ENHANCING WEAK BIOMETRIC AUTHENTICATION BY ADAPTATION AND IMPROVED USER-DISCRIMINATION

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (by Research)
in
Computer Science

by

Kumari Vandana
200407006
vandanaroy@students.iiit.ac.in

International Institute of Information Technology
Hyderabad, INDIA
January 2007

Acknowledgments

Research is not possible in solitude. As any other person, I have been helped a great deal in the course of my Masters journey by the people around me. This thesis is a complex amalgamation of my work and the influence of several people along the way. I shall now set on the impossible task of acknowledging their huge contributions through the following meager number of words.

Foremost, I thank my advisor, Dr. C. V. Jawahar for his valuable guidance. He not only provided the direction for my research with his valuable insights and technical advice, but also supported me morally with general doses of encouragement. His generous feedback have made this work possible.

I thank Dr. Jawahar again, along with Dr. P. J. Narayanan for giving me an opportunity to do an internship at the Center of Visual Information Technology (CVIT). The experience during the internship period aroused an interest in research finally culminating in this thesis. Evidently, this work is built upon the faith they had shown in me. I am also indebted to Dr. Anoop Namboodiri who has played an important role in my research with his extremely useful inputs. His help in understanding the various concepts in biometrics has been invaluable.

Some people influence professionally, some personally, but there are always a few, who have a major impact on almost all parts of life. S. Manikandan is not only a great friend and confidant, he aided me at every step with his acute mathematical and technical skills. His ample contribution in the many derivations throughout this thesis is worth a mention. A. Balasubramanian, a friend since my undergraduate studies, was always around with his moral and psychological support apart from constructive inputs and insightful advice. I would be doing great injustice if I do not express my gratitude to all the members of CVIT for their constant support and encouragement. Particularly, I thank Satyanarayana who was always there to assist in matters related to infrastructure.

The set of friends, I was associated with at IIIT have made the past few years of my life a very memorable experience. All the joy and fun that I had with them, had made life exciting and enjoyable. Each one of them has taught me a lot and left an indelible impression. I shall be forever thankful to each one of them.

Finally, I cannot thank my family enough for their unflinching faith in me. None of this would have been remotely possible if not for their unparalled love. I dedicate this thesis to them.

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "ENHANCING WEAK BIOMETRIC AUTHENTICATION BY ADAPTATION AND IMPROVED USER-DISCRIMINATION" by Kumari Vandana, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Advisor: Dr. C. V. Jawahar

Abstract

Biometric technologies are becoming the foundation of an extensive array of person identification and verification solutions. Biometrics is defined as the science of recognising a person based on certain physiological (fingerprints, face, hand-geometry) or behavioral (voice, gait, keystrokes) characteristics. Weak biometrics (hand-geometry, face, voice) are the traits which possess low discriminating content; they change over time for each individual. Thus they show lower performance as compared to the strong biometrics (eg. fingerprints, iris, retina, etc.). Due to exponentially decreasing costs of the hardware and computations, biometrics has found immense use in civilian applications (Time and Attendance Monitoring, Physical Access to Building, Human-Computer Interface, etc.) other than the forensics ones (e.g. criminal and terrorist identification). Various factors come into picture while selecting biometric traits for civilian applications, most important of which are user psychology and acceptability. Most of the weak biometric traits have little or no association with criminal history as against fingerprints (a strong biometric); data acquisition is also very simple and easy with weak biometrics. Due to these reasons, weak biometric traits are often better accented for civilian applications than the strong biometric traits. Moreover, not much research has gone into this area as compared to strong biometrics.

Due to the low discriminating content of the weak biometric traits, they result in poor performance of verification. We propose a feature selection technique called Single Class Hierarchical Discriminant Analysis (SCHDA) specifically for authentication purpose in biometric systems. The SCDHA recursively identifies the samples which overlap with the samples of the claimed identity in the discriminant space built by the single-class discriminant criterion. If samples of claimed identity are termed "positive" samples, and all the other samples "negative" samples, the single-class discriminant criterion finds an optimal transformation such that the ratio of the negative scatter with respect to positive mean over the positive within-class scatter is maximized, thereby pulling together the positive samples and pushing the negative samples away from the positive mean. Thus SCHDA results in building an optimal user-specific discriminant space for each individual where the samples of the claimed identity are well-separated from the samples of all the other users. Performance of authentication using this technique is compared with the other popular existing discriminant analysis techniques in the literature and significant improvement has been observed.

The second problem which leads to low accuracy of authentication is the poor stability of permanence of weak biometric traits due to various reasons (eg. ageing, the person gaining or losing weight, etc.). Civilian applications usually operate in cooperative or monitored mode wherein the

users can give feedback to the system on occurrence of any errors. An intelligent adaptive framework is proposed which uses the feedback to incrementally update the parameters of the feature selection and verification framework for the individuals. This technique does not require the system to be re-trained to address the issue of changing features.

The third factor which has been explored to improve the performance of authentication for civilian applications is the pattern of participation of the enrolled users. As the new users are enrolled into the system, a degradation is observed in performance due to the increasing number of users. Traditionally, it is required to re-train the system periodically with the existing users to take care of this issue. An interesting observation is that although the number of users enrolled into the system can be very high, the number of users which regularly participate in the authentication process is comparatively low. Thus, modeling the variation in participating population helps to bypass the re-training process. We propose to model the variation in participating population using the Markov models. Using these models, the prior probability of participation of each individual is computed and incorporated into the traditional feature selection framework, providing more relevance to the parameters of regularly participating users. Both the structured and unstructured modes of variation of participation were explored. Experiments were conducted on varied datasets, verifying our claim that incorporating prior probability of participation helps to improve performance of a biometric system over time.

In order to validate our claims and techniques, we used hand-geometry and keystrokes-based biometric traits. The hand-images were acquired using a simple low-cost setup consisting of a digital camera and a flat translucent platform with five rigid pegs (to assure that the images acquired are well-aligned). The platform is illuminated from beneath so as to simplify the preprocessing of the acquired images. The features used for hand-geometry includes lengths of four fingers, and widths at five equidistant points on each finger. Features of thumb are not used as these measurements for thumb show high variability for the same user. This dataset was used to validate the proposed feature selection technique. For keystrokes-based biometrics, the features used were the dwell time (duration of key-press event) and flight time (duration between key-release and next key-press events)of each key, and the number of times backspace and delete key were pressed. Data was collected from subjects who were not accustomed to a particular kind of keyboard (French keyboard). The features extracted from this dataset were time-varying and was used to validate the concept of incremental updation.

In this thesis, we identify and address some of the issues which lead to low performance of authentication using certain weak biometric traits. We also look into the problem of low performance of authentication in large-scale biometrics for civilian applications.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Establishing the identity of a person is becoming critical in our vastly interconnected society. Since the beginning of civilization, humans have used faces to identify known (familiar) and unknown (unfamiliar) individuals. This simple task became increasingly more challenging as population increased. Also, more convenient methods of travel introduced many new individuals into once small communities. There are various evidences as summarized in [3] that show that various other human physiological characteristics were used in early civilizations. For example, numerous hand prints surrounded the paintings in a cave 31,000 years old are felt to have acted as an un-forgeable signature of its originator. There is also an evidence that fingerprints were used as a person's identity as early as 500 B.C. Babylonian business transactions are recorded in clay tablets that include fingerprints. True biometric systems began to emerge in the latter half of the twentieth century, coinciding with the emergence of computer systems. The nascent field experienced an explosion of activity in the 1990s and began to surface in everyday applications in the early 2000s.

1.1 Biometrics

The goal of authentication is to protect a system against unauthorized use. This feature also allows for the protection of individuals by denying the possibility for someone else to impersonate characteristics of authorized users. Authentication procedures are based on the following approaches:

1. Knowledge - Known information regarding the claimed identity that can only be known or produced by an individual with that identity (e.g., passport, password, personal identification number (PIN)). There is high probability that the individual may forget these attributes or share them with others.

2. Possession - The user is authorized by the possession of an object (smart card, optical card, etc). However, these cards can be stolen or broken and can be shared hence threatening the security of the system.

3. Property - Quantitative human characteristics of the individual (e.g., biometrics).

Thus, utilizing biometrics for person authentication is becoming convenient and considerably more accurate than the current and traditional methods. This is because biometrics links the event to a particular individual (a password or token may be used by someone other than the authorized user). Biometric is also convenient (nothing to carry or remember) and accurate. Due to these reasons, it is becoming socially acceptable and inexpensive.

1.2 Types of Biometrics

The word "biometrics" is derived from the Greek words "bio" (life) and "metrics" (to measure). Biometrics is simply the measurement and use of the unique characteristics of living humans to distinguish one from another. Formally, biometrics is defined as the science of establishing identity of persons based on their certain physiological and behavioral characteristics. Some examples are shown in Figure-1.1. A brief comparison of these traits based on features used, distinctiveness (ability to distinguish between two individuals), permanence (invariance of the trait over a period of time) and acceptability are summarized in Table-1.1.

Physiological Biometrics: Physiological biometrics measure the distinct traits that people have, usually (but not always or entirely) dictated by their genetics. Examples of physiological biometrics include advanced techniques like DNA, retinal and iris scans, face, hand and finger geometry, finger and palm prints, etc.

Behavioral Biometrics: The second category of biometrics is behavioral. Behavioral biometrics measure the distinct actions that humans take, which are generally very hard to copy from one person to another. These actions are learned by the individuals over time. Examples of behavioral biometrics include voice printing and gait analysis, which use computers to analyze the sound created by the human voice box or the movement of a person walking. Other common examples include signature, hand-writing, keystrokes, etc.

Strong and Weak Biometrics: The biometric traits can further be classified into two categories: *Strong Biometrics* and *Weak Biometrics*. This classification is based on the characteristics: degree of discriminating content and degree of permanence. Strong biometrics possess high discriminating content and high degree of stability, while weak biometrics have low discriminating content and change over time. Weak biometrics show low performance as compared to the strong biometrics. Examples of weak biometrics include hand-geometry, face, keystroke dynamics, etc. The strong and weak biometrics are discussed later in Chapter 2 in detail.

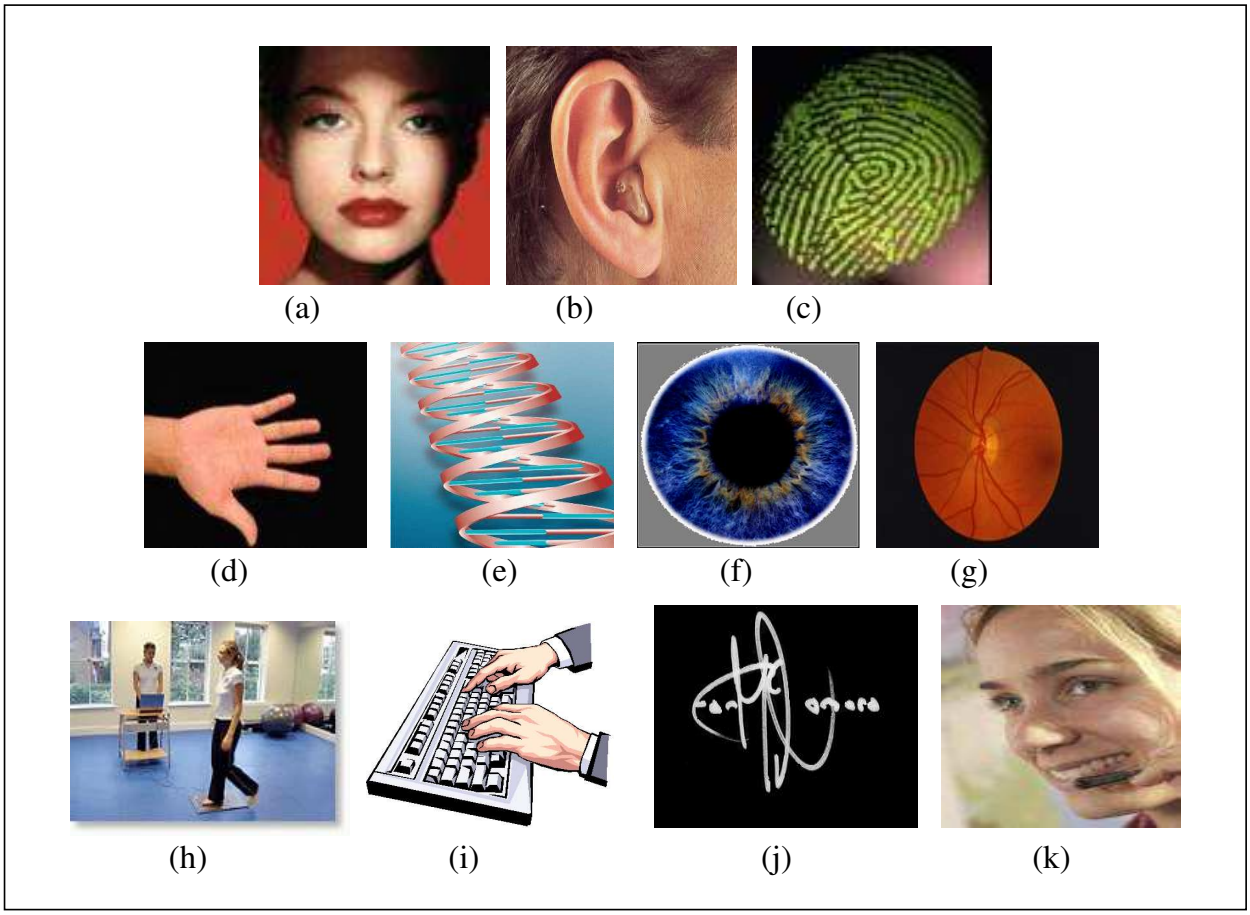


Figure 1.1: Types of biometrics: (a) face, (b) ear, (c) fingerprints, (d) hand-geometry, (e) DNA, (f) iris, (g) retina, (h) gait, (i) keystroke dynamics, (j) signature, and (k) voice.

Biometric Trait	Features Extracted	Distinctiveness	Permanence	Performance	Acceptability
Face	Distance between, eyes, nose, mouth and jaw edges	Low	Medium	Low	High
Ear	Shape of the ear, structure of the cartilaginous tissue of pinna	Medium	High	Medium	High
Fingerprint	Whorls, arches, and loops; pattern of ridges, furrows, and minutiae	High	High	High	Medium
Hand-Geometry	Lengths, widths of fingers; diameter, thickness of palm	Medium	Medium	Medium	Medium
DNA	DNA strands (double helix structure) present in human cell	High	High	High	Low
Iris	Rings, furrows, and freckles in colored tissue around pupil	High	High	High	Low
Retina	Layer of blood vessels at back of the eyeball	High	Medium	High	Low
Gait	Coordinated cyclic, temporal motions	Low	Low	Low	High
Keystroke	Dwell, Flight times, number of mistakes	Low	Low	Low	Medium
Signature	Speed, strokes, curvature, pressure applied	Low	Low	Low	High
Voice	Fundamental frequency, nasal tone, cadence, inflection	Low	Low	Low	High

Table 1.1: Description and comparison of various biometric traits. *Courtesy: A Survey of Biometric Recognition Methods [2].*

Year/Time-Period	Characteristics
1892-1936	Various Biometric Techniques were proposed and patented. For example, Fingerprints (1892), hand-geometry (1858), iris(1936), etc.
1960-1987	Automation of various biometric techniques: Semi-automation of Face Recognition (1960) Automated Signature Recognition (1965) Minutiae extraction using sensors in Fingerprints (1975)
1988	Eigenface technique developed for face recognition
1994	Integrated Automated Fingerprint Identification System (IAFIS) competition held
2000	First Face Recognition Vendor Test (FRVT) held, sponsored by the multiple US Government Agencies. Later, FRVT for fingerprints was organized in 2003 and for iris in 2006.

Table 1.2: Initial developments and major events that occurred in biometrics

Soft Biometrics: The soft biometric traits are the ancillary information, such as gender, age, height, skin and eye color, etc., [4] which are collected during enrollment phase. Although they are not used often, they can be used to complement the identity information provided by the primary biometrics, such as fingerprints, hand-geometry, etc. They provide some evidence about the user identity that could be beneficial. Improvement in performance can be achieved by integrating the soft biometrics with the primary biometrics.

1.2.1 Evolution of Biometric Techniques

Evolution of biometric techniques, traits, and major events that took place in the area of biometrics are summarized in Table-1.2. We discuss the post 2000 research trends in the areas of strong and weak biometrics. Lot of work was already done before 2000 in the areas of fingerprints, iris, retina, etc., and acceptable performance was obtained using these techniques. These traits showed good performance due to their high discriminating content. After 2000, most of the work in the area of strong biometrics is concentrated on areas of image acquisition and fast computing, matching and retrieval. Randolph and Smith [5] highlighted the necessity to preprocess the acquired fingerprint to improve the clarity of the feature details. The authors proposed an approach to fingerprint enhancement using filter banks that operate in a binary finite field. Chen *et al.* [6] proposed a feature extraction technique which uses gradient-direction on wavelet transform as the discriminating texture features. Yuan and Shi [7] stated that phase rather than amplitude information provides the most significant information within an image and proposed to use 2D phase congruency which is invariant to changes in intensity of contrast to extract features from the iris images. An impor-

tant ingredient of iris image preprocessing is image segmentation which consists of approximating iris as a circumference or an ellipse, localize the iris inner (pupillary) and outer (scleric) borders. Proenca and Alexandre [8] analyzed the important contribution of the accuracy of the segmentation algorithm in the error rates of the iris recognition systems. They also proposed a method for identification of translation errors on pupil segmentation. A lot of work has been reported in the area of reducing the complexity of retrieval and matching algorithms in applications based on strong biometrics. Jun Gao *et. al* [9] proposed to use *Synergetic Pattern Recognition* approach to reduce the time taken for fingerprint recognition. Chan *et. al* [10] proposed a method to improve the speed of matching in fingerprint verification. They proposed to shorten the minutiae extraction process in verification phase by using a small subregion of "live" image for minutiae extraction and subsequent matching. Such small subregions were identified based on accurate detection of reference points in the images along with prior knowledge of complete fingerprint obtained during enrollment phase. As iris-based recognition systems are used in large-scale applications, the iris localization is needed to be very fast. Bakry [11] used modular neural network to solve this problem by dividing the data into three groups, thereby reducing the computational complexity and decreasing the time and memory required during test of an image.

On the other hand, the other traits, such as, voice, gait, hand-geometry are still active areas of research for improvement in performance. This is due to the fact that the weak biometrics perform poorly mainly due to their low discriminating content. Most of the work is concentrated on extraction and selection of better features, and using better classifiers to improve performance. Reillo [12] proposed to use Gaussian Mixture Model (GMM) to improve the performance of hand-geometry based verification. Gonzalez *et al.* [13] proposed to extract and reduce the parameters of hand-geometry and used the Multi-Layered Perceptron (MLP) and k nearest neighbor (KNN) classifiers. Erdem *et al.* [14] combined hand-geometry and palm-texture features and used the Independent Component Analysis (ICA) classifier for verification. Ender *et al.* [15] used Hausdorff distance of hand contour and independent component features of hand-silhouette images. There is a lot of scope for improvement in the area of weak biometrics. Also, the research started in the area of fusion of multiple biometric traits to achieve better accuracy. Multimodal biometrics are looked to as a means of (i) lowering the error rates, providing secondary means of enrollment, verification and identification, and (iii) combating attempts to spoof biometric systems using non-live biometrics, such as fingerprints, hand-geometry, etc. Various multimodal biometrics and fusion strategies have been worked on in literature [16, 17, 18, 19, 20, 21, 22, 23, 24, 25].

1.3 Identification vs. Verification

The problem of resolving the identity of a person can be categorized into two fundamentally distinct types of problems with different inherent complexities: (i) verification and (ii) identification. These

tasks are shown pictorially in Figure-1.3.

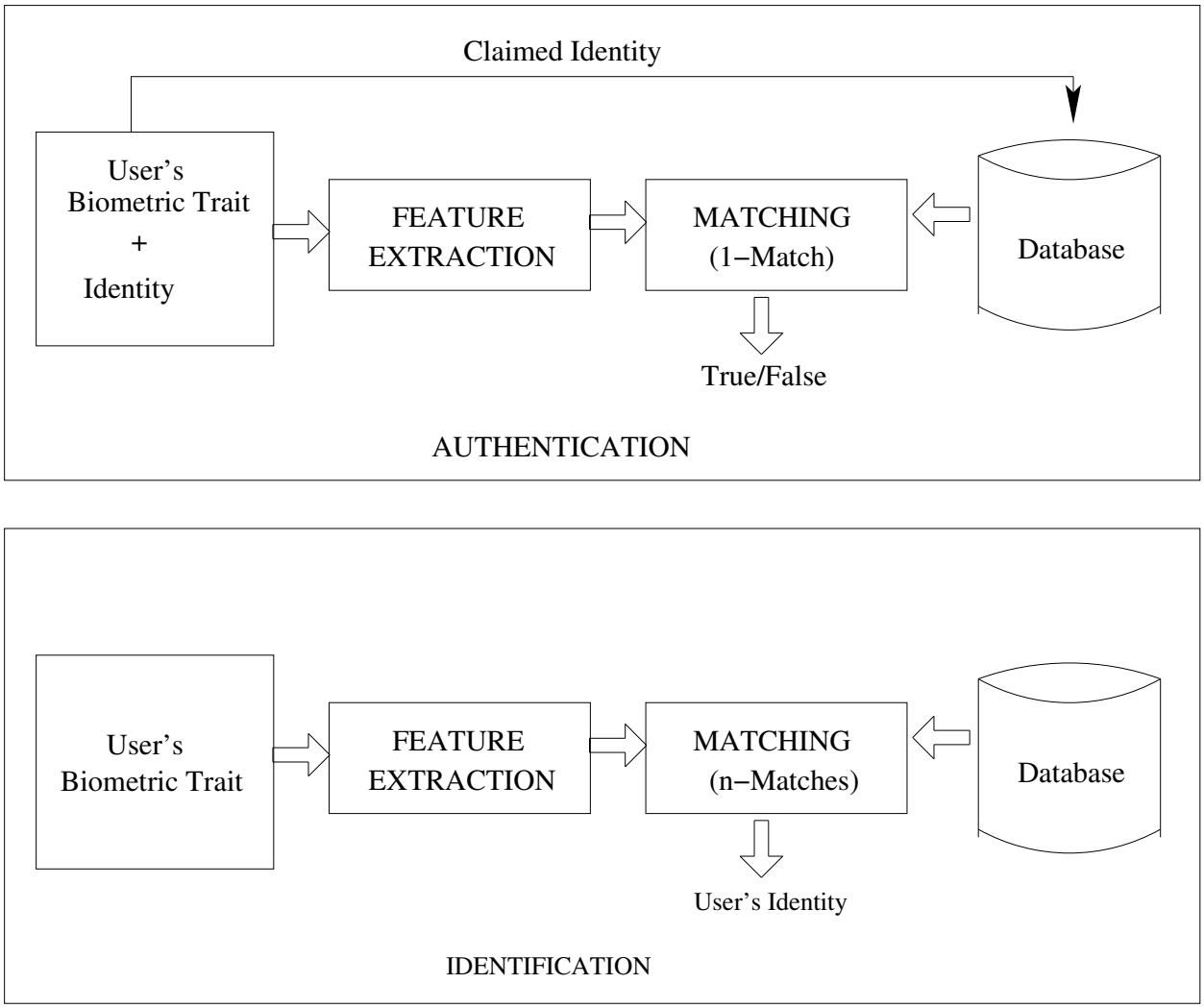


Figure 1.2: Two tasks in biometric systems: Authentication and identification.

Verification: Verification (also called authentication) refers to the problem of confirming or denying a person's claimed identity (Am I who I claim I am?). In the verification mode, the system validates a person's identity by comparing the captured biometric data with his own biometric samples stored in the system database. In such a system, an individual who desires to be recognized claims an identity, usually via a PIN or user name, etc., and the system conducts a one-to-one comparison to determine whether the claim is genuine or not. Identity verification is typically used for positive recognition, where the aim is to prevent multiple people from using the same identity.

Identification: Identification refers to the problem of establishing a subject's identity (Who am I?). In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Hence, the system conducts a one-to-many comparison to establish an individual's identity or reports a failure message if the individual is not enrolled in the database. Identification is a critical component in negative recognition applications where the system establishes whether the person is who he denies to be. The purpose of negative recognition is to prevent a single person from using multiple identities.

1.4 Evaluation of Biometric Systems

No metric is sufficiently adequate to give a reliable and convincing indication of the identification accuracy of a biometric system. The decision made by a biometric system is either a genuine individual type of decision or an impostor type of decision. This can be represented by two statistical distributions called genuine distribution and impostor distribution, respectively. For each type of decision, there are two possible decision outcomes, true or false. Therefore, there are a total of four possible outcomes:

1. A genuine individual is accepted.
2. A genuine individual is rejected.
3. An impostor is rejected.
4. An impostor is accepted.

In general, false accept rate (FAR) defined as the rate of acceptance of an impostor and false reject rate (FRR), defined as rate of rejection of a genuine user are used to evaluate a biometric system. Equal error rate (EER), the value at which the FAR and FRR are equal is also used. The EER makes the comparison of the systems independent of thresholds. In addition, the Receiver Operating Characteristic (ROC) of a biometric system is a graphical depiction of the relationship between the FRR and FAR as a function of the threshold's value. An ROC provides an empirical assessment of the system performance at different operating points which is more informative than FAR and FRR. Percentage accuracy (defined as the percentage of correct verification or identification) is also used to quantify performance of a biometric system.

1.5 Applications of Biometric System

There is a wide range of applications where biometric systems can be used. The applications can be divided into three broad categories:

Civilian Applications These are the low-end security applications. Examples include computer network login, Internet access, physical access control, mobile phone, medical recodes management etc. These applications and the related issues are discussed in detail in Chapter 2.

Government Applications These are the medium security applications, including ID card, driver's license, voter's license, welfare-disbursement, border control, passport control, etc.

Forensic Applications These are high-end security applications. Examples include criminal and terrorist investigation, corpse identification, parenthood determination and many more.

Traditionally, commercial applications have used knowledge-based authentication systems such as smart-card, PIN, passwords. Government applications have used token-based systems and forensic applications have relied on help from human experts to match the features. Selected examples of biometric systems that have been deployed successfully, mainly for civilian applications are shown in Figure-1.3.

1.6 Scope of the Thesis

In this thesis, the emphasis is on the use of biometric authentication for civilian applications, especially in situations where the enrolled population is very high. The processing and pattern recognition issues of weak biometrics are different from the strong biometrics. The performance of identification and verification using strong biometrics has already reached almost the peak. The research in this area is mainly concentrated now on the speed of matching and retrieval. However, there is scope of considerable improvement in performance in case of weak biometrics. In this work, we identify some of the problems which lead to low performance of authentication using weak biometrics. The low discriminating content of weak biometrics can be handled using an appropriate discriminant analysis technique. We propose a user-specific feature selection technique to improve the discriminating content of weak biometrics. The low stability of weak biometrics is addressed by using the users' feedback on the errors occurring during authentication. We propose a learning framework to update the parameters of feature selection framework and hence, adapt to the changing features over time. Further, in case of civilian applications of biometrics, the population size increases over time, resulting in degradation of performance. We use the fact that the size of the regularly participating population is relatively lower than the total enrolled population. We model the variation in participating population and use the *apriori* probability of participation in the feature selection framework. Thus, only the regularly participating population contributes to the authentication process.

Figure 1.3: Examples of biometric systems: (a) The FacePass system used in ATMs(http://www.viisage.com), (b) ExpressCard using hand-geometry , (c) HandReader developed by Recognition Systems (http://www.handreader.com/transition/index.htm) (d) Bio-Pen for handwriting based recognition (http://www.bio-pen.com/), (e) Border-passage system using iris-recognition at Heathrow Airport, (f) INSPASS developed by Recognition Systems, (g) A fingerprint point-of-sale terminal by Indivos Inc., (h) Biometric Cyber Series (Finger Geometry) by Accu-Time Systems(http://www.accu-time.com/cyber_bio-fg.htm)

1.7 Organization of the Thesis

This thesis focuses on enhancing performance of authentication using weak biometrics for civilian applications. So far in this chapter, we provide an introduction to biometrics, various types of biometrics and the metrics used to quantify the performance in biometric systems. We also discuss broadly the various applications of biometrics, and then stated the scope of the thesis. The rest of the thesis is organized as follows:

- Weak and strong biometrics are discussed in detail in Chapter 2. We also discuss in detail the civilian applications of weak biometrics and various issues involved in selecting the biometric trait for weak biometrics. We argue that weak biometrics are better accepted for civilian applications than strong biometrics. We, then analyze various factors which lead to low performance of authentication using weak biometrics. The weak biometric traits, namely hand-geometry and keystrokes, used in this thesis for evaluation and verification of our techniques are then discussed in detail. We explain in detail the generative model for the samples of weak biometrics and then discuss the datasets used in this thesis in detail.

- Chapter 3 focuses on improving the discriminating content of weak biometrics by feature selection. Various popular feature selection techniques used in the context of biometrics are presented. Then, we discuss in detail the novel user-specific feature selection technique called Single-Class Hierarchical Discriminant Analysis (SCHDA) technique for authentication using weak biometrics. We compare the performance of authentication using SCHDA with the well known feature selection techniques and show that SCHDA performs superior to the existing techniques.

- In Chapter 4, the problem of low degree of stability is addressed. An incremental feature selection technique is discussed in detail. This technique relies on the fact that the civilian applications operate in cooperative mode. The incremental feature selection framework uses the feedback provided by the users on every incorrect verification, and performs updation of the parameters of feature selection. Experiments are carried out on the time-varying datasets.

- Chapter 5 discusses in detail the problem of performance degradation due to large enrolled population. The techniques used to model the structured and un-structured variation in population are discussed. Various experiments are conducted to show the applicability of modeling variation in participating population.

- Chapter 6 concludes the work presented in this thesis with the summary of the work, analysis of results and scope for future work.

Chapter 2

Weak Biometrics for Civilian Applications

Selection of a biometric trait for an application depends upon various factors. A biometric trait is selected for an application based on the requirements of the given application, the characteristics of the application, and properties of the biometric traits. The following are the desirable characteristics of a good biometric trait [26].

Universality: Each person should have the biometric characteristic.

Distinctiveness: The biometric characteristic should be able to distinguish between any two persons.

Permanence: The characteristic should be invariant over a period of time.

Collectability: The characteristic can be measured quantitatively.

Desirability: A biometric trait can be selected for an application based on the level of accuracy. Also, user acceptability need to be taken into account for some of the application.

Based on these characteristics and many more, a biometric trait qualifies for use in a given application. Biometric traits vary in the level of perfection on all the above desirable criteria. There is no single trait which meets all the desirabilities perfectly.

2.1 Strong and Weak Biometrics

As stated in Chapter 1, based on the characteristics *degree of distinctiveness* and *degree of permanence*, biometric traits can be broadly classified into two categories: (i) strong biometrics, and (ii) weak biometrics.

Strong biometrics are the biometric traits which possess high discriminating content and have high degree of stability or permanence. On the other hand, weak biometrics are the traits which cannot distinguish between any two users with high accuracy and they change over time, mainly due to physical changes occurring in human body due to aging, gain or loss of weight, etc.

Popular examples of strong biometrics include fingerprints, DNA, iris, retina, etc. while the behavioral biometrics, like keystrokes, signature and some physiological traits, such as hand-geometry, face, etc., are the examples of weak biometrics.

2.2 Civilian Applications

The perception that biometric technologies are hi-tech, high-cost systems and can only be afforded in forensics and high-security military installations, is rapidly changing. Various application domains of biometrics can be found at [27, 28, 29]. Spiraling increase in the availability of inexpensive computing resources, advances in image understanding, better matching strategies, cheaper sensing technologies, increasing demand for the identification needs are forcing biometric technology into new applications requiring positive personal identification.

The biometrics-based authentication is being adopted in a very broad range of civilian applications, requiring medium-to-low security. Some of the examples where biometrics based authentication can be deployed are:

- Low-level Security Applications
    1. Time and attendance monitoring in schools, colleges and offices.
    2. Controlling physical access to buildings.
    3. Public benefit programs such as issue of welfare, aid to dependents.
    4. Membership management in libraries, clubs, etc.
- Medium-level Security Applications
    1. Border control to allow travelers to bypass long immigration check queues.
    2. Voting systems in order to prevent proxy voting.
    3. Banking and ATM in rural areas.
    4. Internet transactions to ensure security, privacy and confidentiality.
    5. Network and PC security to ensure usage by authorized people.

Forensics are the high-end applications of biometrics. Examples include border security, criminal and terrorist identification, etc. These applications require the system to be highly accurate, as the cost of error in forensics is very high. In such a situation, the system cannot work in co-operative mode. Thus, for such applications, the False Accept Rate (FAR) is needed to be very low. In contrast, civilian applications are the medium-to-low security applications of biometrics, hence the cost of error is not high in case of civilian applications. Civilian applications can operate in cooperative mode, which is not possible in the case of forensic applications. These applications require FRR to be very low. In the next section, we discuss the requirements of biometric traits for civilian applications.

2.3 Requirements of Biometric Traits for Civilian Applications

There are various requirements that a biometric trait should satisfy in order to be used for civilian applications. In the following, we discuss various requirements of biometric traits for civilian applications, and argue that weak biometric traits are better accepted for use in civilian applications.

User and Social Acceptability: User acceptability plays a major role in selecting biometric trait for civilian applications.

- User Psychology: As fingerprints have strong historical relationship with criminal identification, users may think that their samples may be used to match against criminals' database. Weak biometrics, such as hand-geometry and voice have no perceived association with criminal history.
- Medical Reasons: Many strong biometric traits may not be socially acceptable due to medical reasons. Iris and retina based recognition systems require exposing eye to the sensor daily which may have adverse affect on them. Weak biometric traits, such as face, voice, hand geometry, etc. cannot usually be affected by exposure to the capture devices like cameras and microphone used for data acquisition.
- Invasiveness: Users may be afraid that the biometric trait (e.g. DNA) may reveal some of the diseases. Weak biometrics, such as face, voice, etc., do not posses any such invasive content.
- Contact/Non-contact type Sensors: Non-contact type sensors are preferred because of social and hygiene issues. Camera based sensors largely used by weak biometric systems are the best options for use in large-scale civilian authentication systems.

Ease of Use: Data acquisition device should be very easy to use so that little or no training is needed to use the system. The users may not be comfortable staring into the capture device

for even 5-10 seconds for acquisition of iris images. However, almost all weak biometrics usually have simple data acquisition device.

Affordability and Cost: Cost is the major issue to be addressed while deploying authentication system at multiple places, particularly in developing countries. Various costs which must be accounted for are:

- Cost of Hardware: Strong biometrics usually need very expensive sensors, e.g., high resolution cameras or special sensors for data acquisition as against weak biometrics which use low-cost sensors (usually simple cameras) for data capture.
- Cost of Supervision: Supervision or inspection is usually required in very high-end security applications. As civilian applications are low-to-medium security applications, they need minimal or no supervision.
- Cost of Maintenance: Highly skilled expert labors are needed for maintenance of expensive and special scanners used in strong biometric systems such as iris and retina sensors. However, for weak biometrics, the capture device is usually a low-cost camera which does not need any special care or maintenance.

Cost of Error: In case of civilian applications, little error does not cause much loss as compared to that in forensic applications. As weak biometrics show low-to-medium accuracy, they can be used for civilian applications.

Cooperative Mode of Operation: Being less security intensive, these applications can work in cooperative mode in which the users can provide feedback to the system which can be exploited by intelligent or adaptive systems.

Thus, weak biometrics have a significant role in civilian applications. They can easily be used with high user and social acceptance and low cost of deployment and maintenance. Our objective is not to find a new biometric trait to meet all the above requirements. We work within the available traits and discuss directions on addressing Large-Scale Biometric problem for civilian applications.

2.4 Large-Scale Weak Biometrics

Weak biometrics show low to medium accuracy. We need to consider various causes of the low accuracy of authentication using weak biometrics. In this work, we have identified three problems which lead to low performance in civilian applications of weak biometrics. The first two deal with inherent problems of weak biometrics while the third problem arises for use in civilian applications.

Low Discriminating Content: Weak biometrics can distinguish between individuals with a low accuracy. This is because traits such as hand-geometry, face, etc. of some individuals may

be very similar to each other. Because of this, the authentication system gets confused between the samples of different individuals, and shows low performance. In order to improve discriminating content of weak biometrics, it is required to select the features so as to improve the uniqueness of the trait.

Low Degree of Stability: Features of weak biometrics change over time due to physical changes (aging, person gaining or losing weight, etc.). Most of these changes are not abrupt and an intelligent system can adapt to the changing features. The civilian applications usually operate in cooperative mode in which the users provide feedback to the authentication system on every incorrect verification. The system learns by updating the parameters of the feature selection framework based on the feedback given by the users.

Varying Participating Population: Although the system has many individuals enrolled, not all of them participate every day. The variation in participating population in a biometric system can be structured or unstructured. For e.g., users can go on leave for a particular duration (unstructured) while a group can visit the company every year for some time-period (structured). It is required to incorporate the prior information of participation of each individual so that authentication can be performed only on the participating population.

2.5 Examples of Weak Biometrics

So far, we argued that weak biometric traits are better accepted for civilian applications, and discussed the issues related to large-scale weak biometric authentication systems. In this section, we discuss in detail the weak biometric traits used in this thesis. In order to validate our techniques we have chosen hand-geometry and keystroke-based biometrics as examples of weak biometrics. While hand-geometry is a physiological trait, keystroke dynamics is the behavioral trait. Each of these biometric traits are discussed in detail in the following subsections.

2.5.1 Hand-Geometry Based Biometrics

Hand geometry refers to the geometric structure of hand, which includes lengths of fingers, widths at various points on the finger, diameter of the palm, thickness of the palm, etc. These features are not as discriminating as other biometric characteristics (such as fingerprints); they can easily be used for verification purpose. The importance of hand geometry and its user acceptability is discussed in detail by Jain *et al.* [30]. The hand images can be obtained using a simple setup including a low-cost digital camera. However, other biometric traits often require specialized, high-cost scanners to acquire the data. User acceptability for hand-geometry based biometrics is very high as it does not extract detail features of the individual. Thus, for applications where the biometric features are needed to be distinctive enough for verification, hand geometry can be used. Further,

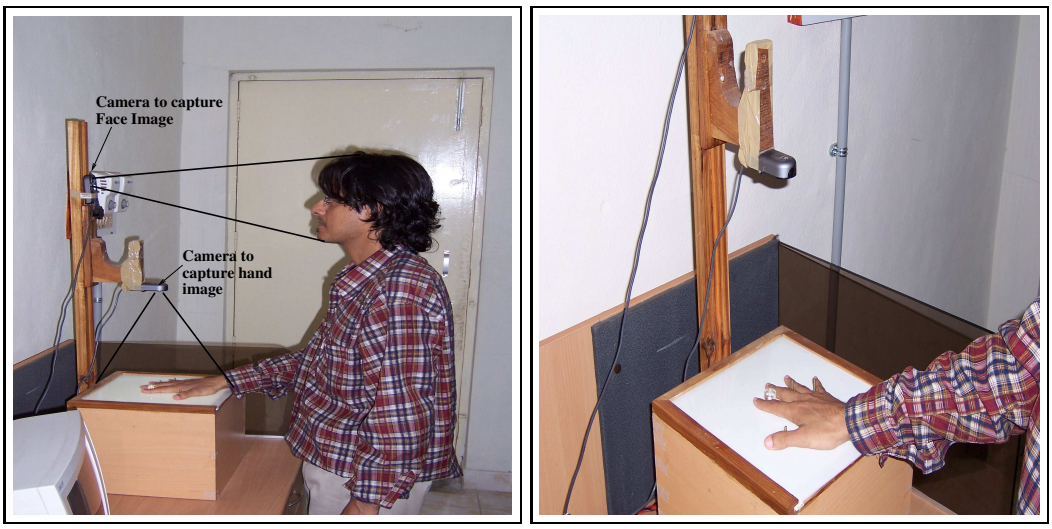
hand geometry features can be easily combined with other biometric traits, such as palm print, fingerprint, etc., in multimodal biometric systems.

The first commercial hand geometry recognition system became available in the early 1970s. They were the first commercially available biometrics device after the early deployments of fingerprinting in the late 1960s [3]. These systems were deployed for three main purpose: physical access control; time and attendance; and personal identification. There has been several hand geometry verification systems published in literature. Jain *et al.* [31] developed a pegged hand geometry verification system for web security. Later Jain and Duta [32] developed another pegged system which aligns the two images and define a metric, Mean Alignment Error as the average distance between corresponding points measured between the images to be verified. Wong and Shi [33] developed system which uses a hierarchical recognition process, with Gaussian mixture model for the one set of features and a distance metric classification for a different set of features.

The hand images are acquired using a setup with two digital cameras (one to capture image of hand and the other to capture face image) and a flat platform with five rigid pegs. The setup is designed for multimodal biometric system in which we are trying to fuse the results of hand-geometry and face-based recognition to obtain better accuracy. The setup is shown in Figure-2.1(a). Both the images (face and hand) can be taken simultaneously using the two cameras shown in the figure. In this work, we address feature selection mechanism for hand-geometry based biometric (unimodal) authentication. The top view of flat surface used to capture hand images is shown in Figure 2.1(b). As we are not using complex image-processing routines to extract features from the image of the hand, we assume that the user places his hand over the flat surface such that the fingers are well separated. The five rigid pegs shown in the figure serve that purpose. The pegs are used to help the user place his hand properly such that the acquired images are well-aligned. The flat surface is translucent, white colored and is illuminated by a light source beneath it to ensure that the background is well separated from the foreground (hand image). This helps to binarize the image and use simple image-processing routines to extract the boundary and hence the features from image of the hand. The image capture for both the unimodal (hand geometry) and multimodal systems is shown in Figure-2.1.

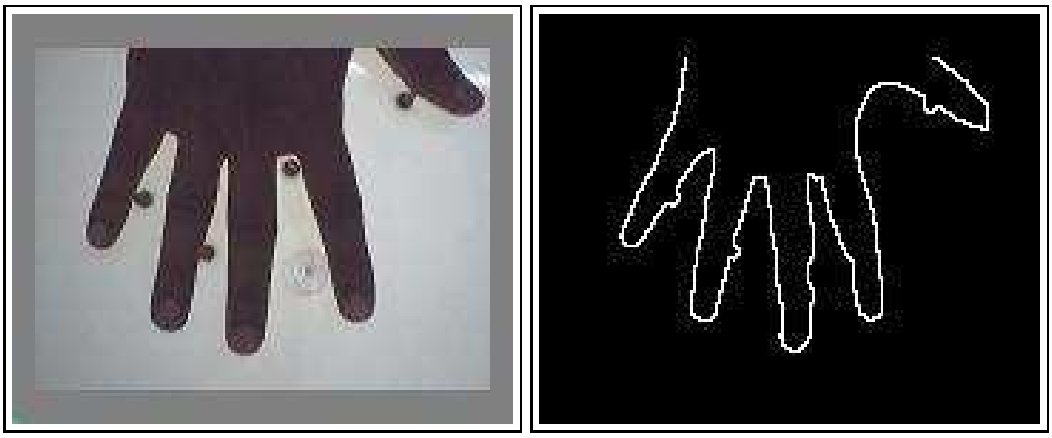
As the hand-image is clearly separated from the background, simple thresholding is used to binarize the image. From this binary image, we obtain the longest contour by using the chain code contour extraction method. The acquired and contour-extracted images are shown in Figure 2.2. The boundary of the hand is defined by the largest contour.

We use a set of features extracted using very simple image-processing algorithms. We use lengths



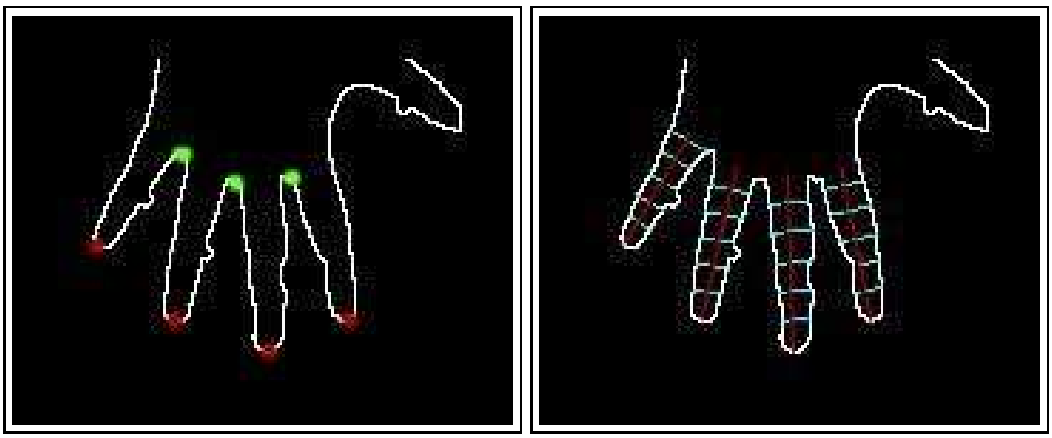
(a) (b)

Figure 2.1: (a) Face and hand image acquisition (b) Hand image acquisition.



(a) (b)

Figure 2.2: (a) The hand-image acquired using the setup (b) Boundary extracted from the hand-image using longest contour.



(a) (b)

Figure 2.3: (a) Boundary of the hand-image with peaks and valleys marked. The green dots show the valleys while the red dots show peaks. (b) The hand-image showing with the raw features marked. The raw features include lengths of four fingers and widths at five equidistant points on each finger.

of four fingers and widths at five equidistant points on each finger as raw features. As these measurements for thumb show high variability for the same person, we do not include the length and widths of thumb in the feature vector for our system. Hence we obtain the feature vector of dimension 24 for each person. The raw features are extracted with the help of landmarks defined as the peaks and valleys. The finger tip points are called peaks and the points joining adjacent fingers are termed valleys. The peaks and valleys of each finger are extracted by traveling along the hand boundary (Figure-2.3(a)). These landmarks are then used to extract raw 24-component feature vector (Figure-2.3(b)). As stated above, these components include lengths of four fingers and widths at five equidistant points on each finger.

2.5.2 Keystrokes Based Biometrics

Psychologists and mathematicians have been experimenting with human actions since as early as the beginning of the 20th century. It has been demonstrated that human actions are predictable in the performance of repetitive, and routine tasks [34]. In the 19th century, observation of the telegraph operators showed that each operator had a distinctive pattern of keying messages over telegraph lines [35]. An operator often recognized who is typing on the keyboard and sending information simply by listening to the characteristic pattern of the Morse code.

Now, the telegraph keys have been replaced by other similar input devices, such as keyboard.

In keystroke biometrics, persons are authenticated or identified based on their keying dynamics which are assumed to be unique for each individual to a large degree. Keystroke dynamics is mostly applicable to the authentication problem. Keystroke dynamics include several different measurements [36, 37, 38, 39] which can be detected when the user presses keys on the keyboard. Most commonly used features include:

- Latency between consecutive keystrokes, *dwell-time*.
- Duration of the keystroke, *hold time*.
- Overall typing speed.
- Frequency of errors.

These features are shown in Figure-2.4.

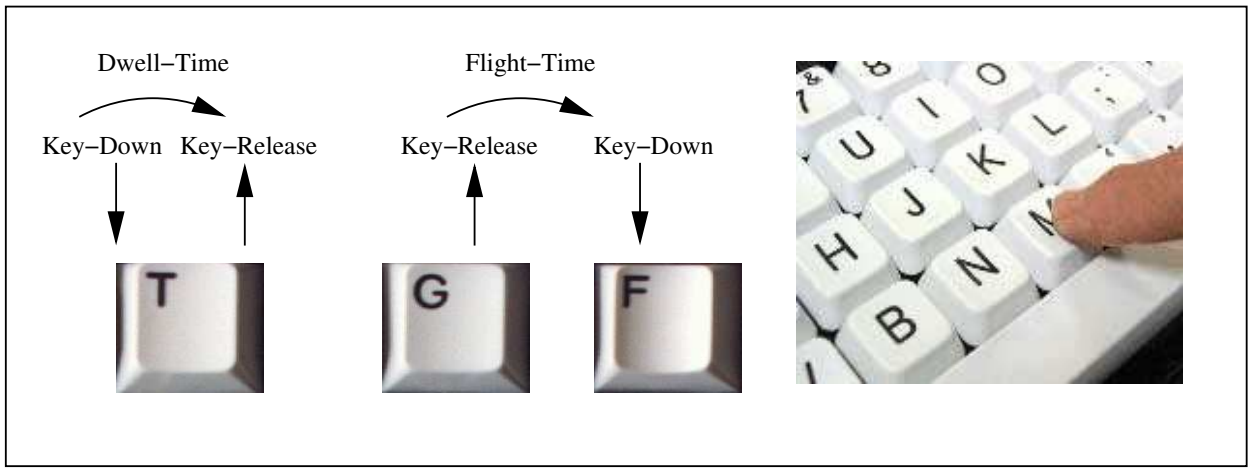


Figure 2.4: Features of keystroke dynamics.

Keystroke based authentication does not require any additional hardware with which to read, scan, view or record the requesting user as every computer is equipped with the keyboard. To authenticate an individual, the system relies solely on the software. To create a user template, the user needs to type a sentence multiple times. User acceptance rate of keystrokes as biometric traits is very high, since the users do not even necessarily notice that such a system is in use.

Keystroke recognition is not considered as an effective single-factor authentication technique because hand injuries, fatigue, and other conditions can affect authentication effectiveness. Also, since keystroke recognition is a relatively new biometric technology, reliable information concerning its effectiveness is not available. However, when combined with other authentication techniques,

such as passwords, it can prove to provide reliable authentication.

Keystroke dynamics continues to be looked at with keen interest in data processing circles and other computer industries that require large amounts of keyboard data entry. For example, network security, PC login, etc. Keystroke dynamics technology is an ideal solution for this type of industry since the employee is already using the keyboard on a consistent basis.

2.6 Generative Model

A generative model as given below is considered for experimental study in addition to the real data is proposed. This models the distribution of individuals and their variations in participation over time.

Characteristics of biometric samples:

1. The samples belonging to an individual are observed to be centered around a representative sample and are scattered within a given range. Hence, Gaussian assumption can be made for the distribution of samples belonging to a particular user. i.e,

$$p(x/\omega_i, \theta_i) \sim N(\mu_{\omega_i}, \sigma_{\omega_i}),$$

   where ω_i is the i^{th} user class and θ_i are the parameters of the samples of user i.

   In order to show that the Gaussian assumption is correct for the data collected, we show the probability distribution of the samples of the hand-geometry data. The probability distribution of the first two components of the raw samples for the hand-geometry data is shown in the Figure-2.5. The bell-shaped curve clearly shows that the samples obey the Gaussian distribution.

2. All weak biometric traits vary over time because of changes in human body. Thus, parameters of a user vary with time.

3. Ideally, the samples belonging to different users should be well separated. However, in the case of weak biometrics, clusters belonging to different users overlap with each other due to low discriminant content.

4. The samples of an individual are not well separated in the original dimension, but get better separated when transformed to a new feature space.

5. The population participating in authentication process changes with time. Let $P_t(\omega_i)$ be the prior probability of participation of the user ω_i at time instant t.

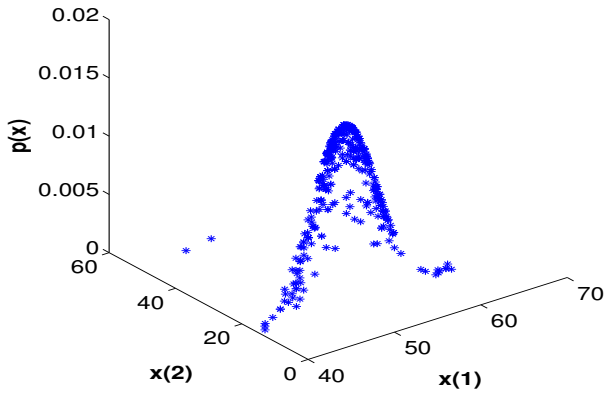


Figure 2.5: Gaussian distribution of samples of weak biometrics.

$$P^t(x/\omega_i) = \sum_{i=1}^{c} P_t(\omega_i) * p(x/\omega_i, \theta_i)$$

where $P^t(x/\omega_i)$ is the probability that the sample x of user ω_i will be generated on t th time instant, say, day.

2.7 Datasets

Our approaches were tested on both real and synthetic data sets. As part of this work we collected the real datasets, including samples from the hand-geometry and keystrokes-based biometrics. These datasets are summarized in Table-2.1.

Dataset	No. Users	No. Samples	Temporal Variation
D1(a)	40	10	No
D1(b)	12	60	Yes
D2(a)	65	10	No
D2(b)	65	30	Yes
D3	150	160	Yes

Table 2.1: Datasets used for various experiments.

Hand-Geometry Dataset (D1): The "raw" 24-dimensional feature vectors extracted from the hand images include lengths of each finger and widths at five equidistant points on each finger.

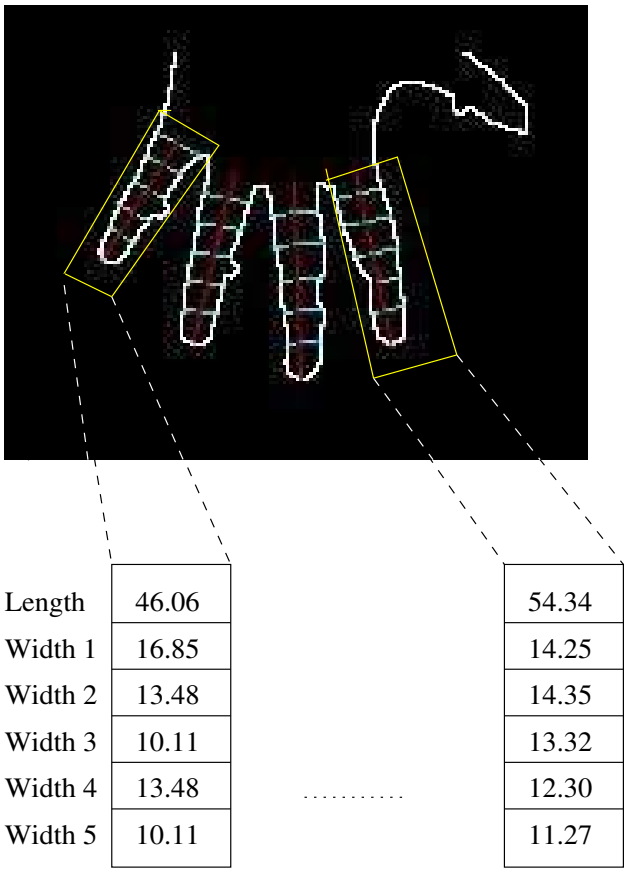


Figure 2.6: Hand-geometry features.

These features are shown in Figure-2.6. These are the lengths and 5 widths values of two fingers, starting from the little finger. Hand images of sixteen different people are shown in Figure-2.7. Two hand-geometry datasets were collected for our experiments: (i) D1(a): This dataset consisted of 10 samples each from 40 users. All the samples were collected on the same day and did not have any temporal variations. (ii) D1(b): This dataset consisted of samples from 12 users over a period of 60 days in order to capture the temporal changes occurring in hand-geometry features over time.

Keystrokes Dataset (D2): The keystroke features consisted of dwell-times (time elapsed between key-down and key-release events of the same key), flight-times (time elapsed between key-release and next key-down), total time taken to type, and number of mistakes. In order to capture the temporal variation in data, we used the French keyboard (with different layout from the popular US-layout keyboard). Users who were not accustomed to French keyboard typed the same sentence 30 times. This allowed us to get the time-varying (learning) samples and the saturated (learned) samples. First 20 samples were used for practice session which constituted time-varying

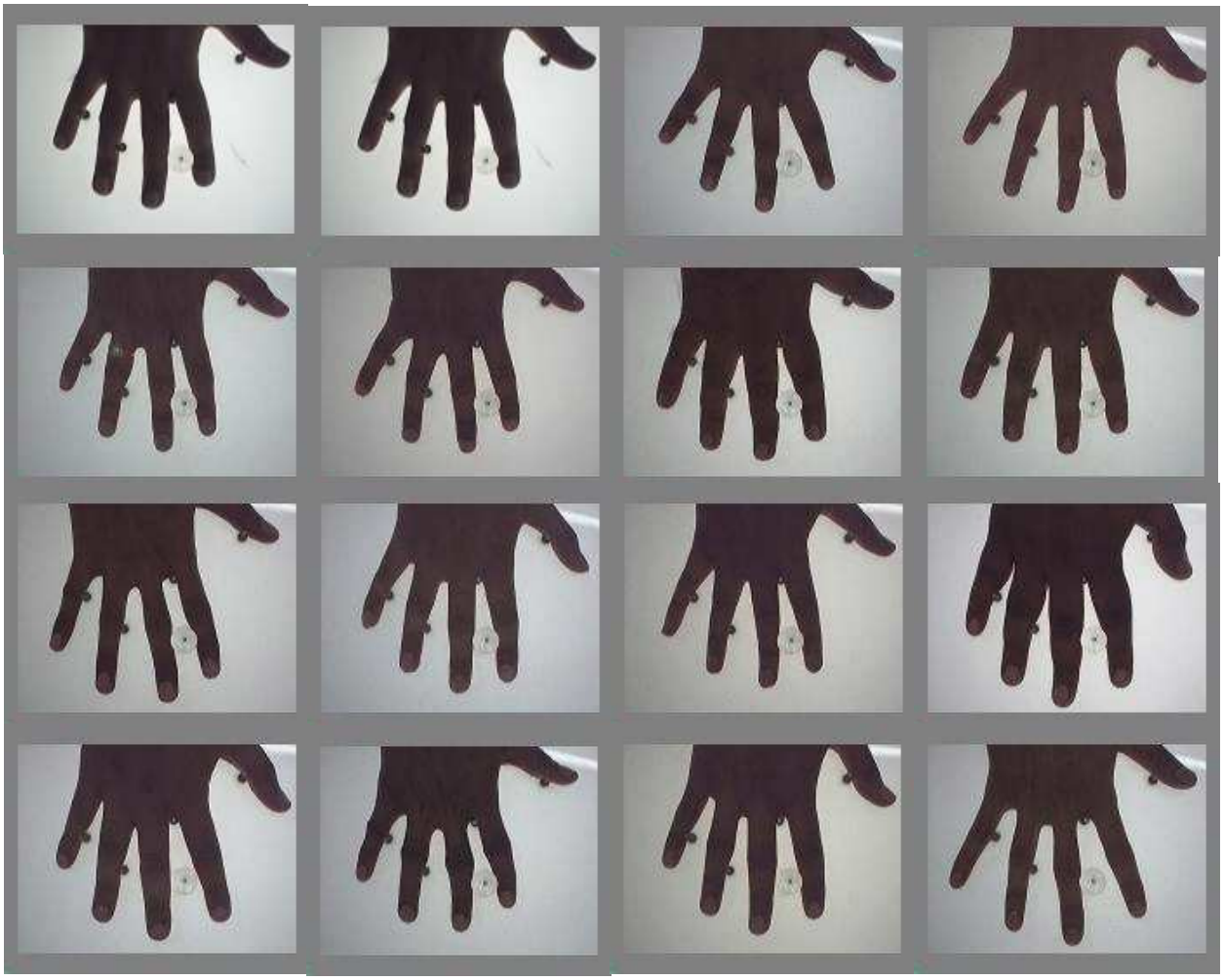


Figure 2.7: Set of hand-geometry images of sixteen different people.

dataset (D2(b)) while the last 10 samples constituted the saturated dataset (D2(a)). A total of 65 users participated in the data collection process, out of which, 46 were accustomed to US-layout keyboard while the rest were new to the computers. A flavour of the raw keystroke-dynamics data for a part of the typed text is shown in Table-2.2. The raw data extracted contains the following information for each entry:

- Character: The character
- Code: ASCII code of character
- Dwell Time: Time elapsed between last key-press and key-release events, in milliseconds.
- Flight Time: Time elapsed between last key-release and key-press events, in milliseconds.

Character	Code	Dwell Time	Flight Time
a	65	141	-
	32	109	1281
q	81	47	1578
u	85	109	391
i	73	125	203
c	67	78	829
k	75	94	531
	32	110	343
b	66	47	1047
r	82	78	328
o	79	94	234
w	87	109	2469
n	78	129	391

Table 2.2: Sample feature values for keystrokes-dynamics.

Synthetic Dataset (D3): Samples were generated for each user by incorporating *apriori* probability using the generative model. Samples were generated for 150 users corresponding to 160 days incorporating the apriori probabilities.

2.8 Summary

In this chapter, we differentiated between strong and weak biometrics. We discussed various low-to-medium security applications of biometrics and the various issues that affect the selection of a

biometric trait for such civilian applications. We argued that weak biometrics are better suited for civilian applications, than strong biometrics. Then, we discussed the issues related to the large-scale weak biometrics. These problems, and their solutions are discussed in detail in the subsequent chapters. The first of these problems, namely, low discriminating content is addressed in the next chapter.

Chapter 3

User-Specific Feature Selection

By definition, weak biometric traits, such as, hand-geometry, face, voice, etc., have low discriminating content. Hence, accuracy of the authentication based on these biometric traits is very low as compared to the other biometrics, like fingerprints, retina, iris, etc. In order to improve performance of the authentication using weak biometrics, it is required to improve the discriminating content of the biometric traits. In this chapter, we discuss a feature selection technique specifically for authentication. Performance of the authentication using the proposed feature selection technique is compared with various other existing discriminant analysis techniques.

In the biometric context, feature selection is very important when the traits used for authentication are very similar to each other. Hence when they are similar, one cannot distinguish between people with high accuracy. Biometric traits unique to every individual show very high accuracy. Jain *et al.* [40] compare various biometric traits based on the accuracy obtained using those traits. The False Accept Rate (FAR) using fingerprints is shown to be less than 0.01%. The FAR using iris-based biometric system is reported to be less than 0.001%. This error rate is reported to be 10% for face, 1.5% for hand-geometry and 3% for voice based biometric systems. However, these biometric traits are being used in a wide variety of day-to-day applications. In this chapter, we discuss various feature selection techniques to improve performance of authentication using weak biometrics. Feature selection scheme transforms the features from original feature space to a new space where the discriminating content is improved, hence, allows the system to distinguish between individuals with better accuracy.

3.1 Related Work

For improving the discriminating content of the features, many Discriminant Analysis technique [41] are popular. In the field of biometrics, discriminant analysis techniques have been extensively used, particularly in the area of face recognition. The Two-Dimensional-Oriented Linear Discriminant

Analysis (2DoLDA) approach was proposed by Muriel Visani *et. al* [42] in which they proposed to apply LDA on face image matrices rather than applying it on vectorized images. 2DoLDA may be implemented in two different ways: Row-oriented LDA (RoLDA) and Column-oriented LDA (CoLDA). Later, they proposed a method to efficiently combine two complementary versions of 2DoLDA, through an iterative algorithm using a generalized bilinear projection based Fisher criterion [43]. They showed a great improvement in performance using their proposed method, while leading to a significant dimensionality reduction. In another work, Zhu and Sung [44] proposed Margin Maximization DA (MMDA) for face recognition. The authors pointed out the problem of unstable performance of face recognition using LDA due to sparse samples. The MMDA technique derives features by maximizing the average margin between the classes. The method does not require the within-class scatter S_w to be non-singular and well-conditioned as it does not involve its inverse term, and the features can directly be derived from the input space. Lu *et al.* [45] pointed out that the distribution of face images, under a perceivable variation in viewpoint, illumination or facial expression, is highly nonlinear and complex. They proposed a kernel machine-based discriminant analysis method, which deals with the nonlinearity of the face patterns' distribution. The proposed method also effectively solves the so-called small sample size(SSS) problem, which exists in most Face Recognition tasks. A recursive non-parametric discriminant analysis technique (RNDA) was proposed by Peng and Quang [46]. Zhang *et al.* [47] proposed Kernerlized Maximum Average Margin Criterion (KMAMC), combining the idea of Support Vector Machine with Kernel Fisher Discriminant Analysis (KFD) for face recognition. A technique to select discriminating features based on strangeness measure (defined as the ratio of the sum of the k nearest distances from the same class to the sum of the k nearest distances from all other classes) was proposed by Li *et al.* [48].

In this chapter, we discuss a novel discriminant analysis technique to improve the discriminating content of weak biometrics, specifically for authentication. This technique iteratively finds the samples which overlap with the samples of the claimed identity and finds the optimal transformation. In the transformed space, the samples of claimed identity are sufficiently away from all the other samples and hence, help to improve accuracy of the overall verification system.

3.2 Popular Discriminant Analysis Techniques

There are various feature selection techniques that can be applied to the problem of biometric authentication. This section describes the discriminant analysis techniques which can be used to improve the discriminative content for better verification in biometric systems.

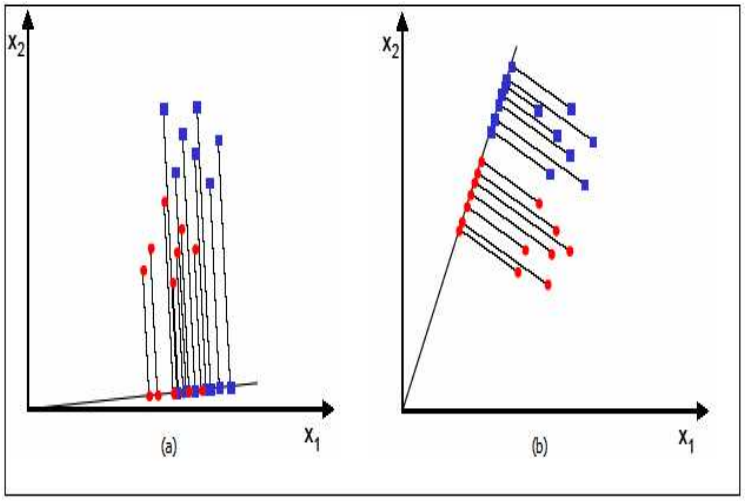


Figure 3.1: Projection of the same set of samples onto two different lines. Figure (a) shows greater separation between the samples of two classes.

3.2.1 Multiple Discriminant Analysis

Multiple discriminant analysis (MDA) is a generalization of Fisher's linear discriminant analysis (FDA) for multiple classes [41]. Fisher linear disciminant is a well known discriminant analysis technique that finds an optimal projection to maximize the distance between the means of the two classes while minimizing the variance within each class. Figure-3.1 illustrates the Fisher Discriminant Analysis technique.

Multiple discriminant analysis seeks to finds a linear transformation matrix that maps the original high dimensional space to a lower dimensional space such that the classes are linearly separable. For a $c-$class problem with $c > 2$, a transformation matrix from a $d-$dimensional feature space to a $m-$dimensional space ($m \leq d$) is determined such that the Fisher criterion of total scatter versus average within-class scatter is maximized. The Fisher criterion function is given by
Maximize

$$J = \frac{\|W^T S_B W\|}{\|W^T S_W W\|} \tag{3.1}$$

with respect to W

S_B and S_W, called the *between-class scatter* and *within-class scatter* respectively, are defined as,

$$S_W = \sum_{i=1}^{c} \sum_{x \in \omega_i} (x - m_i)(x - m_i)^T \tag{3.2}$$

$$S_B = \sum_{i=1}^{c} (m_i - m)(m_i - m)^T \tag{3.3}$$

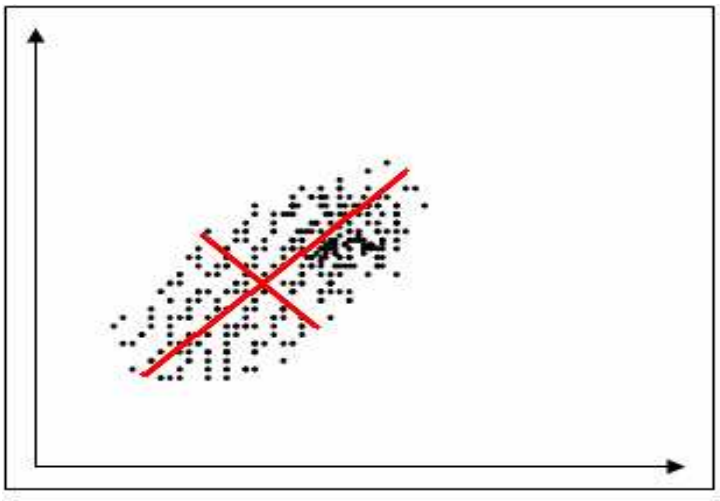


Figure 3.2: A simple two-dimesional dataset, showing the Eigen vectors corresponding to the first two maximum variations in the data.

where x is a sample, m_i is the mean vector of class ω_i, m is the overall mean of all classes. The optimization problem results in finding the Eigenvectors of ${S_W}^{-1}S_B$ corresponding to the m largest Eigenvalues. The transformed features obtained using MDA are shown to be more discriminative than the raw features used for verification.

3.2.2 Principal Component Analysis

Fisher linear discriminant, described in the previous section reduces the dimensionality of the data to improve the discriminating content. Principal Component Analysis (PCA) is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Figure-3.2 shows the Eigenvectors corresponding to the maximum variation in the data. PCA can be used for dimensionality reduction in a data set by retaining those characteristics of the data set that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the "most important" aspects of the data.

The criterion function to reduce the dimensionality from d to $d^{'}$ is given by:
Minimize

$$J = \sum_{k=1}^{n} \| \left(m + \sum_{i=1}^{d^{'}} a_{ki} e_i \right) - x_k \|^2 \tag{3.4}$$

with respect to e_i. The above criterion function is minimized when the vectors $e_1 \ldots, e_{d'}$ are the Eigenvectors of the scatter matrix having the largest Eigenvalues [41]. The scatter matrix is given

by:

$$S = \sum_{k=1}^{n} (x_k - m)(x_k - m)^T. \tag{3.5}$$

In the above equations, n is the total number of samples, m is the mean of all the samples.

3.2.3 PCA-MDA

In biometrics, we are usually posed with small-sample size problem where the number of samples of each user is less than the number of features extracted for the corresponding biometric trait. Thus, it is required to lower the dimensionality of the feature space while increasing the discriminating content of the samples. One of the simplest techniques to achieve this is to apply Principal Component Analysis (PCA) on the samples and then apply MDA to increase the discriminating content of the features. A combination of PCA-MDA technique was earlier applied by Deng and Tsui [49] for appearance based hand-posture recognition. In essence, PCA reduces the dimensionality of feature space by restricting attention to those directions along which the scatter is maximum. Thus PCA can be used to reduce the dimensionality of the feature space and then use MDA (Equation 3.1) to find the features which are most discriminating in the reduced space.

3.2.4 Multiple Exemplar Discriminant Analysis

Kevin and Chellappa [50] proposed Multiple Exemplar Discriminant Analysis (MEDA) to improve the classification results for the problem of face recognition. LDA is a single-exemplar method, in the sense that each class during classification is represented by a single example which is the sample mean of the class. As stated above, a common problem faced by most biometrics based authentication systems is availability of only a small number of samples per user to represent the user. To overcome this drawback, the authors proposed to represent each class using multiple examples. Rather than minimizing the within-class distance while maximizing the between-class distance, MEDA finds the projection directions along which the within-class exemplar distance (i.e. the distances between examples belonging to different classes) is maximized. The criterion function is the same as that used by LDA (Equation 3.1). However, the definitions of S_W and S_B change. Since MEDA uses all the available examples per class, the within-class scatter in LDA becomes the within-class exemplar distance (i.e. the distances between exemplars belonging to the same class).

$$S_W = \sum_{i=1}^{c} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} ({x_j}^i - {x_k}^i)({x_j}^i - {x_k}^i)^T \tag{3.6}$$

The basic element in (Equation 3.6) is a pairwise difference between any two examples belonging to the same class. The space constructed using these basis elements is called the intra-personal space.

Similarly, the between-class scatter in LDA becomes the between-class exemplar scatter (i.e. the distances between examples belonging to different classes):

$$S_B = \sum_{i=1}^{c} \sum_{i=1;j\neq i}^{c} \sum_{k=1}^{N_i} \sum_{k=1}^{N_j} (x_k{}^i - x_l{}^j)(x_k{}^i - x_k{}^j)^T \tag{3.7}$$

This creates the extra-personal space.

3.3 Kernel Machines

Kernal Machines approach the problem of pattern classification by mapping the data into a high dimensional feature space, where each coordinate corresponds to one feature of the data items, transforming the data into a set of points in a euclidean space. In that space, a variety of methods can be used to find relations in the data. Since the mapping can be quite general (not necessarily linear, for example), the relations found in this way are accordingly very general.

The kernel machines provide an elegant way of designing nonlinear algorithms by reducing them to linear ones in some high-dimensional feature space $\mathbb{F}$ nonlinearly related to the input sample space $\mathbb{R}^{\mathbb{J}}$:

$$\phi = \mathbf{z} \in \mathbb{R}^{\mathbb{J}} \rightarrow \phi(\mathbf{z}) \in \mathbb{F} \tag{3.8}$$

The idea can be illustrated by a toy example depicted in Figure-3.3, where two-dimensional input samples, say $z = [z_1, z_2]$, are mapped to a three-dimensional feature space through a nonlinear transform: $\phi(\mathbf{z}) = [z_1, z_2] \rightarrow \phi(\mathbf{z}) = [x_1, x_2, x_3] := [z_1^2, \sqrt{2}z_1z2, z_2^2]$ [1].

It can be seen from Figure-3.3 that in the sample space, a nonlinear ellipsoidal decision boundary is needed to separate classes A and B, in contrast with this, the two classes become linearly separable in the higher-dimensional feature space. However, the dimensionality of the feature space $\mathbb{F}$ could be arbitrarily large, possibly infinite. Fortunately, the exact $\phi(\mathbf{z})$ is not needed and the feature space can become implicit by using kernal machines. The trick behind this method is to replace the dot products in $\mathbb{F}$ with a kernal function in the input space $\mathbb{R}^{\mathbb{J}}$ so that the nonlinear mapping is performed implicitly in $\mathbb{R}^{\mathbb{J}}$. Thus, the central issue to generalize a linear learning algorithm to its kernel version is to reformulate all the computations of the algorithm in the feature space in the form of dot product.

Algorithms capable of operating with kernels include Support Vector Machines (SVM), Fisher's linear discriminant analysis (LDA), principal components analysis (PCA), canonical correlation analysis, ridge regression, spectral clustering, and many others.

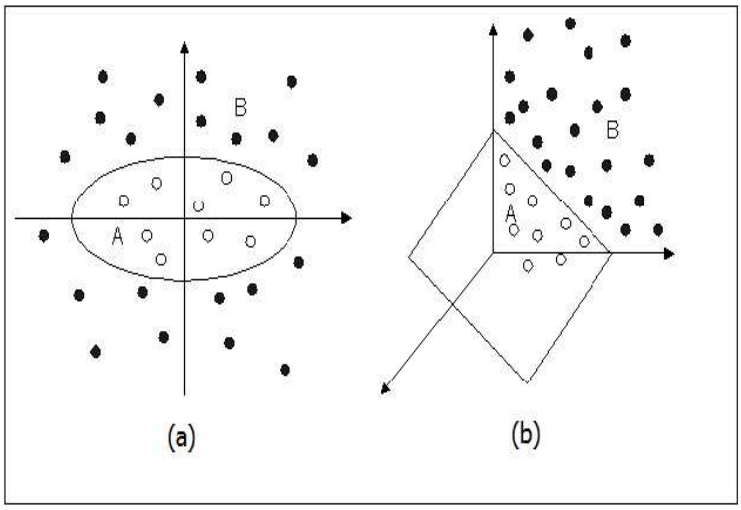


Figure 3.3: A toy example of two-class pattern classification problem [1]. (a) Samples lie in the 2-D input space, where it needs a nonlinear ellipsoidal decision boundary to separate classes A and B. (b): Samples are mapped to a 3-D feature space, where a linear hyperplane can separate the two classes.

3.4 Single Class Hierarchical Discriminant Analysis

In this section, we discuss in detail the discriminant analysis technique for authentication. Our technique is built on the top of two existing discriminant analysis techniques: Single Class Discriminant Analysis and Hierarchical Discriminant Analysis. Each of these techniques is discussed in detail in the following subsections.

3.4.1 Single Class Discriminant Analysis

Verification problem can be best posed as a single-class classification problem. Formally, the single-class classification problem or biased classification problem is defined as the learning problem in which there are an unknown number of classes but the system is only interested in one class. Samples of the claimed identity are termed "positive", while all the other samples are termed "negative". Similarly, in case of authentication, the emphasis is on the samples of the claimed identity only. Single Class Discriminant Analysis (e.g., Biased Discriminant Analysis (BDA)) [51] fits well into the problem of authentication.

It is required to transform the features into a new space such that the discriminative power of the raw features of each user is enhanced. However, the transformation is required to be such that the feature vectors of the claimed user get well separated from all the other feature vectors in the transformed space. In other words, the discriminant should be biased towards the claimed identity.

In transformed space, the feature vectors from the claimed identity are required to get clustered closely while those from the other classes are pushed apart from the features of the claimed identity and hence enhance the performance of the verification algorithm.

The Biased Discriminant finds an optimal transformation such that the ratio of "the negative scatter with respect to the positive centroid" over the "positive with-in class scatter" is maximized.

The biased criterion function is defined as:

Maximize

$$J = \frac{\|W^T S_y W\|}{\|W^T S_x W\|}$$

with respect to W.

Let the training set contains N_x positive and N_y negative samples. Then S_x and S_y are defined as,

$$S_x = \sum_{i=1}^{N_x} (x_i - m_x)(x_i - m_x)^T \tag{3.9}$$

$$S_y = \sum_{i=1}^{N_y} (y_i - m_x)(y_i - m_x)^T \tag{3.10}$$

where x_i denote the positive samples, y_i denote the negative samples, $m_x = \frac{1}{N_x} \sum_{i=1}^{N_x} x_i$ is the mean vector of the positive samples, and W can be computed from the Eigenvectors of $S_x^{-1} S_y$.

Biased Discriminant Analysis works by first minimizing the variance of the positive samples, and then maximizing the distance between the centroid of the positive samples and all the negative samples. In essence, BDA finds the discriminating subspace in which the positive samples are 'pulled' closer to one another while the negative samples are 'pushed' away from the positive ones.

3.4.2 Hierarchical Discriminant Analysis

Hierarchical Discriminant Analysis (HDA)[52] was proposed by Yuichi *et. al* for texture classification. Linear discriminant analysis is very effective for two-class classification problems. However, when extended to multi-class classification problem, the precision of discrimination deteriorates. This is because of the occurrence of overlapped distributions on a discriminant space built by the Fisher criterion. Fisher criterion calculates the between-class scatter by calculating the distance between the overall mean computed from all samples and a mean of each class. All interclass distances are not taken into consideration, and then effective projection space may be lost about each class.

Initially, all classes are assumed to be part of a single cluster on the initial discriminant space built by the Fisher method. The single discriminant space is divided by grouping the overlapped classes recursively. All samples of two classes are projected on the straight line which connects the

center of each class. If the samples which belong to a different class overlap on the straight line, the two classes are considered to belong the same group. Two classes are considered as the same group even if there exists some samples in one class whose distance from its class center is longer than that from the other class center.

3.4.3 Single Class Hierarchical Discriminant Analysis (SCHDA)

With Biased Discriminant Analysis, main cause of errors during authentication is the similarity of some negative samples with the positive samples in the discriminant space built by BDA. This corrupts the computed discriminants and hence, increases the errors. This problem can be addressed by identifying the negative samples close to the positive cluster and then applying BDA again on this subset of samples. The procedure is continued until the positive samples are sufficiently far away from the negative samples or only two samples are left in the group. Figure-3.4 gives more insight into feature selection using the SCHDA technique.

Single Class Hierarchical Discriminant Analysis works in the following manner: Initially samples of classes are assumed to be in a single group and the BDA criterion is applied to this group. Next, the negative samples which are very close to the positive cluster in the BDA discriminant space are identified. This identification is performed on the basis of two thresholds: k and n, where k is the minimum distance required between well-separated negative samples from the positive cluster and n is the minimum number of samples required from a negative class to be included into the group. Then, the samples from these negative classes along with the positive samples form the reduced group. All the other samples are discarded. BDA criterion is again applied on the reduced group. This procedure is repeated until the following termination conditions become true.

1. All the negative samples are at a distance of at least k from the positive cluster.
2. Only two classes are left out in the reduced group.
3. There is no change in groups obtained in successive iterations.

3.5 The SCHDA Algorithm for Authentication

Any biometric system works in two phases: (i) Training, and (ii) Verification. Each of these phases are discussed in the following subsections.

Training: Training of the system requires finding the optimal discriminating space for each user, using the SCHDA method. During the training phase, samples of each user are fed as input to the training algorithm (SCHDA) and optimal weight matrix is computed and stored. The SCHDA algorithm for training for hand-geometry based verification is presented in Algorithm-1.

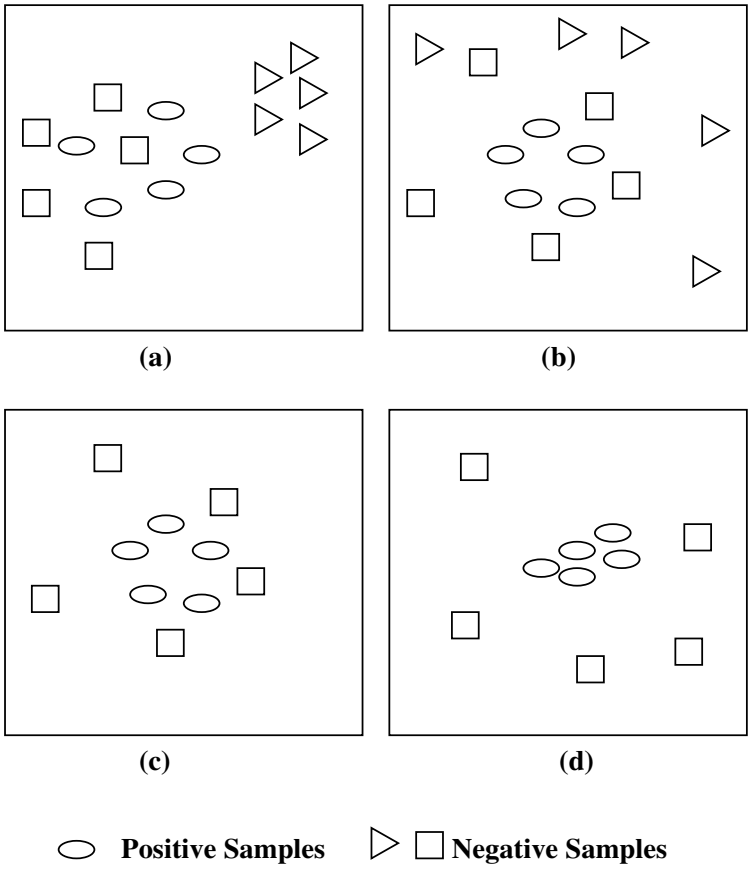


Figure 3.4: Scatter of positive and negative samples (a) With "raw" features (b) After applying BDA (c) The reduced group (d) After applying SCHDA on reduced group. The negative samples very close the the positive cluster in feature space built by BDA (shown in (b)) get well-separated from the positive cluster after applying SCHDA on the reduced group (shown in (c)).

Algorithm 1 Training using SCHDA

1: Data Set : $S = s_i, i = 1..N,$
   where N is the number of samples and s_i is a $d \times 1$ vector.
2: **for** $k = 1..c, c$ is number of users **do**
3: Label all samples in S from user k as positive and rest as negative:
   $X = x_i, i = 1..N_x, N_x$ is the number of positive samples.
   $Y = y_i, i = 1..N_y, N_y$ is the number of negative samples.
4: Calculate mean vector of all the positive samples: m_x.
5: Calculate the scatter matrices: S_x, S_y as per Equations-3.9 and 3.10.
6: Calculate W_k as a $(d \times d^{'})$ matrix whose columns are the Eigen vectors of ${S_x}^{-1}S_y$, where $d^{'}$ is the number of non-zero Eigen values.
7: Apply W_k on S and identify the users close to samples of the positive user. Form a group $s \subset S$ of samples of these classes based on the thresholds k and n.
8: Repeat the above steps with $S = s$ until either of the following become false:
   All the negative samples are at a distance of at least k from the positive cluster.
   Only two classes are left out in the reduced group.
   Same groups are obtained in successive iterations.
9: Store the W_k after performing all the above steps for user k.
10: **end for**

Verification: During the verification phase, the system is presented a new sample v with the claimed identity, l. It is required to verify the claimed identity of the sample using the SCHDA technique. The verification algorithm is presented in Algorithm-2.

Algorithm 2 Authentication using SCHDA

1: Retrieve W_l corresponding to the user l.

2: Apply the transformation to all the samples in the data set and to v:
$s_i{}' = {W_l}^T s_i, i = 1..N$
$v' = {W_l}^T v$

3: Apply k-nearest neighbor algorithm to verify the claimed identity, l

The verification using the transformed features does not add to the computational complexity of the system as the training is done offline. During verification phase only multiplication of the samples with the weight matrix adds to the complexity.

The SCHDA allows to create an optimal discriminant space for each individual, separating well the samples of claimed identity from all the other samples. This feature leads to an improvement in authentication due to increased discrimination of the features.

3.6 Results

Experiments were conducted to compare the performance of authentication using "raw", BDA and SCHDA features. These experiments were carried out on the stationary datasets D1(a) and D2(a). The error rates were observed to reduce considerably when BDA and SCHDA features were used for authentication. The parameters k and n were set empirically to the values 4.5 and 5 respectively. The FAR and FRR of the system using SCHDA-transformed features was observed to be 0.0085 and 0.0769 respectively. Table-3.1 shows the improvement in verification achieved by using BDA and SCHDA features over "raw" features on the dataset D1(a). The results on the dataset D2(a) are shown in Table-3.2. Considerable improvement in performance is observed when features selected by SCHDA method were used for authentication.

	Raw Features	BDA Features	SCHDA Features
FRR	0.204	0.136	0.0769
FAR	0.304	0.017	0.0085

Table 3.1: Improvement in performance of verification using SCHDA features over "raw" and BDA features on dataset D1(a).

We compared the error rates using the proposed framework with the performance using other discriminant analysis techniques earlier discussed in Section-3.2 (Table-3.3). It can be observed that

	Raw Features	BDA Features	SCHDA Features
FRR	24.6	15.6	11.08
FAR	17.9	10.7	8.92

Table 3.2: Improvement in performance of verification using SCHDA features over "raw" and BDA features on dataset D2(a).

the SCHDA performed better than the other discriminant analysis techniques discussed earlier on our dataset. It can be observed that the EER of the system was highest using PCA-MDA features. This is due to loss of discriminating information after performing PCA on the raw features. The BDA-transformed features showed low FAR and FRR values as compared to MDA, MEDA and HDA. This shows that BDA is better suited for verification than other DA techniques. The error rates were found to be lowest with the proposed SCHDA technique.

	MDA	PCA-MDA	MEDA	HDA	BDA	SCHDA
FAR	0.0418	0.0769	0.034	0.0256	0.017	0.0085
FRR	0.1396	0.1496	0.119	0.0769	0.136	0.0769
EER	0.512	0.8547	0.46	0.393	0.389	0.283

Table 3.3: Comparison of error rates of the verification system using various discriminant analysis techniques. The error rates using SCHDA features are observed to be lowest.

Comparison based on Receiver Operating Characteristic: In addition to FRR, FAR and EER, Receiver Operating Characteristic (ROC) curve of the system is also shown in the Figure-3.5. ROC of a biometric system is a graphical depiction of the relationship between the FRR and FAR as a function of the threshold's value. The ROCs obtained using HDA, BDA and HDA features are shown in the figure. The closer the ROC is to the axes, the better the feature selection technique. It can be seen that the ROC curve corresponding to SCHDA is closest to the axes while those corresponding to BDA and HDA are farther from the axes.

Effect of changing dimensions: We observed the effect of changing the dimensionality of the transformed feature space on FRR and FAR of the system. We tested the percentage accuracy of the system with 1, 5 , 10, 15, 20, 24 dimensions. The performance of the system was not observed be very high with 1 and 24 dimensions. This is because only one dimension was not appropriate to discriminate between the features. Highest performance was observed with 10 dimensions. However, after that, the performance of the system was observed to decline with increasing dimension. This is because the discriminating information is present in the first few dimensions corresponding to

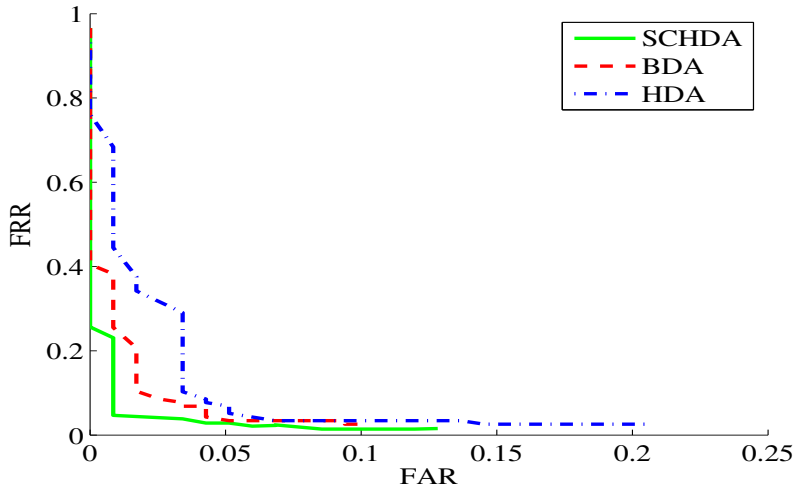


Figure 3.5: ROC curve of authentication with various DA-transformed features. The closer the ROC curve to the origin, the better the feature selection technique.

the larger Eigen values. Thus, as we increase the dimensions, the lower discriminating components also contribute to the distance between the test sample and the training samples and hence the error rate increases. The results are summarized in the Table-3.4.

Dimensions	1	5	10	15	20	24
Performance	82.15%	92.31%	96.58%	84.62%	77.78%	64.10%

Table 3.4: Effect of changing number of dimensions on performance of the system.

Effect of changing number of users: We conducted experiments to study the effect of number of users on the performance of the system. With BDA-transformed features, the performance of the system was observed to decline with increasing number of users. However, with HDA and SCHDA, increasing number of users did not have much impact on the performance of the system. The results are summarized in Table-3.5.

3.7 Summary

A novel technique to select discriminating features from the raw-biometric features for authentication is proposed. The performance of the system was observed to improve considerably using the SCHDA features. We compared performance of authentication using this technique with the popular discriminant analysis techniques. Through various experiments, we showed that this technique

	5	10	15	20
SCHDA	100.00%	96.67%	95.56%	93.33%
HDA	100.00%	93.10%	88.30%	88.33%
BDA	100.00%	93.33%	91.11%	90.00%
	25	30	35	40
SCHDA	92.00%	93.33%	93.33%	92.31%
HDA	89.33%	88.67%	87.29%	87.03%
BDA	88.00%	85.56%	87.62%	86.32%

Table 3.5: Effect of number of users on performance of the system. The performance using BDA decreases with increasing number of users, while the number of users does not have much impact when HDA and SCHDA features are used.

can be used to improve the performance of the biometric-based authentication system. However, the features selected using techniques described in this chapter do not adapt to the changing features of weak biometrics. In the next chapter, we discuss an incremental feature selection technique to adapt to the changing features.

Chapter 4

Incremental Feature Selection

In order to improve performance of the authentication, we need to address the second inherent problem discussed in Chapter 2 with weak biometrics, *stability* or *permanence*. For example, at birth, human hands are nearly symmetrical. As the body ages, the hands change due to natural and environmental changes. Most people become either right or left handed, causing one hand to be slightly larger than the other. The frequently used hand tends to be more susceptible to injury. Young peoples' hands change rapidly as they mature. Older peoples' hands change with the natural aging process, gain or loss of weight, or arthritis. All these factors necessitate that practical hand geometry based authentication systems learn minor hand shape changes and continually update the templates or features as users are verified by the system. Similar variations are observed with other weak biometrics such as face, voice, etc.

4.1 Related Work

Ramanathan *et al* [53], studied the similarity of faces as a function of time. They introduced the notion of *PointFive Faces* - the better illuminated half of a frontal face extracted assuming bilateral symmetry. This helped to overcome the non-uniform illumination across face images. They proposed a Bayesian age-difference classifier that is built on a probabilistic Eigenspaces framework. Later, in [54], they proposed a craniofacial growth model which takes into account anthropometric evidences collected on facial growth. This model characterized growth related shape variations observed in human faces during formative years and is in accordance with the observed growth patterns in human faces across years. They also demonstrated the applicability of their model to predict one's appearance across years and to perform face recognition across age progression.

In Chapter-3, we discussed various feature selection techniques which can be used to improve discriminating content of weak biometrics. However, a major drawback of these techniques is that they are "stationary", in the sense that once selected, the features do not change to adapt to the

changes occurring in the biometric characteristic. This may lead to a major decline in the performance of the system as the features vary over time.

In this chapter we propose an adaptive feature selection framework based on simple statistical techniques. We propose to address the poor stability of weak biometrics by incrementally updating parameters of feature selection framework. Since the weak biometric systems can operate in co-operative mode, users' feedback and knowledge acquired from history can be used for adaptation. Jain *et. al.* [55], proposed a method to learn the genuine and impostor probability distributions of each of the users and hence compute a different threshold as against a common threshold for all the users. This method achieves much higher classification accuracies with user dependent thresholds. A practical limitation of their technique is that large number of samples are needed to train the system. The proposed method overcomes this difficulty as learning takes place over time with new samples by online adaptation. A technique along the lines of the work done by Hall and Martin [56] is discussed in this chapter to adapt to the changing features. We illustrate the updation method by incrementally updating the parameters of BDA.

4.2 Incremental Feature Selection

In this section, we present a scheme to incrementally update the feature vectors of a user in order to improve the performance of the system over time.

With BDA, any user can be characterized by the discriminant Eigenspace model:

$$\Omega = (S_x, S_y, m_x, m_y, N_x, N_y) \tag{4.1}$$

Let X denote the set of positive samples and Y the set of negative samples. Let x_{new} be the incorrectly verified sample of the user. The problem is that of estimating an updated model for all the users using the new sample and the current model. That is, finding

$$\Omega^{'} = ({S_x}^{'}, {S_y}^{'}, {m_x}^{'}, {m_y}^{'}, {N_x}^{'}, {N_y}^{'}) \tag{4.2}$$

using only Ω and x_{new}. Thus, on every incorrect verification, $\Omega^{'}$ is calculated and used for the subsequent verifications. The updation is carried out in two steps:

Step(i): Update the Eigenspace for the positive class, i.e., $x_{new} \in X$
In this case, the set of negative samples remains unchanged. Thus, the number of negative samples and hence the mean of negative samples remain the same. As the new sample belongs to the positive class, number of positive samples increases by 1. i.e.,

$${N_x}^{'} = N_x + 1 \tag{4.3}$$

The updated mean of positive samples becomes

$$m_x{}' = \frac{N_x m_x + x_{new}}{N_x + 1} \tag{4.4}$$

Updated within class scatter of the positive samples $S_x{}'$ will be

$$\sum_{i=1}^{N_x}(x_i - m_x')(x_i - m_x')^T + (x_{new} - m_x')(x_{new} - m_x')^T \tag{4.5}$$

Substituting for updated mean, $m_x{}'$ and using the definition of S_x,

$$S_x{}' = S_x + \frac{(x_{new} - m_x)(x_{new} - m_x)^T}{N_x + 1} \tag{4.6}$$

Updated scatter of negative samples with respect to the positive mean becomes

$$S_y{}' = \sum_{i=1}^{N_y}(y_i - m_x{}')(y_i - m_x{}')^T \tag{4.7}$$

Substituting for the expression for updated mean, $m_x{}'$ and with some simplifications, $S_y{}'$ becomes,

$$S_y - \frac{N_y}{N_x + 1}[(m_y - m_x)(x_{new} - m_x)^T + (x_{new} - m_x)(m_y - m_x)^T] + \frac{N_y}{(N_x + 1)^2}(x_{new} - m_x)(x_{new} - m_x)^T \tag{4.8}$$

Step (ii): Update the Eigenspaces of all the other classes, i.e. $x_{new} \in Y$
Number of positive samples, mean of positive samples and hence the within class scatter of the positive samples remain unchanged. Number of negative samples increases by 1. i.e.,

$$N_y{}' = N_y + 1 \tag{4.9}$$

And updated mean of negative samples is given by:

$$m_y{}' = \frac{N_y m_y + x_{new}}{N_y + 1} \tag{4.10}$$

Updated scatter of negative samples with respect to the positive mean:

$$S_y{}' = \sum_{i=1}^{N_y}(y_i - m_x)(y_i - m_x)^T + (x_{new} - m_x)(x_{new} - m_x)^T \tag{4.11}$$

Rewriting in terms of S_y,

$$S_y{}' = S_y + (x_{new} - m_x)(x_{new} - m_x)^T \tag{4.12}$$

These incremental expressions are used for calculation of the optimal discriminative features at every stage of adaptation and learn the changes in features of each user. As the changes in features are learned incrementally, this technique helps to improve the poor stability of weak biometric traits.

4.3 Training Verification and Learning

The verification using the IBDA-transformed features does not add to the computational complexity of the system as the training and learning is done offline. During the verification phase only multiplication of the samples with the weight matrix adds to the complexity. All the samples incorrectly verified are employed for the updation of the discriminant Eigenspace models. These Eigenspace models for all the users are calculated during learning phase. This model updation does not require explicit storage of large quantity of labeled examples.

The feature vectors from each of the users is obtained and stored in the database. During training phase, samples from each of the users is fed as input to the training algorithm (BDA) and optimal weight matrix and the discriminant Eigenspace model Ω for each user is computed and stored.

During verification phase, the system is presented a new feature vector with the claimed identity. The discriminant Eigenspace model for the claimed user is left unchanged if the user is verified correctly. Otherwise the sample is marked as incorrectly verified and it is stored and later used during the next learning phase for updation of the discriminant Eigenspace model and optimal weight matrix for the claimed user.

During learning phase, the incorrectly verified samples are used to update the corresponding models and the updated models are stored for further testing and subsequent learning phases. Figure-4.1 shows the procedure.

4.4 Results and Discussions

In order to show the significance of incremental feature selection framework on time-varying features, we conducted an experiment to compare the percentage accuracy of authentication using static feature selection and incremental feature selection. The experiment was conducted on the time-varying datasets D1(b) and D2(b). Figure-4.2(a) shows that on dataset-D1(b), initially performance with both statically and incrementally selected features was observed to be almost equal as the features selected did not change. During subsequent verification phases, performance using incrementally selected features improved as the system learned from the errors over time.

The results on the dataset-D2(b) (Figure-4.2(b)) give more insight into the incremental updation framework. As the data was collected from people who were not accustomed to the French keyboard, the initial samples were very similar, resulting in low performance of the authentication system. Thus, with BDA-features, a rapid decline in performance was observed, as the features

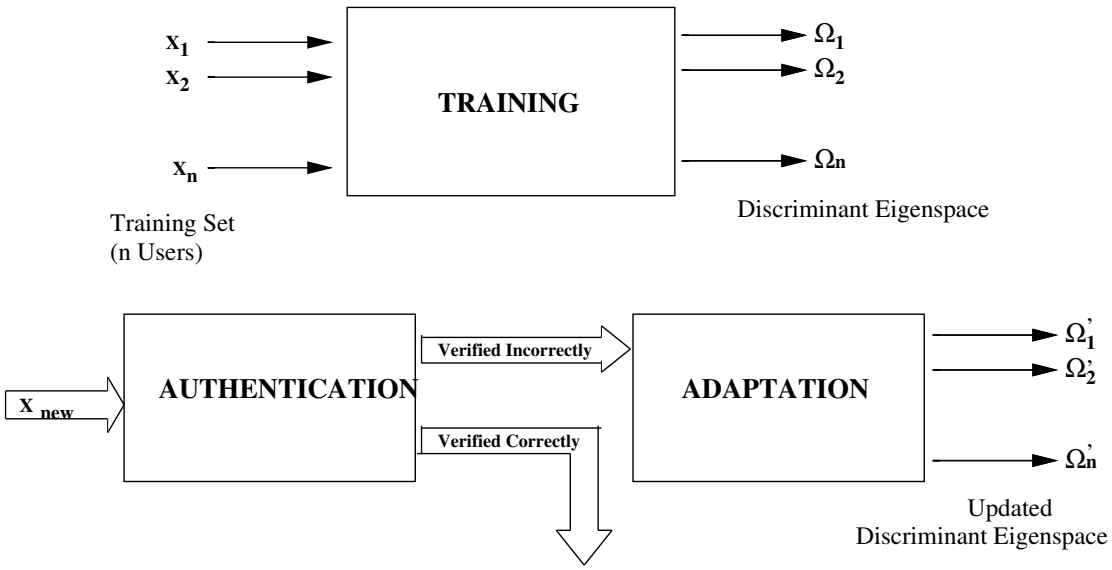


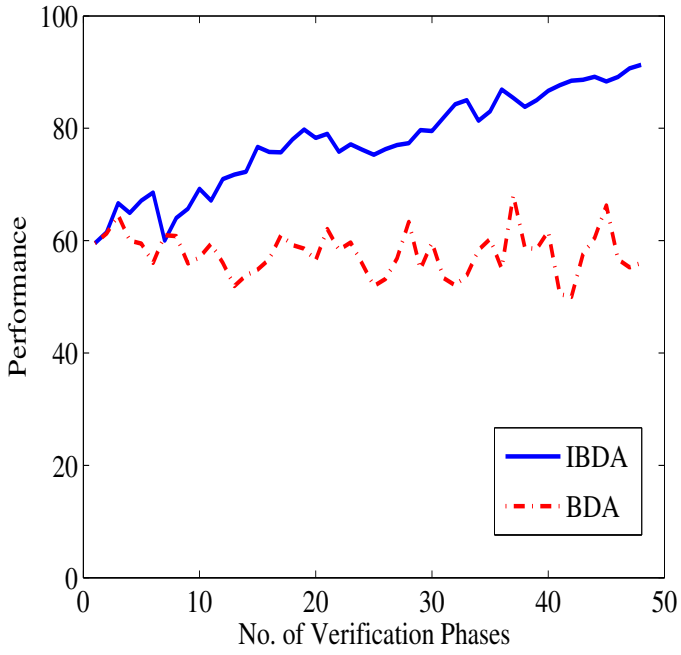
Figure 4.1: The training, verification and adaptation phases using incremental feature selection.

were changing very fast. As the users started getting familiar to the keyboard learning took place and improvement in performance was observed over time. However, a saturation phase was reached when users adapted to the keyboard and hence the features became stationary, showing stationary performance towards the end.

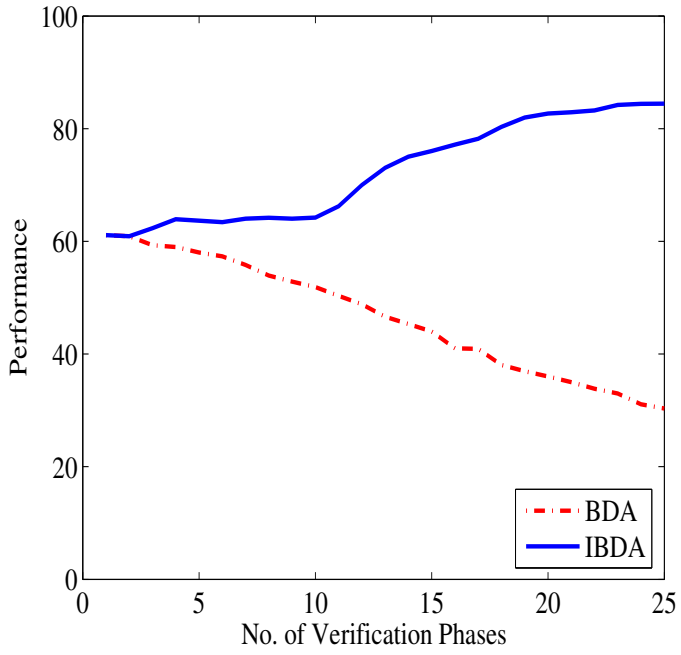
These experiments show that incremental feature selection improves the performance of the system significantly while the traditional feature selection shows degradation in performance over time due to low stability of the features.

4.5 Summary

In this chapter, we gave a statistical solution to the poor stability of features for weak biometrics. An incremental feature selection framework was proposed to adapt the feature selection framework to the changing features over time. This helped to improve the performance of the system even though the traits of the users changed. We showed the incremental feature selection framework by updating the parameters of the Biased Discriminant Analysis and hence, came up with a new technique, called Incremental Biased Discriminant Analysis (IBDA). We compared the performance of authentication using BDA and IBDA over time. The performance was observed to improve considerably when IBDA was used as the feature selection mechanism.



(a)



(b)

Figure 4.2: Performance over time using statically and incrementally selected features on datasets D1(b) and D2(b). The performance is observed to improve over time using the incrementally selected features over the statically selected features.

Chapter 5

Modeling Time-Varying Population for Civilian Applications

Population size plays a major role in determining the performance of any biometric authentication system, particularly when such systems are used for civilian applications. The civilian applications are usually used over a very long period of time and the system keeps enrolling new users. This results in a huge population enrolled into the system over time. It is a well known fact that performance of any pattern-recognition system is inversely proportional to the number of classes. Hence over time, as the system keeps adding new users, the performance of the system degrades. It is, thus, required to improve or maintain the performance of the system with the increasing number of users over time. In this chapter, we throw light on various scenarios in which the regularly participating population is relatively less than the total enrolled population. The variation in participation of users can be modeled and performance of authentication can be optimized over the regularly participating population.

5.1 Related Work

Biometrics is a special class of pattern-recognition system in which the users get added to the system over time. Traditionally, it requires to train the system with every new user enrolled into the system. Also, as the number of classes increases, performance of the system decreases. Thus, over time system shows degradation in performance. Vasconcelos *et al.* [57] proposed a feature selection algorithm that combines information theoretic feature selection and minimum Bayes error solutions to select discriminating features for object recognition problems involving large number of classes. Teddy Ko [58] proposed to use multimodal biometrics by fusion of fingerprint, face and iris based biometrics to improve recognition accuracy for large user population. In another work, Ram *et al.* [59] tried to solve the Large-scale Biometric Pattern (LBP) recognition problem by using

Multi-agent system for LBP (MLBP). The authors proposed to use multi-agent system solution to this problem because of the need to distribute the problem and parallelize it among multiple computational units.

Most of the existing works deal with the usage of strong biometrics and the use of weak biometrics for civilian applications has not been explored in much detail.

5.2 Modeling Time-Varying Population

The variation in participating population in a biometric system can be structured or un-structured. For e.g., users can go on leave for a particular duration (un-structured) while a group can visit the company every year for some time-period (structured). It is required to incorporate the prior information of participation of each individual, so that authentication can be performed only on the regularly participating population. In order to optimize the performance over the participating population, we propose to incorporate the time-varying parameter, i.e., the probability of participation $P_t(\omega_i)$ of each user ω_i on a particular day into the conventional framework for feature selection. Techniques which can be used to compute $P(\omega_i)$ are described in following sections. An intelligent framework is needed, which can select and adapt to the changing features and optimize the performance of authentication over the regularly participating population only. In this section, we propose to model the time-varying population using the Markov models.

Markov model is suitable for modeling the variation in participating population as it is based on finding the probability of observing a sequence of participation of each user. Structured variation in population causes a repetition of sequence of participation over a period of time. Hence, knowing the period of repetition, a HMM can be trained on the sequence and the prior probability of each individual can be estimated. In case of un-structured variation in population, there is no repetition of sequence involved and the variation in pattern of participation of each individual is for a short period of time. For such a sequence, Markov chain model is better suited because of its low computational complexity. Moreover, the use of HMM for modeling such un-structured sequence will require training of the HMM for each user every day. Each of these situations and their solutions are discussed in detail in the following subsections.

5.2.1 Un-Structured Variation in Population

Consider an authentication system being used for physical access to an organization. Some employees may be absent for various reasons, like on-leave, sick, etc. This variation in pattern of participation of every employee is usually un-structured. That is, the employee is absent for usually a short duration of time and then comes regularly later on. It is thus required to calculate the

prior probability of each employee using the information obtained from recent history.

Participation of every individual can be considered as a random variable whose probability on a particular day depends upon his history of participation. The simplest way to model this variation is by Markov chain [41] in which the probability of each element in the sequence depends upon the value of the previous element. Every individual can be in one of the two states: "0" corresponding to absence and "1" corresponding to presence on any particular day. Assuming that participation of an individual depends upon his presence or absence on the previous day, we need to find the prior probability of his participation on the current day given the state he was in on the immediately preceding day.

In order to estimate the prior probability of participation or presence of an individual, we need to find the probability of going to state "1" from the last state which can be either "0" or "1". Mathematically, the prior probability of participation (P_{01} or P_{11}) of an individual can be calculated based on frequency of transitions from last state to state "1" over frequency of transitions from last state to both the states.

$$P_t{}^j(\omega_i) = \frac{N_{j1}}{N_{j1} + N_{j0}}$$

Here, $P_t{}^j$ is the probability of going to state "1" given that the last state was j. N_{j1} is the number of transitions from state j to state "1". This probability is calculated over a window of past M days. Since we are calculating the prior probability of participation based on the state on the previous day, Markov chains are better suited to model the un-structured variation in population.

5.2.2 Structured Variation in Population

Markov chain based technique allows us to predict the participation of a user based on his previous day's participation only. This technique works well for the case when the change in probability of participation of each user is un-structured. In that case, the model accounts for the participating population changing for a short period of time. Consider a situation in which the variation in participating population is structured. For example, assume a huge batch of students divided into multiple groups. Each group attends a particular class on fixed days of a week. Thus, the students participate in authentication process periodically or the variation in participating population is structured. In this case, this variation is structured over a long period of time. Hence, a technique is needed which can exploit the prior information of participating population of each individual on each day of the period and optimize performance based on this information.

For this situation, Hidden Markov Model (HMM) [41] can be used to predict the sequence of states in which a user can be present over a period of time. The Hidden Markov Model is a finite

set of states in which transitions among the states are determined by the transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. Only the outcome, not the state, is visible to an external observer. In an authentication system, there could be various reasons of absence of a user (e.g., working in shifts, belonging to different batch, etc.). The model for varying population should be able to account for these reasons while predicting his presence or absence on any particular day. Thus, the hidden states in case of our problem are the status of the individual and the outcome or observation are symbols "1" for presence and "0" for absence. This is because only whether the person is participating or not is visible to the system but not the reason or the status of the person.

Hidden Markov Model can be applied in the following manner to our problem of estimating prior probability of participation of each individual:

1. Hidden States: An individual can be in one of the states, which determine if he will be able to participate in the authentication process. Let the hidden states be denoted by $H = h_j, j = 1 \ldots M$, assuming M possible hidden states.

2. Visible Symbols: In our case, there are two possible visible symbols: "0"(absence) or "1"(presence). Let $V = v_k$, where $k = 0, 1$

As the variation in participating population is structured, we need to find the probability of participation of each individual on every day during the period. That is,

$$P_T(\omega_i) = p_t(\omega_i), t = 1 \ldots T$$

$P_T(\omega_i)$ is the time varying series of probability of participation of the user ω_i and T is the periodicity (a week or fortnight or a month, etc.). Each user enrolled in the system has a corresponding HMM for period T. It is required to train the HMM based on the sequence of participation observed over past T period of time. Training the HMM returns two probability matrices : $A = a_{ij}$ and $B = b_{jk}$, where a_{ij} denotes the probability of transition from h_i to h_j hidden state and b_{jk} is the probability of emitting v_k given that the system is in h_j state. Input for training is the sequence $V^T(\omega_i) = v_{kt}(\omega_i)$, where ω_i is the ith user and $v_{kt} \in V$ on t^{th} day. It is required to find the probability that the user will be present (i.e, $v_{kt} = 1$) on tth day given that he is in state h_j. That is,

$$P_t(\omega_i) = P(v_{kt} = 1 | h_{jt}) = b_{kj}$$

Now, h_{jt} can be determined as a solution of the decoding problem of HMM which is defined as given the sequence of visible symbols, find the most probable sequence of states which produced the given symbols.

5.3 Experimentation and Results

Various experiments were conducted to show that modeling the variation in participating population helps to optimize the performance of authentication. Both real and synthetic datasets were used for experiments. Incremental Biased Discriminant Analysis (IBDA), discussed in Chapter-4 was used to select the discriminating adaptive features. The prior probability of each user was incorporated into this framework. Markov chain was used to model the unstructured variation in population.

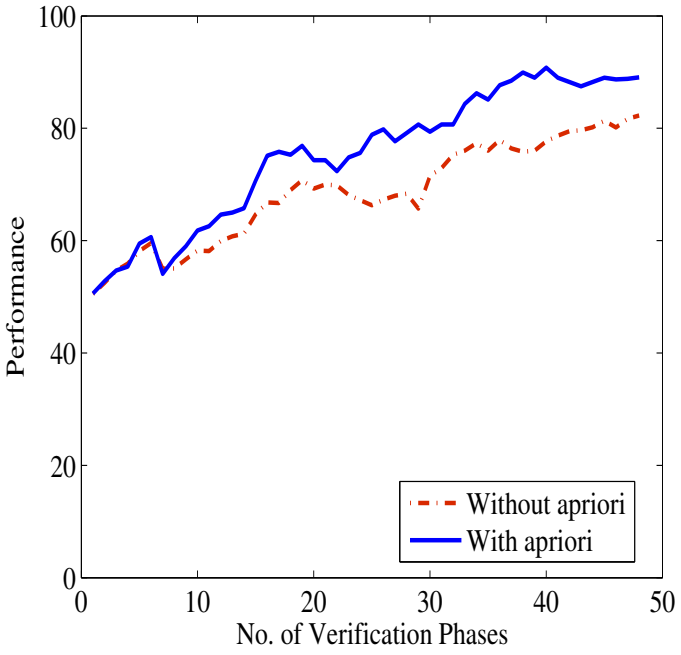
5.3.1 Results on Un-structured Variation in Population

Various experiments were conducted in order to show that incorporating *apriori* probability of participation of each individual helps to improve the performance of authentication. Markov chain was used to model the un-structured variation in population.

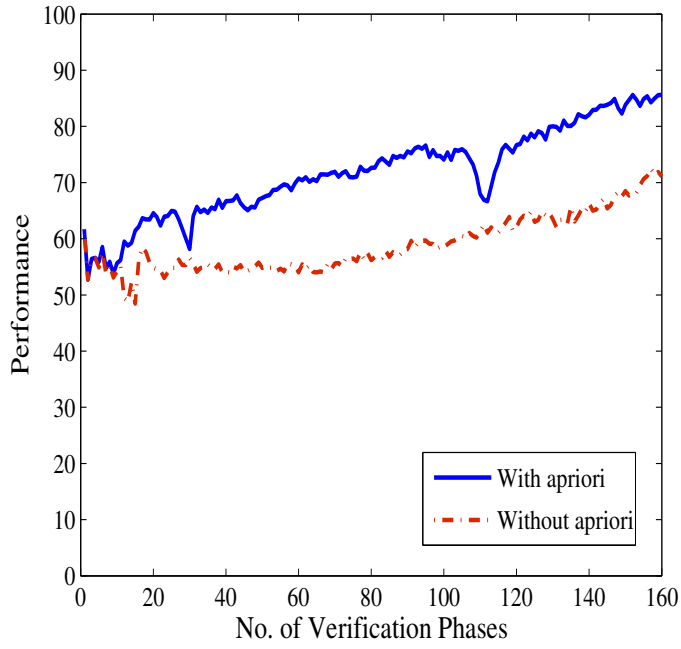
Performance Over Time: Experiments were conducted on both the datasets D1(b) and D3 to show the improvement in performance of the authentication. For these experiments, window of size 6 was chosen for computation of prior probabilities using Markov chain for both the datasets. Figure-5.1 clearly shows the improvement in performance when *apriori* probability of participation of each individual was incorporated over performance with simple IBDA-transformed features. However, on certain days (For e.g., in Figure-5.1(a), on days 16, 21 and 32, and in Figure-5.1(b), on days 46, and 110), performance with *apriori* probability was observed to be almost the same as that without incorporating *apriori* probability. This was because during those periods, prior probability of participation of each individual became almost equal to each other and hence, all individuals contributed equal towards the performance. This is essentially same as using simple IBDA-transformed features as it assumes all users to participate equally-likely on each day.

Effect of Varying Window Size: Our method for computing prior probability of participation of each individual is based on his history of participation for last M days. This value is used in computation of transition probabilities for the Markov chain. Table-5.1 shows the average performance of the authentication during a period of every 10 days on the dataset D3. It can be clearly observed that performance with the windows of sizes 1 and 3 were lower as compared to those with window-sizes 6, 8 and 12. This is because very small window size gives very less information based on which users' participation can be anticipated in future. Optimal performance was observed with window-size 6. Larger windows resulted in degradation of performance because unwanted information used to predict the participation.

Effect of Varying Number of Users: Experiments were also conducted in order to show the effect of number of users on the performance of authentication. Table-5.2 shows the results. When



(a)



(b)

Figure 5.1: Effect of incorporating *apriori* probability of participation for unstructured variation in population on datasets (a) D1(b) and (b) D3. The performance improves when the *apriori* is incorporated as compared to when it is not incorporated.

Period	1	2	3	4	5	6	7	8	9	10
1	51.6	49.7	52.7	52.3	48.4	53.6	51.7	50.4	49.6	50.9
3	64.3	65.9	65.4	66.3	67.5	66.3	69.0	72.1	75.3	79.3
6	68.6	71.7	71.8	73.3	74.2	75.0	77.6	78.2	82.5	85.14
8	67.2	69.3	69.8	69.2	71.9	71.4	72.7	75.1	81.4	83.3
12	57.9	59.2	62.5	64.1	64.2	65.0	67.2	67.7	68.2	68.6

Table 5.1: Effect of varying window size on the performance of authentication.

only IBDA-transformed features were used without incorporating *apriori* probabilities, the performance was observed to decrease with increasing number of users drastically while the degradation was observed to be at a much lower rate. This is because even though the number of users enrolled into the system increased, number of users participating was much lower. Hence the effective population size contributing to performance was low.

Users	10	30	40	60
Without apriori	93.4	92.7	90.4	85.9
With apriori	94.1	92.6	91.9	86.1
Users	80	110	130	150
Without apriori	83.9	79.6	76.9	70.5
With apriori	85.2	83.4	83.1	81.2

Table 5.2: Effect of number of users on the performance of authentication with un-structured variation in population. The rate of degradation of performance is at a much lower rate with increasing number of user when the *apriori* probabilities were incorporated.

Effect of Entry and Exit of Users over Time: We conducted experiments to show the pattern of change in performance when the users enrolled and exited from the system over time. Performance was recorded after a gap of 10 days between each consecutive observation. Table-5.3 shows that the system showed degradation in performance when new users entered into the system. While this degradation was higher when *apriori* was not used, it was at a lower rate when priori probability was incorporated. However, when the users started exiting, the performance did not change much when prior probability of participation was not incorporated while the an improvement in performance was observed when the users started exiting the system. This is because in the latter case, only the participating users contributed to the performance while in the previous case, sample of all the enrolled users were used for feature selection.

Users	30	50	100	120	140
Without apriori	93.6	85.2	81.3	78.9	70.8
With apriori	92.8	90.2	79.7	75.5	74.3
Users	110	90	60	40	20
Without apriori	68.4	67.7	65.3	66.7	65.2
With apriori	77.2	83.9	88.4	90.6	91.4

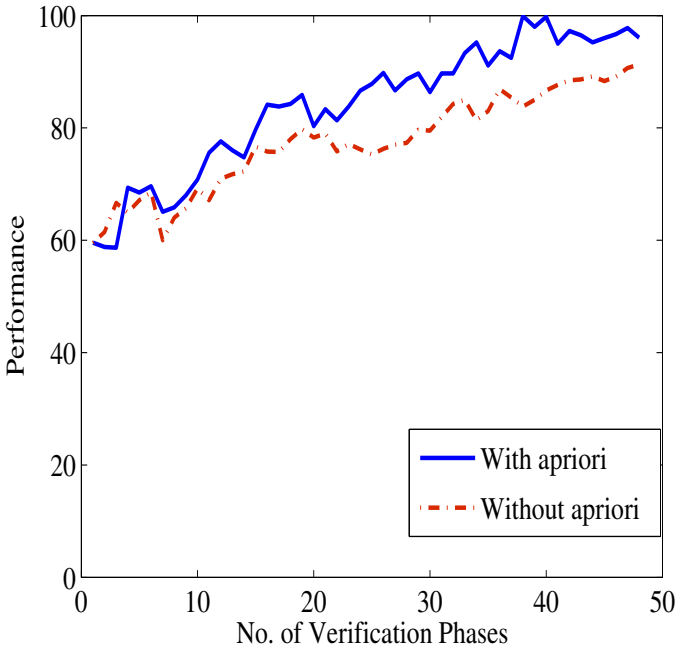
Table 5.3: Effect of users entering and exiting the system on performance on un-structured variation in population. A decline in performance is observed with entry of users in each case. Performance of authentication improved with the exit of users when *apriori* was incorporated, while it decreased when *apriori* was not used.

5.3.2 Results on Structured Variation in Population:

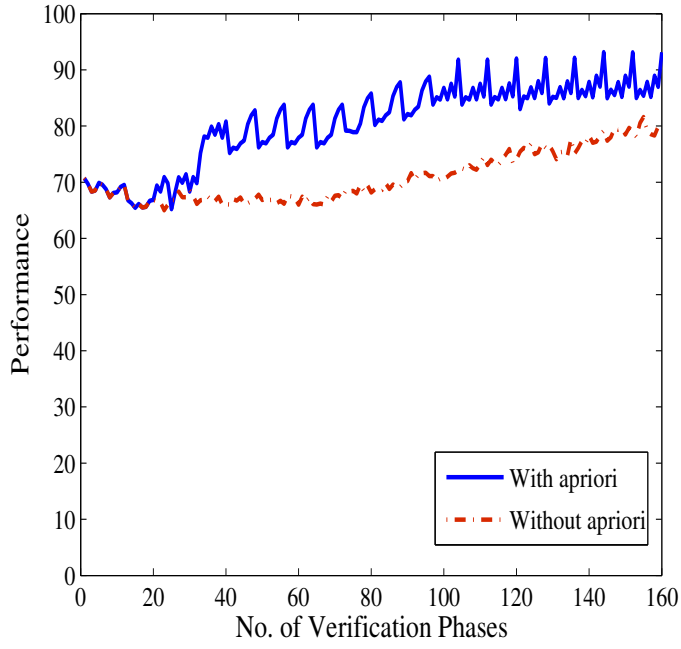
Hidden Markov Model was used to model the structured variation in population. Various experiments were conducted in order to show that modeling the structured variation in population helps to improve the performance of authentication. Period of eight days was used for modeling the population using HMM.

Performance Over Time: Experiments were conducted on both the datasets D1(b) and D3 in order to show that the performance improves over time when prior probability is incorporated into the system. Figure-5.2 shows the comparison of performance of authentication when prior probability was incorporated over when it was not incorporated into the IBDA feature selection framework. Figure-5.2(a) shows the improvement in performance clearly. Figure-5.2(b) shows clearly that the performance changed periodically. This is because on each day of the 8-day period, different number of users participated and this pattern repeated after every 8-day period. This also shows that performance was directly proportional to the number of users actually participating in the authentication process. This periodicity is not clearly visible with the dataset D1 in Figure-5.2 because same number of users participated on all the days of the period.

Comparison with Re-training the System: Traditionally, in order to improve performance of the system with entry or exit of users, pattern recognition systems require the system to be re-trained with existing users. Our method bypasses the need of re-training by considering only the participating users for feature selection and still obtain almost same performance as obtained after re-training the system with existing users. This experiment was conducted with all the 150 users enrolled into the system but different number of actual participating users using our method of incorporating prior probabilities. This was compared with performance when the system was



(a)



(b)

Figure 5.2: Effect of incorporating apriori probability of participation for structured variation in population on datasets (a) D1(b) and (b) D3. Improvement in performance is observed when the *apriori* is incorporated. The performance varies periodically with the periodic variation in population.

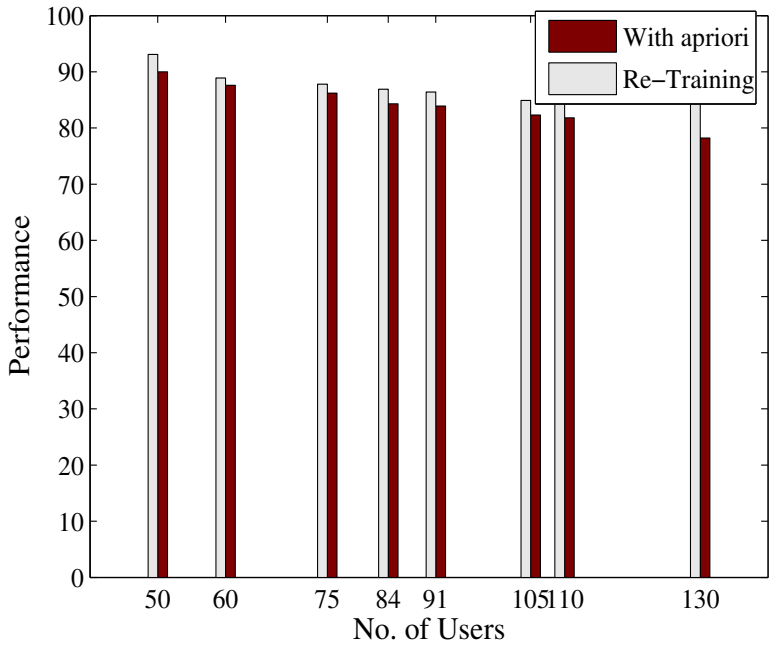


Figure 5.3: Comparison of performance obtained after re-training the system. The performance of authentication using *apriori* is almost same as that when the system is re-trained with the existing users.

re-trained with the existing users. Figure-5.3 shows that our framework showed almost same performance as that obtained after re-training the system.

Effect of Varying Number of Users: Effect of varying number of users was studied for structured variation in population. Table-5.4 shows that a degradation at a rapid rate was observed when population model was not incorporated into IBDA framework. Due to the lower effective participating population, a lower rate of degradation of performance was observed when *apriori* probability was used with IBDA.

Effect of Entry and Exit of Users over Time: Effect of entry and exit of users with time on the performance of authentication was studied for structured variation in population. Performance was observed to decline with entry of users for both the cases, i.e. when prior probability was not used and when it was incorporated into the framework. However the rate of decrease was lower when *apriori* was incorporated. With exit of users, an improvement in performance was observed when *apriori* was used while it remained almost same when *apriori* was not incorporated. Table-5.5 shows the results.

Users	10	30	40	60
Without apriori	93.7	92.9	91.3	87.6
With apriori	94.3	94.0	92.1	88.9
Users	80	110	130	150
Without apriori	85.9	81.8	78.2	74.2
With apriori	87.0	85.3	85.1	84.9

Table 5.4: Effect of number of users on performance with structured variation in population. The rate of degradation of performance using *apriori* is at a lower rate as compared to when the *apriori* is not incorporated.

Users	30	50	100	120	140
Without apriori	92.0	89.3	82.3	80.3	77.3
With apriori	93.9	91.9	84.6	85.9	85.6
Users	110	90	60	40	20
Without apriori	76.3	74.9	74.2	73.3	72.9
With apriori	85.5	84.9	91.7	92.3	92.0

Table 5.5: Effect of users entering and exiting the system on performance on structured variation in population.

5.4 Summary

In this chapter, we proposed to improve performance of authentication using weak biometrics for civilian applications. We argued that the effective population participating in the process is usually very less as compared to the total population enrolled into the system. We proposed to incorporate the prior probability of participation of each user into the adaptive feature selection framework. This results in optimized performance of authentication over time over regularly participating population. Through various experiments we showed that incorporating prior probability helps to improve the performance of authentication.

Chapter 6

Conclusions and Scope for Future Work

In this thesis, we looked into the possibility of using weak biometrics for civilian applications. We presented various techniques to improve the performance of authentication using weak biometrics, specifically for civilian applications. Our contributions were mainly in handling the inherent problems of weak biometrics, namely, low discriminating content, and low degree of stability. We also identified that the large enrolled population leads to the low performance in case of civilian applications of weak biometrics and modeling the variation in participating population improves performance of authentication.

We proposed a novel user-specific feature selection technique, called Single-Class Hierarchical Discriminant Analysis to improve the discriminating content of weak biometrics. The SCHDA projects the samples of each individual to a new feature space where the samples of the claimed identity are well separated from the samples of all the other users. We compared the performance of authentication using the SCHDA technique with the well known feature selection techniques used in biometrics and showed considerable improvement in performance. The experiments were performed on the stationary hand-geometry dataset and the learned keystroke-dynamics dataset.

In order to handle the poor stability of features of the weak biometric traits, we proposed an incremental feature selection technique. This is based on the assumption that civilian applications work in cooperative mode and the users can give feedback to the system on occurrence of errors. The incremental feature selection technique updates the parameters of the feature selection framework on occurrence of errors. This updation is done using only the learned parameters and the sample which caused an error during authentication. We demonstrated the improvement in performance of the system using the incremental feature selection over the statically selected features

on the time-varying datasets of hand-geometry and keystroke dynamics.

We addressed the large population size in civilian applications by exploiting the fact that the size of regularly participating population is relatively less than the total enrolled population. Also, the participating population varies with time. We proposed to by-pass the traditional method of re-training the system with existing users by modeling the variation in participating population. We used the popular Markov models to model the variation in structured and un-structured variation in population. We showed on the datasets of hand-geometry and keystrokes collected over a number of days, and the synthetic dataset, the improvement in performance obtained by modeling the variation.

The work presented in this thesis gives an interesting direction along which lot of other research work can be done. Various weak biometrics can be combined with other weak biometrics and strong biometrics, to provide a better confidence level of identification and authentication. However, there remains the issue of computation and fusion of scores obtained by individual biometrics. Also, the issue of cost effectiveness comes into picture while deploying large-scale biometrics in Indian economy. Apart from the user-psychology (which plays an important role in Indian society), the cost of deployment and maintenance needs to be considered, especially when such systems are needed to be deployed at multiple places. Also, our technique of improving stability, and hence performance of biometric authentication relies on the users' feedback. This feedback needs to be correct so as to allow the system to update the features of the genuine user. For example, verification results of various weak biometrics can be fused to improve the performance and confidence measure of authentication (multimodal biometrics). Better features, for e.g., 3-Dimensional features of hand-geometry can be used to improve performance. Also, if the issues of low discriminating content and low degree of stability in weak biometrics are addressed to a considerable extent, the weak biometric traits can be used for the identification purpose as well.

Related Publications

- Vandana Roy and C. V. Jawahar, "Feature Selection for Hand-Geometry based Person Authentication", in *Proceedings of the International Conference on Advanced Computing and Communications (ADCOM)*, 2005, pp. 143-149.
- Vandana Roy and C. V. Jawahar, "Hand-Geometry Based Person Authentication Using Incremental Biased Discriminant Analysis", in *Proceedings of the National Conference on Communications (NCC)*, 2006, pp. 261-265.
- Vandana Roy and C. V. Jawahar, "Modeling Time-Varying Population for Biometrics", in *International Conference on Computing: Theory and Applications (ICCTA)*, 2007, pp. 361-366.

Bibliography

[1] B. Scholkopf and A. J. Smola, *Learning with Kernels*. MA: MIT Press, Cambridge, 2001.

[2] K. Delac and M. Grgic, "A survey of biometric recognition methods," *in International Symposium Electronics in Marine, ELMAR*, June 2004.

[3] Biometrics History by Subcommittee on Biometrics, March 2006 at: `http://www.biometricscatalog.org`.

[4] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," *in International Conference on Biometric Authentication (ICBA)*, pp. 731–738, July 2004.

[5] T. R. Randolph and M. J. T. Smith, "Fingerprint image enhancement using a binary angular representation," *in International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 1561–1564, May 2001.

[6] W. S. Chen, K. H. Chih, S. W. Shih, and C. M. Hsich, "Personal identification technique based on human iris recognition with wavelet transform," *in International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 949–952, March 2005.

[7] X. Yuan and P. Shi, "Iris feature extraction using 2d phase congruency," *in International Conference on Information Technology and Applications*, vol. 2, pp. 437–441, July 2005.

[8] H. Proenca and L. A. Alexandre, "A method for the identification of inaccuracies in pupil segmentation," *in International Conference on Availability, Reliability and Security (ARES)*, pp. 224–228, April 2006.

[9] J. Gao, H. M. Dong, D. G. Chen, L. Gan, and W. W. Dong, "Research on synergetic fingerprint classification and matching," *in International Conference of Machine Learning and Cybernetics*, vol. 5, pp. 3066– 3071, November 2003.

[10] K. C. Chan, Y. S. Moon, and P. S. Cheng, "Fast fingerprint verification using subregions of fingerprint images," *in IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, pp. 95– 101, January 2004.

[11] H. M. El-Bakry, "Fast iris detection for personal identification using modular neural networks," *in International Symposium on Circuits and Systems*, vol. 3, pp. 581–584, May 2001.

[12] R. S. Reillo, "Hand geometry pattern recognition through gaussian mixture modelling," *in International Conference on Pattern Recognition (ICPR)*, pp. 937–940, September 2000.

[13] S. Gonzalez, C. M. Travieso, J. B. Alonso, and M. A. Ferrer, "Automatic biometric identification system by hand geometry," *in International Carnahan Conference on Security Technology*, pp. 281–284, October 2003.

[14] E. Yoruk, H. Dutagaci, and B. Sankur, "Hand biometrics," *in Image Vision Comput.*, pp. 483–497, 2006.

[15] E. Konukoglu, E. Yoruk, J. Darbon, and B. Sankur, "Shape-based hand recognition," *in IEEE Transactions on Image Processing*, pp. 1803–1815, July 2006.

[16] L. Hong and A. K. Jain, "Integrating faces and fingerprints for personal identification," *in Trans. on Pattern Recognition and Machine Intelligence (PAMI)*, vol. 20, 1997.

[17] A. K. Jain, L. Hong, and Y. Kulkarni, "A multimodal biometric system using fingerprints, face and speech," *in International Conference on Audio Video based Biometric Person Authentication*, pp. 182–187, March 1999.

[18] R. W. Frischholz and U. Dieckmann, "Bioid: A multimodal biometric identification system," *in Computer*, pp. 64–68, February 2000.

[19] A. Ross, A. Jain, and J. Z. Qian, "Information fusion in biometrics," *in International Conference on Audio- and Video-Based Person Authentication (AVBPA)*, pp. 354–359, June 2001.

[20] Y. Wang, T. Tan, and A. K. Jain, "Combining face and iris biometrics for identity verification," *in International Conference on Audio- and Video-based Biometric Person Authentication (AVBCA)*, pp. 805–813, June 2003.

[21] R. Snelick, M. Indovina, J. Yen, and A. Mink, "Multimodal biometrics: Issues in design and testing," *in International Conference on Multimodal Interfaces (ICMI)*, pp. 68–72, November 2003.

[22] M. G. K. Ong, T. Connie, A. T. B. Jin, and D. N. C. Ling, "A single-sensor hand geometry and palmprint verification system," *in ACM Workshop on Biometrics: Methods and Application (WBMA)*, pp. 100–106, Novemeber 2003.

[23] A. K. Jain and A. Ross, "Multibiometric systems," *in Communications Of The ACM*, vol. 47, pp. 34–40, January 2004.

[24] A. Ross and R. Govindarajan, "Feature level fusion in biometric systems," *in Biometric Consortium Conference (BCC)*, September 2004.

[25] K. Nandakumar and A. K. Jain, "Score normalization in multimodal biometric systems," *in Pattern Recognition*, vol. 38, pp. 2270–2285, 2005.

[26] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," in *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, pp. 4–19, January 2004.

[27] Biometrics Resource at: `http://www.findbiometrics.com/`.

[28] Biometrics Consortium: Government Applications and Operations at: `http://www.biometrics.org/REPORTS/CTSTG96/`.

[29] International Biometrics Group at: `http://www.biometricgroup.com/applications.html`.

[30] A. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 4–19, January 2004.

[31] A.K.Jain, A. Ross, and S.Pankati, "A prototype hand-geometry based verification system," *in International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, pp. 166–171, March 1999.

[32] A. K. Jain and N. Duta, "Deformable matching of hand shapes for verification," *in International Conference on Image Processing (ICIP)*, pp. 857–861, October 1999.

[33] L. Wong and P. Shi, "Peg-free hand geometry recognition using hierarchical geometry and shape matching," *in IAPR Workshop on Machine Vision Applications*, pp. 281–284, 2002.

[34] D. Umphress and G. Williams, "Identity verification through keyboard characteristics," in *International Jounal Man-Machine Studies*, vol. 23, pp. 263–273, 1985.

[35] W. L. Bryan and N. Halter, "Studies in the physiology and psychology of the telegraphic language," in *The Psychology of Skill: Three Studies*, pp. 35–44, 1973.

[36] A. Peacock, X. Ke, and M. Wilkerson, "Typing patterns: A key to user identification," in *IEEE Security and Privacy*, vol. 2, pp. 40–47, September 2004.

[37] J. Ilonen, "Keystroke dynamics," in *Lappeenranta University of Technology, Skinnarilankatu 34, 53850 Lappeenranta*.

[38] S. Mandujano and R. Soto, "Deterring password sharing: User authentication via fuzzy c-means clustering applied to keystroke biometric data," in *Mexican International Conference in Computer Science*, vol. 00, pp. 181–187, September 2004.

[39] M. Villani, C. Tappert, G. Ngo, J. Simone, H. S. Fort, and S.-H. Cha, "Keystroke biometric recognition studies on long-text input under ideal and application-oriented conditions," in *Computer Vision and Pattern Recognition Workshop*, pp. 39–39, June 2006.

[40] A. K. Jain, S. Prabhakar, L. Hong, A. Ross, and J. L. Wayman, "Biometrics: A grand challenge," *in International Conference on Pattern Recognition (ICPR)*, vol. 2, pp. 935–942, August 2004.

[41] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley Interscience Publication, 2 ed., 2000.

[42] M. Visani, C. Garcia, and J.-M. Jolion, "Two-dimensional-oriented linear discriminant analysis for face recognition," in *International Conference on Computer Vision and Graphics (ICCVG)*, September 2004.

[43] M. Visani, C. Garcia, and J.-M. Jolion, "Bilinear discriminant analysis for face recognition," in *International Conference on Advances in Pattern Recognition (ICAPR)*, pp. 247–256, August 2005.

[44] Y. Zhu and E. Sung, "Margin-maximization discriminant analysis for face recognition," in *International Conference on Image Processing (ICIP)*, vol. 1, pp. 609–612, October 2004.

[45] J. Lu, K. Plataniotis, and A. Venetsanopoulos, "Face recognition using kernel direct discriminant analysis algorithms," in *IEEE Transactions on Neural Networks*, vol. 14, pp. 117–126, January 2003.

[46] P. Wang and Q. Ji, "Learning discriminant features for multi-view face and eye detection," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 373–379, June 2004.

[47] B. Zhang, X. Chen, S. Shan, and W. Gao, "Nonlinear face recognition based on maximum average margin criterion," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 554–559, June 2005.

[48] F. Li, J. Kosecka, and H. Wechsler, "Strangeness based feature selection for part based recognition," in *Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, pp. 22–22, June 2006.

[49] J. Deng and H. Tsui, "A pca/mda scheme for hand posture recognition," in *International Conference on Automatic Face and Gesture Recognition (AFGR)*, vol. 00, p. 0294, 2002.

[50] S. K. Zhou and R. Chellappa, "Multiple-exemplar discriminant analysis for face recognition," in *International Conference on Pattern Recognition (ICPR)*, pp. 191–194, 2004.

[51] X. S. Zhou and T. S. Huang, "Small sample learning during multimedia retrieval using biasmap," in *Computer Vision and Pattern Recognition (CVPR)*, pp. 1–17, 2001.

[52] S. Yasuoka, Y. Kang, and K. Morooka, "Texture classification using hierarchical discriminant analysis," in *International Conference on Systems, Man and Cybernetics*, pp. 6395–6400, 2004.

[53] N. Ramanathan and R. Chellappa, "Face verification across age progression," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 462–469, June 2005.

[54] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 387–394, June 2006.

[55] A. K. Jain and A. Ross, "Learning user-specific parameters in a multibiometric system," in *International Conference on Image Processing (ICIP)*, vol. 1, pp. 57–60, September 2002.

[56] P. Hall and R. Martin, "Incremental eigenanalysis for classification," in *British Machine Vision Conference (BMVC)*, pp. 286–295, 1998.

[57] N. Vasconcelos and M. Vasconcelos, "Scalable discriminant feature selection for image retrieval and recognition," in *International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 770–775, June 2004.

[58] T. Ko, "Multimodal biometric identification for large user population using fingerprint, face and iris recognition," *in Proceedings of 34th Applied Imagery and Pattern Recognitioin Workshop(AIPR05)*, pp. 219–224, October 2005.

[59] R. Meshulam, S. Reches, A. Yarden, and S. Kraus, "Mlbp:mas for large-scale biometric pattern recognition," *in 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 1095–1097, May 2006.