Investigation of Multimodal Template-Free Biometric Techniques and Associated Exception Handling

A Thesis Submitted to The University of Kent
For the Degree of Doctor of Philosophy
In Electronic Engineering

By

Saad Rashed Aldosary

April 2015

Acknowledgements

First and foremost I would like to thank Allah for bless me with finishing my thesis. After that, I would like to thank my family –my father, my mother, brothers, sisters and importantly my wife for all the support they have given me during my Ph.D. I am eternally indebted to them and without them this thesis would not have been possible.

Secondly, I would like to extend my gratitude to my supervisor Dr Gareth Howells for his assistance, guidance and encouragement throughout my time at the University of Kent. I would also like to wish Gareth all the best with the commercialisation of this project and with all his research endeavours. Likewise, I would like to express my thanks and best wishes to my best friend Dr, Abdullah Aldosary for his support and guidance though out my Ph.D research.

Finally,

To all my friends at the Department of Electronics at the University of Kent, I wish you all the best with your studies and careers, and I especially want to thank all of you that have made my time in Kent highly enjoyable.

Abstract

The Biometric systems are commonly used as a fundamental tool by both government and private sector organizations to allow restricted access to sensitive areas, to identify the criminals by the police and to authenticate the identification of individuals requesting to access to certain personal and confidential services. The applications of these identification tools have created issues of security and privacy relating to personal, commercial and government identities. Over the last decade, reports of increasing insecurity to the personal data of users in the public and commercial domain applications has prompted the development of more robust and sound measures to protect the personal data of users from being stolen and spoofing. The present study aimed to introduce the scheme for integrating direct and indirect biometric key generation schemes with the application of Shamir's secret sharing algorithm in order to address the two disadvantages: revocability of the biometric key and the exception handling of biometric modality. This study used two different approaches for key generation using Shamir's secret sharing scheme: template based approach for indirect key generation and template-free. The findings of this study demonstrated that the encryption key generated by the proposed system was not required to be stored in the database which prevented the attack on the privacy of the data of the individuals from the hackers. Interestingly, the proposed system was also able to generate multiple encryption keys with varying lengths. Furthermore, the results of this study also offered the flexibility of providing the multiple keys for different applications for each user. The results from this study, consequently, showed the considerable potential and prospect of the proposed scheme to generate encryption keys directly and indirectly from the biometric samples, which could enhance its success in biometric security field.

List of Publication

1- Aldosary, Saad and Howells, Gareth (2012) *A Robust Multimodal Biometric Security System Using the Polynomial Curve Technique within Shamir's Secret Sharing Algorithm.* In: 2012 Third International Conference on Emerging Security Technologies. pp. 66-69. ISBN 9781467324489.

Table of contents

List of Figures

List of Tables

List of Abbreviations

- ABIS: Automated Biometric Identification System.
- AFIS - Automated Fingerprint Identification System.
- API: Application Programming Interface.
- Bio API: Biometric Application Programming Interface.
- DET: Detection Error Trade off.
- DNA: Deoxyribonucleic Acid.
- DoD: Department of Defence.
- EER: Equal Error Rate.
- FAR: False Accept Rate.
- FBI: Federal Bureau of Investigation.
- FMR: False Match Rate.
- FNMR: False Non-Match Rate.
- FRR: False Reject Rate.
- FTA: False to Acquire.
- FTC: False to Capture.
- FTE: False to Enrol.
- GLCM: Gary Level Co-occurrence Matrix.
- HD: Hamming Distance.
- PIN: Personal Identification Number.
- RER: Retrieval Error Rate.
- ROC: Receiver Operation Characteristics.
- SSS: Shamir’s Secret Share.

Glossary

- **Automated Biometric Identification System**: “1. Department of Defense (DOD) system implemented to improve the U.S. government's ability to track and identify national security threats. The system includes mandatory collection of ten rolled fingerprints, a minimum of five mug shots from varying angles, and an oral swab to collect DNA” [A, B].
  “2. Generic term sometimes used in the biometrics community to discuss a biometric system” [A, B].
- **Accuracy**: “A catch-all phrase for describing how well a biometric system performs. The actual statistic for performance will vary by task (verification, open-set identification (watchlist), and closed-set identification)” [A, B].
- **AFIS - Automated Fingerprint Identification System**: “A highly specialized biometric system that compares a submitted fingerprint record (usually of multiple fingers) to a database of records, to determine the identity of an individual. AFIS is predominantly used for law enforcement, but is also being used for civil applications (e.g. background checks for soccer coaches, etc)” [A, B].
- **Algorithm**: “A limited sequence of instructions or steps that tells a computer system how to solve a particular problem. A biometric system will have multiple algorithms, for example: image processing, template generation, comparisons, etc” [A, B].
- **Application Programming Interface**: “Formatting instructions or tools used by an application developer to link and build hardware or software applications” [A, B].
- **Attempt**: “The submission of a single set of biometric sample to a biometric system for identification or verification. Some biometric systems permit more than one attempt to identify or verify an individual” [A, B].
- **Authentication**: “1. The process of establishing confidence in the truth of some claim. The claim could be any declarative statement for example: “This individual’s name is ‘Joseph K.’ ” or “This child is more than 5 feet tall.” 2. In biometrics, “authentication” is sometimes used as a generic synonym for verification” [A, B].
- **Behavioural Biometric Characteristic**: “A biometric characteristic that is learned and acquired over time rather than one based primarily on biology. All biometric characteristics depend somewhat upon both behavioural and biological characteristic. Examples of biometric modalities for which behavioural characteristics may dominate include signature recognition and keystroke dynamics” [A, B].
- **Biometrics Application Programming Interface**: “Defines the application programming interface and service provider interface for a standard biometric technology interface. The BioAPI enables biometric devices to be easily installed, integrated or swapped within the overall system architecture” [A, B].
- **Biological Biometric Characteristic**: “A biometric characteristic based primarily on an anatomical or physiological characteristic, rather than a learned behaviour. All biometric characteristics depend somewhat upon both behavioural and biological characteristic. Examples of biometric modalities for which biological characteristics may dominate include fingerprint and hand geometry” [A, B].
- **Biometric Data**: “A catch-all phrase for computer data created during a biometric process. It encompasses raw sensor observations, biometric samples, models, templates and/or similarity scores. Biometric data is used to describe the

information collected during an enrollment, verification, or identification process, but does not apply to end user information such as user name, demographic information and authorizations" [A, B].

- **Biometric Sample**: "Information or computer data obtained from a biometric sensor device. Examples are images of a face or fingerprint" [A, B].
- **Biometric System**: "Multiple individual components (such as sensor, matching algorithm, and result display) that combine to make a fully operational system. A biometric system is an automated system capable of:
  1. Capturing a biometric sample from an end user.
  2. Extracting and processing the biometric data from that sample.
  3. Storing the extracted information in a database.
  4. Comparing the biometric data with data contained in one or more reference references.
  5. Deciding how well they match and indicating whether or not an identification or verification of identity has been achieved" [A, B].
- **Biometrics**: "A general term used alternatively to describe a characteristic or a process.
  As a characteristic: A measurable biological (anatomical and physiological) and behavioural characteristic that can be used for automated recognition.
  As a process: Automated methods of recognizing an individual based on measurable biological (anatomical and physiological) and behavioural characteristics" [A, B].
- **Capture**: "The process of collecting a biometric sample from an individual via a sensor" [A, B].
- **Claim of Identity**: "A statement that a person is or is not the source of a reference in a database. Claims can be positive (I am in the database), negative (I am not in the database) or specific (I am end user 123 in the database)" [A, B].
- **Crypto-Biometric:** "The field of study covering the design, development, evaluation, and analysis of crypto-biometric systems. The research in this field can be dated back since 1998" [A, B].
- **Crypto-Biometric system:** "A system that combines biometric with cryptography in order to remove one or more drawbacks of either of the two techniques" [A, B].
- **Crypto-bio Key:** "A key obtained from or with the help of biometric data" [A, B].
- **Database**: "A collection of one or more computer files. For biometric systems, these files could consist of biometric sensor readings, templates, match results, related end user information, etc" [A, B].
- **Decision**: "The resultant action taken (either automated or manual) based on a comparison of a similarity score (or similar measure) and the system's threshold" [A, B].
- **Biometric Direct Key Generation:** "A system that able to generate a biometric encryption key directly from the biometric samples with a template free which means the template doesn't need to be stored" [A, B].
- **Biometric Indirect Key Generation:** "A system that required a biometric template to be stored in order to generate the biometric encryption key" [A, B].
- **Equal Error Rate**: "A statistic used to show biometric performance, typically when operating in the verification task. The EER is the location on a ROC or DET curve where the false accept rate and false reject rate are equal. In general, the lower the equal error rate value, the higher the accuracy of the biometric system. Note, however, that most operational systems are not set to operate at the "equal

error rate" so the measure's true usefulness is limited to comparing biometric system performance. The EER is sometimes referred to as the "Crossover Error Rate." "[A, B].

- **Encryption:** "The act of transforming data into an unintelligible form so that it cannot be read by unauthorized individuals. A key or a password is used to decrypt (decode) the encrypted data" [A, B].
- **Enrollment:** "The process of collecting a biometric sample from an end user, converting it into a biometric reference, and storing it in the biometric system's database for later comparison" [A, B].
- **Feature Extraction:** "The process of converting a captured biometric sample into biometric data so that it can be compared to a reference" [A, B].
- **Face Recognition:** "A biometric modality that uses an image of the visible physical structure of an individual's face for recognition purposes" [A, B].
- **False Match Rate:** "A statistic used to measure biometric performance when. Similar to the False Acceptance Rate (FAR)" [A, B].
- **False Non-Match Rate:** "A statistic used to measure biometric performance. Similar to the False Reject Rate (FRR), except the FRR includes the Failure to Acquire error rate and the False Non-Match Rate does not "[A, B].
- **False Acceptance Rate:** "A statistic used to measure biometric performance when operating in the verification task. The percentage of times a system produces a false accept, which occurs when an individual is incorrectly matched to another individual's existing biometric. Example: Frank claims to be John and the system verifies the claim" [A, B].
- **Feature(s):** "Distinctive mathematical characteristic(s) derived from a biometric sample; used to generate a reference "[A, B].
- **Fingerprint Recognition:** "A biometric modality that uses the physical structure of an individual's fingerprint for recognition purposes. Important features used in most fingerprint recognition systems are minutiae points that include bifurcations and ridge endings" [A, B].
- **False Rejection Rate:** "A statistic used to measure biometric performance when operating in the verification task. The percentage of times the system produces a false reject. A false reject occurs when an individual is not matched to his/her own existing biometric template. Example: John claims to be John, but the system incorrectly denies the claim" [A, B].
- **Failure to Acquire:** "Failure of a biometric system to capture and/or extract usable information from a biometric sample" [A, B].
- **Failure to Enroll:** "Failure of a biometric system to form a proper enrollment reference for an end user. Common failures include end users who are not properly trained to provide their biometrics, the sensor not capturing information correctly, or captured sensor data of insufficient quality to develop a template" [A, B].
- **Hamming Distance:** "The number of non-corresponding digits in a string of binary digits; used to measure dissimilarity. Hamming distances are used in many Daugman iris recognition algorithms" [A, B].
- **Identification:** "A task where the biometric system searches a database for a reference matching a submitted biometric sample and if found, returns a corresponding identity. A biometric is collected and compared to all the references in a database. Identification is "closed-set" if the person is known to exist in the database. In "open-set" identification, sometimes referred to as a "watchlist," the

person is not guaranteed to exist in the database. The system must determine whether the person is in the database, then return the identity” [A, B].

- **Impostor:** “A person who submits a biometric sample in either an intentional or inadvertent attempt to claim the identity of another person to a biometric system” [A, B].
- **Iris Recognition:** “A biometric modality that uses an image of the physical structure of an individual’s iris for recognition purposes. The iris muscle is the colored portion of the eye surrounding the pupil” [A, B].
- **IrisCode©:** “A biometric feature format used in the Daugman iris recognition system” [A, B].
- **Minutiae Point:** “Friction ridge characteristics that are used to individualize a fingerprint image. Minutiae are the points where friction ridges begin, terminate, or split into two or more ridges. In many fingerprint systems, the minutiae (as opposed to the images) are compared for recognition purposes2 [A, B].
- **Modality:** “A type or class of biometric system. For example: face recognition, fingerprint recognition, iris recognition, etc” [A, B].
- **Multimodal Biometric System:** “A biometric system in which two or more of the modality components (biometric characteristic, sensor type or feature extraction algorithm) occurs in multiple” [A, B].
- **Noise:** “Unwanted components in a signal that degrade the quality of data or interfere with the desired signals processed by a system” [A, B].
- **Performance:** “A catch-all phrase for describing a measurement of the characteristics, such as accuracy or speed, of a biometric algorithm or system” [A, B].
- **Personal Identification Number:** “A security method used to show “what you know.” Depending on the system, a PIN could be used to either claim or verify a claimed identity” [A, B].
- **Pixel:** “A picture element. This is the smallest element of a display that can be assigned a colour value “[A, B].
- **Recognition:** “A generic term used in the description of biometric systems (e.g. face recognition or iris recognition) relating to their fundamental function. The term “recognition” does not inherently imply the verification, closed-set identification or open-set identification (watchlist)” [A, B].
- **Reference:** “The biometric data stored for an individual for use in future recognition. A reference can be one or more templates, models or raw images2 [A, B].
- **Resolution:** “The number of pixels per unit distance in the image. Describes the sharpness and clarity of an image” [A, B].
- **Receiver Operating Characteristics:** “A method of showing measured accuracy performance of a biometric system. A verification ROC compares false accept rate vs. verification rate. An open-set identification (watchlist) ROC compares false alarm rates vs. detection and identification rate” [A, B].
- **Similarity Score:** “A value returned by a biometric algorithm that indicates the degree of similarity or correlation between a biometric sample and a reference2 [A, B].
- **Spoofing:** “The ability to fool a biometric sensor into recognizing an illegitimate user as a legitimate user (verification) or into missing an identification of someone that is in the database” [A, B].

- **Template:** “A digital representation of an individual’s distinct characteristics, representing information extracted from a biometric sample. Templates are used during biometric authentication as the basis for comparison” [A, B].
- **Threat:** “An intentional or unintentional potential event that could compromise the security and integrity of the system” [A, B].
- **Threshold:** “A user setting for biometric systems operating in the verification or open-set identification (watchlist) tasks. The acceptance or rejection of biometric data is dependent on the match score falling above or below the threshold. The threshold is adjustable so that the biometric system can be more or less strict, depending on the requirements of any given biometric application” [A, B].
- **Token:** “A physical object that indicates the identity of its owner. For example, a smart card” [A, B].
- **Verification:** “A task where the biometric system attempts to confirm an individual’s claimed identity by comparing a submitted sample to one or more previously enrolled templates” [A, B].
- **Vulnerability:** “The potential for the function of a biometric system to be compromised by intent (fraudulent activity); design flaw (including usage error); accident; hardware failure; or external environmental condition” [A, B].

Reference

[A] National Science & Technology Council (NSTC), 14 September 06, http://www.biometrics.gov/Documents/glossary.pdf

[B] Biometrics in Support of Identity Management, Joint Capabilities Document (JCD) Glossary, 4 April, 2008.

Chapter 1.

Introduction

1.1. Introduction

As both government and private sector organizations face a growing number and range of security threats, employees are increasingly given controlled access to their respective departments by means of biometric systems. The widespread application of Internet and related technologies has turned the world into a global village, with greater interconnectivity and unprecedented levels of communication giving access to a wealth of information and knowledge. While bringing people from different communities and cultures closer together, this unprecedented reliance on the Internet and digital devices and systems for the storage of sensitive and strategic data has created insecurity in the transfer of data from one system to another as malicious agents seeking to exploit these weaknesses attack databases and steal or misuse data for their own ends. In particular, this has exacerbated issues of security and privacy relating to personal, commercial and government identities. Over the last decade, the insecurity of the personal data of users in public and commercial domain applications has prompted the development of more robust and sound measures to protect the identity of users from being stolen and spoofing [1-6]This vulnerability is not confined to individual users' data but also extends to systems controlled by government and commercial organizations that hold users' data for security checks and other purposes. To reduce the probability of theft and corruption of these public data, both government and commercial organizations employ a range of security measures that include the application of traditional identification systems. These systems work by means of PINs and smartcards linking indirectly to the user’s personal information to access sensitive data stored in the database. However, these PINs and smartcards can be forgotten or stolen; and they can be exploited to misuse the user’s identity [1].

For this reason, a more advanced approach to security based on so- called ‘biometric’ systems was introduced by government and commercial organizations to protect user identity. In public domain applications, these systems use biometric information linked intrinsically to the individual’s personal data–details that are uniquely characteristic of him/her—making biometric systems a powerful instrument for authenticating the user’s identity. In contrast to the PINS/Tokens-based identification system, biometrics-based identification systems uniquely identify the person using their own characteristic features [1].

1.2. Biometric Authentication Background

In the last decade, as many government and commercial organizations introduce e-commerce, e-government and Internet-based communications to the public domain for data sharing and teleconferencing, there has been an accompanying emphasis on the security and confidentiality of the data shared and disseminated through these services. In this regard, authentication systems have been developed to establish the identity of users using these systems and services. As an authentication mechanism, automated authorization system is considered the preferable approach in this context, as it does not require the supervision of these channels for each and every use. These automated authentication systems employ the existing triad of mechanisms, alone or in combination—token-based (something an individual possesses, such as a key or fob), specific knowledge-based (something an individual knows, such as a password or personal identification number/PIN) and characteristic-based (something that defines an individual, such as biometric data). To achieve the high level of security required, the application of biometric systems has been widely recommended for every use of e-commerce and e-government services [1, 2].

User authentication is the process by which a system tries to determine the truth of the claim made by the user by executing a one-to-one match between the query data and the registered data of the user within the system. Normally, the user's data are stored by assignment of a unique identifier number at the time of registration. When a person subsequently attempts to establish their identity, the system matches the query data with the unique identifier linked to the relevant personal data. In this way, authentication is achieved by a successful match between the unique identifier and the query data. This process differs from the identification process executed. During identification, the query data is matched with all items of data stored in the system, one by one, and if a match is found with any of the data, the system displays a message to confirm that identity is established. Passwords, PINS and tokens are not considered suitable for identification purposes, as they can be easily stolen, forgotten, lost or transferred to an unauthorized person and easily copied in collusion with others who have knowledge of these secret keys; they can even be cracked by hackers to obtain access to the user's personal data and are regarded as a weak method of protection when implemented for the purpose of authentication [1, 2].

In light of these issues, researchers have employed a combination of PINS/tokens and knowledge-based authentication to enhance the protection of users' personal data stored within the system, as for example in the 'chip and PIN' system used for most e-commerce services in the United Kingdom (UK). Yet despite these stringent measures, data can still be compromised, owing in part to negligence or collusion. Biometrics-based identification systems are given more credence in terms of security because biometrics, as characteristic features of individuals, can be copied, stolen and transferred to other people; additionally, the claimant must be present at the time of identification. Furthermore, biometrics-based systems automatically perform the functions of extraction and matching between stored and query data, and so it cannot be made known to people in the vicinity of the user. However, systems in the public domain such as fingerprint- or face-based biometric systems, the biometric data can be surreptitiously acquired for fraudulent use of the user's identity. Similarly, biometric data can be stolen from a database in the absence of adequate spoof prevention measures and security of the biometric template, and further research is needed to bridge these gaps in the security of biometric systems [1, 2].

1.2.1. Choosing a Biometric Modality

Biometrics-based authentication systems compare the query biometric data with the stored template of the user. The comparison is strictly made on similarities of characteristics shared by the query data and the stored biometric template linked to the user, enabling the user to be uniquely identified through the authentication system. There are several modalities in use for the development of biometric systems, including face, voice, signature, fingerprint, iris and gait. These can broadly be divided into two categories of modality: physiological and behavioural. Physiological modalities relate to physiological characteristics of the person such as fingerprints, palmprints and retinal and iris patterns. Behavioural modalities are characteristic ways in which a person executes some action, such as a signature or keystroke. The application of these modalities in combination is considered best practice in boosting the security of a biometric system [1, 2].

The choice of modalities for construction of a biometric system depends on four important properties of the modality:

- Universality: the biometric should be obtainable from everybody.
- Uniqueness: It should provide a high degree of discriminatory information.
- Permanence: It should be invariant as it ages.
- Collectability: It should be practical to extract.

Biometric modalities can be extracted in different ways. The extraction methods used will vary according to modality, and it is important to select the appropriate method for the chosen modality. The selected method should have the following features: high performance rate, universally accepted by users and robust to circumvention [1, 2]. Weaknesses and strengths of the various biometric modalities are presented in Table 1.1.

Table 1.1 Biometric Modalities and Property Strengths [1, 2].

	Universality	Uniqueness	Permanence	Collection	Performance	Acceptability	Circumvention
Face	HIGH	LOW	MED	HIGH	LOW	HIGH	HIGH
Fingerprint	MED	HIGH	HIGH	MED	HIGH	MED	MED
Hand Geometry	MED	MED	MED	MED	MED	MED	MED
Iris	HIGH	HIGH	HIGH	MED	HIGH	LOW	LOW
Hand/Finger Vein	MED	MED	MED	MED	MED	MED	LOW
Signature	LOW	LOW	LOW	HIGH	LOW	HIGH	HIGH
Voice	MED	LOW	LOW	MED	LOW	HIGH	HIGH

Some of the ratings in the above table may be contentious and must be used with caution for the target biometric system, taking into account the subjectivity and purpose of the examiner. The construction of a matrix like Table 1.1 enables the examiner to choose the best biometric modality for use in different systems, based on the requirements and utility of the particular situation. A single biometric modality or some combination of them may be used to build a biometric system, but there are four typical requirements that are universally applied: speed, accuracy, discriminative power and intrusiveness.

The research presented in this thesis investigates direct key generation from multimodal biometric samples for implementation in a cryptosystem. To that end, it is important to select a specific modality to be examined; for various reasons, iris, fingerprint and face modalities were chosen for the purposes of this research. With reference to Table 1.1, it is clear that these three modalities rate highly for most of the properties desirable in a biometric modality. These are also the most mature biometric modalities and have been extensively researched [10]. Furthermore, the iris, fingerprint and face modalities also perform relatively well in large-scale applications by virtue of their high degree of inherent individuality and universality, which means that, logically, the extracted key-space should be very large and free of any repetitions [1]. For present purposes, the left and right iris were considered as two different biometric modalities; the third modality considered here for direct key generation is the user password.

1.2.2. Typical Biometric System Framework

Biometric systems are widely used for identification purposes in different settings around the globe such as Internet banking, law enforcement and ATMs. The most commonly employed biometrics in these systems include iris and fingerprint data. Although the biometric modality used may vary, the main working principle behind the operation of these systems is broadly similar, beginning with signal processing (image) and pattern recognition and ending with a message concerning authentication of the claim of the user [1].

The fundamental biometric system processes involved in user enrollment and subsequent classification are shown in Figure 1.1.

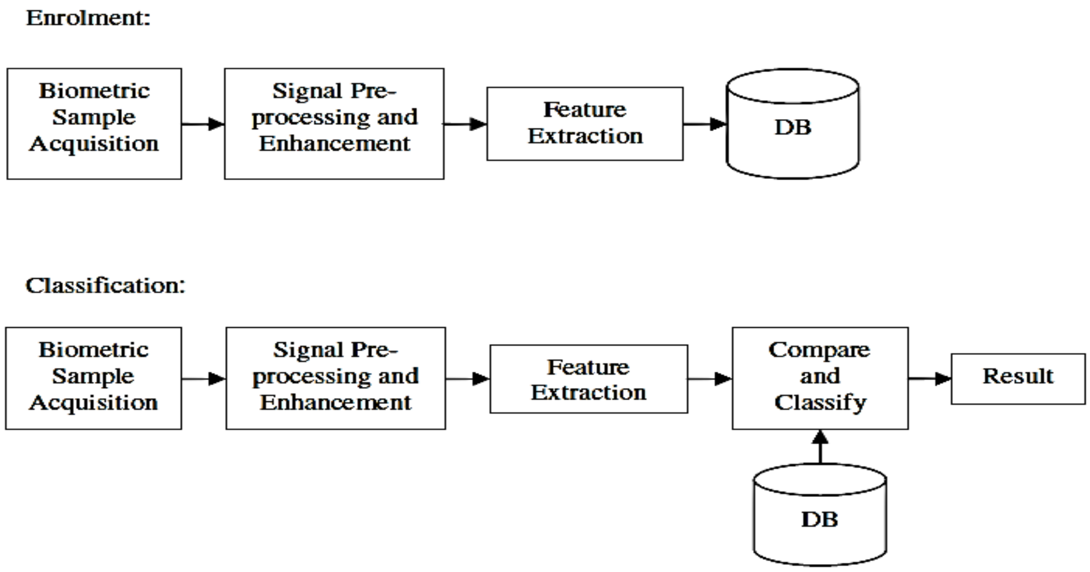

Enrolment:
Biometric Sample Acquisition
Signal Pre-processing and Enhancement
Feature Extraction
DB
Classification:
Biometric Sample Acquisition
Signal Pre-processing and Enhancement
Feature Extraction
Compare and Classify
Result
DB

Figure1.1 Processes of User Enrollment and Subsequent Classification [1]

Biometric sample acquisition, signal processing and enhancement and feature extraction are common to all automated biometric systems. Comparison and classification of the query data are challenging because of certain issues associated with these processes, which will be highlighted in the next section.

1.2.3. Challenges

Signal processing and enhancement: Several issues arise in relation to signal processing and enhancement. The first step executed by the system is to acquire the query data in the form of an sample taken from the analogue (real) world. This sample is captured by means of a dedicated camera. The acquired sample is then presented to the sensors for conversion of the analogue version of the sample into a digitized and discrete version for subsequent processing and analysis. During the conversion of analogue data to discrete digital data, there is always some loss of information.

Furthermore, environmental factors also affect the quality of analogue data derived from the real world—for instance, background light can cause the image to be opaque.

It is therefore useful and necessary to ensure that the signal is captured properly and enhanced effectively for further analysis. Both the signal preprocessing and enhancement stages are essential for normalization and minimization of the variance of signals. The enhancement functions improve image quality by accentuating characteristic features and filtering noise from the background of the sample. Preprocessing is executed to normalize the effects of important features so that all of them can pass equally and consistently through the feature extraction stage. During preprocessing, poorer quality features are highlighted so that they will be considered as noise for feature extraction. Simultaneously, those regions of features that do not belong to the user are also pinpointed through segmentation to exclude them from final processing and feature extraction.

The foregoing steps are of great significance in minimizing intra-sample variations; if these were not minimized, they would continue to accumulate during every stage of the analysis and would be likely to affect the biometric system's pattern recognition performance. For this reason, intra-sample variations are minimized from the outset to disallow variations that might build up at the feature extraction stage. This paves the way for execution of the matching algorithm.

Feature Extraction Stage: This is a very important stage as the performance of the automated biometric system ultimately depends on the quality of the features extracted. The feature extractors should show the characteristic parts of the obtained data clearly and distinctly. The features extracted must also be reproducible, highly discriminative (i.e. displaying the small variations in data obtained from the same source and highlighting greater amount of variations of data obtained from the different sources) and invariant to scale and affine transformations.

However, the extraction of features with characteristic signatures is not straightforward. Discriminative features might be prominent and identifiable with ease, which can affect the performance and quality of decisions made by the biometric system. For example, iris samples taken from the eyes of the same person may show 5–25% variations in their binary codes, as is characteristic of False Reject Rate (FRR) in most iris recognition-based biometric systems. Those taken from different individuals may show similarities of up to 70%, which is alarming and may affect the outcomes of experiments involving iris recognition. Moreover, the comparison between two different iris codes taken from two different users may show similarities of up to 65 to 75 % which may affect the direct key generation from the iris samples within template free scheme, which will be discussed in Chapter 4 later on. Another example of such misrepresentation during the feature extraction stage involves the minutiae of the fingerprints. These are considered highly distinctive and characteristic of every individual, yet their representation may change across different samples collected from the same person at different points in time. The identification of a proper representation of the feature is therefore as important as the identification of the sample itself. The preferred approach is to take the optimum feature selection path for simplification of the classification procedures adopted by the biometric system. On the other hand, the selection of sub-optimal features is likely to increase classifier complexity as it needs to include modelling to reduce feature variations.

1.2.4. Security Issues

Breaches in system security are not always related to the quality of the biometrics and algorithms employed to perform pattern recognition tasks. Another component of the problem relates to the fact that biometric systems are part of the public domain, where user data can be surreptitiously obtained by malicious agents—for example, recording a voice, capturing images of the user's face and latent fingerprints, making it possible for a spoofer to claim someone else's identity by fooling the system. Certain commercial products retain control of the biometric system by patenting to preserve intellectual rights, and users may request some level of transparency to minimize concerns about the storage of their data. These transparency requirements make more data from biometric systems available in the public domain, enabling attacks on the system using various combinations of characters to unlock it. Figure 1.2 illustrates possible attack points within a biometrics-based authentication system [15].

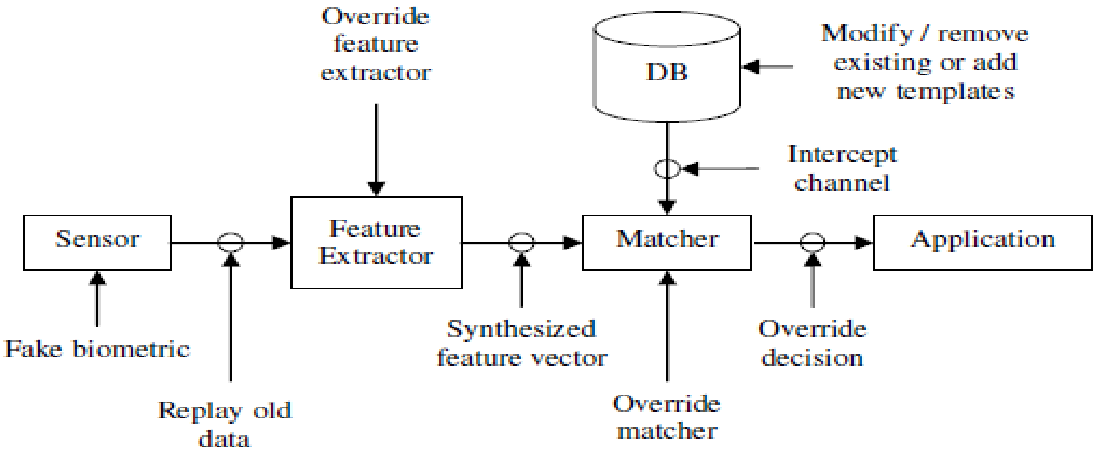


Figure1.2 Attack Points in a Biometric Based Security System [1]

The part of a biometric system most susceptible to attack is the sensor, as biometric data can easily be recovered from this part by virtue of its accessibility to the public for presentation of samples. This high level of sensor accessibility makes the system more vulnerable to spoofing attacks, and prevention of such attacks will require further research [1].

The sensors in biometrics-based systems are also supplied by third party commercial agents, which means that documentation relating to their implementation may be available online, enabling potential attackers to educate themselves about the working principles of these sensors. For example, the attacker can retrieve user data from the transit channels between the sensors and the innards of the biometric system, which can then be relayed for copying or mutation. However, such attacks can easily be circumvented in supervised systems, primarily because of the access requirements of such systems. Furthermore, hardwired sensors connected physically to the biometric system make it still harder to surreptitiously sniff or mutate data [7, 156].

While spoofing attacks can compromise individual security and privacy at the sensor level, spoofing attempts at other levels can wreak havoc on the system database. An attack on the sensors compromises the security and privacy of a single individual by breaching the data transition pathways, but the breaching of security measures at the decision, feature extraction and matching levels will allow unauthorized users to access users' biometric data in more refined forms. Such breaches in the security of feature extraction and pattern matching recognition may allow unauthorized users to change the settings of the biometric framework, denying provision of services to the legitimate users; this state is called 'denial of service' (DoS). Despite the difficulties caused by the DoS state for both system administrators and users, this does not cause serious breaches in the security and privacy of users' data. Nevertheless, circumvention at the database level of the biometric system can compromise the security of the entire database, holding hundreds or thousands of users' data [7, 156].

Clearly, then, protection of the database from spoofing attacks is critical in ensuring the security of all users registered on the biometric system. This can be achieved by use of encryption methods, which prevent attacks that aim to copy templates.

However, infringement of data by unauthorized users remains possible when the attack is launched in the guise of system administrator. In order to prevent this kind of attack, templates are decoupled from the biometric data directly, or using the one-way transformation method. This one-way transformation thwarts attempts by unauthorized persons to access the biometric data of users in the event of mutation or copying of templates [7, 156].

Although the encryption and template manipulation methods can ensure the security of users' biometric data, they are unable to prevent attacks on the database by unauthorized persons. The attack may be designed to delete, modify or overwrite existing templates in the database, which means that the system can be directly controlled by unauthorized persons or can deny services to users (DoS). The weaknesses and loopholes in remotely located communication channels can further undermine the security conditions of the database [7,156].

The templates of users can be stored in the biometric system at any four locations. A template whether be stored remotely in token or smart card, in a central database on a server, on a workstation, or directly on individual biometric sensors. If the templates are being stored on individual tokens or sensors, users must present the personal data locally in order to obtain access, which means that users have no access to applications located at other sites. Furthermore, in the event of theft, the data is vulnerable to tampering by malicious agents [7,156].

The storage of template data on local workstations minimizes the possibility of loss or theft of the data, but again, users cannot avail of authentication at multiple sites. While these issues can be resolved by storing the template data on a centralized database, this approach also creates some security concerns. For instance, the centralized database is normally located remotely, requiring transmission of data from local workstations to the centralized system through the communication channels, and so incoming and outgoing transmission of data can easily be intercepted. Even the transfer of data between the system and the database may be compromised by snooping on the communication channels. This may take several forms—for example, the attacker may alter the data or hijack the communication channels in order to copy certain packets of data. The data obtained by these methods can be analysed to extract required features of the data, which can further be manipulated by the attacker to his own ends. In order to protect the system from such attacks and to eliminate security concerns, some researchers have suggested removing database from the equation so that it cannot be accessible to anyone. However, the pattern recognition and matching stages require the availability of data to authenticate the identity of the user. Further research is required to extend the concepts of biometric encryption [7,156].

1.2.5. Biometric Encryption

This method utilizes a single or combined instance of biometric data for encryption and decryption of data, which is achieved by obtaining a key from the biometric information of the user. Like the biometrics-based authentication process, encryption requires the user to present the claim to be identified. The system will process the request by comparing the submitted samples from the user with the existing enrolled samples in the database. The system then displays the decision in a Yes/No format. However, authentication systems convert the data into 1-bit representations while encryption transforms the data into n-bit representations taken as a unique value (PIN/password). In this way, the method identifies the user of the equation, who does not know the specification of the unique value assigned to his data. Moreover, encryption methodologies relieve the biometric system of the burden of knowing a password/PIN, as authentication is executed through encryption or decryption of data [7].

One widely-used approach to executing the encryption function is by assigning/binding the key to the template, which is released upon successful authentication [16]. Another common approach is to create an unknown key and store it within a metaphorical vault, which unlocks the biometric data on successful matching [17-23].

The detail of these approaches will be elaborated in Chapter 2. However, the storage of data is an essential requirement for these approaches, which is likely to introduce some security loopholes as described earlier. The work presented in this thesis therefore investigates the generation of individual keys directly from biometric samples. By generating keys directly, the need for template creation, storage and matching is eliminated, so reducing the number of attack points. A further investigation into biometric encryption is described in Chapter 5.

1.3. Research Motivation

Despite the advantages of biometric based identification system as described in the above section, unsupervised operation of single-biometric identification systems may offer opportunities for malpractice to the would-be impostors, and it is important that the chosen technology can thwart attempts by malicious agents to spoof or copy the biometric data. Among the technologies that can prevent attackers from penetrating a system and spoofing biometric data is the multimodal biometric system. However, there remain many gaps in this approach requiring further research to make this technology more widely applicable to the needs of both public and commercial organizations. One promising area of development relates to methodologies for the generation of secret keys and template-free biometric systems [1, 2].

Despite the advantages of biometric based identification system as described in the above section, unsupervised operation of single-biometric identification systems may offer opportunities for malpractice to the would-be impostors, and it is important that the chosen technology can thwart attempts by malicious agents to spoof or copy the biometric data. Among the technologies that can prevent attackers from penetrating a system and spoofing biometric data is the multimodal biometric system. However, there remain many gaps in this approach requiring further research to make this technology more widely applicable to the needs of both public and commercial organizations. One promising area of development relates to methodologies for the generation of secret keys and template-free biometric systems [1, 2].

The present study investigates how such a multimodal, template-free biometric system can be based on generating the secret key directly; specifically, it investigates biometric exception handling strategies using Shamir's Secret Sharing technology, which allows an arbitrary subset of biometric modalities to be employed in identifying an individual. This study focuses on solving the exception handling as it is a serious issue in the biometric security, especially in the cases where a user cannot produce the biometric modality at the time of identification due to changes in his/her physical condition , for example, loss of iris due to accident, finger cuts or surgical treatments etc.

In the literature relating to biometrics and security, several researchers have emphasized the application of multimodal biometric systems to enhance the security of users' identities and protect data against spoofing, stealing and other misuses.

One immediate issue with token/passwords/PINs-based identification is the difficulty of remembering PINs, tokens or passwords; the multimodal biometric system relieves the user of this burden, making it easy to use while effortlessly protecting the user's data. However, the literature has also reported some weaknesses and difficulties in using biometric systems in terms of preventing attacks on the system. Some researchers have suggested use of the application encryption technique to develop a stronger and more robustly secured multi-biometric system [1, 7, 8].

In addition, two potential drawbacks inherent in the previously developed encryption biometric security systems scheme are: a) it does not easily lend itself to revoking the generated encryption key and b) it is not robust in the exception-handling scenario, where a user is unable to provide a sample of the given biometric modality. These issues can reduce the potential of these single-modality biometric systems to identify the individual with the damaged features. The present research is trying to investigate the potential of existing techniques for development of a template-free multimodal biometric system.

The security of biometrics-based systems will be greatly enhanced and strengthened by the ability of the proposed system to derive an encryption key directly from biometric samples provided by a given user. Such a system would entail that no copy of the encryption key need to be retained and also that no template or reference sample of the related biometrics would be required. Both of these features add value to the security potential of the system.

1.4. Aims and Objectives of the study

The present research aims to introduce a scheme for integrating direct biometric key generation schemes with Shamir's secret sharing algorithm [11] in order to directly address the two disadvantages of revocability and exception handling.

The thesis will describe techniques for:

- An investigation of template-free multimodal biometric-based encryption in a number of scenarios.
  1. Direct generation of encryption keys from biometric multimodalities, using a suitable technique from the following three:
     i. Linear equation.
     ii. Quadratic equation.
     iii. Cubic curve.
  2. Generation of a secret share from each modality:
- An investigation of biometric exception-handling strategies using Shamir's secret sharing technology, allowing an arbitrary subset of biometric modalities to be employed to identify an individual.

1.5. The Scope of the research

This research investigated on developing the design of two template-free biometric systems, in which biometric data can be mapped onto repeatable unique binary code strings in both cases, opening the key only in the presence of biometric prints. The first system will use the Gray Level Co-Occurrence Matrix method (GLCM) to extract the iris feature, and the second system will use the 2D Gabor Filter. To generate the secret key, Shamir's secret scheme will be applied to the three data points provided (right iris, left iris and password). By use of the linear equation technique, two points of three will then be sufficient to release the biometric key, giving the user the desired flexibility while reducing the FRR. For both designs, a three-factor scheme is proposed, including smartcard, password and biometrics. Here, each component is crucial to reveal the biometric data specific to an individual, with the opportunity of either updating or revoking the key; all three factors would be required to compromise the integrity of key.

1.6. Research Challenges

1.6.1. Inter-Class and Intra-Class Variability

These two types of variation play a fundamental role in the issues related to the pattern recognition stage of the biometric system, as they determine the level of receiver operating characteristics (ROC) performance, which in turn depends on the false-matching rate (FMR) and false non-matching rate (FNMR). Inter-class variations are controlled by the 'randomness' of the given samples—for instance, as the iris sample contains more random characteristics than the fingerprint, the system may perform the recognition task of the iris samples at extremely low FMR [24-28]. Nevertheless, the 'random patterns' factor of the submitted sample increases the intra-class variations determined at high FNMR. For instance, intra-class variations in the iris samples can be introduced by many factors such as pupil contraction/dilation, eyelash/eyelid occlusion and reflected light. These issues can be corrected to some extent by use of quality control measures at the time of sampling. For example, if the eye is overly occluded or blurred, the biometric system will reject the image [26]. However, such tactics undermine the ergonomics of the biometrics-based authentication system. There is also evidence that a less controlled mechanism can be applied to acquire the iris sample, including strategies like capturing the image from a frontal angle and obtaining a level position by setting the appropriate tilt of head and gaze direction.

1.6.2. Generation of Key Directly from Biometrics Samples

It is very challenging to generate the key directly from the given biometric sample without storing the templates, which are considerably reproducible (key stability) with large numbers of effective bits (key entropy). If it is assumed that feature extraction from samples taken from the same modality are likely to be consistent, then the feature vectors of different irises will inevitably overlap in the feature space, especially when the feature space is small. Subsequently, the overlapping feature vectors are likely to generate the same potential binary code patterns. If there is a consistent group of irises and the multiple feature vectors of an iris overlap with them, the number of effective bits can be reduced dramatically because the consistent group will have potentially similar binary code patterns, which will inevitably decrease the unique characteristics associated with the key. This is one potential scenario for irises belonging to same group.

These issues normally appear in relation to the consistent extraction of feature vectors from each iris. Together, these feature vectors will have almost differing degrees of variation, which raises many issues due to instability of the key. The issues of uniqueness and key stability are therefore associated with the method of generating the key directly from the iris sample.

1.7. Thesis Outline

The present research encompasses a number of fields. To minimize cross-referencing between chapters, the content is organized as follows.

Chapter 2 presents an in-depth literature review of the background to basic biometrics, biometric system errors, a comparison of various biometrics and the advantages and disadvantages of biometrics, biometric cryptosystems and template-free biometrics and discussion of current work on direct key generation from the biometric sample.

Chapter 3 outlines a technique for the extraction of a secret key from a multimodal biometric system using Shamir's secret scheme algorithm, which is an established technique for protecting a given item of data by means of information distributed to several participants. This technique has been used for five different biometric modalities: face, index finger, thumb finger, left iris and right iris.

Chapter 4 outlines a number of investigations of direct key generation from the iris modality. These investigations returned negative results and are described as failed schemes. The methodologies are discussed, along with reasons for their failure.

Chapter 5: While the study presented in Chapter 4 returned some negative results, much has been gained in terms of knowledge. In light of these issues, Chapter 5 presents the design of two template-free biometric systems, in both of which biometric data can be mapped onto repeatable unique binary code strings, opening the key only in the presence of biometric prints. The first system uses the Gray Level Co-Occurrence Matrix method (GLCM) to extract iris features, and the second system uses the 2D Gabor Filter. To generate the secret key, Shamir's secret scheme is applied to the three data points provided (right iris, left iris and password). Two points from three are then sufficient to release the biometric key by use of the linear equation technique, giving the user the desired flexibility while reducing the FRR. For both designs, a three-factor scheme is proposed, to include smartcard, password, and biometrics. Here, each component is crucial in revealing the biometric data specific to an individual, with the opportunity of either updating or revoking the key; all three factors would be required to compromise the integrity of the key.

Chapter 6: The final chapter draws general conclusions about the research as a whole along with concluding remarks and recommendations for future work.

.

Chapter 2.

Literature Review

2.1 Introduction

In the modern society, the leakage of the personal details through stealing, spoofing or unauthorized usage has increased to an alarming extent. [4-6]. Due to lack of security, several researchers have focussed on the development of foul-proof biometric systems. The literature on the biometric systems and recent advances in this field have showed that multi-modal biometric systems can play a vital role in securing the personal details from being unauthorized used, stealing and spoofing. This system can free the users from burden of remembering the password or carrying the key in pockets in all the time by providing the keyless solutions to access their personal bio-data in a particular system. However, even the multi-modality of biometric system does not make it fault-free due to some weaknesses and pitfalls; and that it can be subject to external threats of stealing the saved data in the system. Because the risk of these threats, new approach which is called biometric encryption being used to secure the public data by providing the key-sharing solution safely. The properties of this new approach made this system an exciting researchable area for those working on the development of safe and secure biometric solutions for the public and personal data. [1, 10, 29].

Biometric encryption allows the derivation of encryption keys directly from the sample data obtained from the users, thereby eliminating the necessity of the preservation of key or template or the reference sample relevant to the stored biometric data. These features add an element of enhanced security to the stored data, and prevent any attacks on the system. However, like any other system, this method involves two inherent pitfalls in its functioning: difficulty in revoking the generated encryption key and its inability to handle the scenario in which user shows his/her inability to provide the sample having given biometric modality.

These issues led to the development of schemes which would allow the integration of a direct biometric key using existing biometric key generation schemes by applying the Shamir's secret sharing algorithm. These efforts tried to remove the forgoing issues of exception handling and revocability.

The proposed scheme in this thesis requires the provision of arbitrary sets of the biometric modalities, which will further allow the researcher to assign the arbitrary encryption key to each modality, thereby resolving the issue of revocability. This allowed the researcher to build a model in which arbitrary key can be supplied to those scenario which are unable to produce the required user biometric modality. Furthermore, this also allowed the derivation of multiple independent encryption keys from a single sub-set of biometric modality.

Biometric encryption is a method which involves the derivation of keys for encryption and decryption by processing the physiological and behavioural characteristics of the individual supplying the biometric data.

2.2 Background

This section focuses on background information which will be discussed a brief literature in biometric and a multi-modal biometric system of implementing various aspects which is lead to the main point of this research which is Bio-Cryptosystems and template free investigating. Biometrics and cryptography can be combined together into a one system by Bio-Cryptosystems. In this chapter, introduction will be presented on field of multimodal biometrics and bio-cryptosystems. Then, a discussion of a bio-cryptosystem developing was involved by the various methodologies. After that, in bio-cryptosystems field the previous works will be carried out by different researchers. Finally, a template free technique will be carried and discussed in different points views. Moreover, list of previous work will be presented and explained the comparison between them.

2.2.1 History of Biometric Analysis

The human beings have several biological characteristics which help recognize the identity of the individual from each other. For centuries, the distinctive features of individual such as gait, voice, face etc are being used to recognize the people from each other. However, it is the raw concept of recognition. Officially, the use of biological traits was approved by Alphonse Bertillon, chief of criminal identification unit in Paris, for the recognition of criminals in mid 19th century [1, 30]. However, it was far from the use of fingerprints for identification of criminals, and that idea was discovered in the late 1980s following the historic discovery of fingerprinting by Alec Jafferys at the department of Genetics, University of Leicester. The major law-enforcement agencies in the UK and the rest of the Europe embraced this method of genetic fingerprinting to identify the criminals. The methodology involved the collection of genetic fingerprints from the suspects and preservation in the data file also referred as a print. The fingerprints often in fragmentary shape gathered from the site of crime were compared with those in the database in order to infer the identity of the criminal. The accuracy of this method in terms of solving the crime cases produced the best results for the prevention of crimes in Europe. Later on, the popularity of this method increased tremendously, and it was accepted by various countries across the globe. There are some other body measurements which came in use for identification of the people such as voice and face. This concept of identification of the people were not restricted to only law-enforcement agencies, it was also adopted by civilian organizations in different countries for the identification of the people entering the sensitive areas. For example, the method was used to establish the parenthood, to identify the illegal immigrants, security clearance of workers etc [1, 30].

The pertinent question is what kind of behavioural or physiological traits can qualify to be used as a biometric data [1, 30]. The following criteria are required to be met for being qualified as biometric data:

Universality: The selected trait should be universal i.e. each person should have a characteristic feature which would define his /her personality

Distinctiveness: The human physiological and behavioural trait must have distinct signatures, which are recognizable from other human beings.

Permanence. The selected features to be used a biometric data should be invariant, which means it should change over the time or with age.

Collectability: The physiological or behavioural trait used as biometric data should be collectable and quantifiable easily.

There are some requirements for biometric for the practical biometric system which employs biometric information for the recognition of personal details.

Performance: The system should be accurate, fast and achieve the desired recognition level by taking into account different operational and environmental considerations.

Acceptability: The system should inferred minimally to the personal lives of the individuals, so that its use can be accepted by people easily

Circumvention: The system should be made foul-proof from any vicious attacks on the system intended to steak the information or misuse the individual biometric data for fraudulent purpose. It should be robust in enhancing the security of stored data about the individuals.

Harmless: The system should not harm the users physically or health-wise.

Biometric system is the process of collection of behavioural or physiological data from the individuals, which should be able to extract the required set of data on request and compare with the template stored in the system in order to confirm the identity of the individuals.

Depending on the requirements, the biometric system may be used in identification mode or the in the verification mode.

Verification mode:

In this mode, the system verifies the claim of the user which wish to be identified. In order to execute its function, the biometric system captured the biometric data from the user which might be in form of Personal Identification Number (PIN) or the access card smart card, chip card etc, then it makes one-to-one comparison of the individual claim of identity with the stored data about the user. On successful comparison, the system verifies the identity and allows user the access to the required facilities [30, 31].

This mode is applied for the positive recognition, when the motive is to prevent the multiple users from using the same identity.

Identification Mode:

In this mode, the biometric system identify the claim of the user being true or false. This is achieved by capturing the data about the users, the system asks the question “Whose biometric data is this?”. The one-to multiple comparison of the captured claim is made by the system with all biometric data sets stored in the system. IN the case of failure of any match within the database, the system rejects the claim or otherwise accepts it. Basic aim of this mode is negative recognition executed to prevent the user from taking on multiple identities [30].

In contrast to positive recognition, the token, PIN, smart cards can not be used for the negative identification. This is only achieved using the biometric data provided by the user. Nevertheless, the user enrolment is common to both modes of the system [30].

However, the researcher will use the term “verification” throughout the thesis without having any intention to make distinction between identification and recognition. The illustration of verification and identification modes of the biometric system has been presented in form of block diagram (Figure 2.1).

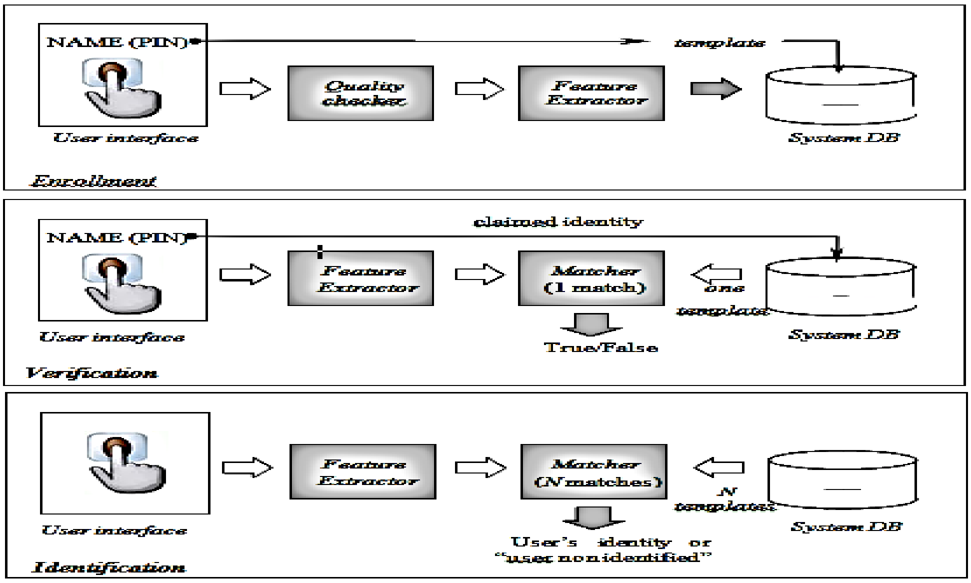


Figure2.1 Block diagrams of enrollment, verification and identification tasks are shown using the four main modules of a biometric system, i.e., sensor, feature extraction, matcher, and system database [30].

The process of recognition of the user by verification mode can be given as follows: Feature vector extraction from the biometric data given as X_Q and a claimed identity I, determine if (I, X_Q) belongs to class w_1 or $w_{2.}$ Where w_1 indicate that the claim is a genuine user and w_2 indicate that the claim is an imposter user. Typically, X_Q is matched against X_I, where X_I is the biometric template corresponding to user I. Thus, the following equation will helped to determine the category of the user claim identity.

$$(I, X_Q) \in \begin{cases} w_1 & if\ S(X_Q.X_I) \geq t. \\ w_2 & otherwise \end{cases} \quad (2.1)$$

Where S represents the extent of similarity between X_1 and the feature vector X_Q, and "t" denotes the pre-defined threshold. The function $S(X_Q, X_1)$ gives the matching score or similarity index between the claimed identity and the biometric data of the user stored in the system [30].

Main Modules of Biometric System:

The biometric system may contain following modules based on the type of the function performed by them:

- Sensor Module: This modules is important for sensing or capturing the user’s biometric data. This modules contain camera to take the image of the given biological feature of the user such as finger-prints. For instance, the sensor module of the finger-print based biometric system takes the image of the ridges and valleys of the finger of the user [1, 30].
- Feature Extraction Module: This application performs the task of extraction of minutiae features such as the local ridges and the valley details of the user’s finger. This processing highlights and specifies the position and orientation of the minute structures in the fingerprint image, which are extracted for further processing [1, 30].
- Matcher Module: This module compares the extracted features of the user with the stored templates in the system in order to execute the recognition function. For example, the fingerprint biometric system make comparison between the minutes structures of ridges and valleys of input finger image to those in the template set in the database, the level of similarity between the input and the template is expressed by means of matching score which is reported by the system. The matcher module also contains the decision module, which takes the decision on the user’s claim as “confirmed (verified)” or the user’s identity is established (identification” based on the matching score [1, 30].
- System database module: This module is also called enrolment module as it enrols the users for the provision of the personal data and its successive storage in the database of the system. During this phase, the biometric traits of the user such as fingerprint are taken by the system through scanning process of the fingers. The scanning is performed by the biometric reader, which may or may not be supervised by the human during its operation. The quality of the image

is checked by various quality control parameters of the image. After ensuring the quality and reliability of the image, the feature extractor become operational in order to extract the minute features of the biometric data, thereby resulting in the compact and illustrative image of the digital representation which is also called ‘template’. Depending on the nature of the function of the biometric system, the templates may be loaded onto the chip-cards, smart cards, issued in form of PINS, tokens or keys. Usually, the system keeps the record of multiple templates against a single individual which serves to compensate the variations occurring in the biometric data over a particular period of time [1, 30].

2.2.2 Comparison between Various Biometrics

There are several biometric characteristics are being used to recognize the identity of the user in both police-cum-military and civilian worlds. Based on the characteristics of this biometrics, the software designers have developed several applications which use them and try to establish the identity of the person in a particular environment. However, it must be recognized that no single biometric meet all the requirements of the existing system and vice versa. Therefore, it comes down to the fact there is non-existent of an optimal biometric application or biometric measurement. The performance of the applications depends on their operational mode and the properties of the biometric measurements for such applications.

This section describes the various biometrics and their properties for the different applications used to develop the biometric system.

Deoxyribose Nucleic Acid (DNA):

DNA is the unique biochemical molecule which is associated with each individual with unique genetic make-up, which constitutes the basis of its utility for identification of a personality from the other. The use of DNA for identification purpose is restricted to the only forensic application such as to establish the parenthood and to identify the criminal in order to resolve the complicated criminal cases [1, 30].

However there are many issues associated with its use, which must be taken into taken during the biometric-based recognition of the people, which are given below:

Contamination and sensitivity: DNA samples can easily be contaminated or stolen from unsuspected person, which can be used to fulfil the ulterior motives [1,30].

Recognition issues: It is really difficult to isolate and identify the DNA samples as it requires the wet techniques and expertise of the forensic scientists to process such samples. Therefore, it can not be used for online applications available for most of applications employed for the other biometrics [1, 30].

Privacy issues: The genetic make-up of an individual contains some private information such the links of hereditary diseases and some other health disorders based on the DNA molecule. The processing of DNA samples without authorization and revealing the DNA based information about the health status of the individual may harm its employment career e.g. during hiring and firing process [1, 30].

Ear:

Ear is used as a biometric to identify the person. The structure of the cartilaginous tissues of the pinna and the shape of the ear are the unique features of ear which distinguish the person from the other. Normally, the distance between the point of ear on the head and the salient point on the pinna are recorded for identification of the people. However, this is not broadly applied across various applications of biometric systems due to its intrusiveness and being less distinctive in nature [1, 30].

Face:

Face is the most commonly used non obstructive biometric which is historically and primordially used by human to recognize the people from each other. The face is used to recognize the individuals through static and controlled mug-shot verification in addition to dynamic and uncontrolled identification especially in the cluttered environment of the airport. The various attributes of face are used for the recognition purpose such as the positions of nose, eyes, eye-brows, lips, chin and their special relationships with each other. Furthermore, the global analysis of the face image is also utilized to recognize the face of the person by using the weighted combination of a number of canonical faces representing the face image [1, 30].

Although the performance of commercially available identification and verification biometric system by using the face image is sufficient, but the biometric applications used for this purpose impose a set of restrictions on the style of images for the accurate recognition of the face image [35]. For example, the biometric system requires the particular illumination method and the background of the image in order to recognize the face. The performance of face recognition based biometric system is drastically affected if they are presented with face images taken in different light and background view. Furthermore, there is also argument about the level of the confidence by which such applications area capable to recognize the person from other people based on the contextual information obtained from the face [37].

The performance of the given biometric system is judged by its capacity to detect automatically the posture of face in its right orientation, locate the face of a requested person if there are many faces available in the database, and recognize the face from its general viewpoint.

Facial, hand and hand vein infrared thermogram:

The pattern of heat generated by the hand, face and other body parts is characteristic to every individual, which could be imaged using the hidden infrared camera. This is non-obstructive method and works like taking the photographs of an individual secretly, as the device does not need the contact with the physical features of the human to execute its imaging operation. However, taking the image in the cluttered environment is really challenging. For example, if the person is surrounded by hear radiators, heat pipes, and heater or some other heat emanating surfaces, it becomes difficult to take the thermogram of the required person. The patterns of heat generated by the person and other heat emanating surfaces in the environment are likely conflict with each other, making it difficult to recognize the thermogram belonging to the required individual. In similar fashion, the near infrared camera takes the images of the pattern of veins in the clenched fist of the individual which is characteristic to each person. This technology can be seen widespread due to prohibitively high cost of infrared and near infrared camera [1, 30].

Fingerprint:

The fingerprints are unique to very individual. Even the fingerprints from the identical twins vary from each other, though they share the identical genetic make-up. The pattern of fingerprints is developed during the seventh month of the fetal development, and consists of ridges and valleys on the surface of the tips of the fingers. Normally, the index finger of the right hand is used for the verification and identification of the person based on the fingerprints. The pattern of fingerprints has been used for centuries to identify the individuals from each other apart from the salient features of the face [32].

The widespread use of this biometric measurement for the identification has become widespread due to the cheap and affordable cost of the fingerprint scanner (~ $20 US when ordered in bulk quantity). The integration cost of these scanners in the system like the laptop or desktop computer is also marginal. The several applications employ the fingerprint scanner for the identification and verification the identity of the people at places like airports, the areas of high sensitivity requiring the restricted access, the national and international civilian institutions etc. This biometric system is highly suitable for both small-to medium scale identification events involving a few hundred users and the large scale recognition events involving millions of identities [30].

The major drawback of such applications is their requirement of extensive computing resources, especially when they are operated in the identification mode. The identification of the small fraction of the population through fingerprints based biometric system is highly challenging due to genetic, environmental, aging and occupational factors. For example, the people related to the manual work may have cuts and bruises on their hands, which make it difficult to identify them by matching their fingerprints with the cognate templates in the database [30].

Gait:

The way people walk on the surface of the ground is called gait which is set of spatial-temporal biometric measurements. The gait is behavioural trait and characteristic to every individual to some extent. However, it changes over the time due to age, injuries, fluctuations in the body weight etc. Therefore, it is not very distinctive trait like fingerprints and face. Consequently, its application for the recognition of the person is restricted to only those low-security applications [30].

The processing of the gait biometric for the verification purpose require the set of steps during the sample walk of the required individual. The application measures the several various joint movements during the walk and save them in video formats, which are recalled for verification on the request. The processing of these biometric measurements require the extensive input of data and the computational labour, which makes it expensive and less feasible to be applied in every situation for the verification of the people [1, 30].

Hand and Finger Geometry:

The hand geometry based biometric system relies on the shape of the hand, the size and width of the fingers, the temporal-spatial relationships of the fingers with each other. The accuracy and reliability of these systems is fairly good, their installations can be found in hundreds of the places around the world for the verification of the people. Thus commercial available hand geometry based biometric systems can be purchased at the affordable cost [1, 30].

Furthermore, the hand geometry based systems can not be used to identify the person from the large population, as the hand geometry is less distinctive compared to voice, face and fingerprints. However, they are highly suitable for the verification of the persons. In addition, the hand geometry may be affected by the diseased condition like arthritis, the jewellery worn on fingers like rings. These limitations pose a challenge is extracting the correct information about one’s hand geometry. Though the sizes of the commercially available fingers based biometric system involving the index and middle fingers are fairly smaller than those for the hand-geometry based systems, but their sizes are still larger than face, fingerprints and voice based biometric systems [1, 30].

Iris:

The Iris is the characteristic feature of every individual, which is bounded by the pupil and the sclera on either side of the eye. The visual texture of the iris is formed during the fetal growth and developmental process, and it is kept in stabilized state for the first two years of life. The complexity lent to the textures of iris makes it characteristic feature to every individual, and that this can be utilized as a basis of recognition of individual from each other. The performance and accuracy of the commercially available iris based biometric system is fairly high [1, 30].

Due to the promising results in terms of speed and accuracy makes this system to be deployed in large scale applications aimed to identify the individuals from the large population. Like fingerprints, the irises belonging to two identical twins are different, and this distinctiveness makes it highly suitable candidate to be used as the biometric characteristic for the identification purposes. The best feature of the iris feature that it can not be tampered with the surgical processes and it is easier to recognize the artificial or implanted iris. In the past, this technology remained very expensive and non-feasible to be deployed for the large scale identifications. However, the modern iris based biometric system are more cost-effective, efficient and user friendly [1, 30].

Keystroke:

It is postulated that everyone has its own characteristic way of keying in the information. This is behavioural characteristic which may not be invariant and very distinctive like other biological traits such as fingerprints; therefore, this feature is not manipulated to identify the individuals. However, this may be useful in the verifying the identity of individual from the sample population. This is non-obstructive discriminatory method which can take the information while the person keying in the information on the computer without his/her knowledge [1, 30].

Odour:

It is well known fact that each body exudes a particular odour due to distinct chemical composition. The human also exudes a characteristic odour from their bodies, which is signature of every individual. This characteristic feature can be used for the identification of individual, during which the whiff of air around the required individual is captured and sprayed over the large array of sensors specialised for the detection of various chemical groups in the composition of the sprayed air. In this way, the variations in the chemical components of the human odour form the basis of the identification by the odour based biometric system. It is still unclear whether the effect of deodorant smells and various other chemicals in the surrounding air affect the accuracy and performance of such systems [1, 30].

Palmprint:

The palm of each individual contains characteristic ridges and valleys structures, wrinkles and principal lines which are characteristics to each individual. Since the palm's area is larger than the finger tips, they are more distinctive than the fingerprints, and carry the better identification and verification value.

The applications employed to capture the bigger image because the area of the palm is larger is than the fingers. Therefore, the high resolution Palmprint based biometric system is bulkier and bigger than the fingerprint based biometric system, which makes them expensive to be installed everywhere. However, being bigger and special and containing more distinctive characteristics, the distinct image can be captured by even the low-resolution scanners which are much cheaper than the high-resolution ones. Nevertheless, the application of high resolution scanner gives much clearer view of the hand geometry, wrinkles, ridges, valleys and principal lines, which increases the accuracy and performance of the Palmprint based biometric system [1, 30].

Retinal Scan:

The retinal vasculature is an important part of the eyes which holds the special structural features characteristic feature of every individual, which makes the basis of its use in the identification and verification biometric system. Like the iris based biometric system, it is also one of the securest biometric system as the user can not manipulate with retina of his or her eye. During the scan process, the user is asked to focus on the special point in the visual field in order to capture the image of the retina vasculature. This involves the users' acceptance and effort to be involved in the scan process; and that these factors seriously affect the acceptance level of the retinal scan among the subjects. Furthermore, the retinal scan can also reveal the presence of some medical condition such as hypertension, which makes it another deterring factor in the acceptance of retinal scan based biometric systems [1, 30].

Signatures:

Signatures are used for the verification person in several legal, government and commercial transactions, since it is known that the way individuals put their signatures on the paper is unique to each individual. Although, the signatures are styled to each individual and considered to be an important behavioural biometric, but they are not invariant. They can be subject to variations due to emotional and physical states of the signatories. Even the signatures put by the same person in successive attempts may include the small variations. Moreover, the forgers can reproduce the same signature in order to fool the system. The signatures also require the acceptance of the user, contact with the writing instruments and cooperation. These factors can deter its versatility in the verification of persons through biometric system [1, 30].

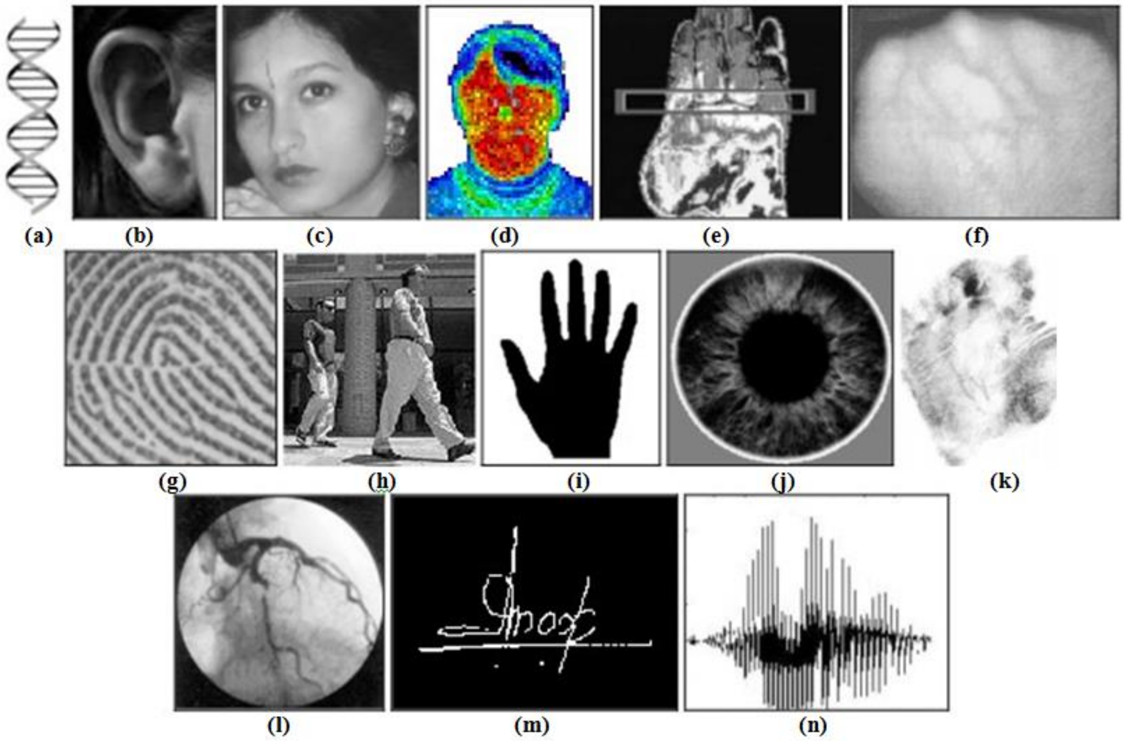


Figure2.2 Examples of biometric characteristics: a) DNA, b) ear, c) face, d) facial thermogram, e) hand thermogram, f) hand vein, g) fingerprint, h) gait, i) hand geometry, j) iris, k) palmprint, l) retina, m) signature, and n) voice [1, 30].

Voice:

Voice is also used as a vital biometric for the recognition of the persons for centuries. In order to speak, the user uses his/her mouth, vocal tracts, lips, tongue, nasal cavities, and these appendages vary in the size in every individual. This makes the voice pattern of a person distinctive of those uttered by the other people. The voice or sound patterns may be affected by weather (cold, hot) and emotional conditions (sadness, happiness), from which it can be inferred that voice is also a behaviour biometric which is not invariant for a person. The voice is also undergoes variations over time, with age and diseased conditions. Because of these factors, voice is not universally used for the large scale identification of persons [1, 30].

There are three types of the voice based biometric system: text-dependent, text-independent, phone based biometric systems. In the text-based biometric system, the subject is asked to read through the pre-determined text in order to verify the voice pattern. This system is considered more accurate but is not safe because the professional forgers can copy the style of reading the text that fool the system. Similarly, the text-independent biometric system recognizes the person based on whatever the subject speaks during normal conversation. Though it is safer method in terms of offering protection against fraudulent practices, but it requires some intricacies in its design which are sometimes difficult to achieve. Phone voice based biometric system may be good alternative to the text-dependent biometric system if the quality of the voice is protected from degradation by using the standard quality of communication channels and microphone [1, 30].

The brief comparison between all the biometrics described above is demonstrated in Table 1.1 using seven factors which are required to qualify a certain biological trait to be a good biometric. From the comparison, it is clear that there is not a single biometric system which meets the seven-point criteria of the effective biometric system, which indicates the non-existent of an optimal system for the perfect optimal level of performance and accuracy. For instance, it is a well-known fact that fingerprints and iris based biometric system are more accurate than voice based biometric system. Nevertheless, the voice-based biometric system is preferred to the fingerprints based biometric system due to seamless integration of customer’s voice into the telephone system during the tele-banking transactions.

2.2.3 Advantages and Disadvantages of Biometrics

In this section, the advantages and disadvantages of two important biometric applications are examined in detail: the commercial positive recognition applications which is capable to work in both identification and verification modes, and the forensic negative recognition application mostly employed by governmental agencies that require identification of the individuals asking access to the highly sensitive areas..

2.2.3.1 Positive Recognition in Commercial Applications

Commonly there are two methods used to execute the positive recognition:

Knowledge-based method: this method involves the allocation of PINS and passwords to the users to access certain facility. The option of setting password is given to the users who use words or digits, or mixed words and digits which can be easily remembered by them. Normally, the names of family members, birthdays, and favourite movies, names of schools or colleges attended by users are taken as passwords. According to a survey of 1200 British office workers, it was found that 50% of the participants chosen their own names, family member’s names, pet’s name as their passwords to login the system. Whilst the rest of participants had selected the names of their favourites stars (Darth Vader, Homer Simpson) to be their passwords. Such passwords can easily crack by password crackers using the simple brute force dictionary attack method [1, 30].

It is considered to be highly advisable to set the different passwords for the different applications and to keep changing the passwords after certain time period. These practices confer the enhanced protection upon the system against any malicious attack. However, most of the people ignore these cautions and set the same passwords for all the applications to avoid headache of remembering different passwords, which provides a weak security to their applications. If the hacker is able to break a password for single application, it is more likely that all applications would be hacked [1, 30].

For instance, the hacker may create a bogus website enticing the users to sign in to achieve some air miles or other monetary benefits. When the user enter the username and password to access the site, the passwords are hacked. The hacker may employ the users' information including user name and password to hack the corporate bank accounts or some other sensitive information which can be used to threaten the user. Furthermore, the hacker need to hack a single password being used to operate the intranet of the company in order to hack the whole systems used by different employees. In this way, the single password option weakens the security of the overall system [1, 30].

Similarly, longer passwords are often considered to be a secure method to protect the system, which can be written by user on some paper or cell phone because of difficulty in remembering them. However, they are not secure and can be stolen easily by some malicious agent. These types of passwords can easily lost or forgotten, so the recovery of these passwords is done by the Help desk. The help face has to incur cost of US$38 to recover a single password for the customer [1, 30]. In short, the short passwords are secure while the long passwords are more secure but easily forgotten or lost [1, 30].

Cryptographic techniques are developed to create the encryption keys using encryption method which creates strong and secure long passwords consisting of long words. The best thing about encryption keys is that they don't need to be remembered by the users, and that they are protected by simple passwords, thereby making the system highly secured and immune to fraudulent attacks [1, 30].

Token-based methods: in this method the special keys and chip cards are issued to the users for the identification purpose. If the keys or tokens are shared among the colleagues, it becomes difficult ti recognize who used the system using the shared key for his/her ulterior purposes. The problems associated with knowledge-based methods are also related to the possession-based personal recognition method. For instance, tokens or keys may be stolen, misused, a master key can be made to open several keys being used to operate a system [1, 30].

Biometrics-based recognition method: In view of insecurity and issued relating to the both foregoing methods, the system is exposed to greater risk for being manipulated by the malicious agents. Therefore, a biometric systems are developed for the identification are verification of the individuals. The biometrics can not be stolen or misused or hacked like the passwords or the tokens. The person requesting the identity or verification need to be present at the time of request processing, which makes very less likely that access is denies to the right person or provided to the false identity. The duplication of the biometrics is also difficult, and this makes the system more secured and reliable. Moreover, the computer network working on the biometric provides an equal level of security to all users, so it is difficult to break the security to all users even using the social engineering methods [1, 30].

In addition, the biometrics based recognition system confer incredible security level on the system, gives users the freedom from remembering the multiple passwords and carrying the keys/tokens in pockets at all times [1, 30].

Probability of brute force attack on the biometric system

The probability of brute force attack on the biometric system is possible. However, the success of such attack is dependent on the matching accuracy of the biometric system being operated in the verification mode in a commercial application. Let us consider a scenario in which the application is working at 0.001% FMR. Under these conditions, most of the biometric systems are capable to deliver FNMR less than one percent (1%). These conditions can lead one to conclude that one of 100,000 brute force attacks are likely to succeed in fooling the system. This level of security is equivalent to that offered by the randomly chosen 5-digit PIN method which requires the 100,000 brute force attacks (50,000 on average) in having success to access the password protect system [1, 30].

The probability of success is further reduced by the stringent requirements of the biometric applications. They normally require thousands of the biometric sample to allow the hacker succeed in breaking the system, which is extremely difficult in comparison with generating thousands copies of the PINs and passwords. Further protection can be granted by lowering the FMR and raising the FNMR values simultaneously of the system at the cost of greater inconvenience to the users. Moreover, the inconvenience is always faced by users whenever an attempt is made to provide higher security to the application, for example, the long passwords can increase security but simultaneously increase the inconvenience of remembering and correctly typing them to have access the system [1, 30].

However, several commercial based applications tried to operate the system in the identification mode rather than in verification mode in order to reduce the inconvenience to the users. They do not require the identity claim from the users. The speed of such applications is the biggest challenge, and furthermore, the accuracy is even worse than the speed in some situations [1, 30].

Suppose a fingerprint based biometric application with 10,000 users operating in an identification mode. It requires the special hardware set up and the fast fingerprint matching algorithm to make identification within few seconds. If the FMR value of the application in verification mode is equal to 0.001%, then value of FMRN in identification mode will be equivalent to 10% (10,000 x 0.001%). This implies that if an imposter uses his/her 10 fingers on both hands for identification, he/she is more likely to succeed in gaining access to the system. This indicates that system is at greater risk of being manipulated by imposters [1, 30].

The small-to-medium scale applications involving few hundred users may use single biometric based identification system, while the large scale applications involving thousands of users need much more sophistication and security precautions in their design. The multi-modal biometric systems are the best options to design large scale applications. In multimodal biometric system, combination of two or more than two biometrics are utilized in order to execute the identification of the user. For example, the combination of index fingerprint and face or the fingerprints from two fingers in combination can be used in multimodal biometric system [1, 30].

2.2.3.2 Negative Recognition in Government and Forensic Applications

The negative recognition in government and forensic applications are designed to prevent the terrorists or hijackers from boarding plane or checking the employee’s background. This is executed by operating the application in an identification mode. It is real challenge to achieve the same level of accuracy with the identification mode of biometric systems as with the verification mode, primarily because of large number of comparisons are required to be performed with the multiple templates in the database [1, 30].

In order to understand this idea let us suppose a scenario in which Federal Bureau of Investigation (FBI) is searching for 100 most wanted criminals, which means that database size is 100. The operating conditions of a commercial biometric system working in verification mode includes 1% FNMR and 0.001% FMR. Given these operating conditions, the biometric system would fail to match the correct persons 1% time and would erroneously report the wrong person to be the correct person 0.001% time [1, 30].

Now consider the scenario when this biometric system is operated in the identification mode. The operating performance of the biometric system will be changed into FNMRN = 1%; FMRN = 0.1% (100 x 0.001%), which indicates that the biometric system has 99% probability to catch the right criminal out of the sample population. Suppose, if this biometric system is deployed as an identification system on US airport which enrols 200,000 passengers per day for the identification purpose, it is likely to generate 200 false alarms. Moreover, if face is used instead of fingerprints for the identification, the number of misses and the false alarms will spike considerably owing to the poor capability of the system to capture the high-quality face image in the cluttered environment of airport with the varying shades of light [1, 30].

This indicates that deployment of automatic biometric system with single biometric option for the negative identification is likely to create problems for the security staff at the airport due to generation of false results. Multimodal biometric systems can significantly improve the accuracy of negative recognition in biometric applications. Basswords and PINs can further jeopardize the security level of biometric systems designed for negative recognition, therefore, the traditional biometric applications are not suitable for the negative recognition [1, 30].

Despite the low efficiency and accuracy of the automatic biometric application for negative recognition, this is the only option for the large-scale identification applications. These automatic biometric systems can be made more effective by converting them into semi-automatic applications. The semi-automatic status can be achieved by hiring staff for manual examination of the false alarms generated by the system in order to make a final decision. For instance, only 5 FBI agents are required to manually examine the 200 false alarms by the system daily. In this way, the semi-automatic biometric application can help to catch the criminal by combining the automatic and manual examinations of the given templates with those stored in the database [1, 30].

Similarly, the application for the negative recognition of the background checks and forensic criminal identification can also be operated in semi-automatic and cost-effective fashion. For instance, in latent search mode, an Automatic Fingerprint Identification System (AFIS) can reduce the number of fingerprint matches executed by staff members from millions to hundreds, and forensic experts examines the data to give final verdict on it [1, 30].

Finally, it can be concluded that biometric applications for the negative recognition are the safest way to identify the required individual; and it does not infringe upon the civil liberties. For example, if the system holds a person’s data in the “criminal database”, it will be recalled, otherwise he/she will not be remembered by it. Furthermore, legislations in every country protect the people from the abuse of such system [1, 30].

2.2.4 Biometric Privacy Protection

One of the important challenges in the biometric application is the protection of the data of users from being attacked or from the being leaked to the third party which can manipulate the data to blackmail or threaten the users. The attacks on the privacy of the users in the biometric applications have been addressed by the legislation in different countries, so that the flow of data within the country or trans-border can be prevented to protect the data privacy of the users. These laws has been developed and enforced by the developed countries to thwart the privacy violations at different data processing phases in the biometric applications [160].

For instance, the legislation has been made in the European countries to prevent the storage of incorrect data of the users, the abusive or unauthorised data disclosure to the third part or unlawful storage of the data by any organization within the country. Different governing measures taken by European governing bodies such as OECD Guidelines, initiatives passed by “Council of Europe’s Convention of the Protection of Individuals” basically govern the privacy protection of the users of biometric applications. OECD Guidelines are related to violation of data disclosure to other countries. Similarly, EU issues tow important directives to prevent the violation of data privacy, which include 95/46/EC and 97/66/EC specifically designed for the tele-communication sector ((European Parliament and Council, 1995, 1997) [160, 164, 165]. Furthermore, the EU also issued another directive 2002/85/EC which prevent the privacy attack on the data of users from the internet-based applications (European Parliament and Council, 2002) [160, 166].

After carefully reviewing the data protections rules and regulations in different countries, it can be argued that there are some common elements in the framework built by the governing bodies concerning the protection of data privacy in the biometric applications. These common elements include the security, anonymity, data leakage, unlawful storage of data, accuracy of data and specification of the data and processing of data fairly for all the registered users to the biometric applications (Biometric Working Group, 2002)[160, 161]. The laws and legislations also permits the users to access the data without any hindrance and to edit the personal data by adding or deleting functions provided on the interface of data for each users. The further obligations have been imposed on the persons involved in data collection, data processing and data storing. Without authorization, the controllers of the data can not amend or replace any part of the personal data of users [160].

According the forging elements of the privacy framework, the controllers of the biometric applications are only allowed to collect the data with specific purpose which should be mentioned along with the file of data. The data should be adequate without having the excessive information regarding the users' personal history (CNIL, 2001) [160, 163]. The data which is declared as incomplete or inadequate need to be erased by the data controllers. Regarding the element of anonymity, the personal data of the users should not be disclosed to the third party, and should not be stored longer than it is required for the processing of certain information (Frankel, 2000) [160, 167].

According to the element of security, the data controllers are required to put in place reasonable technical, organizational and the legislative measures to prevent the unauthorised use of the user data, its disclosure to the malicious agents and occurrence of modification or destruction of the certain parts of the user's data in the biometric applications [160].

At the design level, the biometric systems are required to comply with the legislations and laws regarding the protection of the data by considering all the foregoing elements of the data protection framework. The controllers of the system should define the personal data for the data manipulation purposes (Prabhakar et al., 2003) [160, 169]. The personal data is defined as a set of data which enables the identification of the person through assignment of the special numbers, codes or signs to the person to which the data belongs. It may contain the record of physical injuries, information about the health, employment, physiological and mental characteristics (Matsumoto, 2002) [160, 170]. According to this definition, all the information retrieved from the users are stored in the form of users data templates which should be treated out of scope of the country’s legislation. The design of biometric applications is required to process the data anonymously in complaint with the privacy protection laws (Vaclav, M. and Zdenek,, 2000) [160, 171].

In short, the data protection issues can be resolved by making and complying with the legislations and laws which should be applied the architecture of the system and controllers of the biometric applications [160].

2.2.5 Multimodal Biometrics

In this system, two or more than two uni-modal biometrics are integrated into single system in order to make it multi-modal biometrics, which can be operated either in verification or identification mode. These systems are useful and flexible in terms of allowing the user to set the modality or change the modalities for the given sample required to be identified or verified [38]. There are several advantages of multimodal biometrics, some of them are listed below:

- They can increase the accuracy and performance of the biometrics systems significantly.

- They are flexible to accommodate the unusual traits which otherwise can not be processed by the uni-modal biometric system.
- They are highly resistant to the spoofing attacks because of difficulty in penetrating through the functions and modalities of the system compared with uni-modal biometric system.

Many approaches known as fusion exist for integrating multiple modalities. Fusion approaches can be distinguished in two ways. First of all, fusion could be carried out at different levels. Hong et al. categorized multi-biometric recognition systems into three architectures based on biometric data fusion [38]. Modalities could be combined at the feature level, the score level, or the decision level. Secondly, fusion could be based on rules or based on machine-learning approaches [39, 40].

The rule based approached adopted for the integration of multiple modalities into the biometric system include min score, max score, and simple sum [39, 41]. There are also machine learning approaches named for integration, which include vector machines, decision trees, multi-layer perception, Fisher's linear discriminate and minimum cost Bayesian classifier [39, 42]. In addition, the implementation of the multi-modal biometric systems can be carried out using the combination of biometrics and cryptographic concepts termed as 'Biometric-Cryptosystems'. The biometric-cryptosystems contain cryptographic technique which allow the integration and fusion of the multi-modalities instead of the fusion approaches employed for this purpose.

The fusion approaches require the calculation of weights of the modalities for the classification purpose, while the biometric-cryptosystems do not need to calculate these weights for all modalities, thereby allowing the flexibility to change or use any subsets of modalities. Therefore, the biometric-cryptosystems are more advantageous to perform the flexible functions of recognition than the fusion approach based multi-modal biometrics system. The various methodologies used for implementation of multi-biometrics system are described in section 2.2.6.

2.2.6 Biometric Cryptosystems

These are the biometric systems in which a cryptographic key binds monolithically with the template of the user stored in the system's database in such a way that key can not be revealed unless the authentication of the perfect match is declared [43]. The binding of the key with the template meets the requirements of security, distortion tolerance and discrimination which are defined as follows [16, 44, 45]:

Security: Security is promise of the biometric system that neither the key nor the template will be revealed to the hacker.

Discrimination: Discrimination is the capability of the biometric system that all the users enrolled on it are identified with their original identifies and different keys are issued separately to them.

Distortion tolerance: Distortion tolerance is the ability of the biometric system to tolerate the small variations in the input data at different times the same user claims his/her identity. Due to this ability, the biometric system is supposed to output the same key for the same user under different conditions and timings of the claim of identity.

The working principle of uni-biometric cryptosystem has been illustrated in the Figure 2.3.

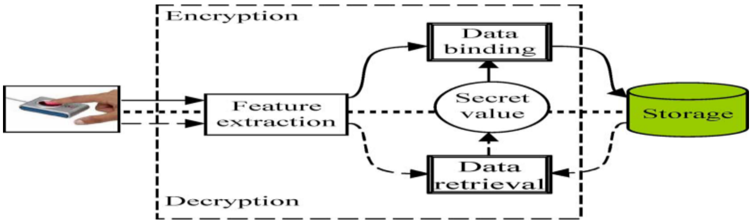


Figure2.3 Uni –Biometric Cryptosystem [16, 44]

Three approaches have been reported to implement the bio-cryptosystem using the binding mechanism of the key with the template, which include biometric based key generation method, biometric based key release method and biometric based key revocation [16, 44].

2.2.6.1 Biometric Based Key Generation

There are normally two established methods of generating the biometric keys from the sample: direct biometric key generation and the indirect key generation. In the direct key generation scheme, the biometric key is generated directly from the biometric samples without having requirement of matching with the template in the database (template free).

However, in indirect key generation scheme, the key is embedded into the template of the biometric data such as iris or fingerprints. Whenever the user wants to gain access to the system, the given biometric data is matched with the biometric template. On successful matching, the key is released. In other words, in this method, the key is generated based on the input biometric information provided to the system. The major challenge in this method is maintaining the entropy of the key and the security of the biometrics data simultaneously [16, 44].

The literature in this area is sparse until now. Handful of empirical evidence can be found in the area of biometric based key generation method. The proposed methodologies can be bifurcated into two categories: Error-coding approach and Shamir's key sharing scheme approach [11, 44, 46].

2.2.6.1.1 Error Correcting Code Scheme

In this scheme, the codeword and decoding functions are assigned to the templates of the user at the time of enrolment. The hash values assigned to codeword are employed directly as key or a seed to implement the key generation function. The biometric data stored in the database is used to calculate or retrieve the hash values or codeword at the time of authentication [44].

Typical error-correcting code based key generation methods are given below:

Biometric Encryption Algorithm (BEA)

Fuzzy Commitment Scheme (FCS)

Fuzzy Vault Scheme (FVS)

BEA utilizes the Fourier transform to match the fingerprint image to the template keys in the database, while FVS and FCS are normally applied to face, iris and fingerprints based biometric systems, and depend on the various unusual biometric traits to measure the closeness of the input data with templates, such as set distance and Hamming distance [38].

2.2.6.1.2 Shamir's Secret Sharing Scheme (SSS)

The SSS, as described in [11], is a well-known and commonly used cryptographic method, which when given a secret, divides the secret into a number of unique shares. The original secret can be retrieved by using a sub-set of these shares. This technique is also referred to as Shamir's *k*-*n* threshold scheme, whereby a secret is divided into *n* shares and *k* is the minimum subset size to retrieve the secret (i.e. at least *k* shares, from a maximum of n shares, are required to retrieve the secret). Being able to reconstruct the original secret from any subset of shares (provided that the subset is greater the k) is a very useful property and is an important aspect of this implementation [11]. A basic overview of Shamir's secret sharing scheme follows.

2.2.6.1.2.1 Secret Division

This scheme was invented by Shamir and Blakely in 1979, which is actually two threshold-based SSS method (k, n). The general logic used behind the construction of this scheme was to distribute the secret information among 'n' number of participants, and to unlock this information any combination of 'k' participants will be required. However, any 'k-1' participants trying to unlock the information would not be able to access the shared secret information. In order to define the set of 'k' participants, Shamir's scheme used the point method to define the mathematical lines and curves. For example, threshold any 2 points, 3 points, 4 points would be required to define the straight line, parabola and cubic curve, respectively. That is it takes k points to define a polynomial of degree k-1. Shamir's secret sharing scheme represents the secret as the coefficients of a $k-1$ degree polynomial [11]. This polynomial is determined using Equation 2.3.

$$f(x) = \sum_{i=0}^{i<k} x^i$$

$$a_i = \begin{cases} \quad i = 0, \; secret \\ i > 0, \quad rand(\,) \in Z_p \end{cases} \qquad (2.3)$$

Where:

Z_p denotes that the value is an element in a finite (Galois) field [0,p) where p is a prime number larger that a_0 and n.

Once the random polynomial is defined, the shares are determined by taking n points sequentially along the polynomial. This will create tuples of [x, *f(x)*] where x = [1, *n*). Therefore the shares [D_1, D_n) can be evaluated by using Equation 2.4 [11].

$$D_1 = f(1), \dots\dots, D_x = f(x), \dots\dots., D_n = f(n) \qquad (2.4)$$

Where:

D_i is the i^{th} share.

Again the secret shares are elements of a finite field [0, p) and are therefore modulo p. This provides tuples of [i, D_i]. Obviously D_0 is the secret value, which should not be public and is hence not a share.

2.2.6.1.2.2 Secret Reconstruction

In order to recover the original secret, interpolation must be performed. As mentioned previously, D_0 = a_0 = secret and therefore can be retrieved by evaluating the polynomial to find *f(0)*. To determine the polynomial and subsequently evaluate it, Lagrange's interpolation algorithm was implemented as it is one of the most common methods to reconstruct the secret. Lagrange's interpolation determines each coefficient of the polynomial from their respective shares to reconstruct the polynomial [11]. Lagrange's interpolation method is presented in Equation 2.3 and is employed in Equation 2.5 to evaluate the polynomial when $x = 0$.

$$D_0 = \sum_{i \in S} D_i \lambda_i (mod\ p)$$

$$\lambda_i = \prod_{j \in S, j \neq i} -x_j / x_i - x_j \qquad (2.5)$$

Where:

S is the subset of shares.

D_i is the *i^{th}* share (i.e. a tuple).

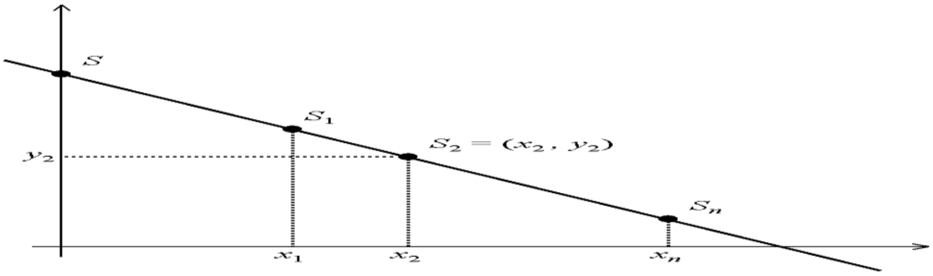


Figure2.4 Plot of SSS Algorithm [11]

Figure 2.4 shows the polynomial curve of the linear equation and the secret is the point where the line intersects with the y-axis. Namely, this point is the point *$(0,f(0)) = (0, D_0)$*. Each share is a point on the line. Any two points determine the line and hence the secret. With just a single point, the line can be any line that passes the point, and hence the secret can be any point on the y-axis. [11]

2.2.6.2 Biometric Based Key Release method

In this method, the biometric key is assigned to the template of the individual at the time of enrolment, which becomes stored permanently as a part of template in the system’s database. The key is hidden into a template in such a way that it can not be extracted without the perfect match of the template with the input biometric data. The keys are normally saved with template along with user name and privileged accessed to the system. Whenever, the user claims the identity, the biometric system, the key is released on successful authentication of the input biometrics information [43]. The illustration of this process is shown in the Figure 2.5.

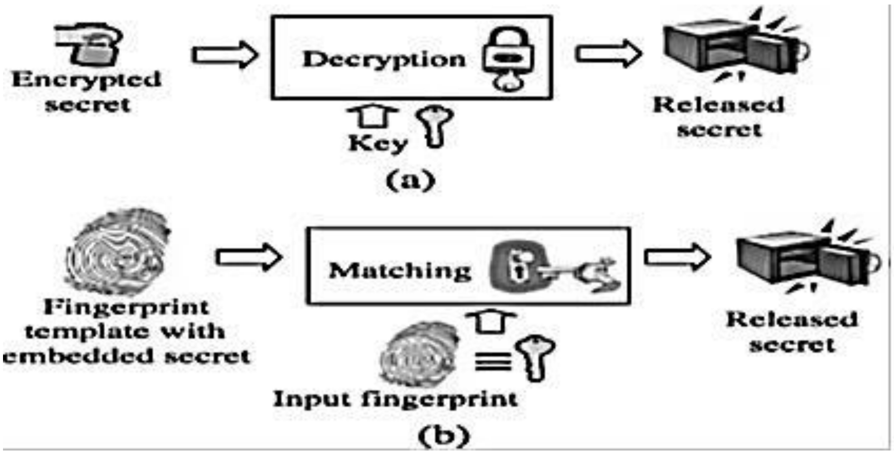


Figure2.5 Biometric Based Key Release System [43].

The main features of the biometric based key release system include:

1. The user’s access to the system so that biometric data can be matched to the templates stored in the database.
2. On successful match and authentication, the key is released.
3. Decoupling of the key release and the authentication of the user.

   The storage of transform version of the original templates in the database. This measure is important as system stores the templates locally, and it raises the security concerns such as higher probability of theft of templates from the database. In order to resolve this issue, the system stores the transform version of the template. In the event of theft, the new transforms are generated from the original templates as shown in Figure 2.6 [43].

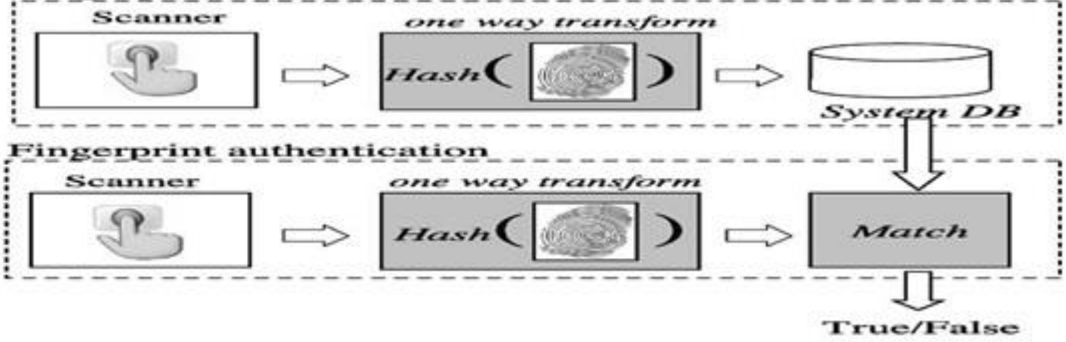


Figure2.6 Hashing Biometric Template [43].

2.2.6.3 Key Revocation

In cryptographic method, the security of the system is highly dependent on the security of the key. If the cryptographic key is comprised, then the whole system will become useless and compromised [172]. The cryptographic keys are generated in random fashion, as explained in the key generation section in 2.2.6.1. The users can nor remember the key generated by system because of length of key. If small keys are generated from the memory point of view, they will become vulnerable to attack. The smart card has been designed to remember these keys for each user. However, the chance of being lost or stolen, the smart cards were found to be insecure [173].

The development of biometric authentication system [10] gave a major impetus to the security of the users' data and made identity secure. The authenticator is linked with direct physical biometric features of the users. The further developments lead to the formation of cryptographic based system which creates a string link between the biometric data of the users and the generation of the cryptographic key [50]. The biometric features are used in two ways: protection of the cryptographic key or generation of the key from the biometric information provided by the users. In the first scenario, the biometric key is protected by applying the different mechanism used to construct the biometric keys such as fuzzy vault scheme [174] or fuzzy commitment [45]. The enrolled biometric information of the user will gain access to the key in the case of complete match [172, 175].

However, in second scenario, the users' biometric data is enrolled on trial basis to generate key, later on the message is encrypted against the key. The key itself is discarded later on. When the user will provide the similar biometric data, the system will revoke the same key against the biometric data after the encryption of the message [20, 130, 175].

Most of the cryptographic based systems face the problem of key irrevocability which is inherent to the biometric data of the users [7]. However the approach of revocable key is highly secured in the case of attack on the biometric system. Unlike the conventional biometric system the cryptographic based system with ability to key revocability can enhance the privacy protection as well as the security of the user's data [20].

In this work, the cryptographic based system is being developed which will work on the principle of key revocability on the presentation of users data to the system for multiple times. This is the novel feature of the proposed cryptographic system developed during this research project. The key could be revoked at the time of decryption of message from the presented biometric data such as iris and fingerprints data without revealing the true data regarding the biometric data.

2.2.6.4 Previous and Related Work in Bio-Cryptosystems

During the last decade, various researchers tried to resolve the issues related to the integration of the biometrics system into the bio-cryptosystems. Though the number of the published works in this field is low and are representative of all the modalities being utilized to build the biometric systems, but they still highlight the important challenges and possible suggestions to combine biometrics and bio-cryptosystems together [43]. The following list of the related work has been presented below, which has been carried out in the domain of bio-cryptosystems. The details of these research works will be presented in the section 2.2.7.1

- Image processing concept was used by Soutar et al [47] to perform biometric encryption.

- David et al [19] performed the error correction and cryptography using an offline biometrics-based identification scheme.
- The construction of ‘fuzzy commitment scheme’ by Jeules and Wattenberg [45].
- Clancy et al [22] developed secure smart-card based fingerprint authentication scheme.
- Manrose et al [21] generated the password hardening protocol using the keystroke dynamics.
- Hoh and Ngo [48] generated the personalized cryptographic key based on Hashing.
- Hao and Chan [38] performed the successful attempt to enhance the security and privacy in biometrics-based authentication systems.
- Dodis et al [44] developed the fuzzy extractors to generate the strong keys from biometrics and other noisy data.

In short, the research works described above reflect the handful approaches made in this field, which tried to use bio-cryptographic approaches to resolve the issues related to the biometrics based identification system. Furthermore, the important limitation observed with these research works was that they only used the short versions of the keys which were less secure due to their susceptibility to the potential attacks [7]. They also could tolerate only small amount of variations in the query biometrics sample and provide the narrow insight into the practicability of the offered solutions [7]. Though many researchers devised the strategies to bind the key to the template, but problems associated with the biometric domain were not resolved sufficiently [7]. Similarly, there are some challenges associated with the use of cryptographic techniques such as the stringent requirement for the exact match of the key with the template.

Consequently the products designed previously only utilized the tamper-resistant hardware for the security to the biometrics systems [49]. This demands more comprehensive protocol based techniques using the marriage of biometrics and cryptographic methods to generate the secured key binding to the template.

Subsequently, the two schemes: exception handling and Shamir secret sharing schemes were combined in this thesis (presented in Chapter 3) in order to implement the secure multi-modal biometric system.

The method used in this thesis for implementing the multimodal biometrics system is based on the hypothesis that biometric modalities can not necessarily be converted to binary sequence of data for generating the key shares as is the case in a biometric-based key generation technique, but instead can be used solely for the purpose of releasing the keys as in the case of a biometric based key release system. I have chosen to use the biometric based key release approach for a number of reasons which I will be discussing shortly. The biometric based key generation technique which is the traditional method of implementing Shamir's Secret Sharing Algorithm in a bio-cryptosystem has a number of disadvantages in contrast to the biometric-based key release method and this is evident from the previous work already carried out in this field which I discussed in section above.

2.2.7 Template Free Biometric

The previous researchers tried to find an interaction between the two important security technologies: biometrics and cryptography [50]. Biometrics holds a great importance in terms of capturing the various aspects of the biometrics samples provided by the users such as iris, fingerprints, voice, face etc. The combination of cryptography into biometrics allows the user to obtain a unique signature to access the data. The unique signatures are produced with higher degree of confidence and accuracy, which can not be forged or stolen like the PINs/token used by the traditional identification systems. The previous research works made attempts to alleviate some issues of security by attaching the IBM Transaction Security System or the signature verification pen to the unique signatures produced by bio-cryptosystems [8].

However, the major loophole of these techniques was their absolute reliance on the tamper-resistance property of the hardware to guarantee security. For instance, if the token is tampered, both template and the associated key will disappear, thereby resulting in the insecurity to the personal data of the users. The hackers can attack the system for breaking the tokens' security level using various methodologies such as API attacks on the software of the token normally associated with IBM design and exploitation of chip-testing technologies [51].

Therefore, an attempt has been made to evolve more reliable combination of three technologies including biometrics, cryptography and tamper-resistance hardware. Nevertheless, the choice of algorithm is a challenging task, as they can increase the background noise of the biometric data which is used to give approximate match to the template data stored in the database.

The users have often expressed their reluctance to enrol on the biometric system, primarily because of their concerns that system's database holding their personal data is liable to attacks, opening some possibility of leakage of personal data of users to the malicious agents. This has limited the implementation of biometrics system as a favourite choice for several organizations working in both public and private domains of the society. However, if the users are provided with guarantee that their personal data will not be stored in the central database of the biometrics-based identification systems, the implementation of these security technologies can find their way into almost every governmental or commercial organization [50, 51].

This demands a research endeavour aimed to design the template-free biometric systems without having requirement of storage of personal data of users within the central database. The previous attempts made in this context utilized the mapping technique to convert the biometric data into the repetitive binary strings [20, 21, 52-54]. Subsequently, these binary strings were used to design the encryption key which in turn was employed to open the direct hashing [54, 55] or look-up table [20, 52, 53]. Despite the inherent flaws and drawbacks of such methods, they offered a bright prospect of generating the template-free biometric system.

The one of the core issues associated with the template-free approach is the high level of unreliability related to the individual bits of template data, which is further exasperated by a noise background generated by biometrics attributes during their measurements. However, the requirement of key demanded by the cryptographic method was another obstacle in the way of creation of template-free biometrics system. In order reduce the noisy background and exactitude demanded by the cryptographic method [21], many attempts had been made to derive the biometric key using key-stroke patterns [21] and various biometrics such as voice [20], fingerprints [52, 56], facial characteristics [53] and hand-written signatures [54].

Nevertheless, the resulting system produced the 20% FRR during verification stage, which reduced the feasibility and practicability of implementation of such systems in biometrics based identification applications. The biometric data contain some unchangeable attributes specific to the individual, therefore, they are resistant to changes in them. Key diversity is another challenge in the way of developing template-free biometric systems, as multiple keys are required to access various attributes of biometric data. The openness of the biometric data also raises some security concerns, for example poor quality images of the iris and fingerprints can surreptitiously obtained by hidden cameras, which can be used by attacker to steal the personal information of the users. Thus the greater level of reliance on the biometric system is more likely to increase the security concerns of both users and administrators [51]. Furthermore, the biometric data can be spread globally due to travelling of the individual to various countries.

Given the public mistrust about the biometric systems and the foregoing issues associated with it, the current study is aimed at the derivation of biometric data from iris by addressing the associated issues with this process, as iris represents the most authentic source of biometric information specific to an individual with its greatest property of distinguishability. Previous study carried out by Davinda and his colleagues [57] used error correction codes methodologies in order to derive the key from the iris code. Daugman et al [50] made a similar attempt to construct the template-free biometrics system by utilizing only the string of error correction data which could be used to acquire the biometric information without derivation of access key. They not only designed the two factor scheme to test the identification of token and biometric data but they also extended this scheme to three factors by including the password as a third factor. The drawback of their two and three factor scheme was that the secret key was embedded in the smartcard which was also susceptible to attacks launched by hackers.

2.2.7.1 Previous and Related Work in Biometric Template Free

In this section, the detailed survey of the previous work related to the construction of template free biometric system is presented [16, 20, 21, 52-54]. Monrose et al [21] conducted a pioneering research work to build a key extraction model using the keystroke dynamics. They derived the binary string from the typing pattern of the person, and subsequently combined it with the user’s password to extract a hardened password. In the model, each bit corresponded to the discrete keystroke property, which accommodated a bit-feature variation by allowing the exercise of error tolerance property to some extent. Furthermore, they applied the concatenation method to generate the short strings of bits. In the follow-up research work, they improvised upon their previous model and constructed the reliable voice biometrics system by applying the discretization method [20]. The major improvements made by them in their feature extraction model included an increase in entropy of biometric key from 12 bits to 46 bits, a fall in FFR from 48.4% to 20% [20].

Nevertheless, another biometric system based on the hand written signatures were developed by Hao and Chan, which produced 43 different hand written signature features defined by taking into consideration of various attributes of the users such as azimuth, velocity, pressure and altitude. Bits were produced for each feature using the feature coding method, and were further used to create a binary string. The salient features of their model of key extraction included an achievement of 40-bit key entropy, 1.2% false acceptance rate, 28% FRR [54].

Fingerprints based biometric system is recognized as one of the historical and mostly used biometric system for the identification of criminals to solve the criminal cases [50, 51]. The pioneering work has been conducted by Soutar et al who developed the first commercial fingerprints biometrics system. They applied Fourier transform and majority coding techniques for the extraction of the phase data from the fingerprints sample and reduction of the feature variation element, respectively. Notably, unlike following the approach of assigning key to the fingerprints sample, they locked the sample with a random key embedded within it, resulting in the creation of phase-phase product. The salient feature of their fingerprints based biometric system was that it required the presence of another genuine sample for the authentication purpose, which revolutionized the world of biometric technology by providing the facility of assignment of random keys to the fingerprints samples taken from the users. However, this system could not be supported by good quality performance data, which restricted its widespread applications.

Similarly, Clancy et al [52] carried out research work on the generation of key from the fingerprints. They recorded the minutiae points of the fingerprints as real points, and created locking set of these points. Furthermore, they derived the locking key by applying the binomial reconstruction method on the locking sets, and hidden it by adding chap points in the locking key. The key was recoverable when the given sample of the fingerprints overlapped substantially with the locking set with the aid of application of the Reed-Solomon code. The peculiar feature of this model was the development of key with 64-bit entropy key, however, it yielded the outcome with 30% FRR code which acted as an obstacle in its wide-spread application.

Goh and Ngo [53] utilized the concepts derived from Soutar et al [16] and developed the face biometrics based identification system. They extracted the 18 projections from the face image by using the biometric locking approach of Soutar et al [16]; and subsequently they generated the single bits for each projection. Ultimately these bits were reduced to the strings by applying concatenation approach, which were further used to extract the binary code. The system was supplemented with the majority coding by applying approach devised by Davinda et al [57]. Moreover, they made use of the error correction approach involving polynomial thresholding in order to reduce the effect of feature variation. Interestingly, the resulting face biometric system demonstrated the 80-bit key entropy and 0.93 FRR. However, they used the video camera to take images in continuous fashion rather than face database, and that it also caused the decrease in the feature variation. This left a question mark on the validity and implementation of system for face images taken from discontinuous fashion. In short, several researchers made attempt to devise the biometric systems with ideal FRR, but all of these works ended up with having 20% FRR with the exception of Goh and Ngo work, which was intolerable level to implement the system practically. The limitation of above works was the length of key which was very short and susceptible to attack.

Furthermore, some theoretical knowledge can also be found in the literature in relation to extraction of key by reducing the fuzzy background and feature variations. Also, the suggestion of application of fuzzy extractor is made by [49] to extract the strong keys from biometrics with noisy data. Dodis et al [49] applied this approach in order to extract the key from the noisy data using the error-correction code and hash function on the sample input data. They reported that leakage of the data from the sample input was negligible after the application of hash function. Hence, this approach is suitable for the biometrics data which involved high noisy data, and required to much hidden for the further processing.

Byron [60] introduced some modifications in the fuzzy extractor scheme in order to apply the fixed permutation to iris before the application of hash function. The major utility of the Byron scheme was that different keys were derived from the biometric samples by using the different permutations, so if one key was promised, it did not affect the integrity of the other biometrics keys belonging to a particular biometric sample. Nevertheless, the drawback of this technique was not able to protect the data in the case of stealth of data. Another important research work was carried out by

Jeules and Wattenberg [45] who suggested the third part theory which gave the biometric key a separate domain independent of the biometric data by introducing the step of allocation of random key into the fuzzy commitment scheme and addition of XORs and redundancy to the iris code.

Similar research work was carried out by Duagman et al who introduced some modifications into the Jeules and Wattenberg approach by inclusion of concrete coding scheme which gave the best outcomes. Moreover, they also generated a surplus secret password by using the token-like tamper-resistant smartcard; and they also generated the 140-bit entropy biometric key [50].

Due to these issues, therefore, we intended to carry out the study to design such a template-free biometric system in which biometric data can be mapped onto the repeatable unique binary codes/strings which will only open up the key in the presence of biometric prints. Moreover, in order to generate the secret key, Shamir secret scheme [11] will be applied in three points provided (Right iris, Left iris, and password). Hence, 2 points out of 3 are sufficient to release the biometric key by using linear equation technique and that gives the user the flexibility which will reduced the FRR. We propose the three-factor scheme including the smart card, password and biometrics. In this scheme, each component holds the crucial position to complement the system required for revealing the biometric data specific to an individual with opportunity of either updating or revoking the key; and all three factors would be required to compromise the integrity of key. In addition, we tried to achieve the goal of development of scheme which provided the flexibility in terms of creating the length (short and long keys) of biometric key; and furthermore, in the event of smartcard loss/stolen, there would not be much information available for attacker to access the information because only Y-coordinates are stored in the smartcard. Finally, the scheme aims to reduce the FRR level of the system to increase its practicability in real life situations and this scheme will be introduced in Chapter 5.

2.2.8 Performance Evaluation Strategies

Two samples of the same biometric characteristic from the same person (e.g., two impressions of a user's right index finger) are not exactly the same due to imperfect imaging conditions (e.g., sensor noise and dry fingers), changes in the user's physiological or behavioural characteristics (e.g., cuts and bruises on the finger), ambient conditions (e.g., temperature and humidity) and user's interaction with the sensor (e.g., finger placement). Therefore, the response of a biometric matching system is the matching score, $S(X_Q, X_I)$ (typically a single number), that quantifies the similarity between the input and the database template representations (X_Q and X_I, respectively). The higher the score, the more certain is the system that the two biometric measurements come from the same person. The threshold time factor '*t*' controls the decision of the biometric system. The biometric sample which generate the matching score equal to or higher than '*t*' are regarded as mate pairs – belonging to the same individual – and similarly those biometric samples which produce the matching score lesser than '*t*' are declared as the non-mate pairs [30]. The distribution of score values generated from the samples given by the same person is called the genuine distribution and those from different individuals is termed as the imposter distribution (Figure 2.7 a) [30].

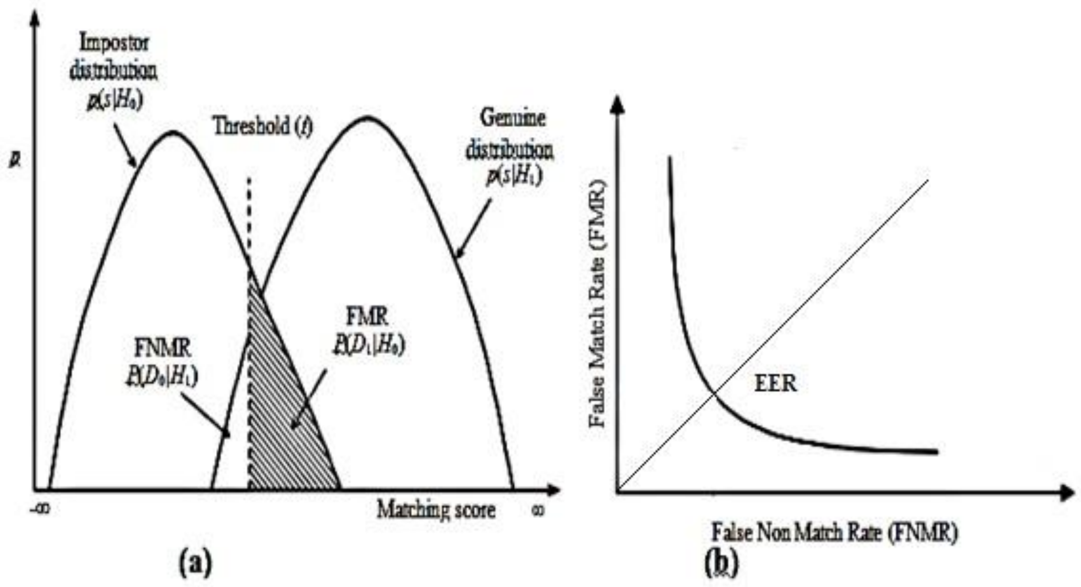


Figure2.7 Biometric system error rates [30].

Figure 2.7 shows the biometric system error rates: (a) FMR and FNMR for a given threshold *t* are displayed over the genuine and impostor score distributions; FMR is the percentage of non-mate pairs whose matching scores are greater than or equal to *t*, and FNMR is the percentage of mate pairs whose matching scores are less than *t*. (b) Choosing different operating points results in different FMR and FNMR. The curve relating FMR to FNMR at different thresholds is referred to as Receiver Operating Characteristics (ROC). Typical operating points of different biometric applications are displayed on an ROC curve. Lack of understanding of the error rates is a primary source of confusion in assessing system accuracy in vendor/user communities alike.

During biometric verification, the system may show two kinds of errors: *False match error* and *False non-match error*. The false match error is displayed when the biometric system mistakenly shows that the two biometric measurements taken from the different individuals belong to the same person. This is also called false accept error. Conversely, when the biometric system mistakenly recognizes that two measurements from the same individual belonging to the different persons. This kind of error is also called false reject error. Every biometric system makes some trade-offs between the false match rate (FMR) and the false non-match rate (FNMR) [30].

The performance of the biometric system taken at the various points of the threshold 't' functions is demonstrated in the form of curve called a Receiver Operating Characteristic (ROC) Curve. Hence, the ROC is obtained by drawing the curve between FMR and FNMR functions for different values of threshold 't' (Figure 2.7 b) [30].

Mathematically, the error in a verification biometric system can be calculated by the following equation:

$$FMR = \int_{t}^{\infty} P\big(S(X_Q, X_I)\big|H_0\big)dS, \qquad (2.6)$$

$$FNMR = \int_{-\infty}^{t} P\big(S(X_Q, X_I)\big|H_1\big)dS, \qquad (2.7)$$

Where

S = The users whose identity need to be verified

X_1 = The stored biometric template

X_Q = The acquired input for recognition

H_0 = The input biometric measurements from the person do not belong to the same person as X_1 template does

H_1 = The input biometric measurements from the person belong to the same person as X_1 template does

dS = distribution $P\big(S(X_Q, X_I)\big|H_0\big)$

Besides the above errors, there are some other error rates which are utilized to evaluate the accuracy of the system including Failure to Acquire rate (FTA) (also known as Failure to Capture rate (FTC)) and Failure to Enrol rate (FTE) [30].

The FTC is applicable to only those biometric systems which contain the automatic image capture facility. The Failure to capture error happens when the system rejects the poor quality image such as the faint fingerprints, occluded face etc. The increased FTE rate shows the increased level of components of the perceived system's accuracy: FMR and FNMR. Similarly, the FTE occurs when the system fail to enrol the person for the identification due to poor quality of biometric measurements. Thus the system quality and accuracy can be increased by increasing the perceived components of the system's accuracy (FMR, FBNR). The improvements in the FMR and FNMR, consequently results in higher rates of FTE and FTC. In this way, the all of the four rates – FMR, FNMR, FTE and FTC – constitute important specifications of any biometric system, and they are normally considered during the performance evaluation of the system [30].

The accuracy of the biometric system in the identification mode can be obtained by taking the consideration of the system in the verification mode under the following assumptions [30]:

- The number of attempt for identification is made only once
- A single biometric template is used to identify a specific person
- The number of identities present in the biometric system is equivalent to N
- Identification false match rate error = FMR_N
- The identification false non-match rate = $FNMR_N$

In the light of the above assumptions, the following approximations can be written:

$\text{FNMR}_N \cong \text{FNMR}$ and $\text{FMR}_N = 1\text{-}(1\text{-FMR})^N \cong \text{N}\cdot\text{FMR}$ (the approximation hold good only when N.FMR<0.1). A detailed discussion on these issues is available in [32] and [33].

For a biometric database in which each entry is classified and indexed properly, the system searches only a small proportion of the database to search for the identity requested by the user. For this, the following formulations can be written:

$\text{FNMR}_N = \text{RER}+(1\text{-RER})\cdot\text{FNMR}$, where RER (Retrieval Error Rate) is the probability that Where

RER = the probability of the template in the database matching with the input biometric measurements (fingerprints) is wrongly rejected. This happens when the biometric system is unable to retrieve the right template from the database, thereby generating the false non-match outcome.

1-RER = the probability of the template in the database matching with the input biometric measurements (fingerprints) is retrieved

FNMR = False non-match rate of the biometric system

The above expression does not show the frequency of non-matches of the templates with given biometric measurements before the right match is made between the input and the stored template in the database [34].

$\text{FMR}_N = 1\text{-}(1\text{-FMR})^{N\cdot P}$; where P (also called the *penetration rate*) is the average percentage of database searched during the identification of an input fingerprint.

The requirements of the biometric system for accuracy and consistency vary from system to system. Some biometric system such as forensic biometric system requires great stringency for FNMR as the authorities show much concerned to identify the right criminal for a specific crime, although they go to the extent to manually examine the potential non-matches generated by the system. On the other hand, for the accuracy of the biometric system designed to control the access of the people to a certain secure facility deepens highly on maintaining high FNMR in the system, even though the legitimate users sometimes face inconvenience due to this high value. However, there are several other civilian applications whose performance and accuracy lie in between FMR and FNMR, therefore, both parameters are considered during the design of such biometric system [30].

For instance in some application used by banks such as the bank card verification, the high FMR means the high probability of false match from various persons' biometric measurements with the same person template and subsequently the loss of hundreds and thousands of the dollars while the high FNMR will likely cause the loss of a valued bank customer. The trade-offs between FNMR and FMR in the various kinds of biometric applications can be observed in the Figure 2.7b.

2.2.9 Data Source and materials

A number of biometric samples (preferably face, fingerprint and iris) are required to establish the identity of system users. Data experiments for fingerprints used the FVC2004 database, which is the Third International Fingerprint Verification Competition database [91]. For iris recognition, the CASIA-IrisV3 database was used, as created by the National Laboratory of Pattern Recognition (NLPR), the Institute of Automation (IA) and the Chinese Academy of Sciences (CAS) [92]. This database was selected because it contains 22034 images taken from 700 individuals collected from 1500 iris samples. For this research the images located in the CASIA-IrisV3-Interval section has been used because it contains the images collected from the two different sessions, which means that variations in the iris images between two sessions can easily be spotted and recognized by using this part of the iris images data [92]. For face recognition, the FERET (Facial Recognition Technology) database was used [93]. No special hardware or software equipment is required for implementation. Software requirements for developing the system are open to interpretation, and any software development tool or programming language can be used (e.g. Matlab).

2.3 Summary

In this chapter a survey of the literature regarding the biometric systems, the history of the biometric system was presented. It was discussed that password methods for the identification of the individual were prone to the attacks from the hackers. The characteristics of the biometric systems were explained with the view of highlighting the importance of the biometric systems in the identification of the individuals. The authentication and verifications functions with their mechanism were included in this chapter. All biometric systems share the following modules in their designs: sensor, feature extraction, matcher and system's database modules. The biometric system's errors were presented. It has been discussed that these errors may downplay the performance of the biometric systems. Further research is required to address the errors for improving the performance of the system. Furthermore, the comparison between various biometric systems showed that various human individualistic properties such as face, voice, fingerprints, iris etc. can be used to develop the different biometric systems. Biometric cryptographic methods have been discussed, and it was found that they can provide better security and privacy to the user’s data compared to the conventional biometric systems. The template free biometric system can provide better security as hackers can manipulate with templates stored in the database, resulting in better performance and strong security and privacy features. The indirect key generation using Shamir Secret Sharing have the great promise to provide the better security. Therefore, the next chapter will present the indirect key generation using Shamir’s secret sharing key.

Chapter 3.

Indirect Key Generation Using Shamir's Secret Sharing Scheme

3.1Introduction

There are multiple challenges faced by the biometric system which are mostly related to the security of biometric data stored or captured by the system. Biometric systems are susceptible to attacks from the malicious agents due to the loopholes in the security of the systems arising from lost or stolen keys mainly due to small length of keys. Several researchers tried to devise a various schemes by integrating the biometric system into the bio-cryptosystems in order to secure the biometric data stored within the system such as fuzzy commitment scheme [45], smart-card based fingerprint authentication scheme [22], password-hardening protocol, the personalized cryptographic key based on Hashing [38] etc. Though, these approaches improved the security to considerable extent by increasing the rejection rates at the cost of performance and key entropy using more than one biometrics, but the important limitation observed was that they only used the short versions of the keys which were less secure due to their susceptibility to the potential attacks [9]. Moreover, they could only tolerate the small amount of variations in the query biometric sample, which hindered their practical applications. Against the background of these challenges, the researcher endeavoured to devise a scheme for indirect key generation using Shamir's Secret Scheme within a multi-modal biometric system.

In comparison with previous approaches in bio-cryptosystems, the proposed scheme will strengthen the security and performance of the multimodal biometric system, with its ability to derive an encryption key indirectly from biometric samples provided by a given user, using an exception-handling scenario. In such a system, no copy of the encryption key would need to be retained, and not all biometric samples of the related biometrics would be required, so enhancing both the system's security potential and its practical usability. In the proposed scheme, an arbitrary subset of biometric modalities needs to be supplied by the user, thereby addressing those situations in which a given modality cannot be supplied. Further, an arbitrary (secret) encryption key may be associated with the scheme, which readily addresses the revocability issue when a key becomes compromised and allows for multiple independent encryption keys to be derived from the same biometric samples [61]. The proposed scheme will also confer the three performance levels based on Shamir's Secret Share with different associated levels of security, which will enable the system's administrators to tailor the multimodal biometric system according to their own requirements.

In this chapter, the proposed structure of the scheme is elaborated in detail. Furthermore, it will also introduce the modalities chosen to test the proposed scheme, and selection and justification of approaches used for the biometric feature extraction.

3.2Generation of Encryption Key using a Multimodal Biometric System

The single biometric system that used one biometric modality requires the query biometric sample to process the key, if the user fails to provide the exact match, the system will not be able to identify the user. Similarly, the multi-modal biometric systems using various schemes have been previously constructed by different researchers; however, they faced the similar problem of the rejection of identity claim by the system in the case of unavailability of a single modality out of multimodalities. In order to resolve this issue to increase the performance of the system, The researcher investigated a technique for the extraction of a secret key from a multimodal biometric system using Shamir’s SSS algorithm. Secret sharing is an established technique for protecting a given item of data by means of information distributed to several participants. The technique plays a fundamental role in protecting data or secret information from impacts such as loss, spoofing or theft. Shamir’s SSS is a cryptographic method devised by Shamir and Blakely in 1979 [3], which divides the secret into parts and gives each unique part to a distinct participant. To reconstruct the secret, a predetermined subset of these components is required [1–4].

However, the main challenge faced the researcher was employ five biometric modalities—face, index finger, thumb finger, right iris and left iris—to generate the points on the required polynomial while meeting the following criteria.

1- The length of the biometric encryption key is unlimited.

2- The number of successful biometric inputs required during the authentication process to release the secret shares is flexible, depending on which of the following versions of Shamir’s style has been used:

    - Linear equation technique (needs at least two successfully biometric modalities out of five to release the secret shares for biometric encryption key extraction purpose).
    - Quadratic technique (needs at least three out of five).
    - Cubic technique (needs at least four out of five).

However, the researcher was able to overcome this challenge to considerable extent, which will be described in detail in the later sections of this chapter. Another issue was faced in the area of the biometric feature extraction, which might have been caused by two factors: the feature extraction method and the low quality biometric sample images. However, the following methods have been used to extract the biometric feature for each modality to generate the encryption key indirectly.

3.2.1 Feature Extraction approaches for Face

Face is considered to be one of the vital method methods used to identify the persons based on the recognition of their facial features, The accuracy of face localization, extraction and tracking of facial features are the critical factors for the implementation of applications like animation, face-based human identification, human-machine interaction (vision-based). As a result of last 30 years, various methods and approaches have been devised to extract the facial features efficiently and effectively. These approaches can be categorised into three main groups. The first of these is the holistic approach, in which the whole facial region is used as an input to the system (e.g. Eigenface [69, 70], Fisherface [157], WISARD [158] etc.). In the second, feature-based approach, only local features of the face such as nose, eyes and so on are segmented and then used as inputs for structural analysis (e.g. hidden Markov model). Finally, the hybrid approach uses both local features and the whole face [64, 65].

From these three approaches, the Eigenface based approach has been applied for the facial feature extraction. The researcher drew on the previous research work for the implementation, evaluation and calibration of this work [69, 70]. Eigenfaces represent the set of vectors used in computer based facial recognition. Furthermore, this approach has been chosen due to the following reasons:

- There is no requirement to identify or locate facial components (eyes, lips etc.).
- Feature size can be reduced significantly with minimal loss of information.

- Partially occluded faces can be reconstructed using Eigenfaces.
- It is an efficient approach in terms of its speed, simplicity and learning capability.

The subsequent sections describe the detailed analysis of Eigenfaces based feature extraction approach.

3.2.1.1 Eigenface Approach

The major goals of face recognition using the Eigenfaces method is to extract the salient features from the face image, encoding the features and compare them with the templates stored in the system's database. This approach is one of the well investigated methods used for face recognition. It is also called eigenpicture, eigenvectors, principal component and Karhunen- Loeve expansion.

Previous work on face recognition ignored the issue of face stimulus and assumed that predefined measurements were relevant and sufficient. This suggests that coding and decoding of face images should emphasise the significant local and global of features, which may or may not be related to facial features such as eyes, nose, lips and hair. However, they might have great implications for the final identification of the facial features [63-70]

The objective is to extract the relevant information in a face image, encode it efficiently and compare one face encoding with a database of similar encoded faces. One simple approach for extracting the information content from an image of a face would be to somehow capture the variation in a collection of face images [64, 67], and ultimately using this information for encoding and comparing the individual face images with the template facial image.

Mathematically, the aim is to locate the principal components of the distribution of facial images, or the eigenvectors of the ‘covariance matrix’ of the collection of the facial images, where each point/vector is treated as an image. Each eigenvector represents the variation among the set training facial images.

Any eigenvector is basically representative of ‘set of features that together characterize the variations between the set of training facial images. This shows that eigenvector is made up of the less or more contributions from each facial image, resulting in the presentation of eigenvector in the form of ghostly face image which is called ‘eigenface’ [69, 70]. The illustration of eignfaces is presented in Figure 3.1. Together, Eigenfaces provide a map of variations among the set of training facial images.

Thus each individual is recognized by a set of eignepictures required to reconstruct and describe his/her face, which is extremely compact form of presentation in comparison with the images themselves.

This approach involves the following initialization the following procedures for the face recognition [69, 70]:

1. The training set of images from the individual is captured.
2. The calculation of Eigenfaces from the training images is performed and only the *M* images having the highest eigenvalues are selected for further processing. The face space is defined from the collection of *M* images to obtain the feel of new faces. As a result of appearance of new faces, the Eigenfaces can be recalculated or updated.
3. For each known individual, the corresponding distribution in M-dimenstional weight space is calculated their projection of their faces on to the “face space”.

Following the initialization steps, the following steps are required to recognize the new facial images [69, 70]:

1. The set of weights from the input facial image and the M Eigenfaces is calculated by projecting the input facial image onto each of the facial image.
2. The closeness of the weights to the "face space" is determined in order to decide upon whether the image is face.
3. If the image is declared as face image then the weight patterns are classified to be known or unknown face.

Each face image can be represented exactly in terms of a linear combination of Eigenfaces. The number of possible Eigenfaces is equal to the number of face images in the training set. The faces can also be approximated by using 'best' Eigenfaces, which have the largest Eigen values and therefore account for most of the variations within the set of images. The primary reason for using fewer Eigenfaces is computational efficiency [69, 70]. The next step is to calculate the Eiginface which is presented in the subsequent section.

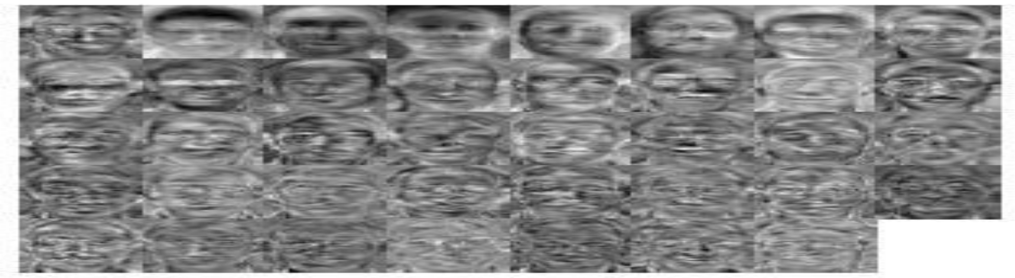


Figure3.1 Eigenfaces [69, 70].

3.2.1.2 Calculation of Eigenfaces

Training set of m images of size N x N are represented by vectors of size N^2. Each face is represented by Γ_1, Γ_2 , Γ_3,… Γ_M. The feature vector of a face is stored in an $N \times N$ matrix. Now, this two dimensional vector is changed to one dimensional vector.

For example:

$$\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} = \begin{bmatrix}1\\ 2\\ 2\\ 1\end{bmatrix}$$

Each face image is represented by the vector Γi.

$$\Gamma 1 = \begin{bmatrix}1\\ -2\\ 1\\ -3\end{bmatrix}, \Gamma 2 = \begin{bmatrix}1\\ 3\\ -1\\ 2\end{bmatrix}, \Gamma 3 = \begin{bmatrix}2\\ 1\\ -2\\ 3\end{bmatrix} \ldots\ldots\ldots . \Gamma \mathrm{M} = \begin{bmatrix}1\\ 2\\ 2\\ 1\end{bmatrix}$$

Average face image is calculated by

$$\Psi = (1/M)\sum_{i=1}^{M} \Gamma i.$$

$$\begin{bmatrix}1\\ -2\\ 1\\ -3\end{bmatrix} + \begin{bmatrix}1\\ 2\\ -1\\ 2\end{bmatrix} + \begin{bmatrix}2\\ 1\\ -2\\ 3\end{bmatrix} + \cdots \ldots\ldots\ldots\ldots . + \begin{bmatrix}1\\ 2\\ 2\\ 1\end{bmatrix} \rightarrow \begin{bmatrix}-1\\ -1\\ 2\\ -3\end{bmatrix}$$

$$\Psi = (\Gamma 1 + \Gamma 2 + \Gamma 3 + \cdots + \Gamma_M)/\mathrm{M}$$

Each face differs from the average by $\Phi_{\mathrm{i}} = \Gamma_{\mathrm{i}} - \Psi$ which is called the mean centred image.

$$\Phi 1 = \begin{bmatrix}2\\ -1\\ -1\\ 0\end{bmatrix}, \Phi 2 = \begin{bmatrix}2\\ 4\\ -3\\ 5\end{bmatrix}, \Phi 3 = \begin{bmatrix}3\\ 2\\ -4\\ 6\end{bmatrix} \ldots\ldots\ldots\ldots\ldots . \Phi_M \begin{bmatrix}2\\ 3\\ 0\\ 4\end{bmatrix}$$

A covariance matrix is constructed as:

$C = AA^T$, where A = [Φ1, Φ2, .., Φ_{M}] of size $N^2 \times N^2$

$$A = \begin{bmatrix}2 & 3\\ -1 & -2\\ -1 & 1\\ 0 & 2\end{bmatrix} \quad A^T = \begin{bmatrix}2 & -1 & -1 & 0\\ 3 & -2 & 1 & 2\end{bmatrix}$$

Size of covariance matrix will be $N^2 \times N^2$ (4 × 4 in this case). Eigen vectors corresponding to this covariance matrix is needed to be calculated, but that will be a tedious task therefore for simplicity we calculate A^TA which would be a 2 × 2 matrix; in this case, $A^TA = \begin{bmatrix} 6 & 7 \\ 7 & 18 \end{bmatrix}$ and the size of the matrix is $M \times M$. Consider the eigenvectors v_i of $A^T A$, such that $A^T AX_i = \lambda_i X_i$. The eigenvectors vi of A^TA are X_1 and X_2, or 2 × 1. Now, multiplying the above equation with both sides, we get

$AA^T AX_i = A\lambda_i X_i$

$AA^T (AX_i) = \lambda_i(AX_i)$

With reduced dimensionality, eigenvectors corresponding to AA^T can now be easily calculated, where AX_i is the eigenvector and λ_i is the Eigen value.

The eigenvectors of the covariance matrix AA^T are AX_I, which is denoted by U^i. U^i resembles facial images that look ghostly and are called Eigenfaces. Eigenvectors correspond to each Eigenface in the face space and discard faces for which Eigen values are zero, so reducing the Eigenface space to some extent. Eigenfaces are ranked according to their usefulness in characterising the variation among the images. A face image can be projected into this face space by the equation

$\Omega_k = U^T (\Gamma_k - \Psi); k=1,....,M$, where $(\Gamma_k\ \Psi)$ is the mean centred image.

Hence, a projection of each image can be obtained as Ω_1 for projection of $image_1$ and Ω_2 for projection of $image_2$, and so on.

The test image, Γ, is projected into the face space to obtain a vector, Ω as

$\Omega = U^T (\Gamma - \Psi)$

The distance of Ω to each face (called the Euclidean distance) is defined as

$\epsilon_k^2 = ||\Omega - \Omega_k||^2; k = 1,,M$, where Ω_k is a vector describing the k^th face class.

A face is classified as belonging to class k when the minimum ϵ_k is below some chosen threshold Θ_c otherwise, the face is classified as unknown. Θ_c is half the largest distance between any two face images:

$$\Theta_c = (1/2)max_{j,k} \, ||\Omega_j - \Omega_k \, ||; \, j,k = 1,.....,M$$

We have to find the distance between the original test image Γ and its reconstructed image from the Eigen face Γ_f

$$\epsilon^2 = ||\Gamma - \Gamma^f \quad ||^2, where \, \Gamma^f = U * \Omega + \Psi$$

If ≥ Θ_c then the input image is not even a face image and is not recognised.

If < Θ_c and $k \geq \Theta$ for all k then the input image is a face image but is identified as an unknown face.

If < Θ_c and k < Θ for all k then the input images are the individual face images associated with the class vector Ω_k.

3.2.2 Feature Extraction Approaches for the Iris

There are various available methods for performing feature extraction, including wavelet encoding, Gabor filters, Log-Gabor filters, Haar wavelet and the Laplacian of Gaussian filter. For present purposes, the 2D Gabor filter approach was chosen for extracting the iris features. This approach has been selected because it efficiently represents the local texture features of the iris with zero-crossing representation [71, 74].

However, the process of feature extraction from iris involves a number of steps which are localization/segmentation, iris normalization/unwrapping [71-74]. The detailed discussion on these steps will be provided in Chapter 5. In this section, the focus is placed on the iris feature extraction/encoding.

The iris code will be extracted by demodulating the unwrapped iris image using complex-valued 2D Gabor wavelets [73], where the 2D Gabor filter equation can be represented as follows:

$$G(x, y, \theta, f) = exp\left\{-\frac{1}{2}\left[\frac{x'^2}{\sigma_{x'}^2} + \frac{y'^2}{\sigma_{y'}^2}\right]\right\} cos(2\pi f x)$$

$$x' = xcos\theta + ysin\theta$$
$$y' = ycos\theta - xsin\theta \qquad (3.1)$$

Where *f* is the frequency of the sinusoidal plane wave along the direction θ from the x-axis, and $\sigma_{x'}$and $\sigma_{y'}$ represent the variance along x- and y-axes, respectively. In this project, frequency values are assigned five different values (2, 4, 8, 16, and 32); orientation values $(\theta)$are assigned four different angles (0°, 45°, 90° and 135°) [72, 73].

Prior to applying a 2D Gabor filter to the unwrapped image, a mask will be created that will be used to identify non-iris pixels, corrupt areas and specular reflections by checking the grayscale intensities of the image. A 2D Gabor filter will then be applied to the unwrapped iris image and its corresponding masks, generating a real and imaginary part for both, which will then be used in the process of feature encoding/extraction.

The extracted features are converted into a number of bits of information known as iris codes. It also produces a corresponding noise mask, which identifies corrupt areas within the iris pattern. This encoding process is achieved through the process of phase quantisation, in which there is a quadrant containing a 2-bit binary number; position in the quadrant cannot be determined by comparing the real and imaginary parts of the filter [73].

The extracted features are converted into a number of bits of information known as iris codes. It also produces a corresponding noise mask, which identifies corrupt areas within the iris pattern. This encoding process is achieved through the process of phase quantisation, in which there is a quadrant containing a 2-bit binary number; position in the quadrant cannot be determined by comparing the real and imaginary parts of the filter [73] which is illustrated in Figure 3.2.

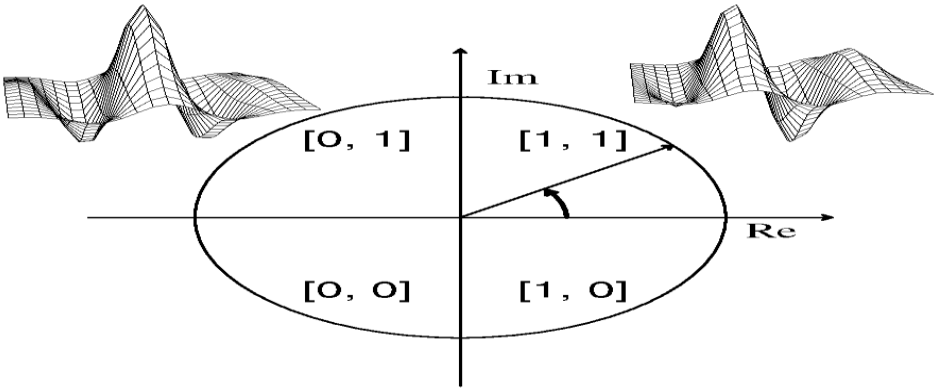


Figure3.2 Phase Quantization [73].

In order to identify the similarity between two iris images, the Hamming distance between two iris images will be calculated. This is the chosen metric for comparison of templates. The formula for computing the Hamming distance is shown in following equation [73].

$$HD = \frac{1}{N}\sum_{j=1}^{N} A_j \bigotimes B_j \qquad (3.2)$$

Where *A* and *B* are the iris codes for both two iris images provides. Whilst, the mask bits prevent non-iris artifacts such as eyelashes, eyelids, specular reflections, or other noise from influencing iris comparisons. To account for rotational inconsistencies, one template is shifted left and right bitwise, and a number of Hamming distance values are calculated from successive shifts. The lowest value will be taken, as this corresponds to the best match between templates.

3.2.3 Feature Extraction Approaches for Fingerprints

The extensive available range of fingerprint detection techniques can be broadly categorised into three groups [75-83]:

- Correlation-based matching: Two fingerprints are superimposed and correlations between corresponding pixels are computed for different alignments (e.g. various displacements and rotations).
- Minutiae-based matching: minutiae are extracted from two fingerprints, and matching decisions are based on the number of pairings.
- Ridge feature-based matching: Features are extracted from the ridge pattern.

For the current research work, the minutiae-based feature extraction and matching technique will be adopted primarily because it constitutes the backbone of several fingerprints recognition studies [83], where it uses the termination points and bifurcation as local features of fingerprints for the accurate recognition of the sample[83].This approach also widely used, computationally inexpensive, and locates the maximum number of minutiae pairings between the input sample and template for matching purpose [83].

2.3.1.1 Minutiae-Based Feature Extraction and Matching Approach

Fingerprint patterns are characterised by the existence of minutiae associated with the formation of ridges; there are more than 52 different types, of which seven are employed by human experts: crossover, core, bifurcation, ridge ending, island, delta and pore. Only two of these are currently used by automated systems: ridge endings, where a ridge terminates, and ridge bifurcation, which is where a ridge splits from a single into a double path [82-90].

Each minutia is typically associated with an (x, y) position, representing a direction or orientation. Matching these minutiae leads to identification or verification of a sample fingerprint [82-90].

During this process, first the fingerprint feature are extracted by the fingerprint grayscale image to a binary image, which will then be thinned (i.e. ridge line thickness reduced to one pixel), and necessary repair works will be performed. A simple image scan can then be used to detect those pixels that correspond to minutiae [82-90].

- A ridge pixel is a ridge ending if the number of ridge pixels in the 8-neighbourhood is 1.
- A ridge pixel is a ridge bifurcation if the number of ridge pixels in the 8-neighbourhood is greater than or equal to 3.
- A ridge pixel is an intermediate ridge pixel if the number of ridge pixels in the 8-neighbourhood is 2.
- [x,y, theta, associated ridge] are stored for each minutia.

A post-processing stage (called minutiae filtering) will subsequently be performed to remove any spurious minutiae, which are generated by corrupted regions or as a result of thinning. Figure 3.3 shows the different stages of fingerprint.

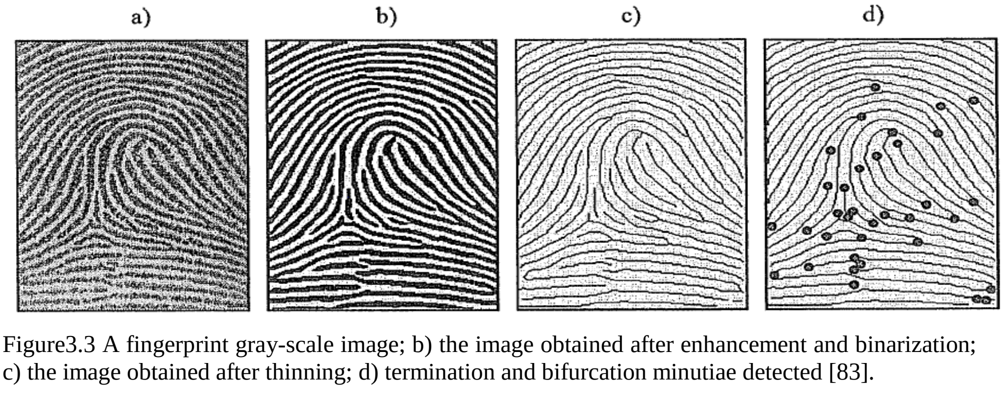


Figure3.3 A fingerprint gray-scale image; b) the image obtained after enhancement and binarization; c) the image obtained after thinning; d) termination and bifurcation minutiae detected [83].

Following the minutiae based feature extraction from the fingerprint image, the matching between the minutiae in the template and the live fingerprint is performed.

Each minutia will be represented by $m = \{x,y,\theta\}$, where (x,y) is the minutia location and θ is the minutia angle. Then, the template $T = \{m_1, m_2, \ldots , m_m\}$ and live fingerprint $I = \{m'_1, m'_2, \ldots , m'_n\}$, where m and n denote the number of minutiae in T and I, respectively.

A minutia m'_j in I and a minutia m_i in T are defined as 'matching' if the spatial distance (sd) between them is smaller than a given threshold r_0 and the direction difference (dd) between them is smaller than an angular tolerance θ_0. The two fingerprints must be aligned in order to maximise the number of matching minutiae. Alignment of two fingerprints involves displacement (in x and y) and rotation (θ) and may involve issues of scale and distortion. The matching score is the maximum number of minutia matches for any of these possible alignments, and this will be used to identify a user [82-90].

3.3 Generation of Secret Key Shares for Biometric Modalities

After following the feature extraction from the biometric modalities by applying the above mentioned approaches and creating biometric templates, the next logical step is to assign the secret shares to each of the modality from the actual biometric encryption key, which will allow the system to operate without storage of the biometric encryption keys in the authentication processes. This will lead to provide higher level of security to the biometric encryption key and the user personal data compared to systems based on template storage requirement.

In order to generate the secret shares, Shamir's Secret Scheme has been applied, which is elaborated in the subsequent sub-section.

3.3.1 Shamir’s Secret Sharing Scheme (SSS) Algorithm

Shamir’s secret sharing scheme algorithm [11, 13, 61, 62] will be used to generate secret key shares, each of which will be associated with an individual biometric modality. The appropriate shares will only be released upon the successful verification of each individual biometric modality. The SSS algorithm will also be used in recombining the secret key shares to reconstruct the secret; this will only be possible if the required numbers of key shares are available. Using Shamir’s secret techniques, the reconstruction process requires linear, quadratic or cubic equations. If, for instance, using a linear equation technique, only one of five biometrics provided at the verification stage was successfully released, the reconstruction process would be impossible and the user would be rejected because the linear equation strategy needs at least two points to be successfully verified in order to release the secret key.

For present purposes, we will first define the threshold values of K, N and P when creating the SSS algorithm program, where

K = Number of shares to be generated

N = Number of shares required to reconstruct the secret

P = A prime number

The number of key shares to be generated will be equivalent to the number of modality options available in the system. The number of shares required to reconstruct the secret will be the number of modalities required for successful authentication.

The program will generate a polynomial function $f(x)$ of $K-1$ degree with random coefficients, where the constant is a secret key to be defined or generated. Once the polynomial is built, K shares (x_i, y_i) will be constructed. At the sample collation stage, these shares will be assigned to the various biometric modalities by linking them together in the database. At the verification stage, upon successful verification of any chosen biometric modality, the key share associated with that biometric modality will be released. The *SSS* algorithm will then perform the necessary computations by combining the secret shares using the Lagrange interpolation method, or simply by solving Shamir's equation [11, 13, 61, 62], depending on the type of technique used to obtain the secret, in effect authenticating the user. If the number of secret key shares required for reconstructing the secret is not released because a user fails to verify the minimum required modalities, the *SSS* algorithm will be unable to reconstruct the secret and the user will not be authenticated by the system.

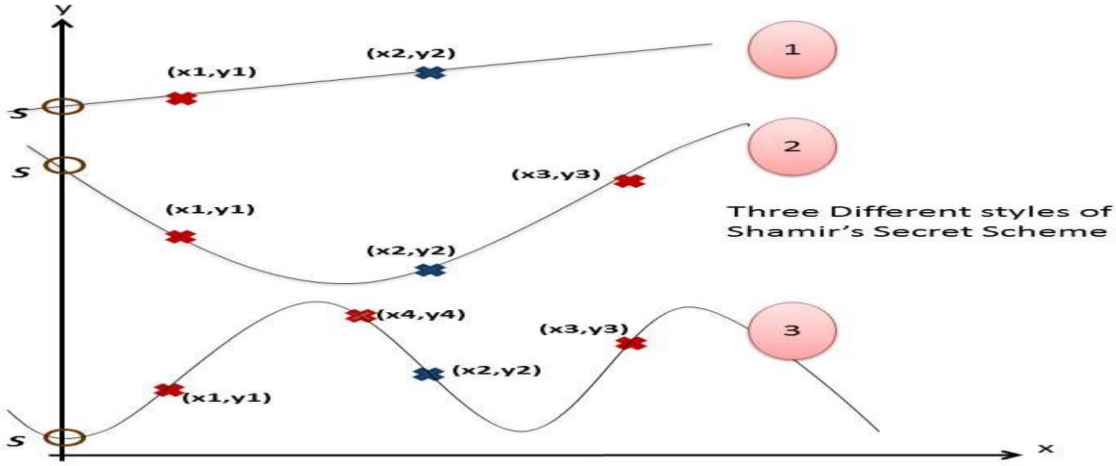


Figure3.4 SSS three different techniques [11].

Figure 3.4 shows three styles of using Shamir's SSS. The linear equation (1) uses two secret shares points to release the secret key (*S*). However, the quadratic equation (2) requires at least three secret shares points to release the secret key, and the cubic equation (3) requires at least four.

3.4 The Proposed Scheme's Architecture for indirect Key Generation and Key Release

The overall architecture of the proposed multi-modal biometric system for indirect key generation is presented in the Figure 3.5 and 3.6. The individual points on a given polynomial are derived from five biometric modalities of a given person: face, iris (right, left), and fingerprint (thumb, index finger). After feature generation, the equation of the polynomial must be solved in order to retrieve the key, shown as the intercept on the vertical axis [11]. The number of modalities required to derive the secret is naturally governed by the order of the polynomial: a linear equation requires two points to solve, a quadratic equation requires three points, a cubic equation requires four, and so on.

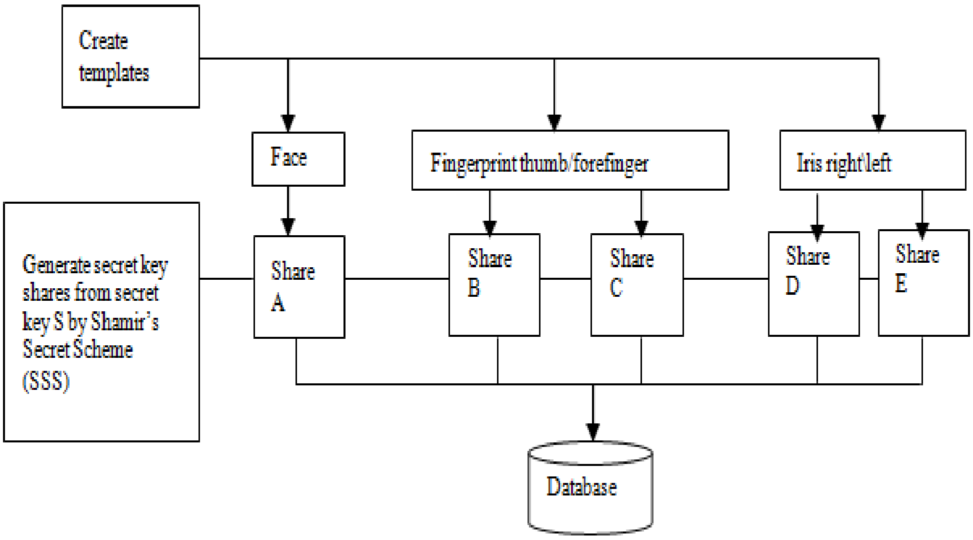

Create templates
Face
Fingerprint thumb/forefinger
Iris right/left
Generate secret key shares from secret key S by Shamir's Secret Scheme (SSS)
Share A
Share B
Share C
Share D
Share E
Database

Figure3.5 Enrolment process

Figure 3.5 illustrates the enrolment process, in which multiple samples of each modality are captured from a given user and a reference template for each modality is created in the database. Each reference template is associated with a point on the desired polynomial, which will be released during authentication. Figure 3.6 shows the authentication structure activated by providing samples from a candidate user. Each modality employs an algorithm for matching with the stored reference database. If a match is found, the corresponding coordinate point is released. The process concludes by employing the generated points to solve the polynomial equation, and to release the key if sufficient points are available.

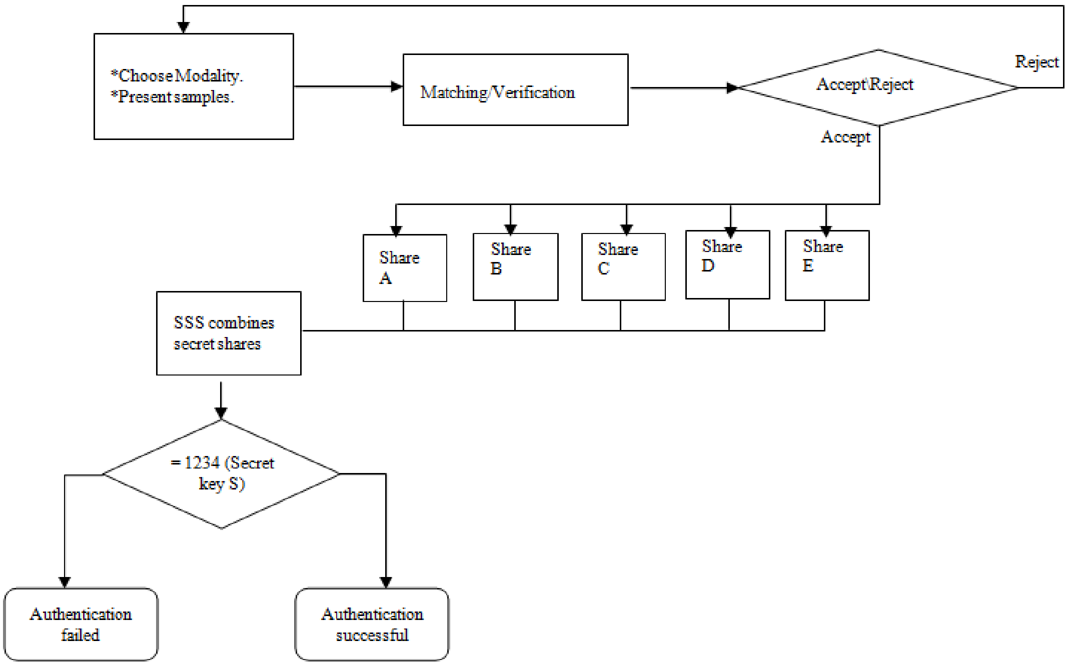


Figure3.6 Authentication process

3.5 Performance and Experimentation Results

The major thrust behind the development of the proposed scheme is to develop a robust multimodal biometric system which may perform without requirement of the storage of encryption key within the system, and perform equally well to authenticate the user's claim even when he fails to provide one/two modalities out of the given set of modalities to the multi-modal biometric system. Shamir's Secret Scheme was found highly suitable to achieve this objective, which allows the generation of secret share and release by the secret key based on meeting two points on linear equation, three points on quadratic equation, and four points on cubic equation. This flexibility of the Shamir's scheme enabled the researcher test various combination of biometric samples for evaluating the performance of the proposed biometric scheme. Based on the experimental conditions, three polynomial orders were used: linear, quadratic and cubic, which are required by the Shamir's secret Scheme approach to test the performance of the proposed scheme when 2 points, three points and four points are provided to define linear, quadratic and cubic equations

This section deals with the experimentation process, detailing the experimental results in terms of the false acceptance and rejection rates for the overall authentication system, for all users in an authentication scenario.

3.5.1 Experimentation Data and Requirements

For the purposes of this investigation, 300 pseudo-users were created from five separate, genuine biometric databases, taking the corresponding values from each database and deeming them to derive from a given virtual user. These databases are FVC2004 represented fingerprint samples [91], CASIA-IrisV3 presented iris samples [92] and FERET presented face samples [93]. For each person, three samples are employed for enrolment, and for testing, each user provided five samples. In the authentication phase, then, five tests were created for each user, and for each test, five samples were provided for all 300 pseudo-users, as shown in Table 3.3.

Table 3.1 Example enrolment data.

User id	Samples	Face	Fingerprint		Iris	
			Thumb	Index Finger	Right	Left
1	S1					
	S2					
	S3					

Passive forgeries were also investigated by requiring each of the 300 users to try to gain access by falsely claiming the identity of each of the other 299 users, which was adopted to check the system’s susceptibility to the attack. Table 3.1 shows the enrolment of a genuine user. There are eight samples per user for each modality (three for training and five for testing), chosen randomly.

Table 3.2 Samples for Testing

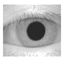
User id	Test	Face	Fingerprint		Iris	
			Thumb	Index finger	Right	left
1	1					
	2					
	3					
	4					
	5					

Table 3.2 presents an example of the testing method, in which five tests were provided for each user; each test could cover five different modalities (face, thumb, index finger, right iris, left iris). The performance of the system for genuine users (Table 3.3) shows five tests for each user, with each test requiring five samples to be provided.

Table 3.3 Sample performance; √ indicates success for given modality or polynomial order; X indicates failure.

User ID	Test	Face	Fingerprint		Iris		Secret Share Techniques		
			Thumb	Index finger	Right	Left	Linear equation	Quadratic equation	Cubic equation
1	1	X	√	√	√	√	√	√	√
	2	X	√	√	√	√	√	√	√
	3	X	√	√	√	√	√	√	√
	4	√	X	√	√	√	√	√	√
	5	√	X	√	√	√	√	x	x

A set of thresholds for each component modality classifier—chosen randomly to test the results of FAR and FRR for each biometric modality—was used to extract the threshold for the polynomial cross curve between FAR and FRR, known as the equal error rate (ERR) (Table 3.4).

Table 3.4 FAR and FRR results for each biometric modality separately.

Biometric Modality	FRR	FAR	ERR
Fingerprint	5.9%	3.2%	4.89%
Face	12.9%	6.2%	10.76%
Iris	1.79%	0.034%	0.25%

3.5.1.1 SSS Algorithm Requirement

The major goal of this project was to develop a working implementation of Shamir's SSS Algorithm [11], representing the cryptographic aspect of a multimodal biometric system. The SSS Algorithm can be used to generate a number of secret key shares from a chosen secret (which could be any number), and these key shares are then bound to each biometric modality. Successful verification of each individual biometric modality should release the appropriate key share. A combination of the minimum key shares required (two) should reconstruct the original secret, leading to authentication. If the required number of key shares is not successfully combined, the secret cannot be reconstructed, and the user will be rejected by the system.

Five different biometric modalities should be used to facilitate implementation of the SSS algorithm. Using the linear equation technique, the algorithm should be incorporated into the system in such a way that a user can only be authenticated following successful verification of at least two biometric modalities, based on their chosen options. If a user fails to verify at least two biometric modalities, they should be rejected by the system (and so on for the other techniques).

3.5.1.2 Verification Algorithms Requirement

In conjunction with the SSS algorithm for recognising a user, three individual verification algorithms should be implemented for use by multimodal biometric systems. These should include a face recognition algorithm, an iris recognition algorithm and a fingerprint recognition algorithm. The algorithms should be able to identify a user by comparing a sample presented by the user against a stored template, making a decision on the basis of degree of similarity.

3.5.1.3 Database Requirement

A database should be created to accommodate storage of biometric templates for different users. Secret key shares created by the SSS algorithm should also be stored in the database and linked to the respective individual biometric modalities in the database. During the verification process, user samples presented to the system should be compared against their respective templates stored in the database, and the appropriate key shares should be released if verification was successful.

3.5.1.4 Testing Requirements

The system should be tested for technical performance of the SSS algorithm and of the individual verification algorithms. From these tests, overall performance of the system can be deduced and, most importantly for authenticating users, false acceptance rates and false rejection rates of the system can be tested.

3.5.1.5 Software implementation of feature extraction algorithms

In this work, three feature extraction algorithms have been used, named as Eigneface algorithm for face feature extraction, 2D-Gabor filter for iris feature extraction, and minutiae based feature extraction for fingerprints. The Shamir secret scheme algorithm also has been used for key generation indirectly form the multimodal biometric system used in this work. All these algorithms have been implemented using the software MATLAB. The codes for implementation of these algorithms have been generated by the researcher himself. In addition, in order to gain the expert knowledge about the coding for feature extracting, the following sources have been consulted: AT&T Laboratories Cambridge (online source) implemented by Luigia Rosa [176] for the implementation of face recognition features and the fingerprints recognition implemented by Wuzhili [177]. However, the implementation of 2D-Gabor filter algorithm has been implemented in MATALB by researcher from the scratch.

3.5.2. Experimentation Results

The performance of the system is illustrated in Tables 3.1 to 3.6. Tables 3.3 and 3.6 show sample results for the authentication system for genuine and impostor users (five tests for each user, each requiring five samples to be provided). In Table 3.3, the observed failure of one modality does not necessarily cause the system as a whole to fail; samples 1 to 3 (face failure) and sample 4 (fingerprint failure) demonstrate the exception-handling potential of the system. As shown in sample 5, multiple failures still lead to overall failure. A summary of the overall performance of the system for all quoted samples, in terms of the false acceptance and false rejection rates of the Shamir component of the system for the three polynomial orders, are shown in Table 3.5.

Table 3.5 FRR and FAR results for authentication.

	Linear equation	Quadratic equation	Cubic equation
FRR	1.8%	4.8%	8,2%
FAR	5.2%	2.59%	0%

Table 3.6 Sample performance for passive forgery.

Impostor user	Claimed Identity for user	Test	Face	Fingerprint		Iris		Secret Share Techniques		
				Thumb	Index finger	Right	Left	Linear equation	Quadratic equation	Cubic equation
1	2	1	x	√	√	x	X	√	X	X
		2	x	√	√	x	X	√	X	X
		3	x	√	√	x	X	√	X	X
		4	x	X	X	x	X	X	X	X
		5	√	X	√	√	√	X	X	X

Table 3.5 shows the results for three different polynomials of Shamir's secret scheme. The linear equation requires at least two points to generate the secret key S. The false rejection rate of 1.8% in the linear equation method is good by comparison with the quadratic and cubic techniques, but it must also be noted that the false acceptance rate is greater (5.2%) by comparison with the other techniques. At 0%, the false acceptance rate for the cubic equation technique is very good, as this technique requires at least four points to generate the secret key S. However, the false rejection rate is relatively high (8.2%). For the quadratic equation technique, the false rejection and acceptance rates are both intermediate. Because this technique requires at least three out of five points to generate the secret key S, it reflects a balanced probability in terms of both FRR and FAR. These results are very interesting, as they demonstrate the relative merits of the polynomial orders when considering the desired performance of the system. However, the outcomes of these tests are in part dependent on the quality of the biometric samples employed and on the algorithms employed for the individual modalities.

3.6 Summary

This chapter has explored the technique of secret sharing to allow an encryption key to be created from multimodal biometric samples. The results show the potential of the system for efficiently deriving encryption keys while also allowing for exception handling, which is currently a significant impediment to the practical deployment of biometric systems. This improved robustness property represents a significant enhancement. A further significant advantage of the proposed technique is that the biometric key itself need not actually be stored, which along with the unlimited length of the biometric encryption key further enhances the security of the system.

Chapter 4.

Investigations of Iris Direct Key Generation

4.1 Introduction

Biometrically constructed security is theoretically strengthened if an encryption key is extracted directly from biometric samples as provided by a given user. Such a system means that the retained encryption key would not be copied and also a template or reference sample would not be required, significantly enhancing potential system security. However, two probable weaknesses intrinsic to this scheme are that the generated encryption key would not easily be revoked and, where a user is unable to provide a biometric sample, the scheme would not be robust in this condition [41, 44].

This chapter introduces several schemes for integrating direct biometric key generation schemes with Shamir's secret sharing algorithm [11] to directly address these two disadvantages of revocability and exception handling. Within the proposed scheme, individual points on a polynomial curve are directly derived from iris samples taken from an individual by applying a user function, which is created for each user to minimise the amount of data stored and enabling Shamir's secret scheme to be applied to derive the required key. The proposal is robust, in that the new technique generates an encryption key from biometric modality samples, using a minimal amount of stored data. The system's potential has been investigated in relation to passive forgeries. The current chapter reports preliminary work on how an encryption key may be generated directly from the biometric modality by extracting points on a parabolic curve derived from actual biometric samples. These schemes returned negative results, indicating that the Equal Error Rate (EER) is high or that the level of performance or security is low. Each scheme will be presented in detail, including an explanation of how it works, why this scheme might be used, and any feedback from it.

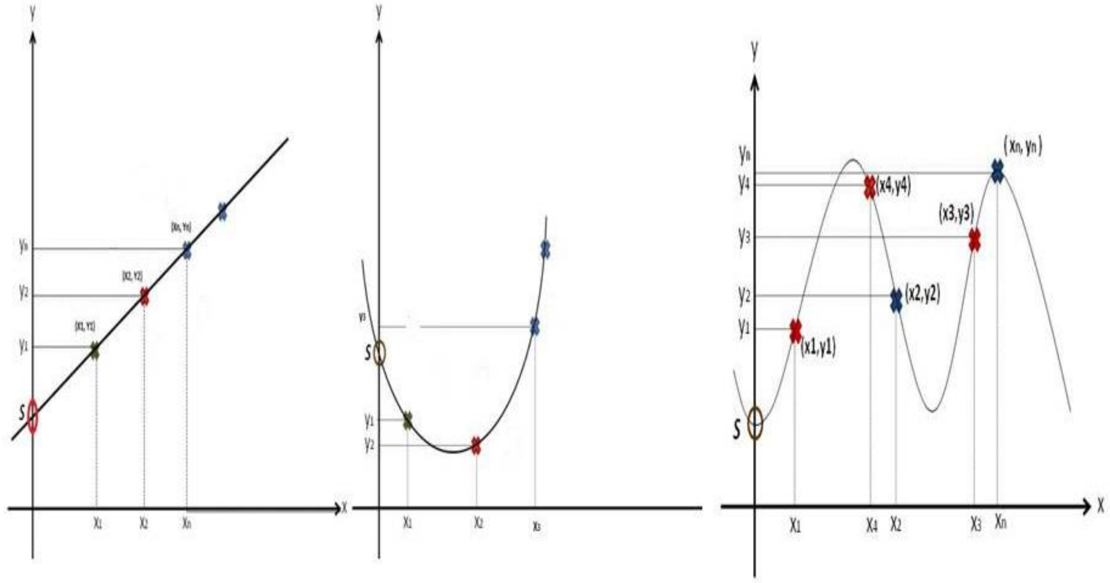


Figure4.1 Three Different Scenarios for Direct Key Generation from Individual Biometric Samples

Figure 4.1 shows how the direct key could be generated from individual biometric samples applying Shamir’s secret scheme. Each point in the curve represents a different biometric modality enabling the biometric secret key generated by Shamir’s secret scheme. In the field of biometric encryption, this proposal must satisfy security considerations as well as level of performance flexibility. From the security point of view, this scheme uses a multimodal, template-free biometric system, and the length of the biometric encryption key is unlimited because, as shown in Figure 4.1, the y-axis extends to infinity. From a performance perspective, three different techniques using Shamir’s secret scheme are presented: Linear, Quadratic, and Cubic equations. For the multimodal biometric system in this proposal, at least two biometric modality points would be sufficient to release the biometric secret key in the linear equation technique, with three points needed in the quadratic and four points in the cubic.

4.2 Biometric modality

Iris recognition is the chosen modality for these investigations, entailing pre-processing, feature extraction and feature encoding and using the same methods in all schemes.

4.2.1 Pre-processing

The process of iris code extraction involves a number of steps: iris localization/segmentation, iris normalisation/unwrapping [71-73], and feature extraction/encoding [71, 73]. These were discussed in detail in Chapter 3 (section 3.3.2).

4.2.2 Feature Extraction

Two methods of iris feature extraction have been used in these investigations.

1- 2D Gabor Filter [73].

In this approach, the iris code was extracted by demodulating the unwrapped iris image with complex-valued 2D Gabor wavelets [73], where the 2D Gabor filter equation can be seen in Eq.(3.1).

Furthermore, Frequencies are assigned five different values (2, 4, 8, 16 and 32). Orientation values (θ) represent four different angles (0°, 45°, 90°, and 135°) [72, 73].]. It can be noticed that different values and orientation angles have been selected for the experiments, because the functions of the Gabor filter to extract the iris features are not optimized, therefore, the parameters for extraction of high quality iris features for each experiment are required to be optimized on trial and error basis using different values.

Encoding iris features by 2D wavelet demodulation, the resulting complex and imaginary parts form the 2D Gabor filter. Eq.(3.1) is then used in the process of phase quantization in generating the iris code [73].

2- Gray Level Co-Occurrence Matrix (GLCM) [94, 95].

As proposed by Haralick et al. [95], GLCM is one of the most widely-used approaches to extraction of textural features. The approach can be defined as follows. Suppose an area has N_c and N_r resolution cells in the horizontal and vertical directions, respectively, and Ng level in the gray tone. Let $L_c = \{1, 2, .., N_c\}$ be the horizontal spatial domain, let $L_r = \{1, 2, ..., N_r\}$ be the vertical spatial domain, and let $G = 1, 2, .., N_g$ be the set of N_g quantised gray tones. The image I can then be represented as a function that assigns some gray tone in G to each resolution cell or pair of coordinates in $L_r \times L_c$; $I : L_r \times L_c \rightarrow G$. Texture-context information is specified in a matrix of relative frequencies P_{ij} with two neighbouring resolution cells separated by distance d occurring on the image, one with gray tone i and the other with gray tone j. P_{ij} can then be described by Eq. (4.1):

$$P(i,j,d,\theta) = \#\{((k,l),(m,n)) \in (L_r \times L_c) \times (L_r \times L_c) \parallel max(|k-m|,|1-n|) = d,\theta,I(k,1) = i, I(m,n) = j\} \qquad (4.1)$$

where θ and d are the directions and distances between two pixels in the image, and # denotes the number of elements in the set. Traditionally, Haralick’ s features are fourteen in number; for the present analysis, we chose the first twelve texture features. Using GLCM to extract textures is sensitive to three factors: selection of window size, number of gray levels, and distance between pixel pairs. Window size was set at 8×8 pixels, gray levels at 256 and the distance between pixel pairs at 4 pixels.

4.3 Investigative Methodologies

4.3.1 System 1: Key Generation from Iris modality using mini-template

The researcher faced the problem of generating the encryption key directly from the iris modality from the template free biometric system. It was tested whether the encryption key can be generated from the system using the mini-template – a template which was reduced to a minimum possible data. The success of this system was supposed to enable the researcher to keep reducing the data stored in database unless the template free system can be generated. The following paragraphs will give overview of working principle of this system.

The overall operation of the system is as follows. The individual points on a given polynomial are derived from the iris modality. After iris feature generation, extraction of the position of "1" characters of the iris feature will be required for each sample, calling a function related to an enrolled user id. This is the novel property of the algorithm. The equation of the polynomial may then be solved to retrieve the key, which is represented as the intercept on the vertical axis [11]. The number of samples that are required to derive the secret is naturally governed by the order of the polynomial; a quadratic equation, as used in this scheme, requires three points to solve.

The enrollment process requires capture of the iris sample from the given user, generating a bespoke function for each user. The explanation of the algorithm is as follows.

- After extracting iris features, these will be represented as a binary matrix of 0 and 1.

Sample 1:

0	1	1	1	0	0	1	1	0	1

- Extraction of ones positions will be as follows:

Sample 1	0	1	1	1	0	0	1	1	0	1
Position	1	2	3	4	5	6	7	8	9	10

So, X

Ones Position	2	3	4	7	8	10

- To provide the Y_n coordinates for three points [Y_0 Y_1 Y_2], the following considerations must be taken into account.

1. As was mentioned before, a quadratic equation requires three points to solve the polynomial curve using Shamir's secret scheme.
2. The secret key should be controlled and flexible (shown as the intercept on the vertical axis).
3. X_n = {X_0 X_1 X_2} coordinates are derived from three iris samples in the authentication process; it will be discussed and explained later how X_n = {X_0 X_1 X_2} is calculated.
4. Several sets will be applied to find the equal error rate, increasing the distance between the three points by adding a Z value taken from Table 4.4 to Y_n coordinates, which will be described later.

- So, Y_n = {Y_0 Y_1 Y_2}, points will be provided by the following algorithm:

$$Y_0 = YY + (Z * 2),\ \ Y_1 = YY + Z,\ \ Y_2 = YY$$

where YY is a limitation point for the vertical axis Y, serving as a secret key point, as shown in Table 4.1 below.

Z is the value added to Y_n coordinates to find the equal error rate, basically increasing the distance between the three points X_n,Y_n.

Table 4.1 Deciding where the Secret Key will be Placed in the Vertical Axis Y (system 1)

YY	Limit points that could be place in the vertical axis to be a final point (Secret key)
(0)^2, (1)^2	-5 – 5
(2)^2	0 – 10
(3)^2	5 – 15

The reason for the differences between *Y_0, Y_1*, and *Y_2* is to avoid errors that might be caused by the line equation curve and to extract the equal error rate (as mentioned in point 4 above) where all three points have the same *Y* coordinate. Controlling the secret key and making it more flexible, as shown in Table 1, means that if the secret key needs to be on the vertical axis between (-5 – 5), by inspection 0 and 1 are in the middle, and by raising them to 2, they are equal to 0 and 1, which are still in the middle of the limit. In addition, if the secret key needs to be between 20 and 30, by raising 5 to 2 it is equal to 25, which still within the limit.

Finally, the function for each user will be created as follows:

User_id_F(x) = {XL_1 , Y_n}.

where *XL_1* references a matrix of ones positions from the sample that has been captured, and *Y_n* references *Y_n* coordinates for the three points {*Y_0*, *Y_1*, and *Y_2*}.

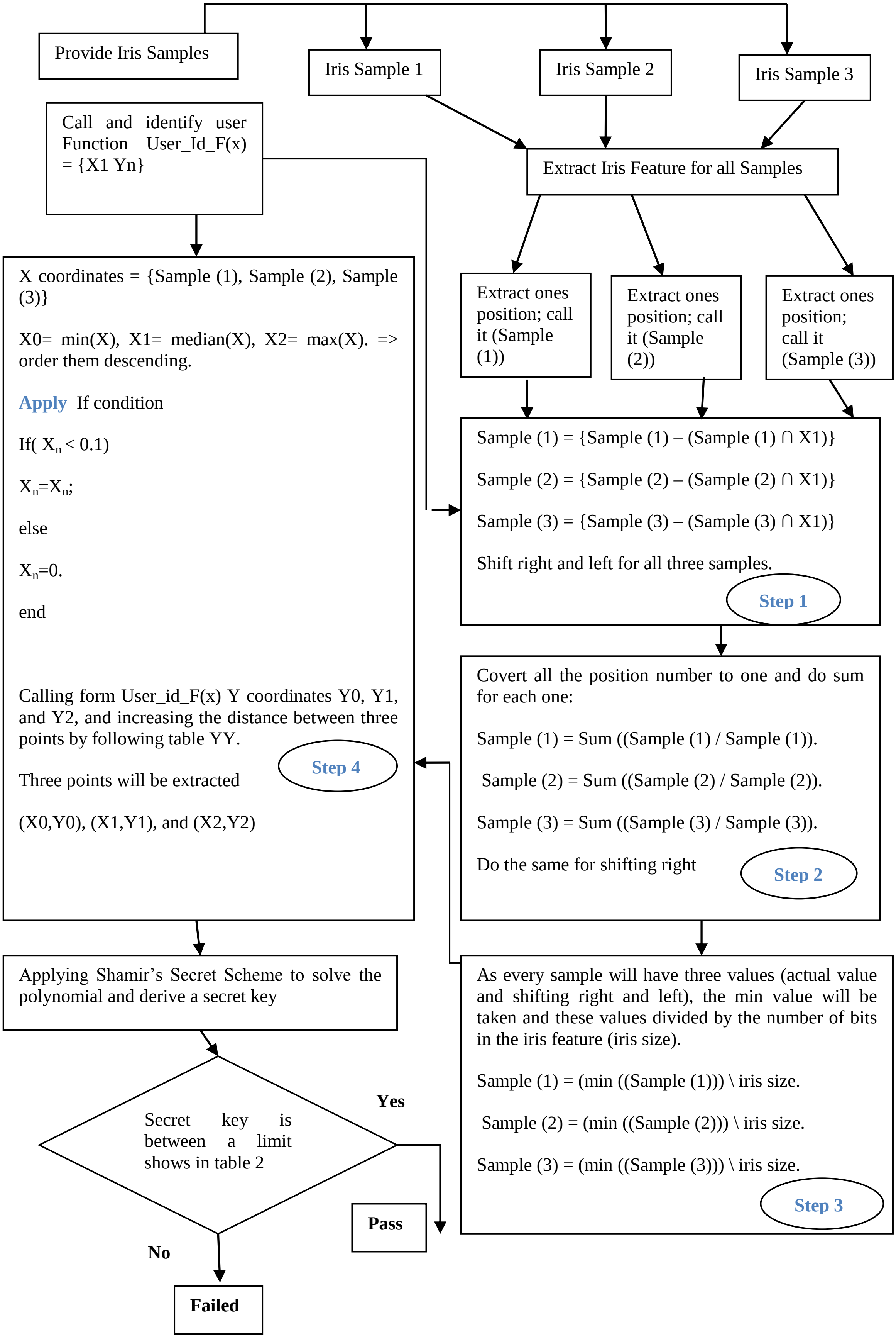

Provide Iris Samples
Iris Sample 1
Iris Sample 2
Iris Sample 3
Call and identify user Function User_Id_F(x) = {X1 Yn}
Extract Iris Feature for all Samples
Extract ones position; call it (Sample (1))
Extract ones position; call it (Sample (2))
Extract ones position; call it (Sample (3))
X coordinates = {Sample (1), Sample (2), Sample (3)}
X0= min(X), X1= median(X), X2= max(X). => order them descending.
Apply If condition
If($X_n < 0.1$)
X_n=X_n;
else
X_n=0.
end
Calling form User_id_F(x) Y coordinates Y0, Y1, and Y2, and increasing the distance between three points by following table YY.
Three points will be extracted
(X0,Y0), (X1,Y1), and (X2,Y2)
Step 4
Sample (1) = {Sample (1) – (Sample (1) ∩ X1)}
Sample (2) = {Sample (2) – (Sample (2) ∩ X1)}
Sample (3) = {Sample (3) – (Sample (3) ∩ X1)}
Shift right and left for all three samples.
Step 1
Covert all the position number to one and do sum for each one:
Sample (1) = Sum ((Sample (1) / Sample (1)).
Sample (2) = Sum ((Sample (2) / Sample (2)).
Sample (3) = Sum ((Sample (3) / Sample (3)).
Do the same for shifting right
Step 2
As every sample will have three values (actual value and shifting right and left), the min value will be taken and these values divided by the number of bits in the iris feature (iris size).
Sample (1) = (min ((Sample (1))) \ iris size.
Sample (2) = (min ((Sample (2))) \ iris size.
Sample (3) = (min ((Sample (3))) \ iris size.
Step 3
Applying Shamir's Secret Scheme to solve the polynomial and derive a secret key
Secret key is between a limit shows in table 2
Yes
Pass
No
Failed

Figure4.2 Authentication Process

Figure 4.2 shows the authentication structure, which begins by providing iris samples from a candidate user. Each iris sample employs a feature extraction algorithm to extract iris features. After that, ones positions must then be extracted for each sample. Function *User_id_F(x)* will be called next. This function holds the XL_1 matrix, which is the position of ones that came from the enrollment iris sample, for the user id for the claimed identity, applying the XL_1 matrix with iris samples as per the following calculation and taking account of shifting the iris feature left and right as well:

Sample (1) = {Sample (1) – (Sample (1) ∩ XL_1)} ➔ *Sample(1)_shift_left , Sample(1)_shift_Right.*

Sample (2) = {Sample (2) – (Sample (2) ∩ XL_1)} ➔ *Sample(2)_shift_left , Sample(2)_shift_Right.*

Sample (3) = {Sample (3) – (Sample (3) ∩ XL_1)} ➔ *Sample(3)_shift_left , Sample(3)_shift_Right.* (4.2)

Convert all these samples from the position number to one and sum each sample separately as per the following calculation:

Sample (1) = Sum ((Sample (1) / Sample (1)).

Sample (2) = Sum ((Sample (2) / Sample (2)).

Sample (3) = Sum ((Sample (3) / Sample (3)). (4.3)

To find the closest sample to the genuine user, the min value from the three values from each sample will be taken and divided by the number of iris feature bits (*K*), as follows:

Sample (1) = (min ((Sample (1), Sample (1)_shift_left, and Sample(1)_shift_Right)) \K

Sample (2) = (min ((Sample (2) , Sample (2)_shift_left, and Sample(2)_shift_Right)) \K.

Sample (3) = (min ((Sample (3) , Sample (3)_shift_left, and Sample(3)_shift_Right)) \K (4.4)

Moreover, X_n coordinate values will be {*Sample(1), Sample(2), Sample(3)*}; in descending order, X_0 will be the min (X_n) while X_1 will be the median (X_n) and finally X_2 will be the max (X_n).

Subsequently, the final step in determining whether these three points belong to the genuine user or to an impostor is to apply the following if condition:

```
$if\ ( X_n < 0.1)$
$X_n = X_n$
$else$                    (4.5)
$X_n = 0.$
$end$
```

This condition is used to test if the point is smaller or greater than 0.1 for the classification process of separating genuine from impostor users. If the point is a genuine user, the point will be left as is. However, if the point is greater than or equal to 0.1, then this point will be equal to zero to separate the point far away from the genuine user's points. In addition, the function *User_id_F(x)* will also have the *Y* coordinates {Y_0, Y_1, Y_2}, based on Table 4.1 above. As a result, Shamir's secret algorithm solves the polynomial curve to retrieve the key, shown as the intercept on the vertical axis. Finally, by checking the key and according to Table 4.1, if the key is between the limit that has been chosen then Pass (otherwise Failed).

4.3.1.1 Performance and Experimental Results

This section details the experimentation process, providing experimental results in terms of false acceptance and rejection rates for the overall authentication system for all users in an authentication scenario.

The investigation pertained to 300 users, using iris modality databases, with 18 samples per user. Taking six samples for each user divided in two tasks, each set of three samples was used to create a user function and employed for testing. There were four tests for each user, each providing three samples and applying them seven times with different limitations, according to Table 4.1, to decide the key's place on the vertical axis.

Table 4.2 Samples Performance (√ Indicates Success for Given Modalities; X Indicates Failure)(system1).

User_ **Id**	**Te** **st**	**Iris limitation key place in the vertical axis Table 2**						
		(0)^2, (1)^2 ***-5 – 5***	***(2)^2*** ***0 – 10***	***(3)^2*** ***5 – 15***	***(4)^2*** ***10 – 21***	***(5)^2*** ***20 – 31***	***(6)^2*** ***30 – 41***	***(7)^2*** ***40 – 51***
1	1	x	√	X	x	x	X	X
	2	X	x	X	X	x	X	X
	3	√	√	√	√	√	√	√
	4	√	√	√	√	√	√	√

Passive forgeries were also investigated by employing 300 impostor users trying to gain access by claiming the identity of 300 genuine users.

The data experiments for iris used the CASIA-IrisV3-Interval database [92]. This justification of the use of this database has been provided in chapter 2,

Table 4.3 Passive Forgery Example (system1)

User_Id	1	2	3	4	. ..30
Claimed Identity for user	300	59	58	57	…. 1

Table 4.4 shows a set of variances applied to increase the distance between three points in the polynomial curve to establish false rejection rate (FRR), false acceptance rate (FAR) and equal error rate (EER).

Table 4.4 Trying Several Sets in the Curve to See Variances in FRR, FAR and Equal Error Rate (system1).

Number of Attempts	1	2	3	4	5	6
Z values	0.25	0.025	0.0025	0.00025	0.000025	0

An illustration of the performance of the system is shown in Tables 4.2 to 4.5 and in Figure 4.3. Tables 4.2 to 4.4 show sample results of the authentication system for genuine and impostor users: two tests for each user, providing three samples for each test. In the strategy in Table 4.4, six different points are applied to the polynomial curve to establish the variances in FRR and FAR. Each point in the strategy is applied in Table 4.2. Here, the result of a failure of one system limitation does not necessarily cause the system as a whole to fail (as shown in sample 1 test 1 with all limitation failure). A summary of overall performance for all samples, in terms of false acceptance and false rejection rate, of the Shamir component of the system for the three polynomial points strategy in Table 4.4 is shown in Table 4.5 and graphically in Figure 4.3.

Table 4.5 Results of FRR and FAR for Authentication (system1).

Number of Attempts	1	2 ERR	3	4	5	6
FRR	55.8 %	19,3 %	7.3 %	3.6 %	3.5 %	1.25 %
FAR	5.4 %	18.06 %	21.1 %	22.2 %	23.7 %	95.4 %

Table 4.5 shows the results for six orders of y-coordinate polynomial points in attempting to generate the key secret. The number of attempts made to deduce FRR and FAR in the Table 4.5 were made based on the z-values showed in table 4.4, in order to test the impact of z-vales on the False accept and False Reject rate (FRR and FAR) and quality of encryption key. The FRR in taking y-coordinates from Table 4.1 is 1.25%, but the FAR is too high. By increasing the distance between the three points, the FRR is increased while the FAR is decreased. As a result, attempt numbers 2 and 3 are the best result so far; in attempt number 2, 19.3 % FRR and 18.06% FAR. However, attempt number 3 showed that the FRR is much lower that attempt 2 at

7.3%, but the FAR has slightly increased as compared to attempt 2. Hence, the equal error rate (EER) shows roughly 20 %

These results are interesting, as they demonstrate the relative merits of the polynomial orders when considering the desired performance of the system. However, the outcomes of these tests are partly dependent on the quality of the biometric samples employed and the algorithms employed for the iris modality.

The FRR and FAR ROC curve, showing how the performance of the system varies across six different techniques by increasing the distance between points in the horizontal axis, is shown in Figure 4.3.

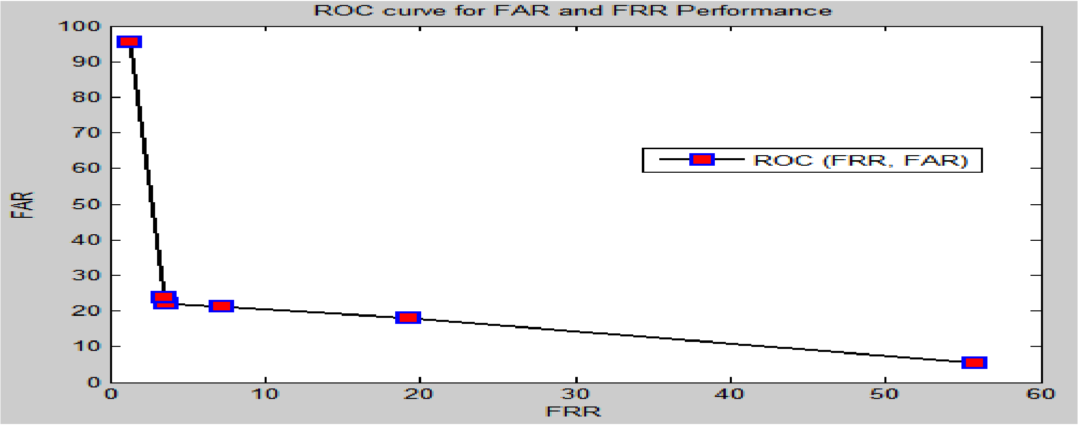


Figure4.3 ROC curve showing FAR and FRR performance of system 1

4.3.1.2 Summary

An attempt has been made to explore a new technique of generating an encryption key from biometric modality samples, using a minimal amount of stored data and examining the system reaction. Within the proposed scheme, the biometric template associated with each user need not be stored; only the matrix of “ones” positions is stored, which does not identify the associated user. The results appeared negative; we considered the scheme to have failed for the following reasons.

- The EER was over 20% which is a considered as a high error rate.
- It seems that the data stored in the mini-template, as we called it in this scheme, needs to be reduced.
- Generation of the biometric encryption key was inaccurate. In this scheme, we extracted it at enrollment, but in the authentication we controlled by limitation in the y-axis, which affected the level of security and performance as reflected by a very high ERR.
- The system is easy to hack. By looking at the authentication structures, a hacker can skip all the steps and go directly to Step 4 in the authentication process in Figure 4.2, providing any number under 0.1.

On the other hand, a lot was learned from this scheme, as summarised in the following points.

- In comparing any two iris samples, whether from the same person or not, there is a huge similarity between them. As evidence of this, any iris recognition system defines a threshold such that a Hamming Distance < 0.2 for example (which is the average threshold in public) identifies the iris provider as genuine or otherwise as an impostor. This indicates that the similarity between any two irises is more than 70%.
- Iris image quality plays a fundamental role in helping any iris recognition system to extract the iris perfectly.
- Iris feature extraction may result in significant similarity between users.

4.3.2 System 2: Template free key generation from Iris modality using 2D Gabor Filter

The template free key generation approach was adopted in this work because it aims to reduce the data stored in the database so that hackers can not attack the privacy of the data. This also intends to reduce the attacks on the security of the biometric system. This system stored the data in form vectors of numbers of the successful blocks against each user's biometric sample, which does not identify the associated users. Hence, the success of this system will prevent the security and privacy attacks on the biometric system, along with increasing the performance of the system.

In this scheme, a new approach was taken to the process of investigation, dividing the iris binary code as extracted by 2D Gabor filter and 2D wavelet demodulation [73], respectively, into many blocks. Each block was converted to a decimal number (binary to integer) by use of conversion tools. Blocks were chosen by dividing the training sets into two tasks, each with a number of training samples, to compare these two tasks and to specify blocks that were the same or close in value. Shamir's secret method was then applied to generate the biometric encryption key for placement on the y-axis. The X_n and Y_n coordinates were specified as follows:

- X_n coordinate: the value of the chosen blocks
- Y_n coordinate: the numbers of the chosen blocks

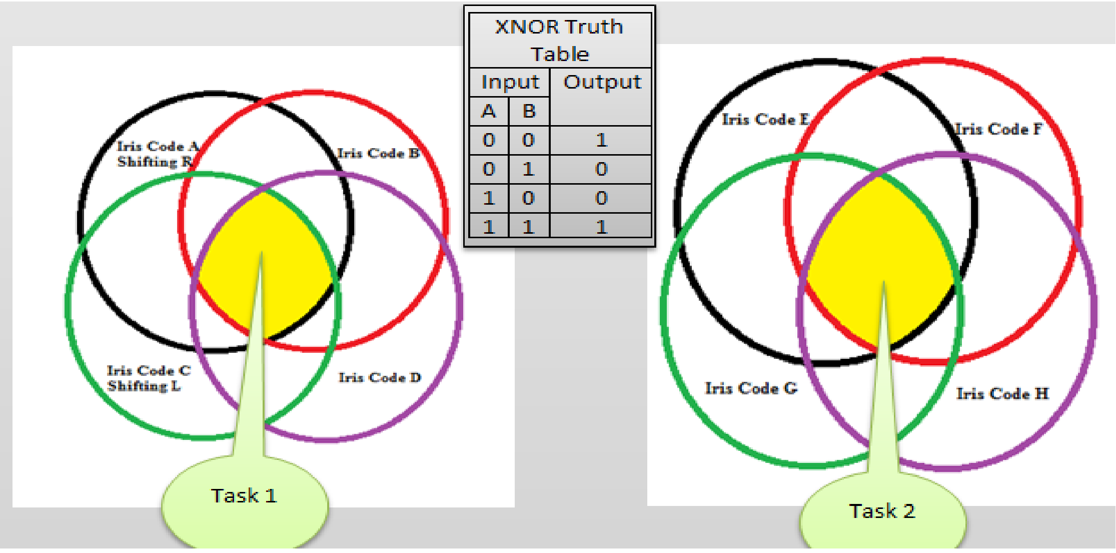
The system architecture was divided into enrollment and authentication processes, which will be explained separately in detail. Each user had 20 iris samples, divided in two, with 8 samples for training purposes divided into two tasks with four samples in each. For the testing process, 12 samples were provided for three tests, again with four samples in each test. The enrollment process can be described as follows.

1- The user is asked to provide 12 iris samples in three tasks, each involving four samples.

2- Feature extraction is performed by the 2D Gabor filter method [73].

3- Iris binary code is generated by 2D wavelet demodulation [73, 96, 97].

4- The logical operation XNOR is then applied for each task as in Figure 4.4. This is used primarily to pinpoint the invariable iris features (codes) in samples taken at different intervals from the same users and secondly to avoid high similarity percentage between two iris codes taken from different codes. This is novel step which has been introduced in the algorithm.


XNOR Truth Table

Input		Output
A	B	
0	0	1
0	1	0
1	0	0
1	1	1

Figure4.4 XNOR Operation for Four Samples in each Task –system 2-

5- To generate the similarity between samples for each task, the result code is sized in 7200 bit lengths, divided into 72 blocks of 100 bits, as it shows in the Table 4.6. This is also novel aspect of this algorithm which led the researcher to find the location of successful blocks.

Table 4.6 Example of dividing the iris code to blocks (system2)

No of Blocks	1	2	3	4	5	6	7	…..	72
Task 1	10010	01100	10010	10100	10001	10111	10101	….	10010
Task 2	10010	01100	10001	10010	10001	10011	10100	….	10010

6- Conversion from binary to integer is applied for each block to generate its decimal number.

7- Suitable blocks (i.e. that are the same or close in value) for Task 1 and Task 2 are identified using the following equation:

$$Suitabl\ block\ (Sb_n) = Task1_n - Task2_n \qquad (4.6)$$

where n is the number of blocks. This shows the variance between two blocks for two tasks, with the suitable blocks achieving a value of 0 or close to 0. The setup threshold for our experiments is $-10^{\wedge 20} < Sb_n < 10^{\wedge 20}$ as shown in Figure 4.5.

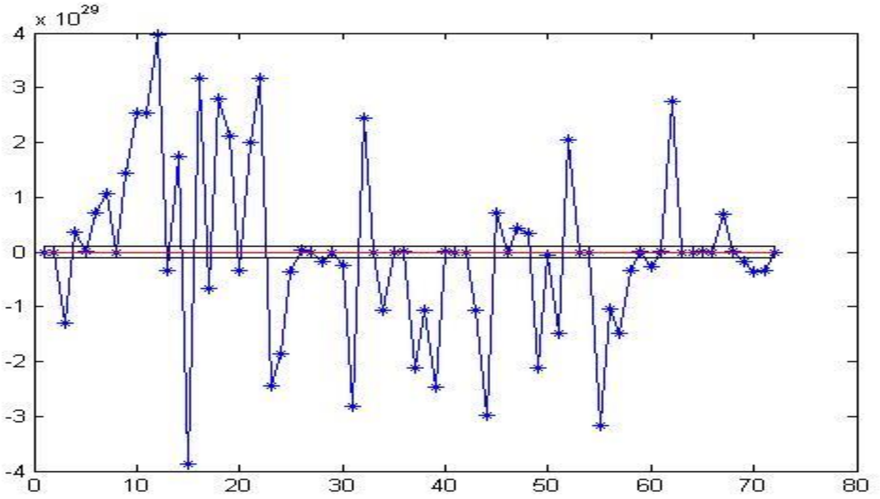


Figure4.5 Blocks of Interest (close to 0) -system 2-

Figure 4.5 shows the number of blocks that can be used for generating the secret key by Shamir's secret scheme.

8- A biometric encryption key is then generated by Shamir's secret scheme (cubic style). As explained in Chapter 3, section 3.3, a cubic equation needs at least four points to create the key, so the successful blocks will be considered as polynomial points. X_n coordinates refer to the number of successful blocks and Y_n coordinates refer to the successful blocks' values, as shown in Figure 4.6.

9- User function will be created number of blocks for X_n. For example, if successful blocks = {1, 2, 4, 6, 7, 33, 63, 72}, these will be stored in the user function to be called in the authentication process.

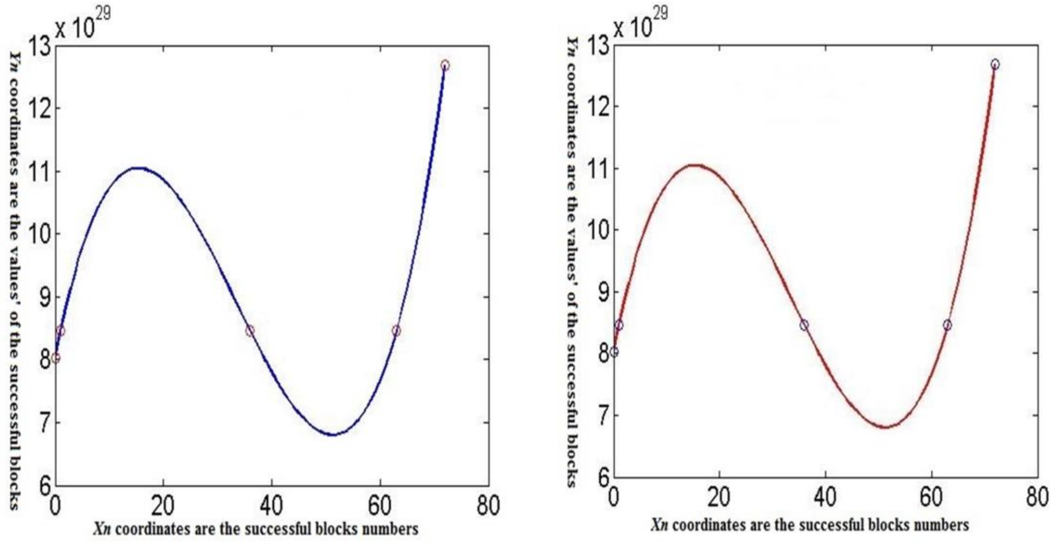


Figure4.6 Training Set Performance of system 2

Secondly, the authentication process can be described in the following steps:

1- The user is asked to provide 8 iris samples, divided between two tests using four samples each.

2- Pre-processing is performed for all testing samples.

3- Feature extraction and encoding is completed.

4- Each test undergoes XNOR to generate similarities between samples for each test.

5- Each test is divided into 72 blocks of 100 bits.

6- A binary converting tool is applied to convert the binary code in each block to a decimal number.

7- The user function is called, carrying the number of the block of interest.

8- Shamir's secret scheme is applied to release the key under cubic style conditions (requiring at least four points to release the biometric encryption key).

4.3.2.1 Experimental Results

In this section, we will cover the experimentation process, wherein we will provide the experimental result by determining the false acceptance and rejection rates for the overall authentication system for all users in an authentication scenario.

The investigation pertained to 300 users, using iris modality databases, with 20 samples per user. Taking eight samples for each user divided in two tasks, each set of three samples was used to create a user function and employed for testing. There were three tests for each user, each providing four samples. In total, 300 users * 3 testing = 900.

Passive forgeries were also investigated by employing 300 impostor users trying to gain access by claiming the identity of each of 300 genuine users. In total, 300 passive forgeries users * 300 genuine * 3 testing = 270,000.

The data experiments for iris used the CASIA-IrisV3-Interval database [92]. This justification of the use of this database has been provided in chapter 2.

Table 4.7 FAR and FRR performance (system2).

Limkey	Key limitation in y axis (Limkey – key < key < key + limkey)						
	1.00e+30	1.00e+28	1.00e+26	1.00e+25	1.00e+24 EER	1.00e+22	1.00e+20
FRR	0.1	4.4	7.3	8.2	16.7	41.8	61.6
FAR	97.9	60.5	29.7	22.7	18	9.1	5.01
No of attempts	1	2	3	4	5	6	7

Table 4.7 and Figures 4.7 and 4.8 illustrate the performance of the system. In Table 4.7, there are seven different conditions of the key limitation strategy, increasing the

biometric key limitation that applies in the polynomial curve to establish the variances in the FRR and FAR. A summary of overall performance of the system for all samples in terms of the FAR and FRR of the Shamir component of the system for the four polynomial points strategy is shown in Table 4.7 and graphically in Figure 4.8. In addition, Figure 4.7 shows training and testing sets for *User1*, resulting in successful extraction of the biometric encryption key.

Table 4.7 shows the result of seven orders of key limitation threshold in generating the key secret. The FRR on increasing key limitation is 0.1%, but the FAR is too high. By decreasing the limitation placed on the y-axis, the FRR is increased while the FAR is decreased. Attempts based on key limitation between -10$^{\wedge 24}$ < biometric key < 10$^{\wedge 24}$ provided the best result so far. In attempt number 5, the equal error rate shows 16.7% FRR and 18% FAR. However, attempt number 4 showed a much lower FRR than attempt 5 (8.2%), but the FAR increased slightly by comparison with attempt 5.

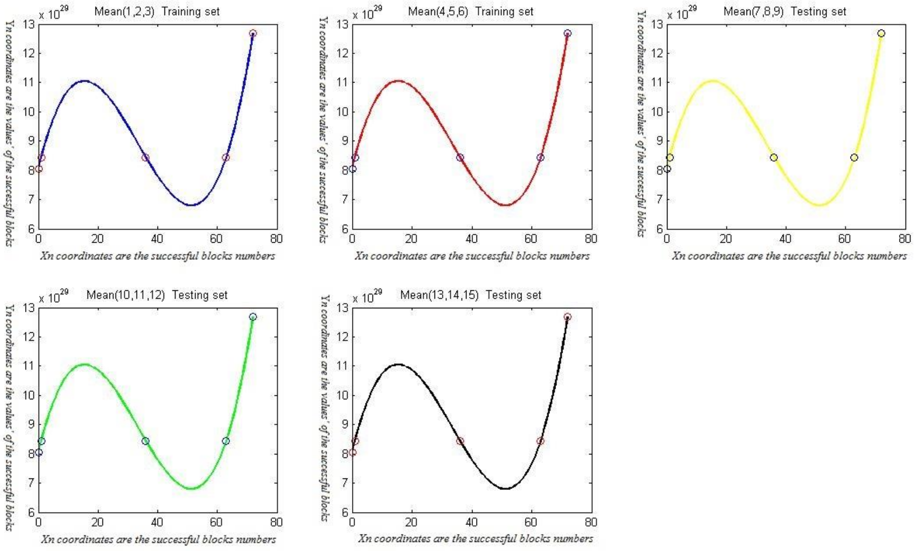


Figure4.7 Performance of Training and Testing Samples for User 1 (system 2)

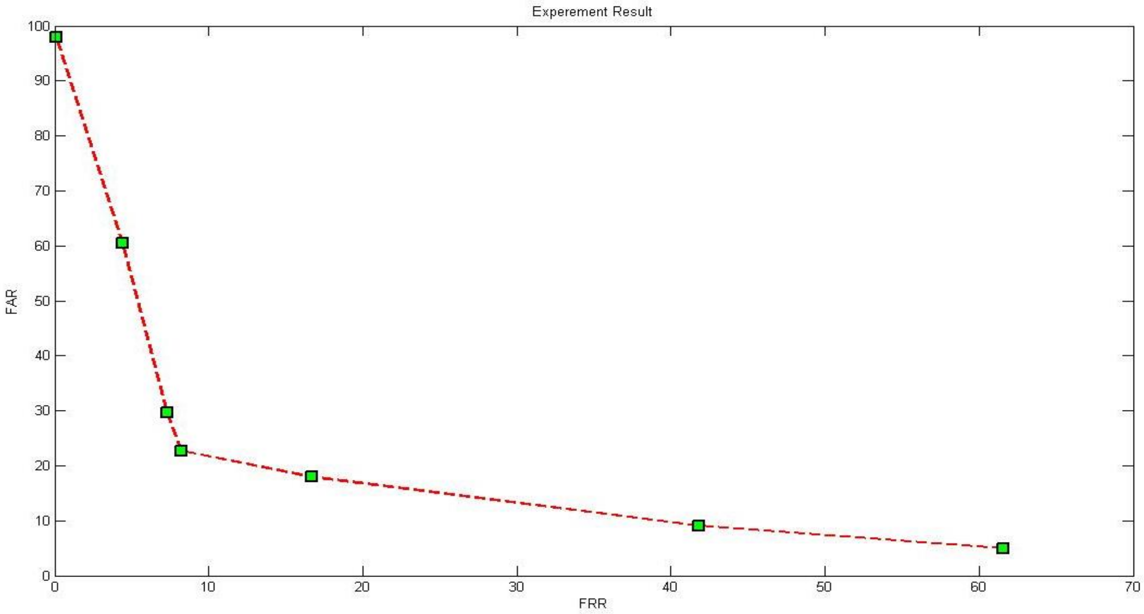


Figure4.8 ROC curve of FAR and FRR performance of system 2

These results are very interesting as they demonstrate the relative merits of the polynomial orders in relation to the desired performance of the system. However, the outcomes of these tests are in part dependent on the quality of the biometric samples employed and the algorithms employed for the iris modality.

The FRR and FAR ROC curve in Figure 4.8 shows how the performance of the system varies across seven different techniques with increasing key limitation in the vertical axis.

4.3.2.2 Summary

An attempt has been made to explore a new technique for generating an encryption key from biometric modality samples, using a minimal amount of stored data and examining the system reaction. Within the proposed scheme, the biometric template identified with each user need not be stored; only the vector of the number of successful blocks is stored, which does not identify the associated user. Although interesting, the results were also disappointing; we considered the scheme to have failed for the following reasons.

- The EER shows roughly 20%, which is unacceptable for an iris recognition system.
- We believe that the data stored ("successful blocks numbers" as we called this scheme) need to be further investigated because of the trick with the iris binary code, which contained just 0 and 1 values. It is not inevitable that the chosen blocks will be the same in every iris sample because iris localization, normalisation, segmentation and unwrapping will definitely affect the order of iris bits. Even if we consider the right/left shifting, the blocks still cannot be considered as standard in every iris sample for each user.
- There was no accuracy in generating the biometric encryption key. In this scheme, the biometric encryption key was extracted during enrollment, but in the authentication phase, the key extracted will be tested by limitation conditions in the y-axis, which affected the level of security and performance, returning a very high EER.

On the other hand, a lot was also learned from this scheme, as summarised below.

- Iris binary code is a vector that depends on pre-processing and feature extraction, making it hard to name a block or digit in the vector that is standard for a particular person, appearing every time as 0 or 1.
- Comparing any two iris binary code from the same person, the variance between them ranged from 5% to 25%, which is typical of the FRR in every iris recognition system. However, comparing two different iris binary codes from two different people, the similarity between them could be over 70%, which is extremely high in terms of these experiments and other approaches to iris recognition.

4.3.3 System 3: Template free key generation from Iris modality using GLCM

Due to failure of the system 2, the new approach has been adopted in which the encryption key was tried to be generated from the iris modality by using different iris feature extraction method called Gray Level Co-occurrence Matrix (GLCM) [94, 95]. In this scheme, feature extraction from iris modality using the GLCM and subsequently extracting the binary code by looking at derivatives, checking every interval in the GLCM matrix against the next interval. Simply put the derivative technique records whether the next interval increases or decreases. The reason for using another feature extraction method instead of the 2D Gabor filter (sections 4.3.1 and 4.3.2) is to recover most of the area in the iris feature extraction field and to establish the system’s reaction and response. To that end, we tested the method with template to assess FRR and FAR results and to decide whether this technique is suitable to be moved to a template-free system.

4.3.3.1 Methodology

This section introduces GLCM feature extraction, which will be applied in this scheme, and a derivative technique for converting the results of GLCM to binary code.

4.3.3.1.1 System Architecture

The architecture for this scheme will be divided into enrolment and authentication architectures, which will be discussed in detail in terms of the enrollment and authentication operations.

4.3.3.1.2 Enrollment operation

There are several steps in the enrollment process.

1- The user is asked to provide iris samples.

2- Iris localization and normalisation methods are applied [71-73].

3- Iris feature extraction using GLCM is applied [94]. As proposed by Haralick et al. [95], GLCM is one of the most widely-used approaches to extracting textural features and can be defined as follows. Suppose an area has N_c and Nr resolution cells in the horizontal and vertical directions, respectively, and N_g levels in the gray tone. Let $L_c = \{1, 2, .., N_c\}$ be the horizontal spatial domain, let $L_r = \{1, 2, \ldots, N_r\}$ be the vertical spatial domain, and let $G = 1, 2, .., N_g$ be the set of Ng quantized gray tones. The image I can be represented as a function that assigns some gray tone in G to each resolution cell or pair of coordinates in $L_r \times L_c$; $I : L_r \times L_c \rightarrow G$. The texture-context information is specified in a matrix of relative frequencies P_{ij} where two neighbouring resolution cells separated by distance d occur on the image: one with gray tone i and the other with gray tone j. P_{ij} can then be described in Eq. (4.1):

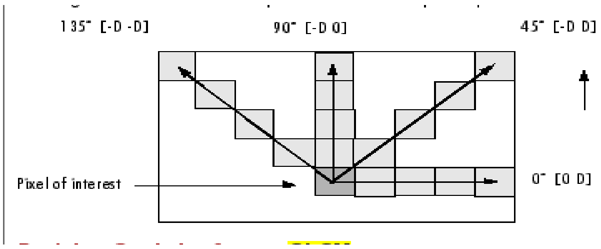


Figure4.9 Co-Occurrence Matrix Directions for Extracting Texture Features [94]

4- In subsequent experiments, the established experimental parameters for iris feature extraction using GLCM were used [95]. For this purpose, the biometric samples were converted into gray tones. The assumption was made that $2^8 = 256$ were enough to make the details of the sample visible to the human eye [95,159], therefore each biometric sample was stored in memory using 8-bit. As a result of GLCM method, the 8 x 8 matrix was generated. The number of gray levels increases the textural information, so is very important factor in

GLCM method. However, the number of gay levels, the more computational cost [95, 159]. Hence, 256 gray levels were assumed to be enough to reveal all nuances of the samples in the current study [95, 159].The established parameters involve the window size set at 8 pixels, gray levels at 256, distance between pixel pairs at 1 pixel, and 4 different angle sets are used: (0°, 45°, 90°, 135°) as shown in Figure 4.9.

5- The resulting matrix of four different angles is added together as per the following equation

$$GLCM_{addDegree} = (0Degree^{o} + 45Degree^{o} + 90Degree^{o} + 135Degree^{o}) \quad (4.7)$$

6- *GLCM*$_{addDgree}$ is reshaped in one vector and the mean is extracted. For more details, in the enrollment process, the user will provide 3 sets, each sets has 3 iris samples. Hence, the mean will be extracted from the 3 training sets for each user. In this step, a test is performed to examine the data after feature extraction as it shows in Figure 4.10 an example of 9 different users.

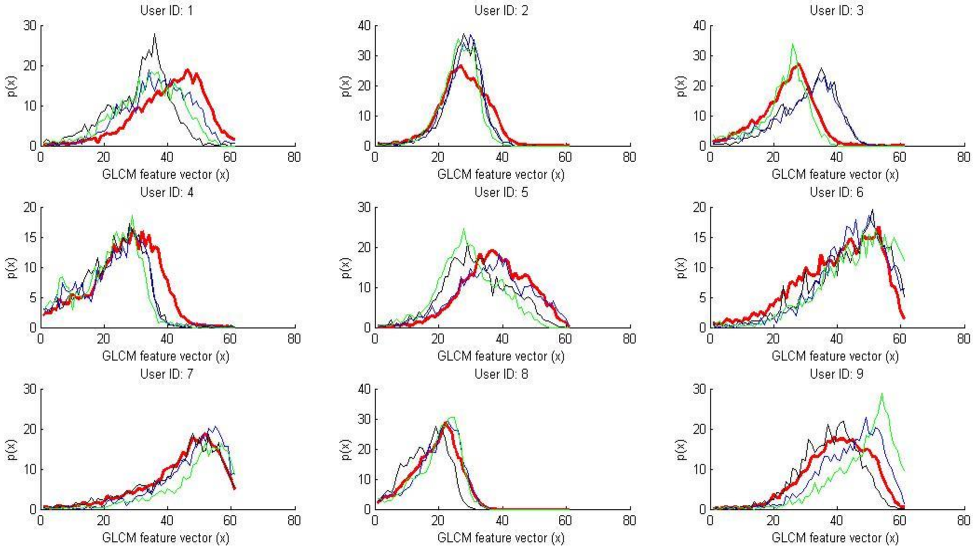


Figure4.10 Training Sets Performance for 9 Users of system 3

As already mentioned, the user is asked to provide nine iris samples in the enrollment phase, to be divided into three sets, each of three samples. Iris features are extracted by GLCM; adding all the results together, the reshaped matrix can be plotted for each set as shown in Figure 4.10, extracting the mean from the three sets provided as plotted in Figure 4.10 (red plot). The results between training sets are quite similar for each user. In addition, the results across different users appeared positive, which afforded a hope that this scheme might work, as shown in Figure 4.10.

7- Looking at derivatives (simply record whether next interval increases or decreases), a binary code was extracted from the mean of the reshaped $GLCM_{addDgree}$ as follows:

$$if\ (Mean\_of\_GLCM_{addDgree}(1,i) < Mean\_of\_GLCM_{addDgree}(1,i+1)$$

$$GLCM_{binarycode}(1,i) = 1$$

$$else$$

$$CLCM_{binarycode(1,i)} = 0 \qquad (4.8)$$

8- Following the limitation increasing rule or 6 bits rule as set in the experiment, if for example $Mean\_of\_GLCM_{addDgree}$ =[110100]; → then, $Mean\_of\_GLCM_{addDgree}$ =[111111], which means the curve values in the chart have increased. However, if $Mean\_of\_GLCM_{addDgree}$ =[110000]; → then, $Mean\_of\_GLCM_{addDgree}$ =[000000], which mean the values have decreased. Figure 4.11 shows the proposed derivative technique.

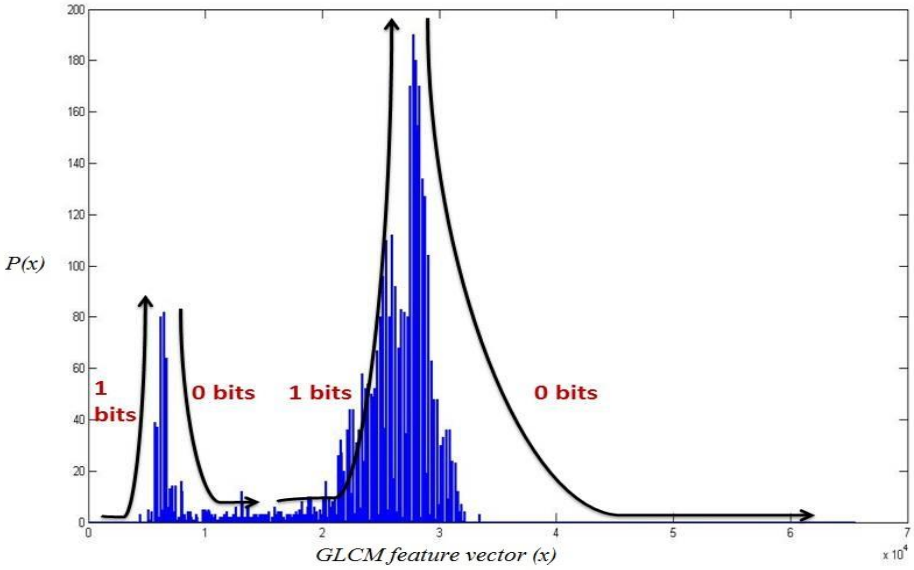


Figure4.11 Derivative Technique Proposal (system 3)

9- Binary code is then extracted from the user's training sets to be stored in the database for authentication purposes.

4.3.3.1.3 Authentication process:

Here, the system is tested by asking the user to provide nine iris samples to be divided into three tests, each using three samples. The pre-processing method is stratified for the iris localization/normalisation and segmentation/unwrapping process [71-73]. Subsequently, the GLCM method is applied to extract iris features, and steps 5 to 8 in the enrollment process are applied in the same way to generate the iris binary code. Finally, the Hamming Distance method is used to check the FRR and FAR against the Hamming Distance template as described in Chapter 3, section 3.2.2 [73, 74].

4.3.3.2 Experimental results

This section deals with the experimentation process, presenting the experimental results in terms of false acceptance and rejection rates for the overall authentication system for all users in an authentication scenario.

For the investigation, there were 300 users, using iris modality databases, with 18 samples per user. Taking nine samples for each user divided across three tasks, each had three samples for training sets and extraction of a template to be saved in the database and used in the authentication test. With three tests for each user, each providing three samples, in total 300 users * 3 testing = 900.

Passive forgeries were also investigated by employing 300 impostor users trying to gain access by claiming the identity of each of 300 genuine users. In total, 300 passive forgeries users * 300 genuine * 3 testing = 270,000.

The data experiments for iris used the CASIA-IrisV3-Interval database [92]. This justification of the use of this database has been provided in chapter 2.

Table 4.8 FAR and FRR Performance (system3)

Threshold	0.01	0.014 (EER)	0.016	0.018	0.02	0.024	0.026	0.028
FRR	42.888	18.777	11.444	7.222	3.666	1.444	0.777	0.777
FAR	3.590	18.964	32.489	47.535	61.505	82.316	88.790	93.367

Table 4.8 and Figure 4.12 present a summary of overall performance of the system for all samples in terms of the FAR and FRR, showing the results for eight orders of

threshold in terms of hamming distance. The result proved that this scheme has many dark points as the ERR with template is 18%, which is unacceptable for an iris recognition system. For more detail, the threshold for the EER is < 0.014, which means that similarities between two iris codes, whether or not they are from the same user, are extremely high at more than 90 %. However, the EER shows, surprisingly, that with a small threshold setting of 0.014, the FAR and FRR are equal at 18%.

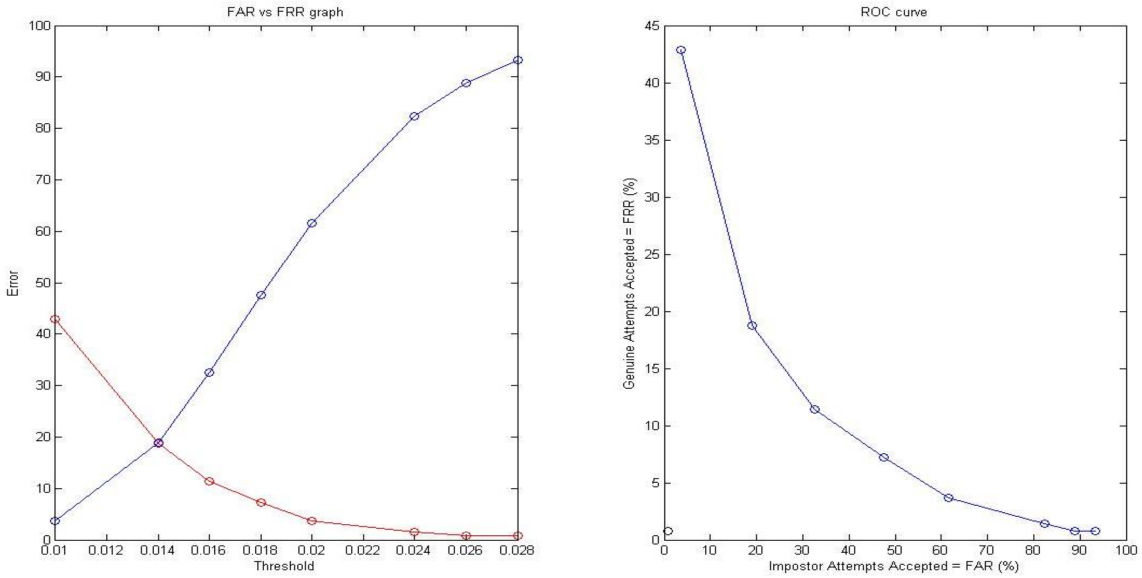


Figure4.**12 ROC curve of FAR and FRR Performance of system 3**

The FRR and FAR ROC curve, showing how the performance of the system varies across eight different techniques by increasing the threshold on the horizontal axis, is shown in Figure 4.12.

4.3.3.3 Summary

In exploring this new technique, using GLCM feature extraction [94, 95], it proved interesting at the outset, but when applied with template stored, the results were disappointing. For that reason, it was not possible to move on in this scheme to build an iris system that could generate the biometric key directly from the iris modality, for the following reasons.

- The ERR appeared close to 18% with a system using a template, which is unacceptable for an iris recognition system.
- We believe that the data stored (template) generated wrongly or that the scheme needs further investigation because the EER is 18% with small threshold set as shown in Table 4.8.
- The false rate is affected by the quality of iris images and implementation algorithms.

4.3.4 System 4: Template free key generation from Iris modality using feature distribution maps

The failures of the generation of encryption key from the iris modality by using the Gabor Filter (System 1) and the GLCM (System 2) showed that there are some variations and gaps in the chosen schemes for these systems which led to the failure of key generation directly from the iris samples. Consequently, it was assumed that introduction of feature distribution maps instead of the derivatives technique used in system 3 could provide the successful results in terms of generating encryption key from the iris modality.

This scheme investigated feature distribution maps using Gaussian distribution (normal distribution) [98-103] after extracting iris features by the GLCM method [94, 95]. Before moving to generate a biometric encryption key from typical user normalisation maps taken from user samples, the normal distribution was tested to decide whether or not to use it by calculating the FRR and FAR.

4.3.4.1 Methodology of Feature Distribution maps

The initial phase in the template-free biometric is the enrollment phase, which yields feature distribution maps for typical users of the system. Enrollment begins with presentation of known biometric samples from all users.

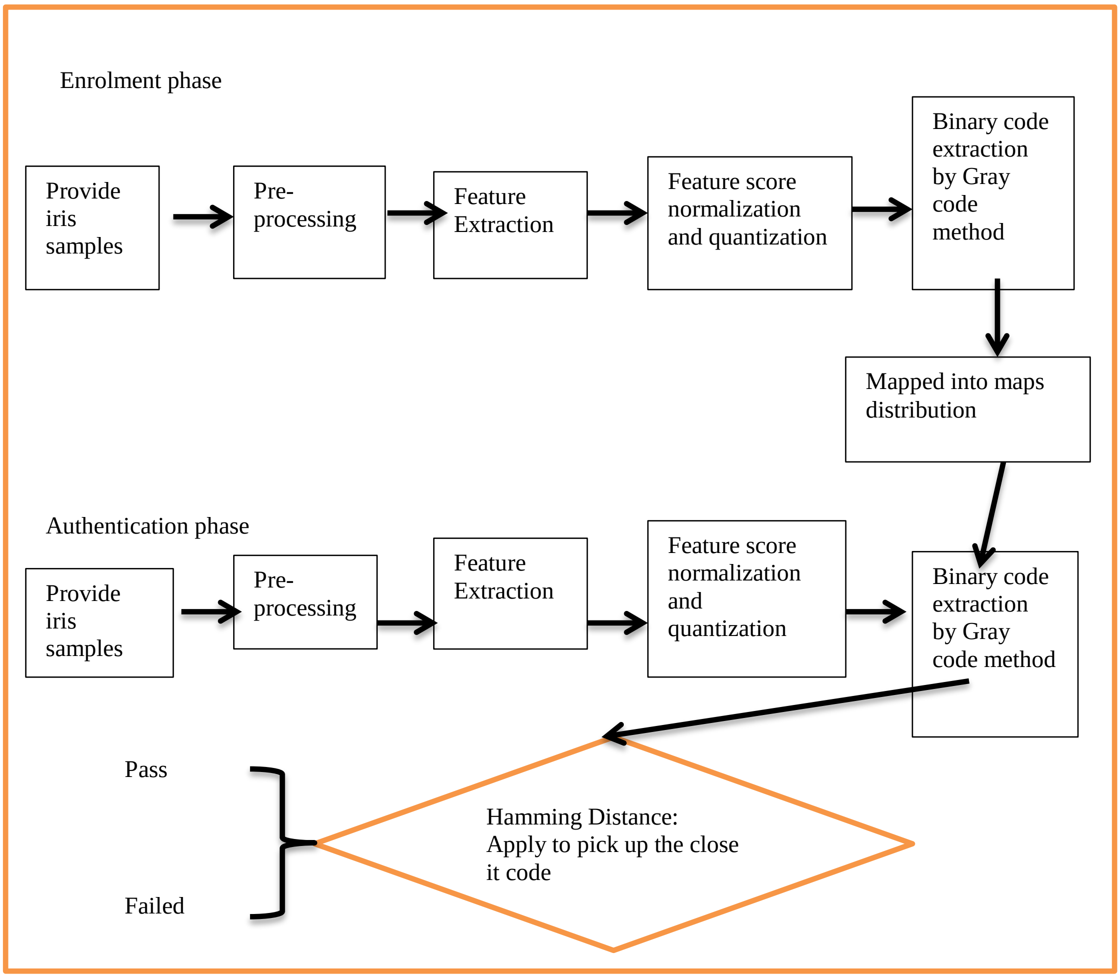

Enrolment phase
Provide iris samples
Pre-processing
Feature Extraction
Feature score normalization and quantization
Binary code extraction by Gray code method
Mapped into maps distribution
Authentication phase
Provide iris samples
Pre-processing
Feature Extraction
Feature score normalization and quantization
Binary code extraction by Gray code method
Pass
Failed
Hamming Distance: Apply to pick up the close it code

Figure4.13 Enrollment and Authentication Process of system 4

Figure 4.13 shows the structures for enrollment and authentication. Enrollment begins with the user providing iris samples followed by pre-processing and feature extraction. Feature score normalisation and quantization are then applied. Feature binary code will be extracted by the Gray code method to be mapped in the distribution map. The authentication process begins with the user presenting his/her iris samples. Feature extraction follows pre-processing and then normalisation and quantization. Feature encoding is generated for checking in the map's distribution using the Hamming Distance algorithm [73, 74].

4.3.4.2 Feature extraction

In this scheme, as mentioned earlier, the GLCM method is used to generate iris features as described in section 4.3.3.1.2 [94, 95]. Traditionally, Haralick's features include fourteen features; we chose the first twelve texture features for our analysis. Using GLCM to extract textures is sensitive to three factors: selection of window size, number of gray levels, and distances between pixel pairs. Specific window size, gray levels and distances of pixels pairs were set through a number of preliminary experiments. For GLCM, in subsequent experiments, the window size was set at 8×8pixels, gray levels at 256, distances between pixel pairs at 4 pixels, and 4 different angles were set (0°, 45°, 90°, 135°), yielding a total of 192 features in one vector for each sample.

4.3.4.3 Normalisation and Quantization

The min-max normalisation technique is employed in all training sets for all users as given by [98, 99]:

$$x^{norm} = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \qquad (4.9)$$

Min and max sample score values are those of users within each feature space, where x_i is the individual sample value, min(x_i) is the smallest value of x in all users for that feature space, and max(x_i) is the largest value of x in all users for that feature space.

A fixed quantization interval between 0 and 1 is used per feature space. For each value in the quantization interval, the mean and standard deviation per user is used to calculate the normal probability distribution function [100-103], given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (4.10)$$

P(x) is plotted against the quantization space Figure 4.14. It should, however, be noted that because there are increasing overlaps in user characteristics, p(x) values are further converted into Gray code [104] to test the system by hamming distance and to establish the false rejection and acceptance rates in deciding whether the scheme has passed or failed.

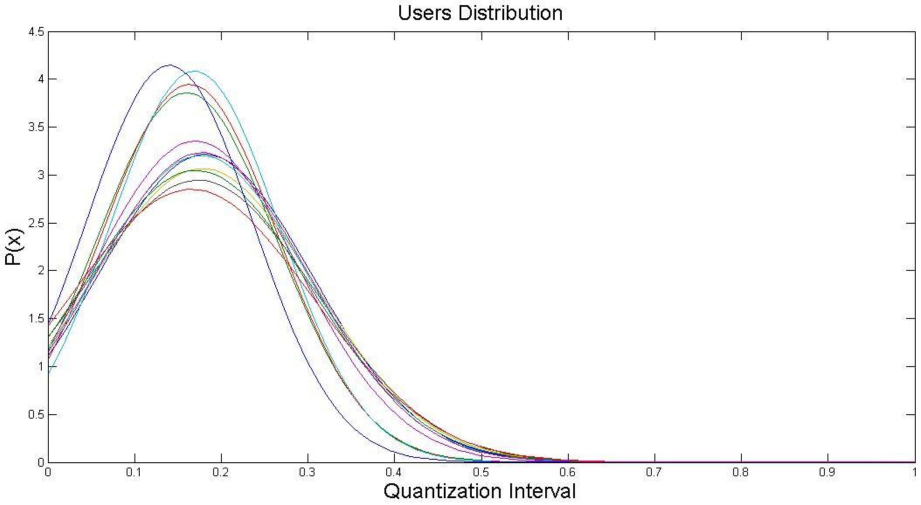


Figure4.14 User Distribution Maps (system 4)

4.3.4.4 Experimental results

This section deals with the experimentation process, providing the experimental results in terms of the false acceptance and rejection rates for the overall authentication system for all users in an authentication scenario.

For the investigation, there were 300 users, using iris modality databases, with 18 samples per user. Taking nine samples for each user divided across three tasks, each had three samples for training sets and extraction of a template to be saved in the database and used in the authentication test. With three tests for each user, each providing three samples, 300 users * 3 testing = 900 in total.

Passive forgeries were also investigated by employing 300 impostor users trying to gain access by claiming the identity of each of 300 genuine users. In total, 300 passive forgeries users * 300 genuine * 3 testing = 270,000.

The data experiments for iris used the CASIA-IrisV3-Interval database [92]. This justification of the use of this database has been provided in chapter 2.

Table 4.9 FAR and FRR Performance (system4)

Threshold	0.39	0.40	0.42 (EER)	0.43	0.45	0.46	0.48	0.49
FRR	89.78	60.00	24.00	2.22	0.33	0.33	0.22	0.00
FAR	0.78	6.33	28.19	72.43	97.99	99.74	99.96	100.00

A summary of overall performance of the system for all samples in terms of FAR and FRR is shown in Table 4.9 and Figure 4.15, summarising the results of 8 orders of threshold in term of hamming distance. The results demonstrate that this scheme has many drawbacks because the EER with template is 26%, which is not acceptable for an iris recognition system. In more detail, the threshold that got the EER is < 0.42, which means that the similarities between two iris codes, whether from the same user or not, is not close to 50% and this extremely low return will increase the FRR. The EER shows an expected result; in looking at Figure 4.14, presenting a plot of 12 users' probability distribution, it should however be noted that there are increasing overlaps in user characteristics.

The FRR and FAR ROC curve show how the performance of the system varies across eight different techniques by increasing the threshold on the horizontal axis as shown in Figure 4.15.

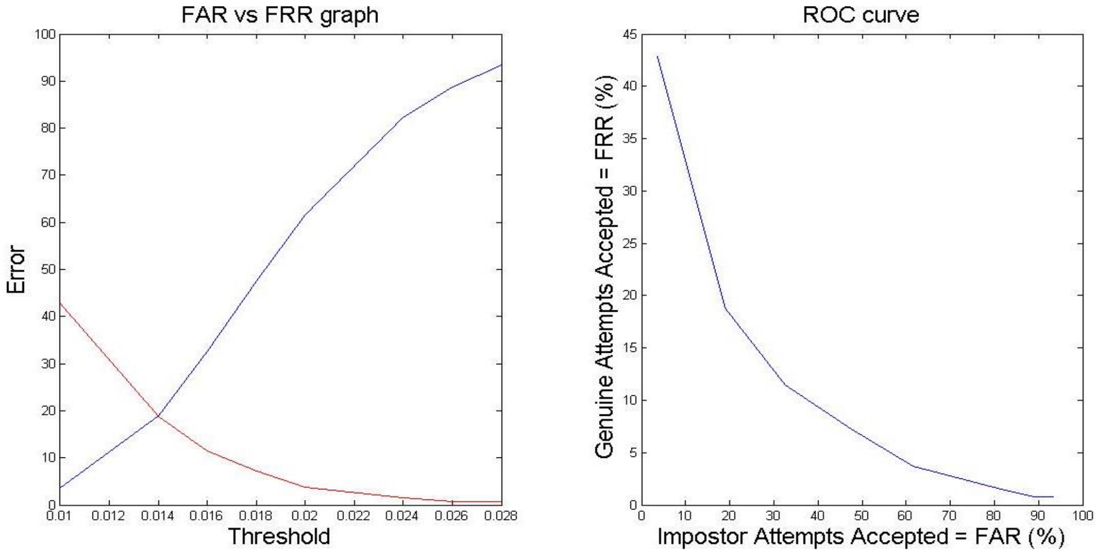


Figure4.15 ROC curve of FAR and FRR Performance of system 4

4.3.4.5 Summary

In attempting this new technique using Gaussian distribution maps with iris feature extraction by GLCM algorithm, the results were disappointing. The EER of 26% was unacceptable, although we did achieve a scheme in which the biometric template does not need to be stored. As a result, this scheme needs to be studied again with a view to building an iris system that can generate a biometric key directly from the iris modality. However, this current scheme has failed, for the following reasons.

- The EER (close to 26%) is unacceptable for an iris recognition system.
- There was a lot of overlap in the normal distribution maps across all users, which caused increasing error rates as shown in Table 4.9.

4.4 Summary

This chapter has introduced several schemes for integrating direct biometric key generation schemes with Shamir's secret sharing algorithm and others [11] in order to directly address the two disadvantages of revocability and exception handling. However, these schemes returned negative results, which means that the false rate (rejecting or accepting) is high or that the level of performance or security is low. Each scheme was presented in detail along with an explanation of how it works, why this scheme was used, and what feedback arose from it.

In summary, the study period has returned some negative results and disappointments, but much has been gained in terms of knowledge. In light of these issues, we opted to explore the design of two template-free biometric systems, in which biometric data can be mapped onto repeatable unique binary codes/strings in both, opening up the key only in the presence of biometric prints. The first system uses the Gray Level Co-Occurrence Matrix method (GLCM) [94, 95] to extract iris features, and the second system uses the 2D Gabor Filter [73]. To generate the secret key, Shamir's Secret Scheme [11] is applied to the three data points provided (right iris, left iris, and password). Two points out of three are then sufficient to release the biometric key by use of the linear equation technique, giving the user the desired flexibility while reducing the FRR. For both designs, a three-factor scheme is proposed, to include smartcard, password, and biometrics. Here, each component is crucial to the system required in revealing the biometric data specific to an individual, with the opportunity of either updating or revoking the key; all three factors would be required to compromise the integrity of the key. These two designs will be introduced in Chapter 5.

Chapter 5:

Biometric Security by a Direct Key Generation Scheme

5.1. Introduction

The main focus of this Chapter is to investigate the template-free biometric system by generating the biometric encryption key directly from the biometric samples. This chapter will address the following research objectives:

- To propose a multifunctional authentication system involving three modalities left iris, right iris and password. However, it should be acknowledged that password is not a standard biometric modality, but it is assumed to be a third biometric modality because the linear equation of Shamir Secret Scheme requires the three points to generate the key. Other biometric traits such as fingerprints could have been selected but their processing may take years to be completed, which is out of the scope of this PhD project. In order to avoid the lengthy and unnecessary prolongation of the project, the password has been assumed as a third biometric modality.
- To apply the Shamir Secret Scheme to generate the biometric encryption key directly from biometric samples (three points).
- To use a three factor scheme for increasing the security of the proposed system involving biometric, smart card and password.

In this chapter, outcomes of experiments to accomplish aforementioned objectives are presented. Furthermore, this chapter will also present the major novel findings and achievements obtained through these experiments including the successful generation of flexibly adjusted encryption biometric key length, enhancement of system's performance and security by using linear equations suggested by Shamir's Secret Scheme, and reduction of the data storage requirement of the system.

Recent work has explored the interaction between cryptography and biometrics, which are two complementary security technologies [50]. Biometrics is unique in the sense that it has the capacity to capture various forms of personal data such as fingerprint, iris and voice. The combination of biometrics and cryptographic techniques enables the user to access a signature that they have themselves created, with a high degree of confidence and accuracy. This makes it difficult for another person to generate the signature by stealing someone's token, or for the user to claim that a signature has been falsely generated from a stolen token when this was not the case.

Abraham et al [8] showed that the approach combining both cryptographic and biometric methods has the potential to generate a digital signature with high level of security and assurance. For instance, if the system will recognized any attempt made by the attacker to use the stolen key to create a signature or the user trying to falsely reject the signature claiming that signature is stolen when actually it is not. Previous research has attempted to address some security-related issues by introducing the signature verification pen and the IBM Transaction Security System with attached signal processor [8]. The main drawback of this technique is its complete dependence on the tamper-resistance of the hardware; in any case of token tampering, both keys and template disappear. However, attackers can break the token by other methods—specifically, API attacks on the token's software (associated with IBM design) and exploitation of chip-testing technology [51]. For that reason, an attempt has been made to evolve a more reliable combination of the three technologies: tamper-resistant hardware, biometrics and cryptography. However, algorithmic combination is rendered problematic by the noisy background of biometric data, offering only an approximate match to the stored data in the template. Similarly, cryptography presents its own problem of requiring exactly the right key. These previously-offered products depend on specific types of hardware device [50]; instead, there is a need to generate a protocol-based and more general strategy for the combination of both metrics and cryptography. Users are often reluctant to have that data stored in a central database, but user resistance to the use of biometric technology can be mitigated by credible assurances that their data will not be stored in this way. This in turn requires investigation of template-free biometric technology. Previous research has made some attempts in this regard by mapping personal biometric information into repetitive binary strings [20, 21, 52-54], which are mappable into the encryption key leading to opening of direct hashing [54, 55] or lookup tables [20, 52, 53].

The great promise of this approach was to create a system free of data storage concerns, but attempts to date at creating a template-free system have inherited several flaws and issues. The core drawback is the association of unreliability factors with the individual bits of information within the template. These issues have been further exacerbated by the noisy background created by biometric measurements of human physical attributes while the cryptographic requirement is to capture the exact key for further processing. Many attempts have been made to use keystroke patterns [21] and other human attributes—including voice [20], fingerprints [52, 56], facial attributes [53] and handwritten signatures [54]—for the derivation of biometric keys, in an effort to fill the gap between the noisy background of biometrics and the exactitude demanded by cryptography [21]. However, these attempts exhibit more than 20% FRR (False Rejection Rate) during verification, which is not practical for security applications [43]. Biometric data contain non-changeable human attributes specific to an individual, and so they cannot be subjected to any alteration. Key diversity is another issue in accessing biometric data; the user may wish to obtain separate keys for his bank and account and to use a computer at his workplace, enjoying the flexibility of revoking any key without affecting any other.

The openness of the biometric data is also a major problem, as the physical attributes of human beings such as low-quality fingerprint or iris data can be caught by any surface or hidden camera, respectively. By implication, a greater dependence on the use of biometric data leads to poorer levels of secrecy [51]. This means that it is not always safe to be reliant on biometric data on its own, especially with global distribution of an individual’s biometric data—for example, those data may be in use in multiple countries because of the user’s travel patterns. Furthermore, the key to biometric technology’s success is its social acceptance among stakeholders [43]. Public mistrust of the use of biometric data is a major issue for implementation; the fear that personal data will be misused arises especially where those data are stored in a central database, which is more prone to attack. Leakage of biometric data related to the health status of an individual and religious constraints in using such data are other obstacles to full application of biometric technology [51]. An earlier study by Davida et al. [57] followed this line of inquiry, attempting to derive a key from the iris code by application of error correction code methodology. A similar attempt was made by Daugman et al [50], in designing template-free biometric data containing only a string of error-correction data that could be used for the acquisition of biometric information but not to derive the key unless biometric prints were detected. They successfully designed a two-factor scheme that included biometric data and token, and they also demonstrated its further extension into a three-factor scheme by including a password as well. The only known issue found with this scheme was that the secret locking key was embedded in the smartcard, which could be easily targeted by hackers.

In light of these issues, we opted to explore the design of two template-free biometric systems, in which biometric data can be mapped onto the repeatable unique binary codes/strings in both, opening up the key only in the presence of biometric prints. The first system will use the Gray Level Co-Occurrence Matrix method (GLCM) [94, 95] to extract the iris feature and the second system will be used 2D Gabor Filter [73]. Moreover, in order to generate the secret key, Shamir's Secret Scheme [11] will be applied to the three data points provided (right iris, left iris and password). Two points out of three are then sufficient to release the biometric key by use of the linear equation technique, giving the user the desired flexibility while reducing the FRR. For both designs, we propose a three-factor scheme, including smartcard, password and biometrics. Here, each component is crucial to the system required for revealing the biometric data specific to an individual, with the opportunity of either updating or revoking the key; all three factors would be required to compromise the integrity of key.

5.2. Methodology

This section will outline two different systems of direct key generation, where the chosen modality is iris modality. The first system will apply the GLCM method for iris feature extraction, while the second system will use the 2D Gabor filter algorithm.

The major objective is to investigate the generation of biometric key from the multi-modal biometric samples directly. The selected modalities are left and right iris and password as three different modalities in order to investigate the multimodal biometric system. For direct generation, the critical step involves the features extraction from the iris modalities, which requires the choice of the sound method for this purpose. Based on the elaboration of feature extraction methods in Chapter 4 section (4.2.2), the researcher has selected two important GLCM and 2D Gabor filter techniques for the iris feature extraction, which are successfully applied by the previous researchers in iris recognition domain of the biometric system [71-74, 95]. For this research work, two methods were applied and tested separately to compare the quality of the biometric encryption key extracted from the samples directly.

5.2.1. System architecture

The existing literature on template-free biometric system indicates the application of two factor scheme (biometric and smartcard)[50], three factor scheme (biometric, smartcard and password) [22, 50], and personal biometric mapping scheme involving the mapping of binary codes into strings/look-up tables [50]. However, for this thesis, the three factor scheme involving biometric, smartcard and password has been proposed and tested.

The architecture of the proposed system based on the three-factor within a multimodal biometric for direct key generation is been presented in Figure 5.1 and 5.2. The important operations of this system are elaborated in the subsequent sections.

5.2.1.1. Enrollment operation

The aim of the current work was to build a multimodal biometric system with ability to generate the key directly from biometric samples. For this purpose, three modalities left and right irises and password were chosen. Iris has been selected for this thesis because it has zero-false accept rate according to several approaches described to construct the authentic and reliable biometric system in the literature [71-74], while password was taken as a third modality in terms of applying Secret Scheme which requires three points to fulfil the linear equation, and furthermore to enhance the performance and security of the multimodal biometric system.

In the enrolment step, as shown in Figure 5.1, the user will be asked to provide iris samples (left and right) and a password as the three modalities for creation of a multimodal biometric system. The iris samples provided will be processed by means of a standard iris pre-processing method that includes iris localization/segmentation and normalization/unwrapping iris imaging methods. Iris features will then be extracted; for present purposes, two systems have been created, each using different methods of iris feature extraction.

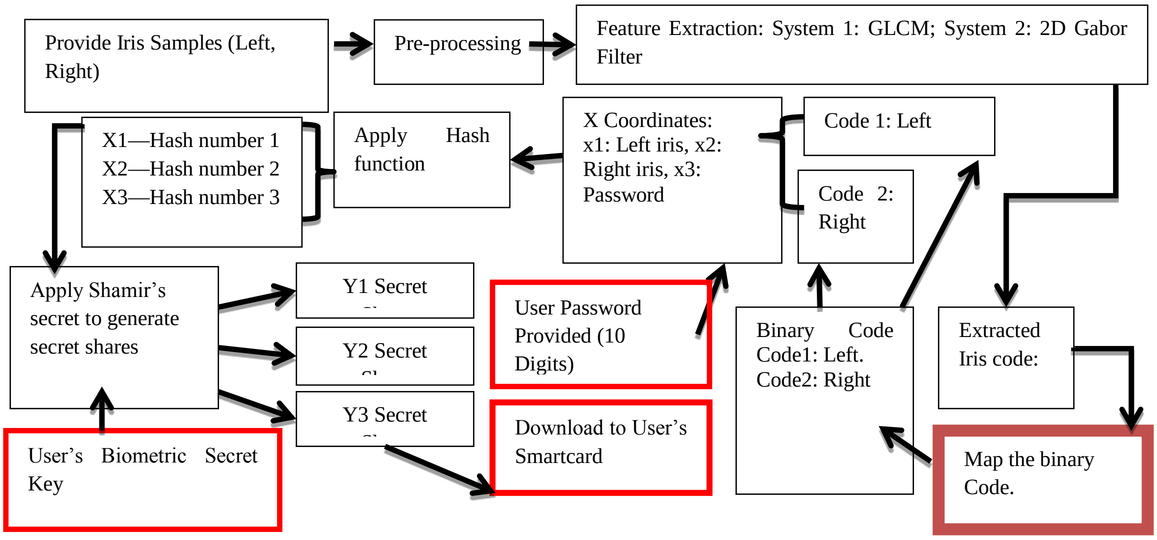

Provide Iris Samples (Left, Right)
Pre-processing
Feature Extraction: System 1: GLCM; System 2: 2D Gabor Filter
X1—Hash number 1
X2—Hash number 2
X3—Hash number 3
Apply Hash function
X Coordinates: x1: Left iris, x2: Right iris, x3: Password
Code 1: Left
Code 2: Right
Apply Shamir's secret to generate secret shares
Y1 Secret
Y2 Secret
Y3 Secret
User Password Provided (10 Digits)
Download to User's Smartcard
Binary Code Code1: Left. Code2: Right
Extracted Iris code:
Map the binary Code.
User's Biometric Secret Key

Figure5.1 Enrollment Operation

The chosen feature extraction approaches are the GLCM algorithm and the 2D Gabor filter algorithm. The iris binary code will be extracted and put through the mapping function to pick up the nearest code, applying the Hamming Distance method, which identifies the code with the lowest variance to the provided iris code in the lookup table [20, 52, 53]. Finally, the biometric secret key will be provided to work with the three features that have been created in the different modalities (left iris, right iris, and password), allowing Shamir's secret scheme to generate the secret shares to be stored in the smartcard after generating hash numbers for the left iris, right iris, and password by applying the hash function [105]. As a result, the smartcard and user biometric modalities (left iris, right iris, and password) in the authentication process will be able to release the biometric secret key by using Shamir's Secret Scheme.

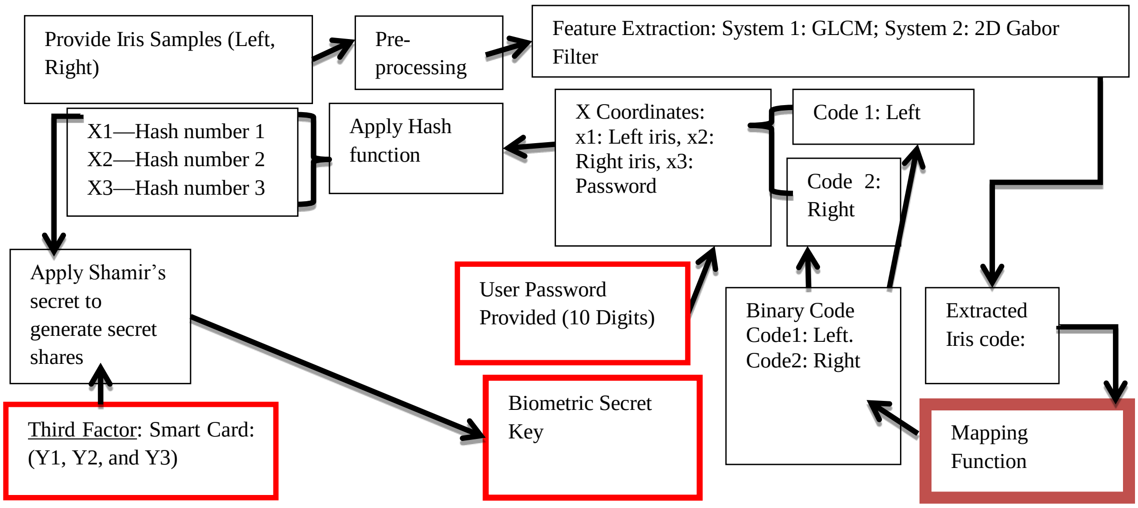

Provide Iris Samples (Left, Right)
Pre-processing
Feature Extraction: System 1: GLCM; System 2: 2D Gabor Filter
X1—Hash number 1
X2—Hash number 2
X3—Hash number 3
Apply Hash function
X Coordinates: x1: Left iris, x2: Right iris, x3: Password
Code 1: Left
Code 2: Right
Apply Shamir's secret to generate secret shares
User Password Provided (10 Digits)
Binary Code Code1: Left. Code2: Right
Extracted Iris code:
Third Factor: Smart Card: (Y1, Y2, and Y3)
Biometric Secret Key
Mapping Function

Figure5.2 Authentication Phase

5.2.1.2. Authentication Process

Figure 5.2 shows the authentication process. First, the user will be asked to provide the three factors: iris modality (left and right), user password and smartcard. As mentioned in the enrollment phase, the smartcard carries the Y_n coordinates (the secret share places in the y-axis), which must be provided, while the biometric iris samples (left and right) and the password will be considered as X_n coordinates (places on the x-axis). Shamir's secret scheme will then be applied to release the biometric encryption secret key. However, in this case, the linear equation technique has been used to release the biometric encryption secret key, using two polynomial points in the curve at least out of three (left iris, right iris and password. As a result, the biometric encryption secret key can be released by the biometric itself (left iris and right iris) in case the user password is forgotten. In addition, if the user fails to provide one of the iris biometrics (left or right), the biometric encryption key can be released from the other two iris samples and the user password to prove experimentally that the combination of the three-factor scheme and Shamir's Secret Scheme provides both high security and performance.

Authentication processing can be summarised as follows.

1. Three factors will be provided by the user (iris biometric, password and smartcard).
2. Iris sample pre-processing is described in section 5.2.2.
3. Iris feature extraction follows:
    - System 1 using GLCM [94, 95]
    - System 2 using 2D Gabor filter [73]
4. Iris feature encoding will be applied:
    - System 1 using global mean as discussed in section 5.2.3.1.2 below
    - System 2 using phase quantization as discussed in section 5.2.3.2.1 below
5. In both systems, the iris code will be checked in the lookup table, using the Hamming Distance method [71, 73] to identify the code with the lowest variance.
6. Results codes for left and right irises from the lookup table, along with the user password, will be hashed by the hashing function to be considered as X_n coordinates.
7. Shamir's secret scheme [11] will then be applied to release the biometric encryption key.

5.2.2. Pre-processing

5.2.2.1. Localization/Segmentation of the iris image

The iris region can be approximated by two circles—one for the iris/sclera boundary and one for the iris/pupil boundary. By using the circular Hough transform approach or Daugman's integro-differential operator, the radius and centre coordinates of the pupil and iris regions can be deduced. In the Hough approach, an edge map is first generated, and from this, votes are cast in the Hough space for the parameters of circles passing through each edge point. A maximum point in the Hough space corresponds to the radius and centre coordinates of the circle as best defined by the edge points. In Daugman's approach, an operator is used to locate the iris and pupil boundaries as well as the arcs of the upper and lower eyelids. Here, the Daugman approach will be used to deduce the pupil and iris regions as it is less computationally intensive and faster than the Hough approach and does not suffer from threshold problems [71, 73]. Figure 5.3 below shows a representation of the iris and pupil regions.

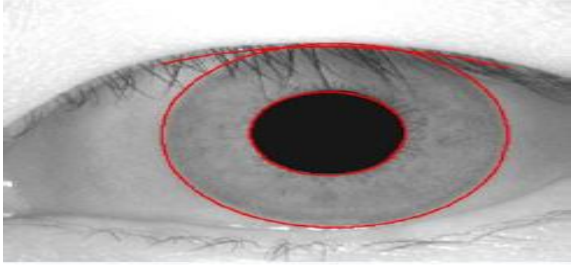


Figure5.3 Iris and Pupil regions [71, 73]

5.2.2.2. Iris Normalization/Unwrapping

Extracted iris images are diverse in dimension and orientation, owing mainly to pupil dilation but also to varying imaging distances, rotation of camera, head tilt, rotation of eye and other factors. The normalization process is an attempt to counteract these forces by remapping each point within the iris to a pair of polar coordinates (*r, θ*), where r is within the interval [0, 1] and θ is the angle [*0, 2π*] as illustrated in Figure 5.4 [71, 73].

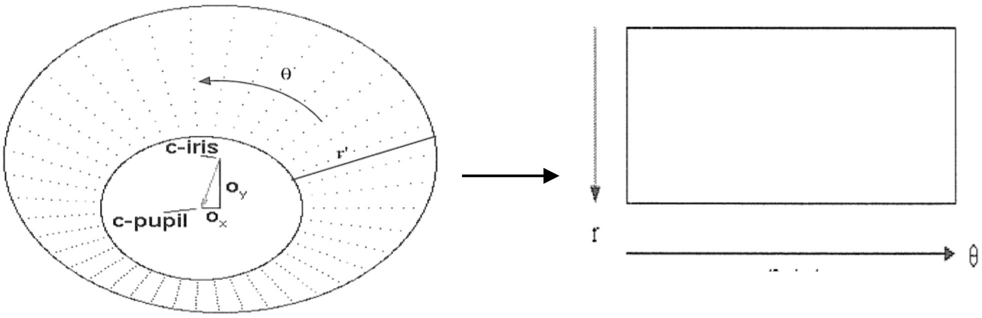


Figure5.**4 Unwrapping iris image [71, 73].**

For the remapping from *(x,y)* to *(r,θ)*:

$I(x, (r,\theta),y((r,\theta)) \rightarrow I(r,\theta)$ with

$x(r,\theta) = (1 - r)xp\ (\theta) + rxl(\theta)$

$y(r,\theta) = (1 - r)yp\ (\theta) + ryl(\theta),$ (5.1)

where *I(x,y)* is the iris image region,

(x,y) are the original Cartesian coordinates,

(r,θ) are the corresponding normalised polar coordinates

and xp, y_p and x_l, y_l are the pupil and iris boundaries along *θ* direction [71, 73].

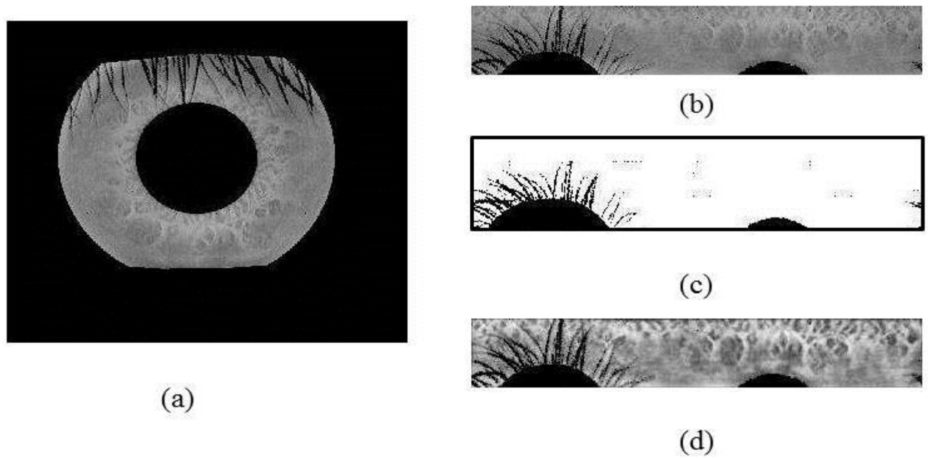


Figure5.5 Iris normalization: (a) Segmented iris image (b) Rectangular polar iris image (c) Noisy iris image (d) Enhanced iris image [71, 73].

Figure 5.5 (a) is a textured polar image. Figure 5.5 (b), applied to the image containing only noise, is a second binary image of the noise, which will be used thereafter to quantify the amount of noise in the image (Figure 5.5 (c)). Since the resulting image lacks contrast, it is preferable to enhance the textured image before analysing its texture (Figure 5.5 (d)).

5.2.3. Feature Extraction

As already mentioned, two different feature extraction approaches have been applied: Gray Level Co-occurrence Matrix (GLCM) [94, 95] and the 2D Gabor filter algorithm [73]. System 1 uses GLCM and System 2 uses the 2D Gabor filter.

5.2.3.1 System 1 Feature Extraction

In this phase, several iris samples are first provided by each user. Pre-processing will be applied to extract the iris and unwrap it. A GLCM scheme will be applied to generate the iris features within different pairings of pixels and angles [95]. Haralick’s first twelve features will be calculated and reshaped in one vector [94]. The features vector will be normalized within global min and max [98, 99] values for all users’ feature spaces, for which global max and min will be applied in the authentication phase and stored. The vector feature is converted to binary code by a decision created form, calculating the mean from each of 12 Haralick features from all users' training sets as the global mean, which will be stored. The global mean will be applied as an If condition during the enrollment and authentication phases for extraction of the binary code.

5.2.3.1.1 Feature Extraction Using GLCM

As proposed by Haralick et al. [95], GLCM is one of the most widely-used approaches to extraction of textural features. The approach can be defined as follows. Suppose an area has N_c and N_r resolution cells in the horizontal and vertical directions, respectively, and N_g level in the gray tone. Let L_c = {$1, 2, .., N_c$} be the horizontal spatial domain, $L_r = \{1, 2, ..., N_r\}$ be the vertical spatial domain, and let G = 1, 2, .., N_g be the set of Ng quantised gray tones. The image I can then be represented as a function which assigns some gray tone in G to each resolution cell or pair of coordinates in $Lr \times Lc; I : Lr \times Lc \rightarrow G$. Texture-context information is specified in a matrix of relative frequencies P_{ij} with two neighbouring resolution cells separated by distance d occurring on the image, one with gray tone i and the other with gray tone j. P_{ij} can then be described by Eq. (4.1).

Traditionally, Haralick' s features are fourteen in number; for the present analysis, we chose the first twelve texture features. Using GLCM to extract textures is sensitive to three factors: selection of window size, number of gray levels and distance between pixel pairs. We will set specific window sizes, gray levels and distances between pixel pairs through a number of preliminary experiments.

5.2.3.1.2 Feature Score Normalisation

As previously reported, [106] normalisation is necessary in order to transform the scores of the various features into a common scale [99, 107]. In this case, normalisation is applied to the feature values of all users within a given feature space in order to reduce the effect of score variability. Therefore, normalisation is best performed when all calibration samples have been acquired in order to ensure that good estimates of each feature’s global maximum and minimum are obtained which can be calculated by Eq.(4.9). Each feature’s global maximum and minimum is stored, along with the global mean of each feature space, so that during operation the feature vectors can be normalised to the standardised range. Normalisation essentially standardises the feature space to a range of [0, 1] [98, 99]. Subsequently, feature vectors can be converted to a binary code by a global mean decision to enable the probabilities to be tested and to release the user secret key directly by using Shamir’s secret scheme.

The feature vector must be converted to binary code according to a global mean decision, as follows equation (5.2):

```
If ($FV(f_n) \geq \text{Global mean Decision}(f_n)$)        (5.2)
    $FV(f_n) = 1;$
Else
    $FV(f_n) = 0;$
End
```

Where FV is the feature vector and n is the number of elements for which size of feature vector is 1*n.

It should, however, be noted that there are overlaps in users' characteristics, increasing the likelihood of one user's unique information being similar to that of another. To overcome this problem, values within the range of an acceptable percentage deviation from the highest probability point are considered most useful for generating the encryption keys for the user from that particular feature.

5.2.3.2 System 2: Feature Extraction using 2D Gabor Filter Algorithm

There are various available methods for performing feature extraction, including wavelet encoding [108], Gabor filters [71, 74], Log-Gabor filters [109, 110], Haar wavelet [111] and the Laplacian of Gaussian filter [112, 113]. For present purposes, the 2D Gabor filter approach was chosen for extraction of iris features. The iris code will be extracted by demodulating the unwrapped iris image using complex-valued 2D Gabor wavelets [73], where the 2D Gabor filter equation can be represented in Eq.(3.1).

For present purposes, frequency values are assigned five different values (2, 4, 8, 16, and 32), and orientation values (θ) are assigned four different angles (0°, 45°, 90° and 135°) [73, 74].

Prior to applying a 2D Gabor filter to the unwrapped image, a mask will be created that will be used to identify non-iris pixels, corrupt areas and specular reflections by checking the grayscale intensities of the image. A 2D Gabor filter will then be applied to the unwrapped iris image and its corresponding masks, generating a real and imaginary part for both, which will then be used in the process of feature encoding.

5.2.3.2.1 Feature Encoding

The encoding process will produce a bitwise template, containing a number of bits of information known as iris codes. It will also produce a corresponding noise mask, which will identify corrupt areas within the iris pattern. This encoding process is achieved through the process of phase quantisation, in which there is a quadrant containing a 2-bit binary number; position in the quadrant cannot be determined by comparing the real and imaginary parts of the filter [73]. However, by using the process illustrated in Figure 5.6, it will be possible to generate the iris codes from the filters produced for the iris image and its mask.

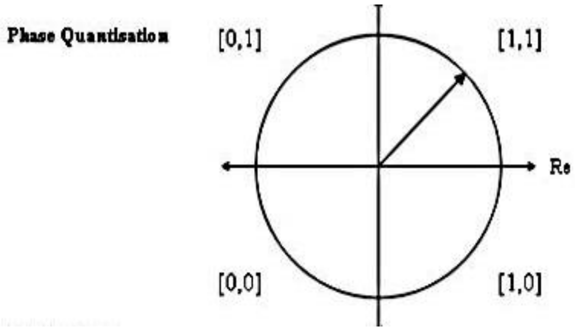


Figure5.**6 An illustration of the feature encoding process [73]**

5.2.4 Mapping Function Creation

This section will explain how the map function has been created and used in both System 1 (using the GLCM algorithm) and System 2 (using the 2D Gabor filter method). The mapping function will provide a lookup table [20, 52, 53] of binary string codes so that each code is of the same length as the user's iris binary code. The main purpose of creating a mapping function like this is to establish how both systems (1 and 2) will react, positively or negatively. However, the binary codes in the lookup table do not identify the associated user; here, the lookup table has been created by unknown iris samples taken randomly to create one lookup table for all users, as shown in the example in Table 5.1.

Table 5.1 Mapping Function

Mapping Function	
Number of codes	String binary codes
1 2 … N	‘101001011110’ ‘101001110100’ … ‘101001000010’

The lookup table could be extended and created in several styles, as follows:

1- One lookup table for all user sets (as in Table 5.1), which is used in this research

2- Separate lookup tables for each user set

5.2.4.1. System 1: Creation and processing of mapping function using GLCM

As mentioned in section 5.2, this research proposes to create a lookup table from unknown iris samples with no relation to experimental training and testing users' sets in order to establish the reaction of the system to the mapping function. The processing steps for mapping function creation in System 1 (using the GLCM method) are as follows.

1- Unknown iris samples collection using irises taken randomly from database

2- Pre-processing of iris samples as mentioned above in 5.2.2, involving:

- Localization and Segmentation of iris images [71-73]
- Normalization and Unwrapping of iris images [73, 74]

3- Feature extraction using GLCM [94, 95]:

- Generation of iris features within different pair of pixels and angles [94]

- Haralick’s first twelfth features calculated and reshaped in one vector [95]
- Feature vector normalized within global min and max values [98, 99]

4- Encoding process to convert the feature vector to binary code by:

- Stored global mean, as mentioned in section 5.2.3.1.2

5- Binary code mapped into the mapping function

5.2.4.2. System 2: Creation and processing mapping function using 2D Gabor Filter

Here, the mapping function will be created as above, from unknown iris samples, using the 2D Gabor filter method as follows:

1- Unknown iris samples collection, using irises taken randomly from database.

2- Pre-processing of iris samples as explained in section 5.2.2.

3- Feature extraction using 2D Gabor filter [73]

4- Encoding of feature vector by phase quantization as discussed in section 5.2.3.2.1.

5- Binary code mapped into the mapping function.

5.2.5. Hamming Distance

Hamming distance is a process for calculating the similarities between two binary codes, in which lower hamming distance denotes higher similarity [73, 74], as in the following equation:

$$\mathrm{HD} = \frac{||(\mathrm{CodeA} \otimes \mathrm{CodeB}) \cap \mathrm{maskA} \cap \mathrm{maskB}||}{||\mathrm{mask\ A} \cap \mathrm{mask\ B}||} \qquad (5.6)$$

where *Code A* and *Code B* refer to the iris codes, and mask A and mask B refer to the iris masks.

Hamming distance is generally used to differentiate between the genuine and imposter users in iris recognition based biometric system. This function is executed by comparing the binary codes coming from input and template iris samples.

For the current research work, Hamming distance is used to obtain the string iris code from the lookup table that achieves the highest similarity with the user iris sample in the enrollment and authentication operations, applying a hashing function to that string from the lookup table to generate the X_n coordinates as used in Shamir's secret scheme portion.

5.2.6. Shamir's Secret Scheme

For the current project, Shamir's Secret Scheme was chosen due to the following two fundamental advantages:

The length of biometric encryption key can be generated with adjustable length due to placement of encryption key of y-axis which allows the researcher to have infinitive space on vertical-space to adjust the length of key. This is the novelty of the system which in contrast to other biometric systems provides the option to administrator to adjust the encryption key of any securable length.

Shamir's Secret Scheme also provided the option of application of linear equation which can be defined by supplying two points. This allowed the user a flexibility to give any two points (biometrics) out of three provided points to release the biometric encryption key

As detailed in Chapters 2 and 3, this technique has three different styles: linear, quadratic and cubic equations. Shamir's secret generally uses three parameters: X_n Coordinates (places in the x-axis), Y_n Coordinates (places in the y-axis) and the biometric secret key (on the y-axis). Figure 5.8 below shows an example for the linear equation technique used here.

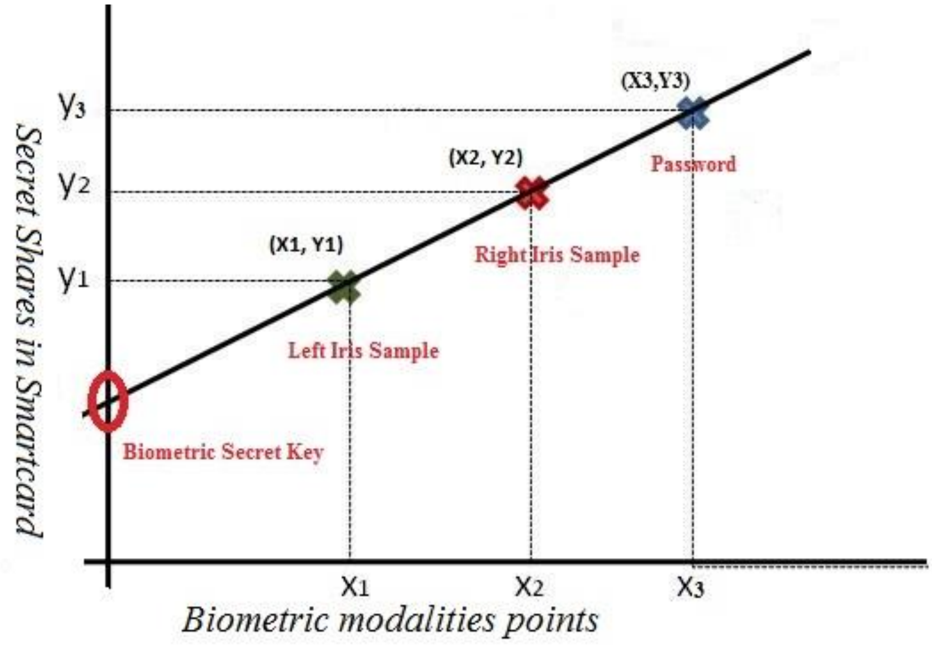


Figure5.**7 Linear Equation style**

Figure 5.7 shows that three points in the line (X_n,Y_n) identify where the biometric secret key crosses the y-axis. The three points are provided by the left iris (x_1), right iris (x_2) and password (x_3) as X_n coordinates while y_1, y_2 and y_3 are the secret shares stored in the user's smartcard. The reasons for using Shamir's secret scheme are as follows.

1- In the enrollment process, the user's biometric secret key will be produced and the X_n coordinates will be generated to allow Shamir's secret to extract the secret share to be downloaded to the smartcard.

2- Once the secret shares are saved in the smartcard, the biometric secret key will be deleted.

3- In the authentication process, the linear equation technique enables release of the biometric secret key by two polynomial points out of the three shown in Figure 5.8 as in any of the following possible combination:

- Left iris and right iris.
- Left iris and password.
- Right iris and password.

For these purposes, Shamir's secret offers flexibility and security performance.

5.2.7. Three-Factor Scheme Combination (Biometric, Smartcard and Password)

This section details the elements of the three-factor scheme (Biometric, Smartcard, and Password), whose purpose is to achieve an acceptable security level as well as enabling less information to be stored on the smartcard.

5.2.7.1. Smartcard

There are three different types of smart card used authentication of the users, which are Mactch-on-Card (MoC), *YesCard* and *NoCard*. They may be susceptible to attack in traditional biometric systems, however by using symmetric encryption primitives, these attacks can be thwarted [114, 119]. Structurally, the smartcard chip contains a communication port for exchanging data and control information with the external world. It is the ideal container for cryptographic secrets such as symmetric secret keys and asymmetric private keys, and the use of a contactless smartcard chip is now mandatory for numerous travel documents [114, 119] and national ID programs. Here, the role of the electronic chip is to authenticate the document (*something-we-have*) using cryptographic tools [116, 119].

We set the goal of reducing the information stored on the smartcard in light of other approaches that used the smartcard with incorporated biometrics. These can be built in two different ways. First, a biometric template can be saved on the smartcard for identification purposes, which is risky in terms of security because of the greater amount of information saved on the smartcard. The second approach saves an unlocking encryption key and a key locked by biometric modality, which still means that important information is available on the smartcard if an attacker succeeds in unlocking the key [116, 119]. In contrast to these approaches, our smartcard stores only secret shares (Y_n coordinates) and does not itself provide any important information to potential hackers, showing the strength of Shamir's secret scheme, as well as the strength of the combination of the three factors (smartcard, password and biometrics).

5.2.7.2. Password

A password is certainly the oldest and best known solution for providing user authentication. Although this sounds simple to use, care must be taken about how the password is communicated. A secure channel must be provided between the authenticator (the system or person controlling the authentication) and the applicant (the candidate user), notably at the primary exchange to set up the shared password. If these minimal precautions are not taken, very simple man-in-the-middle attacks such as eavesdropping become possible. Furthermore, it is also used to demonstrate the secret sharing primarily so that other selected biometric modalities can be considered but for ease of use, the example passwords are used.

One of the most widely used password-based authentications is the PIN (Personal Identification Number) code, authorising the use of a banking card. In this case, precautions must be taken when entering the PIN code, as it is very easy to spy over the shoulder of the user in an attack known as "shoulder-surfing" [119].

For the purposes of this research, it was decided that the user should create their own password under several conditions, as follows:

1- Minimum length 10 digits.
2- Must contain one uppercase letter (A-Z).
3- Must contain one lowercase letter (a-z).
4- Must contain one number (0-9).
5- Must contain one symbol.

5.2.7.3. Biometric

Biometric authentication has the advantage of checking the user’s unique personal characteristics. The use of biometric data is now mandatory for numerous travel documents [115, 119] and national ID programs. Here, the idea was to use multimodal biometric systems to generate the biometric secret key directly. As investigation in each modality would be time-consuming, it was decided to use the iris modality, which can effectively be considered as two different biometric modalities (left and right iris). Because Shamir’s secret first style needs three modalities to work, we decided to use a password as the third modality, as shown in Figure 5.8 above, to explore this new enhancement of biometric security by direct key generation.

5.2.7.4. Three-Factor Scheme

Three factor scheme involving biometric password and smart card has been tested by [73, 119], which is produced the significant high security level that is why the current study followed the same approach to enhance the security and performance of the system. It also allowed the system to perform with the biometric and smart card while missing the third factor (password).

Any combination of two from the three authentication factors will sacrifice at least one of the relevant security criteria. *Something-we-know* with *something-we-are* will sacrifice privacy because no personal device entails the use of a database to centralize all biometric data. *Something we have* with *something-we-are* will sacrifice a secret in the architecture since biometrics are public data. *Something-we-have* with *something-we-know* will sacrifice real user authentication since there is no proof of a link between the user and their card or PIN code [119].

Moreover, some applications need to duplicate one factor in the authentication scheme; sometimes, we need to show both ID card and passport, or we need to present both face and fingerprints, or we need to enter a password to log in to a system and then enter another password to access the application we intend to use. For instance, the use of smartcard, PIN code, fingerprints and facial recognition is still three-factor rather than four-factor authentication (as is sometimes suggested in press releases and marketing messages). In today’s digital world, most communication channels are insecure as the first goal is to provide user convenience. When delivering a password or biometric data, particular attention must be paid to this communication channel to guard against very simple methods of bypassing authentication. The use of cryptographic tools is mandatory to ensure the security of any three-factor authentication; the ultimate solution is to combine three-factor authentication with a Public Key Infrastructure (PKI). However, PKI is difficult and costly to set up, manage and maintain, and simpler solutions must be considered for providing secure communications over insecure channels [118, 119] and ensuring confidentiality and integrity of data [117, 119].

Here, three factors have been used in the following ways:

1- Iris modality (left and right) considered as *something-we-are,* where the user will be asked to provide their iris samples during the enrollment and authentication phases.

2- Password as *something-we-know*, considered for present purposes as a third modality, where the user will be asked to create their password during enrollment and will then be asked to provide it during the authentication phase.

3- Smartcard as *something-we-have* that will carry secret shares (Y_n coordinates), using Shamir's secret scheme first style called linear equation technique, where secret shares will be created in the enrollment phase and saved on the smartcard.

5.3. Performance and Results of Experiments

Experimental evaluation is now introduced to determine false acceptance and rejection rates for the overall authentication system for the iris database. A comparison will also be presented between the proposed systems and other related work.

5.3.1. Experimentation data

The data experiments for iris used the CASIA-IrisV3-Interval database [92]. This justification of the use of this database has been provided in chapter 2.

In this system, 300 users were created; each user had 20 iris samples, which were divided into two parts: 11 samples for the enrollment system and 9 samples for the authentication system [92].

5.3.2. Experimentation Setup

The proposed enhancement of iris security by a direct key generation system, which includes image pre-processing, iris feature extraction in both system 1 using GLCM [94, 95] and system 2 using 2D Gabor filter [73], feature encoding, mapping function and Shamir's secret scheme [11], was implemented on the MATLAB platform. The inner and outer iris area localization algorithm, the GLCM algorithm and the 2D Gabor filter algorithm were implemented from scratch. In the present experiment, 5000 tests were created, each providing 11 randomly selected samples for training and 9 for testing.

In subsequent experiments, using GLCM, the window size was set at 8×8 pixels, gray levels at 256 and the distances between pixel pairs at 4 pixels. For each feature extraction, the first twelve Haralick features are calculated [95], and so the feature vector will be (4 angles * 4 pixel pairs * 12 Haralick features = 192 features vector).

The proposed scheme for System 1 has been tested in standard iris recognition structures by using the template in a database to see the reaction to the proposed scheme (using GLCM) by ascertaining the Equal Error Rate (EER) performance, as shown in Figure 5.8.

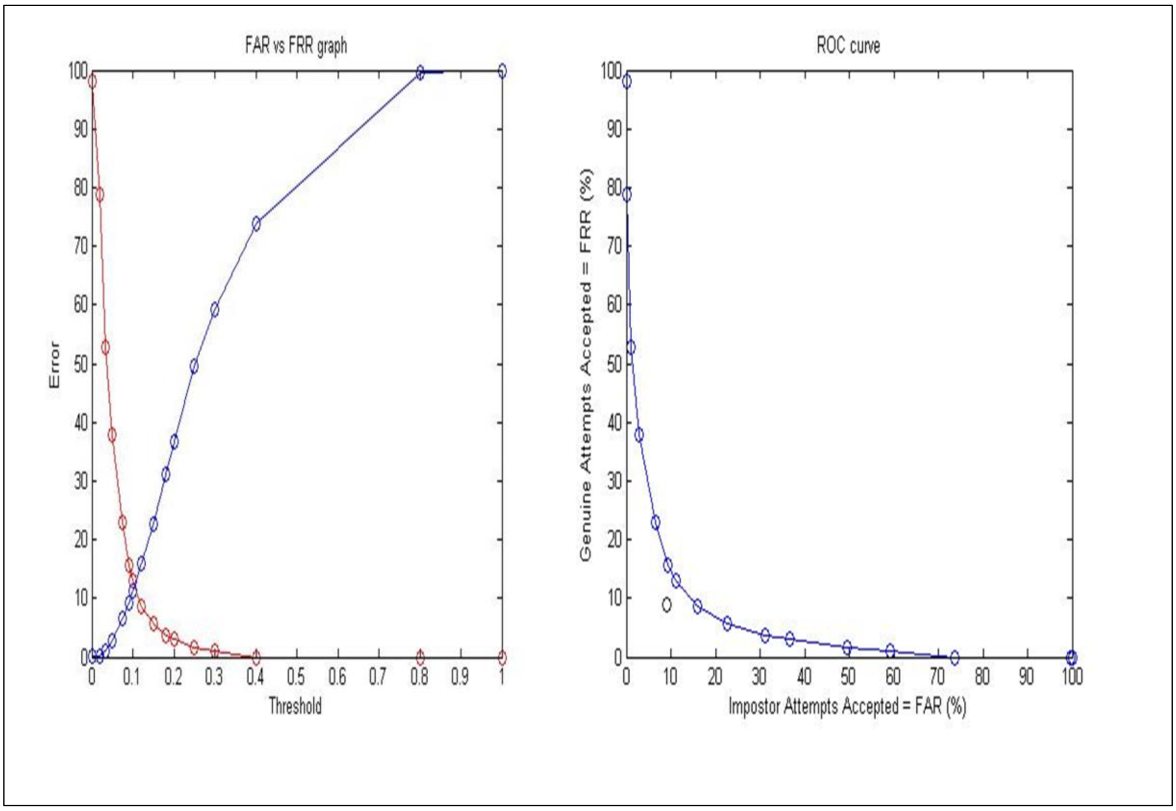


Figure5.**8 FAR vs FRR and ROC curve for FAR and FRR** performance **using GLCM**

Figure 5.8 shows the performance of FAR and FRR for the proposed System 1 (using GLCM methods), which shows that the EER is 13%, giving an accuracy rate of 90%. This seems close to standard iris recognition results using the same GLCM methods as Ying Chen et al. [94], who proposed a combination of GLCM feature extraction with a multi-channel 2D Gabor filter. The GLCM method showed an average accuracy rate (before combination) of 93%, suggesting that the proposed System 1 method seems close to standard iris recognition, using the same feature extraction method as proposed in Yin Chen et al. [94].

For the 2D Gabor filter, in view of its symmetrical character, we set four different direction values for $\theta = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}$, and six values of 2, 4, 8, 16, and 32 for the central frequencies (*f*), giving 4×5 = 20 filters with different frequencies and directions. The average result parameters will be taken to apply in the proposed system 2.

To generate the parameters for the average result, we tested the database for the proposed System 2 (using 2D Gabor filter in standard iris recognition structure) by providing a template in the database to compute FRR and FAR. The average results for FAR and FRR for each of the filters are shown in Table 5.2 below.

Table 5.2 Average results of FRR and FAR for each filter of 2D Gabor filter

2D Gabor filter								
(*f*)	(θ)							
	0^{o}		45^{o}		90^{o}		135^{o}	
2	1.12%	0.02%	1.75 %	0.02%	1.12%	0.02%	1.75 %	0.02%
4	1.91 %	0 %	2.3 %	0%	1.91 %	0 %	2.3 %	0%
8	1.87 %	0.02 %	1.92 %	0%	1.87 %	0.02 %	1.92 %	0%
16	1.34 %	0.02 %	1.6 %	0.02 %	1.29 %	0.02 %	1.6 %	0.02 %
32	1.19 %	0.02 %	1.45 %	0.02 %	1.19 %	0.02 %	1.45 %	0.02 %
	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR
	Average results for each filter							

Table 5.2 shows the parameters of the 2D Gabor filter obtaining the average result to be applied in the proposed System 2 scheme for direct key generation. The average result was 1.34% FRR and 0.02 % FAR leading to the chosen parameters, which are the direction values for θ = 0^{o} and the central frequencies (*f*) = 16. Figure 5.9 below shows the FRR and FAR performance and the ROC curve for FRR and FAR rendering.

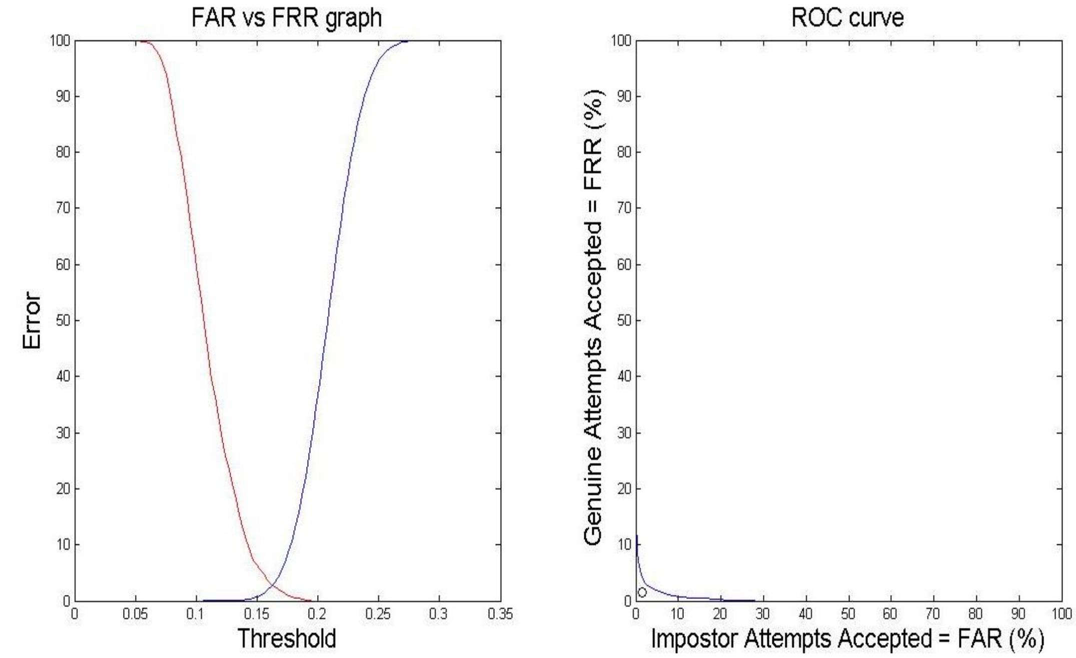


Figure5.9 FRR vs FAR and ROC curve of FAR and FRR performance using 2D Gabor filter

Figure 5.9 shows that the EER for the chosen parameters for 2D Gabor filter is 0.2 %, which compares favourably with the GLCM method's EER of 10%. Both methods were applied in the same database. Table 5.3 highlights the superior performance of the 2D Gabor filter method.

Table 5.3 GLCM vs 2D Gabor filter in proposed scheme with template

	GLCM	2D Gabor
ERR	10%	0.2%

5.3.3. Experimental Results and Discussion

In this section, we have attempted to investigate direct key generation with a free template, using a multimodal biometric system in which we considered iris left and right as two different biometric modalities, with user password as a third modality, to visualise the proposed idea.

5.3.3.1. Experimental Results and Comparison of Systems 1 and 2

In this chapter, two different systems have been used to investigate alternative perspectives on the problem at hand. Simply put, iris modality has been chosen in this investigation of different methods of feature extraction for optimal capture of the subject field. The two chosen variance features were the Gray Level Co-Occurrence Matrix (GLCM) [95] and the 2D Gabor filter algorithm [73], implemented as System1 and System 2, respectively.

The overall performance of both Systems 1 and 2 for all users was assessed in terms of false acceptance and rejection rates by performing several rounds of tests. Table 5.4 shows the average FRR and FAR results of the 5000 tests for all training and testing sets in both systems; as mentioned in section 5.3.2, each user provided 20 different iris samples (11 samples for the training set and 9 samples for the testing set), divided randomly for all 5000 tests. The overall average performance results from these tests are shown in Table 5.4.

Table 5.4 Overall average results: FRR and FAR rates

	GLCM	2D Gabor filter
FRR	13.3999 %	1.7085 %
FAR	9.3333 %	0.0120 %

Table 5.4 shows that the 2D Gabor filter performed well by comparison with the GLCM method (1.7085% FRR versus 13.3999% FRR, respectively). In the 2D Gabor filter method, FAR was also much lower than for GLCM, suggesting that the 2D Gabor filter algorithm is more suitable for iris direct key generation in future work.

5.3.3.2. Comparison and Discussion

The proposed three-factor scheme here can be compared with Daugman et al scheme [50] as discussed in Chapter 2, in which they used a two-factor iris modality and smartcard with freedom of extension to three factors [50]. Both their scheme and ours used a free template and three-factor scheme (biometric, smartcard and password). Their biometric secret key length was 140-bits with 0.47% FRR and 0% FAR. However, with an increase in key length to 196-bits, they achieved 3.65% FRR and 0% FAR. Similarly, when key length was further increased to 224-bits, FRR and FAR were 12.22% and 0%, respectively [50]. This clearly demonstrated a tendency for FRR behaviour to increase with increasing key length values. In our scheme, using the 2D Gabor filter, the length of the biometric secret key was unlimited, with results of 1.7085% FRR and 0.0120% FAR. Additionally, Daugman et al.’s scheme saved the locked biometric secret key by XORing it with the iris code on the smartcard; the decoding process involved XORing the locking key with the presented iris sample [50]. By contrast, the smartcard in both of our schemes holds only the Y coordinates, making it impossible to get any information about the biometric secret key from the smartcard alone and confirming the robustness of Shamir’s secret scheme in enabling use of less information on the smartcard as Y coordinates. In addition, Shamir’s secret scheme is of great help to the user if they fail to set up one of the iris samples, in that they can still access and unlock the key because, as explained above, two points out of three are needed in the authentication phase to generate the biometric secret key, using the linear equation technique. Furthermore, if the user forgets their password, they can still gain access by providing the left and right irises. In this way, all three factors are needed to unlock the biometric secret key using Shamir’s secret scheme [11]—or at least the two factors of smartcard and biometric samples, as the smartcard will provide

Y_n coordinates for three points (left iris, right iris and password). In short, without the smartcard, the user cannot gain access, and one of left or right iris must be provided in addition to the password to allow Shamir's secret scheme to unlock the biometric secret key.

5.3.4. Security Analysis

This chapter has proposed two schemes based on three factors: iris biometric, smartcard and password. In the authentication process, three factors must be provided to release the biometric secret key, which means that two factors or one will not unlock the key, keeping the biometric secret key secure. Furthermore, if the biometric key has been discovered by an attacker, the key can be renewed because it is randomly generated, so blocking the attacker. On the other hand, this scheme makes the lock/unlocking key completely dependent on iris biometrics with support from the password and smartcard, making the iris biometric clearly dominant. For example, the key will not be released unless X_n and Y_n coordinates are provided. X_n coordinates are generated from iris biometrics (left and right) and password. However, Y_n coordinates are created by X_n coordinates and the biometric secret key, which Y_n has downloaded to the smartcard.

From a security perspective, the opportunities for an attacker to hack the system or release the key are as follows.

1. Breaking and obtaining the biometric secret key—this is the most difficult option because, as mentioned above, the key length is unlimited. We tested the system within the key length range of 300–500 bits, which is sufficient.
2. Acquiring the smartcard—this will not enable the attacker to release the key or to access any useful information because, as mentioned above, the novelty of

this system is the minimal information stored in the smartcard. Only Y_n coordinates are stored, and at least two out of three inputs (password with left iris, password with right iris or left and right irises) must be supplied with the smartcard to unlock the key. This confirms the scheme's security, as shown by the FAR of 0 %.

3. Password known by the attacker—again, this will not help to release the key. As mentioned in 2 above, three factors are needed to release the key. In addition, this scheme uses a more difficult form of password creation (up to 10 digits with numbers, capitals and small letters and symbols). Although the password can be captured, the attacker still cannot release the key because they have possession of just one of the coordinates in the x-axis, ensuring a 0% FAR.
4. If both smartcard and password have been captured by the attacker, it is still impossible for them to release the key because, in the linear equation scheme, two points are needed, and the attacker still has only one point from x3 in the password and y3 in the smartcard. The other y1 and y2 coordinates in the smartcard cannot help them to discover the other x1 and x2 coordinates or to unlock the key; the biometric samples (iris left or right) are missing, and it is difficult or impossible to capture the user's iris data without being noticed, again affording a 0% FAR.

5.4. Summary

This chapter has explored the technique of secret sharing and combination of three factors within a free template scheme to allow an encryption key to be created and released from biometric samples with support from a smartcard and password. The results for System 2 confirm the potential for deriving encryption keys efficiently and directly from the three factors (biometric sample, smartcard and password). However, by using the linear equation version of Shamir's secret scheme, this system also enables the key to be released by providing at least two out of three of the required polynomial points (left iris sample, right iris sample and password). Furthermore, the length of the biometric secret key is unlimited, offering the flexibility of multiple keys for different applications for each user, so that an attack on one does not affect the other's applications, given the possibility of revocation. In summary, we believe that the scheme has reached the requisite security level or at the very least encourages further development of the scheme to further reduce the FRR 1.7085%, perhaps involving multimodal biometrics such as face, fingerprint and iris to generate the key directly instead of the iris modality and password investigated here.

Chapter 6.

Conclusion

The biometric systems serve as an identification tools in several organizations, which allow the access to the authorised people in the sensitive areas. The applications of these systems in the security of sensitive areas play an important part in the protection of key information or the material to be protected from the unauthorised access. Similarly, the organizations have provide the authorised access to sensitive strategic data sources which are key to the operations of the business and provide a competent advantage to the companies in the market.

However, these biometric applications are often targeted by the spoofers, attackers and mal-practitioners to attack the security and privacy of these applications, and endanger the privacy and security of the people having their data stored in the system. Furthermore, the issues of key revocability and exception handling cause the biometric system to perform less efficiently and effectively, which are serious drawback of the existing encryption biometric systems In order to avoid these issues, the continuous research efforts are being carried out by the researchers in the field of biometric application using biometric system for generation of key from the biometric modalities directly.

If the biometric key is generated directly from the biometric samples, then there will be less probability of the attack on the system because the hackers will not be able to hack or steal either encryption key or the templates, because of the fact that the encryption key and the template are not required to be stored in the database. Due to these issues associated with biometric systems, the current study has been designed to generate the encryption key directly from multimodal biometric system (left iris, right iris and password).

Consequently, the purpose of the research presented in this thesis was to introduce a scheme for integrating direct biometric key generation schemes with Shamir's secret sharing algorithm [11] in order to address the two disadvantages of revocability and exception handling. In the proposed scheme, an arbitrary subset of biometric modalities needs to be supplied by the user, which will address the situation where a given modality cannot be supplied. Furthermore, an arbitrary encryption key (secret) was generated by the proposed scheme to address the revocability issue and also to allow the multiple independent encryption keys to be derived from the same biometric samples.

6.1 Findings

For the first scenario presented in chapter 3, a technique of secret sharing was applied to allow an encryption key to be created from multimodal biometric samples. The results showed the potential of the proposed system to generate encryption keys efficiently while also allowing exception handling, which was considered a significant impediment to the practical deployment of biometric systems in the previous works. This improved robust property represents a significant enhancement of performance of the proposed system compared with the previous models. A further significant advantage of the proposed technique was that the biometric key itself was not required to be stored in the database which along with the length flexibility of the biometric encryption key further enhanced the security of the system.

Moreover, the proposed system circumvented the main challenge of employing five biometric modalities—face, index finger, thumb finger, right iris and left iris—to generate the points on the required polynomial while meeting the following criteria.

1- The length of the biometric encryption key need to flexible i.e keys of varying lengths should be generated easily depending on the applications' requirements.

2- The number of successful biometric inputs required during the authentication process to release the secret shares should be flexible, depending on the following versions of Shamir's style has been used:

   - Linear equation technique (needs at least two successfully biometric modalities out of five to release the secret shares for biometric encryption key extraction purpose).
   - Quadratic technique (needs at least three out of five).
   - Cubic technique (needs at least four out of five).

3- The main difference between this scheme and other multimodal approaches is that in other existing multimodal systems, if any of the biometric modalities failed to process or to provide the required biometrics, the system will reject the user. However, the proposed scheme presented in Chapter 3 solved this problem.

In this way, the current research work provides the new avenues and prospects to develop more robust and highly secured and performance oriented multimodal biometric system based on the application of Shamir's Secret Sharing Scheme. This scheme was uniquely applied to give the encryption key a flexible length which was not provided by the previous researches conducted in biometric encryption field. However, the limitation of the proposed was that it retained the template within the database like other related system.

On the other hand, second scenario which was created to focus on direct biometric encryption key generation from multimodal biometrics used in this: left iris, right iris, and user password. The proposed technique utilized these three biometric modalities in different combinations to allow an encryption keys to be generated and released from the biometrics samples with support of smartcard and password. In this way, this technique has the potential and prospect to implement the template-free biometric system. The results for *System 2* in section *5.3.3.1* demonstrated the potential of the three-factor scheme (biometric sample, smartcard and password) for direct generation of key from three biometric modalities. However, by using the linear equation version of Shamir's secret scheme, this proposed system also enabled the key to be released from the input biometric samples by providing at least two out of three of the required polynomial points (left and right iris samples and password). Furthermore, the length of the biometric secret key made flexible in terms of increasing the length of key up to securable length, offering the probability of multiple keys for different applications for each user, so that an attack on one application could not affect the others. In summary, the scheme is expected to have reached the requisite security level or at least encourages further development of the scheme to reduce the FRR up to 1.7085%, perhaps involving multimodal biometrics such as face, fingerprint and iris to generate the key directly instead of the iris modality and password investigated here.

However, the researcher tested the potential of several schemes on hit and trial basis whether they can be employed to generate the biometric keys directly from the biometric scheme, but they were proved to be ineffective in this context. The results of these trials have been presented in chapter 4. However, they gave the researcher some important lessons and knowledge about the development of multimodality biometric system to be utilized for the next stages of the research, which is presented through following points:

- In comparing any two iris samples, whether from the same person or not, there is a huge similarity between them. As evidence of this, any iris recognition system defines a threshold such that a Hamming Distance < 0.2 for example (which is the average threshold in public) identifies the iris provider as genuine or otherwise as an impostor. This indicates that the similarity between any two irises is more than 70%.
- Iris binary code is a vector that depends on pre-processing and feature extraction, making it hard to name a block or digit in the vector that is standard for a particular person, appearing every time as 0 or 1.

Finally, the design of two template-free biometric systems researcher has been selected, in which biometric data can be mapped onto repeatable unique binary codes/strings in both, opening up the key only in the presence of biometric prints. The first system used the Gray Level Co-Occurrence Matrix method (GLCM) [94, 95] to extract iris features, and the second system used the 2D Gabor Filter [73]. In order to generate the secret key, Shamir's Secret Scheme [11] was applied to the three data points provided (right iris, left iris, and password). Two points out of three were found to be sufficient to release the biometric key by use of the linear equation technique, which not only gave a user the desired flexibility but also reduced the FRR. For both designs, a three-factor scheme was proposed, to include smartcard, password, and biometrics. Each component within the proposed system was required equally to reveal the biometric data specific to an individual, with the opportunity of either updating or revoking the key. The peculiar feature of the proposed system is that it cannot be compromised until all three factors are provided to it. These two designs are presented in Chapter 5.

6.2 Future Work

The main aim of this research was investigate multimodal biometric system involving three different biometric modalities (Iris, Fingerprint and Face) in terms of generating the biometric encryption key directly from input biometrics samples. Furthermore, this research work considered the left and right iris as two different modalities while the password as the third modality to build the multimodal system capable to generate the biometric encryption key directly by using Shamir's secret scheme algorithm. The further work can be undertaken on a small scale in the same direction on iris or the technique can be widened to include other modalities.

This thesis, especially in chapter 5 presented a clear and robust that will provide further avenues for the future researchers to develop this area, which is given below:

- Achievement of a reasonable level of performance using Shamir's secret scheme which allows the generation of key directly from 2 out of 3/4 modalities in linear equation technique. However, in quadratic 3 modalities are needed to generate the key out of 4 or 5 modalities and so on.
- Attaining a sensible level of security by implementing the scheme with template free and with three factors scheme provided to the system (biometric, smartcard, password).
- The length of the biometric encryption key is set flexible which means that key with substantial secured length can be generated; and that is the main novelty of this thesis. The researchers can apply this technique for generation of flexible keys for other biometric modalities.

- Minimisation of level of data stored in the smartcard, for example, the data in the smartcard is just the Y_n coordinates which are not useful for the attacker to unlock the system because the biometric encryption key cannot be released using only Y_n coordinates themselves. This method may have the potential application in the smart cards for individual identification in military, banks and other organizations which require the access to sensitive areas using the smart card based individual recognition.
- The direct key generation from the biometric samples has been tested in this work, which shows the potential of the system to be applied in the biometric application applications in military and airports which require the greater of level of security and privacy. As the proposed method does not require the storage of template and the encryption keys in the database of biometric systems.

Therefore, this research work can be extended to study multimodal system in details to generate the biometric encryption key directly from different combinations of biometrics including face, fingerprint, iris, and voice using Shamir’s secret to release the key with three different styles of flexibilities (Linear equation, Quadratic, and Cubic) as it shows in Figure 6.1 below while in this thesis, iris modality has been used with support from user password.

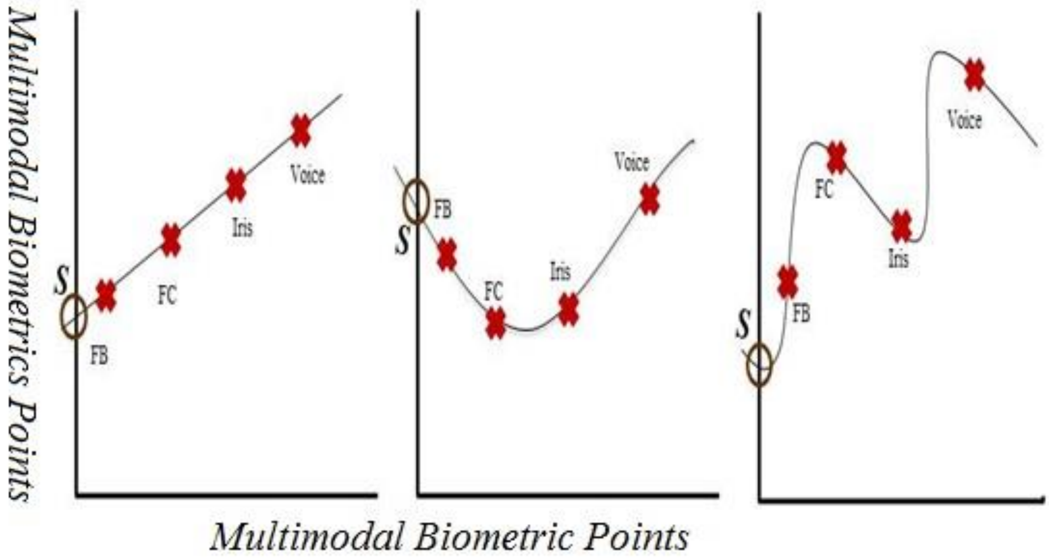


Figure6.1 Direct Key Generation from Multimodal Biometric System by Shamir's Secret Scheme

On the other hand, the new researcher might look at the maps distribution on iris modality in terms of solving the iris feature overlaps in feature space which might be helpful for reducing the number of important data stored in the maps distribution related to the users templates. Moreover, the future work can focus on developing maps distribution which might solve the potential overlaps of binary codes of iris samples obtained from different people, which is considered to be the main obstacle in designing template-free iris recognition biometric system. The researcher showed in the experiments that the similarities between two iris codes from different people may exceed 70% which is considered too high to generate the key directly without having any data stored that not related to user’s templates.

The future research work may also consider different iris feature combinations for extracting the suitable iris feature for key generation with acceptable FRR and key entropy. For example, instead of using 2D Gabor filter alone or GLCM, the combination between these two will be a great idea in template free system which may lead to selection of a feature capable to generate the biometric key from it because this thesis faced some challenges in generating such as the high percentage of similarities between users’ iris binary codes.

References

[1] K. Jain, A. Ross, and K. Nandakumar, "Introduction to Biometrics", *Springer*, 2011 (ISBN 978-0-387-77325-4)

[2] S. Prabhakar, S. Pankanti, and A. K. Jain, “Biometric Recognition: Security and Privacy Concerns”, *IEEE Security and Privacy Magazine*, Vol. 1, No. 2, pp. 33-42, 2003.

[3] A. K. Jain, R. Bolle, and S. Pankanti (editors), ‘Biometrics: Personal Identification in Networked Society’, *Kluwer Academic Publishers*, 1999.

[4] A. Azzini , S. Marrara , R. Sassi , F. Scotti, “ A fuzzy approach to multimodal biometric continuous authentication”, *Published online*: 28 June 2008

[5] Braz, C., & Robert, J. “ Security and usability: The case of the user authentication methods”. *In Proceedings of IHM 2006*, Montreal, QC, April 18–21, 2006.

[6] Krawczyk, S., & Jain, A. K. “ Securing electronic medical records using biometric authentication”. *In Proceedings of AVBPA* 2005, Springer, pp.1110–1119.

[7] U. Uludag, S. Pankanti, S. Prabhakar, and A.K. Jain, “Biometric cryptosystems: Issues and Challenges,” *Proc. IEEE*, vol. 92, no. 6, pp. 948-960, 2004.

[8] D.G. Abraham, G.M. Dolan, G.P. Double, and J.V. Stevens, “Transaction Security System,” *IBM Systems* J., vol. 30, no. 2, pp. 206-229, 1991.

[9] J. Pedraza, M. A. Patricio, A. de As´ıs, and J.M.Molina, “Privacy and legal requirements for developing biometric identification software in context-

based applications," *International Journal of Bio-Science and Bio-Technology*, vol. 2, no. 1, pp. 13–24, 2010.

[10] D. Maltoni, et al., Handbook of Fingerprint Recognition: Springer Verlag, 2003.

[11] A. Shamir, "How to Share a Secret," *Communications of the ACM*, vol. 22, pp. 612-613, November 1979.

[12] Liu, C. L. " Introduction to Combinatorial Mathematics", New York: McGraw-Hill, 1968 .

[13] Dawson, E., Donovan, D., "The breadth of Shamir's secret-sharing scheme", *Computers & Security* 13: 69–78, doi:10.1016/0167-4048(94)90097-3 .,1994.

[14] Knuth, D. E. , "The Art of Computer Programming, II: Seminumerical Algorithms", (3rd ed.), Addison-Wesley, p. 505 ,1997.

[15] M. Faundez-Zanuy, "On the vulnerability of biometric security systems," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 19, pp. 3-8, 2004.

[16] C. Soutar, et al., "Biometric Encryption," ICSA Guide to Cryptography, 1999.

[17] A. Juels and M. Sudan, "A Fuzzy Vault Scheme," Design, Codes and Cryptography, vol. 38, pp. 237 - 257, 2005.

[18] Y. Shenglin and I. Verbauwhede, "Automatic secure fingerprint verification system based on fuzzy vault scheme," 2005, pp. v/609-v/612 Vol. 5.

[19] U. Uludag, et al., "Fuzzy Vault for Fingerprints," *Lecture Notes in Computer Science*, vol. 3546, pp. 310 - 319, 2005.

[20] F. Monrose, et al., "Cryptographic Key Generation From Voice," *IEEE Symposium on Security and Privacy*, May 2001 2001.

[21] F. Monrose, et al., "Password Hardening Based on Keystroke Dynamics," *ACM Conference on Computer and Communications Security,* 1999.

[22] T. C. Clancy, et al., "Secure Smartcard-Based Fingerprint Authentication," *Proc. ACM SIGMM* 2003 Multim., Biom. Met. & App, pp. 45 - 52, 2003. 2 41

[23] Y. Chang, et al., "Biometric-based Cryptographic Key Generation," *Proc. IEEE Conf.* Multimedia and Expo, pp. 2203 - 2206, 2004.

[24] X. Li, "Modeling Intra-Class Variation for Non-Ideal Iris Recognition," *Proc. Int'l Conf. Biometrics,* pp. 419-427, 2006.

[25] J. Daugman, "Biometric personal identification system based on iris analysis", United States Patent, Patent Number: 5,291,560, 2010.

[26] L. Ma, T. Tan, and Y. W. D. Zhang. ''Personal identi.cation based on iris texture analysis''. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(12):1519 - 1533, 2003.

[27] L. Ma, T. Tan, and Y. W. D. Zhang. E.cient ''iris recognition by characterizing key local variations''. *IEEE Trans. on Image Processing*, 13(6):739 – 750, 2004.

[28] R. Wildes. Iris recognition: an emerging technology. *Proc. of IEEE,* 85:1348– 1363, 1997.

[29] Bolle, R. M., Connell, J. H., Pankanti, S., Ratha, N. K., & Senior, A. W. *"*Guide to biometrics". *New York: Springer-Verlag*. 2004

[30] A. K. Jain, A. Ross, S. Prabhakar, "An Introduction to Biometric Recognition", *IEEE Trans. on Circuits and Systems for Video Technology*,Vol. 14, No. 1, pp 4-19, January 2004

[31] J. L. Wayman, “Fundamentals of Biometric Authentication Technologies”, *International Journal of Image and Graphics*, Vol. 1, No. 1, pp. 93-113, 2001.

[32] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, “FVC2002: Fingerprint Verification Competition”, *Proc. International Conference on Pattern Recognition (ICPR)*, pp. 744-747, Quebec City, Canada, August 2002.

[33] United Kingdom Biometric Work Group (UKBWG), “Best Practices in Testing and Reporting Performance of Biometric Devices, Version 2.01”, August 2002, Available from http://www.cesg.gov.uk/technology/biometrics/.

[34] R. Cappelli, D. Maio and D. Maltoni, “Indexing Fingerprint Databases for Efficient 1:N Matching”, *Proc. International Conference on Control Automation Robotics and Vision (6th)*, 2000.

[35] Sharat Chikkerur, et al., "A Systematic Approach For Feature Extraction In Fingerprint Images," *Lecture Notes in Computer Science*, vol. 3072/2004, pp. 344-350, 2004.

[36] P.J. Philips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone, “FRVT 2002: Overview and Summary”, available from http://www.frvt.org/FRVT2002/documents.htm

[37] M. Golfarelli, D. Maio and D. Maltoni, “On The Error-Reject tradeoff in Biometric Verification Systems”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp. 786-796, July 1997.

[38] Bo, F.,S. X. Yang, et al.(2009). "Multibiometric Cryptosystem: Model Structure and Performance Analysis.Information Forensics and Security*”, IEEE Transactions on 4(4*): 867-882.

[39] Fakhreddine Karray, Jamil Abou Saleh, Mo Nours Arab and Milad Alemzadeh, “Multi Modal Biometric Systems: A State of the Art Survey”, *Pattern Analysis and Machine Intelligence Laboratory,* University of Waterloo, Waterloo, Canada.

[40] M. Faundez-Zanuy, “2004, Data fusion in biometrics”, *IEEE Aerospace and Electronic Systems Magazine,* 20(1), pp 34-38 (2005).

[41] M. Indovina, U. Uludag, R. Snelick, A. Mink and A. Jain, “Multimodal biometric authentication methods: a COTS approach”, *Proceeding of the MMUA*, Workshop on Multimodal User Authentication, Santa Barbara, California, pp 99-106 (2003).

[42] S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, “Fusion of face and speech data for person identity verification”, *IEEE Transactions on Neural Networks*, 10(5), pp 1065-1074 (1999).

[43] Uludag, U., S. Pankanti, et al. "Biometric cryptosystems: issues and challenges." *Proceedings of the IEEE 92(6)*: 948-960, 2004.

[44] Zheng, G., W. Q. Li, et al. “Cryptographic key generation from biometric data using lattice mapping”. *18th International Conference on Pattern Recognition*, Vol 4, Soc:513-516. (2006).

[45] A. Jules and M. Wattenberg. “A fuzzy commitment scheme*”. In Proc. ACM Conf. Computer and Communication Security,pages* 28–36, 1999.

[46] S. A. Vanstone and P. C. van Oorshot. “An introduction to error correcting codes with applications*”. Kluwer AcademicPublishers*, 1989.

[47] C.Soutar, D. Roberge, S. A. Stojanov, R. Gilroy, and B. V. K. Vijaya Kumar, “Biometric encryption using image processing,” *in Proc.SPIE, Optical Security and Counterfeit Deterrence Techniques* II, vol. 3314, pp. 178–188, 1998.

[48] A. B. J. Teoh, D. C. L. Ngo, and A. Goh. “Personalised cryptographic key generation based on FaceHashing*”. Computer& Security,* 23:606–614, 2004.

[49] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” *in Proc. Eurocrypt*, pp. 523–540, 2004.

[50] Feng Hao, Ross Anderson, John Daugman, "Combining Crypto with Biometrics Effectively", *IEEE Transactions on Computers*, vol.55, no. 9, pp. 1081-1088, September 2006, doi:10.1109/TC.2006.138.

[51] R.J. Anderson , "Security Engineering: A Guide to Building Dependable Distributed Systems", ISBN 0-471-38922-6, Wiley, 2001.

[52] T.C. Clancy, N. Kiyavash, and D.J. Lin, “Secure Smart Card-Based Fingerprint Authentication,” *Proc: ACM SIGMM Workshop Biometrics Methods and Application (WBMA)*, 2003.

[53] A. Goh and D.C. L. Ngo, “Computation of Cryptographic Keys from Face Biometrics,” *Proc. Int’l Federation for Information Processing*, pp. 1-13, 2003.

[54] F. Hao and C.W. Chan, “Private Key Generation from On-Line Handwritten Signatures,” *Information Management & Computer Security,* vol. 10, no. 2, pp. 159-164, 2002.

[55] K.J. Pawan and M.Y. Siyal, “Novel Biometric Digital Signature for Internet Based Applications,” *Information Management and Computer Security,* vol. 9, no. 5, pp. 205-212, 2001.

[56] Y. Seto, “Development of Personal Authentication Systems Using Fingerprint with Smart Cards and Digital Signature Technologies,” *Proc. Seventh Int’l Conf. Control,* Automation, Robotics, and Vision, Dec. 2002.

[57] G.I. Davida, Y. Frankel, B.J. Matt, and R. Peralta, “On the Relation of Error Correction and Cryptography to an Off Line Biometrics Based Identification Scheme,” *Proc. Workshop Coding and Cryptography,* 1999.

[58] S.S. Agaian, Hadamard Matrix and Their Applications. Springer Verlag, 1985.

[59] D. Wheeler, “Protocols Using Keys from Faulty Data,” *Proc. Security Protocols Workshop,* 2001.

[60] X. Boyen, “Reusable Cryptographic Fuzzy Extractors,” *Proc. CCS,* pp. 82-91, 2004.

[61] Aldosary. S, Howells. G, ‘’A robust multimodal biometric security system using the polynomial curve technique within Shamir’s Secret Sharing algorithm’’, *Emerging Security Technologies (EST), Third International Conference on Lisbon, IEEE,* doi: 10.1109/EST.2012.9, 2012.

[62] M. Bellare and P. Rogaway. ‘’Robust computational secret sharing and a unified account of classical secret-sharing goals’’. *Proceedings version of*

this paper. Proc. of the 14th ACM Conference on Computer and Communications Security (ACM CCS), ACM Press, 2007.

[63] A. Singh, S. Kumar, ‘’ Face Recognition Using PCA and Eigen Face Approach’’, *thesis, Bachelor of Technology, Department of Computer Science and Engineering, National Institute of Technology Rourkela,* Rourkela-769008, India, 2012.

[64] W. Zhao, R. Chellappa, P. Phillips, A. Rosenfeld, “Face recognition: a literature survey”, *ACM Computing Surveys* vol. 35, pp.399–458, December 2003.

[65] Daugman J. ,”Face and gesture recognition: Overview”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 675-676,1997.

[66] J. Li, B. Zhao, and H. Zhang, "Face recognition based on PCA and LDA combination feature extraction," *1st IEEE International Conference on Information Science and Engineering*, pp. 1240-1243, 2009.

[67] Elham Bagherian and Rahmita Wirza O.K. Rahmat, ‘’Facial Feature Extraction for Face Recognition: A Review,’’ *IEEE International Symposium on Information Technology*, Vol. 2, Aug 26-28, pp.1-9, 2008.

[68] A. Bansal, K. Mehta and S. Arora,” Face Recognition Using PCA & LDA Algorithms”, *Second International Conference on Advanced Computing & Communication Technologies IEEE,* 978-0-7695-4640- 7/12, 2012.

[69] M. A. Turk and A. P. Pentland, “Face recognition using eigenfaces,” *in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Maui, Hawaii, pp.586-591, 1991.

[70] M. Turk and A. Pentland, “Eigenfaces for Recognition,” *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[71] J.G. Daugman, “High Confidence Visual Recognition of Persons by a Test of Statistical Independence,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, 1993, pp. 1148-1161.

[72] Jong-Gook Ko, Youn-Hee Gil, Jang-Hee Yoo, and Kyo IL Chung, “A Novel and Efficient Feature Extraction Method for Iris Recognition” *ETRI Journal*, Volume 29, Number 3, June 2007.

[73] J. Daugman, “The Importance of Being Random: Statistical Principles of Iris Recognition,” *Pattern Recognition*, vol. 36, no. 2, pp. 279-291, 2003.

[74] J. Daugman. ‘’How iris recognition works’’. *Proceedings of 2002 International Conference on Image Processing*, Vol. 1, pp 16-22 2002.

[75] M. T. John, "Region-of-interest detection for fingerprint classification," 1994, pp. 48-59.

[76] R. Krishnan, "Fingerprint Capture Challenges and Opportunities," *Presentation for US Department for Homeland Security*, 2006.

[77] N. K. Ratha and R. M. Bolle, "Effect of Controlled Acquisition on Fingerprint Matching," *Proceedings of the International Conference on Pattern Recognition*, vol. 2, pp. 1659 - 1661, 1998.

[78] A. K. Jain, et al., "Filterbank-based fingerprint matching," *Image Processing, IEEE Transactions* on, vol. 9, pp. 846-859, 2000.

[79] L. G. Shapiro and G. C. Stockman, ‘’Computer Vision’’: *Prentice Hall*, 2001.

[80] E.-K. Yun and S.-B. Cho, "Adaptive fingerprint image enhancement with fingerprint image quality analysis," *Image and* Vision *Computing*, vol. 24, pp. 101-110, 2006.

[81] X. Wang, et al., "Adpative Fingerprint Enhancement by Combination of Quality Factor and Quantitative Filters," *Lecture Notes in Computer Science*, vol. 3781, pp. 111 - 118, 2005.

[82] J. Xudong and Y. Wei-Yun, "Fingerprint minutiae matching based on the local and global structures," 2000, pp. 1038-1041 vol.2.

[83] J.Fei Lim, R, Chin., " Enhancing Fingerprint Recognition Using Minutiae-Based and Image-Based Matching Techniques ," *AIMS Conference, IEEE Proceedings*-, doi: 10.1109/AIMS.2013.48, pp. 261-266, 2013.

[84] ISO/IEC and ANSI, "Biometric Data Interchange Formats — Part 2: Finger Minutiae Data," ISO/IEC JTC 1/SC 37, 04-05-2004 2004.

[85] D. Maio and D. Maltoni, "Direct gray-scale minutiae detection in fingerprints," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, pp. 27-40, 1997.

[86] G. Parziale and A. Niel, ‘’A Fingerprint Matching Using Minutiae Triangulation’’, 2004.

[87] C. J. Lee and T. N. Yang, "Direct Minutiae Detection in Raw Fingerprint Images," Signal and Image Processing, 2004.

[88] J.-H. Shin, et al., ‘’Minutiae Extraction from Fingerprint Images Using Run-Length Code’’, 2003.

[89] N. Yager and A. Amin, "Fingerprint verification based on minutiae features: a review," *Pattern Analysis & Applications,* vol. 7, pp. 94-113, 2004.

[90] A. Jain, et al., "Fingerprint matching using minutiae and texture features," *in Image Processing,* Proceedings. 2001 International Conference on, 2001, pp. 282-285 vol.3.

[91] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman and A.K. Jain, “FVC2000: Third Fingerprint Verification Competition”, *IEEE Transactions on Pattern Analysis Machine Intelligence*, 2004.

[92] CASIA Iris Image Database V-3.0, http://www.cbsr.ia.ac.cn/irisdatabase. htm, 2008.

[93] A. P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 22, No. 10, October 2000.

[94] Y. Chen, F. Y. Yang, and H. L. Chen, “An effective iris recognition system based on combined feature extraction and enhanced support vector machine classifier,” *Journal of Information and Computational Science,* vol. 10, no. 17, 2013.

[95] R. M. Haralick, K. Shanmugam, I. H. Dinstein, Textural features for image classification, *IEEE Systems, Man, and Cybernetics Society,* 3 (1973), 610-621

[96] W. W. Boles and B. Boashash, “A human identification technique using images of the iris and wavelet transform,” *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 1185–1188, 1998.

[97] J. Daugman (2003), “Demodulation by Complex valued wavelets for stochastic pattern recognition”, *International Journal of Wavelets,* Multiresolution and Information Processing, 1(1), 1-17

[98] Rumsey, D., ‘’Statistics for Dummies’’. *Wiley publishing Inc*., Indiana, 2003

[99] Nandakumar, K.: ‘Integration of multiple cues in biometric systems’, *PhD thesis, Michigan State University* (2005)

[100] E. Papoutsis, et al., "Integrating Multimodal Circuit Features within an Efficient Encryption System," *Journal of Information Assurance and Security*, vol. 2, pp. 117 - 126, 2007.

[101] E. Papoutsis, et al., "Key Generation for Secure Inter-Satellite Communication," *NASA / ESA Conference on Adaptive Hardware and Systems* (AHS - 2007), 2007.

[102] E. Papoutsis, et al., "Effects of Feature Trimming on Encryption Key Stability for an ICmetric System," *ECSIS Symposium on Bio-Inspired,* Learning and Intelligent Systems for Security, 2008.

[103] Joshua. A, Howells. G, ‘’ Key Generation in a Voice Based Template Free Biometric Security System’’, *LNCS 5707,* pp. 170-177, 2009.

[104] Black, Paul E. ‘Gray code’, *NIST*, 25 February 2004. .

[105] Sedgewick, Robert . "14. Hashing". Algorithms in Java’ *(3 ed.). Addison Wesley.* ISBN 978-0201361209, 2002.

[106] Atah, J.A., Howells, G.: ‘Score Normalisation of Voice Features for Template Free Biometric Encryption’. *In: The multi-conference in computer science, information technology, computer engineering,* control and automation technology, Orlando, FL, USA (July 2008)

[107] Poh, N., Bengio, S.: 'A study of the effect of score normalization prior to fusion in Biometric Authentication Tasks', (December 2004)

[108] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," *IEEE Trans. Image Process*., vol. 13, pp. 739-750, 2004.

[109] P. Yao, J. Li, X. Ye, Z. Zhuang and B. Li, "Iris Recognition using modified Log-Gabor filters," *Proc. ICPR 2006*, Hong Kong, pp. 1-4, Aug. 2006.

[110] L. Masek, 'Recognition of human iris pattern for biometric identification', *B. Eng. Thesis, University of Western Australia*, 2003.

[111] L. Shinyoung, K. Lee, O. Byeon. "Efficient Iris Recognition Through Improvement Of Feature Vector And Classifier", *ETRI Journal*. vol. 23, pp. 61-70, June 2001.

[112] H. Proença and L. A. Alexandre, "Towards noncooperative iris recognition: A classification approach using multiple signatures," *IEEE Trans. Patt*. Anal. Machine Intell., vol. 29, pp. 607-612, Apr. 2007.

[113] L. Ma, T. Tan, Y. Wang and D. Zhang, "Local intensity variation analysis for iris recognition," *Pattern Recognition*, vol. 37, no. 6, pp. 1287-1298, Jun. 2004.

[114] ICAO. Annex 1 - Use of Contactless Integrated Circuits. Technical report, May 2004. Available at http://www.icao.int/mrtd/download/documents/Annexs.pdf.

[115] ICAO. Biometrics deployment for Machine Readable Travel Documents. Technical report, May 2004. Available at http://www.icao.int/mrtd/download/documents.

[116] ICAO. PKI for Machine Readable Travel Documents offering ICC read-only access. Technical report, Oct. 2004. Available at http://www.icao.int/mrtd/download/documents/TR-PKIy

[117] Serge Vaudenay. ‘A Classical Introduction to Cryptography: Applications for Communications Security’. *Springer,* 2005.

[118] Serge Vaudenay. ‘Secure communications over insecure channels based on short authenticated strings’. *In Victor Shoup*, editor, *CRYPTO*, volume 3621 of *Lecture Notes in Computer Science*, pages 309–326. Springer, 2005.

[119] Claude Barral, ‘’ Biometric & Security: Combining Fingerprint, Smart Cards and Cryptography’’, *Ph.D. thesis, Ecole Polytechnique Federale de Lausannel*, June 2010. Cited on page(s) 21.

[120] J. Daugman, “Biometric Decision Landscapes,” *Technical Report UCAM-CL-TR-482,* Computer Laboratory, Univ. of Cambridge, 2000.

[121] F.J. MacWilliams and N.J.A. Sloane, ‘The Theory of Error-Correcting Codes’. *North Holland,* 1991.

[122] R.J. McEliece, ‘The Theory of Information and Coding’. *Cambridge Univ*. Press, 2002.

[123] S. G. Mallat, "A Wavelet Tour of Signal Processing," *Academic Press,* p. 577, 1998.

[124] A. Adler, "Biometric System Security," in *Handbook of Biometrics*, ed: Springer, 2007.

[125] A. K. Jain, *et al.*, "Biometric Template Security," *EURASIP Journal on Advances in Signal Processing (Special Issue on Biometrics),* 2008.

[126] N. Ratha, *et al.*, "Privacy Enhancements for Inexact Biometric Templates," in *Security with Noisy Data*, ed: Springer London, 2007.

[127] A. T. B. Jin and J. Kim, "FuzzyHash: A Secure Biometric Template Protection Technique," *2007 Frontiers in the Convergence of Bioscience and Information Technologies,* 2007.

[128] Y. Sutcu, *et al.*, "A Secure Biometric Authentication Scheme Based on Robust Hashing," *ACM Multimedia and Security Workshop*, 2005.

[129] S. Yang and I. M. Verbauwhede, "Secure fuzzy vault based fingerprint verification system," 2004, pp. 577-581 Vol.1.

[130] B. Chen and V. Chandran, "Biometric Based Cryptographic Key Generation from Faces," *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications,* pp. 394 - 401, 2007.

[131] W. J. Scheirer and T. E. Boult, "Cracking Fuzzy Vaults and Biometric Encryption," *Biometrics Symposium 2007,* pp. 1 - 6, Sept 2007.

[132] A. Adler, "Vulnerabilities in Biometric Encryption System," *Audio- and Video-based Biometric Person Authentication,* vol. LNCS3546, 2005.

[133] A. Bodo, "Method for Producing a Digital Signature with Aid of a Biometric Feature," *German Patent DE 4243908A1,* 1994.

[134] P. K. Janbandhu and M. Y. Siyal, "Novel Biometric Digital Signatures for Internet-based Applications," *Information Management and Computer Security,* vol. 9, pp. 205 - 212, 2001.

[135] Y. Yamazaki and N. Komatsu, "A Secure Communication System using Biometric Identity Verification," *IEICE Transactions on Informations and Systems,* vol. E84-D(7), pp. 879 - 884, 2001.

[136] S. Hoque, *et al.*, "Feasability of Generating Biometric Encryption Keys," *Electronics Letters,* vol. 41, 17 March 2005 2005.

[137] W. Sheng, *et al.*, "Template-Free Biometric Key Generation by Means of Fuzzy Genetic Clustering," *IEEE Transactions in Information Forensics and Security,* vol. 3, pp. 183 - 191, 2008.

[138] N. K. Ratha, *et al.*, "Cancelable Biometrics: A Case Study in Fingerprints," *Proceedings of 18th International Conference on Pattern Recognition,* vol. 4, pp. 370 - 373, 2006.

[139] W. C. Ku, *et al.*, "Further cryptanalysis of fingerprint-based remote user authentication scheme using smartcards," *Electronics Letters,* vol. 41, pp. 240-241, 2005.

[140] E.-J. Yoon and K.-Y. Yoo, *Secure Fingerprint-Based Remote User Authentication Scheme Using Smartcards*, 2005.

[141] Y. Wu, Dong (Kent Ridge Digital Labs), "Method of Using Biometric Information for Secret Generation' *(Patent No: WO 02/078249 A1),"* 2002.

[142] W. G. J. Howells, *et al.*, "A Method and Apparatus for the Generation of Code from Pattern Features' *(Patent No: WO/2008/003945),*" 2006.

[143] F. Gray, "Pulse Code Communication' *(Patent No: 2,632,058),*" 1953.

[144] U. Korte and R. Plaga, "Cryptographic Protection of Biometric Templates: Chance, Challenges and Applications," *Lecture Notes in Informatics,* vol. 108, 2007.

[145] J. Daugman, 'Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons', *Proc. of IEEE, 94 (2006),* 1927-1935

[146] J. Daugman, ‘New methods in iris recognition’, *IEEE Trans. Syst.* Man. Cyb., Part B, 37 (2000), 1167-1175

[147] D. M. Monro, S. Rakshit, D. Zhang, ‘DCT-based iris recognition’, *IEEE Trans. Pattern Anal.*, 29 (2007), 586-595

[148] L. Ma, T. N. Tan, Y. H. Wang, ‘Personal identification based on iris texture analysis’, *IEEE Trans. Pattern Anal.*, 25 (2003), 1519-1533

[149] T. N. Tan, Z. F. He, Z. Z. Sun, ‘Efficient and robust segmentation of noisy iris images for noncooperative iris recognition’, *Image Vision Comput.*, 28 (2010), 223-230

[150] R. Wildes, Iris recognition: ‘An emerging biometric technology’, *Proc. of IEEE,* 85 (1997), 1348-1363

[151] X. F. He, S. J. An, P. F. Shi, ‘Statistical texture analysis-based approach for fake iris detection using support vector machines’, *Lect. Notes Comput.* Sci., 4642 (2007), 540-546

[152] W. S. Chen, R. H. Huang, L. Hsieh, ‘Iris recognition using 3D co-occurrence matrix’, *Lect. Notes Comput.* Sci., 5558 (2009), 1122-1131

[153] W. Yang, L. Yu, G. Lu, ‘Iris recognition based on location of key points’, *Lect. Notes Comput. Sci.* 3072 (2004), 484-490

[154] R. H. Abiyev, K. Altunkaya, ‘Neural network based biometric personal identification with fast iris segmentation’, *Int. J. Control Autom.,* 7 (2009), 17-23

[155] F. N. Sibai, H. I. Hosani, R. M. Naqbi, ‘Iris recognition using artificial neural networks’, *Expert Syst. Appl.,* 38 (2011), 5940-5946

[156] R. Wanderhoof, "Smart cards, biometrics and privacy," Card Technology Today, 2003.

[157] R.A. Fisher, “The Use of Multiple Measures in Taxonomic Problems,” Ann. Eugenics, vol. 7, pp. 179-188, 1936

[158] T.J. Stonham, “Practical face recognition and verification with WISARD” Aspects of Face Processing, pp. 426- 441, 1984.

[159] F. R. Renzetti, L. Zortea, “ Use of a gray level co-occurrence matrix to characterize duplex stainless steel phases microstructure”, Frattura ed Integrità Strutturale, 16 (2011) 43-51; DOI: 10.3221/IGF-ESIS.16.05.

[160] V. Zorkadis P. Donos, "On biometrics-based authentication and identification from a privacy protection perspective", Information Management & Computer Security, 2004 Vol. 12 Iss 1 pp. 125 – 137.

[161] Biometric Working Group , Best Practices in Testing and Reporting Biometric Device Performance, technical report, 2002, available at: www.cesg.gov.uk/technology/biometrics.

[162] Clabrese, C. , “The trouble with biometrics”, J.1999, Login, Vol. 24 No. 4, pp. 56-61.

[163] CNIL, 22nd Annual Report for 2001, Commission Nationale de l’Informatique et des Liberte´s,2001, available at: www.cnil.fr/

[164] European Parliament and Council ,Directive 95/46/EC on the Protection of Individuals With Regard to the Processing of Personal Data and on the Movement of Such Data, 1995, available at: www.eu.int/comm/. . ./media/dataprot/.

[165] European Parliament and Council ,Directive 97/66/EC on the Protection of Personal Data in the Telecommunications Sector, 1997, available at: www.eu.int/comm/. . ./media/dataprot/

[166] European Parliament and Council ,Directive 2002/58/EC on the Protection of Personal Data in Electronic Communications, 2002, available at: www.eu.int/comm/. . ./media/dataprot/

[167] Frankel, Y. , “Biometric identification and authentication with privacy preservation”, Proceedings of the RSA Conference, 2000.

[168] Jain, A., Bolle, R. and Pankanti, S. , Biometrics: Personal Identification in Networked Society, Kluwer Academic Publishers, Dordrecht, 1999.

[169] Prabhakar, S., Pankanti, S. and Jain, A.K., “Biometric recognition: security and privacy concerns”, IEEE Security & Privacy, 2003, pp. 33-42.

[170] Matsumoto, T. et al. , “Impact of artificial ‘gummy’ fingers on fingerprint systems”, Proceedings of the Optical Security and Counterfeit Deterrence Techniques, 2002, available at:http://cryptome.org/gummy.htm

[171] Vaclav, M. and Zdenek, R. , Biometric Authentication Systems, technical report, ecom-monitor.com, 2000, available at: www.ecom-monitor.com/papers/biometricsTR2000.pdf

[172] Stallings, W.: Cryptography and Network Security: Principles and Practice, 5e, Prentice Hall, (2010).

[173] ”Advance Encryption Standard (AES)”, Federal Information Processing Standards Publication 197 United States National Institute of Standards and Technology (NIST) November 26, (2001).

[174] Nandakumar, K., Jain, A., and Pankanti, S.: Fingerprint-based fuzzy vault: Implementation and performance. IEEE Trans.Inf. Forensics Security, 2(4), 744-757, (2007)

[175] Rathgeb, C., and Andreas U.: Context-based biometric key generation for Iris. IET computer vision, 5(6), pp. 389-397. (2011).

[176] AT&T Laboratories Cambridge. The ORL face database, Olivetti Research Laboratory available at http://www.uk.research.att.com/pub/data/att_faces.zip

[177] Fingerprint recognition paper by WUZHILI (Department of computer science and engineering, Hong-Kong Baptist University, 2002).