REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE M'HAMED BOUGARA-BOUMERDES

Institut de **G**énie **E**lectrique **E**lectronique

Mémoire de Magister

Présenté par :

HAFNAOUI Imane

En vue de l'obtention du diplôme de **MAGISTER** en **:**

Filière : Electrical And Electronic Engineering

Option : Electrical And Electronic System Engineering

Multimodal Biometric Fusion using Evolutionary Techniques

Devant le jury composé de :

M^{r}	DAHIMENE Abdelhakim	MC/A	UMBB, Boumerdes	Président
M^{me}	SERIR Amina	Professeur	USTHB, Alger	Examinateur
M^{r}	HAMMOUCHE Kamel	Professeur	UMMTO, Tizi-Ouzou	Examinateur
M^{me}	CHERIFI Dalila	MC/B	UMBB, Boumerdes	Promoteur

Année Universitaire **2013 /2014**

Abstract

The work of this research focuses on fusing multiple biometric modalities at the score level using different combination rules. The research puts an emphasis on employing optimization techniques in order to achieve optimum accuracies.

Due to the limitations that unimodal systems suffer from, such as noisy data, non-universality, and susceptibility to spoof attacks, multibiometric systems have gained much interest in the research community on the grounds that they alleviate most of these limitations and are capable of producing better accuracies and performances. A multibiometric system combines two or more biometric sources in order to overcome their unimodal system counterparts and achieve higher accuracies. One of the important steps to reach this purpose is the choice of the fusion techniques utilized. A thorough study is performed to investigate the different fusion rules and schemes.

In this work, a modeling step based on a hybrid algorithm that includes social rules derived from the swarm intelligence, Particle Swarm Optimization, and the concepts of natural selection and evolution, Genetic Algorithm, is used to combine the two modalities at the score level. This optimization algorithm is employed to select the optimum weights associated to the modalities being fused. The performance of the hybrid GA-PSO is compared to those of classical combination rules. For that purpose, the proposed schemes are experimentally evaluated on publicly available score databases (XM2VTS, NIST and BANCA) which come in clean and degraded conditions.

An analysis of the results is carried out on the basis of comparing the techniques' resulting EER accuracies and ROC curves. Furthermore, the execution speed of the hybrid approach is compared to that of the single optimization algorithms GA and PSO.

Keywords: *Multibiometric, fusion, multimodal, score level, Genetic Algorithm, Particle Swarm Optimization, Hybrid, GA-PSO, optimization techniques, databases, EER, ROC, pattern recognition.*

Résumé

Le travail de cette recherche se concentre sur la fusion de plusieurs modalités biométriques au niveau des scores en utilisant différentes règles de combinaison. La recherche met l'accent sur l'emploi des techniques d'optimisation afin d'obtenir des performances optimales .

En raison des limites des systèmes unimodales, telles que les données bruyant , non - universalité , et la susceptibilité à usurper les attaques , les systèmes multi-biométriques ont gagné beaucoup d'intérêt dans le milieu de la recherche puisqu'ils traitent la plupart de ces limitations et sont capables de produire de meilleures précisions et performances. Un système multi-biométrique combine deux ou plusieurs sources biométriques afin de surmonter les limites des systèmes unimodales et atteindre des précisions plus élevées. Une des étapes importantes pour atteindre cet objectif est le choix des techniques de fusion utilisées. Une étude approfondie est effectuée pour étudier les différentes règles et régimes de fusion.

Dans ce mémoire, une étape de modélisation basé sur un algorithme hybride qui comprend des règles sociaux issues de l'intelligence de l'essaim , essaim de particules d'optimisation , et les concepts de la sélection naturelle et l'évolution , l'algorithme génétique , est utilisé pour combiner les deux modalités au niveau des scores. Cet algorithme d'optimisation est utilisé pour sélectionner les poids optimaux associés aux modalités étant fusionnées. Le rendement de l'hybride GA- PSO est comparé à ceux des règles de combinaison classiques. Les techniques proposés sont évaluées expérimentalement sur des bases de données publiques (XM2VTS, NIST et BANCA) contenant des scores qui viennent dans des conditions de propreté et dégradées.

Une analyse des résultats est effectuée en comparant les valeurs des EERs des techniques testées et des courbes ROC. En outre, la vitesse d'exécution de l'approche hybride est comparée à celle des algorithmes d'optimisation simples GA et PSO.

Mots-clés: *multibiométriques , fusion , multimodal , le niveau de score, Algorithme Génétique , Particle Swarm Optimization , hybride , GA- PSO , les techniques d'optimisation.*

ملخص

هذا البحث يركز على دمج عدة وسائط بيومترية على مستوى درجة التطابق باستخدام قواعد تركيبة مختلفة. يضع البحث التركيز على توظيف تقنيات التحسين من أجل تحقيق الدقة المثلى.

بسبب القيود التي تعاني منها الأنظمة الأحادية الواسطة، مثل البيانات الصاخبة، غير عالمية ، و وإمكانية التعرض للمحاكاة و التقليد، اكتسبت نظم متعددة الوسائط الكثير من الاهتمام في الأوساط البحثية على أساس أنها تخفف معظم هذه القيود و قدرتها على إنتاج الدقة و الأداء الأفضل. نظام متعددة الوسائط يجمع بين اثنين أو أكثر من مصادر البيومترية من أجل التغلب على نظرائهم النظام الأحادي الواسطة وتحقيق الدقة الأعلى. واحدة من الخطوات الهامة لتحقيق هذا الغرض هو اختيار التقنيات المستخدمة للدمج. تم إجراء دراسة وافية للتحقيق في قواعد الدمج المختلفة و المخططات.

هنا، نقترح خطوة إدماجية تعتمد على خوارزمية الهجين التي تتضمن القواعد الاجتماعية المستمدة من استخبارات السرب، سرب الجسيمات الأمثل (PSO)، و مفاهيم الانتقاء الطبيعي و التطور، الخوارزميات الجينية (GA)، ويستخدم في الجمع بين عدة وسائط على مستوى درجة التطابق. ويستخدم هذه الخوارزمية الأمثل لتحديد الأوزان الأمثل المرتبطة بالوسائط التي تندمج فيها. تتم مقارنة أداء الهجين GA-PSO لقواعد الدمج الكلاسيكية. لهذا الغرض، تم تقييم المخططات المقترحة تجريبيا على قواعد البيانات المتاحة للجمهور (XM2VTS، وNIST BANCA) والتي تأتي في ظروف نظيفة و المتدهورة.

ويتم تحليل النتائج على أساس من الدقة بمقارنة EER و منحنيات ROC . علاوة على ذلك، تتم مقارنة سرعة تنفيذ الهجين إلى كل من خوارزميات التحسين GA و سرب الجسيمات الأمثل PSO.

الكلمات الرئيسية : *متعددة الوسائط البيومترية ؛ دمج على مستوى الدرجة ؛ الوجه ؛ الصوت ؛ تقنيات التطور ؛ ذكي هجين GAPSO، GA, PSO.*

Acknowledgement

I would like to take this opportunity to express my thanks for my advisor Dr.CHERIFI Dalila for her unlimited support, devoted guidance and keen encouragement throughout my graduate studies.

I am grateful to Mr. KRIBES Youcef and Mr. BENZEKRI Aziz for guiding me all throughout this journey, motivating me at times when I needed it the most and above all for being thoughtful.

I extend my thanks to my friends, colleagues and students for their understanding and unwavering support, as well as, the Institute of Electrical and Electronic Engineering for providing all the means for the completion of this work.

Finally, I would like to express my utmost gratitude to my parents and siblings for their patience, unconditional love and encouragement to achieve my goals.

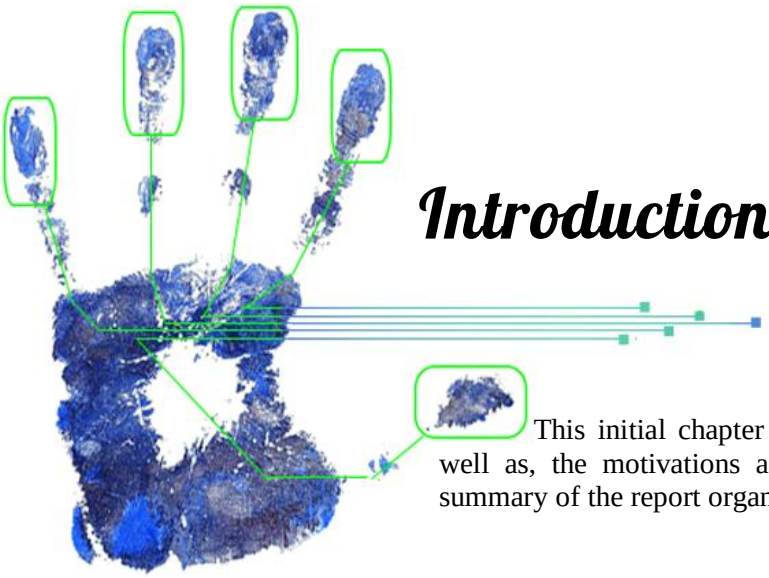
Table of Contents

List of Figures

List of Tables

Glossary of Acronyms

BANCA	Biometric Access Control for Networked and E-Commerce Applications
BFS	Brute Force Search
BSSR1	Biometric Scores Set - Release 1
COTS	Commercial off-the-shelf
CPU	Central Processing Unit
DCT	Discrete Cosine Transform
DNA	Deoxyribonucleic acid
EER	Equal Error Rate
EERG	Equal Error Rate Graph
ET	Evolutionary Techniques
FA	False accept
FAR	False Accept Rate
FR	False reject
FRR	False Reject Rate
GA	Genetic Algorithm
GA-PSO	Hybrid Genetic Algorithm - Particle Swarm Optimization
GMM	Gaussian Mixture Model
gbest	Global Best (PSO)
ID	Identity
K-NN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LFCC	Linear Frequency Cepstral Coefficient
LP	Lausanne Protocol
MLP	Multi Layer Perceptron
NIST	National Institute of Standards and Technology
pbest	Personal Best (PSO)
PCA	Principle Component Analysis
PIN	Personal Identification Number
PSO	Particle Swarm Optimization
ROC	Receiver Operating Characteristics
SS	Simple Sum
SSC	Spatial Synoptic Classification
SVD	Singular Value Decomposition
SVM	Support Vector Machine
W	Weight
WS	Weighted Sum
XM2VTS	Extended Multi Modal Verification for Teleservices and Security applications



Introduction

This initial chapter introduces the theme of this study, as well as, the motivations and objectives to pursue this work. A summary of the report organization is included.

Biometrics is a constantly evolving technology. It has been widely used in many official and commercial identification applications, especially those involving automatic access control. These days, and due to the expansion of the networked society, there is an increasing need for secured and reliable personal identification systems. The need for reliable, simple, flexible and secure systems is a great concern and a challenging issue for several applications that render services to only legitimately enrolled users. Examples of such applications include withdrawal of money from automatic telling machines (ATMs), sharing networked digital resources, access to nuclear facilities, performing remote financial transactions (teleshopping) and physical access control.

Traditionally, passwords, personal cards, PIN-numbers and keys have been used in this context. However, security can easily be breached in these systems when a card or key is lost or stolen or when a password is compromised. Furthermore, difficult passwords may be hard to remember by a legitimate user and simple passwords are easy to guess by an impostor. The use of biometrics offers an alternative means of identification which helps avoid the problems associated with conventional methods.

Biometric authentication, or simply biometrics, is the science of establishing an identity based on the physical, behavioral or chemical attributes of an individual. A biometric-based authentication is basically a pattern recognition problem which makes a personal identification decision in order to determine the authority.

Most biometric systems that are currently in operation typically use a single biometric trait. These traits can be physiological, meaning that which defines *what the user is*, or behavioral, which describe *how the user acts*. Such systems are called unibiometric systems.

The system acquires data from the user, extracts the most important features, compares these features with a stored template and finally returns either an identity or a decision of accept/reject depending on the mode of recognition. Basically, a typical biometric system has four main modules, namely, sensor module, feature extraction module, a matching module, and a database module.

It is becoming increasingly apparent that a single biometric trait, used in a unibiometric system, is not sufficient to meet a number of system requirements, including matching performance, imposed by several large-scale authentication applications [1]. The limitation of unimodal systems, such as noisy sensor data, intra-class variations, non-universality, vulnerability to spoof attacks and more, can lower the performance of the system, and make it more susceptible to refusing a legitimate user and jeopardize personal security.

Multibiometric systems seek to alleviate some of the drawbacks encountered in unibiometric systems by consolidating the evidence presented by multiple biometric sources, such as collecting voice and face or multiple fingerprints of the same person. These systems are expected to significantly improve the recognition performance of a biometric system besides improving population coverage, deterring spoof attacks, and reducing the failure-to-enroll rate.

On the other hand, as a trade-off to increasing the performance of the biometric system, multibiometrics require significantly higher storage space, processing time and computational demand. But the advantages present a compelling case for deploying multibiometric systems in large-scale authentication systems (e.g., border crossing) and systems requiring very high accuracies (e.g., access to a secure military base).

One of the core points of our work revolves around examining whether the performance of a biometric-based authentication system can be improved through integrating complementary biometric traits which come primarily from two different and independent modalities. Therefore, the main aim of the research will be to investigate the effectiveness of the suggested fusion techniques for multimodal biometrics, with the following specific objectives:

- Review the existing multibiometric approaches
- Explore the classical multimodal fusion techniques at the score-level
- Investigate the performance of multimodal system after applying *Genetic Algorithm* and *Particle Swarm Optimization* to the fusion compared to classical methods

This research poses two fundamental questions:

Which fusion scheme can achieve the best performance and which techniques can we apply to improve the suggested fusion scheme?

As a contribution, our work will investigate applying a hybrid algorithm of both Genetic Algorithm and Particle Swarm Optimization to one of the fusion schemes. To the best of our knowledge, there is no reported research work that combined these two algorithms in this context. The objective of this endeavor is to decrease the computational time invested when trying to optimize the process of multimodal fusion.

The report is organized into five chapters, including this one. The content of each chapter is summarized as follows.

Chapter 2

The discipline of biometrics is introduced as well as its evolution towards multimodal biometrics. The key issues with unimodal biometric systems are investigated and the ways multibiometric deals with them. Most used biometric traits are presented in addition to an overview of the biometric recognition system and its modes. The chapter closes with a description of the different multibiometric sources and the distinct levels at which the fusion can be implemented.

Chapter 3

This chapter explores some state-of-the-art fusion schemes and describes their principle in detail. Also, two of the most popular optimization techniques and our proposed hybrid approach are employed to get the most out of the discussed fusion

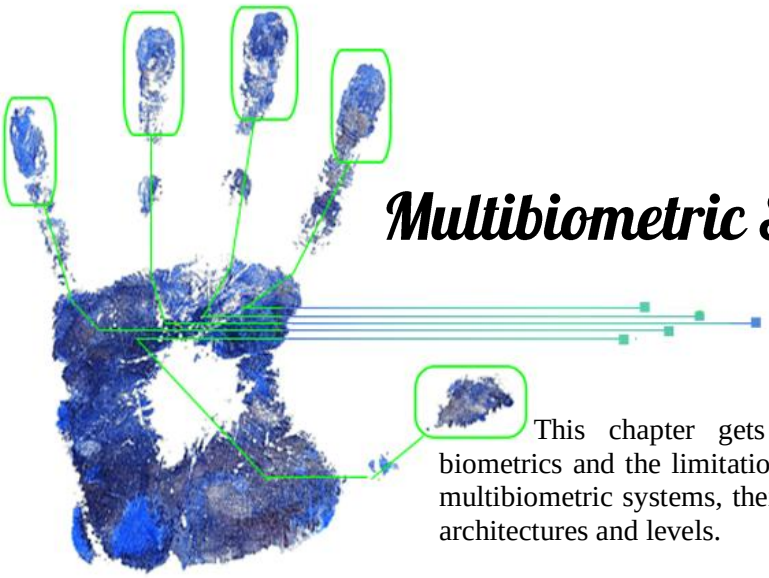
techniques in terms of performance. A thorough description of the algorithms is presented.

Chapter 4

In this chapter, an experimental investigation is conducted on the fusion of two modalities using the classical combinations rules and optimization techniques. To evaluate the performances of the system, the fusion schemes are tested on three public score databases (NIST, XM2VTS and BANCA). Performance metrics used in the evaluation are reviewed here. Afterwards, taking advantage of both Genetic Algorithm and Particle Swarm Optimization, and in an attempt to improve the final authentication performance, we further propose and develop a hybrid algorithm for fusion.

Chapter 5

It concludes the dissertation by a summary of our work and contributions and discussion of directions for future work and recommendations.



Multibiometric Systems and Fusion

This chapter gets us acquainted with the field of biometrics and the limitations of unimodal systems. It introduces multibiometric systems, their advantages and the different system architectures and levels.

2.1. Introduction

Biometric authentication refers to the identification of humans by their physiological, chemical or behavioral characteristics. Historically, biometric science is not a new science as its first appearance was in the 19th century, introduced by a French police officer, Alphonse Bertillon [2], who invented a number of anthropomorphic measurements, called Bertillonage, for identifying criminals. His system was built on the assumption that the body of people do not change in basic characteristics. Bertillon's system involved measuring five primary measurements of body parts such as head length; head breadth; length of the middle finger and the length from elbow to end of middle finger (see *Fig.2.1*).

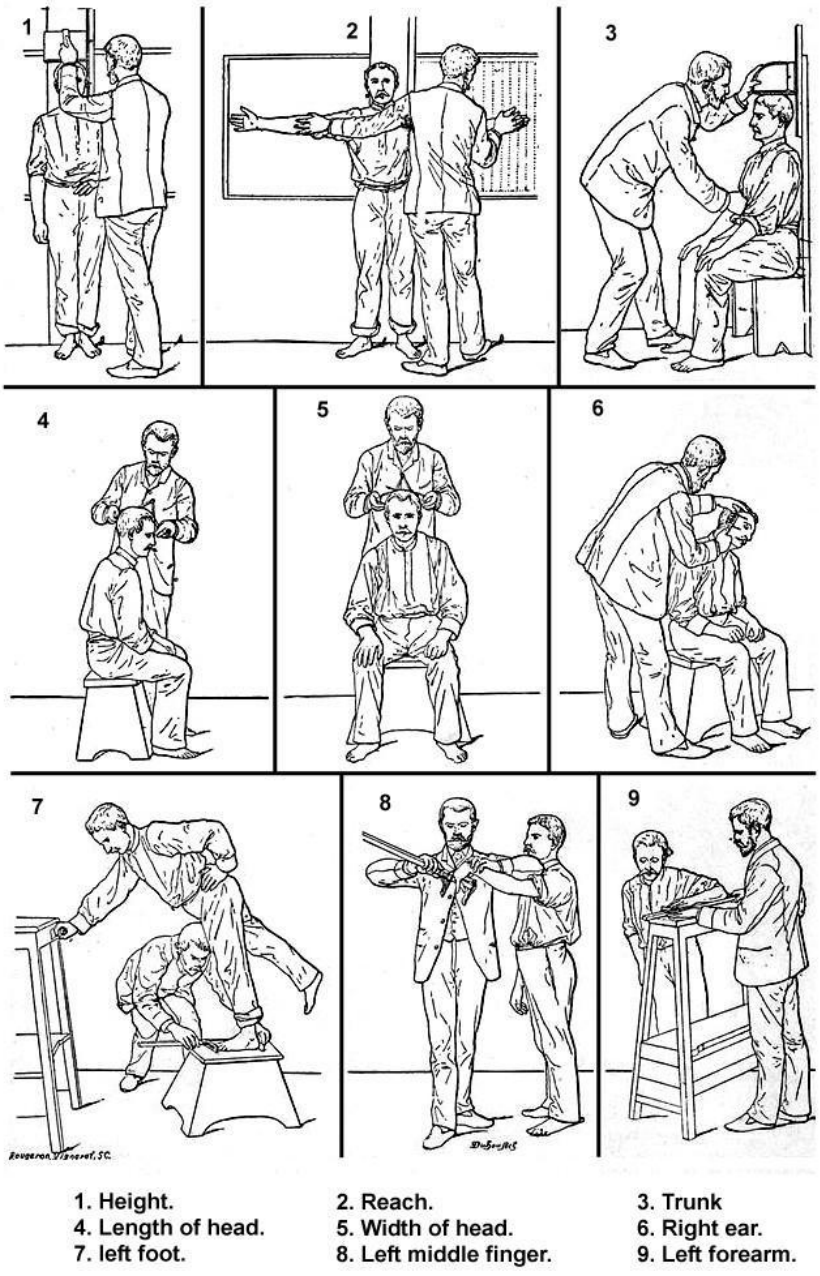


Fig.2.1. A chart from Bertillon's Identification anthropométrique (1893) demonstrating how to take measurements for his identification system [2]

As of recently, biometrics gained quite a lot of interest, especially in the research community. That is due mainly to the unreliability and inconvenience of traditional authentication systems. Biometrics offer a natural and reliable solution to certain aspects of identity management by utilizing fully automated or semi-automated schemes to recognize individuals based on their inherent traits. By using biometrics, it is possible to establish an identity based on *who you are*, rather than by *what you possess*, such as an ID card, or *what you remember*, such as a password [1].

The importance of biometrics in our society has been reinforced by the need for large-scale identity management systems which functionality relies on the reliable determination of an individual's identity in the context of several different applications. Examples of these applications include [3]:

- Sharing networked computer resources
- Granting access to nuclear facilities
- Performing remote financial transactions
- Boarding a commercial flight
- Web-based services (e.g., online banking)
- Customer service centers (e.g., credit cards)

2.1.1. Biometric characteristics

Many different aspects of human physiology, chemistry or behavior can be used for biometric authentication. The selection of a particular biometric for use in a specific application involves a weighting of several factors. Seven such factors have been identified to be used when assessing the suitability of any trait for use in biometric authentication [4].

- **Universality** means that every person using a system should possess the trait.
- **Uniqueness** means the trait should be sufficiently different for individuals in the relevant population such that they can be distinguished from one another.
- **Permanence** relates to the manner in which a trait varies over time. More specifically, a trait with 'good' permanence will be reasonably invariant over time with respect to the specific matching algorithm.
- **Measurability** relates to the ease of acquisition or measurement of the trait. In addition, acquired data should be in a form that permits subsequent processing and extraction of the relevant feature sets.
- **Performance** relates to the accuracy, speed, and robustness of technology used.
- **Acceptability** relates to how well individuals in the relevant population accept the technology such that they are willing to have their biometric trait captured and assessed.

- **Circumvention** relates to the ease with which a trait might be imitated using an artifact or substitute.

2.1.2. Biometric modalities

Biometric systems authenticate a user identity by the means of measuring an individual‘s unique features. Here, we briefly present some of the well recognized modalities which, depending on their nature, belong to one of three categories: Physiological, Behavioral and Chemical.

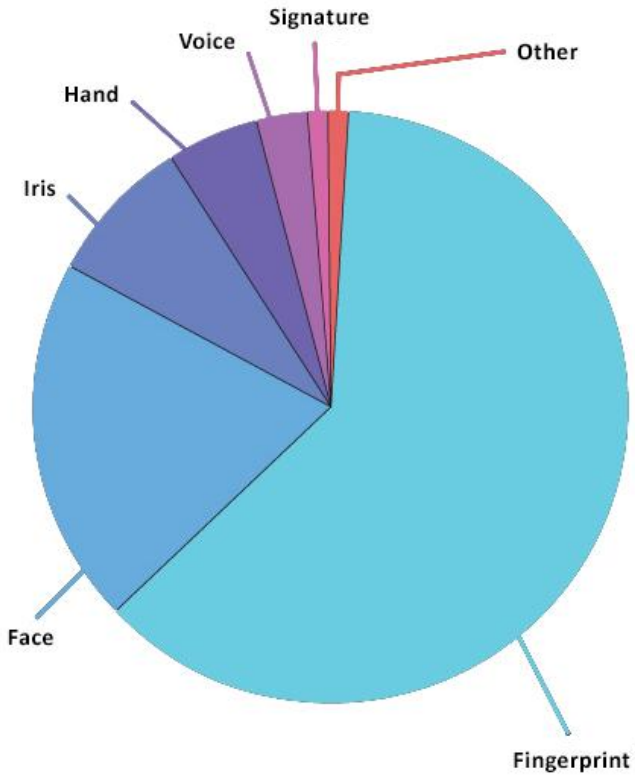


Fig.2.2. Worldwide Biometric Marketing (Revenue by technology 2009)

2.1.2.1. Physiological

i. Face

Face recognition is an active area of research with many applications. It consists of the analysis of facial features or patterns for the authentication or recognition of an individual’s identity. During the past 25 years, a substantial amount of research effort has been devoted to face recognition. A number of face-recognition techniques are widely popular, including:

- Principle Component Analysis (PCA) [5], [6]
- Linear Discriminant Analysis (LDA) [7]
- Singular Value Decomposition (SVD)
- A variety of Neural Network-based techniques [8]

The performance of these approaches is quite impressive and is sufficiently mature that they can be ported to real-time experimental/demonstration systems. Generally, there are two major tasks in face recognition:

a. Locating faces in input images
b. Recognizing the located faces

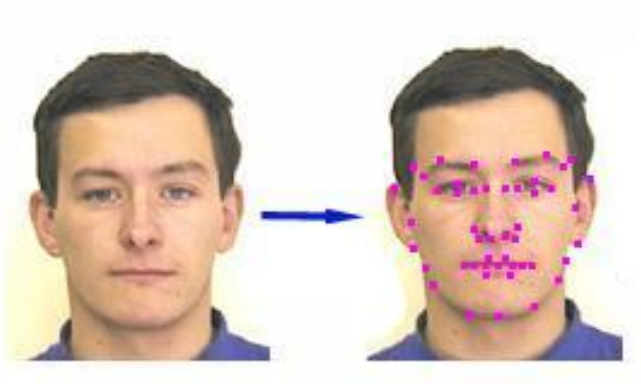


Fig.2.3. Face Modality

ii. Fingerprint

Fingerprint is the oldest form of biometrics. Archaeological evidence exists as a proof that ancient Assyrians and Chinese had used fingerprints as a form of identification. It is also the most widely used in today's biometric authentication systems, as is clearly seen in *Fig.2.3*

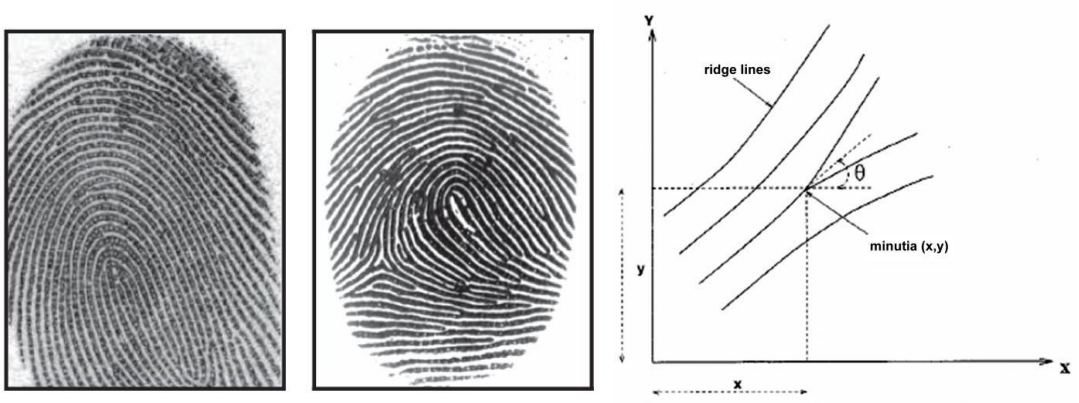
Finger ridges configurations are unique to a person and they do not change throughout the lifetime of an individual except due to accidents. Also, fingerprint recognition has a very low error rate which makes them very desirable in the biometric field.

Fingerprint matching generally depends on the comparison of local ridge characteristics and their relationships. The two most prominent ridge characteristics are called minutiae. A fingerprint typically contains about 40 to 100 minutiae.

Examples of minutiae are shown in *Fig.2.4.* For a given fingerprint, a minutia can be characterized by its type, its x and y coordinates, and its directionθ, which definitions are also shown in *Fig.2.4.*

Fingerprint verification consists of two main stages [9], [10]:

a. Minutiae extraction
b. Minutiae matching



(a) (b) (c)

Fig.2.4. Fingerprint (a)(b) Examples (c) Minutiae Representation

iii. Iris

Some biometric systems use the features found in the iris to recognize individuals. The iris is the annular region of the eye bounded by the pupil and the sclera (white of the eye) on either side. The complex iris texture carries very distinctive information useful for personal recognition. Each iris is distinctive and even the irises of identical twins are different. Choosing the iris as a modality for authentication is promising and has lower rates of error [11].

Fig.2.5. Iris Modality

iv. Hand geometry

Hand geometry recognition systems are based on a number of measurements taken from the human hand, including its shape, size of palm, and the lengths and widths of the fingers. The technique is very simple, relatively easy to use, and inexpensive.

Environmental factors such as dry weather or individual anomalies such as dry skin do not appear to adversely affect the authentication accuracy of hand geometry-based systems. However, the geometry of the hand is not known to be very distinctive and hand geometry-based recognition systems cannot be scaled up for systems requiring identification of an individual from a large population. This, among the fact that hand scanners are big and the

difficulty to extract the correct hand geometry information due to an individual's jewelry or limitations in dexterity, makes this technique less used [1].

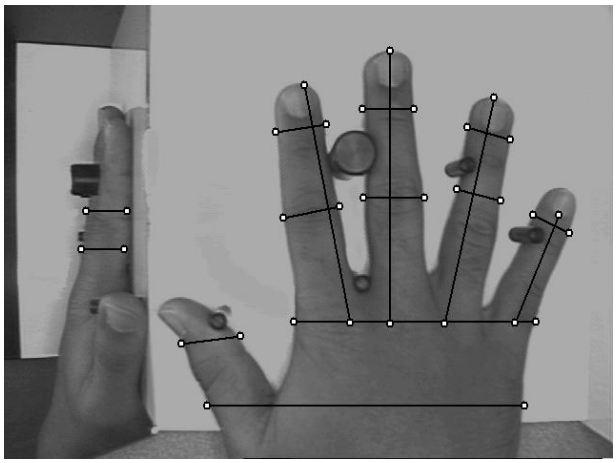


Fig.2.6. Hand Geometry Modality

v. Retina

Research conducted in the 1930s suggested that the patterns of blood vessels in the back of the human eye were unique to each individual [12].

The retina is a thin layer of cells at the back of the eyeball of vertebrates. It is the part of the eye which converts light into nervous signals. The principle of retina biometrics consists of capturing and analyzing the patterns of blood vessels on the thin nerve on the back of the eyeball that processes light entering through the pupil.

Retinal patterns are highly distinctive traits. Every eye has its own unique pattern of blood vessels; even the eyes of identical twins are distinct [12]. Although each pattern normally remains stable over a person's lifetime, it can be affected by disease such as glaucoma, diabetes, high blood pressure, and autoimmune deficiency syndrome.

Fig.2.7. Retina Scan

vi. Hand vein

Vein recognition is a type of biometrics that can be used to identify individuals based on the vein patterns in the human finger or palm.

Though used by law enforcement agencies, this method of identification is still in development and has not yet been universally adopted by crime labs as it is not considered as reliable as more established techniques, such as fingerprinting. However, it can be used in conjunction with existing forensic data in support of a conclusion [13].

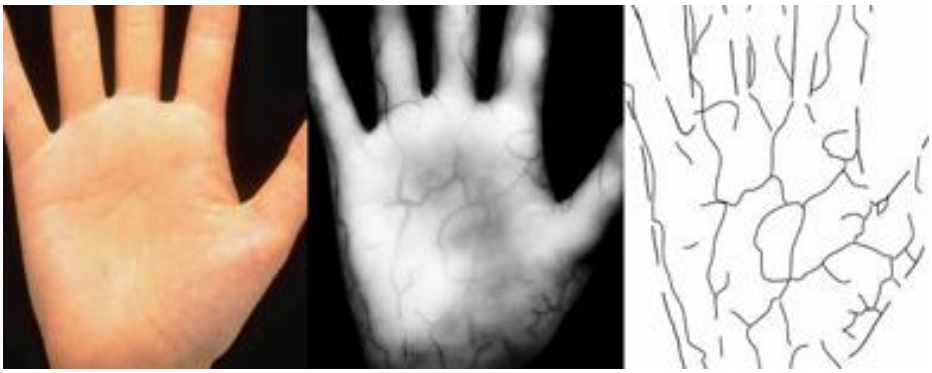


Fig.2.8. Hand Vein Modality

2.1.2.2. Behavioral

i. Voice

Voice recognition, also known as Speaker recognition, is the identification of the person who is speaking by characteristics of their voice.

Strictly speaking, *voice* is also a physiological trait because every person has a different vocal tract, but voice recognition is classed as behavioral as it is affected by a person's mood [13].

Biometric voice recognition, which recognizes *who is speaking*, is separate and distinct from speech recognition, recognizing *what is being said*. These two terms are frequently confused, as is voice recognition. Voice recognition is a synonym for speaker, and thus not speech recognition.

Fig.2.9. Voice Modality

ii. Gait

Gait refers to the manner in which a person walks, and is one of the few biometric traits that can be used to recognize people at a distance. Therefore, this trait is very appropriate in surveillance scenarios where the identity of an individual can be surreptitiously established. Gait-based systems also offer the possibility of tracking an individual over an extended period of time. However, the gait of an individual is affected by several factors including the choice of footwear, nature of clothing, affliction of the legs, walking surface, etc [1].

Fig.2.10. Gait Modality

iii. Keystroke

The behavioral biometric of Keystroke Dynamics uses the manner and rhythm in which an individual types characters on a keyboard or keypad. The keystroke rhythms of a user are measured to develop a biometric template of the users typing pattern for future authentication.

It is hypothesized that each person types on a keyboard in a characteristic way. This biometric is not expected to be unique to each individual but it may be expected to offer sufficient discriminatory information to permit identity verification [14].

This technique works by monitoring the keyboard inputs at thousands of times per second in an attempt to identify the user by his/her habitual typing rhythm patterns

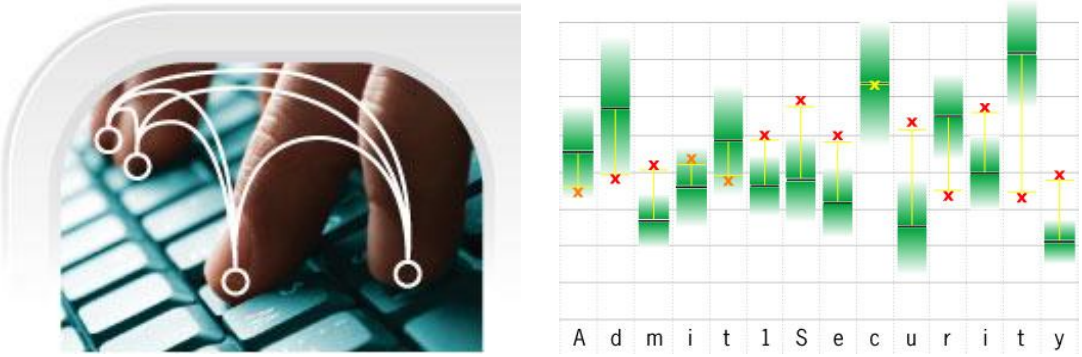


Fig.2.11. Keystroke Dynamics

iv. Signature

The way a person signs his or her name is known to be a characteristic of that individual [15]. This technique has been accepted in government, legal, and commercial transactions as a method of authentication for a while now.

Signature is a behavioral biometric that changes over a period of time and is influenced by the physical and emotional conditions of the signatories. Signatures of some people vary substantially.

Fig.2.12. Signature recognition

2.1.2.3. Chemical

To this point, modalities have been presented and categorized on whether their recognition depends on the physical aspect or the behavior of individuals. This last part deals with a modality that is of the chemical nature: the DNA.

i. DNA

The cells that contain DNA share genetic information through chromosomes. Humans have 23 chromosomes pairs that house a person's DNA and their genes. 0.3% of an individual's DNA is variable repetitive coding unique to an individual. This repetitive coding is the basis of DNA biometrics. DNA recognition uses genetic profiling, also called genetic fingerprinting, to isolate and identify these repetitive DNA regions that are unique to each individual to either identify or verify a person's identity.

The basic steps of DNA profiling include [16]:

a. Isolate the DNA (sample can originate from blood, saliva, hair, semen, or tissue)
b. Section the DNA sample into shorter segments containing known variable number tandem repeats (VNTRs)—identical repeat sequences of DNA
c. Organize the DNA segments by size
d. Compare the DNA segments from various samples

The more repeats of sequences there are for a given sample, the more accurate the DNA comparison will be, thus decreasing the likelihood of the sample matching multiple individuals. A few drawbacks of this technique are the depth of the procedure, the physical invasiveness of obtaining the DNA sample, and the time required to perform a DNA comparison. Also contamination of the sample renders the comparison impossible.

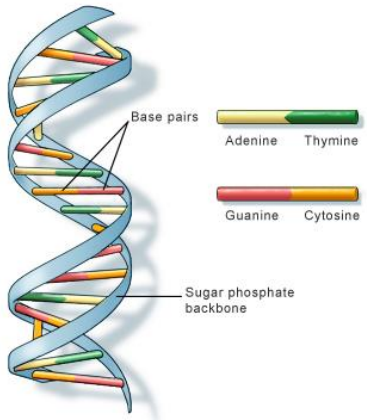


Fig.2.13. A DNA Code

➢ Whereas some biometrics has gained more acceptance then others in range of applications, it is beyond doubt that using biometrics has gained a measure of acceptance. Nevertheless, each biometric modality has its strengths and limitations, and no single biometric modality is likely to meet all the desired performance of every authentication applications.

Biometric Modality	*Universality*	*Uniqueness*	*Permanence*	*Collectability*	*Performance*	*Acceptability*	*Circumvention*
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial thermogram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand geometry	M	M	M	H	M	M	M
Hand vain	M	M	M	M	M	M	L
Voice	M	L	L	M	L	H	H
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Palm-print	M	H	H	M	H	M	M
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Iris	H	H	H	M	H	L	L

Table.2.1. Comparison of biometric modalities based on biometric characteristics (**H** = high, **M** = medium, **L** = low)

2.1.3. Biometric structure

Generally, any typical authentication biometric system comprises four processing modules: Data acquisition, Feature Extraction, Matching Unit and Decision Module.

- *Data acquisition unit* consists of acquiring the biometric signal with a special sensor and converting it to a digital form.
- *Feature extraction unit* extracts key information from the digital representation of the biometric cue.
- *Matching unit* matches extracted features with templates stored in a database and outputs a similarity measure.
- *Decision making unit* issues a binary decision whether to accept or reject the claimed identity.

Before any recognition can be achieved, a necessary primary step must be completed: Enrollment (*Fig.2.14*)

In order to access the biometric system, the user has to be registered. In this stage, an image of the specific biometric trait is captured and an ID assigned to it. This image is then converted to a template, after going through the feature extraction process, to be stored in a template database.

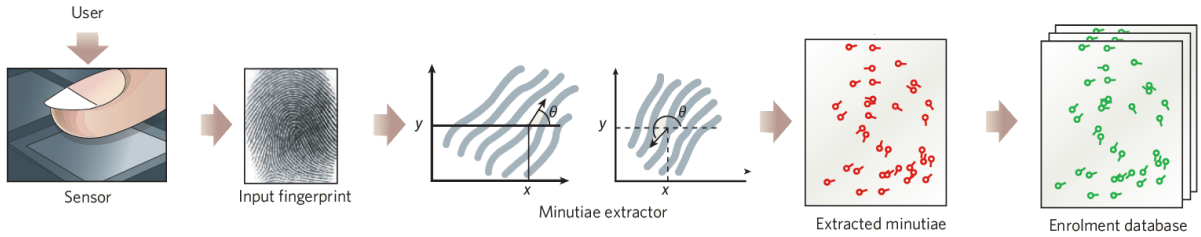


Fig.2.14. Biometric Enrollment

2.1.4. Biometric recognition modes

Depending on the purpose behind its usage, biometrics can be used for identification or for verification.

2.1.4.1. Verification mode

In this mode, the system validates a person's identity by comparing the captured biometric data with their own biometric template(s) stored in the system database. In such a system (*Fig.2.15*), an individual who desires to be recognized claims an identity, usually via a PIN, a user name or a smart card, and the system conducts a **one-to-one (1:1)** comparison to determine whether the claim is true or not.

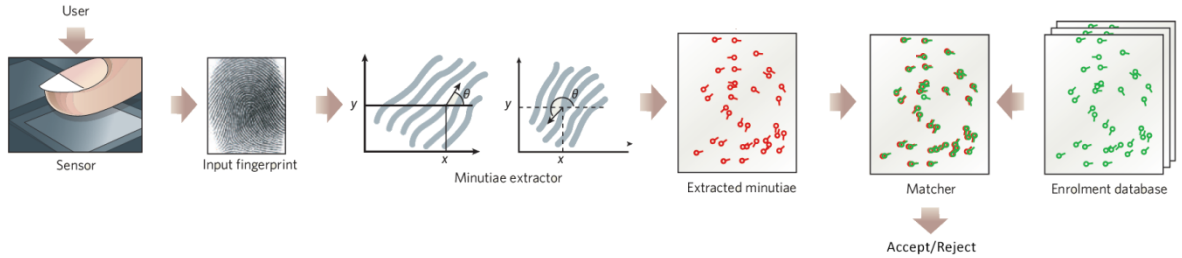


Fig.2.15. Biometric Verification Process

The verification problem may be formally posed as a two-category classification problem as follows [1]:

Given an input (query) feature set X_Q and a claimed identity I_d, determine if (I_d, X_Q) belongs to *gen* or *imp*, where *gen* indicates that the claim is a "genuine" user and *imp* indicates that the claim is an "impostor". Typically, X_Q is matched against X_T, the stored biometric template corresponding to user I_d, to determine its category. The resulting decision rule is,

$$(I_d, X_Q) \in \begin{cases} gen & if\ S(X_Q, X_T) \geq t_s \\ imp & otherwise \end{cases}$$

Where S is the function that computes the similarity between X_Q and X_T and produces a similarity score $S(X_Q, X_T)$.

t_s is a predefined decision threshold. Both terms will be discussed thoroughly in upcoming chapters.

2.1.4.2. Identification mode

In the identification mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Therefore, the system conducts a **one-to-many (1:N)** comparison to establish an individual's identity, or fails if the subject is not enrolled in the system database, without the subject having to claim an identity.

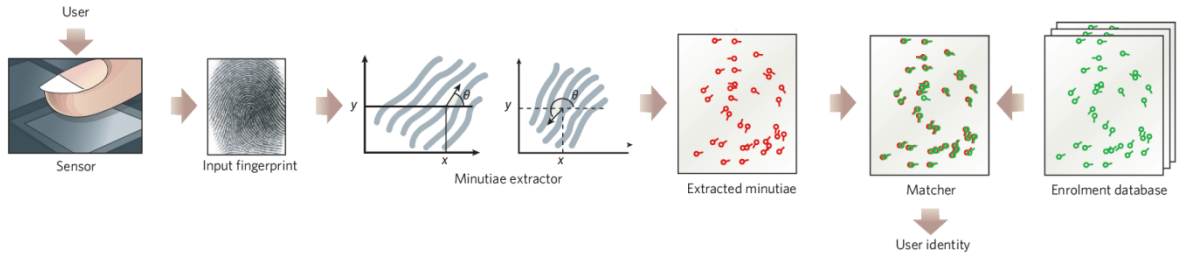


Fig.2.16. Biometric Identification Process

Given an input feature set X_Q, the identification will determine the identity I_K, $k = \{1, 2, \dots, M, M+1\}$ where $\{1, 2, \dots, M\}$are the M identities enrolled in the system, and $\{M+1\}$ indicates the reject case where no suitable identity can be determined for the input [1].

Hence,

$$X_Q \in \begin{cases} I_K & if \ \max_K \{S(X_Q, X_{T_K})\} \geq t_s \\ I_{M+1} & otherwise \end{cases}$$

2.2. Multibiometric systems

A system that consolidates the evidence presented by multiple biometric sources is known as a *multibiometric* system. Multibiometric systems are those which use, or are capable of using, more than one physiological or behavioral characteristic for enrolment either in verification or identification mode.

It is generally believed that by integrating various biometric traits into one single unit, the limitations of unibiometric systems can be alleviated, given that several biometric sources usually compensate for the weaknesses of single biometric. These inherent limitations can be alleviated by fusing the information presented by multiple sources.

For example, the face and gait traits, or multiple images of the face, or the fingerprints of the right and left index fingers of an individual may be used together to resolve the identity of an individual. Fusion in biometrics helps "expand" the feature space used to represent individuals. This increases the number of people that can be effectively enrolled in a certain personal identification system.

2.2.1. Limitations of unimodal

The majority of currently-in-use biometric systems exploit a single biometric trait. Such systems are called unibiometric systems. Regardless of significant scientific and technological advances in the latest years, there are still several limitations derived from using one biometric trait. Some of these limitations are listed below.

a. Noisy Sensor data

An example of noisy data can be in the form of an image of a scarred fingerprint or a voice sample altered by cold. Noisy data may also result from defective or improperly maintained sensors (e.g., accumulation of dirt on a fingerprint sensor) or unfavorable ambient conditions (e.g., poor illumination of a user's face in a face recognition system). Noisy biometric data may not be successfully matched with corresponding templates in the database, resulting in a genuine user being incorrectly rejected.

b. Intra-class variation

The biometric sample obtained from a user throughout the identification or verification phase is not identical to the sample which was collected to generate the reference database from the same user during the enrolment phase. This is known as the "intra-class" variations.

These variations may be to inappropriate interactions of the user with the sensor, as in the case when a user changes his or her pose or facial expression in front of a camera. This can happen when using different sensors at enrolment and verification or due to alterations in the biometric modality, such as the case of developing new wrinkles in face or the presence of new scars in a fingerprint. Intra-class variations are more relevant in behavioral biometrics traits such as voice and signature (see *Fig.2.17*).

Consequently, individuals with large intra-class variability will regularly be falsely rejected as the acquired biometric trait they present does not match with any of the biometric template that they had enrolled with.

Fig.2.17. Three Samples of Signatures from One Session

c. Universality

A biometric modality is called universal as long as every subject of a target population is capable of presenting a valid biometric sample for authentication. This principle of universality is an essential condition in any efficient biometric recognition implementation. However, all biometric modalities are not really universal. One example is of persons who suffer from a particular handicap.

The National Institute of Standards and Technology (NIST) has reported that it was not possible to acquire a good quality fingerprint from about 2% of the population (for instance, people with disabilities related to the hand, people with oily or dry fingertips, etc.) [17]. Consequently, such people cannot be signed up in a fingerprint verification system.

d. Distinctiveness

The biometric characteristics extracted from different persons may be quite similar. For instance, face recognition systems that depend on facial appearance fails in identifying identical twins. This short of distinctiveness usually increases the FAR of a biometric system.

e. Spoof attacks

Spoofing involves the deliberate manipulation of one's biometric traits in order to avoid recognition, or the creation of physical biometric artifacts in order to take on the identity of another person.

Many studies [18], [19] demonstrated that it is possible to spoof a number of fingerprint authentication systems using simple techniques with molds made from range of materials such as plastic, clay, silicon or gelatin (see *Fig.2.18*). As a matter of fact, behavioral biometric modalities, such as voice or signature, are more susceptible to this kind of attacks than physiological biometric modalities.

Spoof attacks, when successful, can severely undermine the security afforded by a biometric system.

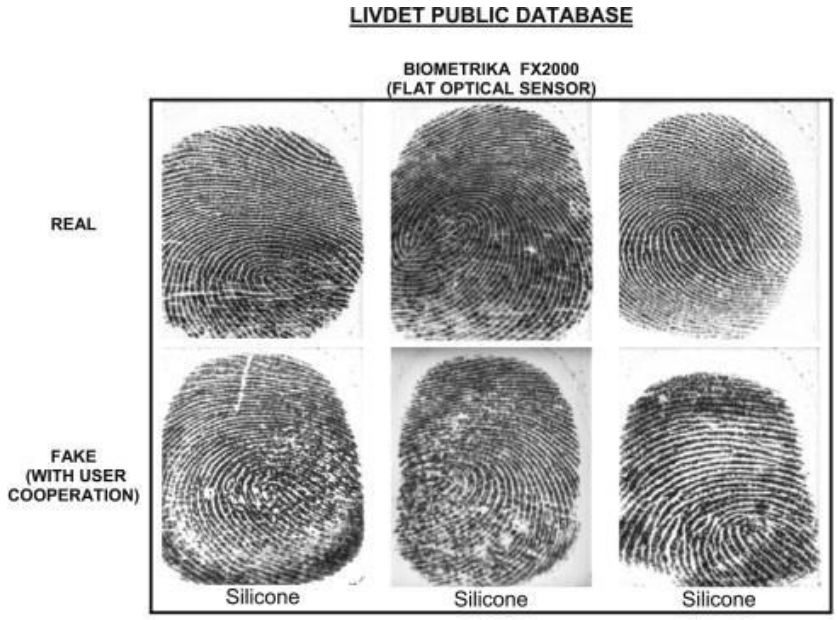


Fig.2.18. Real and Faked Fingerprints from a Public Database Used In a Recent Research [20].

2.2.2. Advantages of Multibiometrics

We established that due to all the practical difficulties in relation to using a unimodal system, the error rate associated with unimodal systems is relatively high. This makes unimodal-based authentication techniques improper for deployment in safety-critical or real-time applications. Some of the aforementioned drawbacks can be overcome by considering a multimodal biometric approach. Multibiometric systems offer the following advantages over unibiometric systems:

i Using an efficient fusion method to combine evidences from different sources can considerably improve the overall accuracy of the authentication system.

ii Multibiometric systems are capable of addressing the problems related to non-universality that unimodal biometrics suffer from. For instance, if someone‘s voice is altered by cold, they cannot be enrolled in a voiceprint recognition system, but they can still be identified using other biometric traits like fingerprint or palm print.

iii Multibiometric systems can add more flexibility to the enrolment procedure during user authentication. Lets us suppose a hypothetical access control application built using the modalities of face, voice and fingerprint. Later on, at the time of authentication, the user has the flexibility to choose all or a subset of available biometrics. This is convenient for users with special needs, users with hand-related disabilities, for example, can enroll to the same system with their voice sample.

iv The noisy data, which usually have a considerable effect on the performance of the authentication process, can be considerably reduced with the availability of multiple sources of information. In such case, if the user failed to enroll using one of the sources due to acquisition conditions, another biometric source with lower error rates will adjust the balance to have a better performance.

v Multimodal systems are more resistant to fraudulent techniques since it is not easy for an imposer to forge several biometric traits at the same time. By asking the subject to present the biometric traits in a random order, the system can detect that the user is present at the acquisition point.

Nonetheless, multibiometric-based systems have some of their own drawbacks. Unfortunately, they are more expensive as they should require more computational and storage resources. In addition, they also require a large number of test samples and additional time for user enrolment which usually cause inconvenience to the user.

2.3. Multibiometric fusion sources

In multibiometric systems, various sources of biometric information can be used for fusion. Depending on the nature of these sources, we can classify multibiometric sources into four categories: multi-sensors, multi-samples, multi-algorithms, multi-modalities. *Table 2.2* illustrates the four categories by the case of using two of something.

Category	Modality	Algorithm	Biometric trait	Sensor
Multi Sensors	1 (always)	1 (usually)[a]	1 (always, and same instance)	2 (always)
Multi Samples	1 (always)	1 (always)	2 samples of 1 biometric trait(e.g. 2 fingerprints of the same finger)	1 (always)
Multi Algorithms	1 (always)	2 (always)	1 (always)	1 (always)
Multi Modalities	2 (always)	2 (always)	2 (always)	2 (usually)[b]

Table.2.2. Comparison between the different multi-biometric systems (categorized on the basis of sources) [21]

a: *It is possible that two samples from separate sensors are processed by using separate "feature extraction" algorithms, and then through a common comparison algorithm, making this "1.5 algorithms", or two completely different algorithms.*
b: *a multimodal system with a single sensor used to capture two different modalities (e. g, a high resolution image used to extract face and iris).*

2.3.1. Multi sensors

In multi-sensor systems a single biometric trait is captured using multiple sensors in order to extract diverse information. For example, Marcialis et al. [22] investigated the multi-sensor fingerprint system employing optical and capacitive fingerprint sensors. They demonstrated that the integration of sources provided by these two sensors substantially outperform the systems using either one of the sensor images.

Implementing multi-sensor verification system demands increase in system cost and user co-operation. Therefore, acceptability of such systems depends on the improvement in verification accuracy achieved compared to that of a single-sensor system.

In addition to improvement in performance accuracy, multi-sensor verification system has other advantages too. A single sensor is not equally suited to capture all type of modalities. Hence, use of more than one sensor also increases the coverage of user population. Moreover, the difficulty of presenting fake fingers or facial shots increases when multiple sensors are used as multiple sensors might require different fake samples. This helps in preventing fraudulent attempts.

Fig.2.19.a. Multiple Sensors

2.3.2. Multi samples

Multi-sample systems involve fusion of information from multiple samples within the same biometric modality. For example, evidence from the left and right irises or the left and right index fingers can be combined for the recognition of an individual. Multi-sample systems are particularly useful for the individuals whose biometric traits cannot be reliably captured due to inherent problems. For example, it might not be possible to acquire sufficient features when the skin is very dry. In such cases, combining information obtained from fingerprints of multiple fingers of an individual provides with better discriminatory information required for recognition. In the same way, different profiles such as the frontal profile, left profile and right profile of a face can be fused to address challenges arising from variations in facial pose.

Multi-sample systems generally do not require additional sensors and also do not necessitate new feature extraction and matching algorithms. However, in some applications a new sensor arrangement might be required to capture various instances simultaneously. This type of system is necessary in applications where the size of database is very huge.

Fig.2.19.b. Multiple Samples

2.3.3. Multi algorithms

Multi-algorithm systems process the same biometric sample using multiple algorithms. They can use multiple feature sets extracted from the same biometric sample or multiple matching schemes operating on a single feature set. Jain et al. [23], integrate the evidence obtained from three different fingerprint matchers in order to improve performance of the proposed fingerprint verification system. The matching algorithms used are Hough transform based matching, string distance based matching and dynamic programming based matching.

However, the authors point out that integrating matching algorithms do not guarantee improved performance. Factors such as correlation between the matching algorithms used, disparity in the efficiencies of those algorithms and the fusion technique employed impact the performance that can be achieved with fusion [1].

These systems employ a single sensor and hence reduce the cost as well as avoid the need for users to interact with multiple sensors. However, multi-algorithm systems require additional feature extractor modules or matching modules.

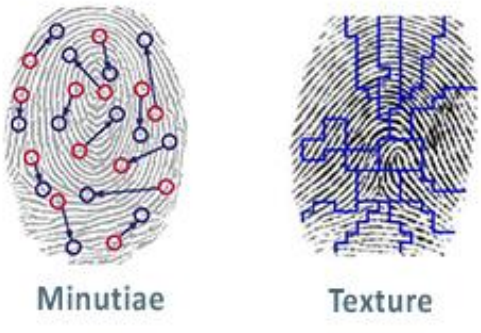


Fig.2.19.c. Multiple Algorithms

2.3.4. Multi modalities

Multimodal systems combine two or more different biometric modalities for establishing identity. There are some problems in deploying multimodal systems. Cost and complexity of added sensors and the appropriate user interfaces are increased. It is also more difficult to control the acquisition environment simultaneously for several traits.

On the other hand, Multimodal systems have several advantages. Better recognition rates can be achieved combining different modalities. They provide very high protection against spoofing as it is quite difficult for an imposter to spoof more than one biometric trait simultaneously. Multimodal systems also address the problems of noisy data. Even if one input is very noisy, input from other biometric trait might aid in recognition process.

It is good to note that higher performance improvement can be expected by using physically uncorrelated traits (e.g., fingerprint and iris) than using correlated traits (e.g., voice and lip movement) [1].

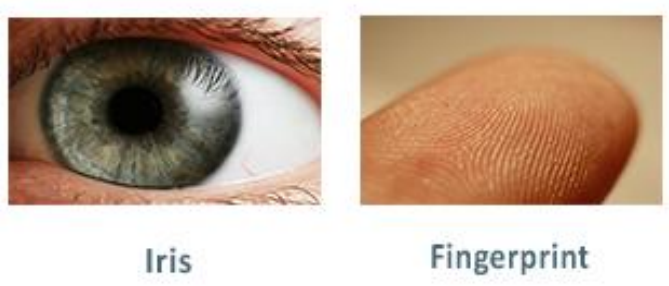


Fig.2.19.d. Multiple Modalities

2.4. Multibiometric fusion levels

Fusion in multibiometric systems can be performed at different stages of the biometric verification/identification system. It can take place at one of these levels: sensor level, feature level, score level, or decision level (*Fig.2.21*).

It is important to determine the type of information that should be consolidated during the fusion process. The amount of information available decreases after each level of processing in the different biometric system modules (*Fig.2.20*). The raw data at the sensor level represents the richest source of information whereas the final decision at the decision module just contains an abstract level of information.

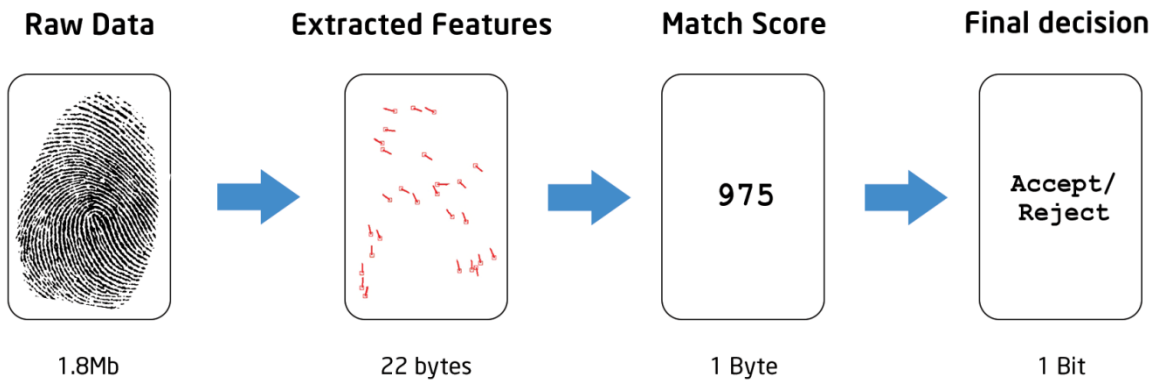


Fig.2.20. Amount of Data throughout the Recognition System Units [1]

The various levels of fusion can be categorized as pre-classification or fusion-before-matching and post-classification or fusion-after-matching [24]. This categorization is based on

the fact that the amount of information available for fusion is drastically reduced once the matcher is invoked. Fusion before matching can take place either at the sensor level or at the feature level. Fusion at score level and decision level occur in the post-classification stage.

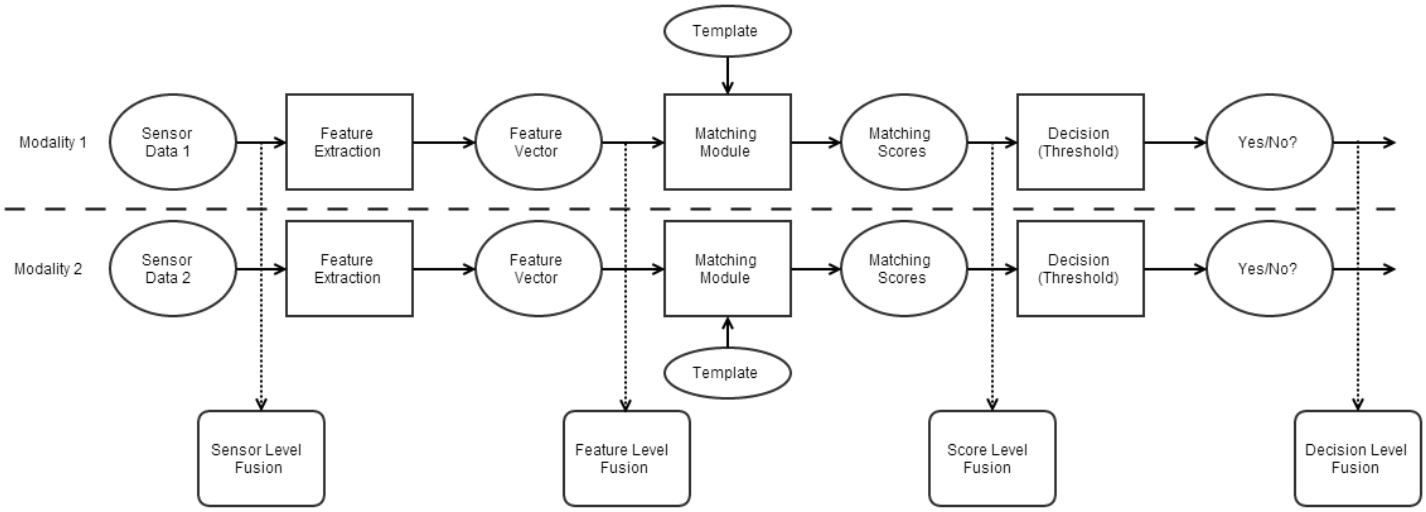


Fig.2.21. Multi-biometric Fusion Levels

2.4.1. Pre-classification

2.4.1.1. Sensor level

Sensor level fusion is the combining of data derived from sensory sources such that the resulting information is in some sense better than would be possible when these sources are used individually.

Sensor level fusion can benefit multi-sample systems which capture multiple snapshots of the same biometric. The technique mostly used is *mosaicing*, where a small fingerprint sensor may capture two or more impressions of a person's fingerprint and create a composite fingerprint image that reveals more of the underlying ridge structure as is the work of [25], [26]. Mosaicing has also been attempted by researchers in face recognition where multiple 2D images representing different poses are stitched to generate a single image [27], [28].

At this module, the data is at its rawest form and it represents the richest source of information. However, it is highly probable that raw data is contaminated by noise, for example, non-uniform illumination, background clutter, etc.

2.4.1.2. Feature level

In feature level fusion, feature sets originating from multiple information sensors are integrated into a new feature set. For non-homogeneous compatible feature sets, such as features of different modalities like face and hand geometry, a single feature vector can be obtained by concatenation. The new feature vector now has a higher dimensionality which

increases the computational load. It is reported that a significantly more complex classifier design might be needed to operate on the concatenated data set at the feature level space [29].

Ross and Govindarajan [30] discuss feature level fusion as applied to three different scenarios: (a) multi-algorithm, (PCA and LDA) (b) multi-sensor (three different color channels of a face image) (c) multimodal (face and hand geometry). Other examples of feature level fusion can be found in Son and Lee [31] (face and iris) and Kumar et al. [32] (hand geometry and palm-print).

The fusion at the feature level is expected to perform better in comparison with the fusion at the score level and decision level. The main reason is that the feature level contains richer information about the raw biometric data. However, because it is very challenging, such a fusion type is less studied in literature as compared to score and decision level and that is for the following reasons:

a. The feature vectors of multiple modalities might be incompatible, as is the case with the minutiae set of fingerprints and Eigen-coefficients of face.

b. The relationship between the feature spaces of different biometric systems may not be known.

c. Concatenation of two feature vectors might result in a feature vector with very large dimensionality leading to the curse-of-dimensionality problem. In such cases, when sufficiently large numbers of training samples are not available, increasing number of features might degrade system performance.

d. Most commercial biometric system vendors do not provide access to the feature sets.

e. More complex matchers might be required to operate on concatenated feature vectors.

2.4.2. Post-classification

2.4.2.1. Score level

In score level fusion, different matchers provide scores indicating the degree of similarity between the input and template vectors. These scores are combined in such a way to reach a better recognition decision. Because it is relatively easy to access information at this level and fuse the scores output by the different matchers, which in turns offers the best trade-off between accessibility and fusion convenience, this scheme is extensively studied in literature.

Snelick and al. [33] discuss the fusion of face and fingerprint at the score level. To achieve the fusion, they employed both the sum and weighted sum rules. The combination of scores for noisy speech and clean handwriting is the subject of investigations in [34]. Verlinde et al. [35] perform experiments and compare the performance of fusion using three different classifiers based on the k-nearest-neighbor (k-NN) classifier, decision trees and logistic regression.

Theoretically, it is guaranteed that the performance after combining matcher scores can be no worse than the performance of individual matchers. The idea is to identify the appropriate combination scheme that will fuse the score reliably and maximizes the matching performance [36].

It is worth noting that match scores generated by individual matchers might not be homogenous. For example, one matcher may produce a similarity score where a high value indicates better match whereas another matcher may produce a distance/dissimilarity score where a smaller value indicates better match. Another matter to consider is that the match scores generated from different matchers may not be in the same numerical range. For such reasons, in some cases, scores are usually normalized prior to fusion [1].

Fusion at score level can be generally categorized into two categories: classification based fusion and combination based fusion. These will be more discussed in the next chapter.

2.4.2.2. Decision level

Decision level fusion is performed using the decisions output by the biometric matcher components. Many Commercial Off-the-Shelf (COTS) biometric matchers provide only the final recognition result which is the match or non-match, accept/reject, decision. When such matchers are used in multibiometric systems, fusion is only possible at decision level.

To state a few, majority voting scheme [37] is a commonly used approach in decision level fusion where the input biometric sample is assigned to that identity on which the majority of matchers agree. When none of the identities is agreed by majority of matchers then “reject” decision is output by the system. This scheme assumes that performances of all matchers are similar. However, when the matchers used do not have similar recognition accuracy, it is reasonable to assign higher weights to more accurate matchers. This is done in weighted majority voting scheme.

For many reasons, one which is the reduced amount of information, fusion at such a level is considered to be the least powerful [38].

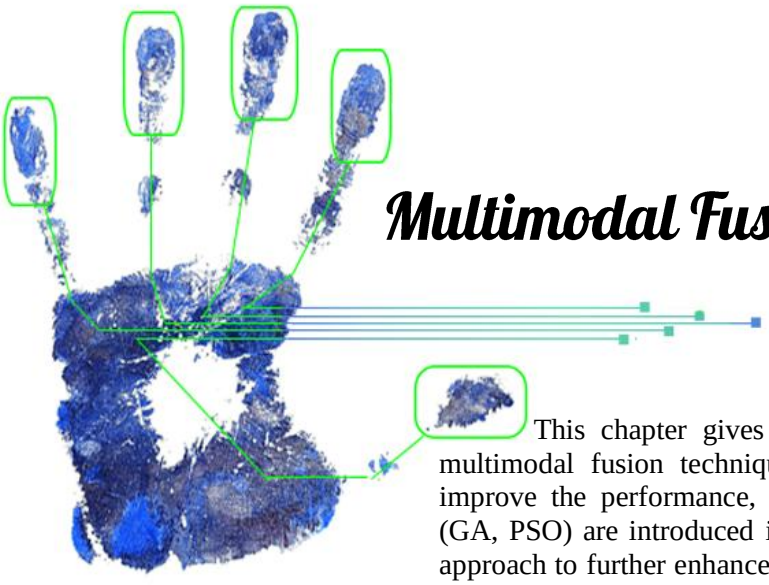
Conclusion

The field of study that this work focuses on, biometrics and multibiometric systems, has been introduced in this chapter. To have a grasp on the concept of biometrics, an overview of the characteristics of a biometric system as well as some of the widely used modalities was given. We also got acquainted with the two modes of a biometric authentication system: identification and verification.

Our investigations led to the conclusion that the unimodal systems have various limitations. We offered the multibiometric systems as a solution to elevate some of said limitations.

As the next step to understand multibiometrics better, the possible biometric sources to go through the fusion were discussed, in addition to the different levels at which the fusion can be performed.

As the scope of this study is narrow and we cannot possibly investigate all the possible scenarios, our work will focus on the fusion of multimodalities at the score level. The next chapter will discuss this model in more details.



Multimodal Fusion at Score Level

This chapter gives an overview of the state-of-the-art multimodal fusion techniques at the score level. In order to improve the performance, well-known optimization algorithms (GA, PSO) are introduced into the fusion. We propose a hybrid approach to further enhance the performance and accuracy of the system.

3.1. Introduction

As a consequence to the limitations of unimodal systems (authentication using a single modality) and the benefits that multibiometric systems (authentication using a combination of biometric sources) discussed in the previous chapter, recent research has been focusing more on multibiometric fusion techniques.

Since we answered the questions as to *what to fuse* (fusion sources) and *when to fuse them* (fusion levels), we now define the *how to fuse*. In here, we delve more into the multimodal fusion at score level.

As has been established in the previous chapter, fusion can be performed at different levels of the authentication system. It has been noted also that fusion at the score level is relatively better as compared to the others. For such reasons, there have been extensive studies in literature on score level fusion.

Fusion techniques are divided into two categories. Fusion can be performed as a classification problem or as a combination problem. The different rules will be introduced as well as some of the widespread optimization techniques.

Since the matching scores resulted from the various matchers are heterogeneous, score normalization is needed to convert them into the same nature, prior to combining them, so some well-known scores normalization methods will be also introduced in this chapter.

3.2. Similarity scores

In any biometric system, there is a bank of stored reference templates. These templates are taken and saved in a database when a user presents his or her biometric attributes for the first time (enrollment process). When a user tries to access the system for verification or identification, he will introduce another biometric sample, which is converted into a template and is then compared to the stored template.

To have a sense of whether we are dealing with a true user or an imposter, the matching module produces a score that quantifies the similarity between the input template and the reference template. In biometrics, two types of scores are acknowledged; distance scores and similarity scores.

A simple yet powerful way to determine similarity is to calculate the Euclidean Distance between two data objects p and q, where $\boldsymbol{p} = (p_1, p_2,..., p_n)$ and $\boldsymbol{q} = (q_1, q_2,..., q_n)$ are two points in Euclidean n-space, using the formula in equ.3.1.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{3.1}$$

This is what is called a distance score. The shorter the distance, the more similar the data objects are. To obtain similarity scores, the distances are subtracted from 1. Hence, the higher the value, the more similar the data objects are.

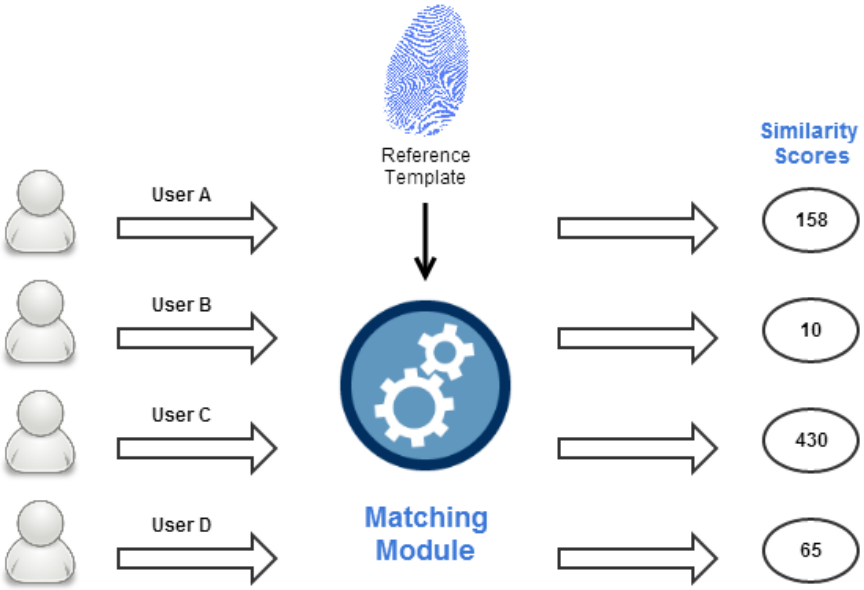


Fig.3.1. A Matcher Generating Similarity Scores

In the domain of databases, the data contained inside these databases is generated in the same manner. As an example, the NIST BSSR1 database (set 1), which will be described in more details in the next chapter, deals with 517 users but, all in all, it contains more than 26K scores. That is because of the generation of, not only genuine scores, but impostor scores too.

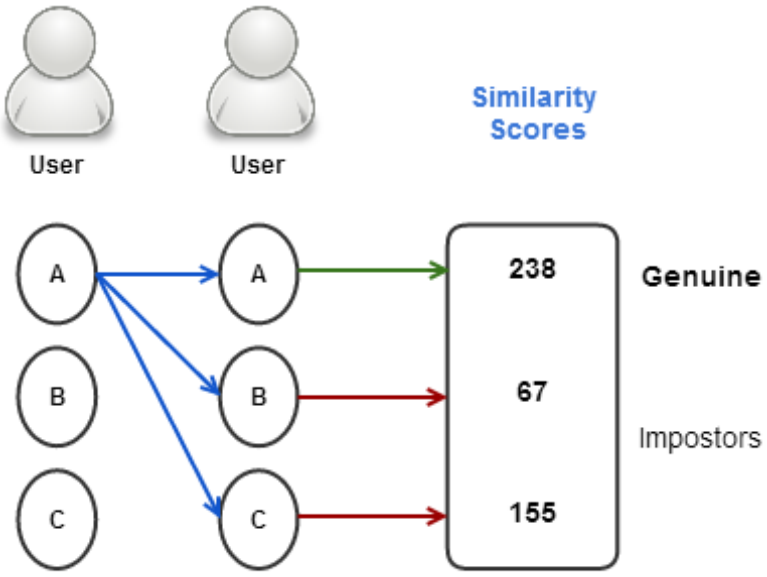


Fig.3.2 Dataset Score Generation

Each user template is compared to all the other templates, generating 1 genuine score and 516 imposter scores per user. *Fig.3.2.* demonstrates this idea. Here, template of User A is compared to the templates of Users A, B and C. The resulting scores represent similarity scores. As expected, the score with the highest value is generated from comparing User A to

User A template which makes it a genuine score. The others have lower values which in turn makes them imposter scores.

3.3. Score level fusion techniques

Multi-biometric fusion can be performed using different methods. Here, we provide an overview of these methods widely used now by multimedia researchers. Categorization of these methods depends on their nature inherently and they are divided into two categories: classification based fusion and combination based fusion.

3.3.1. Classification based fusion

This category of methods includes a range of classification techniques that have been used to classify the multimodal observation into one of the pre-defined classes.

In this approach, a multidimensional vector is formed of the similarity scores generated by several classifiers (matchers). This multidimensional vector is again provided to the classifier for final result. The final result is either “Accept” (genuine individual) or “Reject” (impostor individual).

If we take as an example, scores generated from speech, face, signature, and handwriting based recognition systems are 98, 80, 88, and 72, respectively, a vector {98, 80, 88, 72} is composed and classified as accept or reject. If the matching scores of the face and speech are x1 and x2, then a two-dimensional feature vector (x1, x2) is constructed.

Classifiers such as the support vector machine (SVM), Bayesian inference, Dempster–Shafer theory, dynamic Bayesian networks, neural networks or maximum entropy model are employed for classification. S. Ben-Yacoub and al. [39] report results in their work that SVM and Bayesian classifiers perform better as compared to the other classifiers.

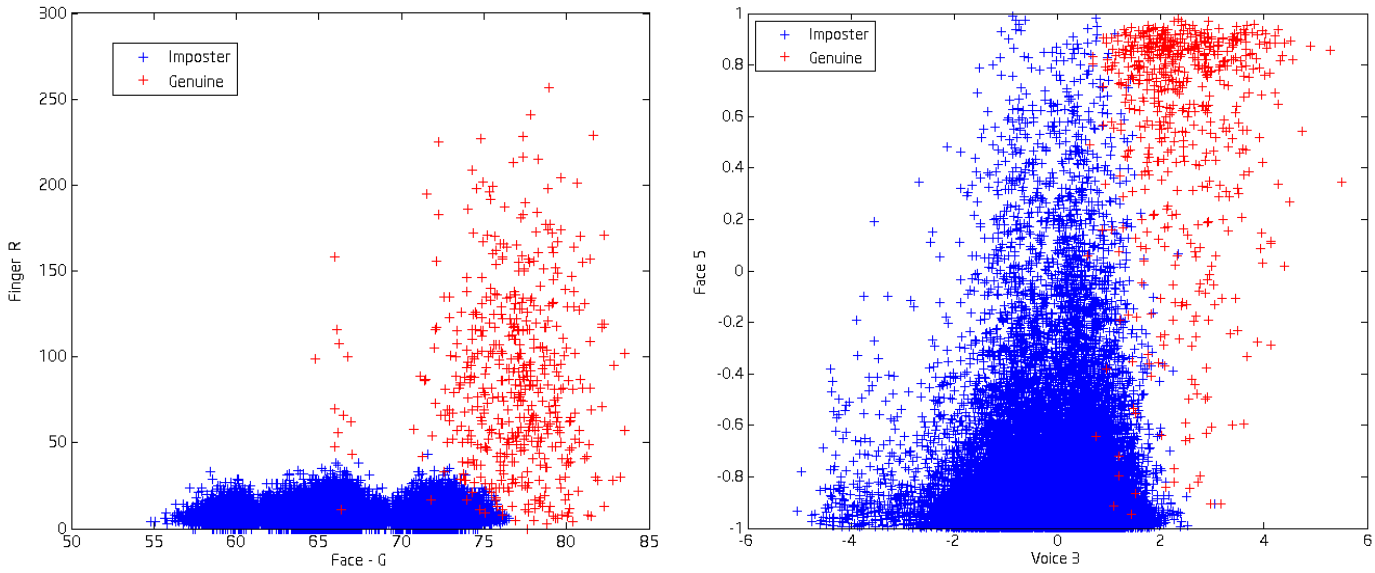
3.3.2. Combination based fusion

The rule-based fusion method includes a variety of basic rules of combining multimodal information. These include statistical rule-based methods such as sum simple, Max rule, Min rule, linear weighted fusion; product rule and weighted sum. With some of these methods, score normalization is required before fusion of scores.

3.3.2.1. Score normalization

Since the matching scores output by the various modalities are heterogeneous, score normalization is needed to transform these scores into a common domain, prior to combining them. For example, in the NIST dataset, the G-matching module for face provides scores that are in the range of {55, 85} and the right finger module in the range of {0, 255}, as is shown in *Fig.3.3.a*. This presents a problem. Depending on the combination rule, whether smaller or larger, the contribution of one score will be less significant than the other, resulting in the former having no effect on the fusion. This is why normalization is necessary, to transform the scores into equal magnitude, by changing their boundaries.

What follows is a list of the normalization techniques widely used in literature when multimodal biometric fusion is discussed.



(a) (b)

Fig.3.3. Score Distribution of (a) Finger – Face from the NIST Database [40] (b) Face – Voice from the XM2VTS Database [41]

It is important to note that a complete grasp of the intricacies of these rules is not mandatory, as that is outside the scope of this dissertation. We only acknowledge that they exist in multimodal techniques.

i. Min-Max

This is the simplest normalization technique. It is best suited for the case where the bounds (maximum and minimum values) of the scores produced by a matcher are known. In this case, we can easily shift the minimum and maximum scores to 0 and 1, respectively. However, even if the matching scores are not bounded, the minimum and maximum values can be estimated for a set of matching scores and then apply the min-max normalization.

Given a set of matching scores $S_i$the normalized scores $\tilde{S}_i$ are given by

$$\tilde{S}_i = \frac{S_i - min}{\max - min} \tag{3.7}$$

Min-max normalization retains the original distribution of scores except for a scaling factor and transforms all the scores into a common range $[0,1]$. Distance scores can be transformed into similarity scores by subtracting the min-max normalized score from 1.

ii. Z-score

The most commonly used score normalization technique is the *z-score* that is calculated using the arithmetic mean and standard deviation of the given data. The normalized scores are given by

$$\tilde{S}_i = \frac{S_i - \mu}{\sigma} \tag{3.8}$$

Where μ is the arithmetic mean and σ is the standard deviation of the given data.

Z-score normalization does not guarantee a common numerical range for the normalized scores of the different matchers. If the input scores are not Gaussian distributed, z-score normalization does not retain the input distribution at the output. This is due to the fact that mean and standard deviation are the optimal location and scale parameters only for a Gaussian distribution. For an arbitrary distribution, mean and standard deviation are reasonable estimates of location and scale, respectively, but are not optimal [36].

iii. Tanh-Rule

The *tanh-estimators* introduced by Hampel et al. [42] are robust and highly efficient. The normalization is given by

$$\tilde{S}_i = \frac{1}{2}\left\{\tanh\left(0.01\left(\frac{S_i - \mu_{GH}}{\sigma_{GH}}\right)\right) + 1\right\} \tag{3.9}$$

Where μ_{GH} and σ_{GH} are the mean and standard deviation estimates, respectively, of the genuine score distribution as given by Hampel estimators.

Anil Jain and al. [36] showed in their work that both min-max and z-score methods are sufficient techniques if the location and scale parameters of the matching scores (minimum and maximum values for min-max, or mean and standard deviation for z-score) of the individual modalities are known in advance, but they are very sensitive to outliers. On the other hand, tanh normalization method is both robust and efficient.

We have selected this normalization procedure because of its stability and it does not use impostors' patterns which can be hard or impossible to obtain in a real application.

The aim of this work is not to analyze the performance of biometric systems depending on the normalization procedure, but to present a new multibiometrics fusion procedure.

3.3.2.2. Classical combination rules

Now, let us suppose we have $i = 1, 2, 3, \dots, N$, where N is the number of users and $m = 1, 2, 3, \dots, M$, where $M$is the number of matchers. We denote S_i^mas the score generated by matcher m as a result for matching user i to the database template. The fused score for user i is denoted as$S_{f\,i}$.

i. Max Rule

The maximum rule chooses the matcher that is producing the higher probability. In other words, the maximum score S_i^m for user i is selected.

$$S_{f\,i} = \max_{m=1..M}(S_i^m) \tag{3.2}$$

ii. Min Rule

The minimum rule chooses the matcher that is producing the lowest probability. The minimum score S_i^m for user i is selected.

$$S_{f\,i} = \min_{m=1..M}(S_i^m) \tag{3.3}$$

The problem with using Max or Min rules is that they are very susceptible to noisy environments.

iii. Product Rule

The product rule fuses scores from different matchers by multiplying them following the equation

$$S_{f\,i} = \prod_{m=1}^{M} S_i^m \tag{3.4}$$

Although this rule gives good results, it is affected by wrongly estimated probabilities. One bad probability estimation (i.e. P=0) results in overall probability of 0 [43].

iv. Simple Sum

Scores generates by all matchers m for user i are summed to produce a fused score.

$$S_{f\,i} = \sum_{m=1}^{M} S_i^m \tag{3.5}$$

The simple sum rule performs better, compared to the previous rules, since no prediction estimation error occurs and it considers scores of all classifiers [43].

v. Weighted Sum

In this method, the information obtained from different matchers is combined in a linear fashion. To combine the information, one will assign weights w_m to different matchers.

$$S_{f_i} = \sum_{m=1}^{M} w_m \times S_i^m \tag{3.6}$$

Where $0 \leq w_m \leq 1$ and $\sum_{m=1}^{M} w_m = 1$.

A lot of focus in literature has been put on developing methods to assign the weights w_m to different matchers.

To mention a few of these methods, exhaustive search, also known as Brute Force Search (BFS), has been employed in previous works to search through a vector of possible weights and pick the ones that give the better performance. As can be expected, this method is very computationally expensive. For this reason, other methods have been used one of which is assigning weights to all the matchers based on their EERs. It is interesting to note that the weights for more accurate matchers will be higher than those of less accurate matchers.

3.3.2.3. Combination using Optimization Techniques

In this work, the focus is on finding the optimum weights for fusion by weighted sum. Optimization techniques, such as Genetic Algorithm and Particle Swarm Optimization, will be employed to try and define the optimum weights that produce better performance compared to classical rules.

I. Genetic Algorithm

Genetic algorithm (GA) is a well-known and frequently used evolutionary computation technique. This method was originally developed by John Holland [44] and his PhD students Hassan et al. [45]. The idea was inspired from Darwin's natural selection theorem which is based on the idea of the survival of the fittest. The GA is inspired by the principles of genetics and evolution, and mimics the reproduction behavior observed in biological populations.

In GA, a candidate solution for a specific problem is called an individual or a chromosome and consists of a list of genes. These chromosomes, and for the purpose of mathematical computations, are represented in some sort of encoding; Binary, real, characters, list of rules, etc. GA begins its search from a randomly generated population of designs that evolve over successive generations (iterations), eliminating the need for a user-supplied starting point. To perform its optimization-like process, the GA employs three operators to propagate its population from one generation to another.

1. **GA operators**
   - **a. Selection**

In the nature, the selection of individuals is performed by survival of the fittest. The more one individual is adapted to the environment; the bigger are its chances to survive and create an offspring and thus transfer its genes to the next population. In GA, the selection of the best individuals is based on an evaluation of fitness function; a function that needs to be maximized or minimized depending on the application. These individuals will be chosen so that they will be applied to the other genetic operations and will create the offspring population.

Some of the widely used selection methods, to state a few, are: Roulette-Wheel selection, Tournament Selection, Stochastic Universal selection, etc. To give a chance for bad individuals in the reproduction process, which Roulette-Wheel does not allow, Tournament selection was chosen for the experiments in this work.

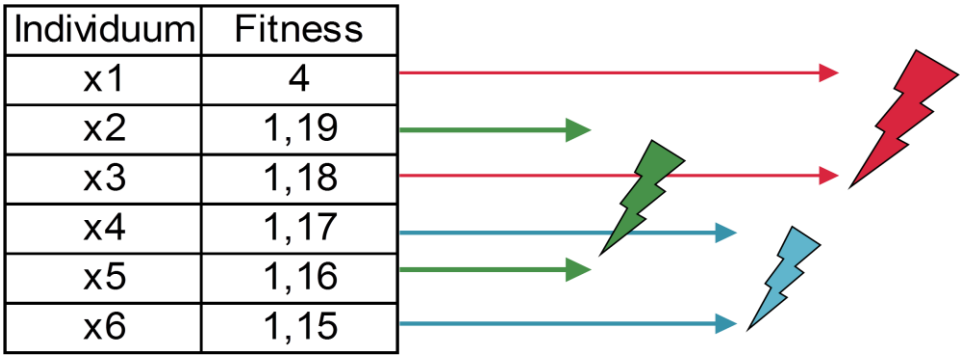


Fig.3.4. Tournament Selection with $k = 2$

In Tournament selection (*Fig.3.4.*), a randomly chosen subset of *k individuals* is picked and the individual in the subset with the best fitness is selected. This procedure is repeated until a population of size N is chosen to be applied to the genetic operators.

- **b. Crossover**

The first step in the reproduction process is the crossover, also known as recombination. The crossover operator mimics the mating process in biological populations. It propagates features of good surviving designs from the current population into the future population, which will have a better fitness value on average. In it, the genes of the parents are used to form entirely new chromosomes. There are many methods to achieve this. To state a few, there is the 1-point crossover; which, in itself is a special case of the n-point crossover (*Fig.3.5.*). This scheme is more suited for when the chromosomes are binary coded. When we have real encoded chromosomes, the arithmetic, heuristic and other crossovers are better suited.

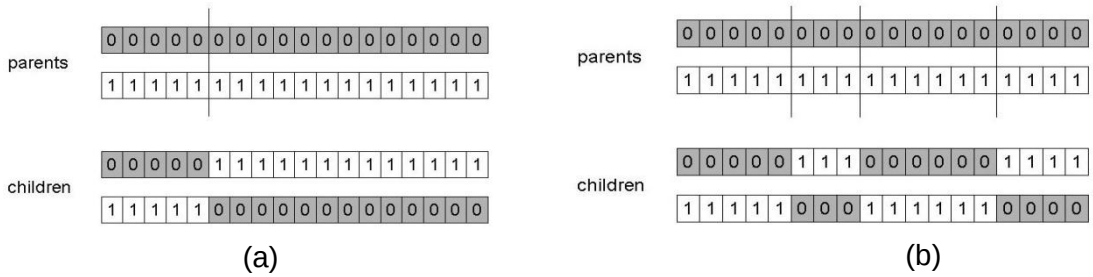


Fig.3.5. (a) 1-point (b) n-point Crossover Representation

In here, the arithmetic crossover, described by the formula in equ.3.10, is chosen.

$$Child_1 = r.Parent_1 + (1-r).Parent_2$$
$$Child_2 = (1-r).Parent_1 + r.Parent_2 \quad (3.10)$$

Where r is a randomly chosen number.

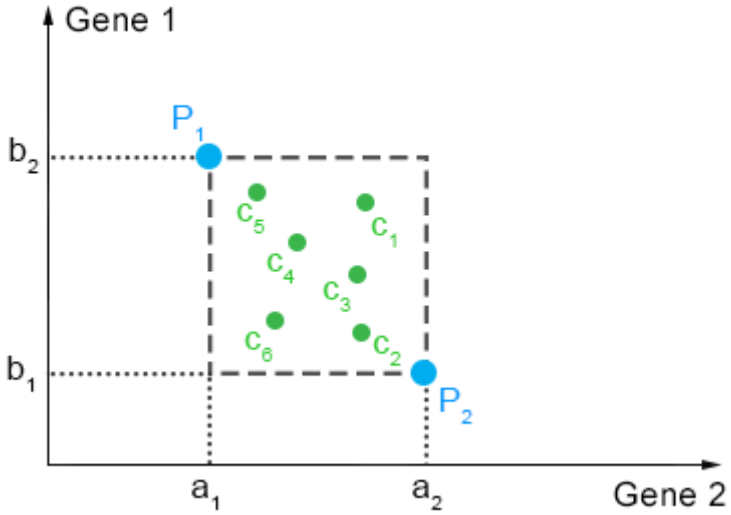


Fig.3.6. Graphical Representation of Arithmetic Crossover

In crossover, there is a probability P_C called *Crossover Probability*. It is a user-defined parameter that determines the size of the selected population that will go through the crossover process to generate children. Choosing a $P_C = 1$ means that all the population will go through crossover. A $P_C = 0$ means no individuals will reproduce, which defies the purpose of Genetic Algorithm.

c. Mutation

The newly created by means of selection and crossover population can be further applied to mutation. Mutation means that some elements of the DNA are changed. Those changes are caused mainly by mistakes during the copy process of the parent's genes.

In the terms of GA, mutation means random change of the value of a gene in the population (*Fig.3.7*). The chromosomes, which gene will be changed and the gene itself, are chosen by random as well.

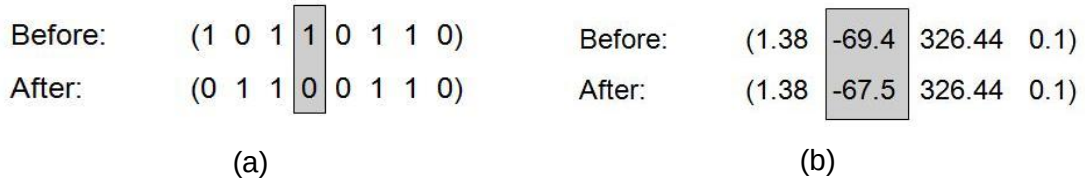


Fig.3.7. Mutation Operation on (a) Binary (b) Real Encoded Chromosomes

This operator promotes diversity in population characteristics. It allows for global search of the design space and prevents the algorithm from getting trapped in local minima.

Similarly to crossover, there is a probability P_m called the mutation rate. It is a user defined parameter, usually very small, that specifies how many times the mutation is performed.

Fig.3.8. summarizes all of the above operations in a numerical example.

Fig.3.8. Example of Genetic Algorithm operations

To illustrate the working process of genetic algorithm, the steps to realize a basic GA are listed:

1. Represent the problem variable domain as a chromosome of fixed length; choose the size of the chromosome population N, the crossover probability P_C and the mutation probability P_m.
2. Define a fitness function to measure the performance of an individual chromosome in the problem domain. The fitness function establishes the basis for selecting chromosomes that will be mated during reproduction.
3. Randomly generate an initial population of size N.
4. Evaluate the fitness of each individual chromosome.

5. Select a pair of chromosomes for mating from the current population. Parent chromosomes are selected with a probability related to their fitness.
6. Create a pair of offspring chromosomes by applying the genetic operators.
7. Repeat Step 5 until the size of the new population equals that of initial population, N.
8. Replace the initial (parent) chromosome population with the new (offspring) population.
9. Go to Step 4, and repeat the process until the termination criterion is satisfied.

2. **Advantages and disadvantages of Genetic Algorithm**

- **Advantages:**

- Ability to scan large search spaces
- Concept fairly easy to understand and there is no use of derivatives
- Good for "noisy" environments
- Allows Parallelism
- Flexible building blocks for hybrid applications
- Always produces a solution, solution gets better with time

- **Disadvantages:**

- High computational cost
- Difficulty in defining fitness function, population size, stopping criteria
- Does not assure the obtaining of the global optimum

II. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is one of the recent nature-inspired optimization methods. This technique was originally developed by Kennedy & Eberhart [46] in order to solve problems with continuous search space. PSO is based on the metaphor of social interaction and communication, such as bird flocking and fish schooling. This algorithm can be easily implemented and it is computationally inexpensive, since its memory and CPU speed requirements are low [47]. PSO shares many common points with GA. It conducts the search using a population of particles which correspond to individuals in GA. Both algorithms start with a randomly generated population.

PSO uses social rules to search in the design space by controlling the trajectories of a set of independent particles. The position of each particle x_i, representing a particular solution of the problem, is used to compute the value of the fitness function to be optimized.

Each particle may change its position, and consequently may explore the solution space, simply varying its associated velocity. In fact, the main PSO operator is the **velocity update v_i**, computed in *equ.3.11*, that takes into account the best position, in terms of fitness value reached by all the particles during their paths during its search g_{best}, and the best position that the agent itself has reached p_{best} , resulting in a migration of the entire swarm towards the global optimum.

The **updating of the position** of a particle is represented in *Fig.3.9.* The effect of the global best and personal best, hence the velocity of the particle, is very well demonstrated.

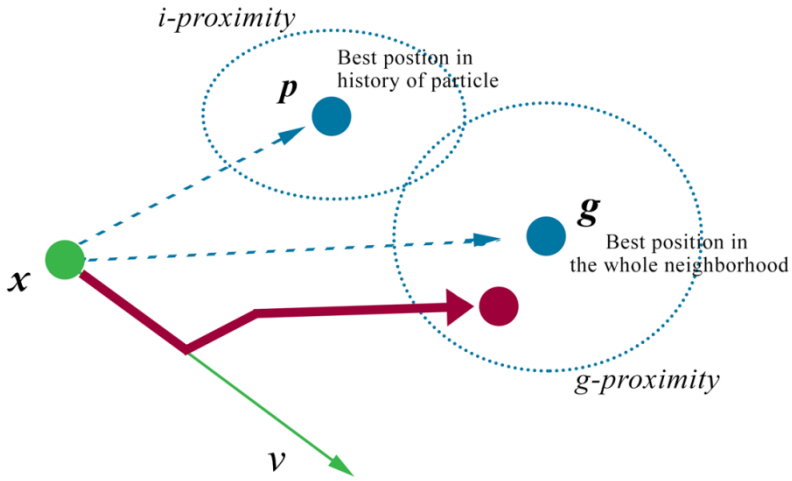


Fig.3.9. Illustration of velocity and position updates of a particle

At each iteration, the particle moves around according to its velocity and position; the cost function to be optimized is evaluated for each particle in order to rank the current location. The velocity of the particle is then stochastically updated according to

$$v_i^{t+1} = k\left(\omega v_i^t + C_1 r_1 (p_{best} - x_i^t) + C_2 r_2 (g_{best} - x_i^t)\right) \quad (3.11)$$

After, the particle position is updated according

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (3.12)$$

ω is a parameter controlling the flying dynamics.
k is called a constriction parameter
$r_1$and r_2 are random variables in the range $[0, 1]$.
C_1 and C_2 are factors controlling the related weighting of corresponding terms.

The inclusion of random variables endows the PSO with the ability of stochastic searching. The weighting factors, C_1 and C_2, compromise the inevitable tradeoff between exploration and exploitation.

Originally, the simple PSO [46] has no constriction coefficient. The algorithm suffered from bouts of divergence. In the works of Clerc and Kennedy [48], they suggested a more generalized PSO, where a constriction coefficient k is applied to both terms of the velocity formula. Clerc shows that the constriction can force PSO into convergence. To achieve such results, $C_1$and C_2 are often set so $C_1 + C_2 = 4.1$; and the constriction factor set $k = 0.729$.

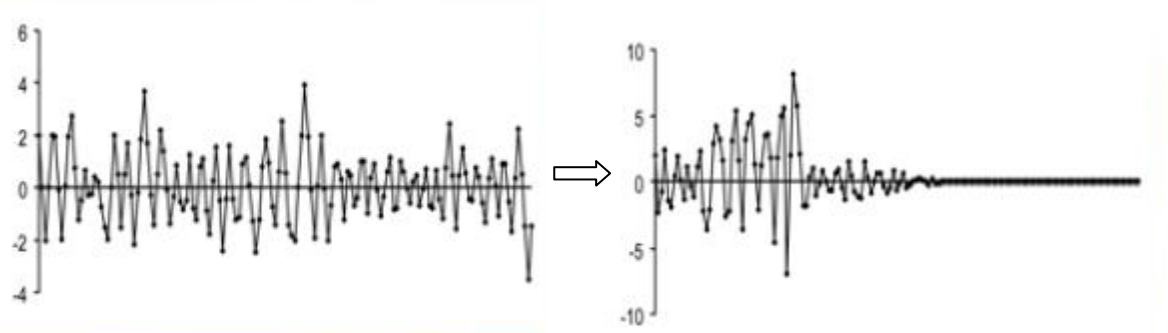


Fig.3.10 The effect of the constriction parameter *k* on assuring convergence

By using the constriction coefficient, the amplitude of the particle's oscillation decreases, resulting in its convergence over time (*Fig.3.10*).

The detailed operation of Particle Swarm Optimization is given below:

1. **Initialization.** Generate a random initial population of particles. The velocity and position of all particles are randomly set to within pre-defined ranges.

2. **Velocity Updating.** At each iteration, the velocities of all particles are updated according to:

$$v_i^{t+1} = k\left(\omega v_i^t + C_1 r_1 (p_{best} - x_i^t) + C_2 r_2 (g_{best} - x_i^t)\right)$$

After updating, v_i^{t+1} should be checked and secured within a pre-specified range to avoid violent random walking.

3. **Position Updating.** Assuming a unit time interval between successive iterations, the positions of all particles are updated according to:

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

After updating, x_i^{t+1} should be checked and limited to the allowed range.

4. **Memory updating.** Update p_{best} and g_{best} when condition is met.

$$p_{best} = x_i^{t+1} \qquad \text{if } f(x_i^{t+1}) < f(p_{best})$$

$$g_{best} = \{x_i^{t+1} \mid \min [\, f(x_i^{t+1}) \,]\}$$

Where $f(x)$ is the objective function subject to minimization.

5. **Termination Checking.** The algorithm repeats Steps 2 to 4 until certain termination conditions are met, such as a pre-defined number of iterations or a failure to make progress for a certain number of iterations.

i. Advantages and disadvantages of Particle Swarm Optimization

- **Advantages:**

- Fast and stable Convergence
- Easy to implement and few parameters to adjust
- Less sensitive to the nature of the objective function compared to other heuristic methods.
- As characterized by its fast convergence behavior, PSO has an in-built ability to adapt to a changing environment.
- PSO is effective for locating and tracking optima in both static and dynamic environments.

- **Disadvantages:**

- Dependency on initial points and parameters
- Weak Global search ability
- Lack of a solid mathematical foundation for analysis
- Slow compared to the other classical mathematical approaches
- PSO sometimes is easy to be trapped in local optima, and the convergence rate decreases considerably in the later period of evolution; when reaching a near optimal solution, the algorithm stops optimizing, and thus the accuracy the algorithm can achieve is limited [49].

III. Hybrid Genetic Algorithm/Particle Swarm Optimization Algorithm

Although GAs have been successfully applied to a wide spectrum of problems, using GAs for large-scale optimization could be very expensive due to its requirement of a large number of function evaluations for convergence. This would result in a prohibitive cost for computation of function evaluations even with the best computational facilities available today. Compared with GA, PSO has some attractive characteristics. It has memory, so knowledge of good solutions is retained by all the particles; whereas in GA, previous knowledge of the problem is discarded once the population changes. It has constructive

cooperation between particles; that is, particles in the swarm share information among themselves.

On the other hand, a drawback of PSO is that the swarm may prematurely converge. The underlying principle behind this problem is the fast rate of information flow between particles, resulting in the creation of similar particles with a loss in diversity that increases the possibility of being trapped in local optima.

The idea behind GA is due to its genetic operators; crossover and mutation. By applying crossover operation, information can be swapped between two particles to have the ability to fly to the new search area.

To deal with all these misgivings, and seeing as both GA and PSO work with an initial population of solutions and combining the searching abilities of both methods seems to be a reasonable approach, we propose a new algorithm (denoted as GA-PSO) that combines the evolutionary natures of both algorithms.

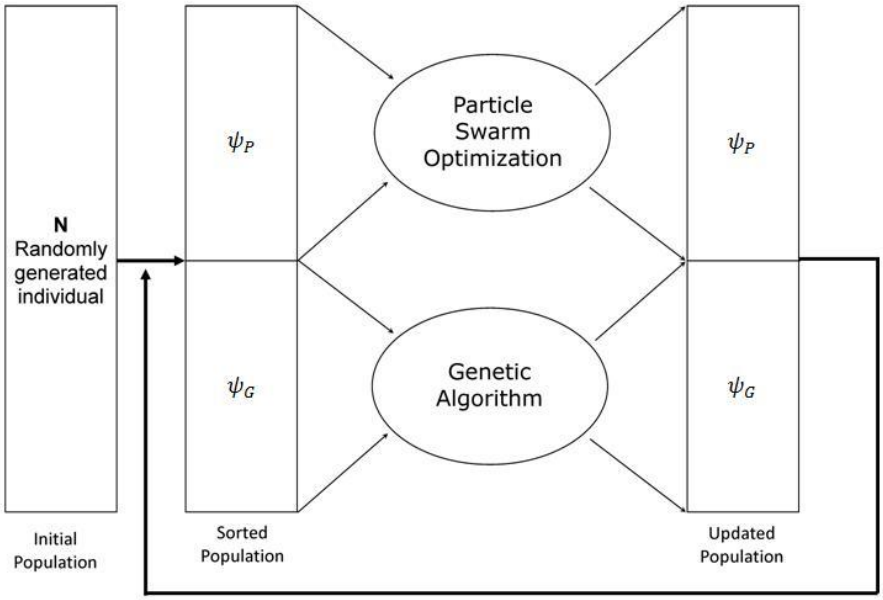


Fig.3.11. Schema Representation of the GA-PSO Hybrid Algorithm

To understand the workings of the algorithm, *Fig.3.11* depicts the schematic representation of the proposed hybrid GA-PSO. As can be seen, GA and PSO both work with the same initial population. The hybrid approach takes N individuals that are randomly generated. These individuals may be regarded as chromosomes in the case of GA, or as particles in the case of PSO.

The N individuals are sorted by fitness, and, according to a user defined probability P_k, the set is divided into two sets$\{\psi_G, \psi_P\}$.

- The top set ψ_P is used to adjust the particles using the PSO algorithm. The procedure of adjusting the particles in the PSO method involves selection of the global best particle, selection of the neighborhood best particle and then updating velocities and

positions. Knowing that PSO is an efficient algorithm in converging toward the best solutions, the resulting particles are expected to perform better.

- The other set ψ_G is fed into the real-coded GA to create new individuals by crossover and mutation operations, as shown in *Fig.3.12*. The crossover operator of the real-coded GA is implemented using the arithmetic rule, with a crossover probability P_C. The random mutation operator proposed for the real-coded GA is to modify an individual with a random number in the problem's domain with a probability P_m.

Both resulting individuals are combined into one single population of N individuals, which are then sorted in preparation for repeating the entire run. The hybrid algorithm terminates when it satisfies the stopping criterion.

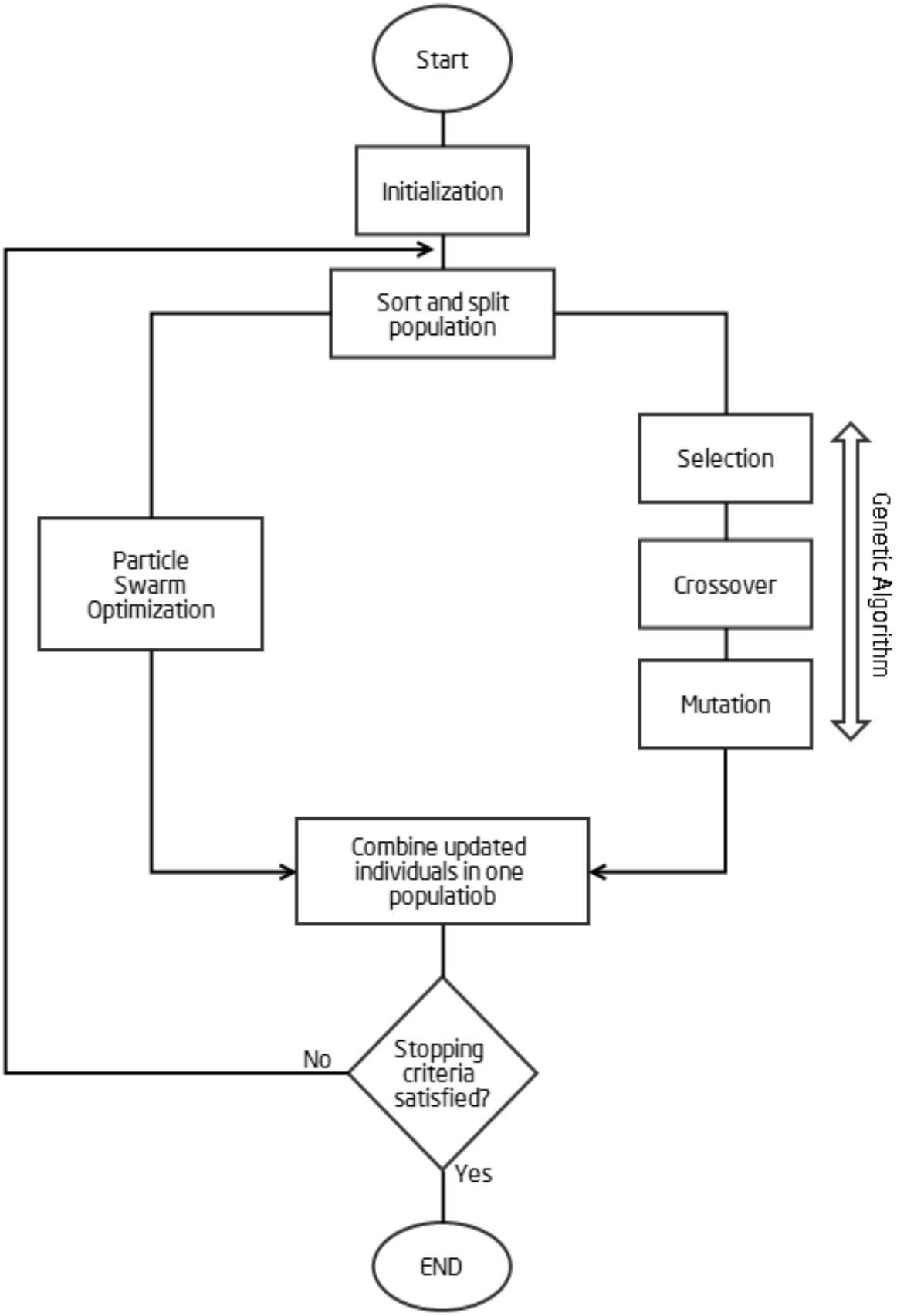


Fig.3.12. Hybrid GA-PSO Algorithm Flowchart

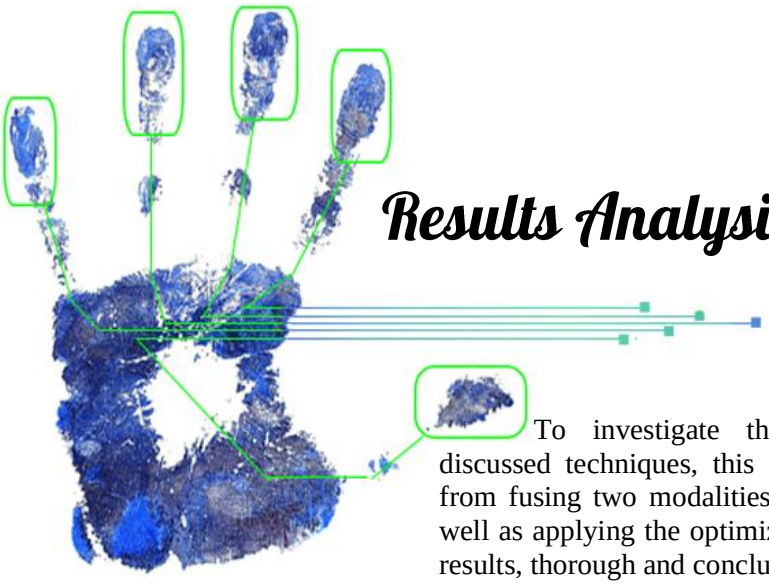
Conclusion

In this chapter, we discussed the fusion at score level. Before going further into the idea of fusion, we first explained what a score is, as it is at the base of this study. Furthermore, it has been shown that combining the scores as they are (raw scores) is not a wise idea. For the purpose of unifying all scores into one range, we presented the most used normalization schemes in literature. It was noted that although min-max and z-score normalizations perform well enough, tanh-normalization is a more efficient scheme. It was concluded that the best scheme to perform is the *tanh-normalization* rule.

More importantly, it was established that the fusion techniques come in two categories: classification-based and combination-based. The classification methods were briefly mentioned as well as a description of the various combination rules that can be used to perform the fusion.

Also, the techniques that will be used to optimize the fusion of multimodals were detailed. Last, our own approach for optimization was introduced. The hybrid GA-PSO as its name depicts, combines both GA and PSO, merging the benefits of both algorithms for the purpose of increasing its performance and robustness.

In the next chapter, we will investigate the performance of the combination rules and the optimization techniques. We will evaluate their competence at combining the different modalities from public databases at varied conditions efficiently.



Results Analysis and Discussions

To investigate the performance of the previously discussed techniques, this chapter presents the results obtained from fusing two modalities using classical combination rules as well as applying the optimization techniques. A discussion of the results, thorough and conclusive, is included at the end.

4.1. Introduction

To consider the deployment of multibiometric systems in real life, these systems need to be evaluated to estimate their performance in real use. According to the application specificity, three types of evaluation were differentiated in [50] to estimate the performance of a biometric system: technological evaluation, the evaluation of scenario and operational evaluation. The first one tests the performance of algorithmic parts of the system (features extraction, comparison and decision) using a publicly available database (benchmark). The evaluation of scenario tests a more complete system also including the sensors, the environment and the population specific to the application (scenario) tested. The operational evaluation tests the biometric system in the real conditions of use. In this work, we are concerned with the first type of evaluation.

In this chapter, experiments conducted on public databases to verify the performance of multimodal fusion techniques and their results are discussed. A brief description of the databases is included as well as the different evaluation metrics used to compare the combination rules and optimization techniques.

4.2. Databases

Databases allow researchers to test their algorithms and compare them with those from the state of the art. A very good analysis on multi-biometric approach, one that can compare them, and quantify their performance, can only be done if a database is possible to acquire or access to some quality public dataset is available. Public databases should be used for evaluating the performance of any biometric system. If the performance of biometric system is evaluated on local database, then it cannot be compared with other techniques that perform on standard database. Other techniques can be used on local database to compare, but it is an exhaustive process. It is better to use standard databases, as they contain a large number of users that captures a number of their biometric data and variations.

We give in this section an overview of the used datasets. *Table.4.1* presents a summary of these datasets.

4.2.1. NIST BSSR1

NIST Biometric Scores Set-Release 1 (BSSR1) [40] is a public database that consists of a set of raw output similarity scores.

The release includes 3 different partitions. Set 1 comprises similarity scores from comparisons of faces and fingerprints of the same individuals. It comprises of sets of raw output similarity scores from two face recognition systems (matchers C and G) operating on frontal faces, and one finger-print system, both left and right index live-scan fingerprints.

Set 2 is comprised of fingerprint scores from one system run on images of 6000 individuals. For each individual, the set contains one score from the comparison of two left index fingerprints, and another from two right index fingerprints.

Set 3 contains scores from two face systems run on images from 3000 individuals. For each individual, the set contains one score from the comparison of face A with a later face, B, and a score from face A and another later face, C.

The release is intended to facilitate interested parties to investigate problems in the fields of biometrics. The data is particularly suitable for study of score level fusion in different scenarios. In our case, since we are interested in multimodal fusion, we will be only using Set 1 of this release.

For the experiments, the scores are divided into a training set and a test set. This will be more explained in the next section.

4.2.2. XM2TVS

This database is built on the XM2VTS face and voice multimodal database [41], respecting the Lausanne Protocols I and II (LP1 and LP2). LP1 has 8 baseline systems and LP2 has 5 baseline systems.

In our experiments, we work with the LP1 set of scores. This dataset includes eight biometric scores per claim, five for face images and the remaining three for speech. The benchmark of LP1 includes two files, one is *dev.label*, the training set, which is used to optimize the fusion parameters, and then there is *eva.label*, the test set, which is used to assess the performance. In total the training set has 600 client and 40k impostor claims, and the test set has 400 client and around 112k impostor claims.

Face 1: (FH, MLP)	Speech 1: (LFCC, GMM)
Face 2: (DCTs, GMM)	Speech 2: (PAC, GMM)
Face 3: (DCTb, GMM)	Speech 3: (SSC, GMM)
Face 4: (DCTs, MLP)	
Face 5: (DCTb, MLP)	

4.2.3. BANCA

The BANCA database contains matcher scores of face and speech [51]. There are 7 different protocols: Mc, Md, Ma, Ua, Ud, G and P. In these protocols, two groups of users are distinguished and are labeled by g1 and g2. We use g1 as a development (training) set and g2 as an evaluation (test) set.

The selected scores to be fused are the following labeled as:

S_{b1}: IDIAP_voice_gmm_auto_scale_33_25.scores
S_{b2}: UC3M_voice_gmm_auto_scale_18_300.scores
S_{b3}: UC3M_voice_gmm_auto_scale_18_32.scores
S_{b4}: SURREY_face_svm_auto.scores
S_{b5}: SURREY_face_svm_man_scale_0.18.scores
S_{b6}: UCL_face_lda_man.scores

Databases	Training		Test	
	Clients	Imposters	Clients	Imposters
NIST BSSR1	259	132869	257	132871
XM2VTS	600	40000	400	111800
BANCA (P – G)	234	312	234	312
BANCA (M – U)	78	104	78	104

Table.4.1. Summary of Databases

4.3. Evaluation metrics

Performance of a biometric system can be degraded due to many reasons, ones of which are the condition of sensors, alterations in user's biometric characteristics, changes in ambient conditions and user's interaction with the sensor. Consequently, a biometric matching system outputs a similarity score that quantifies similarities between the enrolled and input templates. At this stage, two types of similarity scores are identified (*Fig.4.1*).

i. Genuine Scores:

Pairs of biometric samples generating a similarity score and belonging to the same person are genuine pairs. The distribution of these pairs is called a genuine distribution.

ii. Imposter Scores

Pairs of samples generating a similarity score and belonging to different persons are referred to as imposter pairs, generating a distribution called an imposter distribution.

In the decision module, the system decides then whether to accept or reject the identification depending on a threshold t_s.

In order to compare the performance of biometric systems, many metrics are used in the field. In here, we are concerned with these main ones:

4.3.1. False Acceptance Rate (FAR)

It represents the ratio of impostors accepted by the authentication system. In other words, it is the rate at which the non-authorized persons are falsely recognized during matching process as genuines.

4.3.2. False Rejection Rate (FRR)

It represents the ratio of genuine users rejected by the system. Meaning, it is the rate at which authorized people are falsely rejected and labeled as imposters during matching process.

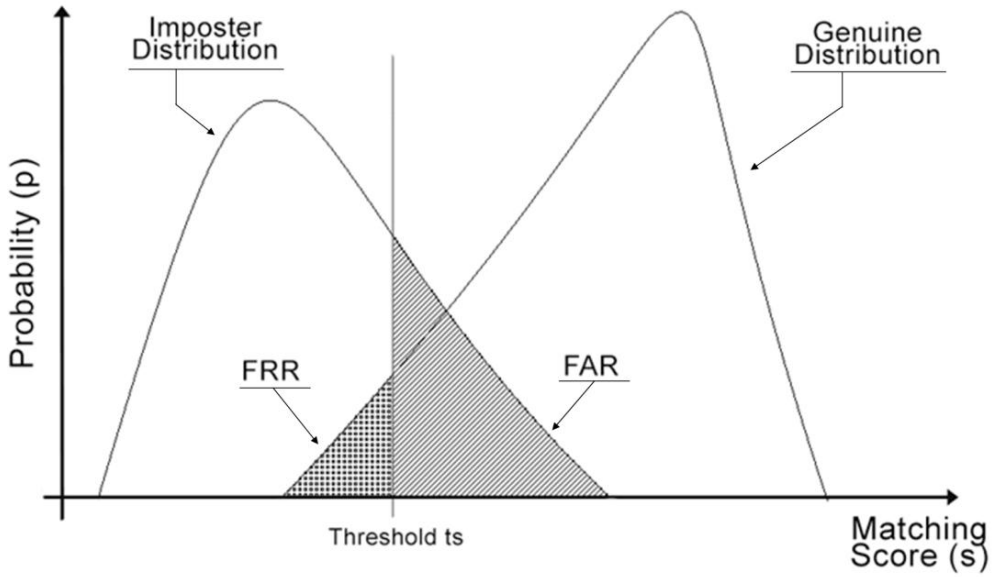


Fig.4.1. FAR and FRR Representations on Score Distribution

To quantify these rates, it is important to define two properties that are known in the biometric authentication, which in itself is a case of detection process. Computing these error rates is based on the comparison of the scores against the threshold t_s.

- **False Accept (FA)**

It is the case where the biometric system identifies an imposter as a genuine and hence accepts it as a true user. *Equ.4.1* represents this, with S_{imp}being a score in the imposter distribution.

$$\mathrm{FA} = \sum \#\left(S_{imp} > t_s\right) \tag{4.1}$$

- **False Reject (FR)**

It is the case where the biometric system identifies a genuine client or true user as an imposter and is consequently rejected. *Equ.4.2* represents this, with S_{gen}being a score in the genuine distribution.

$$\mathrm{FR} = \sum \#\left(S_{gen} \leq t_s\right) \tag{4.2}$$

The FAR and FRR are directly related to the FA and FR respectively. As a result, they are both affected by the change in the threshold t_s, as is depicted in *Equ.4.3* and *Equ.4.4* respectively.

$$FAR = \frac{\mathrm{FA}}{Number\ of\ imposters} \tag{4.3}$$

$$FRR = \frac{\text{FR}}{Number\ of\ genuines} \tag{4.4}$$

4.3.3. Effect of threshold

It has been established that the choices of the threshold directly impacts the decision process, and the error rates FAR and FRR.

In case a higher security is required of a system, t_s is set to be large which in turn decreases the FAR. On the other hand, if the threshold is decreased, it is the FRR that decreases. This kind of setting is used when the system is more tolerant of input variations and noise.

As a consequence of this, a biometric system needs to make a tradeoff between FAR and FRR. The Equal Error Rate (EER) is often used as a performance measure.

4.3.4. Error Equal Rate (EER)

It is the error rate when the system is configured in order to obtain a FAR equal to the FRR which offers a compromise when choosing a threshold.

$$EER = \frac{FAR + FRR}{2} \tag{4.5}$$

For performance analysis, it is good to know that the smaller the EER, the better is the system performance. The focus of our experiments is to obtain the smallest EER using optimization algorithms.

```
ROC ←[]
EER←1.0
DIFF ←1.0
START ←min(scores)
END ←max(scores)
for t_s START to END in N steps do
        FAR← compute FAR for t_s
        FRR← compute FRR for t_s
        append (FAR,FRR) to ROC
        if abs(FAR−FRR) < DIFF then
        DIFF ←abs(FAR−FRR)
        EER←(FAR+FRR)/2
        end if
end for
return EER, ROC
```

Fig.4.2. Classical EER Computing Algorithm

4.3.5. Equal Error Rate Graph

This curve, sometimes called the **FAR vs. FRR curve**, is the most often used by researchers trying to understand the performance of their recognition system. It shows the evolution of both error rates (FAR and FRR) at all thresholds (*Fig.4.3*).

The EER value will pinpoint the score at which the threshold is optimal, in the sense of the best trade-off between the FAR and FRR.

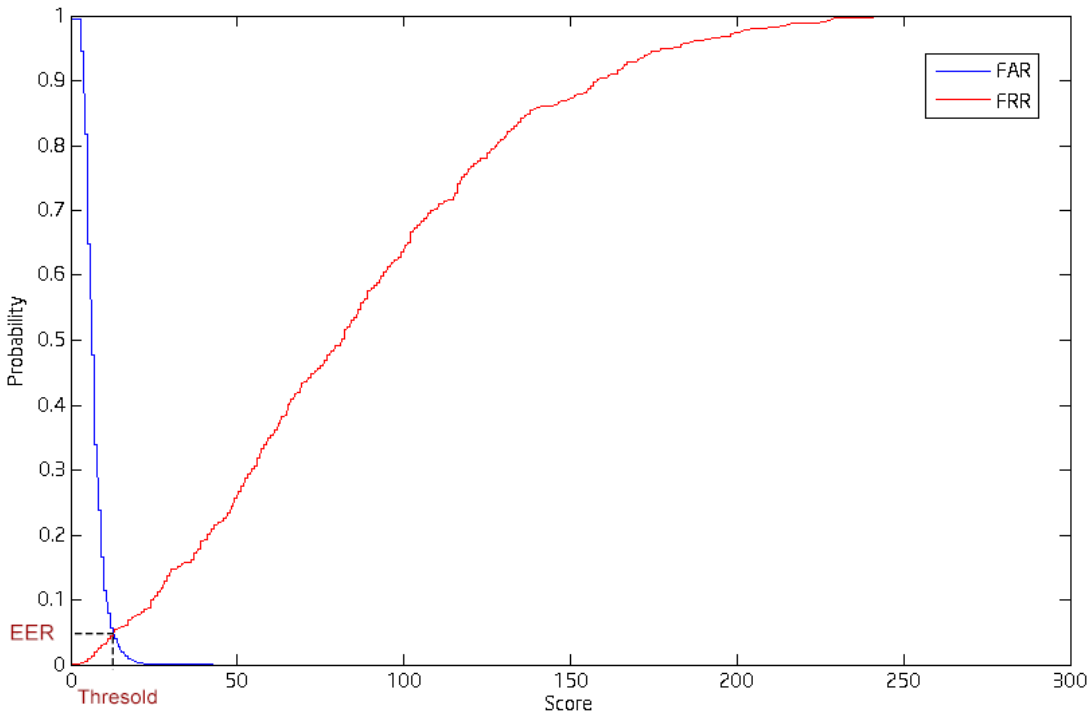


Fig.4.3. Equal Error Rate Graph

Minimizing the area under the Crossover of the two plots is generally the goal of the researcher. The user of an authentication system uses this curve to calculate where to set their operating threshold. The graph will show the expected FAR and FRR at any chosen threshold.

4.3.6. The ROC curve

The system performance at all operating points (thresholds) can be depicted by Receiver Operating Characteristics (ROC) Curve.

It plots the FRR against the FAR. The aim of this curve is to present the tradeoff between FAR and FRR and to have a quick overview of the system performance and security.

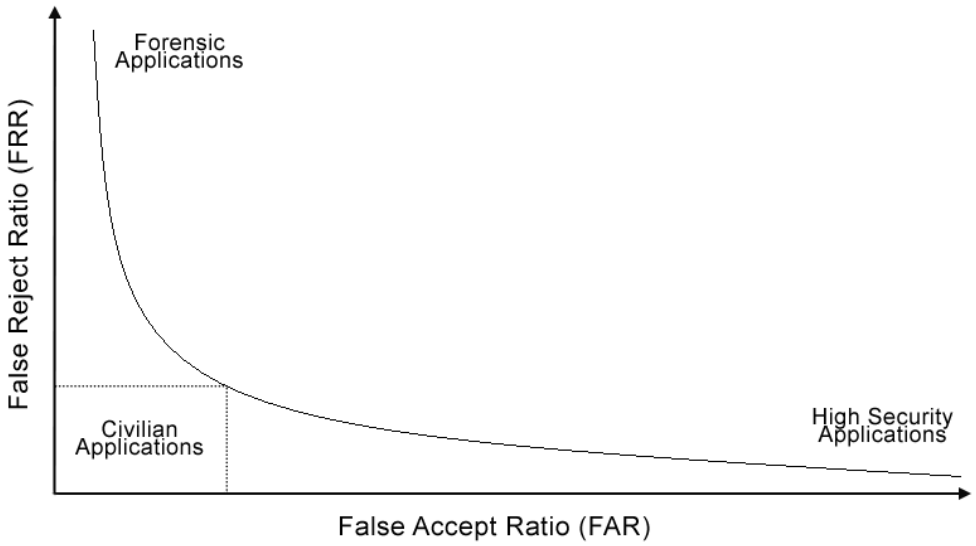


Fig.4.4. Receiver operating characteristic (ROC) curve

4.4. Experiments setup

At this stage, and after getting a grasp of the theoretical concepts, we look into the specifics of the experiment. All the tests have been done on a Pentium Dual machine with 3 GB of RAM with the Matlab programming language.

In order to validate the aforementioned algorithms and their effectiveness when dealing with multibiometrics, we split the databases into two separate sets:

- **The training set** which serves to compute the biometric reference of each matcher. In other words, we train the algorithms to attain the optimal weights for each matcher.
- **The testing set** which serves to validate the results of the training by computing the performance of the fusion with the obtained weights.

When it comes to the evolutionary techniques, there exists a pre-processing step where the parameters for each algorithm are set. The following offers a summary of this setup.

i. Genetic Algorithm

To run the genetic algorithm, there are six parameters to be set; the population size, maximum iteration, fitness function, selection method, crossover and mutation probabilities. The table sums up the genetic algorithm configuration.

Parameter	Value
Population size	50
Maximum iterations	50
Fitness function	EER of fused scores
Selection	Tournament with $k = 4$
Crossover function	Arithmetic
Crossover probability	$P_C = 0.7$
Mutation probability	$P_m = 0.05$

In our experiments, and in terms of multimodal fusion, the algorithm generates an initial population of size *N* which consist of the weights w_i in *Equ.4.6.* In this work, we will fuse two modalities at a time to create fusion scores.

$$S_{f_i} = w_1 * S_i^{modality1} + w_2 * S_i^{modality2} \quad (4.6)$$

Where w_j is the weight, chosen in the interval {0, 1}, and associated to normalized score $S_i^{modality\ j}$.

The fitness function is defined as the EER error rate. For every set of chromosomes (w_1, w_2), the EER of the fused scores S_f is computed.

As a result, the chromosomes are arranged and sorted to go through the selection, crossover and mutation processes. Evaluation of the fitness costs of the "offspring" is once again run and the weights to produce the minimum EER value is picked as optima.

If the stopping criteria are not satisfied yet, this procedure is repeated until one of the conditions is met.

What follows is a pseudo-code for a classical Genetic Algorithm.

```
{ Initialize population of weight pairs (w1, w2) ;
Evaluate population by computing EER;

While Termination Criteria Not Satisfied

        { Select parents for reproduction;
          Perform recombination and mutation;
          Evaluate offspring by computing EER;
        }
}
```

Fig.4.5. Pseudo-code for the Genetic Algorithm

ii. Particle Swarm Optimization

Similarly to GA, Particle swarm optimization requires that some parameters be preset to be able to run properly. Among the swarm (population) size and maximum iterations, the other parameters are:

Parameter	Value
Swarm size	50
Maximum iteration	50
Fitness function	EER of fused scores
Inertia factor	$w = 1$
$C_1$and C_2	$C_i = 2.05$
Random numbers	$r_i = uniform\ distribution\ (0,1)$
Constriction factor	$k = 0.729$

w, the inertia factor, which controls the contribution of the previous velocity on the next velocity of the particle.

C_1 and C_2 determine the relative influence of the cognitive component and social component, respectively.

$r_1$and r_2 are random numbers, which are used to maintain the diversity of the population, and are uniformly distributed in the interval $\{0, 1\}$.

k is the constriction factor that constrains and controls the velocity. When chosen properly, with accordance with $C_1$and C_2, it prevents the algorithm from diverging [48].

The process in Particle Swarm Optimization is quite similar to that of Genetic Algorithm. But instead of dealing with chromosomes and genes, the weights are considered to be particles. After generating an initial population of particles, the algorithm will evaluate the fitness functions, EERs, to detect the particle's best p_{best} and general best g_{best}. These values are used to update the velocities of each particle using *Equ.3.11* and then their positions according to *Equ.3.12*.

The updated population will once again be evaluated to determine the best pair of weights (w_1, w_2) that produces the minimum EER.

The whole process is repeated if the stopping criteria are not satisfied.

What follows is a pseudo-code for Particle Swarm Optimization.

```
{ Initialize population of weight pairs (w1, w2) ;
Evaluate population by computing EER;

While Termination Criteria Not Satisfied
        {Determine personal and general best (p_best, g_best)
        Update particle's velocity (equ.3.11.)
        Update particle's position (equ.3.12.)
        Evaluate offspring by computing EER;
      }
 }
```

Fig.4.6. Pseudo-code for the Particle swarm Optimization

iii. Hybrid Genetic Algorithm/Particle Swarm Optimization

As the name denotes, the hybrid algorithm runs with all the previously mentioned parameters with few additions and alterations as shown in the table below. We introduce a splitting probability P_k that will determine the size of populations that will be subjected to PSO and GA.

Parameter	Value
Initial Population size	50
Maximum iterations	50
Fitness function	EER of fused scores
Selection	Tournament with $k = 4$
Crossover function	Arithmetic
Splitting probability	$P_k = 0.6$
Crossover probability	$P_C = 1$
Mutation probability	$P_m = 0.05$
Inertia factor	$w = 1$
C_1 and C_2	$C_i = 2.05$
Random numbers	$r_i = uniform\ distribution\ (0,1)$
Constriction factor	$k = 0.729$

The only difference is that $P_C = 1$ which, as explained in earlier section means that the whole subset population dedicated for the Genetic Algorithm will go through the crossover operator.

Now that the algorithms are tuned and parameters set, and in order to evaluate the performances, the experiment will go through different stages. These stages are summarized in the flowchart of *Fig.4.7.*

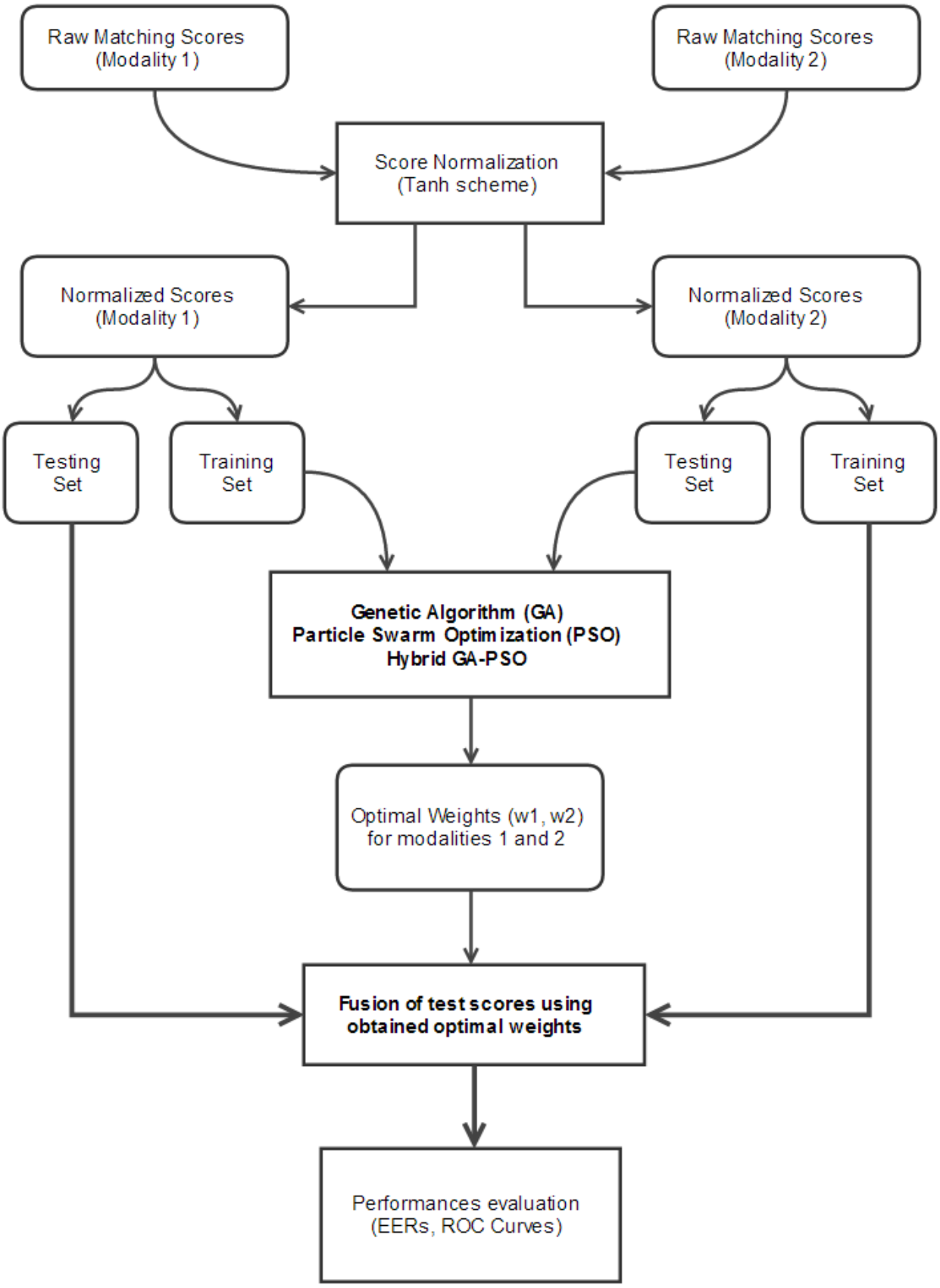


Fig.4.7. Multimodal Biometric Score Fusion Flowchart

4.5. Results and discussion

In this section, the combination rules presented previously, to fuse biometric scores at the matching level are tested. To compare them, the performance metrics mentioned at the beginning of this chapter are used. The equal error rate (EER) is used to tune and test, when applicable, and compare the different techniques.

As a reminder, the lower the EER, the better the system performs. Also, to have a better visual of the performance, the **receiver operating characteristic** (ROC) curves of fused scores are displayed. When it comes to understanding how to read the ROC curve for performance evaluation, the closer the curve is to the edges (axes), the better the performance.

It is important to note that, although the use of chimeric databases (scores from different users) to evaluate the performance of the system has been proven to have no effect on the evaluation as compared to real users databases [52], in this experiment, and for the purpose of keeping the experiments as close to real life applications as possible, the databases used in this work are all non-chimeric. Meaning, the scores of different modalities come from the same persons.

Before applying the rules on these scores, they have all been put under the same range {0, 1} using the *tanh normalization* scheme.

Tables 4.2 summarize the performance (EER) of the single modalities (unimodal) of every database. The best performance in each of the modalities is shown in **bold**.

Modality		**EER (%)**
Face	G	5.69
	C	**4.39**
Fingerprint	Right	5.52
	Left	7.91

(a) NIST (BSSR1) Face and Finger Modalities

	Modality		**EER (%)**
	Feature	**Classifier**	
Face	FH	MLP	1.81
	DCTs	GMM	4.23
	DCTb	GMM	1.67
	DCTs	MLP	3.89
	DCTb	MLP	6.57
Voice	LFCC	GMM	**1.23**
	PAC	GMM	6.61
	SSC	GMM	4.51

(b) XM2VTS Face and Voice modalities

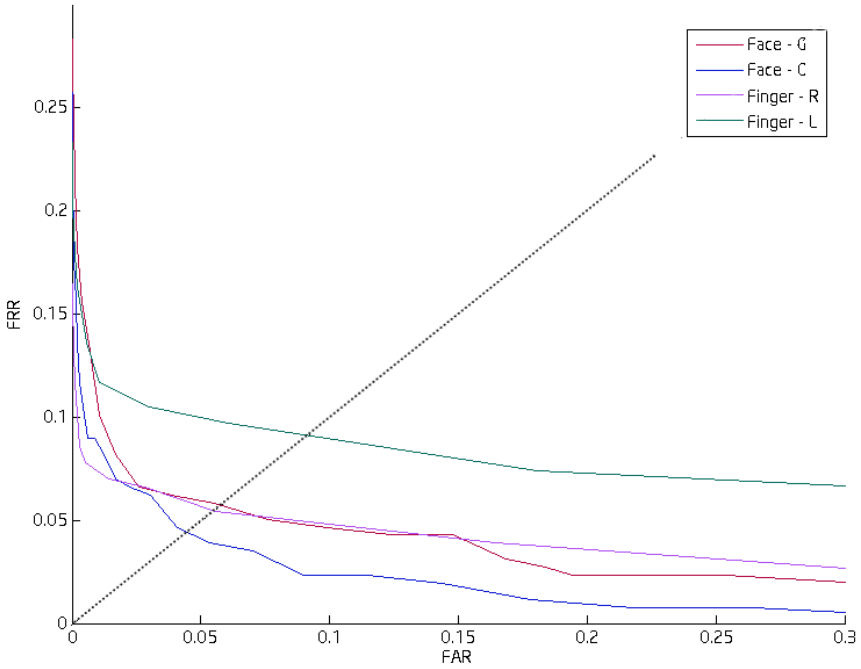
Subset	Modality		EER (%)
G	Face	S_{b1}	11.32
	Voice	S_{b4}	**1.98**
Md	Face	S_{b2}	10.58
	Voice	S_{b5}	4.33
Ua	Face	S_{b3}	28.5
	Voice	S_{b6}	15.1

(c) BANCA Face and Voice Modalities

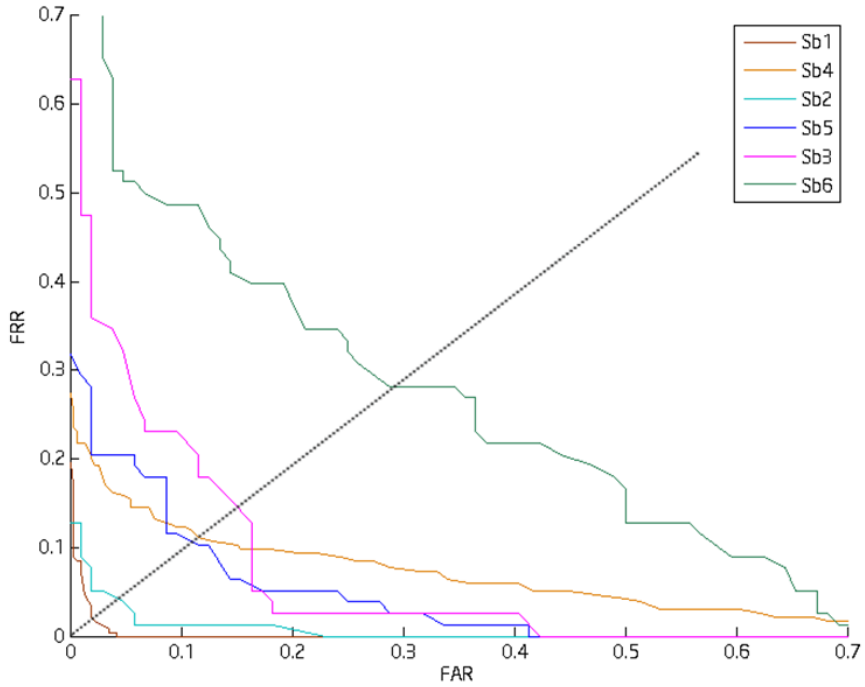
Table.4.2. Performance (EER %) of Single Modalities from (a) NIST (b) XM2VTS (c) BANCA

In every database, there is one modality that performs better compared to the other modalities. Such is the case in the NIST database where matcher C for face gives an EER = 4.39%. Also, in the BANCA dataset, the subset G voice modality (S_{b4}) gives a better EER (=1.98%) which is expected, seeing as the G subset is comprised of clean data scores.

To have a better visualization of the performance of these modalities, the ROC curves are plotted. The clean and degraded data of the BANCA are clearly seen and distinguished (See *Fig.4.8.b*) with the worst EER (28.5%) coming from the Ua Face modality (S_{b3}). It is important at this stage to specify that in this database, and specifically the Ma and Ud subsets, we deal with scores obtained in not so optimal conditions. With Ua, Unmatched Adverse, the scores were obtained under challenging conditions whereas Md, Matched degraded, are scores from different devices. As for the NIST (*Fig.4.8.a*), it can be noticed that, although Matcher C gives better EER values, for a lower FAR, the **Right** finger performs better that the C matcher.



(a)



(b)

Fig.4.8 ROC Curves of Single modalities in the (a) NIST database and (b) BANCA database

4.5.1. Score fusion using classical rules

The first set of experiments with score level fusion is to apply the classical combination rules on the public databases. *Tables.4.3* show the EERs in % of the scores in the NIST (BSSR1), XM2VTS, and BANCA testing sets respectively, and those of the fused scores from the same set after applying the different classical combination rules. To avoid congestion of information when dealing with both XM2VTS and BANCA, and for the sake of clarity, only few of the possible combinations to fuse different modalities are included.

Before applying the rules on these scores, they have all been put under the same range {0, 1} using the *tanh normalization* scheme. The best performance in each of the fused modalities is shown in **bold**.

Combination Rule	Single Modality	Max	Min	Product	Simple Sum
Face – G	5.69	5.49	5.52	2.70	**1.21**
Finger – R	5.52				
Face – C	4.39	3.66	7.91	4.77	**1.00**
Finger – L	7.91				

(a) NIST (BSSR1)

Combination Rule	Single Modality	Max	Min	Product	Simple Sum
Face 1	1.81	1.45	1.81	1.64	**1.24**
Voice 2	6.61				
Face 2	4.30	1.98	4.40	1.65	**1.57**
Voice 2	6.61				
Face 5	6.57	**3.06**	6.46	5.66	3.67
Voice 3	4.51				

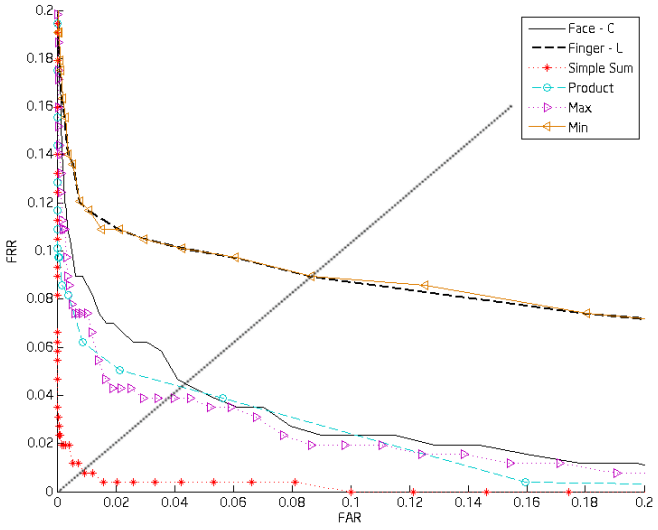
(b) XM2VTS

Combination Rule		Single Modality	Max	Min	Product	Simple Sum
G	Face S_{b1}	11.32	2.19	7.32	2.35	**1.82**
	Voice S_{b4}	1.98				
Md	Face S_{b2}	10.58	5.45	5.61	**3.37**	**3.37**
	Voice S_{b5}	4.33				
Ua	Face S_{b3}	28.5	15.4	28.5	16.9	**10.4**
	Voice S_{b6}	15.1				

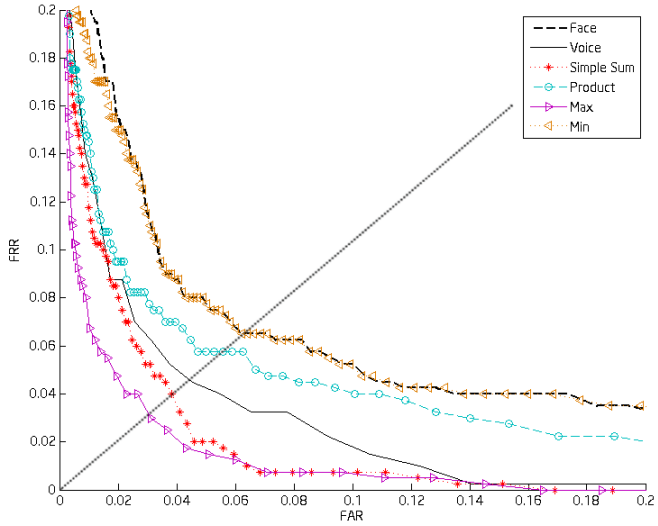
(c) BANCA

Table.4.3. Resulted EERs (%) of fused scores from (a) NIST (b) XM2VTS (c) BANCA Databases using Classical Combination Rules

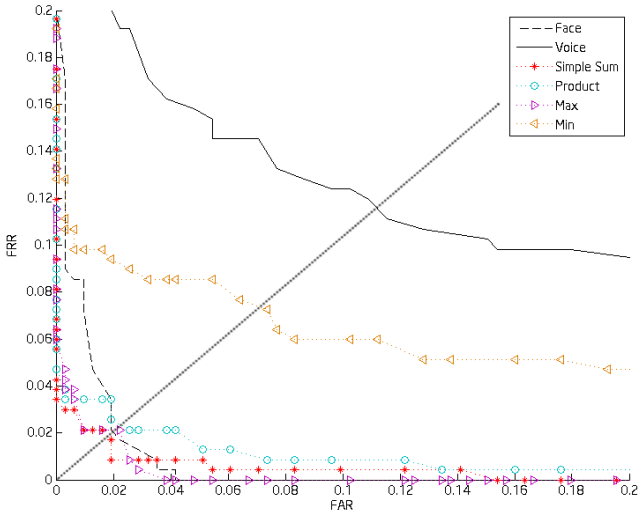
From the first look, an improvement in accuracy is clearly observed between unimodal and multimodal systems, regardless of the fusion rule applied. The ROC curves depicted in *Fig.4.9* demonstrate this, where the single biometrics are outperformed by their multimodal counterparts. In *Table.4.2.a*, even the best matcher, Face-C with EER = 4.39%, is outperformed by a simple max-rule, with an EER = 3.66%.



(a)



(b)



(c)

Fig.4.9 ROC Curves of fused scores from (a) NIST (b) XM2VTS (c) BANCA Databases using Classical Combination Rules

In all of the experiments done on the three databases, the **max-rule** performance is much better than that of the min-rule. In fact, from their ROC representations, the min-rule score distributions follow that of the worst modality that is fused. The fact that the scores in all the databases are *similarity* scores is the reason.

When dealing with this type of scores, the higher the value of the score, the more similar are the data acquired to the template, which explains the results obtained. Sometimes, instead of similarity score, *distance scores* are used. These scores indicate how much the data acquired is closer to the template. For a distance score, the lower the value, the more similar are the data acquired to the template. For such reasons, it is expected that the min-rule will perform better if the scores are distance scores.

Among all the classical combination rules, **simple summation** has the better performance even when dealing with degraded data such is the case in the BANCA Ua subset with (EER = 10.4%). This is due to the fact that Simple Sum handles situations, where a high level of noise is present, better, leading to ambiguity in the classification problem.

Product rule should have a good performance if the different modalities had statistically independent representations, which is *usually* not the case [53].

4.5.2. Score fusion using Optimization Techniques

The last of the combination rules is the weighted sum. In order to obtain the maximum performance out of the system, the goal of the experiments in this section is to optimize the system by locating the optimum weights of the weighted sum rule. GA, PSO and the hybrid GA-PSO are employed for such purpose. *Table 4.4* summarizes our findings.

In here, we notice that the weighted sum, with weights optimized through optimization techniques, performs best among the other rules. Compared to the best EER obtained from S.S. in the previous section for XM2VTS (F2-S2) pair (EER = 1.25 %), the ET give a better optimized EER (= 0.87%).

Combination Rule	Single Modality	Genetic Algorithm	Particle Swarm O.	Hybrid GA-PSO
Face – G	5.69	0.44	0.62	**0.43**
Finger – R	5.52			
Face – C	4.39	**0.75**	**0.75**	**0.75**
Finger – L	7.91			

(a)

Combination Rule	Single Modality	Genetic Algorithm	Particle Swarm O.	Hybrid GA-PSO
Face 1	1.81	**0.87**	**0.87**	**0.87**
Voice 2	6.61			
Face 2	4.30	1.49	**1.25**	1.32
Voice 2	6.61			
Face 5	6.57	1.88	**1.85**	**1.85**
Voice 3	4.51			

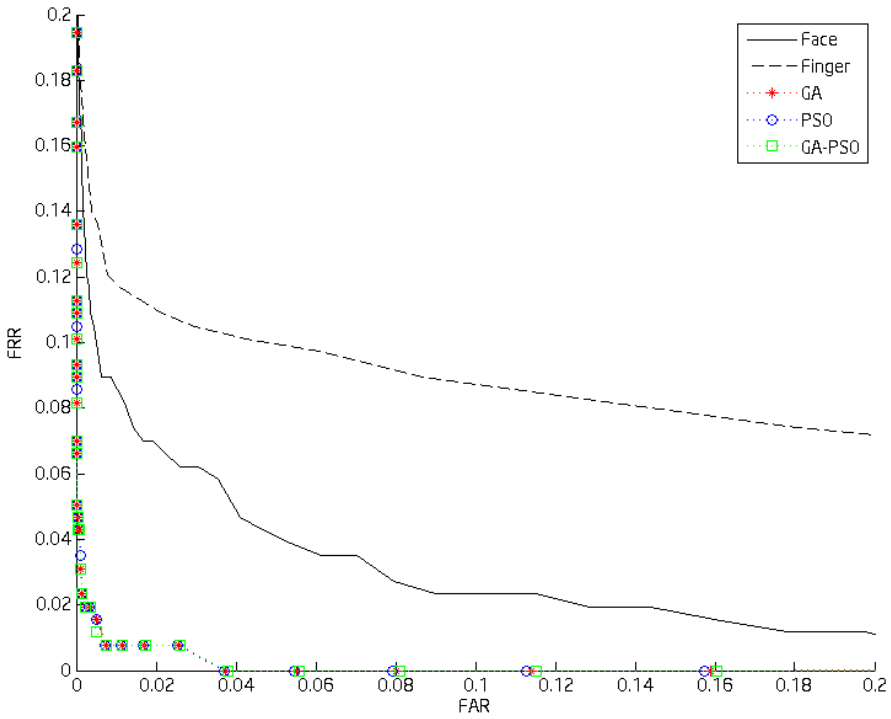
(b)

Combination Rule		Single Modality	Genetic Algorithm	Particle Swarm O.	Hybrid GA-PSO
G	Face S_{b1}	11.32	1.07	1.07	**0.91**
	Voice S_{b4}	1.98			
Md	Face S_{b2}	10.58	**2.24**	**2.24**	**2.24**
	Voice S_{b5}	4.33			
Ua	Face S_{b3}	28.5	11.1	**10.4**	**10.4**
	Voice S_{b6}	15.1			

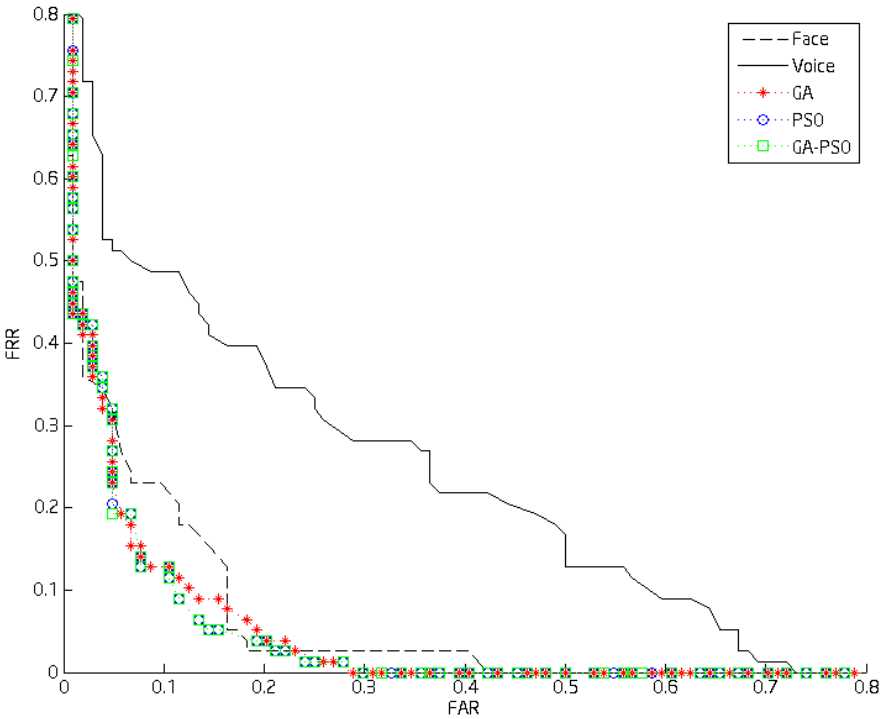
(c)

Table.4.4. Resulted EERs (%) of fused scores from (a) NIST (b) XM2VTS (c) BANCA Databases using Optimization Techniques

On a different note, the performances of GA, PSO and the hybrid GA-PSO are closely similar in most datasets as is observed in NIST (C-L) pair, XM2VTS (F1-S2) pair and BANCA Md subset. The ROC curves representations in *Fig.4.10* give a better general observation. We notice that even with the degraded data (*Fig.4.10.b*), the execution of all three optimization techniques give a good and similar performance rates.



(a)



(b)

Fig.4.10. ROC Curves of fused scores from (a) NIST (b) BANCA Databases using Optimization Techniques

Sometimes, we notice that one of the optimization algorithms gives a better accuracy than the other as is the case in *Table.4.4.a* with the G-R pair ($EER_{GA} = 0.44\%, EER_{PSO} = 0.60\%$). These slight differences might be explained by the stagnation phenomena that PSO suffers from due to its fast convergence property (Converges to local minima).

On the other hand, the results with the Ua pair ($EER_{GA} = 11.1\%, EER_{PSO} = 10.4\%$) in *Table.4.4.c* and XM2VTS (F2-S2) pair ($EER_{GA} = 1.45\%$, $EER_{PSO} = 1.25\%$) in *Table.4.4.b* show that PSO performs better than GA. Taking into consideration the fact that the data is degraded in both cases, this might have had a role in getting these results. We can hypothesize that the reason for this is the slow search of a large space conducted by GA compared to the fast pace of PSO. These techniques having a stochastic factor to their searches (random initial populations) might have helped too. We expect GA to reach the same optimum points as PSO if it was left to run longer (increase maximum iterations) which was the case.

4.5.3. Evaluation of Optimization techniques

Since these are stochastic search techniques, it is difficult to say that one algorithm performs better than the other in a definite and conclusive way. Regardless, an approximate study can be done. When it comes to comparing the optimization techniques to each other in terms of biometric performance, it is clear from the results discussed in the previous section that they mostly result in the same best accuracy (smallest EER). But they differ in other aspects such as the computational cost. Genetic algorithm, due to the fact that it covers large search spaces, has a large computational time. On the other hand, we have PSO that, as a consequence of its fast operations, consumes less computational time but converges quickly to local minima.

The hybrid GA-PSO takes advantage of both algorithms where it gains in computational time, by adding the benefit of fast search property of PSO, and still covering the large search space efficiently. This is demonstrated in *Fig.4.11*, where the cost function is plotted against the number of iterations run by all three algorithms. It can clearly be noted, with the XM2VTS dataset, that GA-PSO takes much few iterations (iter = 2) to reach the global optimum than either PSO (iter = 9) or GA (iter = 38).

	Genetic algorithm	**Particle swarm optimization**	**Hybrid GA-PSO**
Time to run 50 iterations	105	220	315
Time to reach a global minimum	76	38	12

Table.4.5. Running time of Genetic Algorithm, Particle Swarm Optimization and Hybrid GA-PSO

Table.4.5 puts in value the amount of time in CPU-time that each algorithm takes to be executed for 50 iterations and to reach a global minimum. It seems, from a first look, that the

hybrid algorithm gives the least favorable running time. That is quite logical when we know that GA-PSO computes the cost function three times in one iteration while PSO computes it twice, and GA, once. But when taking into consideration that it takes much fewer iterations to reach a global, it is actually faster than the two other algorithms. Same thing can be said when comparing GA and PSO.

After many runs of these programs, it has been noticed that, although GA and PSO mostly give good results, they would occasionally get stuck in local minima, as is the case in *Table.4.4.c* and *Fig.4.11.a* with the NIST dataset. On the other hand, the hybrid GA-PSO is observed to always converge to a global point in the shortest time.

We concluded that GA-PSO is more robust and efficient. It can be said that the reason behind this performance is that PSO converges towards better solutions whereas GA guides the search away from bad ones.

- NIST - - XM2VTS -

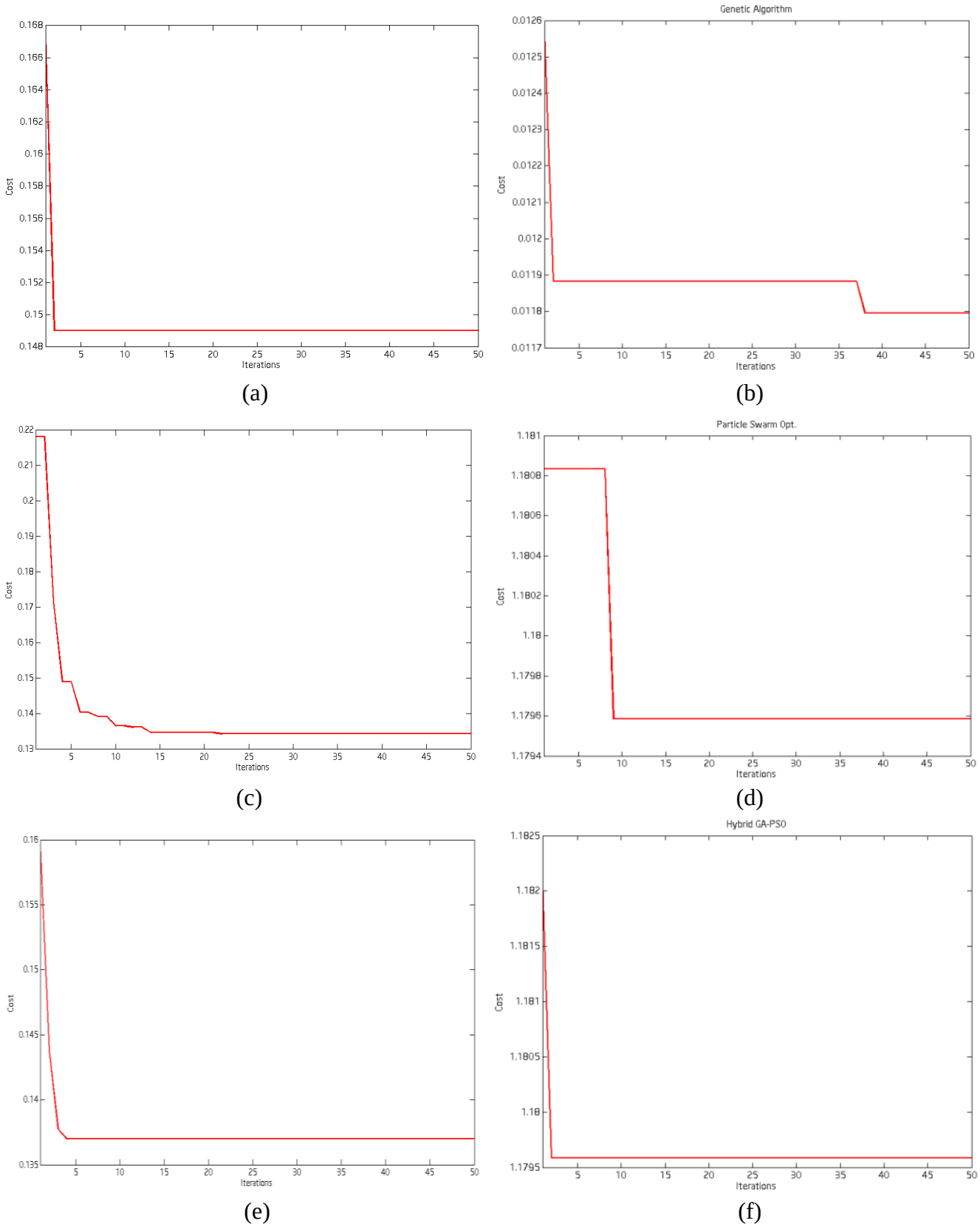


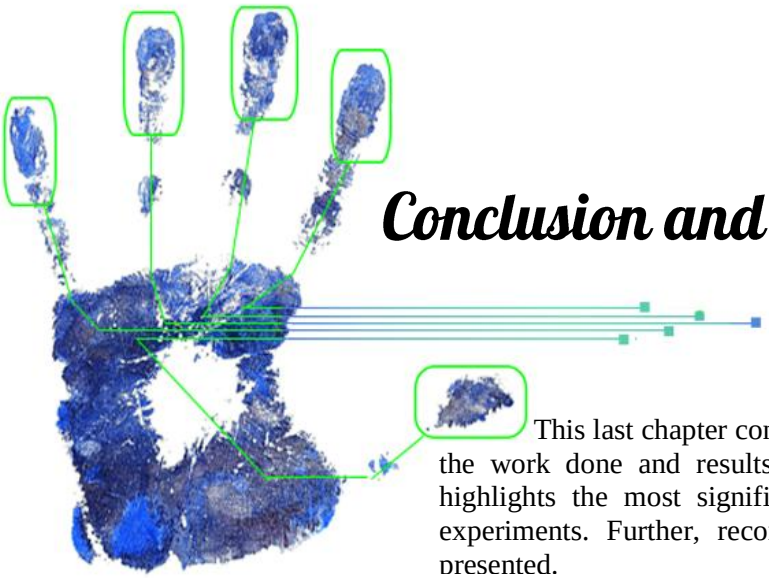
Fig.4.11. Cost Function vs. Number of Iterations for (a)(b) Genetic Algorithm (c)(d) Particle Swarm Optimization (e)(f) Hybrid GA-PSO

Conclusion

The goal in this chapter was to investigate the integration of two biometric modalities for the purpose of achieving a better performance than that of their unimodal counterparts. The experiments have been concerned with the fusion of face and fingerprint from the NIST BSSR1 database, face and voice from the XM2VTS and BANCA databases at the score level. The basic idea was to fuse and evaluate the different modalities using the combination rules; simple sum, product, max and min rules. The results showed that among these rules, simple sum performs better in most cases whereas min-rule gives the worst performance.

In order to improve the performance of the fusion, the optimization algorithms Genetic Algorithm and Particle Swarm Optimization were introduced into the multimodal score level fusion. They were employed to optimize the weights associated to single modalities in the weighted sum rule. In an attempt to further enhance the performance of these latest techniques, we proposed a hybrid GA-PSO algorithm.

Our obtained results demonstrated that whereas fusing the modalities using some classical rules gives better accuracy than using other rules, they all outperform their unimodal counterparts. The inclusion of optimization techniques in the weighted sum rule gives a much better performance as compared to the classical combination results. We remark that our proposed hybrid GA-PSO was able to achieve good accuracy recognition results whilst still attaining better computational time as compared to GA.



Conclusion and Future works

This last chapter concludes the thesis by summarizing the work done and results obtained. More importantly, it highlights the most significant points concluded from the experiments. Further, recommendation for future work is presented.

Multibiometric systems have shown to be a very promising trend. They can easily overcome the limitations that unimodal systems suffer from such as non-universality, noisy data, spoof attacks, etc. They offer larger population coverage which deals with the non-universality problem. Also, multimodal biometric systems make it difficult for an intruder to simultaneously spoof the multiple biometric traits of a registered user.

Multibiometrics consist of combining multiple biometric sources for the purpose of improving the performance of user authentication. Different scenarios and sources combinations are possible. The Fusion can occur at different levels in the recognition system as well. For this work, and after performing a thorough literature review, the focus of our investigation was on fusing multi-modalities at the score level.

A critical question was how to integrate the scores. As part of this study, a review of well-established fusion methods has been carried out. When it comes to classical rules, they regard the modalities as the same in all scenarios. The fact is that some data have better accuracy and contain more information than others. To give more value to one modality over the other, weights are associated to each modality. We proposed Genetic Algorithm and Particle Swarm Optimization as techniques to optimize these weights in the fusion of modalities in order to have the "best" performance. In addition, we proposed a hybrid algorithm combining the benefits of GA and PSO to further enhance the performance of optimization and deal with some of the limitations that GA and PSO possess.

To investigate the performance of these techniques, we combined two modalities from public databases (face and fingerprint from NIST (BSSR1), face and voice from BANCA and XM2VTS). The choice of these specific databases stemmed from the fact that the scores they offer are all *non-chimeric*, meaning they come from the same person, which makes the experiments that much closer to real-life applications. We used data with different degrees of degradation from these databases in order to observe whether the techniques deal with them similarly to clean data. The performance of the optimization techniques was compared to that of the classical combination rules using the fused scores EER and ROC curves.

In order to avoid having one modality shadowing the effect of another modality due to large difference in ranges, we mapped the scores to a single range using the *tanh-normalization* scheme.

Following the presentation of motivations and objectives of the work, we introduced the biometric system, its main characteristics and challenges. Further, we presented some of the unimodal system limitations and the ways a multimodal system can elevate most of them. The different fusion sources scenarios and levels were introduced next, with related works done in the literature to support our choices for fusion in this study.

Subsequently, the classical combination rules and optimization techniques as well as our proposed hybrid approach were defined and illustrated in preparation for the experiments to be performed afterwards. The performance metrics and databases that would evaluate these techniques were also presented.

The results we obtained by comparing the EERs and ROC curves of the fused modalities using the rules defined previously lead us to conclude the following:

- Multimodal fusion outperforms the unimodal performance, regardless of the fusion rule.
- Introducing the optimization techniques (GA, PSO, GA-PSO) into the fusion provided a considerable increase in the accuracy of the multibiometric system and outdid the performance of the classical rules.
- Although, the optimization techniques generally give the same performance with all tested-on data, GA is observed to be slower. PSO, on the other hand, is observed to sometimes converge prematurely due to its fast operations.
- Our proposed hybrid GA-PSO provided a faster and more accurate results by taking advantage of the large space search of GA and fast pace of PSO.

Future work

So far, the issue of fusing multimodalities at the score level using optimization techniques has been discussed. The following section presents possible directions for future work.

- To begin with, the results obtained in this research could be furthermore validated by a real-life application. Considering that the Algerian government is pursuing and putting the ground works for biometric passports, it is an opportunity to apply the methods discussed in this work to combine the face, fingerprint and signature modalities.
- Other optimization techniques, such as Ant Colony Optimization (ACO), or other hybrids can be applied to further enhance the performance of the system. A compromise between quality and cost should be the focus of these studies.
- Since most of the fusion schemes have been implicitly designed for the verification scenario and cannot account for missing data (missing modalities or incomplete score lists) that is commonly encountered in multibiometric identification systems, the theme of this study could be further explored in the identification recognition mode.
- On the same note, dynamic fusion algorithms should be designed to address the problem of incomplete or missing input data.
- The performance metrics of a biometric system can be estimated with a high degree of confidence only when the system is tested on a large representative database. However, current multimodal systems have been tested only on small databases containing fewer than 1, 000 individuals. To compensate, the idea of using virtual data was introduced. Although some studies [54], by assuming independence between modalities, validated the use of virtual datasets, others [55] seem to indicate that the recognition performance of a multimodal biometric system evaluated on a virtual multimodal

database is significantly different from the results obtained on a true multimodal database. The issue, hence fore, is yet to be further investigated in future studies.

- Quality-dependent fusion algorithms aim to dynamically combine several classifier outputs as a function of automatically derived biometric sample quality. The issue is the absence of benchmarks specific for multibiometric fusion algorithms that can be used to compare the results and performances of the authentication system. For future work, it would be interesting to investigate in a unified framework for multimodal biometric fusion incorporating quality measures.

References

[1] Ross, A., Nandakumar, K. and Jain, A.K., "Handbook of Multibiometrics", *Springer-Science + Business Media, LLC*, 2006.

[2] H. T. F. Rhodes. Alphonse Bertillon, Father of Scientific Detection. Abelard-Schuman, New York. 1956.

[3] A.K. Jain, P. Flynn, A. Ross, ***Handbook of Biometrics,*** Springer, New York, USA, 2008.

[4] Jain, A.K.; Bolle, R.; Pankanti, S., eds. (1999). *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publications. ISBN 978-0-7923-8345-1.

[5] L. Sirovich and M. Kirby, "Low Dimensional Procedure for Characterization of Human Faces," J. Optical Soc. Am., vol. 4, no. 3, pp. 519–524, 1987.

[6] M. Turk and A. Pentland, "Eigen-faces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991.

[7] D.L. Swets and J. Weng, "Using Discriminant Eigen-features for Image Retrieval," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 831–836, Aug. 1996.

[8] D. Valentin, H. Abdi, A.J. O'Toole, and G. Cottrell, "Connectionist Models of Face Processing: A Survey," Pattern Recognition, vol. 27, no. 9, pp. 1,209–1,230, 1994.

[9] A. Jain, L. Hong, and R. Bolle, "On-Line Fingerprint Verification," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 4, pp. 302–314, Apr. 1997.

[10] A. Ross and A. Jain, "Multimodal biometrics: an overview," *Proceedings of the 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria)*, pp.1221-1224, 2004.

[11] Jain, A. K. Biometric recognition, NATURE, Vol.449, pp. 38 – 40, September 2007.

[12] Biometric Technology Application Manual, Volume1, National Biometric Security Project, 2008.

[13] Wikipedia Website

[14] Monrose, F. and Rubin, A. (1997). Authentication Via Keystroke Dynamics. In *Proceedings of Fourth ACM Conference on Computer and Communications Security*, pages 48-56, Zurich, Switzerland.

[15] Nalwa, V. S. (1997). Automatic On-Line Signature Verification. *Proceedings of the IEEE*, 85(2):215-239

[16] Biometrics Newsportal.com. "DNA Biometrics." Biometric News Portal. 04/03/2008. <http://www.biometricnewsportal.com/dna_biometrics.asp>

[17] Report to the United States Congress. Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability, September 2011.

References

[18] D. W. Aha, D. Kibler, M. K. Albert. Instance based learning algorithms. Machine Learning, Vol.6, pp. 37-66, 1991.

[19] H. Kang, B. Lee, H. Kim, D. Shin, J. Kim. A study on performance evaluation of the liveness detection for various fingerprint sensor modules. Proceedings of KES, pp. 1245–1253, 2003.

[20] J. Galbally, F. Alonso-Fernandez, J. Ortega-Garcia. A high performance fingerprint liveness detection method based on quality related features. Future Generation Computer Systems, Vol. 28(1), pp. 311-321, 2012.

[21] Soh, J., Deravi , F. and Triglia, A., "Multibiometrics and data fusion standardization", in *Encyclopedia of Biometrics*, S.Z. Li and A.K. Jain, New York, Heidelberg: Springer, 2009.

[22] Marcialis, G.L., and Roli, F., "Fingerprint verification by decision-level fusion of optical and capacitive sensor", *Lecture Notes in Computer Science*, vol. 3087/2004, pp. 307-317, 2004.

[23] Jain, A.K., Prabhakar, S., Chen, S. "Combining multiple matchers for a high security fingerprint verification system", *Pattern Recognition Letters*, vol. 20, 11-13, pp 1371–1379, November 1999.

[24] Sanderson, C., Paliwal, K.K., "Information fusion and person verification using speech and face information", *Research Paper IDIAP-RR 02-33*, IDIAP, September 2002.

[25] Zhang, Y.-L., Yang, J., and Wu, H. (2005). A Hybrid Swipe Fingerprint Mosaicing Scheme. In *Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 131-140, Rye Brook, USA.

[26] Choi, K., Choi, H., and Kim, J. (2005). Fingerprint Mosaicking by Rolling and Sliding. In *Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 260-269, Rye Brook, USA.

[27] Yang, R, Paindavoine, M., Abdi, H., and Monopoli, A. (2005). Development of a Fast Panoramic Face Mosaicing and Recognition System. *Optical Engineering*, 44(8).

[28] Liu, X. and Chen, T. (2003). Geometry-assisted Statistical Modeling for Face Mosaicing. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 883-886, Barcelona, Spain.

[29] M. Faundez-Zanuy, "Data fusion in biometrics," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, pp. 34-38, 2005.

[30] Ross, A. and Govindarajan, R. (2005). Feature Level Fusion Using Hand and Face Biometrics. In *Proceedings of SPIE Conference on Biometric Technology for Human Identification II*, volume 5779, pages 196-204, Orlando, USA.

[31] Son, B. and Lee, Y. (2005). Biometric Authentication System Using Reduced Joint Feature Vector of Iris and Face. In Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication

(AVBPA), pages 513-522, Rye Brook, USA.

[32] Kumar, A., Wong, D. C. M., Shen, H. C, and Jain, A. K. (2003). Personal Verification Using Palm-print and Hand Geometry Biometric. In *Fourth International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, pages 668-678, Guildford, UK

[33] Snelick, R., Uludag, U., Mink, A., Indovina, M., and Jain, A. K. (2005). Large Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):450-455.

[34] C. Vielhauer, S. Schimke, V. Thanassis, and Y. Stylianou, "Fusion Strategies for Speech and Handwriting Modalities in HCI," SPIEEI 2005 – Conference on Multimedia on Mobile Devices, pp. 63-71, 2005.

[35] Verlinde, P. and Cholet, G., "Comparing decision fusion paradigms using k-NN based classifiers, decision trees and logistic regression in a multi-modal identity verification application," in Proc. Int. Conf. Audio and Video-Based Biometric Person Authentication (AVBPA), Washington, DC, pp. 188–193, March 1999.

[36] Jain A., Nandakumar, K., Ross, A., "Score normalization in multimodal biometric systems", Pattern Recognition, vol. 38, no. 12, pp. 2270–228, December 2005.

[37] Kinnunen, T, Hautamaki, V, and Franti, P. (2004). Fusion of Spectral Feature Sets for Accurate Speaker Identification. In Ninth Conference on Speech and Computer, pages 361-365, Saint-Petersburg, Russia.

[38] A. Ross and A. Jain, "Multimodal biometrics: an overview," *Proceedings of the 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria)*, pp.1221-1224, 2004.

[39] S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz, "Fusion of face and speech data for person identity verification," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 10, pp. 1065-74, 1999.

[40] National Institute of Standards and Technology, NIST biometric score set (2006). <http://www.itl.nist.gov/iad/894.03/biometricscores/>

[41] N. Poh and S. Bengio, "Database, Protocol and Tools for Evaluating Score-Level Fusion Algorithms in Biometric Authentication", Pattern Recognition, vol. 39, no. 2, (2006), pp. 223-233.

<http://personal.ee.surrey.ac.uk/Personal/Norman.Poh/web/fusion>

[42] F. R. Hampel, P. J. Rousseeuw, E. M. Ronchetti, W. A. Stahel, Robust Statistics: The Approach Based on Influence Functions, John Wiley & Sons, 1986.

[43] Sahoo, S. K., Choubisa, T., & Prasanna, S. R. M. (2013). Multimodal Biometric Person Authentication : A Review, *29*(1).

[44] Holland, J. (1975). Adaptation in natural and artificial systems. Ann Arbor, MI., The University of Michigan Press.

[45] Hassan, R., Cohanim, B. & Weck, O. (2005). A comparison of particle swarm optimization and the genetic algorithm. Proceedings of 46th

References

AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference, Austin, Texas, pp. 18-21.

[46] Kennedy, J. & Eberhart, R. C. (1995). Particle swarm optimization. Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, USA, pp. 1942-1948.

[47] Eberhart, R. C., Simpson, P. K. & Dobbins, R. W. (1966). Computational intelligent PC tools. Boston, MA, Academic Press Professional.

[48] Clerc M and Kennedy J, The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation, 2002, 6(1):58-73.

[49] X. Yang, J. Yuan, J. Yuan, H. Mao, A modified particle swarm optimizer with dynamic adaptation, Journal of Applied Mathematics and Computation, Vol. 189 , pp. 205-1213, 2007.

[50] P.J. Phillips, A. Martin, C.L. Wilson, and M. Przybocki, An Introduction to Evaluating Biometric Systems, IEEE Computer Magazine, pp.56-63, February 2000.

[51] N. Poh, BANCA score database. <http://info.ee.surrey.ac.uk/Personal/Norman.Poh/web/banca_multi>

[52] A. K. Jain, A. Ross, S. Pankanti. Biometrics: A tool for information security. IEEE Transaction on Information Forensics and Security, Vol.1 (2), pp.125-143, 2006.

[53] S. Prabhakar, A.K. Jain, Decision-level fusion in fingerprint verification, Pattern Recognition. 35 (4) (2002) 861–874.

[54] Garcia-Salicetti, S., Mellakh, M. A., Allano, L., and Dorizzi, B. (2005). A Generic Protocol for Multibiometric Systems Evaluation on Virtual and Real Subjects. In *Fifth International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, pages 494-502, Rye Brook, USA.

[55] Tobias Scheidat and Claus Vielhauer, "Analyzing a multimodal biometric system using real and virtual users", Proc. SPIE 6505, Security, Steganography, and Watermarking of Multimedia Contents IX, 650512 (February 27, 2007).