Expression hash: A new approach for cancelable biometrics

By
Ayesha Riaz
Fall 2015-MS(IS)-8 00000117640

Supervisor
Dr. Naveed Riaz
Department of Computing

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Information Security MS(IS)

In
School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.

(January 2019)

Approval

It is certified that the contents and form of the thesis entitled "**Expression hash: A new approach for cancelable biometrics**" submitted by **Ayesha Riaz** have been found satisfactory for the requirement of the degree.

Advisor: **Dr. Naveed Riaz**
Signature: _____

Date: _____

Committee Member 1: **Dr. Hassan Tahir**

Signature: _____
Date: _____

Committee Member 2: **Dr. Imran Mahmood**

Signature: _____
Date: _____

Committee Member 3: **Dr. Shahzad Saleem**

Signature: _____
Date: _____

Abstract

Biometric authentication has garnered considerable attention from researchers in the past two decades. A biometric system is vulnerable to variety of attacks compromising the integrity of the process. One of the major attacks is on the stored template to extract information about the individual. Unlike passwords and pins, The biometric data is permanent and once stolen, cannot be canceled and re-generated. This has lead to the development of so-called cancelable biometrics scheme which supports the cancellation and re-generation of biometric templates. The schemes have already been successfully implemented for faces, palm prints, and fingerprints.

The idea of this work is to propose a new cancelable biometric approach which has tentatively been named "Expression Hash". The idea is to hash the expression templates with a set of pseudo-random keys which would provide a unique code (expression hash). This code can then be serve as a template for verification. Different expressions would result in different sets of expression hash codes, which could be used in different applications. This will greatly improve the privacy and security of applications. In case of compromise, the existing code can be revoked and can be directly replaced by a new set of expression hash code.

Dedication

I dedicate this thesis to MY MOTHER, TARRANUM RIAZ for her unconditional love, devotion and support.

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: **Ayesha Riaz**
Signature: _____

Acknowledgment

All praises be for you ALLAH; Ar Rahman, Ar Raheem and Al Qaadir

I am grateful to Dr. Naveed Riaz as my supervisor for his patient and forbearing guidance and support through out my thesis. Without his motivation and encouragement, completion of my dissertation was not possible.
I would also like to thank my committee members; Dr. Hassan Tahir, Dr. Imran Mahmood and Dr. Shahzad Saleem for their support and encouragement.
I am also obliged to Mr. Sajid Ali and Mr. Muhammad Nazir who guided me through out the whole process and enhanced my research.
Last but not the least, I am grateful to my husband, Murtaza Qamar for his constant support, encouragement and motivation throughout the thesis phase.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Human being is a social animal and interaction with its own is a basic need. To fulfill this purpose, humans invented languages evolving from signs, symbols and pictures to highly sophisticated languages governed by grammatical laws. However, privacy and security had been the major issues of communication i.e. only the intended recipient gets to read the privileged information without any modifications in the way. Authentication is a process which identifies that the claimant is the intended recipient or not.
Authentication process consists of two steps [1]; enrollment and authentication. During enrollment, the enrolling user is given a system specific identifier and a userś unique attribute is provided to the system called authenticators. When the user wants to access the privileged information, it presents its allotted identifier and corresponding authenticator in the authentication phase. The authenticators are something a user; knows e.g. PIN numbers and passwords, possesses e.g. chips, tokens, memory cards and smart cards, is e.g. fingerprints, iris and retina or does e.g. speech, gait and handwriting.
Passwords and PIN numbers etc. are easily stolen, forgotten or guessed using various known and easy attacks like password attack, guessing passwords, hijacking workstations, electronic monitoring, dictionary attacks, exploiting mistakes of users and multiple password usage attack [2]. Likewise, tokens, memory cards or smart cards have inherent drawbacks of being easily stolen or lost. So, there is a necessity of such authenticators which are not stolen, lost or forgotten, guessed or forged easily. The answer to these tribulations is using biometrics.
Biometrics is the measurement, calculation and analysis of a humanś distinctive identifiers which are either physical or behavioral in nature. Biometrics is divided into two categories; Permanent or static traits, soft or dynamic traits. Static biometric traits are mostly visual biometrics which dont́ or have a very little tendency to change over time such as fingerprints [3], palm

prints [4], face dynamics [5], retina [6], iris [7], tongue shape [8], etc. Soft biometric traits are mostly behavioral characteristics which tend to change quickly over life span such as gait [9], keystroke dynamics [10, 11], handwriting [12, 13], lip movement [14, 15], voice/speech [16], blinking [17] etc.
A biometric authentication system consists of two phases; enrollment phase and authentication phase. In enrollment phase, the user enrolls him to the system by presenting his identity and some unique authenticator i.e. biometric trait in this system. The biometric information is stored in the form of template in the database. The authentication phase is of two types; verification phase or identification phase. In verification phase, the user to be verified presents his identity and query biometrics, if the query biometrics matches with the one stored against his given identity, user is verified to be true. In identification phase, the user to be identified only presents his particular biometrics to the system, the system database is searched for the matching identity to the query biometrics.
There are four basic modules of the biometric authentication system which are used during the two phases of the system. They are; sensor, feature extractor, template database and decision module [18]. Sensor provides HCI (human computer interface) where the user can present its authenticating trait. The quality of the obtained trait is ensured by this module. Feature extractor module extracts information in numeric form called the feature vector from the presented biometric trait. Template databases are the physical storage devices to store the biometric templates in the form of feature vectors against their userś identity in the enrollment phase which are then used in the authentication phase for matching. The decision module calculates the match score of the stored template to the query template, the response of the system is in accordance with the match score.

1.1 Motivation

Technology is growing rapidly, effortlessly available, and inexpensive at the same time; same is with the biometric systems. They have been becoming inexpensive and are easily implementable in diverse user devices. However, biometric template databases are the most vulnerable as they are always present in the system and are subjected to easiest and most damaging attacks. They are needed to be secured physically and electronically.
In Pakistan, there is little or almost no training regarding the security and sensitivity of the databases and they are treated raw. When a biometric verification system is set up no care is taken to secure the database as it is considered an overhead which leads to great risks of security breaches

resulting in increased identity thefts.

1.2 Problem Statement and Objectives

Biometrics provides a promising solution for authentication. However, there are various problems related to the use of biometrics. Biometric characteristics are permanently associated to the user, once compromised cannot be used again i.e. no revocability and replacement, using cross matching all applications using the same compromised biometric trait are vulnerable too i.e. no diversity, the compromised template leaks the information about original biometric trait enabling attacker to make a physical spoof.
The present techniques to secure biometric templates require the usage of numerous large keys for diversity which leads to the problem of key storage and security.
Objectives of this thesis are to design a biometric template transformation framework such that:

- The compromised template can be revoked and updated easily using the same biometric trait i.e. large numbers of templates are obtainable using one biometric trait.
- Create diverse templates using same biometric trait such that cross-matching is not allowed if one of the templates gets compromised.
- The transformed template does not provide information about the original biometric trait.
- The transformation function does not degrade the overall recognition performance of the biometric verification system.

1.3 Solution Statement

A new user authentication method or technique is to be designed which is multifactor but is cost efficient as well and the supplementary information that is to be added in the form of a token is small and manageable as well. For this purpose, supplementary information in some other form is needed. In using face biometrics, this information is taken from the expressions of the face. Seven different expressions; happy, sad, angry, nervous, disgust, surprise and neutral. Thus, the token needed to add diversity is very small in size and easily remembered by the user.

1.4 Contributions

Following are the contributions made in this thesis:

- Multifactor authentication using a unique combination of permanent and dynamic biometrics i.e.; face identity and facial expressions.
- Instead of one biometric feature each user will have seven different biometric feature using seven facial expressions, HAPPY, SAD, ANGRY, NERVOUS, NEUTRAL, SURPRISE and DISGUST.
- Numerous templates are produced for one biometric feature using biohashing which provides protection, diversity, cancelability and revocability.
- A part of supplementary information that is added in the biohashing process is extracted from the facial expressions hence, shortening the length of the token number making it more easily memorable.
- A biometric template transformation framework based on biohashing is designed for face biometrics.
- A large set of transformed templates are obtained for a single person by using seven expressions and a small 4-digit token number for easy revocability and replacement.
- Increased diversity in templates is achieved by significant low match scores for two transformed templates of a single user.
- Original biometric template is hard to recover without the knowledge of token number.
- Other multifactor authentication systems need separate sensors to capture biometric instances and separate software modules to extract the useful information called biometric feature vector. Proposed authentication system makes use of the same sensor i.e.; camera and software modules are embedded together hence more cost efficient and performance efficient.
- Increased recognition accuracy of the biometric verification system.

1.5 Thesis Structure

Inclusion of the first chapter, Introduction, this dissertation has five chapters in totality. The highlights of following chapters are given as follows:

- Chapter 1: This chapter gives a brief introduction to the biometric cryptosystems, the problems associated to it, solution proposed in the thesis, objectives of the thesis work, its motivation and the contributions made to the field of biometric security. In the last section, an overview of the thesis structure arrangement is given.

- Chapter 2: It covers the literature review related to the dissertation. The first half gives a comprehensive overview of the various techniques in use to protect the biometric template databases. The overview includes the history of the diverse techniques used by various researchers. The second half throws light on the techniques used to extract facial expressions and methods developed for facial expression recognition.

- Chapter 3: The proposed methodology is discussed in detail. First the proposed algorithm for feature vector extraction for facial expression recognition is discussed. Then, the methodology to secure the extracted feature vector by creating expression hashes using biohashing is proposed. The biometric authentication system is designed based on facial expressions with a secure template database.

- Chapter 4: In this chapter, the designed experiments and the corresponding results are presented in detail to highlight the increased accuracy of the proposed system.

- Chapter 5: This chapter precisely summarizes the results, deduce conclusions and give future directions.

A detailed bibliography is given at the end for the facilitation of the reader.

Chapter 2

Literature Review

Biometrics is being used as a source of personś identification since the advent of time. Humans primarily used its five senses; see, hear, smell, touch and taste to distinguish between different identities. Documentation of biometrics for the purpose of user identification was first done by a French policeman named Alphonse Bertillon in late nineteenth century called Bertillonage [19]. It included the documentation of biometric traits like; height, reach, head length, head circumference, ear length, forearm length etc. as shown in the figure.

Figure 2.1: Measurement Diagram for Bertillonage

With the improvements in the technologies and techniques to acquire and process better biometric traits for user authentication and identification, biometrics are effectively and extensively used in governmental, forensic, medical, educational and commercial sectors.

2.1 Biometric System Vulnerabilities

Attacks on the biometric authentication systems lead to illegitimate access of the sensitive data like biometric templates stored in biometric template database to the attacker. Typically the security lapse is categorized as; inherent failures, administrator rights, insecure infrastructure, overt biometric traits.

2.1.1 Inherent Failures

The difference between the biometric instances of the same biometric trait of the same user is called intra-user variability which sometimes results in falsely rejecting the genuine user. The rate at which the system rejects the genuine user is referred as false rejection rate (FRR). The similarity between the biometric instances of same biometric trait of different users results in falsely accepting the illegitimate imposter. The rate at which the system accepts the imposter is referred as false acceptance rate (FAR). An adversary can leverage these failures by using zero effort attack in which it presents already available biometric trait to the system and expect to be accepted as a legitimate user with a non-zero probability.
Appropriately tuned threshold level and multifactor authentication can reduce the inherent failures of biometric authentication systems.

2.1.2 Administrator Rights

Administrator usually has the rights to accept the legitimate users whose biometric traits cannot be acquired because of some injury etc. The attacker can collude with the administrator to enroll it or be falsely accepted by the system.
Anonymous administrator, frequent audit trails and continuous user authentication can limit the breaches exploiting administrator rights [20].

2.1.3 Insecure Infrastructure

Infrastructure of the biometric system can be targeted by the intruders. The user interface in the biometric system is usually called the sensor which collects the biometric samples from the user. An adversary can physically damage the sensors, masquerade or alter its biometrics. Robust sensors, liveliness detection and biometric alteration detection can avoid the attacks on the user interface.
Communication between the modules of the system can be intercepted and replaced by the intruder to steal or replace biometric template to access the system or bar a user from accessing. Traditional cryptography techniques like TDEA, AES, and RSA [21–23] can be used to protect the communicating channels.
A computer virus or Trojan horse can replace any software module leveraging the loopholes in the system. Enforcement of secure code execution practices and thorough analysis can protect system against this vulnerability [24]. Template databases of the system can be read, replaced and modified by the attackers to achieve illegitimate acceptance to the system. Spoof of biometric trait can be produced using stolen biometric template using hill climbing attack [25] which is then offered to the system sensor module to acquire access. The stolen template can also be replayed at the matcher. Stolen template of a user from a compromised system can be used for exploitation for some other purposes in different systems [26]. A number of techniques to limit these attacks are discussed in the following sections.

2.1.4 Overt Biometric Traits

Biometric traits like fingerprints, face, voice, signatures, and iris are not secret. They can be covertly captured without the knowledge of the victim. However, the digital identity related to the victim is difficult to determine. So, system databases are more productive targets to get large amounts of biometric information with the digital identities. Thus, the biometric template database security is of high importance.

2.2 Requirements for Protection Techniques of Biometric Template Databases

Following are the required properties for the techniques adopted to protect biometric template databases [27];

- Diversity/non-likability: the biometric templates stored in the databases of two different systems should be diverse enough i.e. they cannot be linked to a specific single user resulting in the release of any information and does not permit cross matching.
- Revocability/cancelability: the revocation of the compromised template should be easy and quick as well based on some new biometric instance of the same biometric feature of the attacked user.
- Security/non-invertibility: the template stored in the database should not reveal any information about the unique biometric feature extracted so no physical spoofs can be constructed.
- Performance/Efficiency: the efficiency of the system to correctly recognize the user should not be degraded using the protection techniques.

2.3 Biometric Template Protection Techniques

Classically, passwords or PIN numbers are used to protect the sensitive digitalized information. However, password protection is not a good idea for biometric template security as the user needs to remember the complex passwords and provide it to the system for each time it needs to get authenticated and is subject to be forgotten and the security provided is insufficient.
Hence, a number of hardware and software solutions are presented to solve the problem of biometric template protection. Hardware solution include smart cards, match on card or card-on-system to protect the biometric template data. All the modules and interfaces of the system are present in the card, so the advantage of card-on-system is that the sensitive information never leaves the card. Commercial example of such a system is privaris PlusID [28]. However, for large scale applications, they are unsuitable because of their costly hardware; the security will be pricier for users or the company. The user has to carry and safeguard the card all the time for valid accessibility but the card might get lost or stolen. The attacker can attempt to get biometric template from these stolen cards leaving the user at high risk.

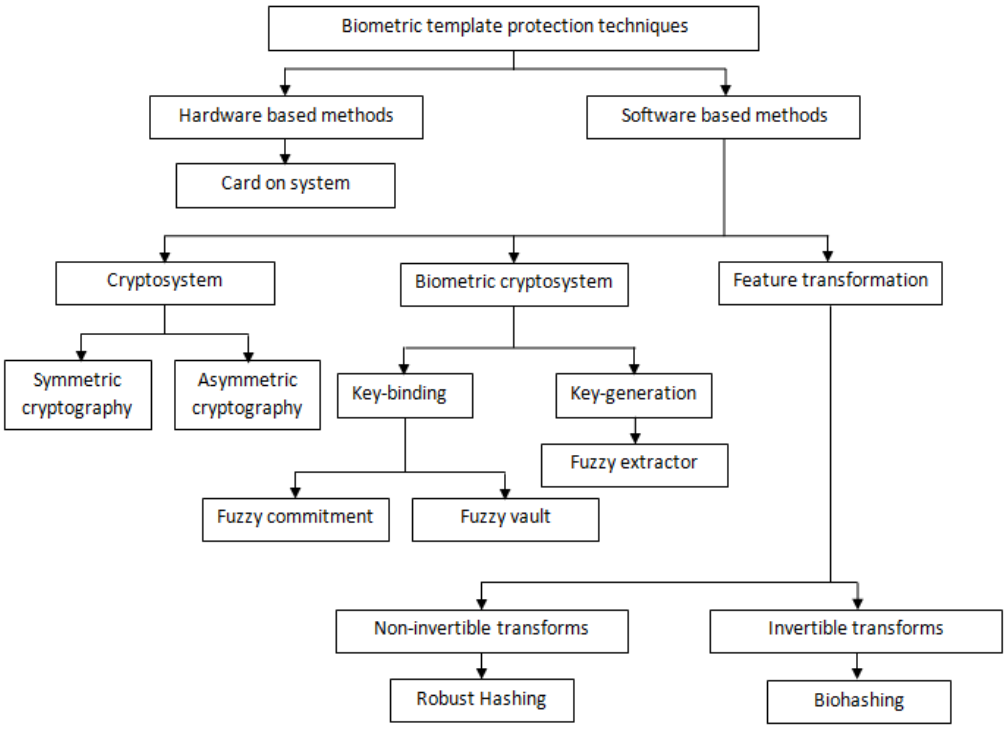

Figure 2.2: Protection Techniques for Biometric Templates

Software solution involves combining the biometric templates with a key or system generated random numbers such that the stored templates expose little or no information regarding the original templates. Software based biometric template protection techniques are classified into subsequent categories:

- Cryptosystem
- Feature transformation
- Biometric cryptosystem

2.3.1 Cryptosystem

Cryptosystems [29] are made up of two processes; encryption and decryption. The conversion of legible messages commonly called plaintext is converted to an illegible format called cipher text is called encryption and decoding the unintelligible cipher text back to the original plaintext message is called decryption.

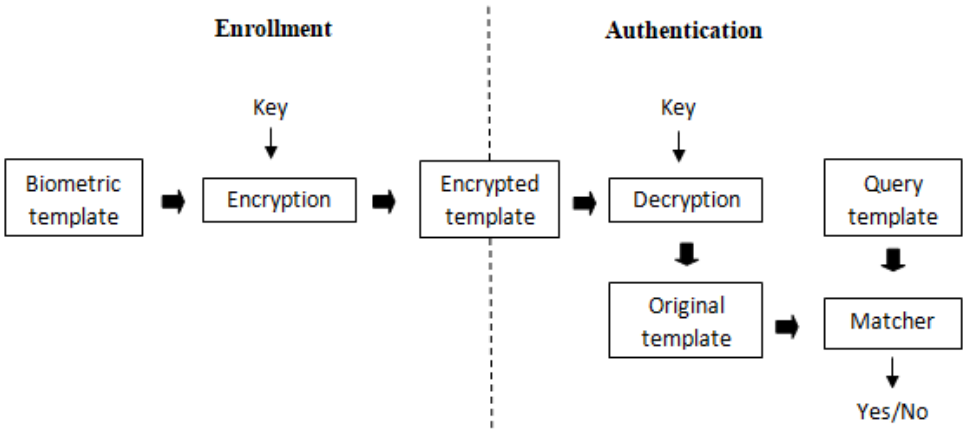

Figure 2.3: Cryptosystem Framework

Symmetric and asymmetric cryptography are the two techniques, used for encryption and decryption. In symmetric cryptography, a secret key is used to encrypt the biometric template; the same secret key is then used to decrypt the template each time in the authentication phase. In asymmetric cryptography, the system has a pair of keys; public key is known to everyone and private key is only known to the system only, during enrollment phase, the template is encrypted with the public key and then can be discarded, the template is then decrypted during the authentication phase with the private key for matching purposes. The attacker cannot get the original template unless he has the correct secret key or private key of symmetric or asymmetric cryptography respectively.
However, the unbreakable cryptosystems and keys used in cryptography are still facing many challenges, two of which immediate attention are:

- Management of key: The standard and approved symmetric and asymmetric cryptosystems like TDEA, AES, and RSA etc. rely on the assumption that the secret shared key in symmetric and private key in the asymmetric cryptosystems is secure and uncompromised. If an unauthorized illegitimate user can get hold of the key, the intended security provided by the cryptosystem is entirely broken. Hence, secure management of the key is a challenge.

- Size of the key: The keys used in symmetric and asymmetric cryptosystems are 128-256 bits and 1024-2048 bits long respectively. These are very long and random with high entropy and are a problem because

they are not possible to be memorized by humans. So, they are stored in a password protected place [30].

2.3.2 Biometric Cryptosystems

In biometric cryptosystems, keys are generated and matched in encrypted form at the time of enrollment and authentication. These keys generated might vary for the same legitimate user because of the variation in the biometric instances of the same biometric trait. to counter the intra-user variations of the system, some public information is drawn out of the biometric feature called helper data. The helper data assists in the accurate reformation of the key for the legitimate users.
Helper data based protection techniques are divided into two different models as:

- Key binding model: the biometric template obtained from the sensor is bound with a unique key at the time of authentication and helper data is generated. The correct unique is to be reconstructed with the helper data and the query biometric instance for legitimate access.
- Key generation model: a unique key and helper data is generated from the biometric template at the time of authentication and helper data storage is unnecessary. On authentication, helper data and query biometric instance is used to retrieve the generated unique key for accessibility.

Key Binding Model

In key binding model, the biometric template collected from the user is cryptographically etched with a special secret key chosen by the system itself. The distinct resultant entity is called the helper data and is stored in the database. The stored obscure helper data does not reveal any information about the original template of the user. On authentication, the query biometric template differs with the one used in enrollment due to intra-user variations. However, these variations are with the tolerance limit already defined by the system. The key generated with the query biometric template and the helper has the same amount of variation as between the enrollment and the query templates. Successful matching occurs when exact key is obtained using error correcting codes.

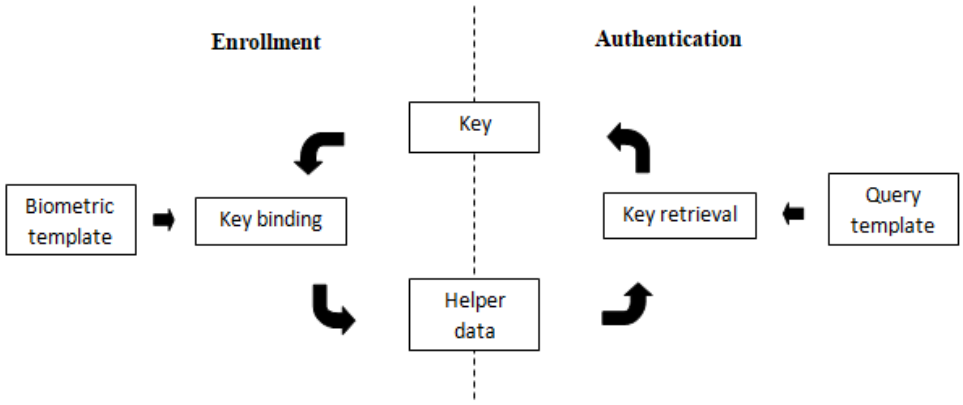

Figure 2.4: Key Binding Framework

- Advantages:
  - Intra-user variability is tolerated by the use of error correcting codes in the system.
- Disadvantages:
  - Use of error correcting codes for the retrieval of keys inhibits the use of highly sophisticated matchers decreasing the overall performance of the system.
  - Key binding models does not provide diversity and revocability or cancelability. Helper data is designed cautiously to cater for intra-user variations.

Key binding models include fuzzy commitment and fuzzy vault methods.

Fuzzy Commitment

The model of fuzzy commitment was first pioneered by Juels and Wattenberg [31]. In this model, cryptographic techniques and error correcting codes are combined. A code word is generated using the error correcting codes. A hash of this code word is stored in the database. The user provides a biometric template and a bit string called witness is extracted from it. The code word generated and the witness extracted are bound together to produce a difference vector and stored in the database as helper data. The helper data, difference vector and query biometric template is used to regenerate the codeword. Hash of the query codeword is matched with the stored hash. User is granted access on correct match of hashes.

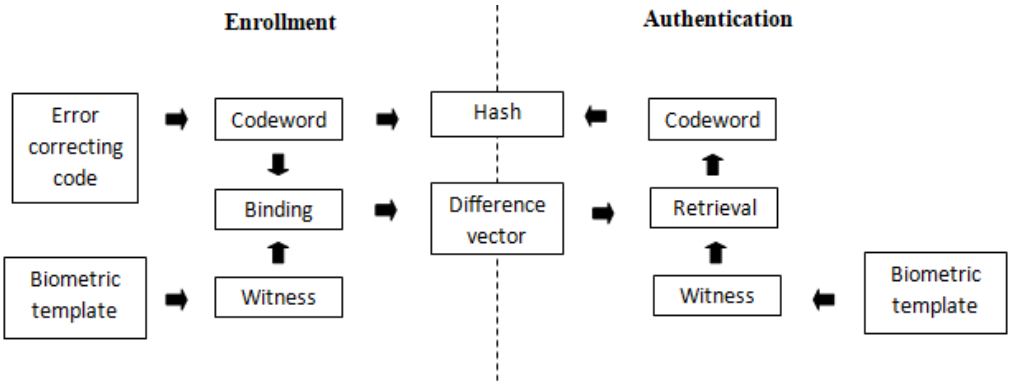

Figure 2.5: Fuzzy Commitment Framework

Fuzzy commitment method was proposed by Teoh and Kim [32] for fingerprints in which Garbor filter is used to acquire fingerprints and are transformed into binary bit string using randomized dynamic quantization. A user specific token is used for the generation of a random matrix. The feature vector and the random matrix determine the transformation function. Nandakumar [33] extracted the binary stings of the minutia set of fingerprints by quantizing the Fourier phase spectrum. For alignment of the minutia set, high-curvature regions are used.
Hao et al. [34] proposed iris code based fuzzy commitment method. Error originated by biometric template variance of types; burst errors and bit errors are corrected or eliminated by applying Reed-Solomon and Hadamard error correcting codes.. Bringer et al. [35,36] used binary Reed-Muller codes to create more efficient decoding matrix. Systematic analysis was carried out by Rathgeb and Uhl [37,38] on error distribution in different algorithms of iris recognition. Components are selected by applying context based information which is bound with Hadamard codewords for the generation of keys. Performance of the fuzzy commitment for iris is improved making use of various techniques in [39,40]. Iris templates have one dimensional circular shifts so, template alignment is practicable.
Van der Veen et al. [41] proposed a face features based fuzzy commitment scheme in which the real values are converted into binary strings by a simple threshold function. Witness is obtained by discriminative feature bit selection. A consistent key biding scheme is presented by Ao and Li [42] binding near-infrared face features with error correcting codes. Lu et al. [43] suggested that binary strings are obtained using PCA (principal component analysis) which is used in fuzzy commitment method.
Maiorana et al. [44] presented fuzzy commitment model for online signatures in which error corrections are made using intra-user variability. An additional

model was suggested for online signatures in [45].
Zheng et al. [46] employed tolerant lattice functions to remove biometric variance in place of conventional error correction codes. Error correction code syndrome was introduced by Dodis et al. [47] in which is correction code is stored in the template to reconstruct original biometric input during authentication phase.
Comparison of fuzzy commitment model experimental results:

Table 2.1: Experimental Results Comparison of Fuzzy Commitment Model

Authors	Biometric Character	FRR/FAR
Teoh and Kim	Fingerprints	0.9/0
Nandakumar		12.6/0
Hao et al.	Iris	0.47/0
Bringer et al.		5.62/0
Rathgeb and Uhl		4.64/0
Van der Veen et al.	Face	3.5/0
Ao and Li		7.99/0.11
Lu et al.		∼30/0
Maiorana et al.	Online Signatures	13.07/4

Fuzzy Vault
In 2002, Ari Juels and Madhu Sudan [48] initiated the idea of fuzzy vault model for biometric template protection. The feature set is extracted from the biometric trait provided by the user. A pre-decided secret key is used in the model often provided by the user himself. A polynomial is created from the secret key and the set of bits chosen from the extracted feature set. Polynomial encoding and error correcting codes are applied on the polynomial to create fuzzy vault. This vault is stored in the database. Actual points of the polynomial are hidden the vault using Chaff points. When authenticating, polynomial is created from the feature vector and the vault stored. The generated polynomial is then used to get secret key. If the secret key matches with the one used, user is successfully authenticated.

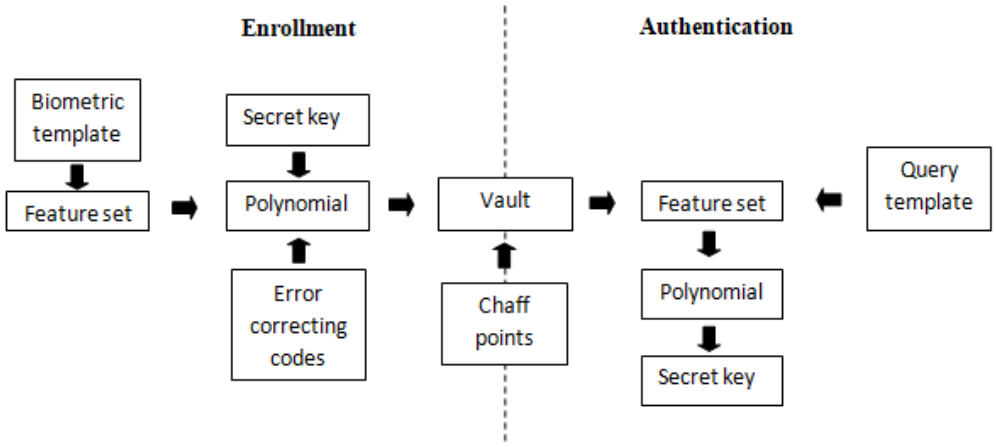

Figure 2.6: Fuzzy Vault Framework

Clancy et al. [49] anticipated the first practical implementation of fuzzy vaults using fingerprints. A set of minutiae were mapped on the polynomial to create the fuzzy vault. After the generation of the vault, Chaff points are added to hide the original points. During authentication, polynomial is recreated using Reed Solomon codes from which key are obtained. Nandakumar et al. [50] suggested that helper data is required to assist the alignment process as points chosen from the feature vector are pre assumed to be already aligned. Uludag et al. [51] proposed line-based minutiae representation for fingerprints based fuzzy vault model which was tested on 450 pairs of fingerprints. Many approaches for alignment improvement of fingerprints in fuzzy vaults are suggested in [52]. In [53] invariant minutiae representation of the rotation and translation form is proposed. Nagar et al. [54] presented the enhancement of the security and accuracy of the fuzzy vaults based on fingerprints by creating feature set with orientation information of minutiae points. Li and Hu [55] made used of highly discriminative P-P minutiae structures for proposing an alignment-free fuzzy vault for fingerprints. This superior quantization permitted its direct usage in the matcher module. To enhance the security of the P-P minutiae structures correlation is abolished by transformation and encoding.
Lee et al. [56] suggested a fuzzy vault model for iris biometrics; iris biometrics is naturally aligned so ICA (independent component analysis) is used to acquire a feature set. Wu et al. [57] anticipated iris biometrics based fuzzy vault model, in which the unordered feature set is obtained by evaluating the gray scale value of the blocks. Blocks are made of the image after the image is acquired and preprocessed. Feature set is normalized for reducing noise.

Reed-Solomon and hash functions are used to create chipper keys. Reddy and Babu in [58] proposed the security enhancement of fuzzy vault scheme for iris biometrics by vault and secret key by password hardening.
Wu et al. [59] suggested fuzzy vault method for palm-prints. Kumar and Kumar in [60, 61] proposed fuzzy vault model for palm-prints in which real-valued DCT (discrete cosine transform) coefficients are used to construct unordered feature set.
Wu et al. [62] presented fuzzy vault method for face biometrics in which PCA (principal component analysis) is used to extract feature set is and Reed-Solomon codes are used to counteract variance. Wang and Plataniotis [63] suggested fuzzy vault scheme for face biometrics. Distance vectors are calculated between the facial biometric features and a pair of random vectors. These distance vectors are then 2-D quantized. Variations are handled using windowing process. This method results in zero error rate as both biometrics and key are variable.
Kholmatov and Yanikoglu [64] presented fuzzy vault model for online signatures while Eskander et al. [65] anticipated a biometric cryptosystem for images of offline signatures using fuzzy vaults.
Comparison of fuzzy vault model experimental results:

Key Generation Model

In key generation model, helper data is directly produced by the biometric template of the biometric trait presented by the user. The key which is a random bit string is generated by the template itself. The helper data is not necessarily stored in the database, but its storage enables the systems to revoke and update. While authenticating query biometric template and helper data is used to generate the key. The newly generated key is matched with the stored key. If the match score is high, access is granted.

- Advantages:
  - Direct generation of keys from biometric templates is tempting and is valuable in various cryptographic applications.
- Disadvantages:
  - Generation of keys with high stability and high entropy is not an easy task.

Table 2.2: Experimental Results Comparison of Fuzzy Vault Model

Authors	Biometric Character	FRR/FAR
Clancy et al.	Fingerprints	20-30/0
Nandakumar et al.		4/0.04
Uludag et al.		27/0
Nagar et al.		5/0.01
Li and Hu		30.91-0/0.28-0, 21.63-5.78/0.28-0
Lee et al.	Iris	∼7/0
Wu et al.		5.55/0
Reddy and Babu		9.8/0
Wu et al.	Palmprints	0.93/0
Kumar and Kumar		∼1/0.3
Wu et al.	Face	8.5/0
Wang and Plataniotis		0.5/7.38
Kholmatov and Yanikoglu	Online signatures	8.33/2.5
Eskander et al.	Offline signatures	14-16/2.39-1.41

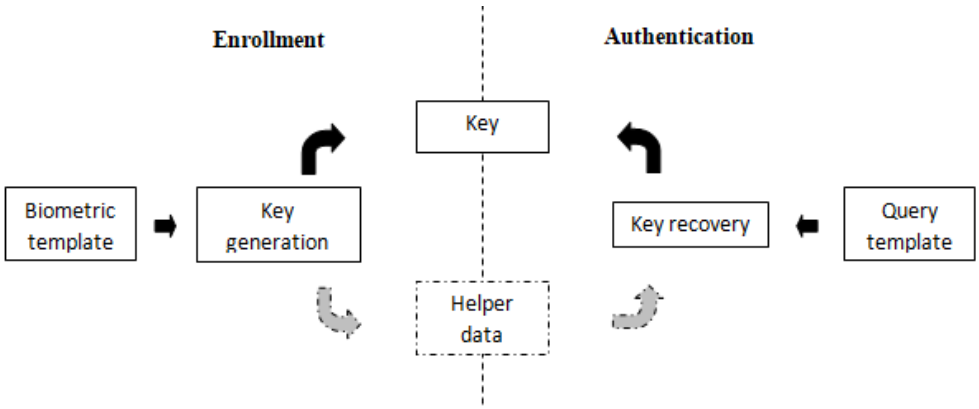

Figure 2.7: Key Generation Framework

For key generation model secure sketches- fuzzy extractor method is used. Blanton and Aliasgari [66, 67] explained the present problems in acquiring multiple sketch constructions from a single biometric trait. Chi et al. [68] presented fuzzy extractor model for multiple biometric traits in which a highly stable codeword is created using various biometric characteristics. Modalities are then used to extract the original codeword. Key entropy and accuracy is enhanced when compared to the uni-biometric cryptosystem.
The biometric information related to the user is in the generated biometric key so biometric template storage is unnecessary. Original biometric template can never be exposed even if the generated key is compromised. Thi and Tran [69]united the concepts of the biometric authentication with Kerberos. Nazatul et al. [70] proposed the online secure voting system making use of biometric and password based security.
Taniguchi et al. [71] proposed and practically implemented fuzzy extractors based on soft decision to produce a consistent and extensive bit string using output of physically non-replicable functionsúsing CMOS (complementary metal oxide semiconductor) technology. Kang et al. [72] presented a well-organized implementation of key generation model using PUF and fuzzy extractors replacing the output of hash functions with the BCH (Bose-Chaudhuri-Hocquenghem) code syndrome. Herder et al. [73] suggested enhanced capability of error correcting codes by avoiding noise using trapdoors and security is limited to LPN (learning parity with noise) and PUF which are secure cryptographically.
Xi et al. [74]achieved high verification performance using near-equivalent dual layer structure check. Verification method based on dual-layer structure check verification is used for local structures of minutia for alignment free fingerprint fuzzy extractor model. Yang et al. [75] presented fuzzy extractor model for fingerprints which is registration free using Delaunay triangle.

2.3.3 Feature Transformation

In feature transformation scheme, the user provides its biometric traits to the system at sensor. The received biometric template is processed into a biometric feature vector. The generated biometric feature vector is then passed through a transformation function triggered by some external information defined by the user itself or the processing system. The transformed template is then stored in the database for authentication. When user provides his query template for verification, the query template is processed and transformed using same functions and parameters. The query transformed template is then matched with the one stored in the database. If the match

score is above the pre-defined threshold, successful authentication is carried out.

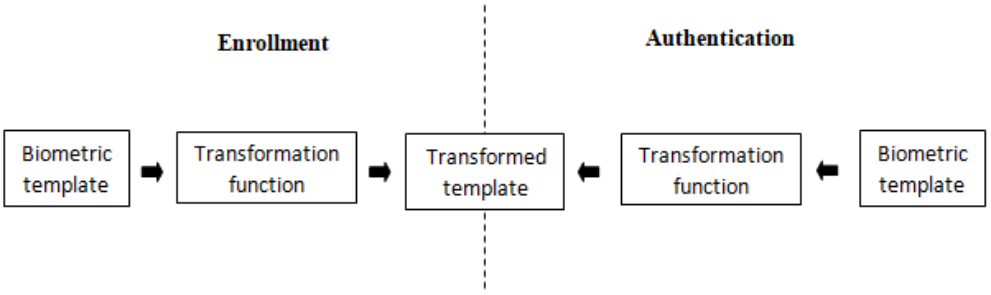

Figure 2.8: Feature Transformation Method Framework

Following two models are categorized as feature transformation schemes:

- Invertible transformation models: the transformation functions used are invertible in nature i.e. if the key and the transformed template are known; original template of the user can be obtained. These models are also known salting schemes.
- Non-invertible transformation models: the transformation functions used are non-invertible in nature i.e. they are one-way functions and the original template of the user is hard to recover when key and transformed template are known.

Invertible Transformation Models

In invertible feature transformation models, biometric features are transformed using invertible functions. The invertible functions use a user defined random key or a small token. The supplementary information addition makes it hard to recover the original template as entropy is enhanced.

- Advantages:
  - User defined random key addition results in low false acceptance ratio.
  - Multiple transformed templates can be obtained using same biometric feature with various keys.
  - Compromised templates can be revoked and updated easily using different random keys.

- Disadvantages:
  - Original template can be compromised if the random key is compromised as function used is invertible in nature.
  - Performance of the system can be degraded as the query template is transformed and then matched.

Biometric Salting or Biohashing is used as invertible transformation model.

Biohashing

Toeh et al. [76, 77] anticipated biohashing method for face biometrics. FDA (fisher discriminant analysis) is used to extract discriminative features. The feature vector obtained is projected randomly in orthogonal direction defined by the user random projection. The resultant vector is then binarized to counter intra-user variations and it is done in a way such that vector has equal number of zeros and ones thus achieving maximum entropy. Results show FAR to be zero and FRR to be 0.93%.

Biohashing scheme is implemented to numerous biometrics like iris biometrics [78, 79], fingerprints in [80, 81] and palm-prints [82]. The use of hashes generated after biohashing in key-binding models is proposed in [83,84]. Kong et al. [85] proposed implementation of face hashing and achieved zero error rate by the addition of unique tokenized random numbers. Teoh et al. [86] suggested that multi stage random projection can resolve the stolen token issue. Various improvements of the biohashing scheme are presented by Lumini and Nanni [87, 88].

Non-Invertible Transformation Models

In non-invertible transformation models or robust hashing schemes, the transformation functions used are non-invertible in nature. The extracted biometric features from the biometric templates are first transformed and then stored. The transformation functions are one way functions which are computed easily in one direction but are very hard to invert.

- Advantages:
  - Original template is hard to recover if the transformed template is compromised thus providing better security than invertible transformation models.
  - Diverse transformed templates for various platforms can be obtained using different keys.
  - The compromised template can be revoked and updated as multiple templates can be obtained using same biometric trait.

- Disadvantage:
  - It is difficult to design a transformation function which preserves both non-invertibility and discrimination i.e. the transformed templates should have large variations for different users (inter-user variability) as presented by the un-transformed templates.

Sutcu et al. [89] presented a non-invertible transformation model for face images. Biometric feature privacy is achieved by many transformed versions of the same biometric template. Cryptographic hash function ensures the properties of confusion and diffusion.
Ratha et al. [90, 91] proposed three non-invertible transformation schemes; polar, folding and Cartesian to generate secure transformed fingerprint templates to store in the database.during authentication phase, matching is done in the transformed state. The original template is hard to recover as it as computationally hard as guessing randomly.
Zuo et al. [92] demonstrated diverse algorithms for creating cancellable iris biometrics. Hammerle-Uhl et al. [93] applied classical algorithms to create non-invertible templates for iris biometrics.

2.4 Facial Expression Recognition

Facial expression recognition has applications in industry, commerce, behavioral and medical sciences. Classification of facial expressions into the categories of happy, sad, angry etc. is called facial expression recognition. Accurate and immediate detection and recognition of expressions by humans is unproblematic but expression recognition is still a challenge in the field of computer vision i.e. machines are not smart enough to accurately and precisely distinguish different expressions.

2.4.1 Issues in the Facial Expression Recognition

Extensive researches and experiments have been carried out to discover a consistent and error-free expression recognition system which addresses the issues of illumination variations, face orientations and expression differences based on region and culture and image resolution. The computation time for matching should be small i.e. close to real time. A system for smart phones [94] was proposed using deep convolution neural network consisting of five layers and sixty-five thousand neurons for real time application.

The existing facial expression classification frameworks require images of faces to be acquired in a restricted environment using high-resolution cameras [95]. When these techniques are used with real-time applications like video conferencing, visual surveillance and smart meetings perform poorly. These traditional approaches do not take orientation difference and lower resolution in to consideration resulting in poor performance.

2.4.2 Descriptors

Effective expression recognition frameworks fundamentally require the extraction of information from face images. The contents of the images are described by visual descriptors. They give elementary information about the shape, color and texture of the image. An ideal feature descriptor should be: robust against rotation, scale or photometric variations, distinctive, low-dimensional and efficient.
The descriptors can be divided into two categories: general information descriptors or global descriptors and specific domain information descriptors or local descriptors. Global descriptors generalize the whole image and describe it entirely including its shape, texture and contours. Low level applications like object detection and classification global descriptors are used. Shape matrices, invariants moments, HOG features are examples of global descriptors.
Local descriptors like SIFT, LBP, WLD describe the image dividing it into small blocks or local patches. Information is extracted from these patches to describe the image. For high level applications like object recognition and facial expression recognition local descriptors are used.

2.4.3 Types of Feature Extraction Techniques for Facial Expression Recognition

Feature extraction techniques are categorized as appearance based and geometric based [96]. In appearance based extraction, multi scale coefficients are obtained by applying image on whole image or portions of the image to detect changes in the appearance of the result. However, multi scale coefficient extraction is time consuming and memory exhaustive [97].
Geometric based extraction detects the location and shape of facial features which work relatively superior than appearance based extraction techniques. Consistent and accurate tracking and detection of facial features is required which in real-time environment is not easy to handle [98]. Li et al [99] proposed an unsupervised method to extract significant features for a global

multi layer satellite image feature categorization technique. Features are extracted using dual layer approach integrating features for scene categorization from both layers using SVM (support vector machine).

2.4.4 Related Work

Numerous researchers are using local descriptors like LBP and WLD recently for the extraction of texture information from images. Several variants of LBP have been anticipated to deal with the issue of low resolution images. Boosted LBP with SVM is proposed by Shan et al. [100] using multi scale images for increased performance.
LBP (local binary patterns) and WLD (weber local descriptor) are used as local descriptors which conserve the local information but are also robust against occlusions, pose variations and misalignment in comparison to holistic methods [101]. LBP and WLD extract features by calculating histograms of the whole image resulting in loosing spatial information. Local information preservation and multi scale analysis is achieved by down-sampling image on diverse scales forming pyramid of image [101]. Feature vector is calculated by concatenating the histograms calculated for each block at different scales. Increased number dimensions of redundant data results in performance degradation of local descriptors, hinders facial expression recognition system for real time because of added computational cost. Optimization algorithms are proposed [102, 103] to decrease dimensionality but it is difficult to optimally reduce the number of distinguishing features because of intra-user and inter-user variability of facial expressions.
Khan et al. [104] demonstrated a novel method called PLBP (pyramid local binary pattern) in which the image is cultivated into diverse portions at various resolutions to obtain spatial layout. Spatial layout information and local texture stimuli is preserved extracting features from explicit regions of face. These features are then combined to get unique feature vector. Decision level fusion of features is done in [105] with sparse representation based classification. Features are extracted using HOG and LBP descriptors which are found to be robust against facial occlusions.
AWELBPP (adaptively-weighted extended local binary pattern pyramid) is modified face descriptor proposed by Gao et al. [106]. Transformation into a pyramid is done on the face image to characterize it into diverse multi-resolution images, these images are then separated into horizontal sub-images and ELBP (extended local binary pattern) is applied to get features vectors. Entropy based useful information is generated for each sub image and then they are fused to high performance feature vectors.
In [107] a new technique is presented to classify emotions maximizing the class

independency. An extension of facial descriptor es-LBP (expression specific local binary pattern) gathers information about specific facial points. Link between the expression classes and facial features is improved by preserving projection presentation of regularized class locality. Other descriptors than LBP have been proposed by researchers.

Spatio temporal descriptor is proposed in [108]. Spatial and dynamic facial expression information is obtained by expanding histogram of gradients of spatial pyramid to spatial temporal domain integrated using dense optical flow. In [109], Guo et al. used ICA (independent component analysis) and PCA (principal component analysis) to extract features. Tong et al. [110] fully described the deformation, facial muscles textures and wrinkles using LGC (local gradient coding) algorithm. A comparison of results of LBP, LCG, HOG and LDP is presented by Kumari et al. [111].

Wang et al. [112] presented a innovative descriptor for recognition of expression called SLFDA (sparse local fisher discriminator) exploiting the sparse property of FDA. Framework for automated facial expression recognition is developed by Happy et al. [113] based on appearance based features. SVM is used for one on one classification purpose. Efficient and reliable performance is claimed by the authors for diverse resolutions.

Ucar et al. [114] provided accuracy rates of 95.15% and 94.65% for CK+ and JAFFE databases using integrated features from local curvelet transform and statistical features. Eleftheriadis et al. [115] used DS-GPLVM (discriminative shared Gaussian process latent variable model) for expression recognition. Experiments are performed for multiple views and the technique is proved to be robust against real time various views. Concept of multimodal learning for expression categorization was applied by Zhang et al. [116]. To achieve robustness, structured regularization is used and correlation is represented between landmark and texture modalities.

Chan et al. [117] proposed a framework to deal with both controlled and uncontrolled environment images using HOG-TOP (histogram of gradients from three orthogonal planes) which extracts dynamic texture features which are then blended with geometric features. STTM (spatio-temporal texture map) is used by Kamarol et al. [118] to capture the spatial and temporal variations of face image. Dynamic features are extracted by dividing image into small blocks and distinguished information is stored in form of histograms.

Chapter 3

Proposed Methodology

In this thesis, a novel approach to cancelable biometrics is proposed called expression hashing. Expression hashing is basically biohashing of biometric feature vector classifying facial expressions. The feature vectors extracted for a person are capable of identifying that person and are essentially diverse enough for seven expressions that they are detectable easily. Thus, feature vector extracted is different for different people and different for different expressions of the same person. Instead of storing these unique feature vectors in the database as plain text, they are transformed using a small token into expression hashes.
Transforming expression feature vectors into expression hashes has inherent advantages including original biometric template protection against stealing and replaying, revocability and easy replacement due to diversity and non-linkability and additional security as token information is included.
Following sections are discussing the proposed methodology of this thesis work.

3.1 Expression Feature Vector Extraction

Khan et al. [119] proposed a novel descriptor using existing techniques of LBP, WLD and DCT called WLBI-CT (Weber Local Binary Image Cosine Transform) in effective manner. Local level details are obtained using block LBP and robustness against real time variations is covered by WLD transform. Information about dominant orientation and spatial layout textures is present in Weber local binary image. Local representation information is obtained using orientation components in both horizontal and vertical directions. To enhance the local representation binary image is divided into blocks.

However, classification performance is degraded using LBP and WLD techniques because of the redundant features. Images are transformed in frequency domain to increase the discriminative power of the descriptor. DCT (discrete cosine transform) features are investigated by Dabbaghchian et al. [120] where high variance features are extracted by DCT in zigzag manner. After applying DCT, middle features contain high recognition power, so upper left corner and middle features are used.
Hence, classification performance is increased using relatively lesser number of features than the previously used state of art techniques. LBP image is divided into blocks preserving texture information and exhibiting more discriminating power.
Following are the steps to calculate the feature vectors from an image.

- image preprocessing
- conversion to LBP orientation image
- conversion to WLD orientation image
- feature extraction using DCT in zigzag manner
- generate final feature vector

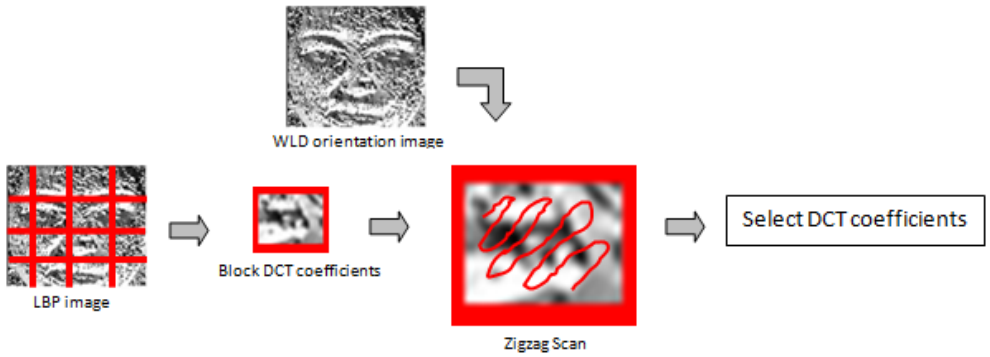

Figure 3.1: Proposed Framework for Feature Vector Extraction

3.1.1 Image Preprocessing

Image preprocessing means training or preparing image for the algorithm application. Face is detected from the image to locate the face region using face object detection algorithm [121]. After face detection, the face region is cropped to get only the required portion of the image. To normalize the illumination effects on the image histogram equalization operation is applied.

3.1.2 Conversion to LBP Orientation Image

The normalized image is then converted to local binary image for texture classification. LBP (Local binary pattern) was introduced by Ojala et al. [122] which surpasses rest of techniques in different applications [123, 124]. Ahonen et al. [125] used LBP for face recognition. Silva et al. conducted a survey on the improvements of the LBP [126].
To extract features with LBP image the image is first divided into N regions or blocks. Each cell contains equal pixels of the image. In each cell or block, each pixel is compared to its neighboring pixels in a circular manner i.e. clockwise or anti-clockwise. The value of central pixel is considered threshold value in each comparison. Values above the threshold value are one and below it is zero. Hence, an eight binary number is produced reading it in invariant manner which is often converted into decimal for convenience.

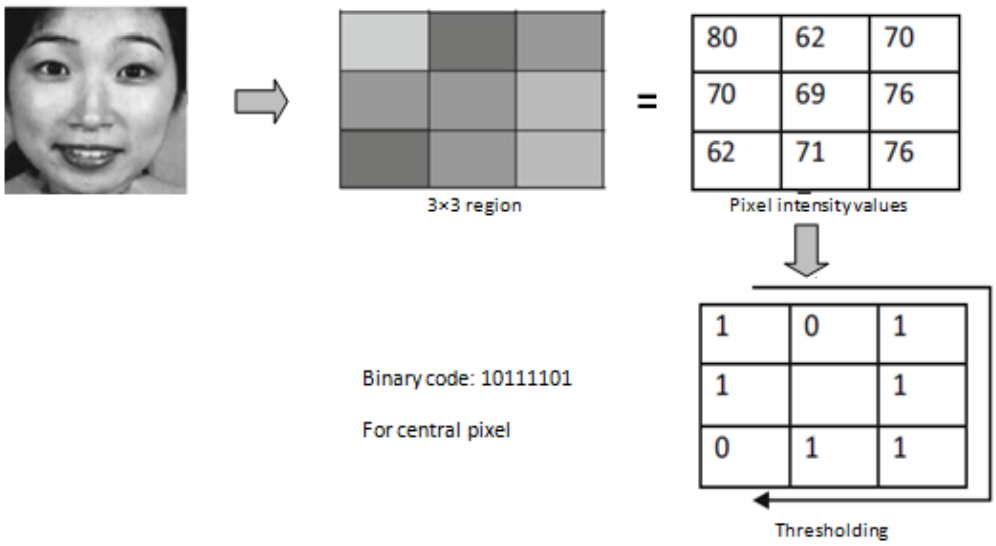

Figure 3.2: Binary Value Calculation for LBP images

In the next step, for each pixel, extensions are made in the neighborhood size uniformly and rotation invariably [124]. LBP operator is defined by following equation:

$$LBP_{X,R} = \sum_{i=1}^{X-1} 2^i S(X_i - X_C)$$

Where, X is the total number of neighboring pixels, R is the radius and X_C is the central pixel.

The binary code is obtained using the following equation:

$$S(X_i - X_C) = \begin{cases} 1 & \text{if} X_i - X_C > 0 \\ 0 & \text{if} X_i - X_C < 0 \end{cases}$$

Uniformity, rotation invariance i.e. clockwise or anticlockwise and circular neighborhood is the parameters used in LBP [124]. Histogram of the computed values over the block is calculated showing the combinations of which pixel values are greater or smaller than the central pixel value. For a 3 neighborhood, eight binary numbers, $2^8 = 256$ dimensional histogram is obtained. The histogram of 256 LBP code is used for local texture features description of expression images.

3.1.3 Conversion to WLD Orientation Image

Chen et al. [127] developed WLD (Weber local descriptor) for capturing texture information exploiting the pixel intensity change ratio. WLD consists of two major components: differential excitation and orientation. Differential excitation $\xi(p_c)$ is the ratio of relative intensity difference to the neighboring pixels and current pixel intensity. The differential excitation is given by:

$$\xi(p_c) = arctan[\sum_{i=0}^{x-1}(\frac{p_i - p_c}{p_c})]$$

Where, x denotes the neighbors and pi is the i^{th} neighbor of the p_c. The results are smoothed out using the arctan function. Differential excitation component captures the local patterns; high value indicates a spot or edge. Orientation component of WLD is represented by the gradient orientation of the current pixel $\theta(p_c)$. Following equation is used for gradient orientation computation [127].

$$\theta(p_c) = arctan(\frac{p_7 - p_3}{p_5 - p_1})$$

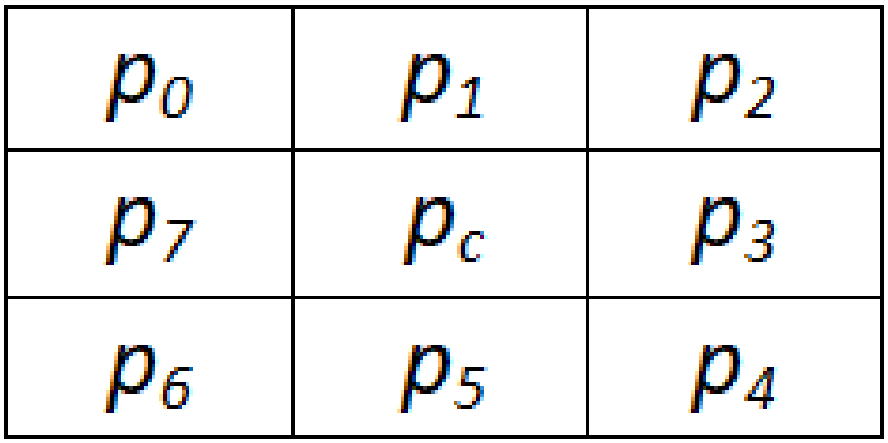

Figure 3.3: WLD Based Feature Extraction Neighborhood Pixel Arrangement

Where, p_c is represented by p_1, p_3, p_5, p_7 as shown in the above figure.
Information is lost when differential excitation is calculated by averaging in an interval. To, tackle this issue, LBP excitation component and WLD orientation component are combined by Khan et al. [119].

3.1.4 Feature Extraction Using DCT in Zigzag Manner

Discrete cosine transform is a popular transformation function for signal and image processing [128]. DCT (discrete cosine transform) and DFT (discrete Fourier transform) [129] is widely used in expression recognition [130]. DCT transforms the image in the frequency domain where maximum information is stored in low frequencies hence providing strong energy compaction and high computational efficiency. The highly variant components of the transformed DCT image are present in the top left corner of the image. Unlike PCA and LDA, frequency domain methodologies are data independent and result in better analysis as compared to spatial domain LBP.
The DCT transformation equation for N image is given by;

$$F(u,v) = \alpha_u \cdot \alpha_v \cdot \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [cos(\frac{\pi u}{2N}(2x+1)) \cdot cos(\frac{\pi v}{2M}(2y+1)) \cdot f(x,y)]$$

Where, $f(x,y)$ is the pixel intensity of x^{th} row and y^{th} column $u = 0, 1, \ldots, N-1, v = 0, 1, \ldots, M-1$.

$$\alpha_u = \begin{cases} \sqrt{(\frac{1}{N})} & \text{for} u = 0 \\ \sqrt{(\frac{2}{N})} & \text{for} u > 0 \end{cases}$$

$$\alpha_v = \begin{cases} \sqrt{(\frac{1}{M})} & \text{for} v = 0 \\ \sqrt{(\frac{2}{M})} & \text{for} v > 0 \end{cases}$$

The highlight of DCT transformation is that it converts useful information into fewer coefficients with no decrease in recognition accuracy and increase in computational efficiency.
DCT transformation is applied to the WLD operated image wholly. For LBP, block division is applied to the image and DCT is applied for each block of LBP image.

3.1.5 Generate Final Feature Vector

DCT filters out the important information by image compression. The feature vectors generated by the zigzag scan of the image are selected on the basis high variance reducing the training time for classification as compared to the feature selection algorithms like GA (genetic algorithm) and PSO (particle swam optimization). Hence, the process proposed is computationally inexpensive.

3.2 Biohashing

Biohashing is an invertible feature transformation technique. Sensor collects the biometric characteristics of the user, the system then extracts the required biometric information in form of a template or feature vector. This template is then transformed using a random key given to the user by the system at the time of enrollment. A set of orthonormal vectors are created using this random key. The dot product or the inner product of the feature vector and the orthonormal vectors is calculated. Finally, a predefined threshold is applied to all the product outputs and a Biohash is calculated. Following are the steps for biometric hashing:

- Generate a set of random numbers based on the token (seed) provided by the user.
- Create a set of orthonormal vectors using Gram-Schmidt process applied on the set of random numbers.

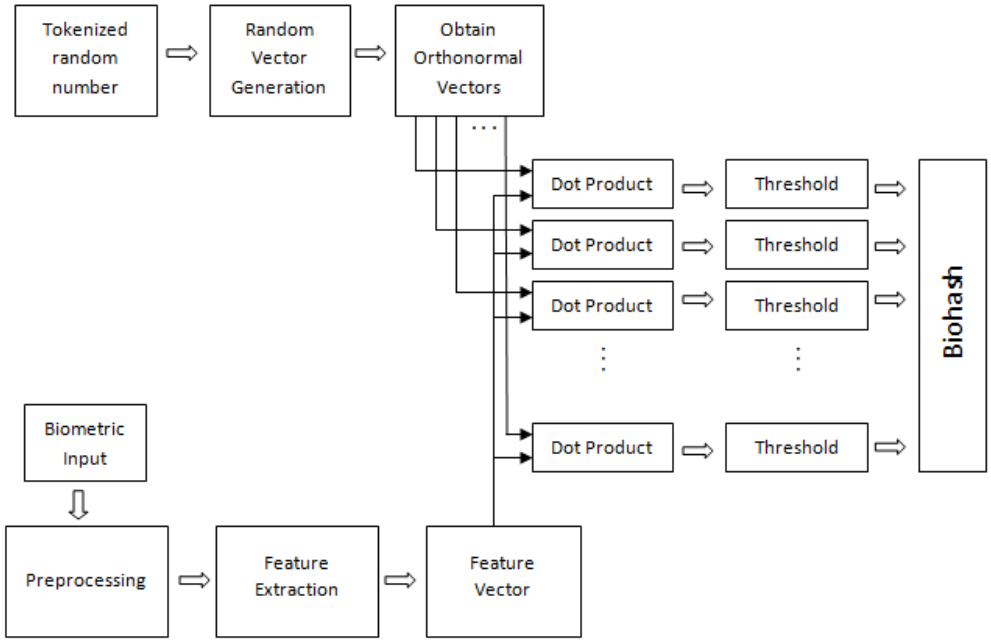

Figure 3.4: Steps to Biohashing

- Compute inner product of the feature vector and the orthonormal vectors.
- Apply a preset threshold to generate a string of 0ś and 1ś called the biometric hash or the Biohash.

3.2.1 Generate Random Numbers

Random numbers is a string of unpredictable numbers which are independent of predictable events but are based on natural random phenomenon. However, producing random numbers based on naturally occurring phenomenon is long, tiresome, expensive and non-repeatable. Thus, for daily applications pseudo random numbers are used instead of real random numbers so the results produced can also replicable.

Pseudo random Numbers and Their Properties

Pseudo means imitating or pretending, so pseudo random numbers are such sequences whose properties are near to the real random numbers. Pseudo random numbers are also called deterministic random numbers as the start

from an arbitrary value or initial value known as seed, the sequence of the random numbers generated can be reproduced with the knowledge of seed. Properties of pseudo random numbers are;

- Uncorrelated The sequence of numbers produced should be mutually independent and unrelated to each other i.e.; not predicting previous or future sequence.
- Very Long Period The generator should not repeat any sequence of random numbers but it not practically possible however repetition of the sequence should occur after producing long periods of numbers.
- Uniform The random numbers produced should be evenly distributed, uniform and unbiased on the given sample space i.e.; equal fractions of random numbers should be in equal fractions of sample space.
- Efficient The overhead of the pseudo random generator should be very low so it can be used in fast processing with multiple or parallel processing.

Methods to Produce Pseudo random Numbers

There are numerous methods to produce pseudorandom numbers which are known as pseudorandom generators (PRNGs) or deterministic random bit generators (DRBGs).
Four criteria for quality of deterministic random number generators are established by The German Federal Office for Information Security [131]. They are summarized as:

- The generated sequences of random numbers are different with high probability.
- A sequence of numbers pass specified statistical tests which are mono bit test, poker test, runs test, long-runs test and auto correlation test such that probability of next bit to be zero or one is one half.
- Any previous or future sequence or any inner state of generator is not practically possible for attacker to be calculated or guessed with any sub-sequence.
- Any previous sequence or inner state of generator is not practically possible for attacker to be calculated or guessed from an inner generator state.

Linear Congruential generators are the oldest pseudo random number generators based on discontinuous linear equations with inadequate quality of sequences produced. With the advancement of techniques based on linear recurrences related to linear feedback shift registers produced advanced PRNGs like Mersenne Twister and xor shift generator family.
Cryptographically secure pseudo random number generators (CSPRNG) require that an adversary without the knowledge of seed cannot determine the output sequence with the help of a random known sequence. Class of CSPRNGs include stream ciphers, block ciphers, Yarrow algorithm, Micalli-Schnorr generator, Blum Blum Shub etc.
In our proposed methodology, we will be using Blum Blum Shub algorithm.

Blum Blum Shub

Lenore Blum, Manuel Blum and Michael Shub [132] proposed a cryptographically secure pseudo random generator which efficiently produces long well distributed sequences with small seed values. The pseudo random generator is based on computationally hard quadratic residuary problem, given integers a and N, it is to decide if a is a quadratic modulo N or not [133].
Generate two large prime numbers p and q that are congruent to $3 mod 4$, $p \equiv q \equiv 3 mod 4$ called the Blum prime numbers, it ensures that each quadratic residue has a square root which is also a quadratic residue. Then N is a multiple of p and $q, N = p.q$. The seed s is chosen from the set $[1, N-1]$ such that p and q are not factors of s, i.e.; s is co-prime to N. Then,

$$x_0 = s^2 \text{mod N}$$

$$x_i = x_{i-1}^2 \text{mod N}$$

Output is generated at each step of the algorithm from x_i in form of even or odd parity or least significant bits of x_i.

$$z_i = \text{parity}(x_i)$$

The output sequence $z_1, z_2, z_3, \ldots \in 0, 1$.Without the knowledge of the initial p and q, it is computationally impossible to calculate or predict the future or previous values of the pseudo random sequence. Any value of x_i can be directly calculated with the knowledge of s or x_0, p and q from Carmichael function $\lambda(N) = \lambda(p.q) = \text{lcm}(p-1, q-1)$ using Eulers formula;

$$x_i = (x_0^{2i \text{mod} \lambda(N)}) \text{mod} N$$

3.2.2 Generate Orthonormal Vectors

The generated set of random numbers is converted into a set of orthonormal vectors. Orthonormal vectors are linearly independent, normalized and orthogonal in nature. The number of orthonormal vectors is equal to the number of feature vectors extracted. The set of random numbers is orthonormalized using Gram-Schmidt process.

Gram-Schmidt Process In 1907, Jorgen Pedersen Gram and Erhard Schmidt [134] proposed a method to orthonormalize a set of linearly independent vectors called Gram-Schmidt Process. In a set of linearly independent vectors $B = \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ all the vectors are linearly independent of each other i.e.; one vector cannot be written in the form of other vector,$\vec{v_i} \neq c\ \vec{v_j}\ \forall i \neq j$ for some constant c, then B is said to be linearly independent set.

Orthonormalize means orthogonal and normalized at the same time. A set of vectors $C = \vec{w_1}, \vec{w_2}, \ldots, \vec{w_k}$ is orthogonal if $\vec{w_i} \cdot \vec{w_j} = 0 \quad \forall i \neq j$ and $\vec{w_i} \cdot \vec{w_j} = 1 \quad \forall i = j$ If the lengths of all vectors in C are one,$\left\|\vec{w_i}\right\| = 1 \quad \text{for} i = 1, 2, \ldots, k$, they are called unit vectors or normalized. Vectors in C are orthogonal and normalized so, C is called orthonormal set.

Suppose $B = \vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ are basis for some subspace V, $V = \text{span}(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k})$, then orthonormal basis for V be $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_k}$ which can be obtained by Gram-Schmidt Process such that,

$$\text{span}(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}) = \text{span}(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_k})$$

Then,

$$\vec{u_1} = \frac{\vec{v_1}}{\left\|\vec{v_1}\right\|}$$

$$\vec{u_2} = \frac{\vec{v_2} - (\text{proj}_{v_1}\ \vec{v_2})}{\left\|\vec{v_2} - (\text{proj}_{v_1}\ \vec{v_2})\right\|}$$

$$\vec{u_3} = \frac{\vec{v_3} - (\text{proj}_{v_1}\ \vec{v_3} + \text{proj}_{v_2}\ \vec{v_3})}{\left\|\vec{v_3} - (\text{proj}_{v_1}\ \vec{v_3} + \text{proj}_{v_2}\ \vec{v_3})\right\|}$$

up to $\vec{u_k}$

$$\vec{u_k} = \frac{\vec{v_k} - \sum_{j=1}^{k-1}(\text{proj}_{v_j}\ \vec{v_k})}{\left\|\vec{v_k} - \sum_{j=1}^{k-1}(\text{proj}_{v_j}\ \vec{v_k})\right\|}$$

3.2.3 Calculate Inner Product

Inner product is calculated so the token based supplementary information can be added in the feature vector. The inner product of the feature vector extracted and the orthonormalized random vectors is calculated to get scalar terms equal to the number of feature vectors.

Inner Product

Inner product is the generalization of the dot product of Euclidean spaces to vector spaces of infinite dimensions which assigns each ordered pair of vectors a scalar quantity [135]. A function $\langle \cdot, \cdot \rangle : X \times X \to K$ that assigns a scalar $\langle x, y \rangle$ to each ordered pair (x, y) of vectors in X, where X is a linear space over the field K, such that it satisfies following axioms is called inner product [136].

- Positive definite

$$\langle x, x \rangle \geq 0 \qquad \forall x \in X$$

$$\text{and} \qquad \langle x, x \rangle = 0 \qquad \text{when} \qquad x = \vec{0}$$

- Conjugate symmetry

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \qquad \forall x, y \in X$$

- Linearity

$$\langle \langle ax, y \rangle = \langle x, y \rangle \qquad \forall x, y \in X \qquad \text{and} \qquad a \in K$$

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \qquad \forall x, y, z \in X$$

3.2.4 Applying Threshold

Threshold is the value or magnitude which creates a border line between two quantities. In our methodology threshold defines if the value is zero or one in the output, below threshold value, value is zero and above threshold value is one. Let $X = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$ be a matrix of order m then threshold is determined by calculating the mean M of all the values in the matrix $M = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}}{m}$. The values above Mare marked as one while values below Mare marked as zero.

3.2.5 Final Expression Hash

The output of the thresholding in the form of 0ś and1ś is called the expression hash. The expression hash is a combination of expression feature vector and the supplementary token information in the form of random number vectors. The transformed biometric template is stored in the database against its userś identity.

3.3 Biometric Authentication Process

Authentication process is the process of determining the claimed identity to be true or not. Authentication process consists of two phases; enrollment and authentication. During enrollment, the users provide their identity and some unique authenticator which is stored in the database. During authentication phase, the claimant provides its identity and the related authenticator which is matched with the stored ones. Decision is made if the claimant is true or not. Following sections explain the proposed authentication process.

3.3.1 Enrollment Phase

The proposed enrollment phase is divided into following steps. First, the user provides his username as fist name and last name, and its biometric trait in the form of its facial expression. Then, the modifier converts the username into user ID, a token number is generated on the basis of user ID. Both user ID and the token number are provided to the user and user ID is stored in the database.
The biometric template is extracted in the form of feature vector from the biometric trait obtained from the sensor using the expression feature extraction method mentioned above. The token number generated based on user ID and the extracted feature vector are then biohashed to generate expression hash. The generated expression hash is then stored in the database next to where the user ID is stored.

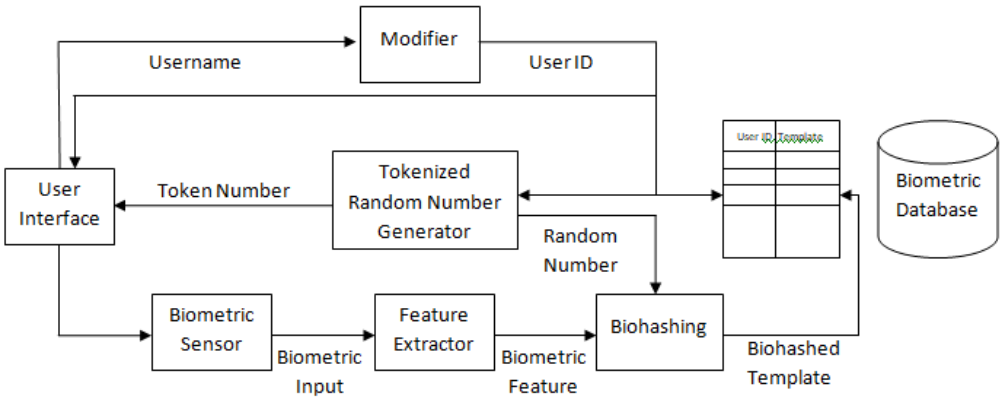

Figure 3.5: Enrollment Phase

3.3.2 Authentication Phase

In the proposed authentication phase these subsequent steps are followed. The user provides its user ID, token number and biometric trait. Expression feature vector is extracted from the obtained biometric trait. The feature vector is then transformed to expression hash using the provided token number.

The matcher compares the expression hash stored against the given user ID and the newly created expression hash from the give biometric instance and outputs a match score. In the decision module, it is checked if the match score is above the predefined threshold or not. If the match score is above the threshold, then access is granted else it is denied.

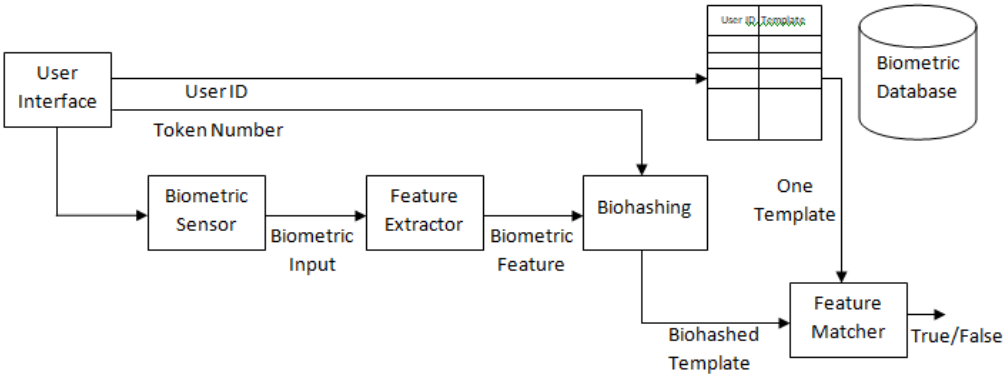

Figure 3.6: Authentication Phase

Chapter 4

Experiments and Results

Access is the prospect of admittance to a closed or private data and should only be granted to the legitimate user. Any other, an imposter or attacker should not be allowed to get the access to secret information. Access is given to a user through an authentication process where the user provides its identifier given by the system at the time of enrollment and the authenticators which verify if the claimant is legitimate or not. Here, identifier is User ID and authenticators are token number, persons face and its expression.

4.1 Database

the database used for experimentation is JAFFE which contains images of ten female Japanese female models. it contain total of 213 images of 7 facial expressions that are happy, sad, surprise, angry, disgust, fear and neutral [137].

4.2 Proposed Biometric System Parameters

Following are the parameters of the proposed biometric authentication system.

4.2.1 Identifier

A unique identifier User ID is extracted from the userś full name consisting of five alphabets and one digit. There are twenty six alphabets in English language, so for one alphabet place there are twenty six possibilities. For five alphabet places there will be $26^5 = 11881376$ possibilities. Digit place can be filled with digits from $0 - 9$, so there will be $10^1 = 10$possibilities. Total

number of possibilities of valid User IDs is $26^50 = 118813760$. User ID is stored in the database as plain text.

4.2.2 Token Number

Token number is given to the user at the time of enrollment chosen from a pre-decided set of 4-digit numbers based on the User ID. One digit place has 10 possibilities from 0-9, so 4-digit number will have $10^4 = 10000$ possibilities. Token number assigned to the user is never stored in the database but the expression hash calculated contains the hidden information related to token number.

4.2.3 Personś Identity and Expressions

Feature Vector of 128 elements is extracted from face biometrics using local binary patterns and discrete cosine transform containing information about the identity of the person and the expression as well. There are seven expressions used; happy, sad, surprise, angry, disgust, fear and neutral. So, for one person number of possible feature vectors is 7. Feature vector of the same person is different for each expression.
Expression hash is calculated using the token number and the feature vector extracted. Expression hash is different for the same person when token number or the expression is different. So, total number of possible expression hashes is $10^4 = 70000$.

4.3 Performance Measures

A biometric verification system is basically a pattern recognition system which attempts to recognize similar patterns from the data set of an entity and dissimilar information from the data sets of different entities. A good recognition system will have small intra-class variation and large inter-class variation. The extracted information called the feature vectors are then transformed using biohashing into expression hashes.

4.3.1 Match Score Distributions

Performance of the system depends on the quality and method of the feature vector produced and the matching scheme. Accuracy of the system can be measured by the matching score of the stored expression hash and the given expression hash and a predefined threshold level. Matching score or similarity

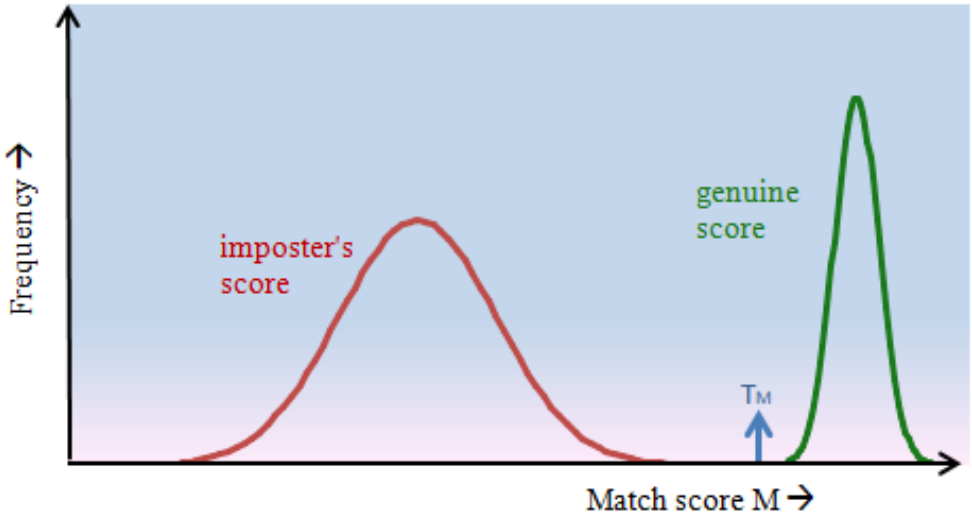

Figure 4.1: Ideal Behavior of Biometric Verification System

score should be above the threshold level for legitimate user and below for imposters.
Let M be the match score and T_M be the pre-defined threshold level then decision for a query for a claimed identity I with expression hash E_q will be;

$$(I, E_q) \in \begin{cases} \text{granted}, & \text{if} M(E_q, E_I) \geq T_M \\ \text{denied}, & \text{if} M(E_q, E_I) < T_M \end{cases}$$

Ideally, the match score of the genuine users is above the threshold level T_M and when an imposter tries to verify itself its score comes out to be less than T_M.
Genuine match score distribution $P_m(M)$ for a set $A = \{A_1, A_2, \ldots, A_X\}$ of X genuine match scores is calculated as;

$$P_m(M) = \frac{1}{X} \sum_{i=1}^{X} 1(A_i = M) = \frac{1}{X}(A_i = M), \forall M$$

Imposter match score distribution $P_n(M)$ for a set $B = B_1, B_2, \ldots, B_Y$ of Y genuine match scores is calculated as;

$$P_n(M) = \frac{1}{Y} \sum_{i=1}^{Y} 1(B_i = M) = \frac{1}{Y}(B_i = M), \forall M$$

par

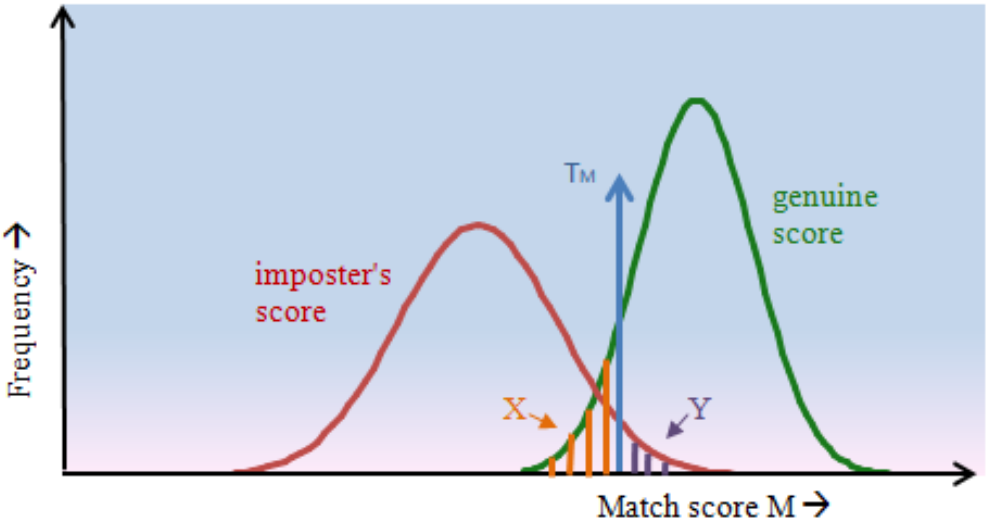

Figure 4.2: Non-Ideal Behavior of Biometric Verification System

4.3.2 Error Rates

But in real-time scenarios, there are instances when the system shows errors. There are two types of errors; false rejection rate (FRR) i.e.; genuine users are rejected falsely and false acceptance rate (FAR) i.e.; imposters are accepted falsely.

False Rejection Rate

When a genuine user is rejected by the system falsely because its match score M is less than threshold level T_M, S_M, it is called false rejection rate (FRR). It is shown by the orange-lined region X in the figure below. The area X under the genuine score frequency curve when S_M, is expressed as function of T_M as;

$$FRR(T_M) = \int_{-\infty}^{T_M} P_m(M)dM$$

False rejection rate FRR for a set $A = A_1, A_2, \ldots, A_X$ of X genuine match scores is calculated as;

$$FRR(T_M) = \frac{1}{X}\sum_{i=1}^{X} 1(A_i \leq T_M) = \frac{1}{X}(A_i \leq T_M)$$

False Acceptance Rate

When an imposter is falsely accepted by the system because its match score M is greater than threshold level T_M, it is called false acceptance rate (FAR). It is shown by the purple-lined region Y in the figure below. The area Y under the genuine score frequency curve when $S > T_M$, is expressed as function of T_M as;

$$FAR(T_M) = \int_{T_M}^{\infty} P_n(M)dM$$

False acceptance rate FAR for a set $B = B_1, B_2, \ldots, B_Y$ of Y imposter match scores is calculated as;

$$FAR(T_M) = \frac{1}{Y}\sum_{i=1} Y1(B_i > T_M) = \frac{1}{Y}(B_i > T_M)$$

Equal Error Rate

When FRR decreases FAR increases and vice versa. However, both the error rates are equal at one point called equal error rate ERR. At this point, FAR=FRR. FAR and FRR are calculated for all points of threshold as shown in the graph below. ERR is the point of intersection of the two curves. The smaller the value of ERR, the better the system. The corresponding value of threshold to ERR is called the optimum value of threshold T_O. At this threshold both FAR and FRR will be equal.

Receiver Operating Characteristic Curve

Another performance indicator of biometric verification system is receiver operating characteristic curve ROC. When FRR is plotted against FAR the resultant curve is called ROC. The closer the curve is to the origin, the better the system.

4.4 Experiments and Results

Experiments are designed to replicate all the possible real life scenarios which can occur. We have one identifier and three authenticators. There could be total of 16 possible scenarios. We divide them into two cases that are when entered User ID is correct and when User ID is incorrect.
Following are the different scenarios that can occur in the proposed biometric authentication system.

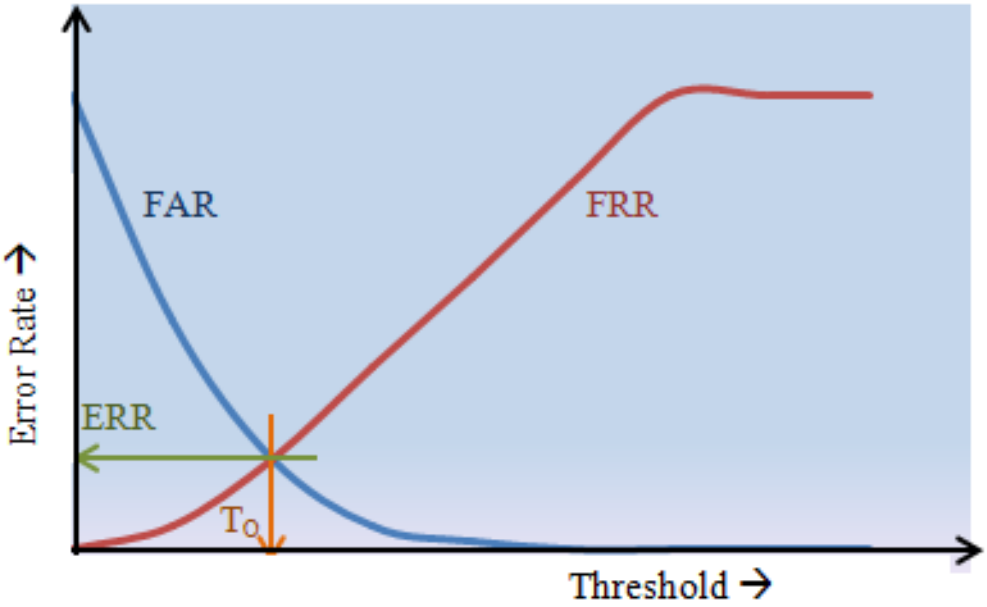

Figure 4.3: FAR and FRR at All Values of Threshold, ERR and T_O.

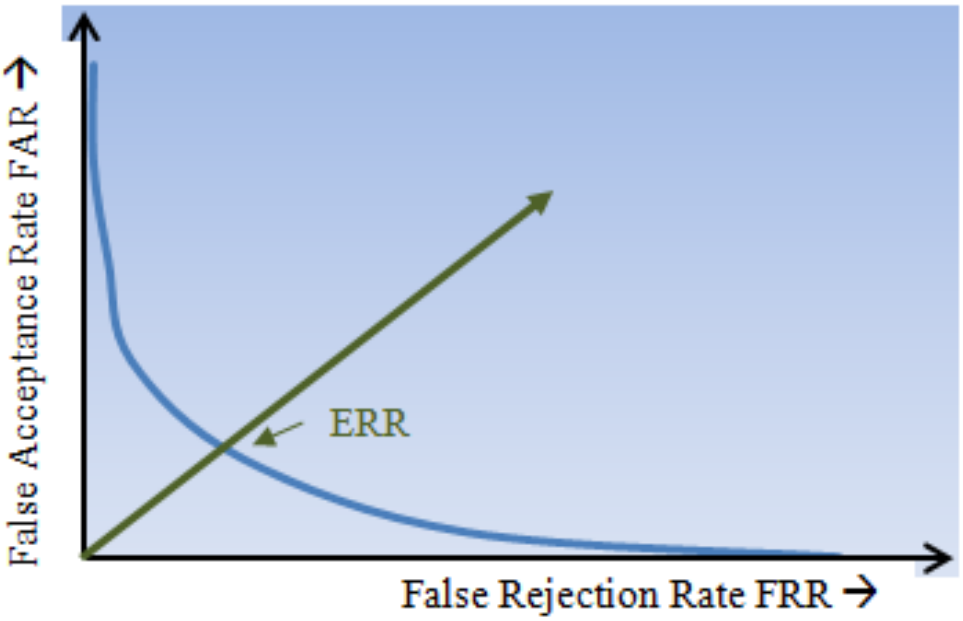

Figure 4.4: Receiver Operating Characteristic Curve ROC.

Table 4.1: Different Scenarios of Expression Hash Authentication System

scenario	Token no.	Identity	Expression	Access
1.	✔	✔	✔	Granted
2.	✔	✔	✖	Denied
3.	✔	✖	✔	Denied
4.	✔	✖	✖	Denied
5.	✖	✔	✔	Denied
6.	✖	✔	✖	Denied
7.	✖	✖	✔	Denied
8.	✖	✖	✖	Denied

4.4.1 Token Is Correct, Persons Identity Is Correct, Expression Is Correct

In this scenario, a genuine user provides his user ID, assigned token number, its face with the expression it enrolled with. This scenario is considered as the legitimate attempt to access the system. A curve of frequency distribution $p_m(M)$ is obtained instead of Dirac delta function because of intra-user variability. In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. So, total pairs to be tested are $1 \times 3 \times 7 \times 10 = 210$. Mean of the match score values is 116.822 with a standard deviation of 9.805. 95.556% of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely rejecting the genuine user will be FRR. At the pre-set threshold T_M, error rate percentage is 4.444%. Error rate increases as the threshold values increase.

4.4.2 Token Is Correct, Personś Identity Is Correct, Expression Is Incorrect

In this scenario, a genuine user provides his user ID, assigned token number, its face with the wrong expression. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n2}(M)$ is obtained is shown in the figure below.
In our database, there are 10 users with 7 different expressions. With one

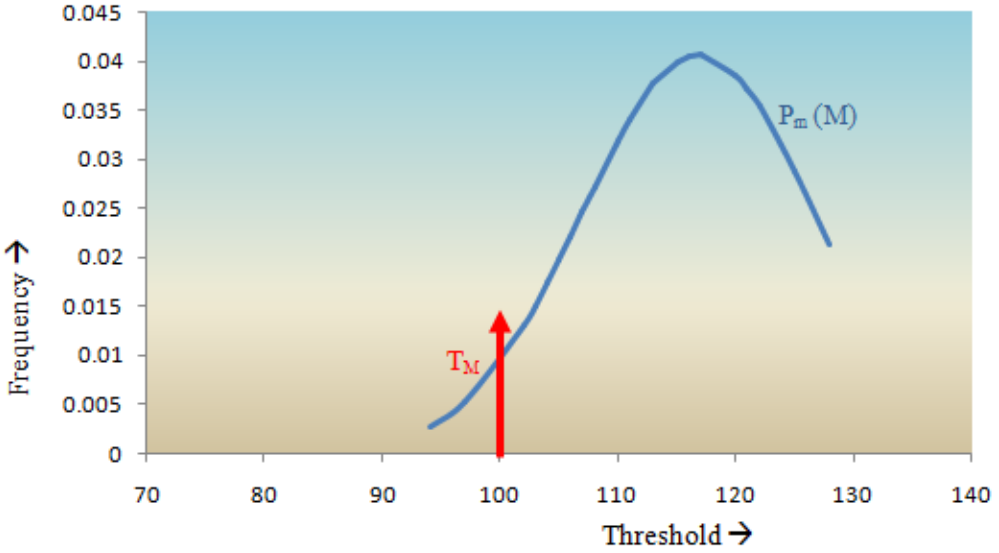

Figure 4.5: Frequency Distribution Curve When Token Is Correct, Persons Identity Is Correct, Expression Is Correct

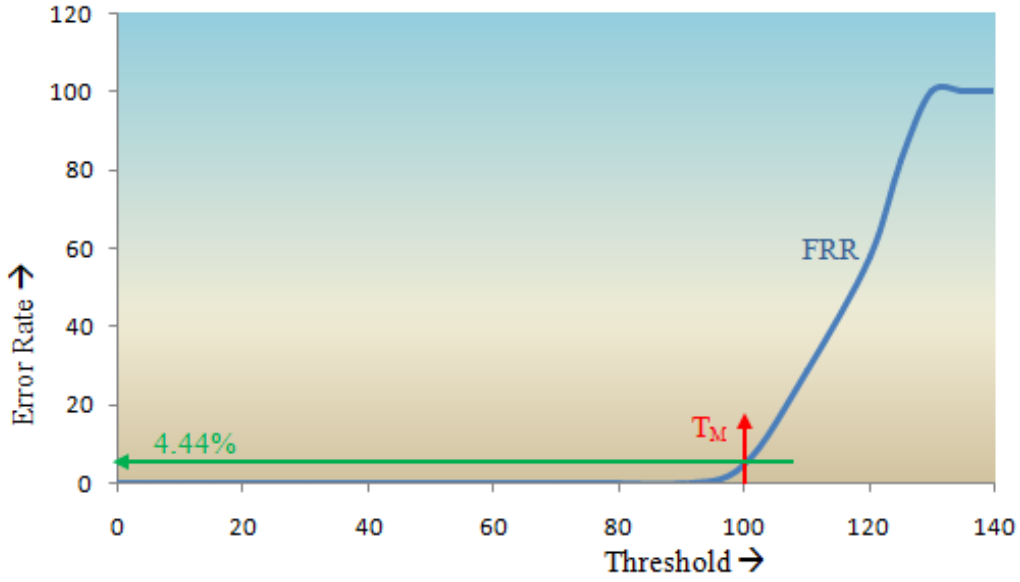

Figure 4.6: FRR When Token Is Correct, Personś Identity Is Correct, Expression Is Correct

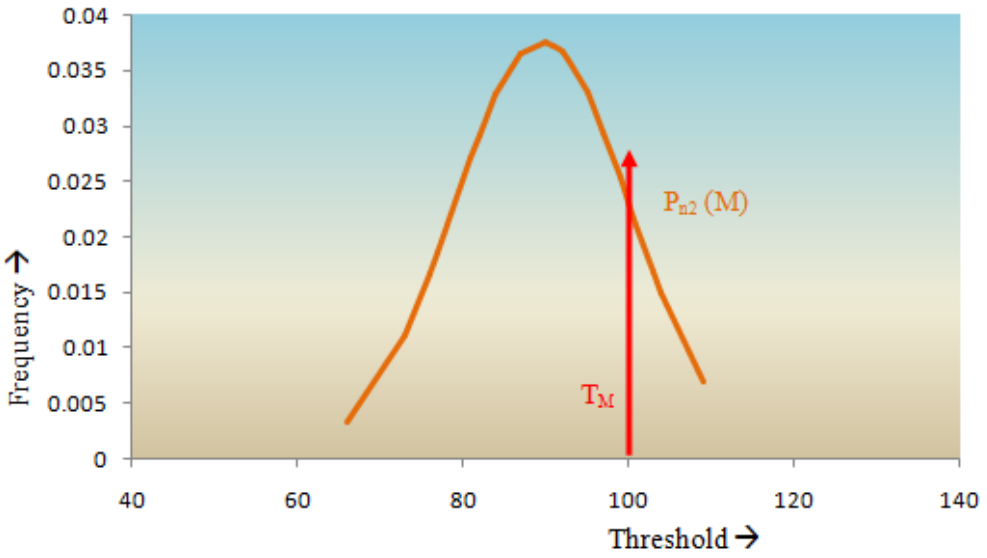

Figure 4.7: Frequency Distribution Curve When Token Is Correct, Personś Identity Is Correct, Expression Is Incorrect

expression, there are 3 images of a user. One expression of a user is tested against its 6 other expressions. So, total pairs to be tested are $1 \times 3 \times 6 \times 7 \times 10 = 1260$. Mean of the match score values is 89.595 with a standard deviation of 10.611. Only 19.047% of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 19.047%. Error rate is decreasing as the threshold values increase.
Equal error rate ERR for scenario 1 and 2 is almost 8% at optimum threshold value of T_O at the intersection of two curves FAR and FRR.
Receiver operating characteristic curve ROC for scenario 1 and 2 is shown in the figure below. The curve is very close to the origin indicating that the verification system is very good considering scenario 1 and 2.

4.4.3 Token Is Correct, Personś Identity Is Incorrect, Expression Is Correct

In this scenario, an imposter provides correct user ID, correct token number, wrong face identity but with correct expression. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n3}(M)$ is obtained is shown in the figure below.

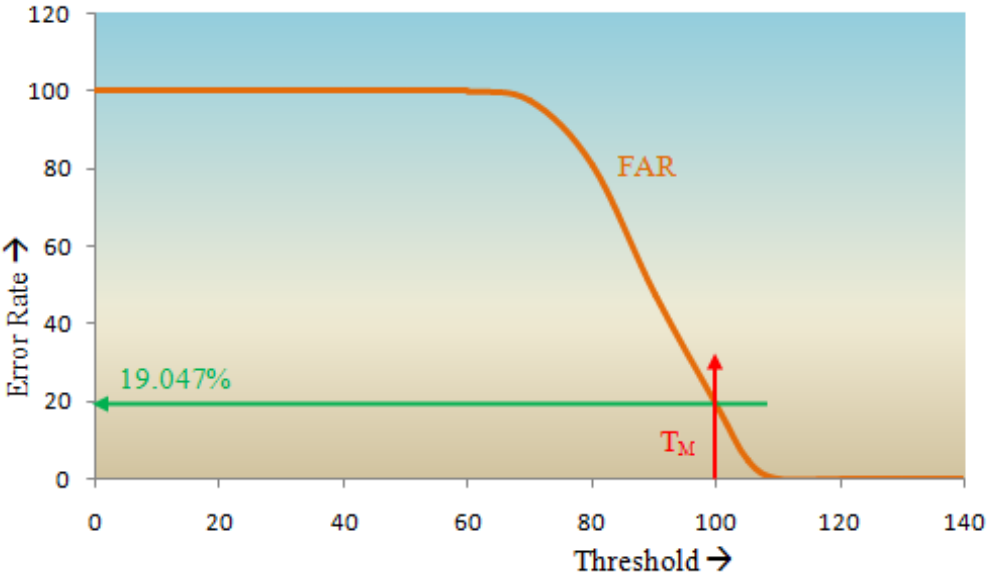

Figure 4.8: FAR When Token Is Correct, Personś Identity Is Correct, Expression Is Incorrect

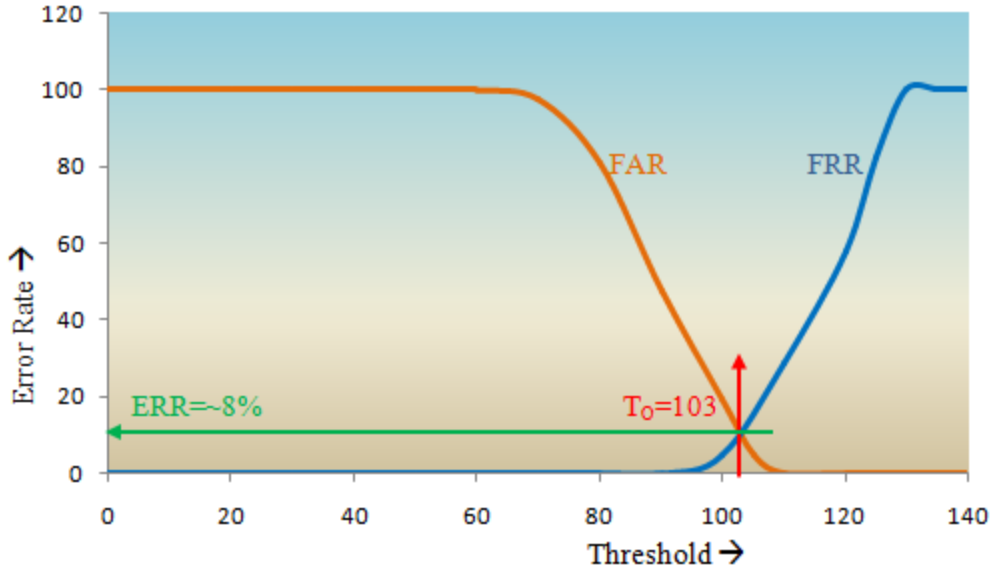

Figure 4.9: EER for Scenario 1 and Scenario 2

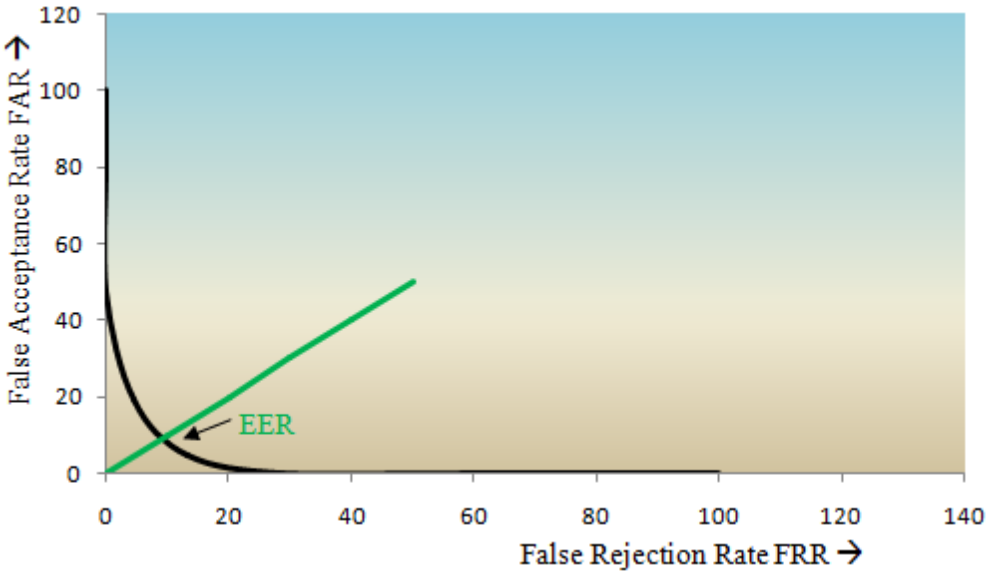

Figure 4.10: ROC Curve for Scenario 1 and 2

In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested against 9 other users with the same expression. So, total pairs to be tested are $1 \times 3 \times 9 \times 7 \times 10 = 1890$. Mean of the match score values is 89.667 with a standard deviation of 8.399. Only 6.349% of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 6.349%. Error rate is decreasing as the threshold values increase.
Equal error rate ERR for scenario 1 and 3 is almost 5% at optimum threshold value of T_O at the intersection of two curves FAR and FRR.
Receiver operating characteristic curve ROC for scenario 1 and 3 is shown in the figure below. The curve is very close to the origin indicating that the verification system is very good considering scenario 1 and 3.

4.4.4 Token Is Correct, Personś Identity Is Incorrect, Expression Is Incorrect

In this scenario, an imposter provides correct user ID, correct token number, wrong face identity but with wrong expression. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n4}(M)$ is obtained is shown in the figure below.

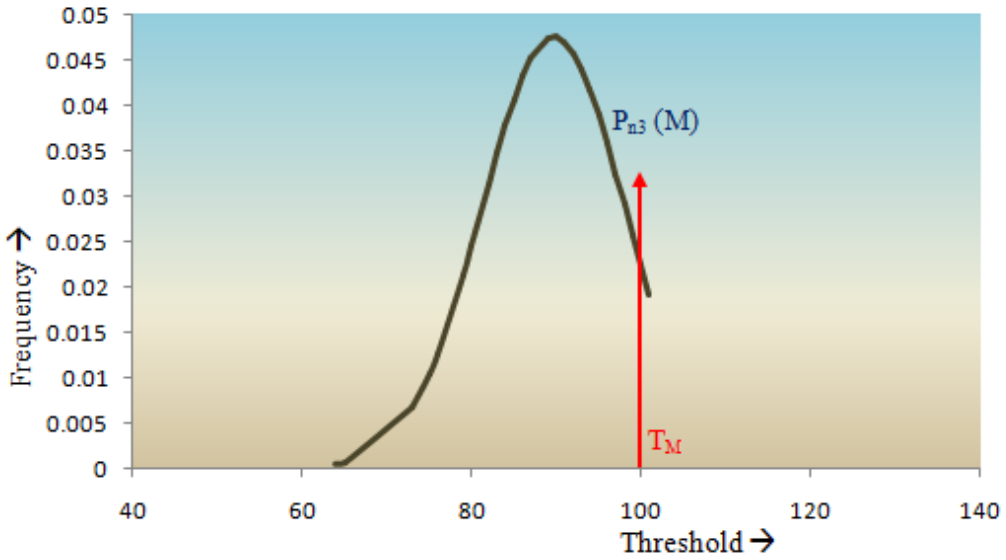

Figure 4.11: Frequency Distribution Curve When Token Is Correct, Personś Identity Is Incorrect, Expression Is Correct

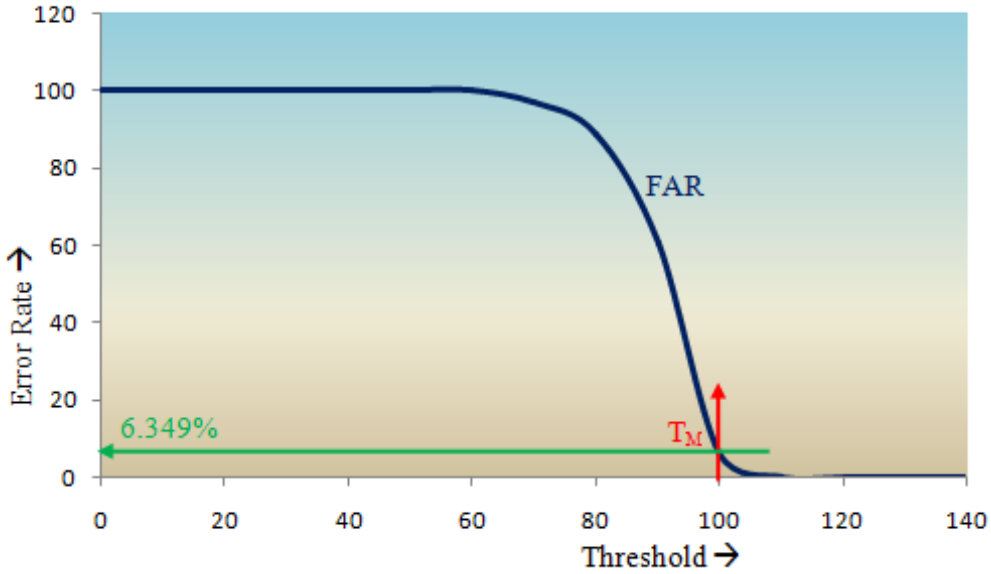

Figure 4.12: When Token Is Correct, Personś Identity Is Incorrect, Expression Is Correct

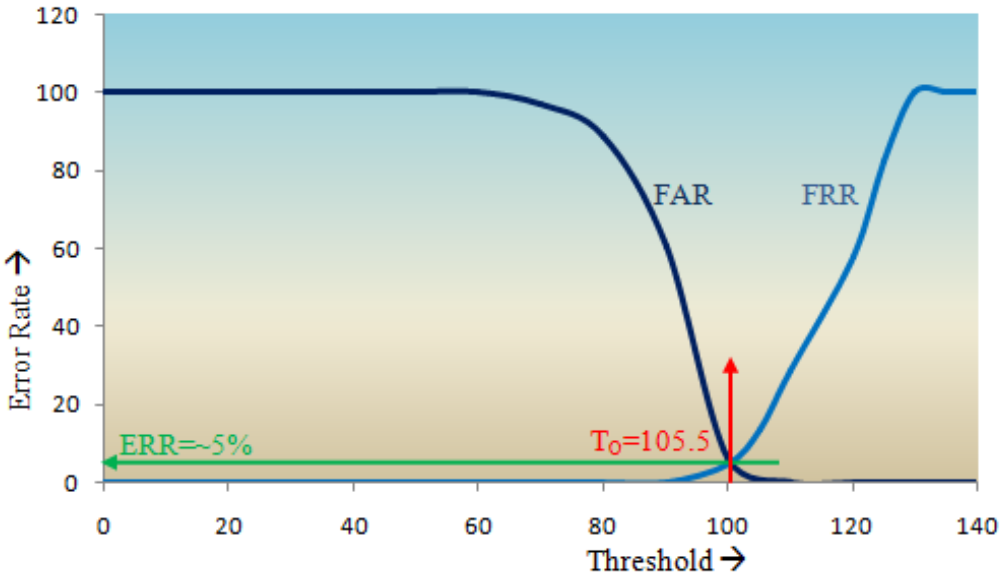

Figure 4.13: EER for Scenario 1 and Scenario 3

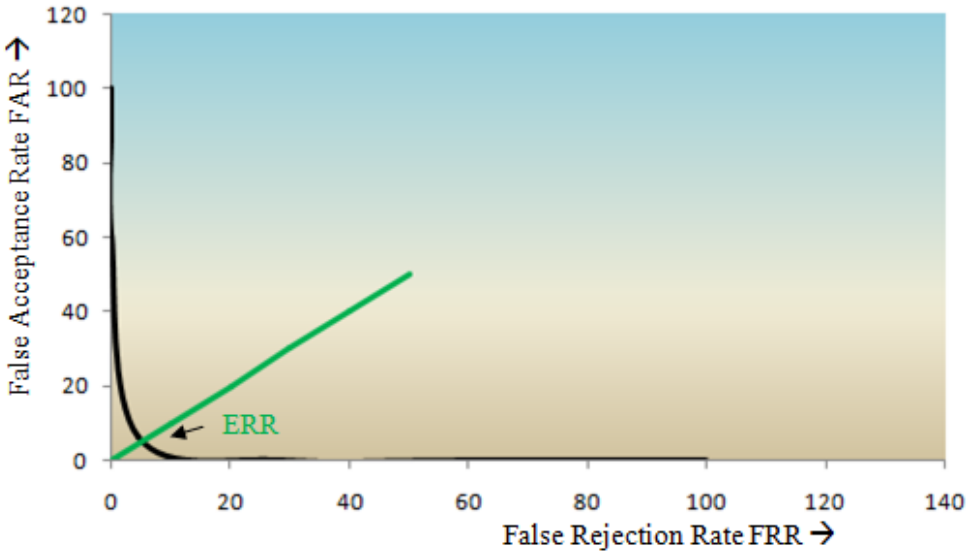

Figure 4.14: ROC Curve for scenario 1 and 3

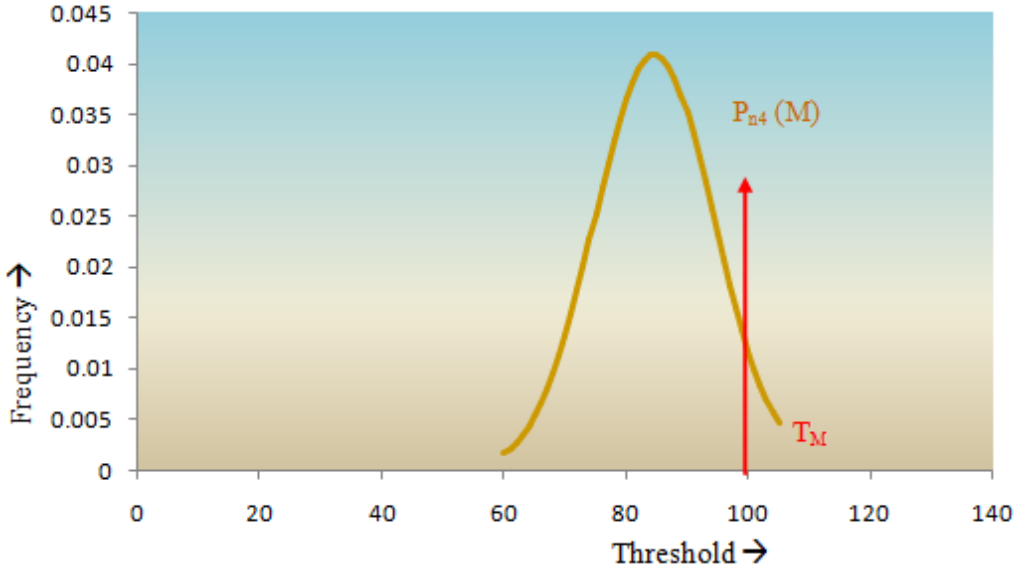

Figure 4.15: Frequency Distribution Curve When Token Is Correct, Personś Identity Is Incorrect, Expression Is Incorrect

In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested against 9 other users with the 6 other expressions. So, total pairs to be tested are $1 \times 3 \times 9 \times 6 \times 7 \times 10 = 11340$. Mean of the match score values is 84.622 with a standard deviation of 9.745. Only 4.762% of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 4.762%. Error rate is decreasing as the threshold values increase.
Equal error rate ERR for scenario 1 and 4 is almost 4.7% at optimum threshold value of T_O at the intersection of two curves FAR and FRR.
Receiver operating characteristic curve ROC for scenario 1 and 4 is shown in the figure below. The curve is too close to the origin signifying that the verification system is very good considering scenario 1 and 4.

4.4.5 Token Is Incorrect, Persons Identity Is Correct, Expression Is Correct

In this scenario, an imposter provides correct user ID, correct face identity, correct expression but token number is incorrect/unknown. This scenario is considered as the illegitimate attempt to access the system. A curve of

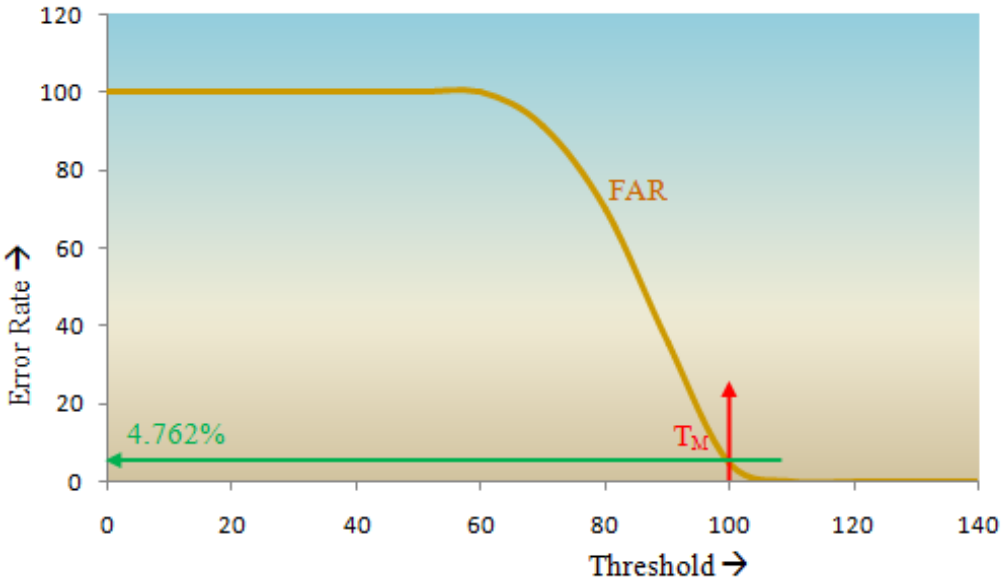

Figure 4.16: FAR When Token Is Correct, Personś Identity Is Incorrect, Expression Is Incorrect

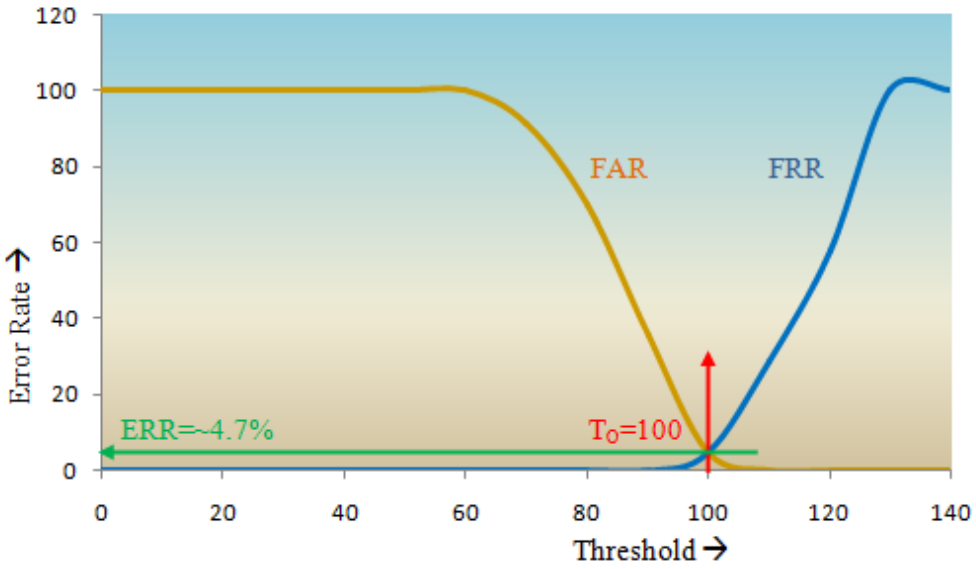

Figure 4.17: EER for Scenario 1 and Scenario 4

Figure 4.18: ROC Curve for Scenario 1 and 4

frequency distribution $p_{n5}(M)$ is obtained is shown in the figure below.
In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested for $10^4 = 1000$ tokens. At least, half of the tokens are to be tested by the attacker to get the correct token number (Brute-Force attack). So, total pairs to be tested are $1 \times 3 \times 7 \times 10 \times 1000/2 = 105,000$. Few of the pairs are tested whose mean of the match score values is 59.826 with a standard deviation of 5.193. None of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 0%. Error rate is zero for threshold values greater than 70.
Equal error rate ERR for scenario 1 and 5 is 0% for threshold values between 70 and 93. Optimum threshold value of T_O can be any between 70 and 93.
Receiver operating characteristic curve ROC for scenario 1 and 5 is shown in the figure below. The curve coincides with the axis showing that it is an ideal case. No legitimate user will be falsely rejected and no attacker can access the system when he has a wrong token number will all the rest true credentials.

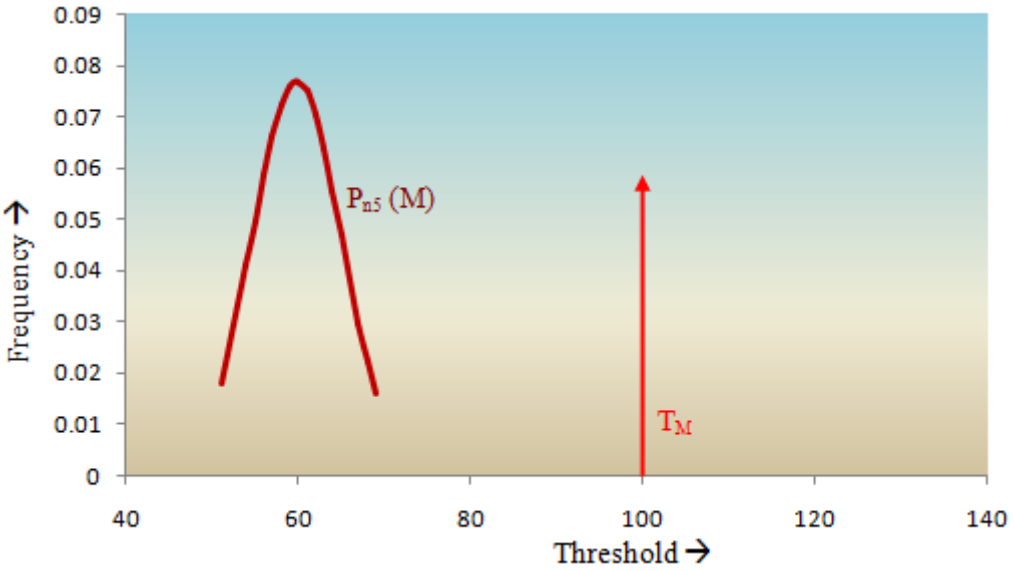

Figure 4.19: Frequency Distribution Curve When Token Is Incorrect, Personś Identity Is Correct, Expression Is Correct

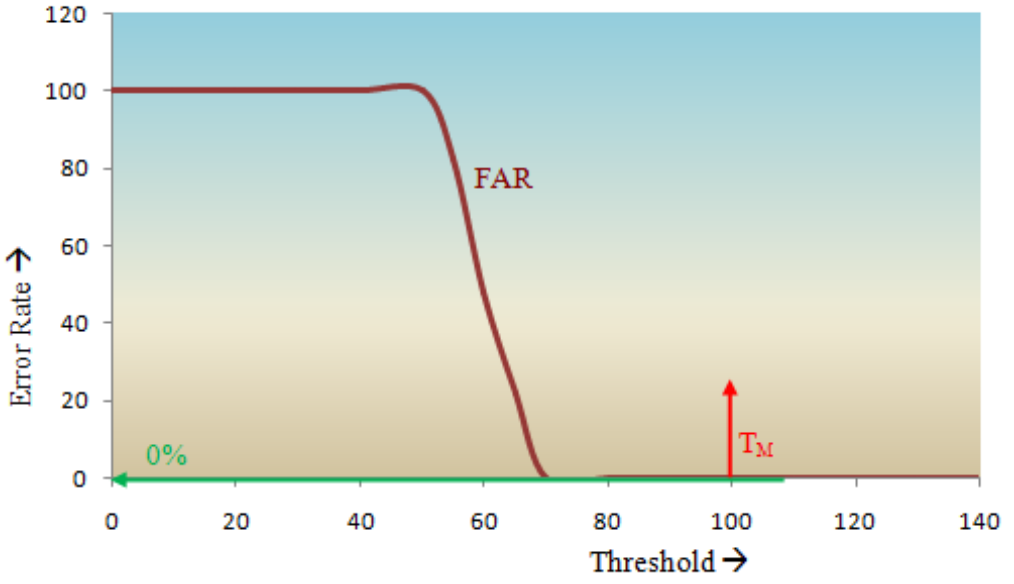

Figure 4.20: FAR When Token Is Incorrect, Personś Identity Is Correct, Expression Is Correct

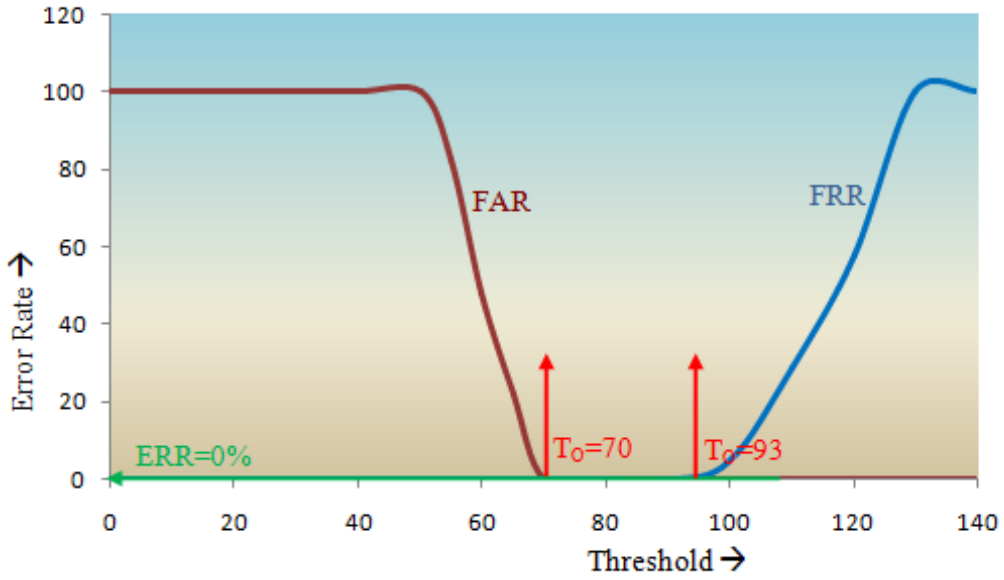

Figure 4.21: EER for Scenario 1 And Scenario 5

Figure 4.22: ROC Curve for Scenario 1 And 5

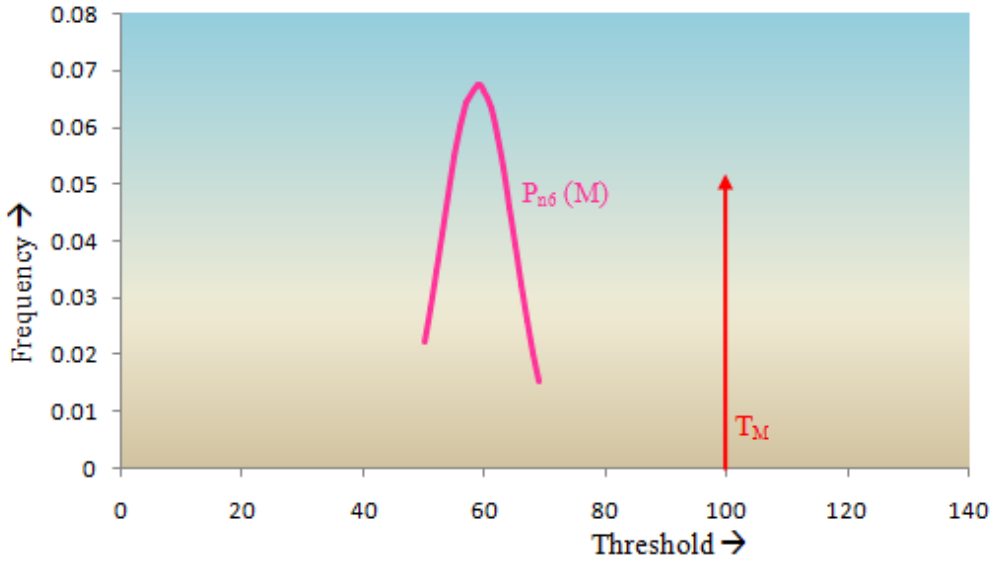

Figure 4.23: Frequency Distribution Curve When Token Is Incorrect, Personś Identity Is Correct, Expression Is Incorrect

4.4.6 Token Is Incorrect, Persons Identity Is Correct, Expression Is Incorrect

In this scenario, an imposter provides correct user ID, correct face identity but incorrect expression and the token number is incorrect/unknown. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n6}(M)$ is obtained is shown in the figure below.

In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested for its six different expressions with $10^4 = 1000$ tokens. At least, half of the tokens are to be tested by the attacker to get the correct token number (Brute-Force attack). So, total pairs to be tested are $1 \times 3 \times 6 \times 7 \times 10 \times 1000/2 = 630,000$. Few of the pairs are tested whose mean of the match score values is 58.829 with a standard deviation of 5.909. None of the match scores are above the pre-set threshold value T_M of 100.

In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 0%. Error rate is zero for threshold values greater than 70.

Equal error rate ERR for scenario 1 and 6 is 0% for threshold values between 70 and 93. Optimum threshold value of T_O can be any between 70 and 93.

Receiver operating characteristic curve ROC for scenario 1 and 6 is shown

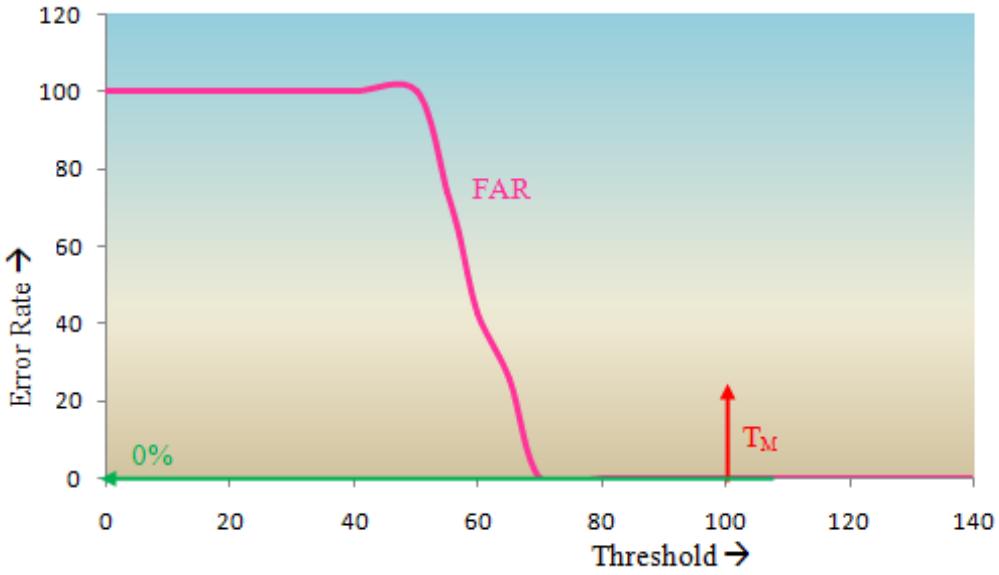

Figure 4.24: FAR When Token Is Incorrect, Personś Identity Is Correct, Expression Is Incorrect

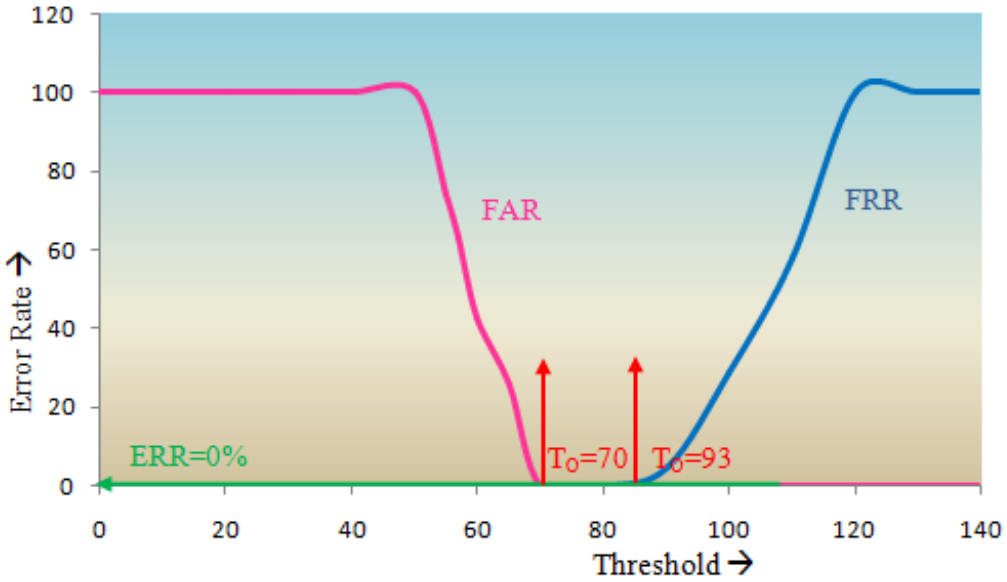

Figure 4.25: EER for Scenario 1 and Scenario 6

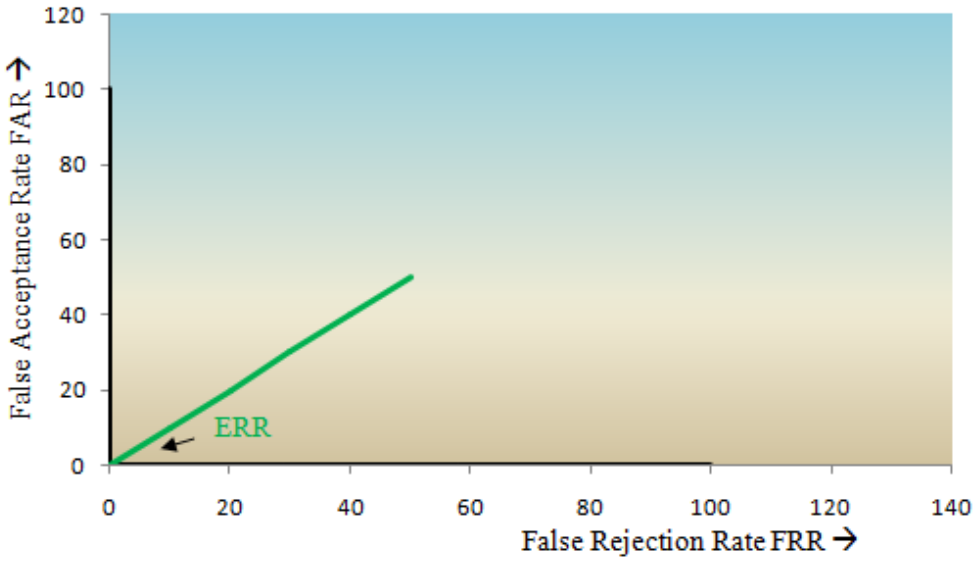

Figure 4.26: ROC curve for scenario 1 and 6

in the figure below. The curve coincides with the axis showing that it is an ideal case. No legitimate user will be falsely rejected and no attacker can access the system when token number and expression are unknown.

4.4.7 Token Is Incorrect, Personś Identity Is Incorrect, Expression Is Correct

In this scenario, an imposter provides correct user ID, correct face identity but incorrect expression and the token number is incorrect/unknown. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n7}(M)$ is obtained is shown in the figure below.

In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested for 9 other users with the same expression with $10^4 = 1000$ tokens. At least, half of the tokens are to be tested by the attacker to get the correct token number (Brute-Force attack). So, total pairs to be tested are $1 \times 3 \times 9 \times 7 \times 10 \times 1000/2 = 945,000$. Few of the pairs are tested whose mean of the match score values is 57.125 with a standard deviation of 5.488. None of the match scores are above the pre-set threshold value T_M of 100.

In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 0%. Error rate is zero for threshold values

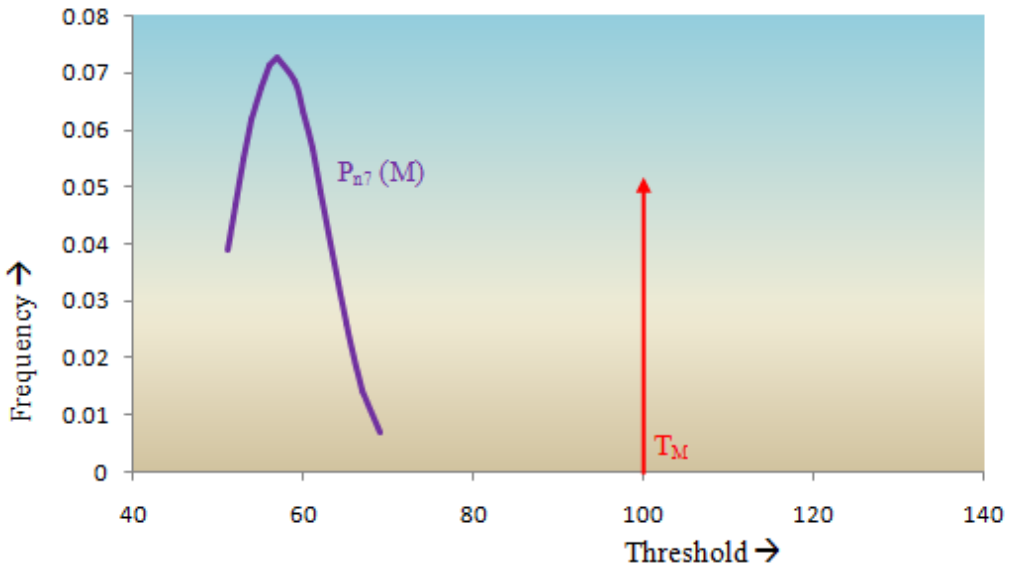

Figure 4.27: Frequency Distribution Curve When Token Is Incorrect, Personś Identity Is Incorrect, And Expression Is Incorrect

greater than 70.
Equal error rate ERR for scenario 1 and 7 is 0% for threshold values between 70 and 93. Optimum threshold value of T_O can be any between 70 and 93. Receiver operating characteristic curve ROC for scenario 1 and 7 is shown in the figure below. The curve coincides with the axis showing that it is an ideal case. No legitimate user will be falsely rejected and no attacker can access the system.

4.4.8 Token Is Incorrect, Personś Identity Is Incorrect, Expression Is Incorrect

In this scenario, an imposter provides correct user ID but incorrect face identity, incorrect expression and the token number is incorrect/unknown. This scenario is considered as the illegitimate attempt to access the system. A curve of frequency distribution $p_{n8}(M)$ is obtained is shown in the figure below.
In our database, there are 10 users with 7 different expressions. With one expression, there are 3 images of a user. One user with one expression is tested for 9 other users with the 6 different expressions with $10^4 = 1000$ tokens. At least, half of the tokens are to be tested by the attacker to get the correct token number (Brute-Force attack). So, total pairs to be tested

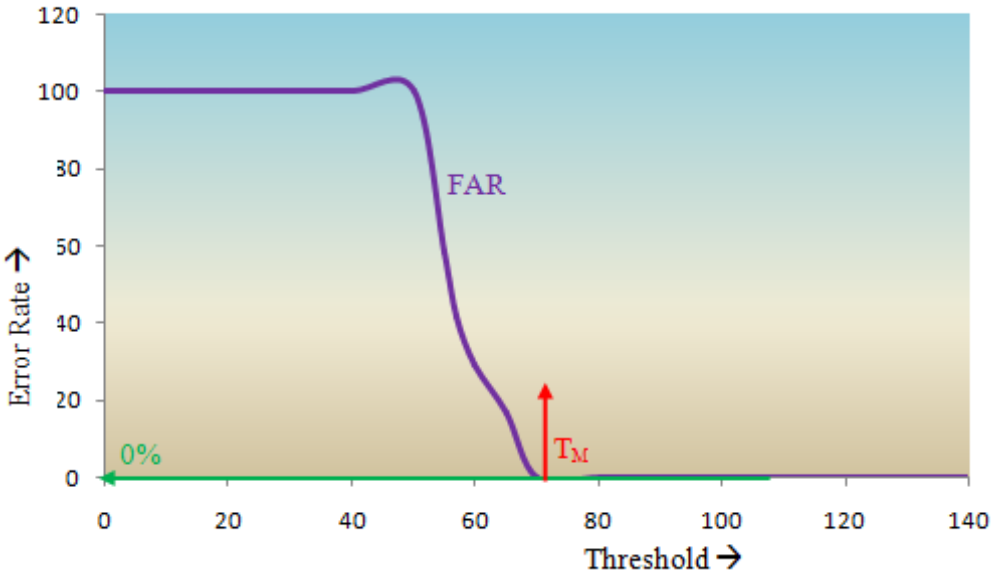

Figure 4.28: FAR When Token Is Incorrect, Personś Identity Is Incorrect, Expression Is Correct

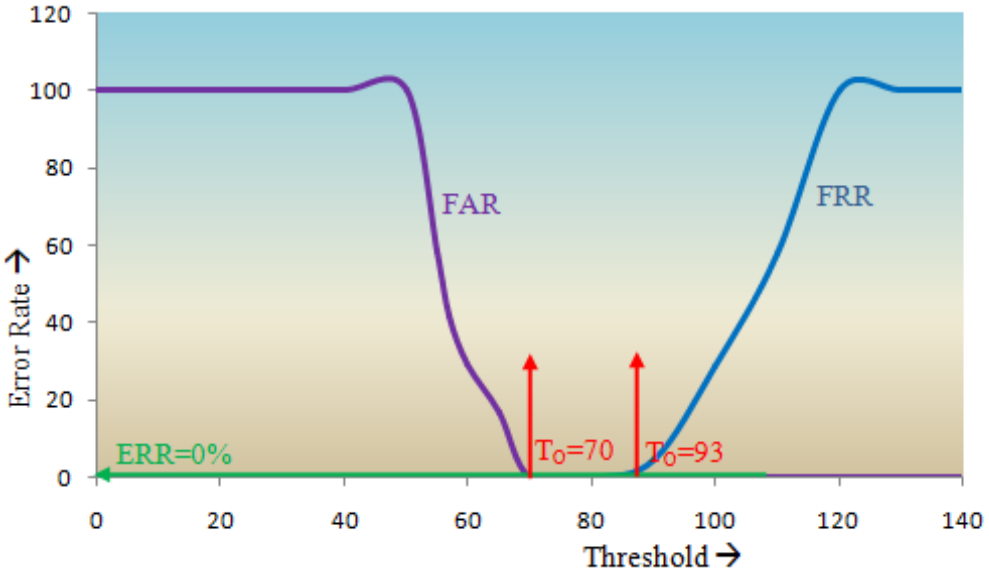

Figure 4.29: EER for Scenario 1 and Scenario 7

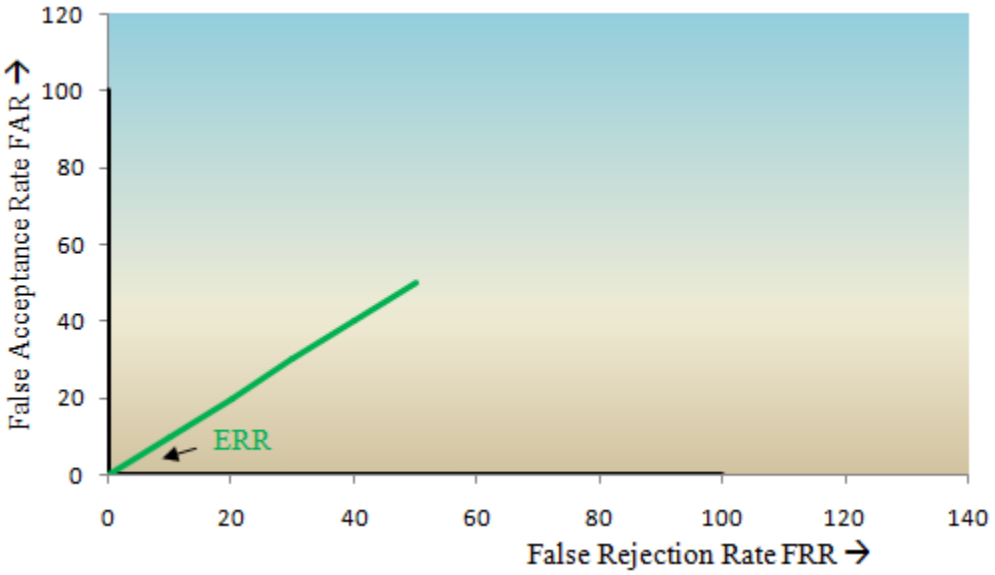

Figure 4.30: ROC Curve for Scenario 1 and 7

are $1 \times 3 \times 9 \times 6 \times 7 \times 10 \times 1000/2 = 5,670,000$. Few of the pairs are tested whose mean of the match score values is 58.619 with a standard deviation of 5.518. None of the match scores are above the pre-set threshold value T_M of 100.
In this scenario, falsely accepting the user will be FAR. At the pre-set threshold T_M, error rate percentage is 0%. Error rate is zero for threshold values greater than 68.
Equal error rate ERR for scenario 1 and 8 is 0% for threshold values between 68 and 93. Optimum threshold value of T_O can be any between 68 and 93.
Receiver operating characteristic curve ROC for scenario 1 and 8 is shown in the figure below. The curve coincides with the axis showing that it is an ideal case. No legitimate user will be falsely rejected and no attacker can access the system.

4.4.9 Cumulative Performance of The Biometric Authentication System

The scenario 1 is only the legitimate case where a genuine user is accessing the system rest of the scenarios from 2 to 8 is illegitimate where attacker is trying to access the system. The cumulative mean of the illegitimate attempts match scores is 80.949 with a standard deviation of 13.860. Only 4.45% of the match scores are above the pre-set threshold level T_M.

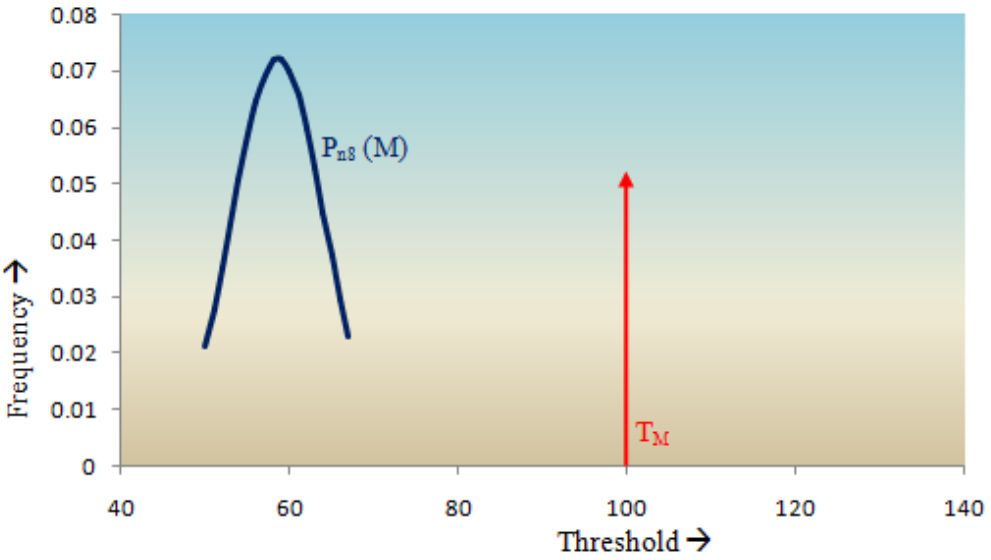

Figure 4.31: Frequency Distribution Curve When Token Is Incorrect, Personś Identity Is Incorrect, And Expression Is Incorrect

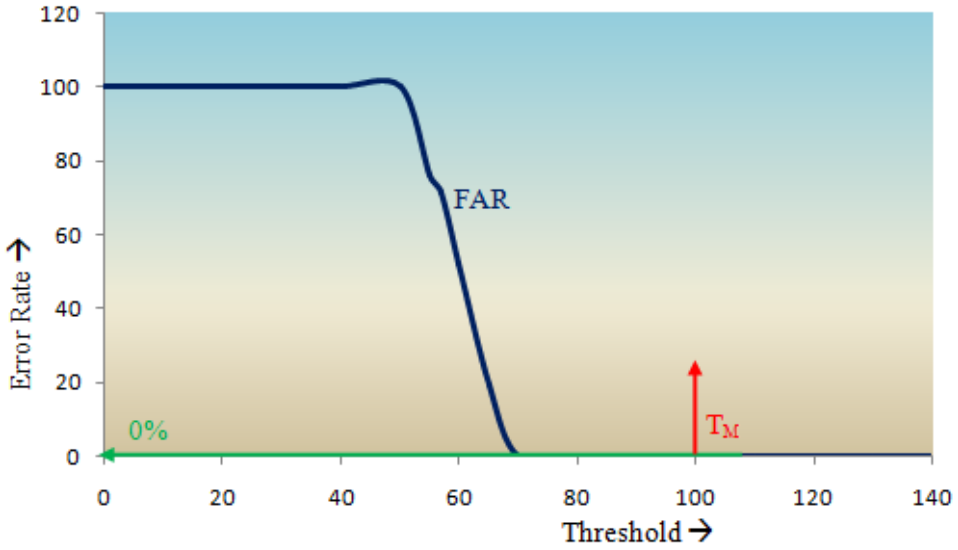

Figure 4.32: FAR When Token Is Incorrect, Personś Identity Is Incorrect, Expression Is Incorrect

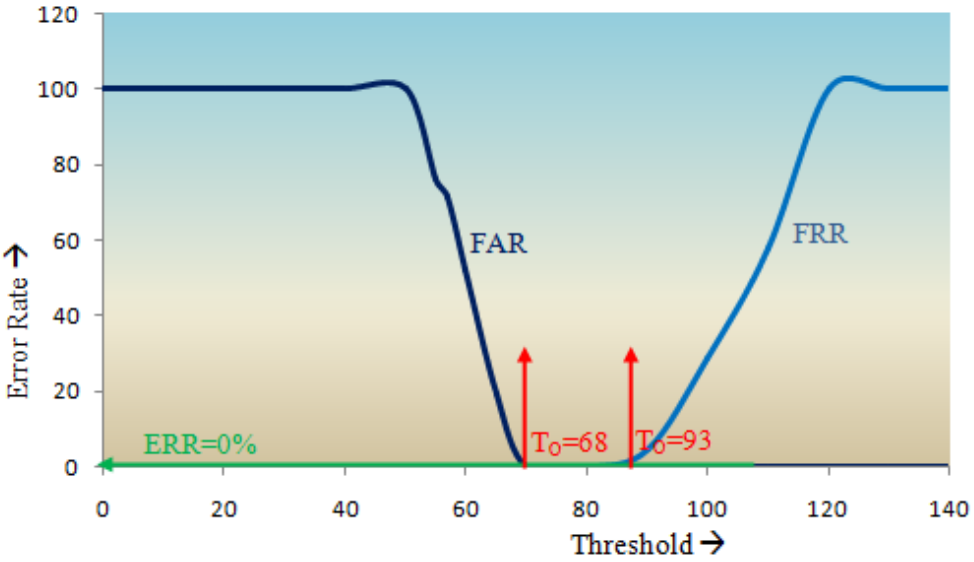

Figure 4.33: EER for Scenario 1 and Scenario 8

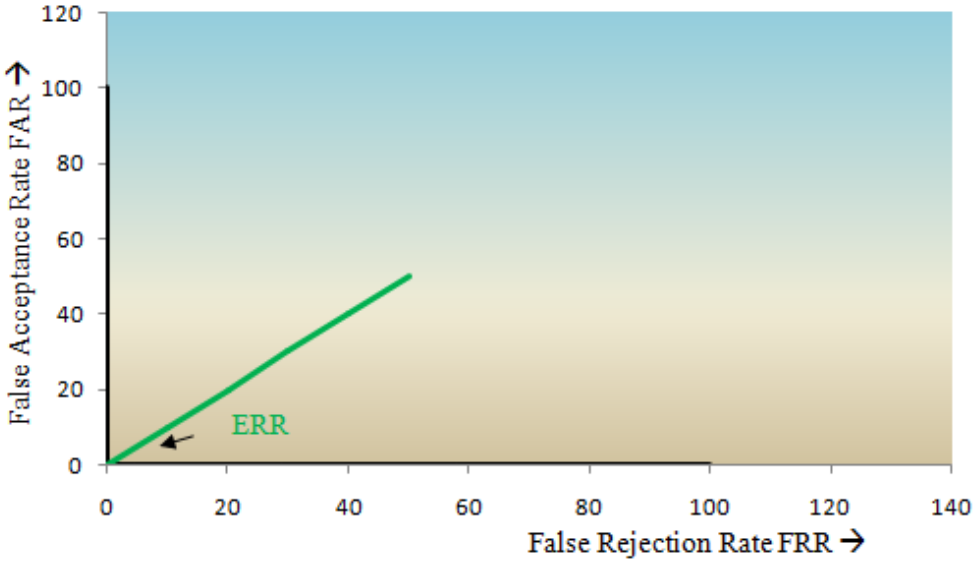

Figure 4.34: ROC Curve for Scenario 1 and 8

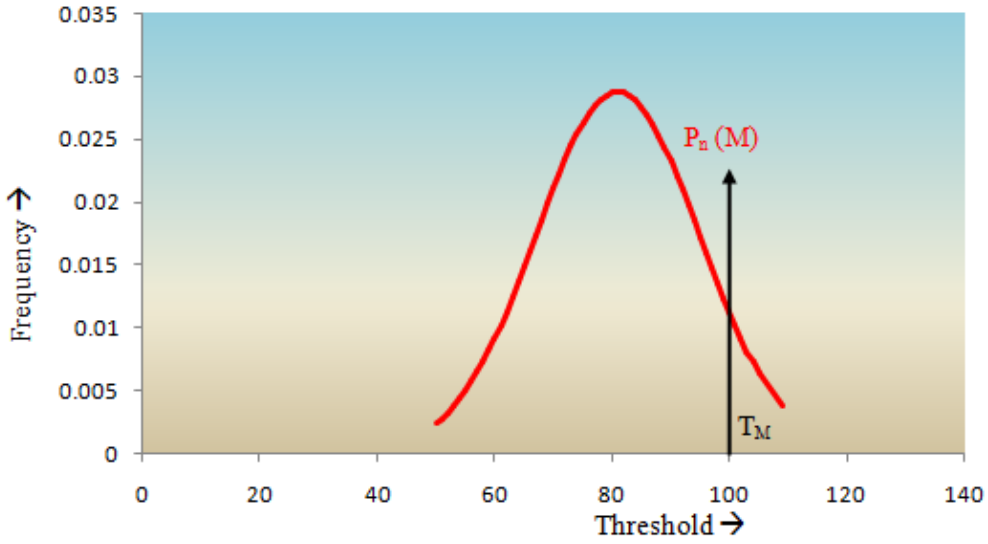

Figure 4.35: Cumulative Frequency Distribution Curve for Illegitimate Attempts

Frequency distribution curves for both legitimate attempts and illegitimate attempts are shown in the figure below. Both the curves are intersecting showing FAR and FRR.
At the pre-set threshold T_M, cumulative false acceptance rate percentage is 4.45%. Error rate is decreasing as the threshold values increase.
Equal error rate ERR almost 4.45% at optimum threshold value of T_O of 100 at the intersection of two curves FAR and FRR.
Receiver operating characteristic curve ROC is shown in the figure below. The curve is very close to the origin indicating that the verification system is excellent.

4.4.10 When User ID Is Incorrect

When the user ID is incorrect, the user (genuine or imposter) is not able to verify itself. There is total$26^5 0$ = 118813760 number of possible user IDs. For an imposter, not knowing the user ID, there are $\frac{118813760}{2} = 59,406,880$ number of user IDs (Brute Force attack); imposter tries before him having a chance to verify himself.
Finding the correct user ID, attacker needs to enter the correct token number associated to the user ID. There are $10^4 = 1000$ token numbers. So, attacker attempts $\frac{1000}{2} = 500$ times (Brute Force attack) before getting the correct token number. As, seen by the scenarios 5 to 8, incorrect token number

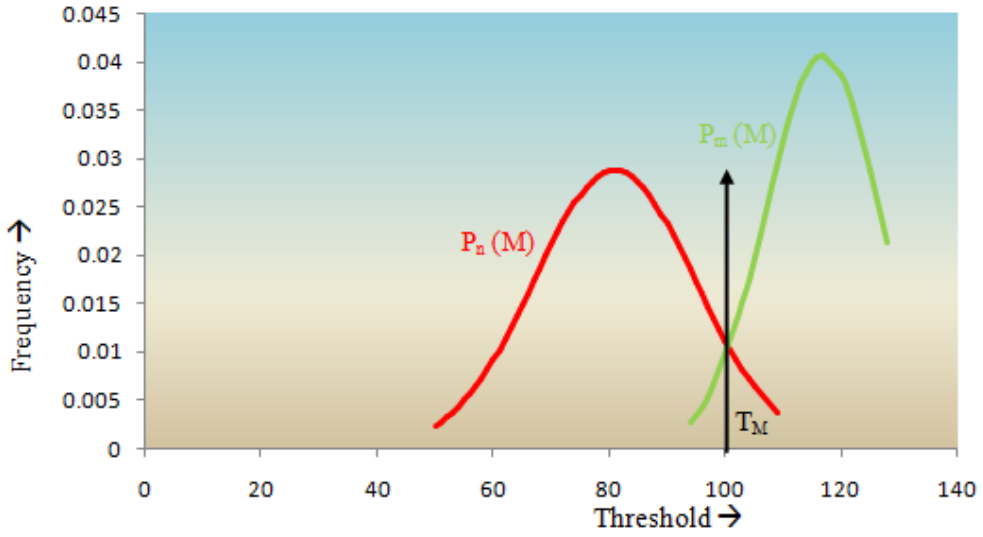

Figure 4.36: Frequency Distribution Curve for Cumulative Illegitimate Attempts and Legitimate Attempts

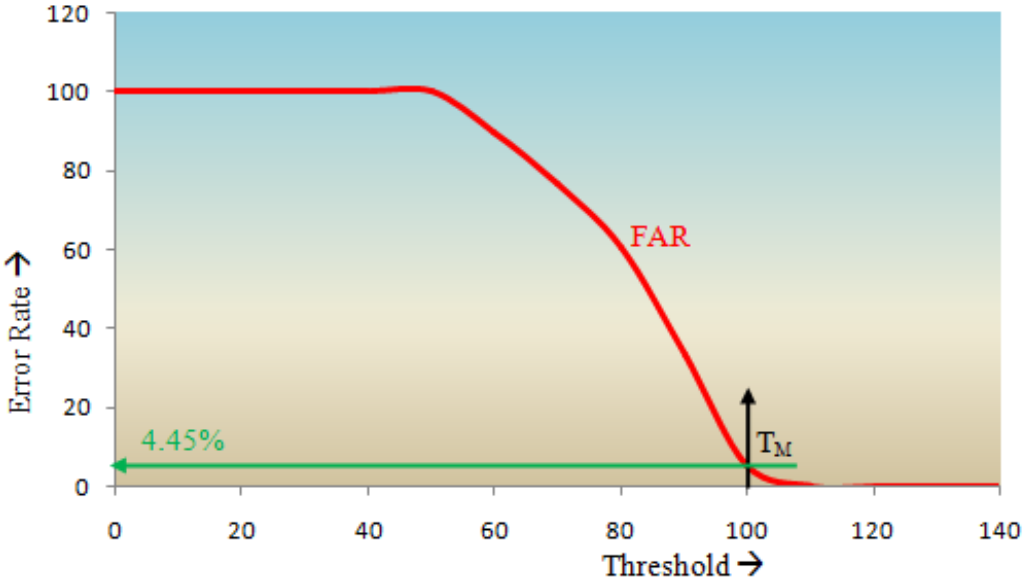

Figure 4.37: Cumulative FAR

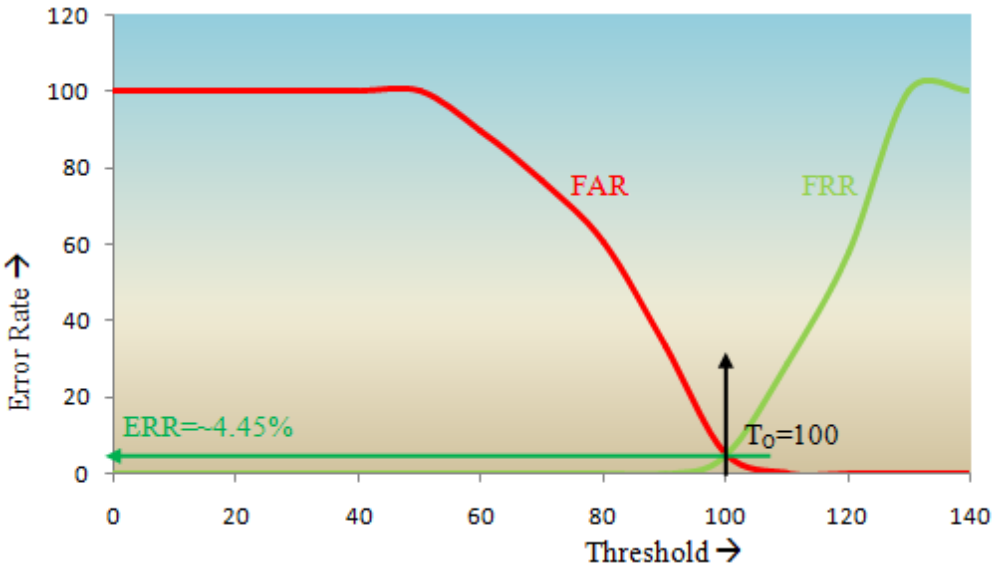

Figure 4.38: Cumulative EER

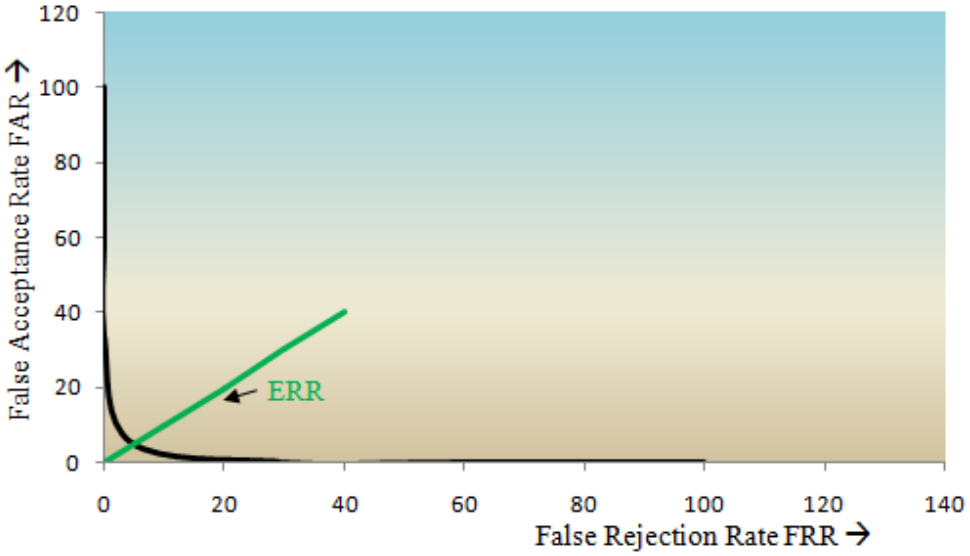

Figure 4.39: Cumulative ROC curve

Table 4.2: Results Of Biometric Authentication System

	Performance Measures						
	Frequency Distribution		FRR	FAR	EER	T_O	ROC
Scenario	Mean	SD					
1.	116.822	9.805	4.444	-	-	-	-
2.	89.595	10.611	-	19.047	8	103	good
3.	89.667	8.399	-	6.349	5	105.5	very good
4.	84.622	9.745	-	4.762	4.7	100	excellent
5.	59.826	5.193	-	0		70-93	ideal
6.	58.829	5.909	-	0		70-93	ideal
7.	57.125	5.488	-	0		70-93	ideal
8.	58.619	5.518	-	0		68-93	ideal
Total(2-8)	80.949	13.860	-	4.45	4.45	100	excellent

results in a very low match score even when the database used have users of same geographical area.
Total number of attempts for an imposter will be $59406880 \times 500 = 29,703,440,000$. One match score generation takes 40.976 seconds. So,

$$\begin{aligned} 29,703,440,000 \times 40.976 &= 1,217,128,157,44 \quad \text{seconds} \\ &= 20,285,469,290.667 \quad \text{minutes} \\ &= 338,091,154.844 \quad \text{hours} \\ &= 14,087,131.452 \quad \text{days} \\ &= 38,568.464 \quad \text{years} \end{aligned}$$

This is computationally exhaustive and practically infeasible as well.

4.5 Comparison of Accuracy

The accuracy of the system is improved when local binary images are bio hashed because of the supplementary token information results in increased

entropy of the transformed template that is saved in the database. However, the classification time is increased significantly as a number of processes are in line before the final decision is to be made.

Table 4.3: Results Of Biometric Authentication System

Technique	Accuracy(%)	Classification time (sec)
WLD-LBI-CT [119]	94.5	0.45
Proposed	95.5	40.976

Chapter 5

Conclusions and Future Work

5.1 Conclusion

In this thesis, a new approach to multifactor user authentication using facial identity and facial expression is discussed along with its biometric template security. Seven different expressions and 4- digit token numbers are used in conjunction with the personś face identity which creates 70,000 expression hash templates that can be stored for one person.
The results show that different templates created with; same token number and different expression are 7, different token number and same expression are 10,000, different token number and different expression are 70,000 for a single person that are easily distinguishable. So, any person can store different templates for different systems and they cannot be cross matched if one of the databases gets compromised, i.e. more diversity is achieved.
The large set of expression hash templates allows the person to easily revoke the compromised stored template and create a new one by just changing the token number or expression using his face identity, i.e. easy revocability.
Expression hashing increases the entropy of the template with the additional information of expressions and token number. So, original template is hard to recover preventing the creation of a physical spoof, i.e. increased security.
The introduction of the token numbers increases the overall accuracy of the biometric verification system with 95.5% accuracy rate in comparison to the expression recognition system previously designed with 94.5% accuracy rate i.e. increased accuracy rate.
However, the performance time of the system is degraded from 0.45 seconds to 40.97 seconds as features extracted from the query image are first transformed into binary sequence called expression hash and then matched sequentially with the one stored in the database.

5.2 Future Directions

In this section, a few improvements in the system are proposed.

5.2.1 Feature Extraction

Features are extracted from the images using local binary patterns and discrete cosine transforms [119]. With the advancement in the field of image processing better methods may be devised which are used to extract features for identification and expression recognition.

5.2.2 Enrollment

The generation of the user IDs is based on a very simple method which may be enhanced using some complex method which include birth dates etc. to create more unique and more user specific IDs. Token numbers that are used are of four digits only; increasing the length of the token number may increase the accuracy of the system. It is therefore, highly recommended to explore the effect of length of token number on accuracy.

5.2.3 Pseudo Random Number Generator

Blum Blum Shub, PRNG is used to generate a sequence of random numbers which are then transformed into orthonormal vectors. There are many PRNGs available which may be used and analyzed solely to check and compare the performance of the system in terms of accuracy and time.

5.2.4 Transformation Rounds

The extracted feature vector is transformed by taking vector product with the orthonormal vectors which are created by the random number sequence generated by the token number as seed once. Feature vector may be transformed multiple times [138] by generating multiple sets of orthonormal vectors generated by multiple token numbers. It is suggested to detect accuracy difference by increasing transformation cycles.

5.2.5 Threshold Method in Biohashing

Threshold is applied by calculating the mean of the values of the transformed feature vector. There are numerous mathematical functions or may be a new

algorithm is designed to calculate threshold which can be applied to check the accuracy of the produced expression hash.

5.2.6 Matcher

The matcher module matches each binary value of the 128 bit long query expression hash with the stored expression hash sequentially. All the bits can be matched in a parallel process decreasing the matching time to 128 times.

Bibliography

[1] Zdenek Riha et al. Toward reliable user authentication through biometrics. *IEEE Security & Privacy*, (3):45–49, 2003.

[2] Kevin Beaver. *Hacking for dummies*. John Wiley & Sons, 2007.

[3] Anil Jain, Arun Ross, and Salil Prabhakar. Fingerprint matching using minutiae and texture features. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 282–285. IEEE, 2001.

[4] Hengjian Li, Jiashu Zhang, and Zutao Zhang. Generating cancelable palmprint templates via coupled nonlinear dynamic filters and multiple orientation palmcodes. *Information sciences*, 180(20):3876–3893, 2010.

[5] Jean-François Connolly, Eric Granger, and Robert Sabourin. An adaptive classification system for video-based face recognition. *Information Sciences*, 192:50–70, 2012.

[6] K Saraswathi, B Jayaram, and R Balasubramanian. Retinal biometrics based authentication and key exchange system. *International Journal of Computer Application*, 19(1), 2011.

[7] Carmen Sanchez-Avila and Raul Sanchez-Reillo. Two different approaches for iris recognition using gabor filters and multiscale zero-crossing representation. *Pattern Recognition*, 38(2):231–240, 2005.

[8] Bo Huang, Jinsong Wu, David Zhang, and Naimin Li. Tongue shape classification by geometric features. *Information Sciences*, 180(2):312–324, 2010.

[9] Amit Kale, Aravind Sundaresan, AN Rajagopalan, Naresh P Cuntoor, Amit K Roy-Chowdhury, Volker Kruger, and Rama Chellappa. Identification of humans using gait. *IEEE Transactions on image processing*, 13(9):1163–1173, 2004.

[10] Francesco Bergadano, Daniele Gunetti, and Claudia Picardi. User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC)*, 5(4):367–397, 2002.

[11] Francesco Bergadano, Daniele Gunetti, and Claudia Picardi. User authentication through keystroke dynamics. *ACM Transactions on Information and System Security (TISSEC)*, 5(4):367–397, 2002.

[12] Falko Ramann, Claus Vielhauer, and Ralf Steinmetz. Biometric applications based on handwriting. In *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, volume 2, pages 573–576. IEEE, 2002.

[13] Lucas Ballard, Daniel Lopresti, and Fabian Monrose. Evaluating the security of handwriting biometrics. In *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft, 2006.

[14] Juergen Luettin, Neil A Thacker, and Steve W Beet. Speaker identification by lipreading. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, volume 1, pages 62–65. IEEE, 1996.

[15] Olga Shipilova. Person recognition based on lip movements. *retrieved July*, 15(2006):03–04, 2006.

[16] Conrad Sanderson and Kuldip K Paliwal. Information fusion for robust speaker verification. In *Seventh European Conference on Speech Communication and Technology*, 2001.

[17] T Westeyn, P Pesti, K Park, and T Starner. Biometric identification using song-based eye blink patterns. *Human Computer Interaction International (HCII), Las Vegas, NV*, 2005.

[18] Václav Matyáš and Zdeněk Říha. Biometric authenticationsecurity and usability. In *Advanced Communications and Multimedia Security*, pages 227–239. Springer, 2002.

[19] Elizabeth C Ambs. The relationship between forensic art and criminal investigations. 2015.

[20] Koichiro Niinuma, Unsang Park, and Anil K Jain. Soft biometric traits for continuous user authentication. *IEEE Transactions on information forensics and security*, 5(4):771–780, 2010.

[21] NIST FIPS Pub. 197: Advanced encryption standard (aes). *Federal information processing standards publication*, 197(441):0311, 2001.

[22] William C Barker, Elaine Barker, et al. Recommendation for the triple data encryption algorithm (tdea) block cipher: Nist special publication 800-67, revision 2. 2012.

[23] Ronald L Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[24] Robert C Seacord. *Secure Coding in C and C++*. Pearson Education, 2005.

[25] Andy Adler. Vulnerabilities in biometric encryption systems. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 1100–1109. Springer, 2005.

[26] Anil K Jain, Arun Ross, and Umut Uludag. Biometric template security: Challenges and solutions. In *Signal Processing Conference, 2005 13th European*, pages 1–4. Citeseer, 2005.

[27] Anil K Jain, Karthik Nandakumar, and Abhishek Nagar. Biometric template security. *EURASIP Journal on advances in signal processing*, 2008:113, 2008.

[28] Kelly Owen and Paul Anthony Howell. Portable device and methods for performing secure transactions, August 24 2010. US Patent 7,780,080.

[29] Whitfield Diffie and Susan Landau. *Privacy on the line: The politics of wiretapping and encryption*. MIT press, 2010.

[30]

[31] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *Proceedings of the 6th ACM conference on Computer and communications security*, pages 28–36. ACM, 1999.

[32] Andrew Beng Jin Teoh and Jaihie Kim. Secure biometric template protection in fuzzy commitment scheme. *IEICE Electronics Express*, 4(23):724–730, 2007.

[33] Karthik Nandakumar. A fingerprint cryptosystem based on minutiae phase spectrum. In *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*, pages 1–6. IEEE, 2010.

[34] Feng Hao, Ross Anderson, and John Daugman. Combining crypto with biometrics effectively. *IEEE transactions on computers*, 55(9):1081–1088, 2006.

[35] Julien Bringer, Hervé Chabanne, Gérard Cohen, Bruno Kindarji, and Gilles Zémor. Optimal iris fuzzy sketches. In *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on*, pages 1–6. IEEE, 2007.

[36] Julien Bringer, Hervé Chabanne, Gerard Cohen, Bruno Kindarji, and Gilles Zemor. Theoretical and practical boundaries of binary secure sketches. *IEEE Transactions on Information Forensics and Security*, 3(4):673–683, 2008.

[37] Christian Rathgeb and Andreas Uhl. Systematic construction of iris-based fuzzy commitment schemes. In *International Conference on Biometrics*, pages 940–949. Springer, 2009.

[38] Christian Rathgeb and Andreas Uhl. Context-based texture analysis for secure revocable iris-biometric key generation. 2009.

[39] Christian Rathgeb and Andreas Uhl. Adaptive fuzzy commitment scheme based on iris-code error analysis. In *Visual Information Processing (EUVIP), 2010 2nd European Workshop on*, pages 41–44. IEEE, 2010.

[40] Christian Rathgeb and Andreas Uhl. Adaptive fuzzy commitment scheme based on iris-code error analysis. In *Visual Information Processing (EUVIP), 2010 2nd European Workshop on*, pages 41–44. IEEE, 2010.

[41] Michiel Van Der Veen, Tom Kevenaar, Geert-Jan Schrijen, Ton H Akkermans, and Fei Zuo. Face biometrics with renewable templates. In *Security, Steganography, and Watermarking of Multimedia Contents VIII*, volume 6072, page 60720J. International Society for Optics and Photonics, 2006.

[42] Meng Ao and Stan Z Li. Near infrared face based biometric key binding. In *International Conference on Biometrics*, pages 376–385. Springer, 2009.

[43] Haiping Lu, Karl Martin, Francis Bui, KN Plataniotis, and Dimitris Hatzinakos. Face recognition with biometric encryption for privacy-enhancing self-exclusion. In *Digital Signal Processing, 2009 16th International Conference on*, pages 1–8. IEEE, 2009.

[44] Emanuele Maiorana, Patrizio Campisi, and Alessandro Neri. User adaptive fuzzy commitment for signature template protection and renewability. *Journal of Electronic Imaging*, 17(1):011011, 2008.

[45] Emanuele Maiorana and Patrizio Campisi. Fuzzy commitment for function based signature template protection. *IEEE Signal Processing Letters*, 17(3):249–252, 2010.

[46] Gang Zheng, Wanqing Li, and Ce Zhan. Cryptographic key generation from biometric data using lattice mapping. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 4, pages 513–516. IEEE, 2006.

[47] Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *International conference on the theory and applications of cryptographic techniques*, pages 523–540. Springer, 2004.

[48] Ari Juels and Madhu Sudan. A fuzzy vault scheme. *Designs, Codes and Cryptography*, 38(2):237–257, 2006.

[49] T Charles Clancy, Negar Kiyavash, and Dennis J Lin. Secure smartcardbased fingerprint authentication. In *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 45–52. ACM, 2003.

[50] Karthik Nandakumar, Anil K Jain, and Sharath Pankanti. Fingerprint-based fuzzy vault: Implementation and performance. *IEEE transactions on information forensics and security*, 2(4):744–757, 2007.

[51] Umut Uludag, Sharath Pankanti, Salil Prabhakar, and Anil K Jain. Biometric cryptosystems: issues and challenges. *Proceedings of the IEEE*, 92(6):948–960, 2004.

[52] Umut Uludag and Anil K Jain. Fuzzy fingerprint vault. In *Proc. Workshop: Biometrics: Challenges arising from theory to practice*, pages 13–16, 2004.

[53] Umut Uludag and Anil Jain. Securing fingerprint template: Fuzzy vault with helper data. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pages 163–163. IEEE, 2006.

[54] Abhishek Nagar, Karthik Nandakumar, and Anil K Jain. A hybrid biometric cryptosystem for securing fingerprint minutiae templates. *Pattern Recognition Letters*, 31(8):733–741, 2010.

[55] Cai Li and Jiankun Hu. A security-enhanced alignment-free fuzzy vault-based fingerprint cryptosystem using pair-polar minutiae structures. *IEEE Transactions on Information Forensics and Security*, 11(3):543–555, 2016.

[56] Youn Joo Lee, Kwanghyuk Bae, Sung Joo Lee, Kang Ryoung Park, and Jaihie Kim. Biometric key binding: Fuzzy vault based on iris images. In *International Conference on Biometrics*, pages 800–808. Springer, 2007.

[57] Xiangqian Wu, Ning Qi, Kuanquan Wang, and David Zhang. A novel cryptosystem based on iris key generation. In *Natural Computation, 2008. ICNC'08. Fourth International Conference on*, volume 4, pages 53–56. IEEE, 2008.

[58] E Srinivasa Reddy and I Ramesh Babu. Performance of iris based hard fuzzy vault. In *IEEE 8th International Conference on Computer and Information Technology Workshops*, pages 248–253. IEEE, 2008.

[59] Xiangqian Wu, Kuanquan Wang, and David Zhang. A cryptosystem based on palmprint feature. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[60] Amioy Kumar and Ajay Kumar. A palmprint-based cryptosystem using double encryption. In *Biometric technology for human identification V*, volume 6944, page 69440D. International Society for Optics and Photonics, 2008.

[61] Amioy Kumar and Ajay Kumar. Development of a new cryptographic construct using palmprint-based fuzzy vault. *EURASIP Journal on Advances in Signal Processing*, 2009:13, 2009.

[62] Lifang Wu, Peng Xiao, Songlong Yuan, Siyuan Jiang, and Chang Wen Chen. A fuzzy vault scheme for ordered biometrics. *Journal of Communications*, 6(9):682–690, 2011.

[63] Yongjin Wang and KN Plataniotis. Fuzzy vault for face based cryptographic key generation. In *Biometrics Symposium, 2007*, pages 1–6. IEEE, 2007.

[64] Alisher Kholmatov and Berrin Yanikoglu. Biometric cryptosystem using online signatures. In *International Symposium on Computer and Information Sciences*, pages 981–990. Springer, 2006.

[65] George S Eskander, Robert Sabourin, and Eric Granger. A biocryptographic system based on offline signature images. *Information Sciences*, 259:170–191, 2014.

[66] Marina Blanton and Mehrdad Aliasgari. Analysis of reusability of secure sketches and fuzzy extractors. *IEEE transactions on information forensics and security*, 8(9):1433–1445, 2013.

[67] Marina Blanton and Mehrdad Aliasgari. On the (non-) reusability of fuzzy sketches and extractors and security in the computational setting. In *Security and Cryptography (SECRYPT), 2011 Proceedings of the International Conference on*, pages 68–77. IEEE, 2011.

[68] Chi Chen, Chaogang Wang, Tengfei Yang, Dongdai Lin, Song Wang, and Jiankun Hu. Optional multi-biometric cryptosystem based on fuzzy extractor. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2014 11th International Conference on*, pages 989–994. IEEE, 2014.

[69] Thi Ai Thao Nguyen and Tran Khanh Dang. Combining fuzzy extractor in biometric-kerberos based authentication protocol. In *Advanced Computing and Applications (ACOMP), 2015 International Conference on*, pages 1–6. IEEE, 2015.

[70] Nazatul Haque Sultan, Ferdous Ahmed Barbhuiya, and Nityananda Sarma. Pairvoting: A secure online voting scheme using pairing-based cryptography and fuzzy extractor. In *Advanced Networks and Telecommuncations Systems (ANTS), 2015 IEEE International Conference on*, pages 1–6. IEEE, 2015.

[71] Masato Taniguchi, Mitsuru Shiozaki, Hiroshi Kubo, and Takeshi Fujino. A stable key generation from puf responses with a fuzzy extractor for cryptographic authentications. In *Consumer Electronics (GCCE), 2013 IEEE 2nd Global Conference on*, pages 525–527. IEEE, 2013.

[72] Hyunho Kang, Yohei Hori, Toshihiro Katashita, Manabu Hagiwara, and Keiichi Iwamura. Cryptographie key generation from puf data using efficient fuzzy extractors. In *Advanced Communication Technology (ICACT), 2014 16th International Conference on*, pages 23–26. IEEE, 2014.

[73] Charles Herder, Ling Ren, Marten van Dijk, Meng-Day Yu, and Srinivas Devadas. Trapdoor computational fuzzy extractors and stateless cryptographically-secure physical unclonable functions. *IEEE Transactions on Dependable and Secure Computing*, 14(1):65–82, 2017.

[74] Kai Xi, Jiankun Hu, and Fengling Han. An alignment free fingerprint fuzzy extractor using near-equivalent dual layer structure check (nedlsc) algorithm. In *Industrial Electronics and Applications (ICIEA), 2011 6th IEEE Conference on*, pages 1040–1045. IEEE, 2011.

[75] Wencheng Yang, Jiankun Hu, and Song Wang. A delaunay triangle-based fuzzy extractor for fingerprint authentication. In *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications*, pages 66–70. IEEE, 2012.

[76] Andrew BJ Teoh, David CL Ngo, and Alwyn Goh. Personalised cryptographic key generation based on facehashing. *Computers & Security*, 23(7):606–614, 2004.

[77] David CL Ngo, Andrew BJ Teoh, and Alwyn Goh. Biometric hash: high-confidence face recognition. *IEEE transactions on circuits and systems for video technology*, 16(6):771–775, 2006.

[78] Thian Song Ong, Andrew Teoh Beng Jin, and David Chek Ling Ngo. Application-specific key release scheme from biometrics. *IJ Network Security*, 6(2):127–133, 2008.

[79] Andrew Teoh Beng Jin, David Ngo Chek Ling, and Alwyn Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 37(11):2245–2255, 2004.

[80] Chong Siew Chin, Andrew Teoh Beng Jin, and David Ngo Chek Ling. High security iris verification system based on random secret integration. *Computer Vision and Image Understanding*, 102(2):169–177, 2006.

[81] Siew Chin Chong, Andrew Beng Jin Teoh, and David Chek Ling Ngo. Iris authentication using privatized advanced correlation filter. In *International Conference on Biometrics*, pages 382–388. Springer, 2006.

[82] Tee Connie, Andrew Teoh, Michael Goh, and David Ngo. Palmhashing: a novel approach for cancelable biometrics. *Information processing letters*, 93(1):1–5, 2005.

[83] Thian Song Ong, Andrew Teoh Beng Jin, and David Chek Ling Ngo. Application-specific key release scheme from biometrics. *IJ Network Security*, 6(2):127–133, 2008.

[84] Andrew Beng, Jin Teoh, and Kar-Ann Toh. Secure biometric-key generation with biometric helper. In *Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference on*, pages 2145–2150. IEEE, 2008.

[85] Adams Kong, King-Hong Cheung, David Zhang, Mohamed Kamel, and Jane You. An analysis of biohashing and its variants. *Pattern recognition*, 39(7):1359–1368, 2006.

[86] Andrew BJ Teoh, Yip Wai Kuan, and Sangyoun Lee. Cancellable biometrics and annotations on biohash. *Pattern recognition*, 41(6):2034–2044, 2008.

[87] Alessandra Lumini and Loris Nanni. An improved biohashing for human authentication. *Pattern recognition*, 40(3):1057–1065, 2007.

[88] Loris Nanni and Alessandra Lumini. Random subspace for an improved biohashing for face authentication. *Pattern Recognition Letters*, 29(3):295–300, 2008.

[89] Yagiz Sutcu, Husrev Taha Sencar, and Nasir Memon. A secure biometric authentication scheme based on robust hashing. In *Proceedings of the 7th workshop on Multimedia and security*, pages 111–116. ACM, 2005.

[90] Nalini K Ratha, Sharat Chikkerur, Jonathan H Connell, and Ruud M Bolle. Generating cancelable fingerprint templates. *IEEE Transactions on pattern analysis and machine intelligence*, 29(4):561–572, 2007.

[91] Nalini Ratha, Jonathan Connell, Ruud M Bolle, and Sharat Chikkerur. Cancelable biometrics: A case study in fingerprints. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 4, pages 370–373. IEEE, 2006.

[92] Dongdong Zhao, Wenjian Luo, Ran Liu, and Lihua Yue. Negative iris recognition. *IEEE Transactions on Dependable and Secure Computing*, 15(1):112–125, 2018.

[93] Jutta Hämmerle-Uhl, Elias Pschernig, and Andreas Uhl. Cancelable iris biometrics using block re-mapping and image warping. In *International Conference on Information Security*, pages 135–142. Springer, 2009.

[94] Inchul Song, Hyun-Jun Kim, and Paul Barom Jeon. Deep learning for real-time robust facial expression recognition on a smartphone. In *Consumer Electronics (ICCE), 2014 IEEE International Conference on*, pages 564–567. IEEE, 2014.

[95] Ying-li Tian. Evaluation of face resolution for expression analysis. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*, pages 82–82. IEEE, 2004.

[96] Y Tian, T Kanade, and J Cohn. Handbook of face recognition, chapter 11. facial expression analysis, 2005.

[97] Caifeng Shan, Shaogang Gong, and Peter W McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and vision Computing*, 27(6):803–816, 2009.

[98] Michel Valstar and Maja Pantic. Fully automatic facial action unit detection and temporal analysis. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pages 149–149. IEEE, 2006.

[99] Yansheng Li, Chao Tao, Yihua Tan, Ke Shang, and Jinwen Tian. Unsupervised multilayer feature learning for satellite image scene classification. *IEEE Geoscience and Remote Sensing Letters*, 13(2):157–161, 2016.

[100] Caifeng Shan, Shaogang Gong, and Peter W McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and vision Computing*, 27(6):803–816, 2009.

[101] Kaimin Yu, Zhiyong Wang, Li Zhuo, Jiajun Wang, Zheru Chi, and Dagan Feng. Learning realistic facial expressions from web images. *Pattern Recognition*, 46(8):2144–2155, 2013.

[102] Wujun Chen, Xiaobo Lu, Yijun Du, and Wenqi Tian. Boosting local gabor binary patterns for gender recognition. In *Natural Computation (ICNC), 2013 Ninth International Conference on*, pages 34–38. IEEE, 2013.

[103] Mahesh Kumbhar, Ashish Jadhav, and Manasi Patil. Facial expression recognition based on image feature. *International Journal of Computer and Communication Engineering*, 1(2):117, 2012.

[104] Rizwan Ahmed Khan, Alexandre Meyer, Hubert Konik, and Saida Bouakaz. Framework for reliable, real-time facial expression recognition for low resolution images. *Pattern Recognition Letters*, 34(10):1159–1168, 2013.

[105] Yan Ouyang, Nong Sang, and Rui Huang. Accurate and robust facial expressions recognition by fusing multiple sparse representation based classifiers. *Neurocomputing*, 149:71–78, 2015.

[106] Tao Gao, XL Feng, He Lu, and JH Zhai. A novel face feature descriptor using adaptively weighted extended lbp pyramid. *Optik-International Journal for Light and Electron Optics*, 124(23):6286–6291, 2013.

[107] Wei-Lun Chao, Jian-Jiun Ding, and Jun-Zuo Liu. Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection. *Signal Processing*, 117:1–10, 2015.

[108] Xijian Fan and Tardi Tjahjadi. A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences. *Pattern Recognition*, 48(11):3407–3416, 2015.

[109] XiaoHui Guo, Xiao Zhang, Chao Deng, and Jianyu Wei. Facial expression recognition based on independent component analysis. *Journal of Multimedia*, 8(4), 2013.

[110] Ying Tong, Rui Chen, and Yong Cheng. Facial expression recognition algorithm using lgc based on horizontal and diagonal prior principle. *Optik-International Journal for Light and Electron Optics*, 125(16):4186–4189, 2014.

[111] Jyoti Kumari, R Rajesh, and KM Pooja. Facial expression recognition: A survey. *Procedia Computer Science*, 58:486–491, 2015.

[112] Zhan Wang, Qiuqi Ruan, and Gaoyun An. Facial expression recognition using sparse local fisher discriminant analysis. *Neurocomputing*, 174:756–766, 2016.

[113] SL Happy and Aurobinda Routray. Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, 6(1):1–12, 2015.

[114] Ayşegül Uçar, Yakup Demir, and Cüneyt Güzeliş. A new facial expression recognition based on curvelet transform and online sequential

extreme learning machine initialized with spherical clustering. *Neural Computing and Applications*, 27(1):131–142, 2016.

[115] Stefanos Eleftheriadis, Ognjen Rudovic, and Maja Pantic. Discriminative shared gaussian processes for multiview and view-invariant facial expression recognition. *IEEE transactions on image processing*, 24(1):189–204, 2015.

[116] Wei Zhang, Youmei Zhang, Lin Ma, Jingwei Guan, and Shijie Gong. Multimodal learning for facial expression recognition. *Pattern Recognition*, 48(10):3191–3202, 2015.

[117] Siti Khairuni Amalina Kamarol, Mohamed Hisham Jaward, Jussi Parkkinen, and Rajendran Parthiban. Spatiotemporal feature extraction for facial expression recognition. *IET Image Processing*, 10(7):534–541, 2016.

[118] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2001.

[119] Sajid Ali Khan, Ayyaz Hussain, and Muhammad Usman. Reliable facial expression recognition for multi-scale images using weber local binary image based cosine transform features. *Multimedia Tools and Applications*, 77(1):1133–1165, 2018.

[120] Saeed Dabbaghchian, Masoumeh P Ghaemmaghami, and Ali Aghagolzadeh. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *Pattern recognition*, 43(4):1431–1440, 2010.

[121] Nasir Ahmed, T_ Natarajan, and Kamisetty R Rao. Discrete cosine transform. *IEEE transactions on Computers*, 100(1):90–93, 1974.

[122] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002.

[123] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In *European conference on computer vision*, pages 469–481. Springer, 2004.

[124] Huawen Liu, Jigui Sun, Lei Liu, and Huijie Zhang. Feature selection with dynamic mutual information. *Pattern Recognition*, 42(7):1330–1339, 2009.

[125] Jie Chen, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikainen, Xilin Chen, and Wen Gao. Wld: A robust local image descriptor. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1705–1720, 2010.

[126] Caroline Silva, Thierry Bouwmans, and Carl Frélicot. An extended center-symmetric local binary pattern for background modeling and subtraction in videos. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISAPP 2015*, 2015.

[127] Fadi Dornaika, Elena Lazkano, and Basilio Sierra. Improving dynamic facial expression recognition with feature subset selection. *Pattern Recognition Letters*, 32(5):740–748, 2011.

[128] Xiao-Yuan Jing, Yuan-Yan Tang, and David Zhang. A fourier–lda approach for image recognition. *Pattern Recognition*, 38(3):453–457, 2005.

[129] Zhang Yankun and Liu Chongqing. Efficient face recognition method based on dct and lda. *Journal of Systems Engineering and Electronics*, 15(2):211–216, 2004.

[130] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[131] Werner Schindler. Functionality classes and evaluation methodology for deterministic random number generators. *Anwendungshinweise and Interpretation (AIS)*, pages 5–11, 1999.

[132] Lenore Blum, Manuel Blum, and Mike Shub. A simple unpredictable pseudo-random number generator. *SIAM Journal on computing*, 15(2):364–383, 1986.

[133] Andrey Sidorenko and Berry Schoenmakers. Concrete security of the blum-blum-shub pseudorandom generator. In *IMA International Conference on Cryptography and Coding*, pages 355–375. Springer, 2005.

[134] Ward Cheney and David Kincaid. Linear algebra: Theory and applications. *The Australian Mathematical Society*, 110, 2009.

[135] Gregory H Moore. The axiomatization of linear algebra: 1875-1940. *Historia Mathematica*, 22(3):262–303, 1995.

[136] PK Jain and K Ahmad. Definitions and basic properties of inner product spaces and hilbert spaces, 1995.

[137] Michael J Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba, and Julien Budynek. The japanese female facial expression (jaffe) database. In *Proceedings of third international conference on automatic face and gesture recognition*, pages 14–16, 1998.

[138] Alessandra Lumini and Loris Nanni. An improved biohashing for human authentication. *Pattern recognition*, 40(3):1057–1065, 2007.