Administrating a questionnaire is a practical way to obtain large amount of information about the system's usability. Producing a questionnaire follows several phases.

These are:

1. Questionnaire follow a design-evaluate-redesign cycle in order to be sure that it seeks the responses that matches the aim of the questionnaire.
2. Questionnaires have clear hypotheses or research questions, i.e specific targets.
3. Questionnaire are analyzed after the evaluation and the implementations are updated according to the required improvements extracted from the responses.

Questionnaires can be interviewer-administered or self-administered. The questions of the questionnaire can be open or closed. Open questions enable the evaluator to gain information on a broad range.

The problem with the open questions is that analyzing diverse information is extremely difficult. Closed questions are a list of questions that propose answer options to the respondents. Although the closed questions can be regarded as limited, this approach enables the users to see the different aspects of the system. While designing the questionnaire, the following recommendations should be taken into consideration:

1. Closed questions are better than the open ones.
2. The flow of the questions should be from the general ones to the specific ones.
3. Two sides of A4 is enough for the questionnaire.
4. Attractive questionnaires are more appealing to the users.
5. Before presenting them to the users, evaluator should try to fill the questionnaire.

The most commonly used questionnaire types are:

1. Multi choice questions and checklists
2. Scalar questionnaire

Multi choice questions are the questions that include as Yes/No/Don't know type of answer options. Checklist enable the users to check from a list of answer choices. Scalar questionnaires are based on a pre-defined scale [45].

5.2.4. Results

After the data have been acquired from 40 people, a short questionnaire to evaluate the ease of use, speed and user interface of the system has been provided to them. Besides, their suggestions on the system and how to improve it have been gathered. The user interface and the data gathering environment have been changed due to suggestions of the user.

The changes have been made on the number setup of the hand image acquisition and taking two images per screen.

In the questionnaire 5 represents Very Good, 4 represents Good, 3 represents Neutral, 2 represents Quite Bad and 1 represents Very Bad in terms of Ease of Use, Speed and User Interface. The results of the questionnaire were as follows:

Table 5.28. Average of the responses to the questionnaire

	Face Results	**Face (%)**	**Hand Results**	**Hand (%)**
Ease of Use	4.075	81.5	3.55	71
Speed	4.125	82.5	3.9	78
User Interface	4.2	84	4.15	83

6. CONCLUSIONS

6.1. Conclusion

In this thesis, six applications have been developed that verify and identify face, hand and fusion of two modalities. All of the applications have similar user interfaces so that the user operates on familiar environments. The applications have been designed as wizard applications and the tasks have been divided into several screens. All of the user interfaces have been developed in Java.

After the face recognition algorithm has been implemented and the hand recognition algorithm has been integrated to the system, tests have been run on the system. To test the recognition performances of the algorithms and to evaluate the usability of the interface, face and hand data have been acquired from 40 people. Also, in order to test the system with a larger database, FERET database and already acquired hand data have been combined. This dataset contains data from 146 people.

Besides testing the recognition performance of the algorithms, the user interface and data acquisition environment have been evaluated by the users after the data acquisition session. The problem of collecting the data has been solved by adding an extra screen to the interface and by taking two snapshots of the user at each screen. The problem with hand data gathering environment has been solved by arranging the positions of the cameras, so the users are able to use the system easily on their own.

The main difficulty in developing the user interface has been changing the data acquisition package. Since JMF (Java Media Framework) package does not support two USB cameras, package has been changed to FMJ (Freedom for Media in Java). Also, it has been modified in order to handle requirements of the thesis.

The main outcome of this thesis is several applications working in real-time for unimodal and multimodal cases. Also, testing the system with different datasets and getting feedback from the users has enabled us to fix the possible artifacts and problems that have not been noticed in the implementation phase.

6.2. Future Work

Although the system's face recognition part works fast, hand recognition part works slowly and this causes a performance degradation in the overall system. In order to speed up the hand recognition, Matlab code can be rewritten in C. By compiling the C files as Mex-files, the system can work with the same way, with better performance.

The final tests that include adverse conditions have shown that changes in physical appearance of the people decrease the recognition performance. To overcome this problem, an update module can be added to the system where users can update their already saved pictures so that the dataset becomes up-to-date. This module would increase the recognition performance.

One of the adverse conditions that affect the performance the worst has been face image acquisition with the user too far from the camera. A performance measure that detects this condition and instructs the user to bring his face closer would increase the face recognition performance.

APPENDIX A: QUESTIONNAIRE

Name Surname:

Department:

FACE RECOGNITION

1. You have provided Name, Surname, Department and Email on the information retrieval screen. Did you find those fields sufficient?

   ○ Yes ○ No ○ Don't Know

   If you choose No, what other fields should be retrieved?

2. Rate the face recognition system on the following criteria.

   Ease of Use

   (a) Very Easy
   (b) Quite Easy
   (c) Neutral
   (d) Quite Difficult
   (e) Very Difficult

   Speed

   (a) Very Fast
   (b) Quite Fast
   (c) Neutral
   (d) Quite Slow
   (e) Very Slow

User Interface

(a) Very Friendly
(b) Quite Friendly
(c) Neutral
(d) Quite Unfriendly
(e) Very Unfriendly

3. The main difficulty on the face recognition part is :

○ Too much information on the Information Screen
○ Placing my face in the ellipse
○ Collecting the data 5 times

4. What are your suggestions to improve the user interface of the face recognition system?

APPENDIX B: DETAILED TEST RESULTS

Table B.1. Results of fusion with Borda Count (Dataset-1)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Score**	**Result**	**Person**	**Score**	**Result**
Person 1	12	Person 1	Person 21	6	Person 21
Person 2	7	Person 2	Person 22	11	Person 22
Person 3	12	Person 3	Person 23	10	Person 23
Person 4	8	Person 4	Person 24	6	Person 24
Person 5	11	Person 5	Person 25	11	Person 25
Person 6	9	Person 6	Person 26	8	Person 26
Person 7	3	Person 7	Person 27	5	Person 27
Person 8	11	Person 8	Person 28	8	Person 28
Person 9	8	Person 9	Person 29	10	Person 29
Person 10	11	Person 10	Person 30	11	Person 30
Person 11	8	Person 11	Person 31	10	Person 31
Person 12	4	Person 12	Person 32	11	Person 32
Person 13	11	Person 13	Person 33	5	Person 33
Person 14	12	Person 14	Person 34	8	Person 34
Person 15	9	Person 15	Person 35	10	Person 35
Person 16	9	Person 16	Person 36	8	Person 36
Person 17	10	Person 17	Person 37	11	Person 37
Person 18	9	Person 18	Person 38	11	Person 38
Person 19	8	Person 19	Person 39	8	Person 39
Person 20	9	Person 20	Person 40	7	Person 40

Table B.2. Results of fusion with Score Level (Dataset-1)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Distance**	**Result**	**Person**	**Distance**	**Result**
Person 1	1.5	Person 1	Person 21	0.67	Person 21
Person 2	0.53	Person 2	Person 22	0.87	Person 22
Person 3	0.97	Person 3	Person 23	0.92	Person 23
Person 4	0.36	Person 4	Person 24	0.34	Person 24
Person 5	1.12	Person 5	Person 25	1.07	Person 25
Person 6	0.66	Person 6	Person 26	0.58	Person 26
Person 7	**1.3**	**Person 6**	**Person 27**	**1.13**	**Person 18**
Person 8	1	Person 8	Person 28	0.78	Person 28
Person 9	1.08	Person 9	Person 29	0.92	Person 29
Person 10	0.79	Person 10	Person 30	0.66	Person 30
Person 11	0.53	Person 11	Person 31	0.4	Person 31
Person 12	0.53	Person 12	Person 32	0.66	Person 32
Person 13	1.07	Person 13	Person 33	0.52	Person 33
Person 14	0.98	Person 14	Person 34	0.4	Person 34
Person 15	0.95	Person 15	Person 35	0.99	Person 35
Person 16	0.62	Person 16	Person 36	0.51	Person 36
Person 17	0.94	Person 17	Person 37	1.02	Person 37
Person 18	0.68	Person 18	Person 38	0.93	Person 38
Person 19	0.72	Person 19	Person 39	0.63	Person 39
Person 20	0.56	Person 20	Person 40	0.42	Person 40

Table B.3. Results of fusion with Confidence Level (Dataset-1)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Conf. (%)**	**Result**	**Person**	**Conf. (%)**	**Result**
Person 1	93	Person 1	Person 21	75	Person 21
Person 2	91	Person 2	Person 22	33	Person 22
Person 3	80	Person 3	Person 23	50	Person 23
Person 4	27	Person 4	Person 24	91	Person 24
Person 5	83	Person 5	Person 25	94	Person 25
Person 6	83	Person 6	Person 26	33	Person 26
Person 7	60	Person 7	Person 27	81	Person 27
Person 8	68	Person 8	Person 28	94	Person 28
Person 9	87	Person 9	Person 29	82	Person 29
Person 10	88	Person 10	Person 30	50	Person 30
Person 11	42	Person 11	Person 31	60	Person 31
Person 12	**57**	**Person 7**	Person 32	60	Person 32
Person 13	43	Person 13	Person 33	80	Person 33
Person 14	75	Person 14	**Person 34**	**88**	**Person 2**
Person 15	50	Person 15	Person 35	67	Person 35
Person 16	93	Person 16	Person 36	83	Person 36
Person 17	80	Person 17	Person 37	60	Person 37
Person 18	67	Person 18	Person 38	70	Person 38
Person 19	80	Person 19	Person 39	25	Person 39
Person 20	87	Person 20	Person 40	92	Person 40

Table B.4. Results of fusion with Borda Count (Dataset-2)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Score**	**Result**	**Person**	**Score**	**Result**
Person 1	10	Person 1	Person 21	6	Person 21
Person 2	7	Person 2	Person 22	7	Person 22
Person 3	10	Person 3	Person 23	7	Person 23
Person 4	8	Person 4	Person 24	5	Person 24
Person 5	6	Person 5	Person 25	8	Person 25
Person 6	4	Person 6	Person 26	8	Person 26
Person 7	3	Person 7	Person 27	6	Person 27
Person 8	5	Person 8	Person 28	8	Person 28
Person 9	8	Person 9	Person 29	8	Person 29
Person 10	10	Person 10	Person 30	6	Person 30
Person 11	5	Person 11	Person 31	8	Person 31
Person 12	8	Person 12	Person 32	7	Person 32
Person 13	7	Person 13	Person 33	8	Person 33
Person 14	6	Person 14	Person 34	8	Person 34
Person 15	10	Person 15	Person 35	10	Person 35
Person 16	6	Person 16	Person 36	8	Person 36
Person 17	8	Person 17	Person 37	10	Person 37
Person 18	6	Person 18	Person 38	6	Person 38
Person 19	6	Person 19	Person 39	8	Person 39
Person 20	5	Person 20	Person 40	7	Person 40

Table B.5. Results of fusion with Score Level (Dataset-2)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Distance**	**Result**	**Person**	**Distance**	**Result**
Person 1	1	Person 1	Person 21	0.66	Person 21
Person 2	0.71	Person 2	Person 22	0.92	Person 22
Person 3	0.86	Person 3	Person 23	0.62	Person 23
Person 4	0.26	Person 4	Person 24	0.47	Person 24
Person 5	0.64	Person 5	Person 25	0.62	Person 25
Person 6	0.32	Person 6	Person 26	0.65	Person 26
Person 7	1.21	Person 7	Person 27	0.59	Person 27
Person 8	0.68	Person 8	Person 28	0.96	Person 28
Person 9	1.62	Person 9	Person 29	0.45	Person 29
Person 10	**1.12**	**Person 38**	Person 30	0.57	Person 30
Person 11	0.68	Person 11	Person 31	0.45	Person 31
Person 12	0.54	Person 12	Person 32	0.49	Person 32
Person 13	0.49	Person 13	Person 33	0.62	Person 33
Person 14	0.75	Person 14	Person 34	0.51	Person 34
Person 15	0.43	Person 15	Person 35	1.17	Person 35
Person 16	0.63	Person 16	Person 36	0.64	Person 36
Person 17	0.42	Person 17	Person 37	1.04	Person 37
Person 18	0.66	Person 18	Person 38	0.64	Person 38
Person 19	**1.12**	**Person 5**	Person 39	0.8	Person 39
Person 20	0.48	Person 20	Person 40	0.65	Person 40

Table B.6. Results of fusion with Confidence Level (Dataset-2)

Person 1 - Person 20			**Person 21 - Person 40**		
Person	**Value (%)**	**Result**	**Person**	**Value (%)**	**Result**
Person 1	85	Person 1	Person 21	50	Person 21
Person 2	88	Person 2	Person 22	70	Person 22
Person 3	29	Person 3	Person 23	78	Person 23
Person 4	92	Person 4	**Person 24**	**79**	**Person 11**
Person 5	50	Person 5	Person 25	80	Person 25
Person 6	95	Person 6	Person 26	33	Person 26
Person 7	**31**	**Person 13**	Person 27	89	Person 27
Person 8	79	Person 8	Person 28	95	Person 28
Person 9	36	Person 9	Person 29	36	Person 29
Person 10	81	Person 10	Person 30	33	Person 30
Person 11	79	Person 11	Person 31	71	Person 31
Person 12	**50**	**Person 5**	Person 32	50	Person 32
Person 13	54	Person 13	**Person 33**	**71**	**Person 21**
Person 14	47	Person 14	Person 34	88	Person 2
Person 15	91	Person 15	Person 35	86	Person 35
Person 16	97	Person 16	Person 36	60	Person 36
Person 17	33	Person 17	Person 37	83	Person 37
Person 18	86	Person 18	Person 38	50	Person 38
Person 19	**67**	**Person 5**	Person 39	82	Person 39
Person 20	92	Person 20	Person 40	57	Person 40

Table B.7. Responses to the questionnaire's face recognition section

Responses (1-20)					**Responses (21-40)**				
Q1	**Q2-1**	**Q2-2**	**Q2-3**	**Q3**	**Q1**	**Q2-1**	**Q2-2**	**Q2-3**	**Q3**
Y	4	3	3	c	Y	5	5	5	b
Y	4	5	4	b	Y	3	4	3	b
Y	2	4	3	c	Y	5	4	4	b
NA	4	4	4	b	Y	4	5	4	b
Y	4	3	5	c	Y	4	5	5	b
Y	5	4	5	b	Y	5	4	5	c
NA	4	4	5	c	Y	5	5	4	c
NA	3	3	3	b	Y	3	2	4	c
Y	4	3	5	b	Y	5	5	5	b
Y	3	4	3	b	Y	5	5	5	c
Y	4	3	5	b	Y	3	3	3	c
Y	4	5	5	b	Y	4	5	5	b
Y	5	5	4	b	Y	4	5	5	b
Y	5	5	4	a	Y	3	3	4	c
Y	5	5	5	na	Y	4	5	3	c
Y	5	4	5	c	Y	3	4	4	b
Y	2	3	3	b	Y	5	5	5	b
Y	4	3	5	b	NA	4	3	3	b
Y	5	5	3	c	Y	4	5	5	na
Y	5	4	4	c	Y	4	4	4	na

Table B.8. Responses to the questionnaire's hand recognition section

Responses (1-20)					**Responses (21-40)**				
Q1	**Q2-1**	**Q2-2**	**Q2-3**	**Q3**	**Q1**	**Q2-1**	**Q2-2**	**Q2-3**	**Q3**
N	4	4	4	b	Y	5	4	4	b
Y	2	3	4	b	Y	2	3	3	b&c
Y	1	1	1	b	Y	5	4	4	b
NA	2	4	4	b	Y	3	5	4	b
Y	3	4	5	b	Y	3	5	5	b
Y	3	4	5	b	Y	5	4	5	b
NA	4	4	5	b	Y	3	3	3	b
NA	3	3	3	b	Y	4	3	4	c
Y	5	5	5	b	Y	5	5	3	b
Y	3	4	4	b	Y	4	5	4	b
Y	5	5	5	b	Y	3	3	3	b
Y	5	5	5	c	Y	4	5	5	b
Y	5	5	4	b	Y	4	5	5	c
Y	4	5	4	c	Y	3	3	4	c
Y	3	2	5	b	Y	3	4	4	c
Y	5	4	5	c	Y	2	4	4	b&c
Y	2	3	3	b	Y	5	5	5	b
Y	3	3	4	b	NA	2	3	3	c
Y	2	2	5	b	Y	4	4	5	b
Y	5	5	5	c	Y	4	4	4	b

REFERENCES

1. Jain, A. K., A. Ross, and S. Prabhakar, “An Introduction to Biometric Recognition”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 14, No. 1, January 2004.

2. O’ Gorman L., “Securing Businesss Front Door: Password, Token, and Biometric Authentication”, in *Guarding Your Business: An Architecture for Security*, edited by S. Ghosh, M. Malek, E. A. Stohr, Kluwer Press, Ch. 8, 2004.

3. Jain, A. K., L. Hong, and S. Pankanti, “Biometric Identification”, *Communications of the ACM*, pages 91-98, Feb 2000.

4. Ross, A. and A. K. Jain, “Multimodal Biometrics: An Overview”, *Proceedings of 12th Signal Processing Conference (EUSIPCO)*, pp. 1221-1224, 2004.

5. Jain A. K. and A. Ross, “Multibiometric Systems”, *Communications of the ACM*, v.47 n.1, January 2004.

6. Delac K. and M. Grgic, “A survey of biometric recognition methods”, *46th International Symposium Electronics in Marine*, ELMAR-2004, Zadar, Croatia, 16-18 June 2004.

7. Latifi S. and N. Solayappan, “A Survey of Unimodal Biometric Methods”, *Security and Management*, pages 57-63, 2006.

8. Kennedy G., “Thumbs Up for Biometric Authentication!”, *Computer Law Review and Technology Journal*, Vol. 8, pp. 379-407 , 2004.

9. Phillips P. J., A. Martin, C.L. Wilson and M. Przybocky, “An Introduction to Evaluating Biometric Systems”, *Computer*, Feb. 2000.

10. Sedgwick N., “The Need for Standardization of Multi-Modal Biometric Combination”, *Algorithmica Limited*, Cambridge, 2003.

11. Kim K., “Face Recognition using Principal Component Analysis”, Department of Computer Science, University of Maryland, College Park. MD 20742, USA.

12. Zhang R. and H. Chang, “A Literature Survey of Face Recognition And Reconstruction Techniques”, December 2005.

13. Zhao W., R. Chellappa, A. Rosenfeld and P.J. Phillips, “Face Recognition: A Literature Survey”, *CVL Technical Report*, Center for Automation Research, University of Maryland at College Park, October 2000.

14. Lu X., “Image analysis for face recognition”, Department of Computer Science and Engineering, Michigan State University, 2003.

15. Ju Q., “A Literature Review of Image-based Face Recognition”, “http://www-users.cs.york.ac.uk/ juquan/ALiteratureReviewofFaceRecognition.ppt”.

16. Yang J., Y. Yu, and W. Kunz, “An Efficient LDA Algorithm for Face Recognition”, *The Sixth International Conference on Control, Automation, Robotics and Vision*, Singapore, 2000.

17. Active Appearance Models, “http://www.isbe.man.ac.uk/ bim/Models/aam.html”.

18. Edwards G. J., C. J. Taylor, and T. Cootes, “Face recognition using the active appearance model”, *In 5th European Conference on Computer Vision*, 1998.

19. Searching and Fitting, “http://www.schestowitz.com/Research/Literature _Report/HTML/node8.html”.

20. National Science and Technology Council (NSTC), “Face Recognition”, “http://www.biometrics.gov/docs/facerec.pdf”.

21. Elastic Bunch Graph Matching, “http://itb.biologie.hu-berlin.de/ w̃iskott/Projects/ElasticGraphMatching.html”.

22. Wiskott L., J. M. Fellous, N. Kruger, and C. von der Malsburg, “Face Recognition by Elastic Bunch Graph Matching”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 7, pp. 775-779, July 1997.

23. Gökberk B., M. O. Irfanoglu, L. Akarun, and E. Alpaydin, “Selection of Location, Frequency, and Orientation Parameters of 2D Gabor Wavelets for Face Recognition”, *Lecture Notes in Computer Science*, Volume 3161, p. 138, 2005.

24. Gökberk B., L. Akarun, and E. Alpaydin, “Gabor Wavelet-based Pose Estimation for Face Recognition”, *Proceedings of the 16th International Symposium on Computer and Information Sciences (ISCIS)*, pp. 275-280, Antalya, Turkey, November 2001.

25. Sanchez-Reillo R., C. Sanchez-Avilla, and A. Gonzalez-Marcos, “Biometric Identification through Hand Geometry ”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1168-1171, 2000.

26. Bulatov Y., S. Jambawalikar, P. Kumar, and S. Sethia, “Hand recognition using geometric classifiers”, Manuscript, 2002.

27. Jiang X., and W. Xu, “Contactless Hand Recognition”, Project Report, School of Computer Science, Carnegie Mellon University, Fall 2006.

28. Jain A. K. and N. Duta, “Deformable Matching of Hand Shapes for Verification”, *Proceedings of IEEE International Conference Image Processing*, 1999.

29. Jain A. K., A. Ross, and S. Pankanti, “A prototype hand geometry-based verification system”, *in Second International Conference on Audio and Video-based Biometric Person Authentication*, (Washington, D.C., USA), pp. 166171, Mar 1999.

30. Kumar A., D. C. Wong, H. C. Shen, and A. K. Jain, “Personal verification using palmprint and hand geometry biometric”, *4th International Conference Audio- and Video-based Biometric Person Authentication*, Guildford, U.K., June 911, 2003

31. Oden C., A. Ercil, V.T. Yildiz, H. Kirmizita, and B. Buke, “Hand Recognition Using Implicit Polynomials and Geometric Features”, *Proc. Third Int'l Conf. Audio-and-Video-Based Biometric Person Authentication*, AVBPA-2001, June 2001.

32. Yöruk E., E.Konukoglu, B. Sankur and J. Darbon, “Shape-based hand recognition”, *IEEE Transactions on Image Processing*, vol.15, no.7, July 2006.

33. Yöruk E., E.Konukoglu, B. Sankur and J. Darbon, “Person Authentication Based On Hand Shape”, *Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference*, 2004.

34. Yoruk E., H. Dutagaci and Bulent Sankur, “Hand biometrics”, *Image and Vision Computing 24 (2006)*, pp. 483497, 2006.

35. Java Standard Edition, “http://java.sun.com/javase/”.

36. The Mathworks - MATLAB and Simulink for Technical Computing, “http://www.mathworks.com/”.

37. FMJ (Freedom for Media in Java), “http://fmj.sourceforge.net/”.

38. MySQL, “http://www.mysql.com/”.

39. JGoodies (Java User Interface Design), “http://www.jgoodies.com/”.

40. JMatIO package, “http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=10759&objectType=File”.

41. JUnit, Testing Resources for Extreme Programming, “http://www.junit.org/index.htm”.

42. Log4j, Logging Services Project, Apache Software Foundation, “http://logging.apache.org/log4j/docs/download.html”.

43. Gökberk B., “Three Dimensional Face Recognition”, PhD Thesis, Boğaziçi University, 2006.

44. Dix A., J. Finlay, G. Abowd, and R. Beale, “Human-Computer Interaction/”, Chapter 11, Prentice-Hall, 1998.

45. Kraemer E., “Evaluating the User Interface”, “http://www.cs.uga.edu/ eileen/4900/Notes/evaluation/index.htm”.