**People's Democratic Republic of Algeria**
**Ministry of Higher Education and Scientific Research**
**Djillali Liabes University of Sidi Bel Abbes**
**Faculty of Technology**
**Department of Computer Science**

# A Hierarchical Fusion Strategy in Multibiometric Authentication Systems

By

Youssef Elmir

A doctorate thesis in Computer Science
submitted in filfullment of the requirements
for the degree of Doctor of Sciences

Defended in front of the members of jury composed of:

| | **First Name** | **Last Name** | **University** | **Quality** |
|---|---|---|---|---|
| Pr. | Abderrahmane | Yousfate | Sidi Bel Abbès | President |
| Pr. | Mohamed Kamel | Faraoun | Sidi Bel Abbès | Examiner |
| Pr. | Ghalem | Belalem | Oran - Es Senia | Examiner |
| Dr. | Abdelmalek | Amine | Saida | Examiner |
| Dr. | Réda Mohamed | Hamou | Saida | Examiner |
| Pr. | Zakaria | Elberrichi | Sidi Bel Abbès | Supervisor |
| Dr. | Réda | Adjoudj | Sidi Bel Abbès | Invited |

**2014-2015**

# *Abstract*

Multi-biometric authentication systems that fuse information from multiple biometric sources, have gain more space, in the field of security and more precisely in the field of recognition and verification of person identities, this, due to their ability to overcome the limitations of uni-biometrics as the non-universality of the biometric traits, the noise at biometric sensors level and the large intra-user variation ... etc.

In this thesis, the case of the fusion of biometric data is inspected in all these circumstances in order to release a multi-biometric system based on biometric fusion of face, fingerprint, voice, online signature or finger vein. The majority of multi- biometric systems proposed in the state of the art of authentication systems are based on the fusion or integration in serial mode or parallel mode, however, this thesis consider to explore hierarchical fusion strategy to benefit from the advantages of both of serial and parallel modes and to improve the overall recognition rate of the authentication system.

In addition, biometric performance enhancement is a chalange. In this thesis, a multimodal biometric system based on hierarchical strategy of fusion, is presented. This strategy combines several biometric traits based on multi-level biometric fusion hierarchy. The multi-level biometric fusion includes a prior-to-matching fusion with optimal feature selection and an after-matching fusion based on the similarity of minimum of distances. The proposed solution enhances the overall recognition performances based on feature selection and reduction using principal component analysis (PCA) or Linear discriminant analysis (LDA).

## *Résumé*

Les systèmes d'authentification multi-biométriques qui fusionnent les informations de plusieurs sources biométriques, ont gagné plus d'espace dans le domaine de la sécurité et plus précisément dans le domaine de la reconnaissance et de vérification de l'identité des personnes, ce, en raison de leur capacité à surmonter les limites de la biométrie uni-modale comme la non-universalité des traits biométriques, le bruit au niveau des capteurs biométriques et la grande variation intra-utilisateur ... etc.

Dans cette thèse, le cas de la fusion de données biométriques est inspecté dans toutes ces circonstances afin de réaliser un système multi-biométrique basé sur la fusion biométrique du visage, l'empreinte digitale, la voix, la signature en ligne ou la veine de doigt. La majorité des systèmes multi-biométriques proposées dans l'état de l'art des systèmes d'authentification sont basés sur la fusion ou l'intégration en mode série ou en mode parallèle, cependant, nous considérons dans cette thèse à explorer la stratégie de fusion hiérarchique pour bénéficier des avantages des deux modes; série et parallèle et améliorer le taux de reconnaissance global du système d'authentification.

En outre, l'amélioration de la performance biométrique est une tâche difficile. Dans cette thèse, un système biométrique multimodal basé sur une stratégie de fusion hiérarchique, est présenté. Cette stratégie repose sur une combinaison de plusieurs caractéristiques biométriques en utilisant une hiérarchie de fusion biométrique multi-niveaux. La fusion biométrique multi-niveaux comprend une fusion de pré-classification avec la sélection optimale des caractéristiques et une fusion de post-classification basée sur la similitude de maximum de scores. La solution proposée améliore les performances de reconnaissance biométrique basée sur la sélection et la réduction appropriée de caractéristiques telles que l'analyse principale des composantes et l'analyse discriminante linéaire, autant que pas tous les composants des vecteurs de caractéristiques prennent en charge le degré d'amélioration des performances.

ملخص


أنظمة التحقق أو التعرف على الهوية متعددة القياسات الحيوية و التي تدمج المعلومات من مصادر حيوية متعددة، استفادت من مساحة أكبر في مجال الأمن وخصوصا في مجال التعرف والتحقق من هوية الأشخاص، و هذا نظرا لقدرتها على التغلب على القيود المفروضة على القياسات الحيوية الأحادية مثل عدم شمولية الصفات الحيوية، التشويش على مستوى أجهزة الاستشعار الحيوية و التغيرات الحيوية الكبيرة لنفس المستعمل ... الخ

في هذه الأطروحة، تم البحث في حالة دمج البيانات الحيوية في جميع هذه الظروف من أجل انجاز نظام التحقق و التعرف على الهوية بقياسات حيوية متعددة استنادا الى دمج القياسات الحيوية للوجه، بصمة الاصبع، الصوت، الامضاء أو وريد الأصبع. وتعتمد غالبية أنظمة التحقق من الهوية متعددة القياسات الحيوية المقترحة في أحدث التقنيات الصناعية على أنظمة الدمج أو التكامل في نمط تسلسلي أو متوازي. و عكس ذلك فقد قمنا في هذه الأطروحة بالاعتماد على استكشاف استراتيجية الدمج الهرمية للاستفادة من مزايا النمطين التسلسلي والمتوازي وتحسين معدل التحقق الشامل لنظام التعرف على الهوية.

بالإضافة إلى ذلك، تحسين أداء القياس الحيوي هو تحد صعب. في هذه الأطروحة، تم تقديم نظام التعرف و التحقق من الهوية متعدد القياسات الحيوية استنادا على استراتيجية دمج هرمية. وتعتمد هذه الاستراتيجية على مجموعة من عدة صفات حيوية باستخدام دمج هرمي متعدد المستويات للقياسات الحيوية. يشمل دمج القياسات الحيوية متعدد المستويات دمج قبلي مع اختيار الخصائص المثلى و دمج بعدي على أساس أقصى درجات التشابه. الحل المقترح يعزز أداء التعرف الحيوي استنادا إلى اختيار و تحديد خاصية مناسبة مثل تحليل المكون الرئيسي والتحليل الخطي المميز، تماشيا مع انعدام دعم كل مكونات الخصائص لدرجة تحسين الأداء.

*To my parents*

*To my brothers and sisters*

*And to the apples of my eyes*

*My wife and my little girl Fatima Tasnim*

# *Acknowledgment*

First and foremost, I would like to give thanks to my lord, Allah the almighty, for giving me the opportunity to complete this thesis. I would like to take this opportunity to acknowledge these people who gave me encouragement, cooperation and support along all this hard work, perseverance and continuous efforts. I would like firstly to express my deepest gratitude, especially to my supervisor Pr. Zakaria Elberrichi for giving me the opportunity to pursue this scientific research at Djillali Liabès University of Sidi Bel Abbès. I would like also to express my appreciation for my co-supervisor Dr. Réda Adjoudj, who helped me to understand the concepts of scientific research, especially in biometrics field. I am fortunate to have them as my supervisors who are never tired in giving me invaluable support, guidance and encouragement.

Without forgetting to thank the president of the jury Pr. Yousfate, and all members of the jury Pr. Belalem, Pr. Faraoun, Dr. Amine and Dr. Hamou, who have accepted to review this thesis, made efforts and spent precious time to evaluate and enhance its final version.

I also like to take this opportunity to acknowledge Pr. Abbes Amira, Dr. Somaya Al-maadeed and Dr. Abdelaâli Hassaïne for their support, advices and assistance.

Last but not least, my deepest thankfulness goes to my parents, wife for their endless love, understanding, sacrifice, care and support.

Thank you very much!

## *List of figures*

*List of tables*

# *List of publications*

## *1 Journal papers*

1. Y. Elmir, Z. Elberrichi, R. Adjoudj, “Multimodal Biometric Using a Hierarchical Fusion of a Person’s Face, Voice, and Online Signature,” Journal of Information Processing Systems (JIPS), vol. 10, no. 4, December 2014.

## *2 Conference papers*

1. Y. Elmir, Z. Elberrichi, R. Adjoudj, “A Hierarchical Fusion Strategy based Multimodal Biometric System,” the $14^{th}$ International Arab Conference on Information Technologies (ACIT’13), December 2013, Khartoum, Sudan.

2. Y. Elmir, O. Ghazaoui, F. Boukenni, “Multimodal Biometrics System's Resistance to Noise,” Conférence nationale sur l'informatique et les Technologies de l'Information et de la Communincation(CTIC’12), November 2012, Adrar, Algeria.

3. Y. Elmir, S. Al-Maadeed, A. Amira, A. Hassaine, “A Multi-Modal Face and Signature Biometric Authentication System Using a Max-of-Scores Based Fusion,” the 19th International Conference on Neural Information Processing (ICONIP’12), November 2012, Doha, Qatar.

4. Y. Elmir, S. Al-Maadeed, A. Amira, A. Hassaine, “Multi-modal biometric authentication system using face and online signature fusion,” Qatar Foundation Annual Research Forum (QFARF'12), October 2012, Doha, Qatar.

5. Y. Elmir, Z. Elberrichi, R. Adjoudj, “Score level fusion based multimodal biometric identification,” Sciences of Electronics, Technologies of Information and Telecommunications (SETIT’12), March 2012, Sousse, Tunisia.

# *Table of content*

# *General introduction*

High security systems are based on the automatic authentication by the recognition of people using images containing the face, fingerprint, iris and the voice signal; otherwise, it is the recognition of persons by their biometric/biophysical characteristics.

Recently, automatic verification of person's identity becomes a tool increasingly important in many applications such as access to automated services (Automated Teller Machines ATMs …) and protected areas (banks …).

Different techniques are available widely used in this context, for example: passwords, swipe cards, personal identification numbers (PIN), but, the only way that really check the identity is a best combination of some possession (possession of the good magnetic card) and some knowledge (password).

It is well known, these simple mechanisms of access control can easily lead to abuse (misuse), induced by lost or stolen magnetic card and its corresponding PIN.

Hence, a new type of methods is being born, based on the so-called characteristics or biometrics, such as voice, face (profile included), the eye (iris, retina), fingerprint, palm of the hand, the shape (geometry) of the hand or other physiological or behavioral characteristics of the person to be checked preferably unique and measurable.

In general, biometric measurements, and in particular the biometric measures that are non-invasive/easy to use (voice, video) are very interesting because they have the advantage that they cannot be lost or forgotten, and they are really personal (they cannot be given to someone else), since they are based on the measurement of physical aspect of the person.

Additionally, applications using a standard technique (password, magnetic card) claiming a certain identity can then be checked using one or multiple biometric traits.

If only one biometric trait is used, the obtained results could be not good enough. This is because the ease to use traits tends to change with time for the same person. This is especially true for the voice that shows a significant variability of intra speaker. A possible solution to overcome this problem is to use not one but multiple biometric traits.

## *1 Why hierarchical fusion?*

Although biometric techniques promise to be very powerful, currently they cannot guarantee an excellent recognition rate with uni-modal biometric systems based on a unique biometric signature. In addition, these systems are often affected by the following problems [1]:

- Noise introduced by the sensor: the noise may be present in the acquired biometric data, this being mainly due to a faulty sensor or poorly maintained. For example, the accumulation of dust on a fingerprint sensor, a bad camera focus blurs resulting in images of face or iris … etc. The recognition rate of a biometric system is very sensitive to the quality of the biometric sample and noisy data can seriously affect the accuracy of the system [2],
- Non-universality: if each individual of a target population is able to present a biometric modality for a given, then this method is called universal

system. The principle of universality is one of the basic conditions necessary for biometric recognition module. However, not all biometric modalities are really universal. The National Institute of Standards and Technology (NIST) reported that it was not possible to obtain a good fingerprint for about 2% of the population (persons with disabilities related to the hand of individuals performing repeated many handicrafts ... etc.). [3]. Thus, such people cannot be enrolled in a verification system fingerprint. Similarly, people with very long eyelashes and those with abnormalities of the eyes or eye diseases (such as some glaucoma and cataracts) cannot provide iris or retina pictures of good quality for recognition stations. The non-universality leads to enrollment errors ("Failure to Enroll" or FTE) and / or capture errors ("Failure to Capture" or FTC) in a biometric system,

- Lack of individuality: the features extracted from biometric data from different individuals can be relatively similar. For example, a certain portion of the population may have an almost identical facial appearance due to genetic factors (father and son, twins, etc...). This lack of uniqueness increases the false acceptance rate ("False Accept Rate" or FAR) of a biometric system,
- Lack of invariant representation: the biometric data acquired from a user during the recognition phase are not identical to the data that were used to generate the model of the same user during the enrollment phase. This is known as the "intra-class variations." These variations may be due to poor user interaction with the sensor (e.g., changes in pose and facial expression when the user stands in front of a camera), the use of different sensors during the enrollment and verification, changing conditions of the ambient environment (e.g., changes in illumination for a facial recognition system) or to changes inherent in biometric modality (e.g., wrinkles due to age, presence of hair in the face image, the presence of scars in a fingerprint ... etc.).. Ideally, the features extracted from the biometric data should be

relatively invariant to these changes. However, in most biometric systems, these features are not invariant and, therefore, complex algorithms are required to account for these variations. Large intra-class variations typically increase the false rejection rate ("False Reject Rate" or FRR) of a biometric system,

- Sensitivity to attacks: although it seems very difficult to steal the biometric modalities of a person, it is still possible to bypass a biometric system using biometric modalities usurped. Studies [4] [5] showed that it was possible to fabricate false fingerprints rubber and use it to counter a biometric system. Behavioral biometric modalities such as signature and voice are more susceptible to this kind of attack than physiological biometric modalities.

Thus, because of all these practical problems, the error rate associated with uni-modal biometric systems is relatively high. This makes them unacceptable for deployment of safety-critical applications. To overcome these drawbacks, a solution is the use of multiple biometric modalities in one system; it is called a multimodal biometric system.

In this thesis, the choice was made to use a hierarchical fusion strategy of face fingerprint, finger vein, voice and/or online signature. Why these modalities? First, the modality of the fingerprint is certainly a well known and acceptable biometric, but, it can be intrusive, it means it can reach the privacy of the individual, however, it is considered one of the most accurate, and by combination with finger vein modality, it will be more secured, this modality is proved in the 1990's by researchers that have found that the venous system was unique to each individual [6]. The veins of the finger are a network of blood vessels under the skin and they can be used for biometric identification [7] [8]. Additionally, the modality of the voice is thought to be one of the most useful biometrics in some situations where the voice is the only collectable data (e.g., phone communications), but its performances dramatically decrease in non-cooperative situations. In the other hands, the modality of the face is non-intrusive, it is one of the most natural ways to recognize a person, it allows to perform work on the fly and its cost of deployment is relatively low: a single camera connected to a computer may suffice. However, face recognition is still relatively sensitive to the surrounding environment to provide a very high recognition rate. Furthermore, the modality of the online signature is very accurate, but its performances depend on the quality of the dynamic collected data. The choice of combination of methods is confirmed by Zephyr analysis (See fig. 1).

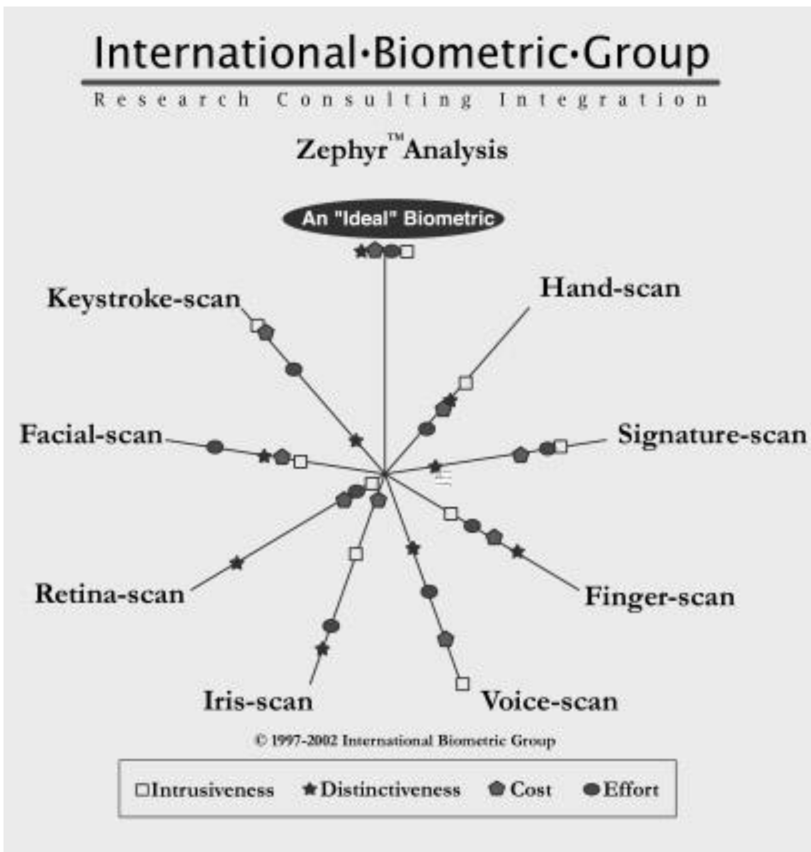


Figure 1. Comparison of different methods on four main criteria: the intrusiveness, the discriminating power, cost and effort.

## 2 Biometric systems and its operating modes

### 2.1 Characterizing of a biometric system

A typical biometric system can be represented by four main modules [9]:

1. The capture module is responsible for the acquisition of biometric data of an individual (this can be a camera, a fingerprint reader, a security camera ... etc,

2. The feature extraction module takes as input the biometric data acquired by the capture module and extract only the relevant information to form a new data representation. Ideally, this new representation is supposed to be unique for each person and relatively invariant to changes in intra-class,

3. The matching module compares all the extracted features with the stored or enrolled ones in the database of the system model and determines the degree of similarity (or difference) between the two ... etc,

4. The decision module verifies the identity stated by a user and determines the identity of a person based on the degree of similarity between the extracted features and the stored or enrolled model(s).

### 2.2 Operating modes of a biometric system

Biometric systems can provide three modes of operation, namely, enrollment, authentication (or verification) and identification. In what follows, fig. 2, fig. 3 and fig. 4 illustrate an example of a biometric system using fingerprint as modality.

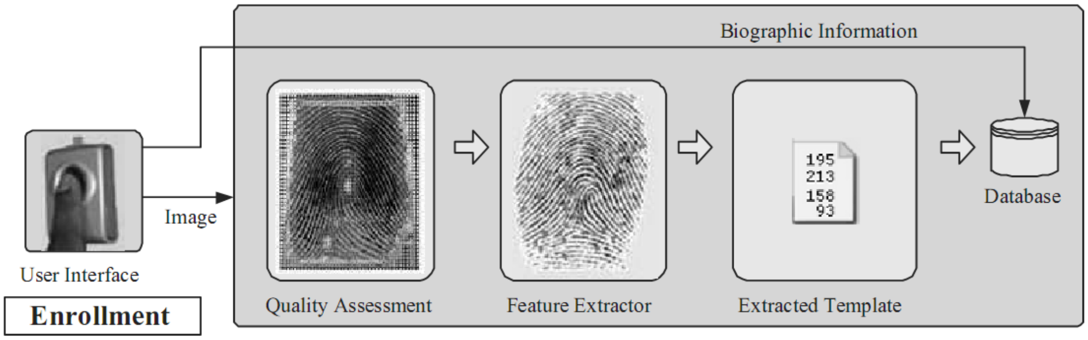


Figure 2. Enrollment mode of a biometric system.

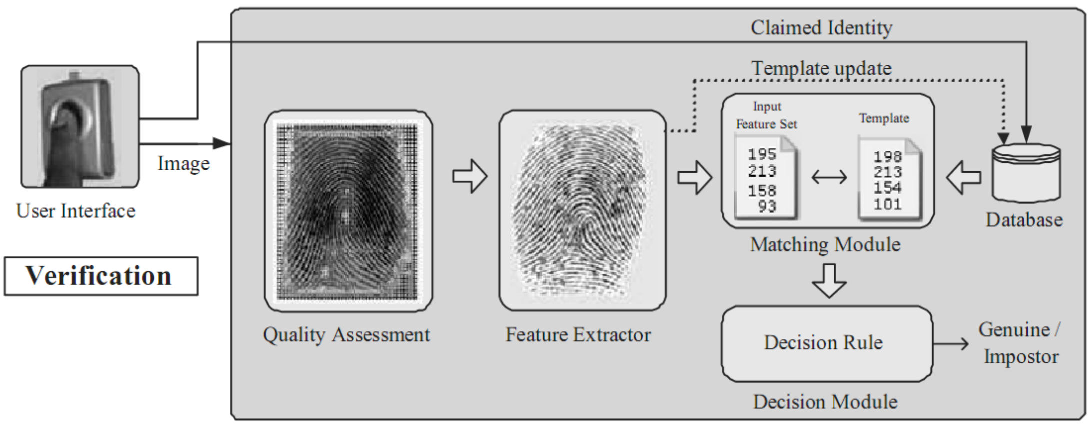


Figure 3. Verification mode of a biometric system.

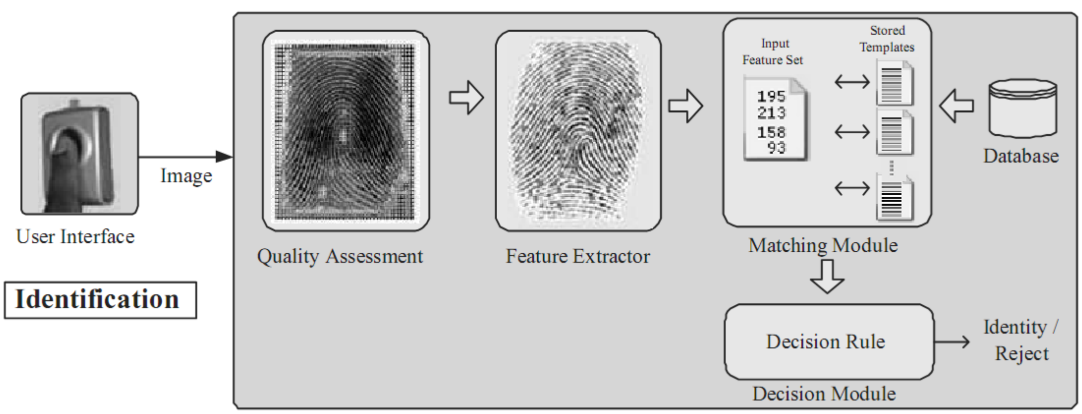


Figure 4. Identification mode of a biometric system.

The quality assessment module determines if the sensed data can be effectively used by the feature extractor. Note that the process of quality assessment in itself may entail the extraction of some features from the sensed data [**10**].

Enrollment (See fig. 2) is the first phase of any biometric system, it is the stage during which a user is registered in the system for the first time and where one or more biometric modalities are captured and stored in a database. This recording can be accompanied by the addition of biographical information in the database.

Depending on the application context, a biometric system may operate in the verification or identification mode (see fig. 3 and fig. 4). In verification, the system checks a person's identity by comparing the captured biometric data with her own biometric template(s) stored in a database. An individual who claims an identity via a user name or a smart card, and the system conducts a one-to-one comparison to decide if the claim is true or not (e.g., "Does this biometric data belong to that individual?"). Verification is typically used for positive recognition, where the aim is to prevent multiple people from using the same identity [10].

For identification purposes, the system identifies an individual by searching the reference templates of all the registered individuals in the database for a match. Therefore, the system makes a one-against-all comparison to determine an individual's identity (or fails if the individual have no reference templates in the system database) without the individual having to claim his identity (e.g., "Whose biometric data is this?"). Identification is a critical stage in negative identification system where it decides if the person is who he (implicitly or explicitly) has to be. The purpose of negative recognition is to prevent a unique person from using more than one identity. Identification can be used also in positive recognition for convenience (the individual has not to claim his identity). While traditional methods of personal recognition such as passwords, PINs, keys, and tokens may work for positive recognition, negative recognition can only be established through biometrics [10].

## *3 Performance of a biometric system*

First, to understand how to evaluate the performance of a biometric system, three main criteria must be clearly defined;

1. The first criterion is called the false rejection rate ("False Reject Rate" or FRR). This rate represents the percentage of people expected to be recognized but are rejected by the system,

2. The second criterion is the false acceptance rate ("False Accept Rate" or FAR). This rate represents the percentage of people not expected to be recognized but they are still accepted by the system,

3. The third criteria is known as the equal error rate ("Equal Error Rate" or EER). This rate is calculated from the first two criteria and is a point of current measurement performance. This is where FAR is equal to FRR, that is to say, the best compromise between false rejection and false acceptance.

Fig. 5 shows the FRR and FAR from distributions genuine and impostor scores while the EER is shown in fig. 6.

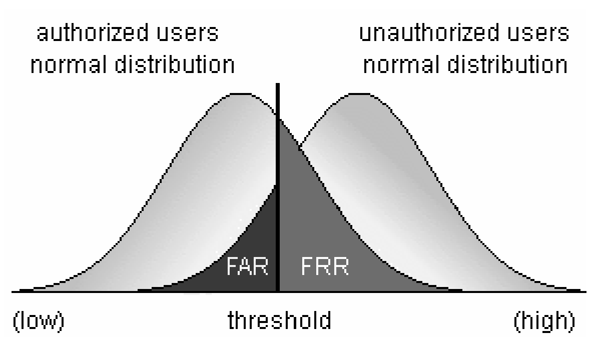


Figure 5. Illustration of FRR and of FAR.

Depending on the nature (authentication or identification) of the biometric system, there are two ways to evaluate the performance:

- When the system operates in authentication mode, it uses what is called a Receiver Operating Characteristic (ROC) curve. The ROC curve (See fig. 6) provides a false rejection rate based on the false acceptance rate [11]. Moreover, the system is more efficient when it has a high overall rate of recognition.

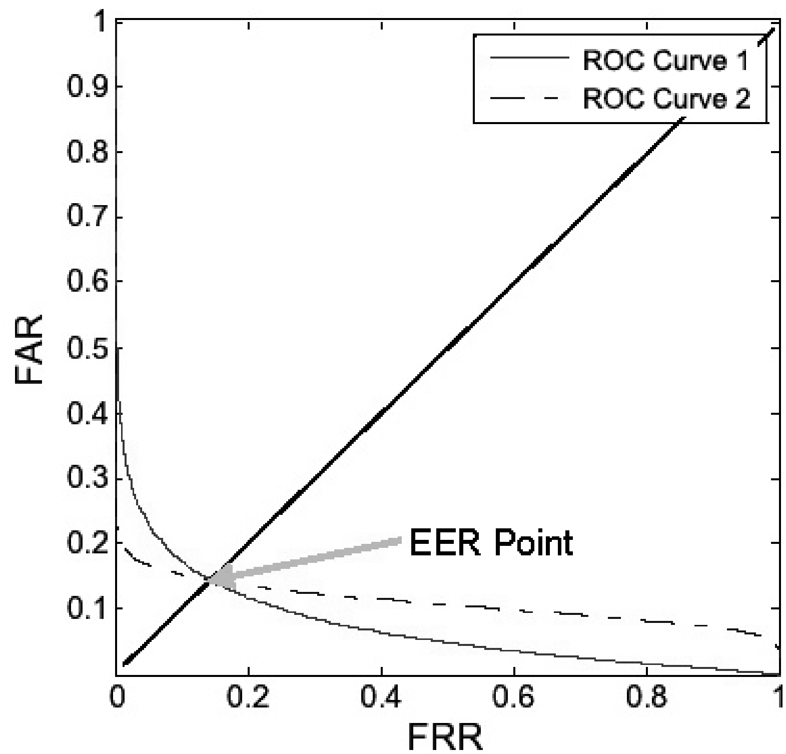


Figure 6. ROC curves.

- However, in the case of a system used in identification mode, it uses what is called a Cumulative Match Characteristic (CMC) curve. CMC curve (See fig. 7) shows the percentage of people recognized on the basis of a variable that is called rank [12]. A system recognizes the rank 1 when it chooses the closest image as a result of recognition. And a system recognizes at the rank 2, when it chooses among two images that best match the input image … etc. So increasing the rank linked the correspondent recognition rate to a low level of security.

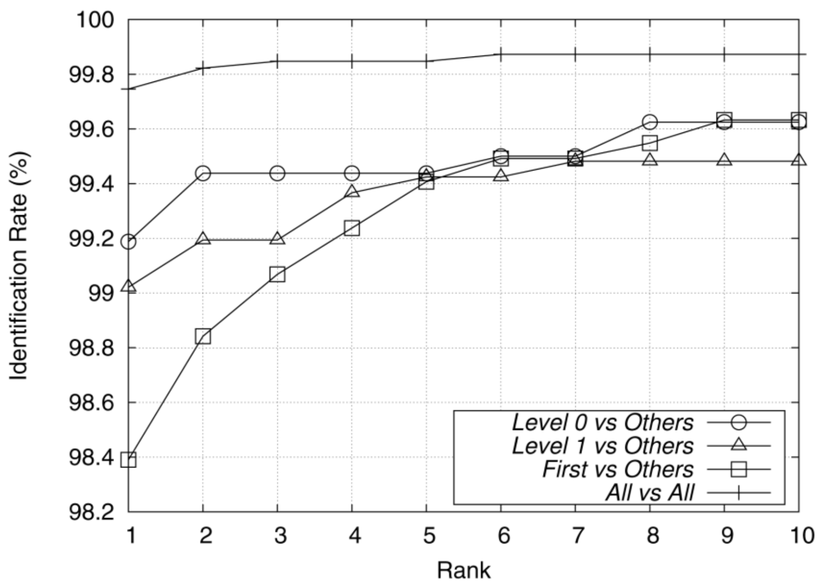


Figure 7. CMC curves.

Finally, the CMC curve is just another way to display the performance of a biometric system and can also be calculated from the FAR and FRR. A comparative study clarifying the relationship between CMC and ROC curves can be found in [13].

## *4 Problematic*

The challenge that must be faced is to improve the performance of the biometric security system by finding an original method of combiner as the recognition rate is higher than the fused modalities maximum recognition rate rules adopted separately. Care should be taken to analyze the execution speed of the application and the overall complexity of the calculations that are supposed to be respectively slower and heavier than a uni-modal biometric system. This can be overcome by using hierarchical fusion strategy to obtain very closer recognition rates in lower duration of time, this is because that this strategy of biometric fusion reduces the number of biometric traits that are necessary for authentication and it uses all of them only if the authentication is not confirmed.

Throughout this thesis, the present authors will strive to be critical and objective in reporting the difficulties that have been encountered with certain methods or by offering alternative solutions to certain techniques. This thesis tries to show the scientific progress that has taken place since the impregnation of the state of the art to the development of a new approach of multimodal fusion.

## *5 State of the art*

Fierrez et al [14] experimentally compare some fusion strategies and use as a monomodal platform, they proposed a global face appearance representation verification system, their minutiae-based fingerprint verification system, and their on-line signature verification system based on Hidden Markov Model (HMM) modeling of temporal functions; all models are applied to the MCYT multimodal database. They proposed and discussed a new strategy to generate a multimodal combined score based on Support Vector Machine (SVM) classifiers from which they derived and evaluated fusion of user-independent and user-dependent schemes. In another study, Dorizzi [15] has presented different types of score fusion methods discussed their complexity when used to model the systems-scores and proposed to a comparison of score fusion methods using a large multimodal database, BioSecure DS3.

Zhang et al [16] proposed a hierarchical fusion scheme for low quality images under uncontrolled situations. In training, they adopted canonical correlation analysis (CCA) to construct a statistical mapping from face to iris in pixel level. In testing, firstly the questioned face is used to obtain a subset of candidate reference templates via regression between the questioned face and irises templates, then ordinal representation and sparse representation are done on these candidate templates for iris recognition and face recognition respectively. Finally, score level fusion based on min-max normalization is done to make the final decision. Their experimental results show the outperforming performance of their proposed approach that achieves 100% of accuracy when the population used is 58 individuals.

Singh et all [17] present a two level hierarchical fusion of face images captured under visible and infrared light spectrum to improve the performance of face recognition. They proposed to combine two face images captured from different spectrums using DWT based algorithm. The amplitude and phase features are extracted from the combined image using 2D log polar Gabor wavelet. An algorithm based on adaptive SVM selects either the amplitude or phase features to generate a fused feature set to improve the recognition. The performance is observed under the worst case scenario of using unique training images. Experiments using Equinox face database proves that the fusion of visible light and short-wave IR spectrum images achieved the best recognition performance with an equal error rate of 2.86%.

Table 1 summarizes these works and their performance on well-known databases.

Table 1. Review of related works

| Reference | Method | FAR (%) | FRR (%) | EER (%) | Database | Population |
|---|---|---|---|---|---|---|
| [14] | Face + sign (sum rule based) | ≈ 40 | ≈5 | | XM2VTS | 50 |
| [14] | SVM (face,sign) user-indep | ≈ 30 | ≈10 | | XM2VTS | 50 |
| [15] | SVM classifier (face and sign) | | | 5.54 | Biosecure | 500 |
| [16] | Hierarchical Fusion of Face and Iris | | | | Their own database | 82 |
| [17] | A two level hierarchical fusion of face images + SVM classifier | | | 2.86 | Equinox face database | 91 |

## 6 Thesis contributions

In this thesis, the field of biometrics is identified, exploring some strategies of fusion of multiple biometric sources in physiological recognition systems and their relations with the different paradigms of artificial intelligence, while keeping an overview and a reference point on the different fusion levels and on the geometric techniques, belonging to the state of the art.

In addition, the goal of this work is to improve the efficiency and the rate of recognition by the use of a hierarchical combination/fusion of two or three different biometric traits in one biometric system in order to make it more robust.

Multimodal biometric systems have been continually developed and many strategies of fusion have been proposed in literatures. Most fusion strategies focus on the score level, but less information can be used at this level. At feature level, some strategies use concatenation methods by using feature vectors which cause dimensionality problems. This thesis has provided a novel methodology and pioneered a direction for research that will enable the development of a new fusion approach to multimodal biometrics capable of yielding better performance in identification and verification process. In addition, the proposed methods overcome the limitations associated with conventional feature fusion approaches in multimodal as well as uni-modal biometrics.

## 7 Thesis organization

The rest of this thesis is organized as follows:

- Chapter 1 deals with the feature acquisition in real environments, different used methods and stages of feature extraction, selection, reduction and representation.

- Chapter 2 presents the state of the art of multimodal fusion strategies, focusing primarily on fusion at feature level and at score level because these two methods have been chosen to be combined.

- The experimental scenarios and results obtained on some datasets using a similarity distance are described in Chapter 3.

- Finally, an overall conclusion of this work will be drawn and some future prospects.

# *Chapter 1*

# *Feature extraction, selection and representation*

This chapter presents feature extraction techniques that are used to extract and represent features of face, fingerprint, finger vein, voice and online signature, these features will be used as one fused feature. Feature extraction is important for the success of the recognition and classification process, and should extract useful information while reducing noise and avoiding redundant data with fast computation. The features given by the extraction process are used to gain statistical information using supervised and parametric statistic techniques. In this chapter, feature extraction is proposed to represent local features that can be used in the fusion method subsequently proposed in Chapter 2.

## *1 Biometric modalities*

### *1.1 Face*

Face recognition algorithms can be classified into two broad categories according to feature extraction schemes for face representation: feature-based methods and appearance-based methods [18]. Properties and geometric relations such as the areas, distances, and angles between the facial feature points are used as descriptors for face recognition. On the other hand, appearance-based methods consider the global properties of the face image intensity pattern. Typically appearance-based face recognition algorithms proceed by computing basis vectors to represent the face data efficiently. In the next step, the faces are projected onto these vectors and the projection coefficients can be used for representing the face images. Popular algorithms such as PCA, LDA, ICA, LFA, Correlation Filters, Manifolds and Tensorfaces are based on the appearance of the face. Holistic approaches to face recognition have trouble dealing with pose variations. Building image face mosaics like those in [19] [20] have been introduced to deal with the pose variation problem. Several of the popular face recognition algorithms are reviewed as well as Elastic Bunch Graph Matching (EBGM) approach [21].

Figure 8. The first six basis vectors of Eigenfaces [**10**].

### *1.2 Fingerprint*

A fingerprint is the representation of the epidermis of a finger: it consists of a pattern of interleaved ridges and valleys [22]. In a fingerprint image, valleys are bright whereas ridges (lines of ridge) are dark. Ridges and valleys often run in parallel; sometimes they bifurcate and sometimes they terminate. When checked at the global level, the fingerprint exhibits one or more points where the ridge lines assume distinctive forms. These regions (called singularities or singular points) may be classified into three typologies: loop, delta, and whorl. Singular points belonging to loop, delta, and whorl types are typically characterized by ∩, Δ, and O forms, respectively. The core point (used usually to pre-align fingerprints) corresponds to the center of the north most (uppermost) loop singularity. At the local level, other features, called minutiae can be extracted from the fingerprint patterns. Minutia relies to the various ways in which the ridge lines can be discontinuous. For example, a ridge may come to an end (termination), or can be divided into two ridges (bifurcation) (see fig. 9). Although several kinds of minutiae can be considered, usually only a coarse classification (into these two types) is adopted to deal with the difficulty in automatically discerning the different types with high performance.

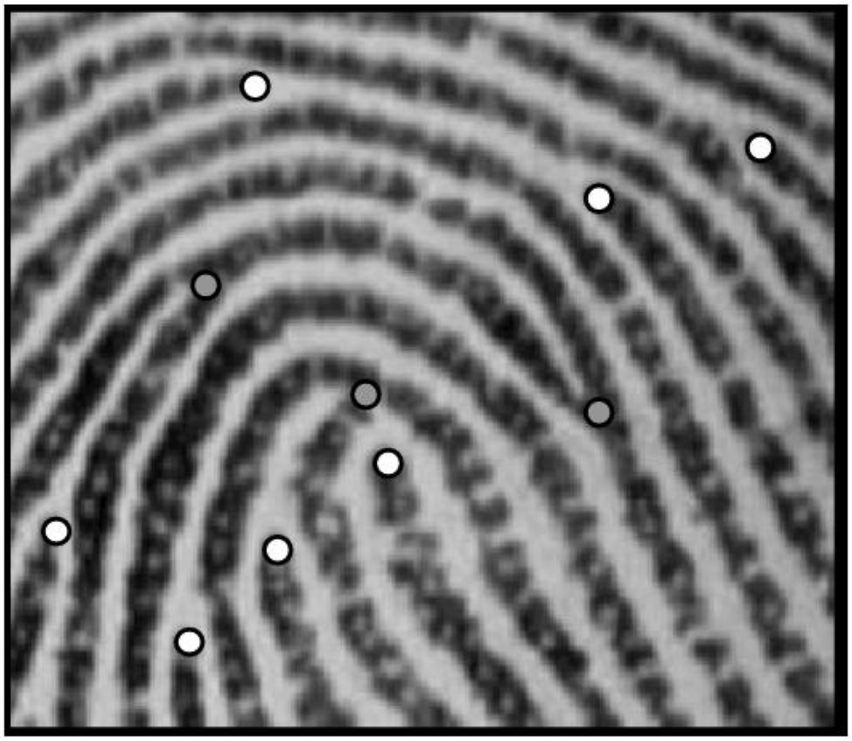

Figure 9. Termination (white) and bifurcation (gray) minutiae in a sample fingerprint [**10**].

Jain et al. [23] proposed a local texture analysis technique where the fingerprint area of interest is tessellated with respect to the core point (see fig. 9).

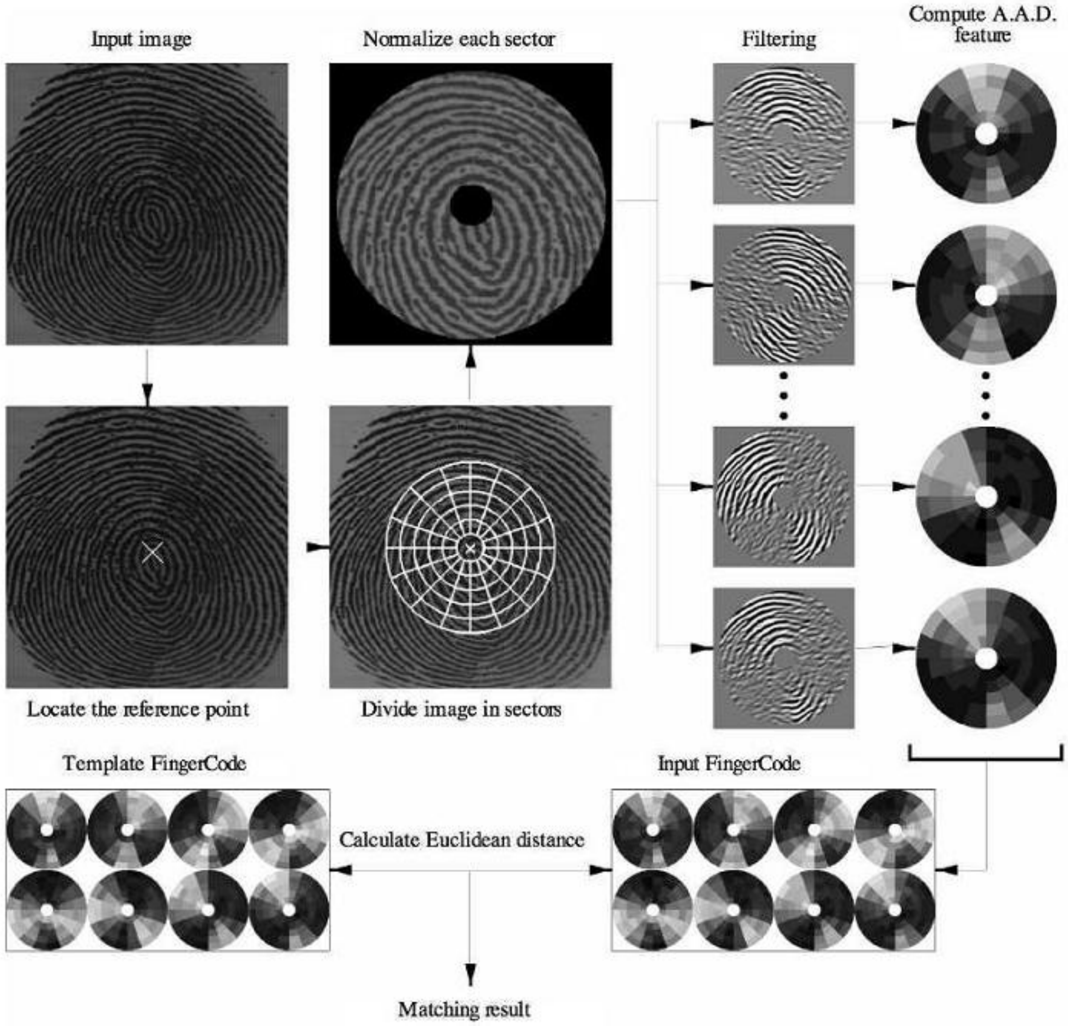


Figure 10. System diagram of Jain et al.'s FingerCode approach [25].

A feature vector (called the FingerCode) is composed of an ordered enumeration of the features extracted from the local information contained in each sector specified by the tessellation. Thus the feature elements capture the local texture information and the ordered enumeration of the tessellation captures the global relationship among the local contributions. Matching two fingerprints is then translated into matching their respective FingerCodes, which is simply performed by computing the Euclidean distance between two FingerCodes. Several approaches have been recently proposed in the literature where non-minutiae features such as spatial relationship of the ridge lines [24], local orientation [25]

[26] and local density [27] [28]are used in conjunction with the minutiae to improve the overall system performance (see fig. 10).

### *1.3 Finger vein*

As a newly emerging modality, finger vein proved that is an efficient biometric for human recognition [29]. Finger veins are located in the interior of the living body; biometric system that uses finger vein identification is protected from forgery and can be less affected by the outer pattern surroundings (dirtiness, humidity, etc). By comparing finger vein with the other biometric modalities (e.g. fingerprint, palm print, iris, face), it has the advantages as : the ease of collection with contactless operation using small size of sensor and low cost [30] [31]. Thus, this modality became widely considered as very promising biometric trait [29]. Most of the finger vein based recognition approaches use the features extracted from the segmented blood vessel network [32] [33] [34]. For example, Miura et al proposed a repeated line tracking based finger vein extraction method [32]. Song et al. proposed the mean curvature method, which considered the vein image as a geometric form and then find the valley-like structures with negative mean curvatures [34]. These methods are good in term of accuracy if the blood vessel networks are well and properly segmented. Nevertheless, errors of segmentation can occur in the process of feature extraction because of the degrading quality of images caused by skin scattering or optical blurring [35]. Moreover, the performance of segmentation of finger vein depends on the image conversion, scale, rotation and uneven illumination. Hence, the global performance of these methods can be degraded when the networks are not well segmented. To overpass the difficulty in the segmentation stage, there are some local feature-based methods [35] [36] used for finger vein identification. For example, local binary pattern set is used in [35] [36] for feature representation. Since the finger veins are beneath finger, LLBP [37] is used for finger vein recognition in [38]. Experimental results proved that the better performance of LLBP than LBP. Hence, LLBP is a good feature coding method which can collect the directional and local features.

Nevertheless, LLBP is limited in extracting only vertical and horizontal lines. Thus, the effective data in a veins image cannot be extremely used. Moreover, the veins are randomly existed in the finger in a lot of type of orientations. These lines may not contain a discriminative data for template comparison [39].

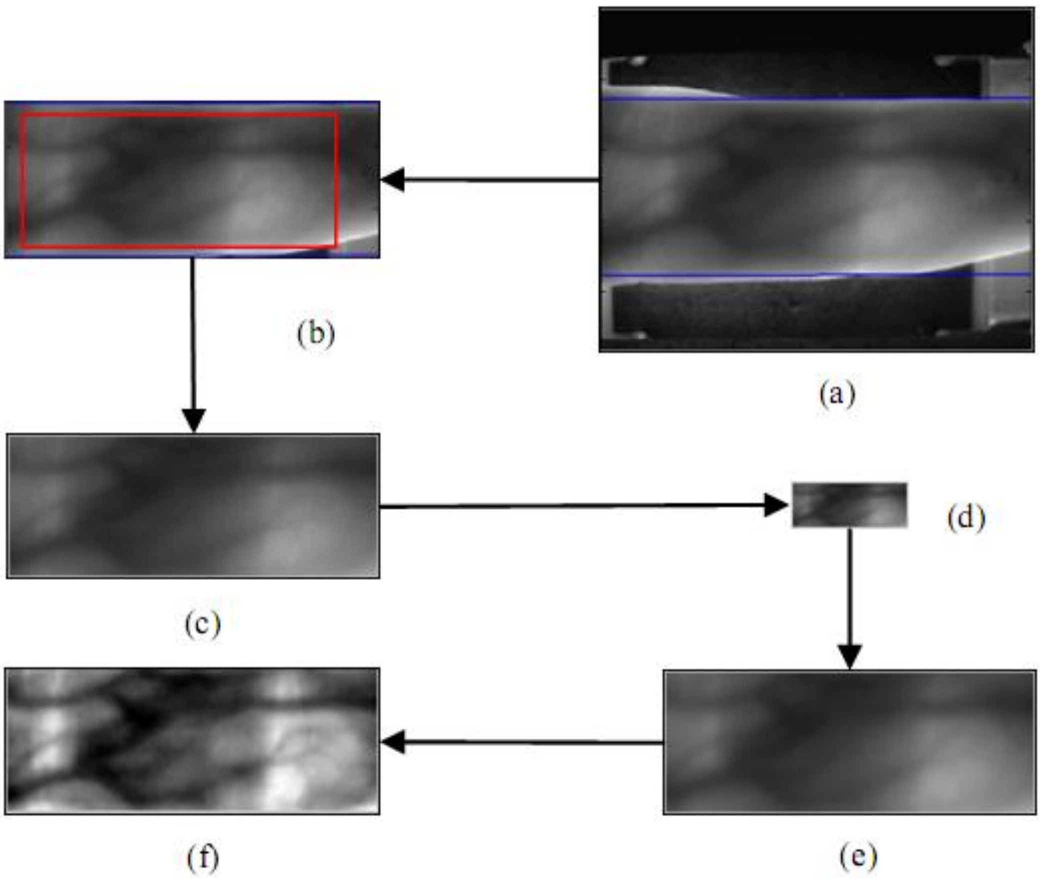


Figure 11. Block diagram for finger vein image preprocessing. (a) acquired image, (b) truncated image, (c) ROI, (d) image with 1/4 size of ROI, (e) denoised image, (f) enhanced image [**39**].

## *1.4 Voice*

Recent data on mobile phone users all over the world, the number of telephone landlines in operation, and recent VoIP (Voice over IP networks) deployments, confirm that voice is the most accessible biometric trait as no extra acquisition device or transmission system is needed. This fact gives voice an overwhelming advantage over other biometric traits, especially when remote users or systems are taken into account. However, the voice trait is not only related with personal characteristics, but also with many environmental and sociolinguistic variables, as voice generation is the result of an extremely complex process. Thus, the transmitted voice will embed a degraded version of speaker specificities and will be influenced by many contextual variables that are difficult to deal with. Fortunately, state-of-the-art technologies and applications are presently able to compensate for all those sources of variability allowing for efficient and reliable value-added applications that enable remote authentication or voice detection based just in telephone-transmitted voice signals [39], [16]. The first step in the construction of automatic speaker recognition systems is the reliable extraction of features and tokens that contain identifying information of interest. This short-time hamming/hanning windowed signals have all of the desired temporal/spectral information, albeit at a high bit rate (e.g. telephone speech digitized with sampling frequency 8 kHz in a 32 ms. window means 256 samples x 16 bits/sample = 4096 bits = 512 bytes per frame). Linear Predictive Coding (LPC) of speech has proved to be a valid way to compress the spectral envelope in an all-pole model (valid for all non-nasal sounds, and still a good approximation for nasal sounds) with just 10 to 16 coefficients, which means that the spectral information in a frame can be represented in about 50 bytes, which is 10% of the original bit rate. Instead of LPC coefficients, highly correlated among them (covariance matrix far from diagonal), pseudo orthogonal cepstral coefficients are usually used, either directly derived as in LPCC (LPC-derived Cepstral vectors) from LPC coefficients, or directly obtained from a perceptually-based mel-filter spectral analysis as in MFCC (Mel-Frequency based Cepstral Coefficients). Some other related forms are described in the literature, as PLP (Perceptually based Linear Prediction) [40], LSF (Line Spectral Frequencies) [41] and many others, not detailed here for simplicity. By far, one of the main factors of speech variability comes from the use of different transmission channels (e.g. testing telephone speech with microphone-recorded speaker models). Cepstral representation has also the advantage that invariant channels add a constant cepstral offset that can be easily subtracted (CMS.- Cepstral Mean Subtraction), and non-speech cepstral components can also be eliminated as done in RASTA filtering of cepstral instantaneous vectors [42]. In order to take coarticulation into account, delta (velocity) and delta-delta (acceleration) coefficients are obtained from the static window-based information, computing an estimate of how each frame coefficient varies across adjacent windows (typically between ±3, no more than ±5).

### *1.5 Online signature*

Automatic signature verification is an important modality area due its social and legal acceptance and widespread handwritten signature use as a personal authentication [43] [44] [45]. Another advantage of the handwritten signature as a biometric modality is that it is easily acquired either with an inking pen over a sheet of paper or by electronic means with a number of existing pointer-based devices (e.g., pen tablets, PDAs, Tablet PCs, touch screens, etc.). Several approaches have been considered for extraction of discriminative data of on-line signature data [43]. The actual techniques are divided into two classes: feature-based, in which a holistic set representation as a vector of global features is obtained using the signature trajectories [46] [47], and function-based, whereas, the recognition process use time sequences describing local properties of the online signature [48] [49] [50] [51], e.g., position trajectory, acceleration, velocity, pressure or force [52]. Recent works prove that feature-based methods are competitive with respect to function-based ones in particular situations [53], the latter approach has traditionally yielded better results. The set of features used can be a result of a feature selection process [54] during a development phase [46] [52] [53], or can be adapted during the enrollment phase to the specificities of the user at hand. The latter approach is believed to be better suited to the problem of signature verification [55] [56], mainly because of the large differences in information content and complexity between signers [57] [58]. However, the user-specific approach encounters challenges of training data scarcity.

In this thesis, feature extraction modules are based on Gabor filters for images of face, fingerprint and finger vein, MFCC for signals of voice and dynamic features for online signature.

## *2 Feature extraction*

At this stage, the biometric data is processed and a set of salient discriminatory features extracted to represent the underlying trait. For example, the position and orientation of minutia points (local ridge and valley anomalies) in a fingerprint image are extracted by the feature extraction module in a fingerprint-based biometric system. During enrollment, this feature set is stored in the database and is commonly referred to as a template [10].

### *2.1. Gabor filters for face, fingerprint and finger vein*

The Gabor filter bank is used to represent images of face, fingerprint or finger vein as vector codes. The defaults parameters rely to the most common parameters used for conjunction with localized face, fingerprint or finger vein images. The function may return a filter bank structure which contains the spatial and frequency representations of the constructed Gabor filter bank.

Images are filtered using a bank of Gabor filters constructed using the construct Gabor filters function of the Pretty Helpful Development (PhD) toolbox [59] [60]. The constructed filters are applied on the input image, to compute the magnitude responses. Then, these magnitude responses are down-sampled. And finally, the resulted down-sampled magnitude responses are concatenated into one unique feature vector. Note that these feature vectors are produced such as those produced in [59] [61] [62].

In fig. 12 a Gabor filter bank of 40 filters (8 orientations x 5 scales) has been constructed; a sample of the face image of the 26th subjects of SDUMLA-MHT face database is used for filtering operation.

(1) Original Image (2) Localized face (3) Magnitudes responses

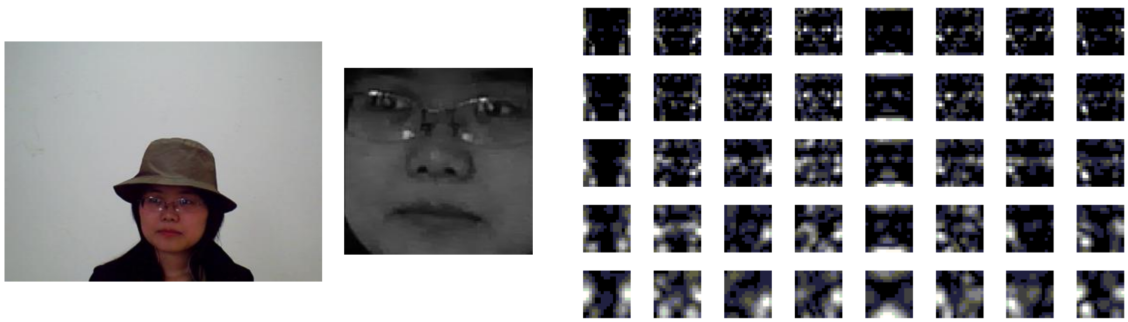

Figure 12. Magnitude responses of the filtering operation with the Gabor filter bank (no down-sampling) using the 26th subject face in SDUMLA-HMT database.

(1) Original Image (2) Localized Face (3) Magnitudes responses

Figure 13. Magnitude responses of the filtering operation with the Gabor filter bank (no down-sampling) using the 24th subject face in BIO database.

In the case BIO, VidTimit and SDUMLA-HMT face database, the region of face is localized using face detection library of [63]; before applying down-sampling operation.

(1) Original Image (2) Localized Face (3) Magnitudes responses

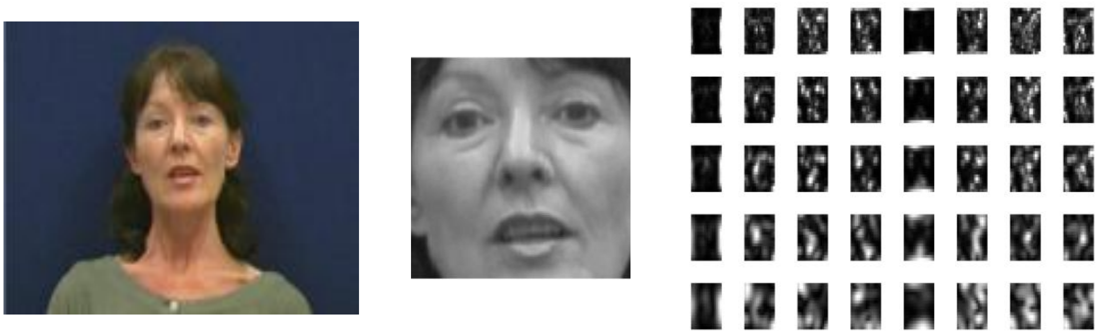

Figure 14. Magnitude responses of the filtering operation with the Gabor filter bank (no down-sampling) using the first subject face in VidTimit database.

In fig. 13 and fig. 14, samples of the face images of the $24^{th}$ and the first subjects of BIO and VidTimit databases, respectively; are localized then used for the same filtering operation.

In fig. 15 a sample of the fingerprint image of the $24^{th}$ subjects of BIO database is used for the same filtering operation.

(1) Original Image (2) Magnitudes responses

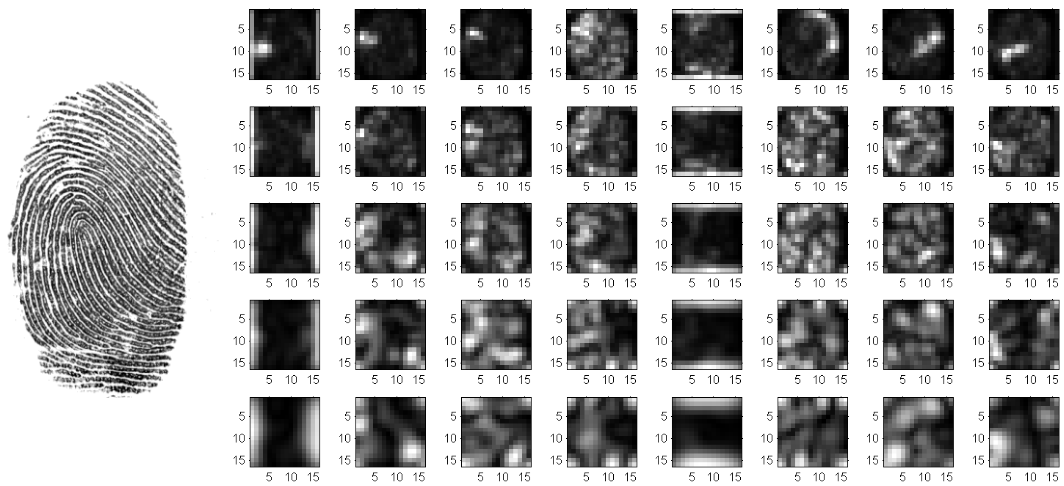


Figure 15. Magnitude responses of the filtering operation with the Gabor filter bank (no down-sampling) using the 24$^{th}$ subject fingerprint in BIO database.

In fig. 16 a sample of the finger vein image of the right index finger of the 91$^{st}$ subject of SDUMLA-HMT database is used for the same filtering operation.

(1) Original Image (2) Magnitudes responses

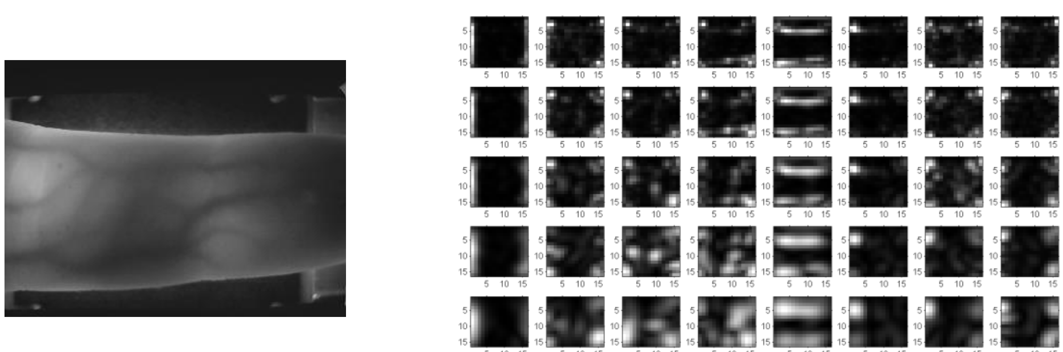


Figure 16. Magnitude responses of the filtering operation with the Gabor filter bank using the 91$^{st}$ subject finger vein in SDUMLA-HMT database.

### 2.2. *Mel frequency cepstral coefficient for voice*

At first, voice signal is converted and represented as parametric way for further processing and analysis. Any speech can be considered as a slowly timed varying signal (quasi-stationary). If it is examined in a sufficiently small period of time (from 5 to 100 ms), its characteristics are considered stationary. Nevertheless, in long periods of time (on the order or more than 200 ms) the characteristics of the signal change to reflect the different sounds of speech being spoken. Hence, analysis of short-time spectral is the best way to characterize the signal of speech. Many possibilities exist for parametrically representing the signal of speech for the speaker identification, like Gaussian mixture models (GMM), Linear Prediction Coding (LPC) [64], Mel-Frequency Cepstral Coefficients (MFCC), and many others.

MFCC's are used the known variation of the ear's of human critical bandwidths with frequency; to collect the phonetically interesting characteristics of speech signal, a linearly spaced filters at low frequencies and at high frequencies, logarithmically have been employed. This is articulated in the Mel-frequency scale, which is linear frequency and logarithmic spacing above 1000 Hz and spacing below 1000 Hz. Computing MFCCs process can be found in more details in [65] [66].

### 2.3. *Dynamic feature for online signature*

When dealing with signatures, two types of modalities are considered, the offline modality, in which scanned images of the signatures are accessible for the matching, and the online modality, in which the signatures are captured by digital tablets. The online modality is believed to provide more data about the signatures (speed, trajectory, pressure … etc.) and reached consequently better accuracy than the offline signature. This is why the online modality is used in the context of this thesis.

Signature verification is performed by comparing a questioned/unknown signature with a reference signature; this comparison is performed at the feature level after extracting characterizing features from both signatures.

Online signatures contain a set of instances; each instance corresponds to the coordinates of the point on the tablet along with the related pressure $X_t$, $Y_t$ and $P_t$ where t is time. Those features (or signals) are used to extract several other features can be calculated:

- Distances: The Euclidian distance can be calculated between each coordinates off successive X and Y of signature: $d_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}$;
- Angles: It is located between the X axis and the line shaped with the first point of signature and its current point $\alpha_t = atan\frac{y_t - y_0}{x_t - x_0}$;
- Speeds: The difference between successive distances $S_t = d_t - d_{t-1}$;
- Angular speeds: The difference between successive angles $AS_t = \alpha_t - \alpha_{t-1}$.

## *3 Feature selection and reduction*

The use of multiple number of orientations and scales in the filtering stage of images using Gabor filters, and multiple number of MFCC's scales for voices and the full number of dynamic features of signatures obtained in the extraction phase, produces output vectors of big dimensionality, this is why a reduction and pertinent data selection are necessary, especially, if considering the fusion at feature level of these vectors.

### *3.1 Principal component analysis*

In order to approximate the original information with lower size feature vectors, it is possible to use principal component analysis (PCA). The principle approach is to compute the feature vectors of the covariance matrix of the original information, and approximate it by a linear combination of the leading feature vectors [67]. By using PCA process, the test vector can be recognized by projecting the original vector onto the feature vector space to obtain the related set of weights, and then using the to be compared with the weights of the vectors in the dataset of training [68] [69]. Low-dimensional feature representation problem may be declared as follows: Let X= ($x_1$ ,$x_2$,..., $x_i$,..., $x_n$) represents the n×N data matrix where $x_i$ is a vector code of n elements, concatenated from for example a fingerprint and finger vein feature vectors. Here n is the entire number of elements in the fingerprint and finger vein feature vectors and N represents the number of reference templates of individual in the training set. The PCA may be well thought-out as a linear transformation (2) from the original vector code to a projection feature vector code.

$$Y = W^T X \tag{1}$$

Where Y is the m × N feature vector matrix, m represents the dimension of the feature vector and transformation matrix W is an n×m transformation matrix whose columns are the feature vectors corresponding to the m largest training feature values computed according to the equation (2):

$$\lambda e_i = S e_i \tag{2}$$

Where $e_i$, $\lambda$ are feature vectors and feature values matrix respectively. Here the total scatter matrix S and the mean image of all samples are defined as,

$$S = \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T, \mu = \frac{1}{N} + \sum_{i=1}^{N} x_i \tag{3}$$

After applying the linear transformation $W^T$, the scatter of the transformed feature vectors {$y_1$, $y_2$... $y_N$} is $W^TSW$. In PCA, the projection $W_{opt}$ can be chosen to maximize the determinant of the total scatter matrix of the projected samples, i.e.

$$W_{opt} = \arg\max_w |W^T S W| = [w_1 w_2 \dots w_m] \tag{4}$$

Where {$w_i$ | I = 1, 2,..., m} is the set of n-dimensional feature vectors of S corresponding to the m largest feature values. In other words, the input vector in an n-dimensional space is reduced to a feature vector in an m-dimensional subspace.

### *3.2 Linear discriminant analysis*

LDA is a dimensionality reduction technique which is used for classification problems. LDA is also known as Fisher's Discriminant Analysis and it searches for those vectors in the underlying space that best discriminate among classes (rather than those that best describe data as in PCA) [70] [71] [72].

LDA creates a linear combination of independent features which yields the largest mean differences between the desired classes. The basic idea of LDA is to find a linear transformation such that feature clusters are most separable after the transformation which can be achieved through scatter matrix analysis [72]. The goal of LDA is to maximize the between-class scatter matrix measure while minimizing the within-class scatter matrix measure [71].

The basic steps in LDA are as follows:

• Calculate within-class scatter matrix, $S_w$ :

$$S_w = \sum_{j=1}^{C} \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \tag{5}$$

where $x_i^j$ is the $i$th sample of class $j$, $\mu_j$ is the mean of class $j$, $C$ is the number of classes, $N_j$ is the number of samples in class $j$.

• Calculate between-class scatter matrix, $S_b$

$$S_b = \sum_{j=1}^{C} (\mu_j - \mu)(\mu_j - \mu)^T \tag{6}$$

where $\mu$ represents the mean of all classes.

• Calculate the feature vectors of the projection matrix:

$$W = \mathrm{eig}\,(S_w^{-1} S_b) \qquad (7)$$

• Compare the test image's projection matrix with the projection matrix of each training image by using a similarity measure. The result is the training image which is the closest to the test image.

# 4 Summary

Feature extraction methods to extract local features in face, fingerprint, finger vein, voice and online signature are presented in this chapter. The proposed method is based on multiresolution analysis using the Gabor transform to extract representation information, MFCCs and dynamic features of signature. The extracted features vector has low dimensionality due PCA/LDA based feature reduction and it contains low frequency information from the source, and thus is suitable to be used in the feature fusion framework proposed in chapter 3.

# Chapter 2

## Multimodal fusion strategies

This chapter investigates multimodal fusion strategies in multibiometric systems which are now a reality due accuracy problem; Multimodality increases the accuracy and robustness because it incorporates several independent sources of biometric information and it can achieve accuracy levels targeted by multiple applications. There is also the problem of lack of universality; and by using multimodality; the Failure to enroll (FTE) can be reduced by the ability to effectively capture a larger proportion of the population. Another problem is the problem of vulnerability to frauds; however, it is more difficult to imitate multiple biometric sources simultaneously and in the authentication process; only a random subset of features can be requested.

Fig. 17 shows a hypothetical mobile banking application where the user has the flexibility to choose all or a subset of available biometric traits (e.g., face, voice and fingerprint) for authentication depending on his convenience. Research is under way to perform iris recognition based on images captured using the camera on the mobile phone (see fig. 17).


Figure 17. Mobile banking application [73].

## *1 Multibiometrics*

Systems that combine evidence of many biometric data sources to decide and find the identity of an individual, are known as multibiometric systems [74].

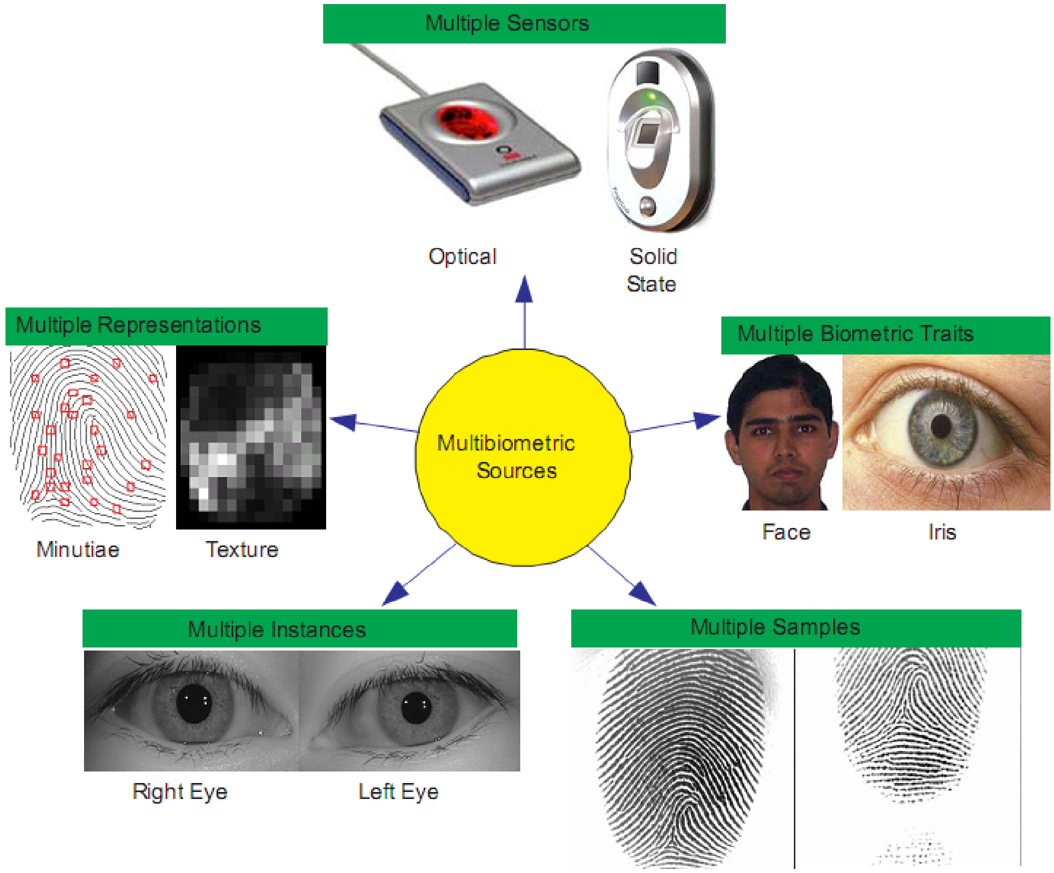


Figure 18. Various sources of information that can be fused in a multibiometric system [75].

In four of the five scenarios (multiple sensors, representations, instances and samples), multiple sources of information are derived from the same biometric trait. In the fifth scenario, information is derived from different biometric traits and such systems are known as multimodal biometric systems see (fig. 18).

Multibiometric systems can alleviate many of the limitations of uni-biometric systems because the different biometric modalities or sources frequently recompense for the natural limitations of the other modalities or sources [76].

Sources of information in a multibiometric system (see fig. 18) may include (i) multiple sensors to capture the same biometric trait (e.g., face captured using optical and range sensors), (ii) multiple representations or multiple algorithms for the same biometric trait (e.g., texture and minutiae-based fingerprint matchers), (iii) multiple instances of the same biometric trait (e.g., left and right iris), (iv) multiple samples of the same biometric trait (e.g., two impressions of a person's right index finger), and (v) multiple biometric traits (e.g., face and iris) [75].

In the first four scenarios, multiple sources of information are derived from the same biometric trait. In the fifth scenario, information is derived from different biometric traits and these systems are known as multimodal biometric systems. In fact, biometric fusion can also be released on any arbitrary consolidation of the above five sources and such systems can be referred to as hybrid multibiometric systems [77]. Brunelli et al proposed an example of a hybrid multibiometric system is the system in [78] where the results of two speaker recognition algorithms are combined with three face recognition algorithms at the match score and rank levels using a HyperBF network. Hence, this system is multi-algorithmic as well as multimodal in its design.

## *2 Operating modes of a multimodal system*

### *2.1 Serial mode*

One biometric information source is use at once; cascaded acquisition measures sources at different times useful as indexing technique in an identification system (with large database) reduces the number of identities that are explored with the next source, this allows to converge to a single identity with the latest biometric source.

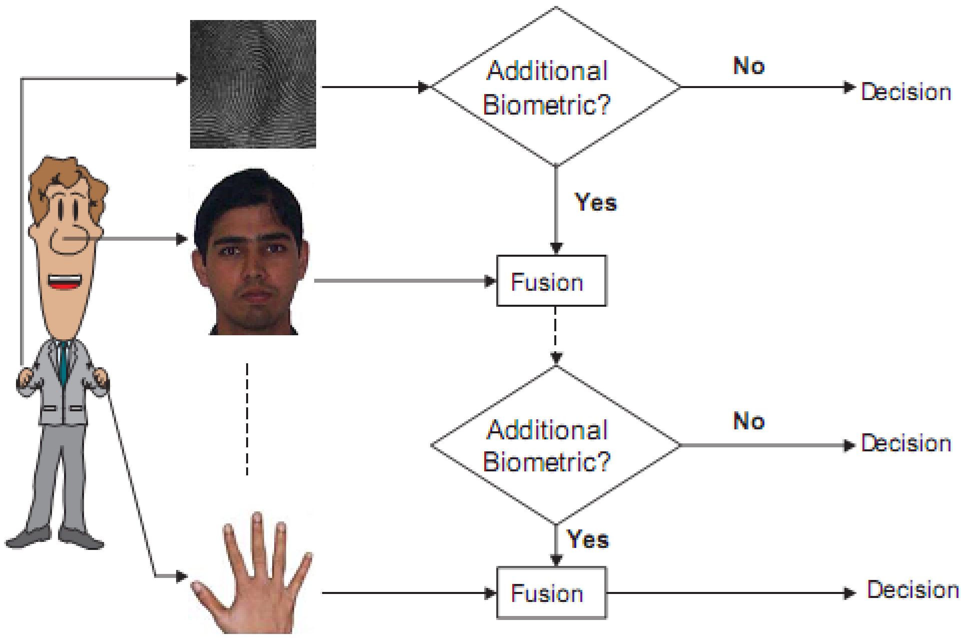


Figure 19. Acquisition and processing architecture of a multimodal system in serial (cascade or sequential) mode [73].

### *2.2 Parallel mode*

Each source is used simultaneously by an independent system. It acquires all the biometric sources simultaneously. The fusion of the responses of all sources leads to a final decision.

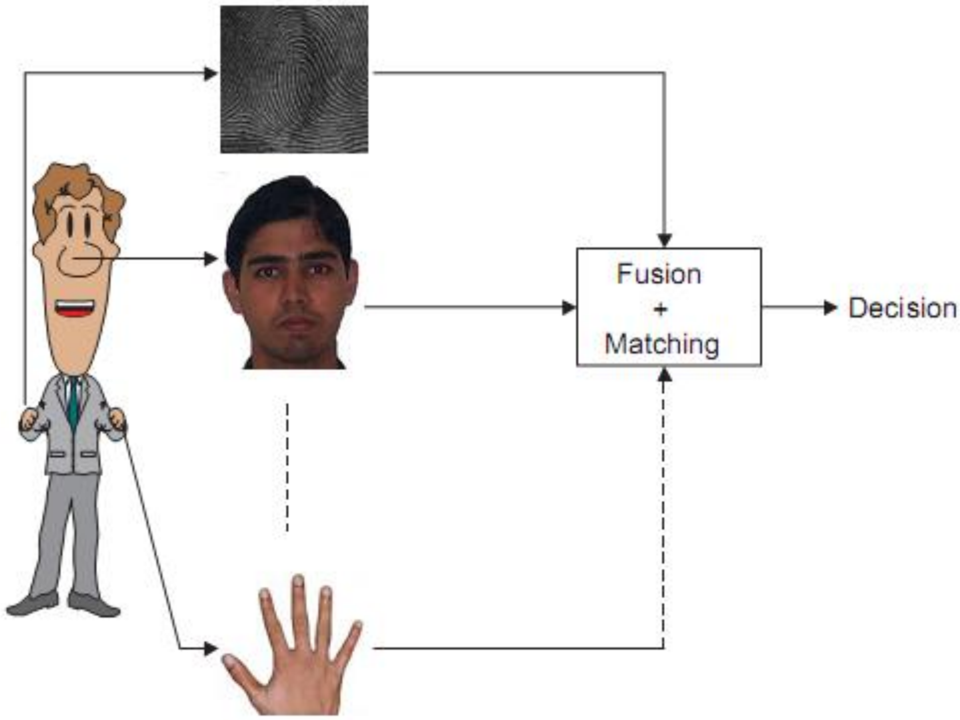


Figure 20. Acquisition and processing architecture of a multimodal system in parallel mode [73].

### *2.3 Hierarchical mode*

Each source is used by an independent system. It is an extension of the parallel mode. A large number of systems are combined in a tree structure to benefit from advantages of both cascade and parallel architectures. This hierarchical architecture can be made dynamic so that it is robust and can handle problems like missing and noisy biometric samples that often arise in biometric systems [79]. However, the design of a hierarchical multibiometric system has not yet received adequate attention from researchers. This is what pushes us to investigate this strategy of fusion; the proposed framework is described in fig. 21.

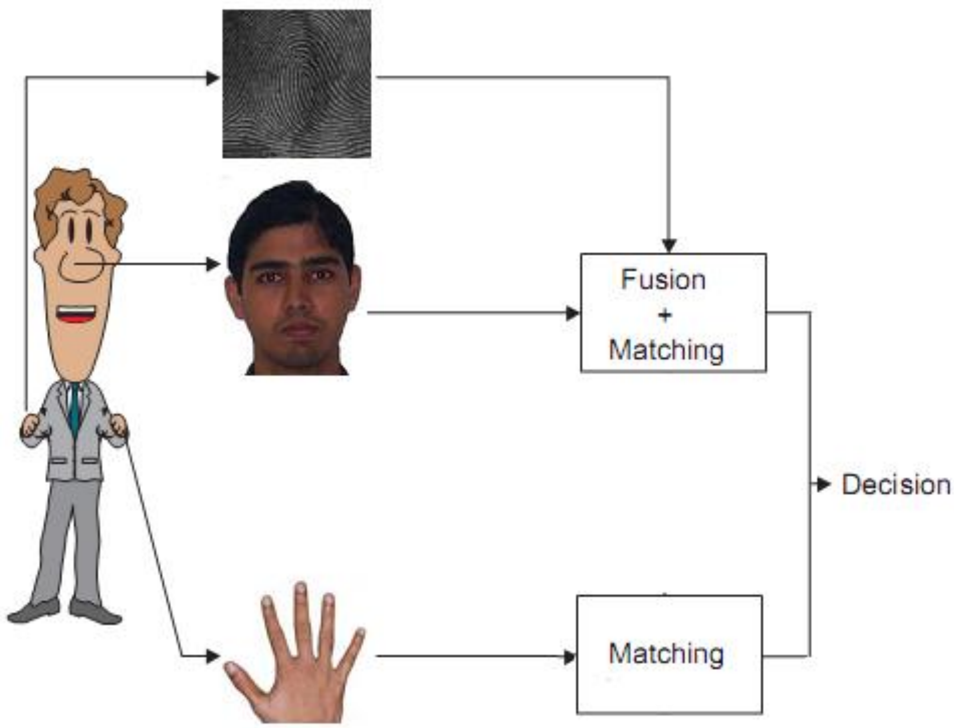


Figure 21. Acquisition and processing architecture of a multimodal system in the proposed hierarchical mode.

### *3 Fusion prior to matching*

Prior to matching, integration of information from multiple biometric sources can take place either at the level of sensor or feature [75]. The fig. 16 presents the framework of multimodal biometric system based on feature fusion as prior to matching.

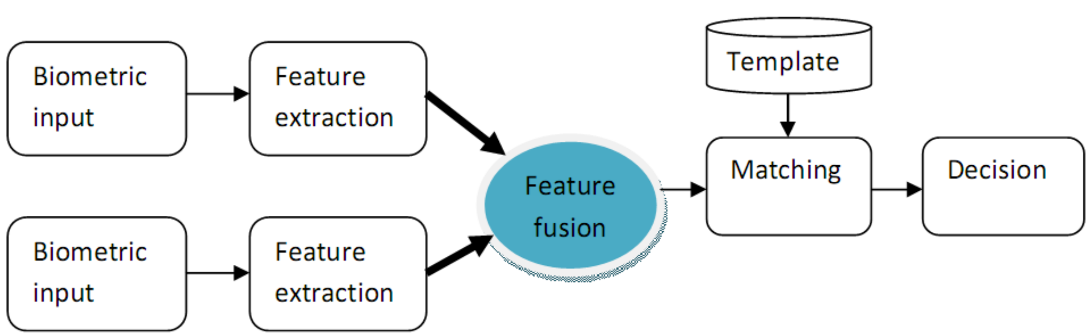


Figure 22. Fusion in multimodal biometrics systems **at f**eature level [80].

Sensor level fusion involves combining raw data from sensors, and it can be achieved if the multiple sources represent samples of the same biometric trait obtained using single or different compatible sensors. In this method, the multiple modalities must be compatible with feature level in the raw data and must be known in advance, for example in the construction of 3D face images by integrating raw images captured from several cameras [81]. Jain [81] integrates information at the sensor level by forming a mosaic of multiple fingerprint impressions in order to construct a more elaborate fingerprint image. Feature level fusion (see Fig. 22) refers to combining the features obtained from multiple modalities into a single feature vector. It is reasonable to combine two feature vectors into one unique new vector if these features extracted from multiple biometrics are independent of each other and involve the same type of measurement scale.

The new combined feature code will have higher size, the fact that increases the discriminating power in feature space. Feature selection techniques or feature reduction schemes may then be used to extract a small number of significant and pertinent features from a larger set of features [82] [83] [84].

### *4 Fusion after matching*

Schemes for integration of information after the classification/matcher stage can be divided into four categories: dynamic classifier selection, fusion at the decision level, and fusion at the score level and fusion at the rank level.

Fusion at score level (see Fig. 23) refers to the combination of similarity scores provided by a matching module for each modality when the input features are compared against templates in the database [85]. This method is also known as fusion at the measurement or confidence level. The matched score output generated by biometrics matchers contain rich information about the input pattern after the feature extraction module. Fusion at matching score level can be categorized as involving two different approaches depending on how the matching score given by matching module is treated [86]. In the first approach, the fusion can be considered as a problem of classification where a feature vector is generated using the matching score output by the individual matchers. Then, the constructed feature vector is classified into two of the classes whether to accept or reject the claim user. In the second approach, fusion is viewed as a combination approach where individual matching scores are fused to generate one unique scalar score using normalization techniques and fusion rules. The new single scalar score is then used to make a final decision. The combination approach to the fusion of matching scores has been extensively studied, and Ross [87] concluded that it performs better than the classification approach. Fusing matching scores using the combination approach has some issues arise during computing a single fusion score given by different modalities. A normalization technique is required to transform comparison scores into a common domain prior to fusing them, since the matching scores generated from different modalities are heterogeneous [86]. Several different kinds of normalization technique have been proposed, such as min-max, median and sigmoid normalization.

Integration of the information at the decision level is performed when each of the individual biometric matchers decides the best match based on the input features presented to the matching module. Various methods such as majority vote [88], behavior knowledge space [89], AND and OR rule [90] can be used to make the final decision. This kind of fusion uses binary information to derive a final decision, and thus fusion at decision level is not effective because only a restricted quantity of data is available at this level. Hence, the integration of the information at feature and matching at score level is commonly chosen due to the richness of data available at the fusion stage.

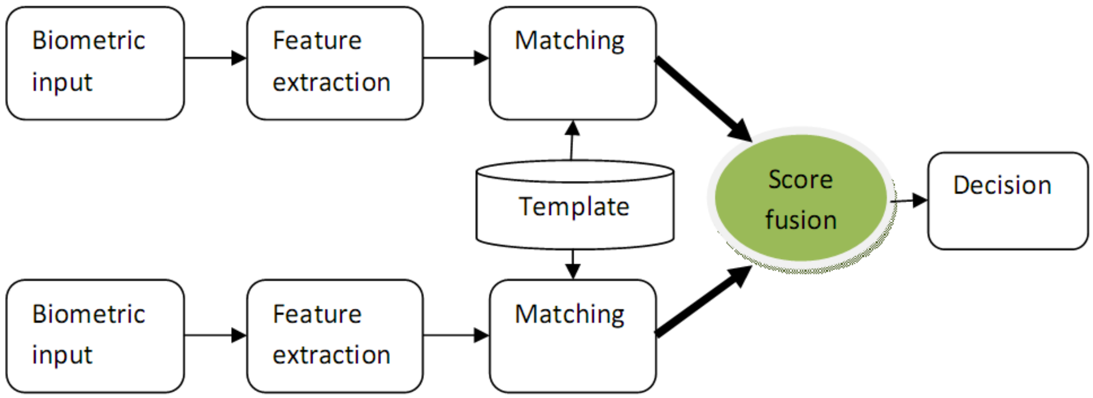


Figure 23. Fusion in multimodal biometrics systems at score level [80].

### *5 Summary*

A new method to fuse the information based on hierarchical strategy is presented in this chapter.

Multimodal biometrics systems that fuse information at an early stage are supposed to be more efficient than those that integrate it at a later stage. This is due to the rich information which exists at feature level compared to that at matching and decision level. As the features contain richer data about the input modality, learning the distribution of the fused feature using statistical models to capture the relevant statistical properties is likely to give better classification performance. In most existing feature fusion methods, a concatenation feature vector is classified using a distance classifier [80].

On the other hand, in most of the multibiometric systems, it is comparatively effortless to access and fuse the scores generated by different biometric matchers. Therefore, information fusion at the score level gives the best tradeoff in terms of data content and simplicity in fusion. Consequently, score level fusion is the most commonly used approach in multibiometric systems [75] [91].

Thus, the aim in this thesis is to propose a combination of both features and similarity scores fusions to enhance the overall performance of the identification/verification based biometric system. According to the best of the present author's knowledge, this is the first study of multimodal biometrics to propose the hierarchical fusion of the similarity scores of fused feature vectors. As well as the first to specifically use local features in multimodal biometrics based on face, fingerprint, finger vein, voice and/or online signature. Theoretically, double fusion should be better to give best performance due to the utilizing several fusion level at once.

The proposed method is based on feature and score fusion of three modalities at the same time where local feature vectors of the first two modalities are combined to form a new fused feature vector. After matching, the score of matching of the fused feature vector is fused with the score of matching of the third modality. The proposed hierarchical fusion has several advantages compared to classic strategies. These advantages are given and investigated in the chapter 3.

# Chapter 3

# Databases, matching, experiment scenarios and results

In this chapter, a process of evaluation is presented. Evaluation of the proposed hierarchical fusion strategy is done on three scenarios, using three different data sets of biometric databases using five biometric modalities; three modalities by scenario.

The proposed strategy is based combination of fusion at two levels, feature and score, these two fusion methods are evaluated and results are given as well as the proposed method. The performances are evaluated using recognition rates and equal error rates.

Curves of CMC, ROC and DET are plotted for all scenarios and databases.

## 1 Databases

### 1.1 QU-PRIP

This dataset is offered from Qatar University; it includes 138 individuals with three reference templates of online signatures, and some of the individuals have as six reference templates [92].

Figure 24. Six online signature samples for the 51st, 63rd and 85th subjects from QU-PRIP database.

### *1.2 BIO*

BIO [93] database is collected at the administrative department (Daïra) of Adrar by Master students using the same biometric equipment that is used for collecting biometric data for national biometric passport and identity cards. It contains two different images of face and ten images of fingerprints of 25 distinct subjects. For some subjects, their face images were taken with varying lighting, facial expressions and facial details. All the images were taken against a light homogeneous background with the subjects in a frontal position (with tolerance for some side movement). The fig. 25 shows some samples from BIO database.

The collection is carried out with images of fingerprints and faces of volunteers (A group of Bachelor students, Master students and professors from university of Adrar) using biometric devices (fingerprint sensor “3M Cogent”, digital camera “canon EOS”). Face images were acquired by a canon EOS digital camera (see fig. 26) This camera takes pictures in jpg format with maximum resolution (960 × 1280) pixels, where, fingerprint images were taken with 3M Cogent (see fig. 27) which is an enrollment photographic device of fingerprint images. This device takes pictures in BMP format with a maximum resolution of 320 × 480 pixels.

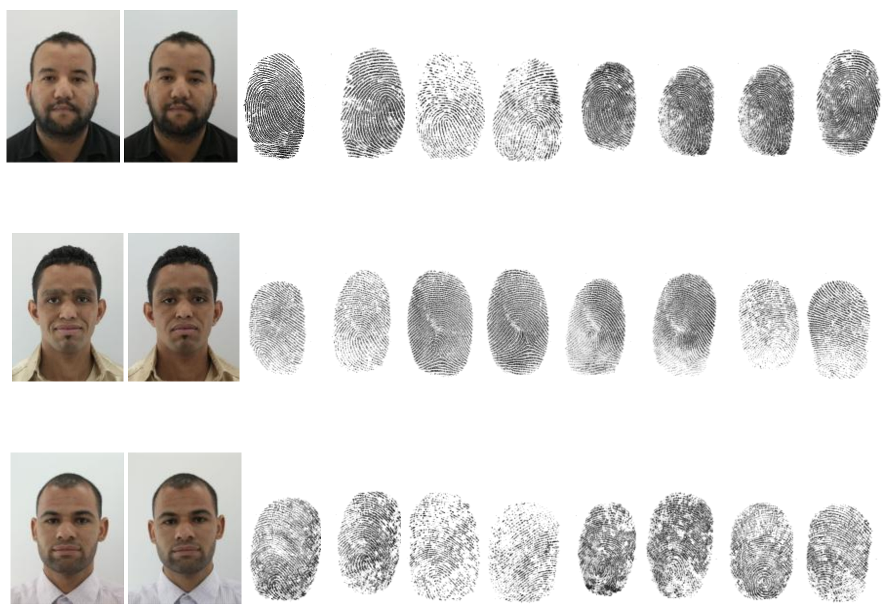

Figure 25. Samples from BIO database.


Figure 26. Digital camera “canon EOS”.


Figure 27. Fingerprint sensor “3M Cogent”.

### *1.3 VidTimit*

#### *1.3.1 Overview*

The VidTimit [94] database contains video and corresponding audio recordings of 43 people who have been asked to recite short sentences. It may be used for research areas as multi-view face recognition, automatic lip reading, and multi-modal speech recognition.

The database was recorded in three sessions, with a mean delay of seven days between the first session and the second, and six days between the second and the third session. The spoken expressions were picked from the test set of the TIMIT corpus[1]. There are ten expressions for each person. The first session contains the first six expressions. The next two expressions are assigned to the second session with the remaining two to the third session.

The first two expressions for all individuals are the same, but the lasting eight in general are different for each individual.

Figure 28. Samples from VidTimit face database.

[1] http://www.ldc.upenn.edu/doc/TIMIT.html

Additionally to the expressions, each individual performed a *head rotation cycle* in each session. The cycle consists of the individual moving the head in a specific way.

The recording was done in an environment of office using a broadcast quality digital video camera. The video of each person is registered as a numbered sequence of images with a resolution of 512 x 384 pixels. 90% quality configuration was used during the creation of the images. The corresponding voice record is stored as a mono, 16 bit, 32 kHz WAV file.

### 1.4 SDUMLA-HMT

Biometrics fusion recognition is a newly arisen and active research topic in recent years. In 2010, the Group of Machine Learning and Applications, Shandong University (SDUMLA) [95] released the Homologous Multi-modal Traits Database called SDUMLA-HMT Database. The SDUMLA-HMT database contains face images captured from seven angles, images of finger vein obtained from 6 fingers, gait videos captured from six angles, images of iris collected using an iris sensor, and fingerprint images captured using five different devices. The database contains real multimodal data of 106 real individuals.


Figure 29. Samples from face database of SDUMLA-HMT.

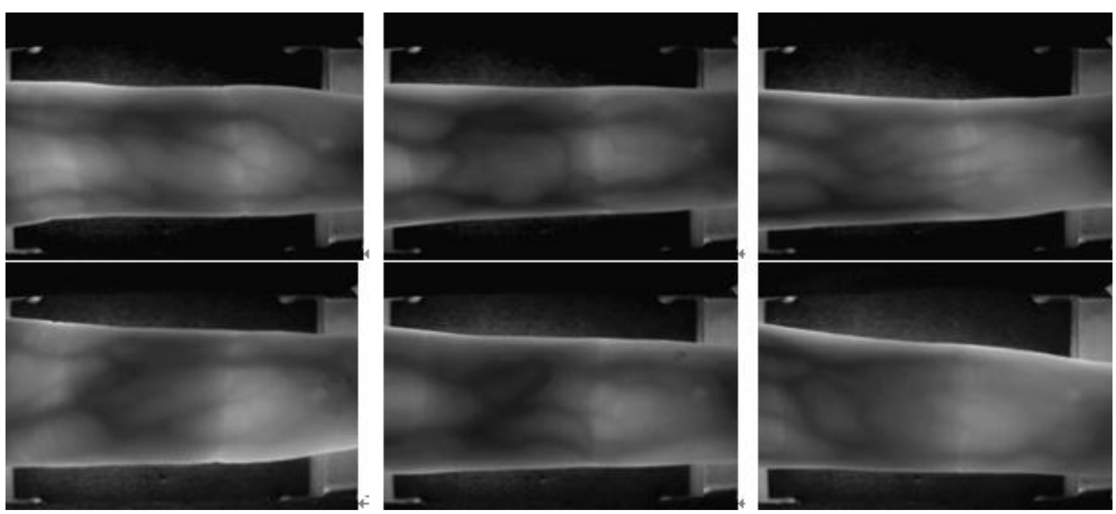

Figure 30. Samples from finger vein database of SDUMLA-HMT.

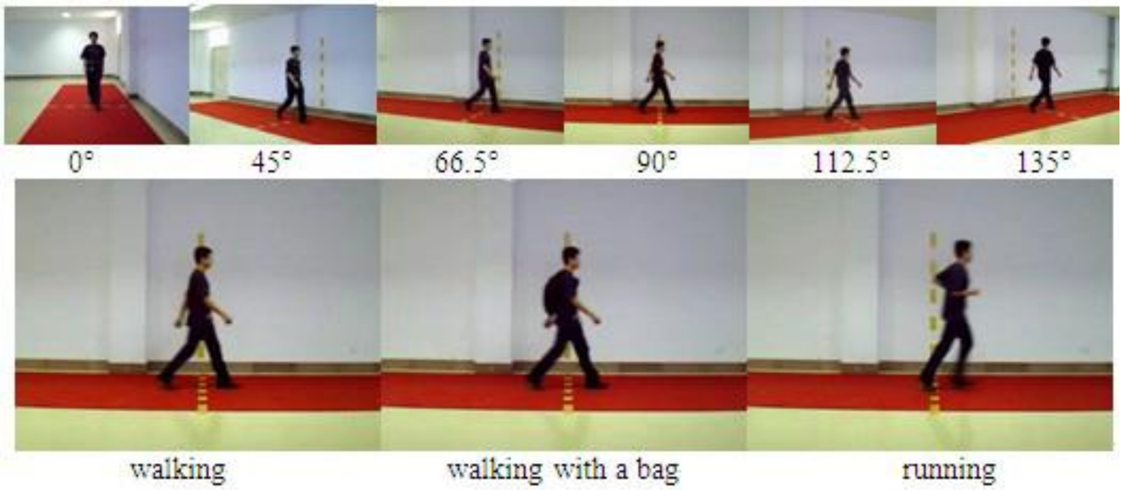


Figure 31. Samples from gait database of SDUMLA-HMT.

Figure 32. Samples from iris database of SDUMLA-HMT.

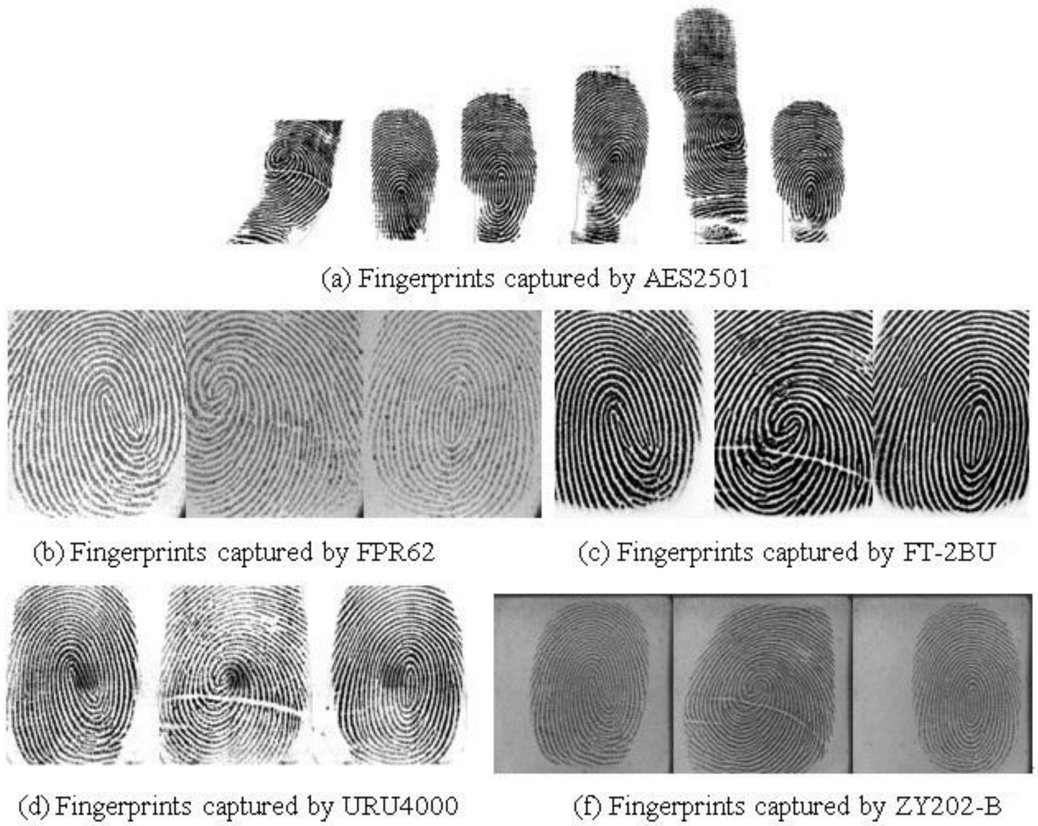


Figure 33. Samples from multisensory fingerprint database of SDUMLA-HMT.

## *2 Computing matching and similarity scores*

This section describes the method for computing the various parameters used to compute the threshold and our image denoising algorithm. The wavelet transform approach is used for the recovery of the corrupted image with optimal filter.

To compute the similarity measure, a bank of face, fingerprint, finger vein, voice and/or online signature vector codes, is adopted. As in the bank of references the vector codes of face fingerprint, finger vein voice and/or online signature signals are available. Five feature spaces are generated. The distance measure between the vector codes of test and the training codes is chosen as the Cosine Mahalanobis distance [96] between the projection of the test vector code and the projections of the gallery vector codes. The choice of Cosine Mahalanobis distance is motivated by the encouraging results obtained in [97] [98] [99] [100].

Let $\Gamma_1$, $\Gamma_2$, $\Gamma_3$... $\Gamma_N$ be the vector selected from the gallery. Let $\Theta = \frac{1}{N} + \sum_{i=1}^{N} \Gamma i$ be the average vector code. Let $\Phi_i = \Gamma_i - \Theta$ be the mean subtracted vector codes. Let the data matrix A is defined as $A = [\Phi_1 \Phi_2 \ldots\ldots \Phi_N]$. The feature vectors of $A^T A$ can be computed as $A^T A v_i = \mu_i v_i$. Pre-multiplying both sides by A, $AA^T A v_i = \mu_i A v_i$. Thus $Av_i$ are the feature vectors of $AA^T$. If $w_i$ is the projection of the mean subtracted vector code on the $i^{th}$ feature vector, then the projection coefficients of the vector code are $u = [w_1, w_2, \ldots\ldots w_N]$. The Cosine-Mahalanobis distance is used to measure the similarity between projection coefficients.

The feature vectors span the vector space. The feature values correspond to the variance along each feature vectors. It is important to know the transformation between the vector space and the Mahalanobis space before calculating the Cosine Mahalanobis distance. The Mahalanobis space has unit variance along each dimension. Let $u$ and $v$ be two vectors in the feature space. Let $\mu_i = \sigma_i^2$ be the variance along the $i^{th}$ dimension. Let $m$ and $n$ be the corresponding vectors in the Mahalanobis space. The relationship between the vectors is defined as:

$$m_i = \frac{u_i}{\sigma_i}, \; n_i = \frac{v_i}{\sigma_i} \tag{8}$$

Mahalanobis cosine is the cosine of the angle between the projections of the vectors on the Mahalanobis space. So, the Cosine Mahalanobis distance between u and v is computed in terms of m and n.

$$D_{MahCosine(u,v)} = \cos\theta_{(mn)} = \frac{mn}{|m||n|} \tag{9}$$

# 3 Fusion at feature level

Researchers believe that biometric systems that fuse information at the first part of the treatment are more effective than those performing the fusion later. Since the features contain rich information about biometric input than the score of similarity or the decision of a classifier [101].

### 3.1 Principle

In the first level, the raw data from the sensor(s) are combined in sensor level fusion [102]. Sensor level fusion can be performed only if the sources are either samples of the same biometric trait obtained from multiple compatible sensors or multiple instances of the same biometric trait obtained using a single sensor. For example, multiple 2D face images obtained from different viewpoints can be stitched together to form a 3D model of the face [103]. Whereas, Feature level fusion refers to combining different feature sets that are extracted from multiple biometric sources. When the feature sets are homogeneous (e.g., multiple fingerprint impressions of a user's finger), a single resultant feature set can be calculated as a weighted average of the individual feature sets (e.g., mosaicing of fingerprint minutiae [104]). When the feature sets are non-homogeneous (e.g., feature sets of different biometric modalities like face and hand geometry), they can be concatenated to form a single feature set. Feature selection schemes can then be applied to reduce the dimensionality of the resultant feature set [82]. By consequence, a stage of normalization is needed to make concatenation between non homogenous feature vectors.

### 3.2 Feature normalization and representation

After the extraction of features, the feature vectors may exhibit important variations both in distribution and range. In experiments min-max normalization is employed to combine modalities feature vectors. The feature normalization goal is to change the location (mean) and scale (variance) of the values of features in order to guarantee that the contribution of each element to the final score is comparable [86]. The simple min-max normalization is tested in this work.

Let $x$ and $x'$ represent a feature value before and after normalization, respectively. The min-max technique computes $x'$ as,

$$x' = \frac{x - \min(F_x)}{\max(F_x) - \min(F_x)}, \tag{10}$$

Where $F_x$ is the function which produces $x$. The min-max is efficient when the minimum and the maximum of the element feature values are identified beforehand. If such information is not on hand, an estimate of these parameters should be obtained from the accessible training data of the sample. The estimate can be affected by the existence of outliers in the training samples and this makes min-max normalization sensitive to outliers [105].

The fusion of features may produce large data that can penalize the whole system, so one more extended stage, is feature reduction using PCA or LDA as it is shown in the framework in fig. 34.

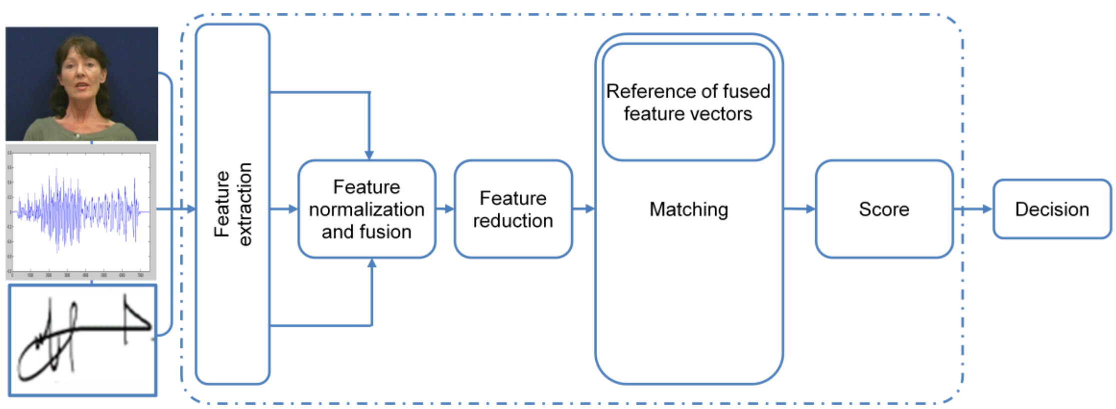


Figure 34. The proposed framework for feature fusion.

## 4 Fusion at score level

Match score is a measure of the similarity between the input and template biometric feature vectors. When match scores output by different biometric matchers are consolidated in order to arrive at a final recognition decision, fusion is said to be done at the match score level. This is also known as fusion at the measurement level or confidence level. It must be noted that the match scores generated by the individual matchers may not be homogeneous. For example, one matcher may output a distance or dissimilarity measure (a smaller distance indicates a better match) while another may output a similarity measure (a larger similarity value indicates a better match). Furthermore, the outputs of the individual matchers need not be on the same numerical scale (range). Finally, the match scores may follow different probability distributions and may be correlated. These factors make match score level fusion a challenging problem.

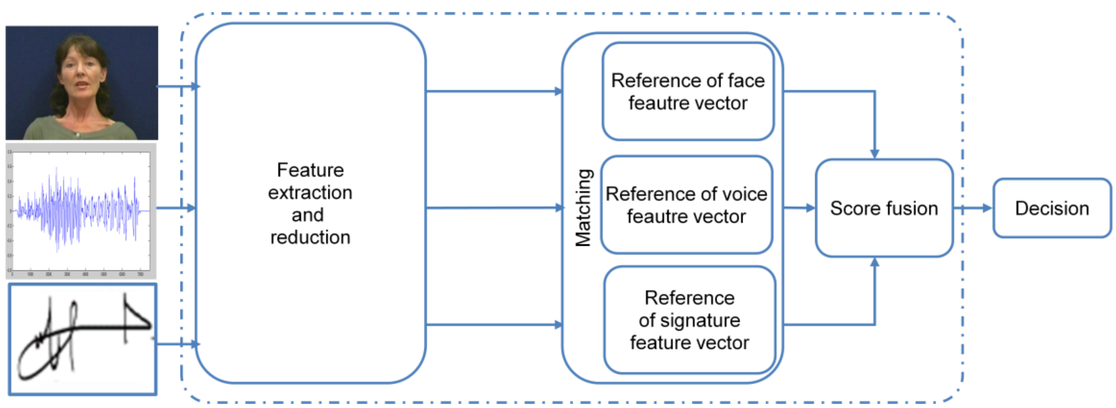


Figure 35. The proposed framework for score fusion.

### 4.1 Rules

The scores used are obtained from three identification/verification systems (see fig. 35). The first system is based on face verification using the combination of Gabor for feature extraction and cosine Mahalanobis distance for classification.

The second is based on voice recognition using MFCC's features.

The third system is based on online signature verification using Nalwa's method [48]. Various strategies are used to fuse face, voice and online signature scores such as the simple sum, minimum and maximum of scores.

## 5 Hierarchical fusion

Here, the proposed hierarchical strategy of fusion based on multimodal biometric system is presented. This strategy consists to combine several biometric modalities using a multi-level biometric fusion hierarchy. The multi-level biometric fusion includes a prior-to-matching fusion with optimal feature selection and an after-matching fusion based on the similarity of minimum of distances (scores). The proposed solution improves biometric recognition performances based on feature selection and reduction using principal component analysis (PCA) or Linear discriminant analysis (LDA).

A hierarchical face, voice and online signature fusion framework is proposed, as illustrated in fig. 36. In the training phase, extraction of features is done from voice using Mel-frequency cepstral coefficients, from face using Gabor filters, and from online signatures a set of samples; each sample relies to the point coordinates on the digitizing tablet along with the corresponding pressure $X_t$, $Y_t$ and $P_t$ where t corresponds to time. In the testing phase, the questioned face, voice and online signature are utilized to find a subset of reference templates in the databases of the same modalities. Lastly, the score level fusion of face, voice and online signature is employed on the candidate subset for personal recognition.

### 5.1 Framework

The proposed multimodal biometric system architecture is shown in fig. 18. This architecture is based on a hierarchical approach for biometric data fusion and classification. The system includes two main identification components:

1. The face and voice identification sub-system. This component is relying on a feature-level biometric fusion scheme;

2. The online signature identification subsystem;

3. The post-classification fusion scheme, which combines the two previous components scores;

4. The decision module which provides the final decision based on the global score.

The proposed multimodal approach is featured by multi-level integration, as shown in fig. 36.

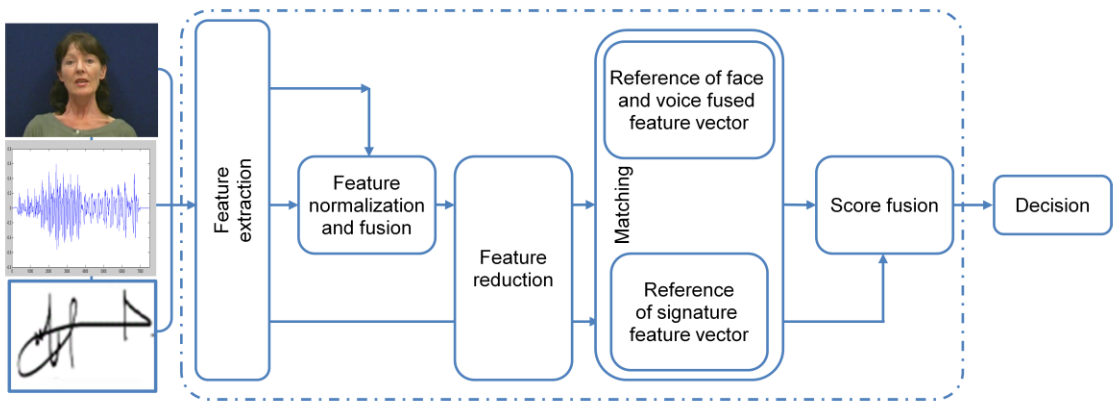


Figure 36. The proposed framework for hierarchical fusion.

## *6 Scenarios*

To evaluate our proposed fusion method, three scenarios are proposed. Starting by BIO bimodal database that contains face and fingerprint images of 25 persons, VidTimit bimodal database with face images and voice records of 43 persons; these two databases contain just two modalities; in order to release the hierarchical fusion, a third modality is indispensable; QU-PRIP database is used as a third modality, it contains online signatures of 138 persons, but this number is reduced to 25 and 43 to be adaptable to the number of persons in BIO and VidTimit respectively.

To generalize the obtained results, a real multimodal database is used, this database is SDUMLA-HMT, and this last database comes with face, 106 persons.

### *6.1 Scenario 1*

The BIO bimodal database is used. It is comprised of face and fingerprint images of 25 volunteers. The second database for the third modality is QU-PRIP database of online signature [92].

Considering the number of persons in the BIO database, online signatures of just 25 persons have been selected from the QU-PRIP. From both subsets, and by taking advantage of the independence of face and voice from online signature trait, 43 virtual subjects have been created.

The following training and testing process for monomodal systems has been established:

For training purposes, each face has been modeled using four samples, and each fingerprint and online signature has been modeled using the same number of samples and one sample of each trait for validation purposes.

For validation and testing, for each client one more sample of each modality (face, fingerprint and online signature) were also selected for validation and one more sample for testing; the same 25 clients are used as impostors, except that each client claims an identity different from his own. Each client has been considered and, from each impostor, one sample has been selected.

Consequently, the sub-corpus for the experiments consists of 25 clients, and 24x25x2=1200 multimodal impostor attempts.

### *6.2 Scenario 2*

The VidTimit bimodal database is used. It is comprised of video and corresponding audio recordings of 43 volunteers (19 female and 24 male), reciting short sentences [106]. In order to apply hierarchical strategy of fusion, it is necessary to use at least three biometric traits, VidTimit is a bimodal database (face and voice), as scenario 1, a second database for the third modality is used; It is QU-PRIP database of online signature [92].

Considering the number of subjects in the VidTimit database, 43 subjects have been selected from the QU-PRIP signature verification datasets. From both subsets, and by taking advantage of the independence of face and voice from online signature trait, 43 virtual subjects have been created.

The following training and testing process for monomodal systems has been established:

For training purposes, each face has been modeled using four samples, and each voice and online signature has been modeled using the same number of samples and one sample of each trait for validation purposes.

For validation and testing, for each client one more sample of each trait (face, voice and online signature) were also selected for validation and one more sample for testing; the same 43 clients are used as impostors, except that each client claims an identity different from his own. Each client has been considered and, from each impostor, one sample has been selected.

Consequently, the sub-corpus for the experiments consists of 43 clients, and 42x43x2=3612 multimodal impostor attempts.

### *6.3 Scenario 3*

The SDUMLA-HMT multimodal database is used. It is comprised of face, finger vein, iris and fingerprint images and gait videos of 106 persons.

The following training and testing process for monomodal systems has been established:

For training purposes, each face has been modeled using four samples, and each fingerprint and finger vein has been modeled using the same number of samples and one sample of each trait for validation purposes.

For validation and testing, for each client one more sample of each modality (face, fingerprint and finger vein) were also selected for validation and one more sample for testing; the same 106 clients are used as impostors, except that each client claims an identity different from his own. Each client has been considered and, from each impostor, one sample has been selected.

Consequently, the corpus for the experiments consists of 106 clients, and 105x106x2=22260 multimodal impostor attempts.

## *7 Experiment results and discussion*

The experiments demonstrate that hierarchical fusion-based method improves the efficiency of the multimodal authentication compared with both score and feature based fusion methods.

Table. 2 demonstrates that, for identification purposes, the best recognition rate (RR) is obtained by hierarchical and scores fusion strategies (100%), furthermore, without making errors based on LDA, but the results obtained by features fusion strategy is not so far. Additionally, for verification purposes, the lowest equal error rate (EER) is obtained by using the hierarchical fusion and score fusion strategies (0.00%).

Table 2. Comparison of performance metrics of scenario 1.

| | PCA | | | | | | LDA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) |
| Face | 40.00 | 0.17 | 0.17 | 100 | 96.00 | 4.00 | 96.00 | 0.17 | 0.17 | 100 | 96.00 | 4.00 |
| Fingerprint | 44.00 | 12.00 | 6.83 | 84.00 | 84.00 | 4.00 | 84.00 | 12.00 | 6.33 | 88.00 | 84.00 | 4.00 |
| Online signature | 80.00 | 0.25 | 0.25 | 100 | 92.00 | 4.00 | 100 | 0.00 | 0.00 | 100 | 100 | 4.00 |
| feature level fusion | 60.00 | 3.25 | 1.25 | 92.00 | 92.00 | 4.00 | 92.00 | 3.75 | 1.75 | 92.00 | 92.00 | 4.00 |
| score level fusion | 92.00 | 0.00 | 0.00 | 100 | 100 | 4.00 | 100 | 0.00 | 0.00 | 100 | 100 | 4.00 |
| hierarchical fusion | 96.00 | 0.00 | 0.00 | 100 | 100 | 4.00 | 100 | 0.00 | 0.00 | 100 | 100 | 4.00 |

Table. 3 demonstrates that, for identification purposes, the best recognition rate (RR) is obtained by hierarchical and scores fusion strategies (100%), furthermore, without making errors based on LDA, but the results obtained by features fusion strategy is not so far. Additionally, for verification purposes, the lowest equal error rate (EER) is obtained by using the hierarchical and score fusion strategies (0.00%).

Table 3. Comparison of performance metrics of scenario 2.

| | PCA | | | | | | LDA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) |
| Face | 88.37 | 2.41 | 1.99 | 95.35 | 95.35 | 2.33 | 95.35 | 2.19 | 1.02 | 97.67 | 95.35 | 2.33 |
| Voice | 16.28 | 25.58 | 23.03 | 20.93 | 4.65 | 2.33 | 30.23 | 23.26 | 22.37 | 25.58 | 9.30 | 2.33 |
| Online signature | 83.72 | 1.77 | 0.61 | 97.67 | 88.37 | 2.33 | 95.35 | 0.3 | 0.3 | 100 | 86.05 | 2.33 |
| feature level fusion | 88.37 | 2.19 | 1.05 | 95.35 | 93.02 | 2.33 | 95.35 | 0.39 | 0.39 | 100 | 95.35 | 2.33 |
| score level fusion | 97.67 | 0.00 | 0.00 | 100 | 100 | 2.33 | 100 | 0.00 | 0.00 | 100 | 100 | 2.33 |
| hierarchical fusion | 97.67 | 0.00 | 0.00 | 100 | 100 | 2.33 | 100 | 0.00 | 0.00 | 100 | 100 | 2.33 |

The obtained recognition rate and equal error rate are perfect due to the use of small databases BIO (25 persons) and VidTimit (43 persons), this fact pushed us to generalize these results and evaluate the performances using a large database that contain 106 persons (SDUMLA-HMT).

Table 4. Comparison of performance metrics of scenario 3.

| | PCA | | | | | | LDA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) | RR at Rank one (in %) | EER (in %) | MHTER (in %) | VR at 1% FAR (in %) | VR at 0.1% FAR (in %) | VR at 0.01% FAR (in %) |
| Face | 74.53 | 6.69 | 6.55 | 85.85 | 74.53 | 62.26 | 82.08 | 5.67 | 5.52 | 87.74 | 73.58 | 66.98 |
| Fingerprint | 66.98 | 5.68 | 5.15 | 86.79 | 71.70 | 66.98 | 81.13 | 3.77 | 3.50 | 89.62 | 74.53 | 66.98 |
| Finger vein | 66.04 | 1.10 | 0.98 | 96.23 | 90.57 | 83.96 | 93.40 | 0.39 | 0.39 | 100 | 93.40 | 85.85 |
| feature level fusion | 97.17 | 0.03 | 0.03 | 100 | 100 | 97.17 | 99.06 | 0.05 | 0.05 | 100 | 99.06 | 97.17 |
| score level fusion | 96.23 | 0.02 | 0.02 | 100 | 100 | 99.06 | 99.06 | 0.01 | 0.01 | 100 | 100 | 99.06 |
| hierarchical fusion | 96.23 | 0.01 | 0.01 | 100 | 100 | 99.06 | 99.06 | 0.01 | 0.01 | 100 | 100 | 98.11 |

Table. 4 demonstrates that, for identification purposes, the best recognition rate (RR) is obtained by hierarchical and scores fusion strategies (96.23%, 99.06%). Additionally, for verification purposes, the lowest equal error rate (EER) is obtained by using the hierarchical fusion strategy (0.01%), whereas, those of score fusion and feature fusion are 0.02, 0.03 respectively. Additionally, features and scores fusion strategy demonstrate respectively low equal error rates in comparison with the monomodal systems. The same observation is noted for the minimal half total error rate (MHTER) and for the verification rate (VR).

It is clear that results obtained by fusion strategies are much better than those obtained by monomodal biometric systems, and the results obtained using LDA are much better than those obtained using PCA.

Furthermore, as shown in figures from fig. 37 to fig. 42, the Cumulative Match Characteristic (CMC) curve is used to evaluate the identification performance as a comparison for scenarios 1, 2 and 3.

Hierarchical and scores fusion strategies generate the better results (100%).

Using LDA, hierarchical and scores fusion have identical curves, but using PCA, it is clear that the best curve is the one obtained by hierarchical fusion.

To evaluate the performance of verification, Receiver Operating Characteristic (ROC) curve and Detection Error Trade-off (DET) curve are used and the results are shown in figures from fig. 43 to fig. 54. The ROC and the DET curves of face, fingerprint, finger vein, voice, online signature, features fusion, scores fusion and hierarchical fusion based verification as comparisons are given in these figures.

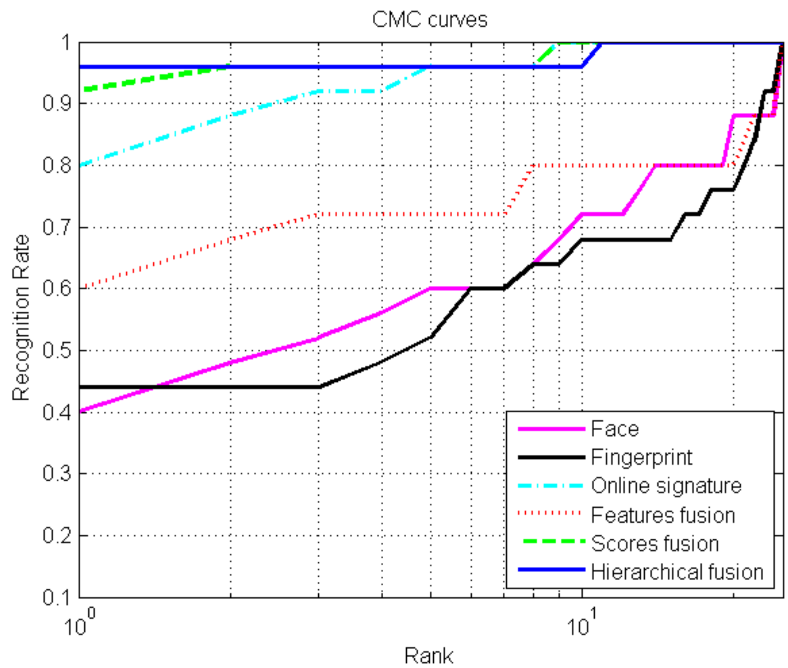


Figure 37. PCA based CMC curves of scenario 1.

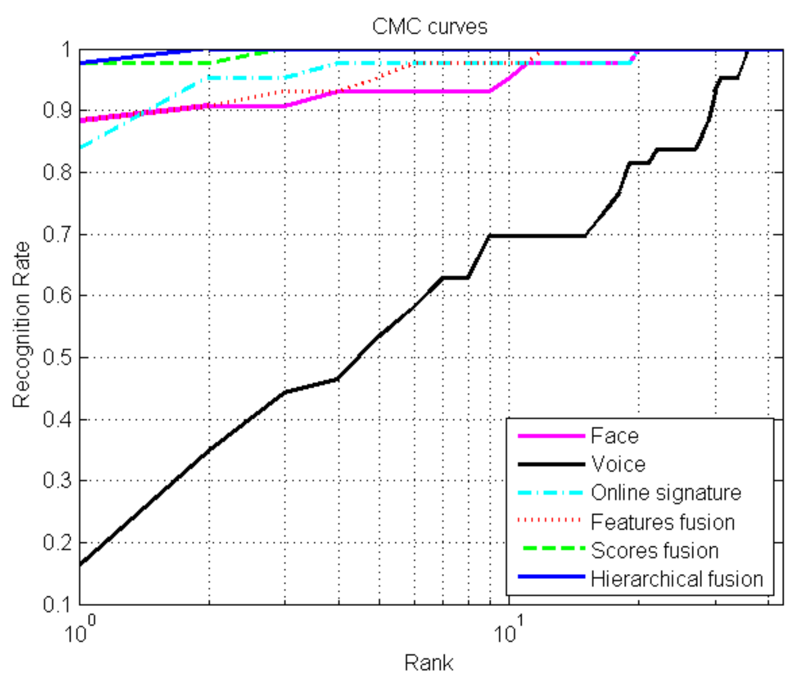


Figure 38. PCA based CMC curves of scenario 2.

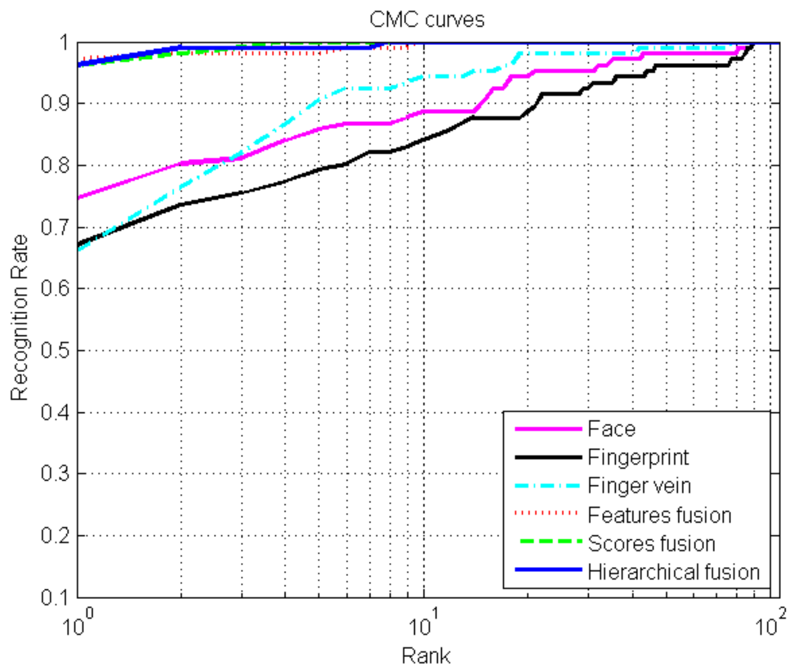


Figure 39. PCA based CMC curves of scenario 3.

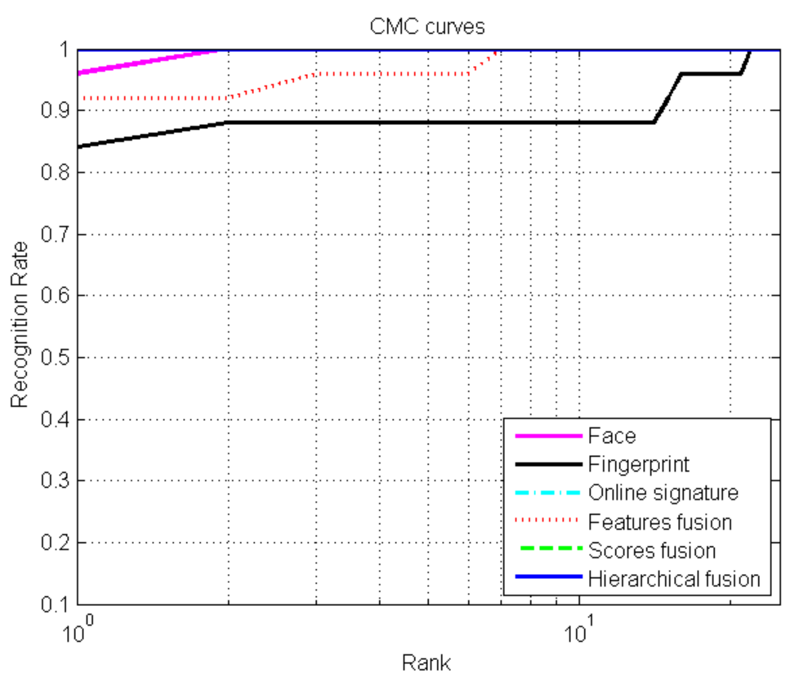


Figure 40. LDA based CMC curves of scenario 1.

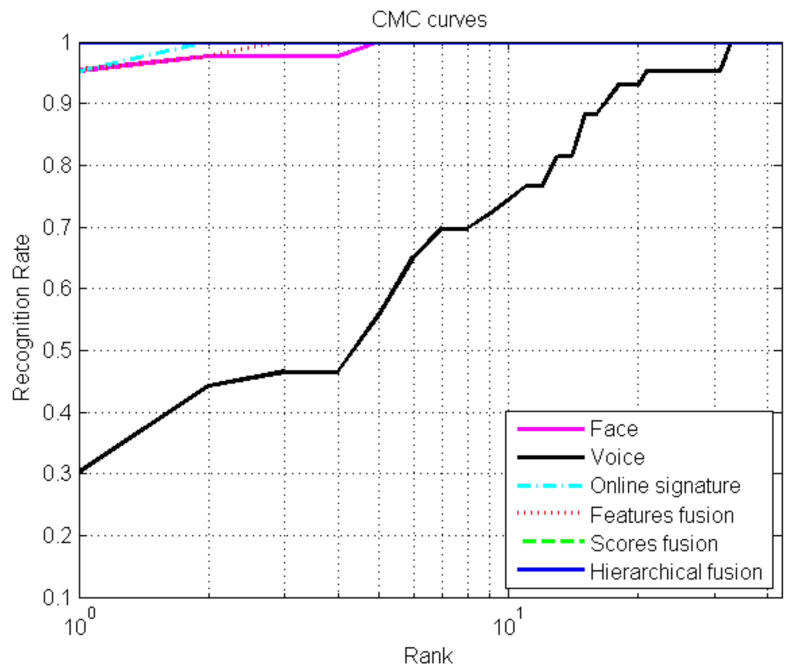


Figure 41. LDA based CMC curves of scenario 2.

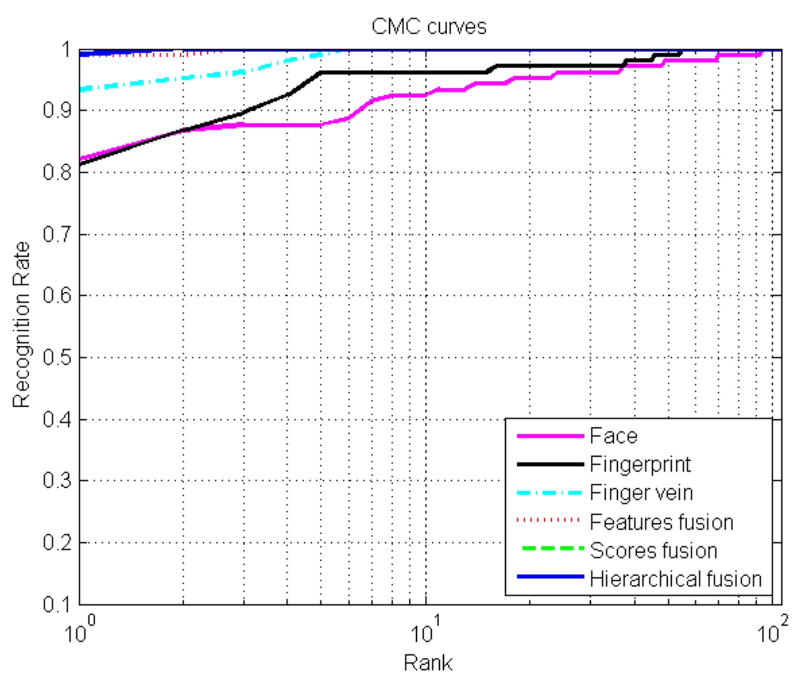


Figure 42. LDA based CMC curves of scenario 3.

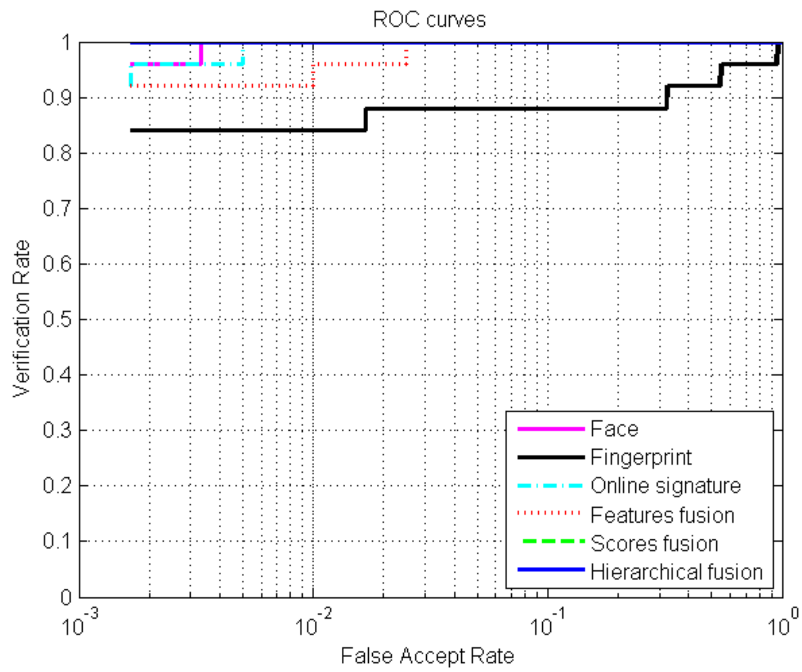


Figure 43. PCA based ROC curves of scenario 1.

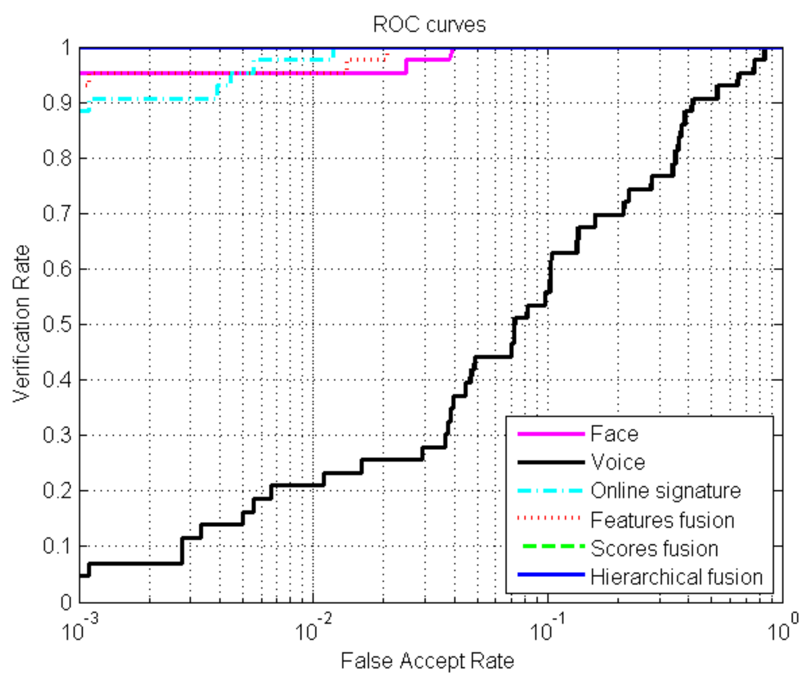


Figure 44. PCA based ROC curves of scenario 2.

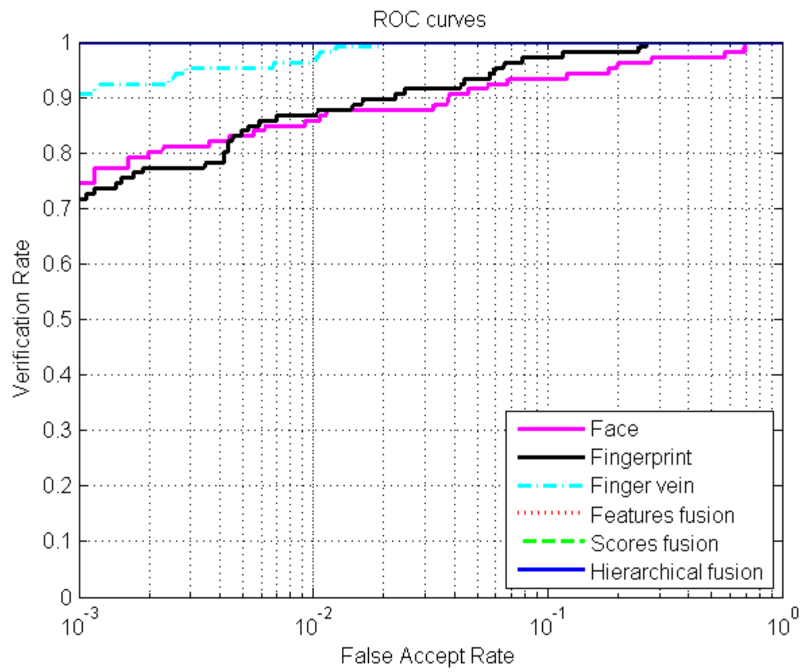


Figure 45. PCA based ROC curves of scenario 3.

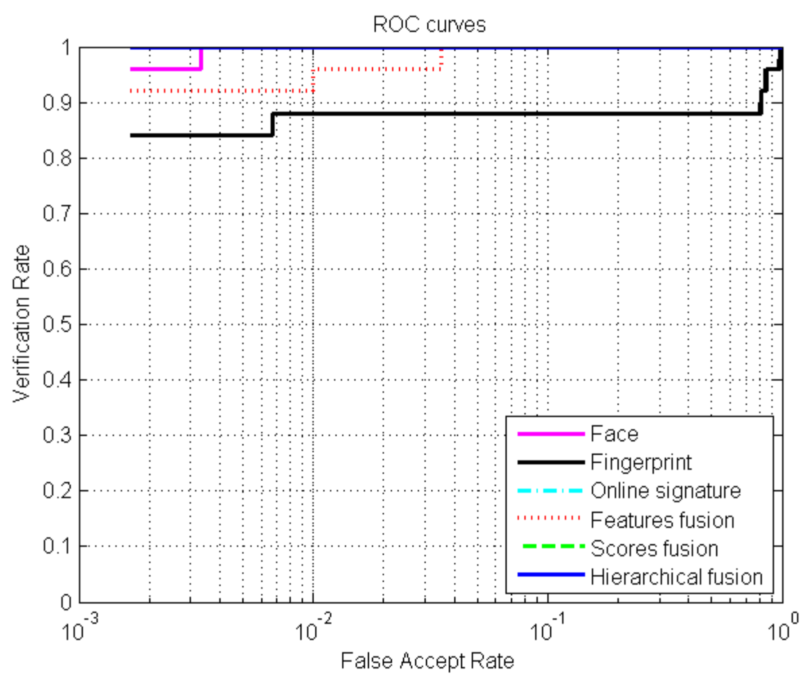


Figure 46. LDA based ROC curves of scenario 1.

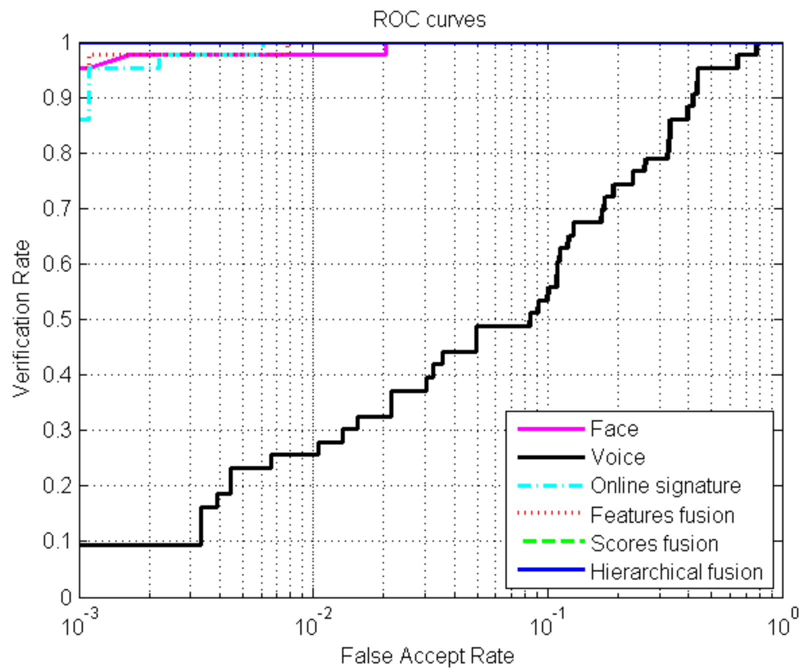


Figure 47. LDA based ROC curves of scenario 2.

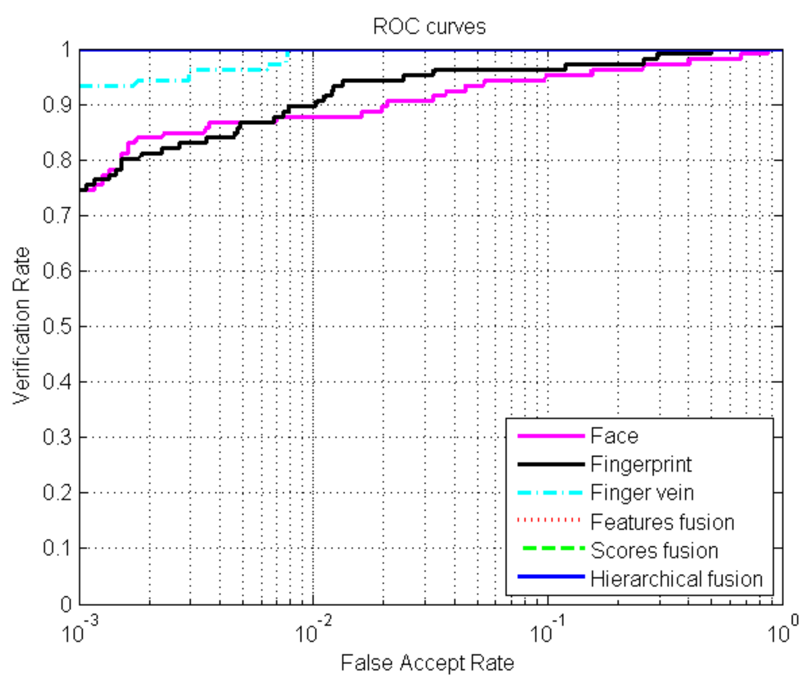


Figure 48. LDA based ROC curves of scenario 3.

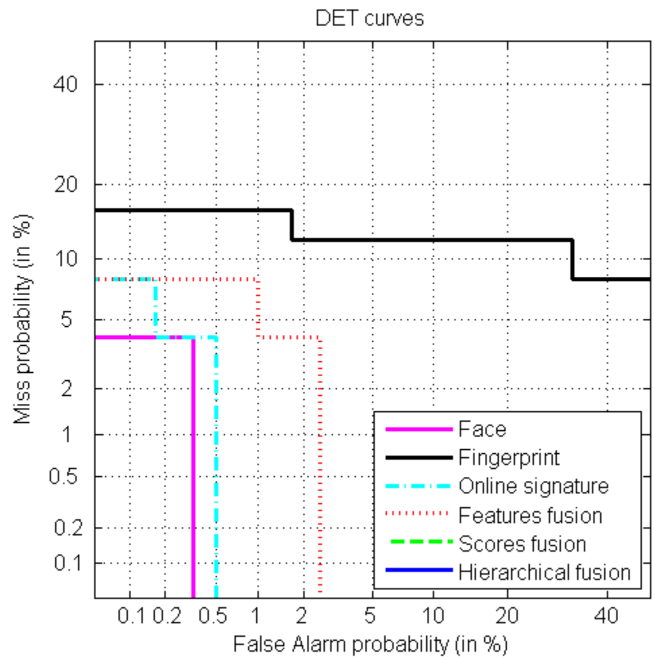


Figure 49. PCA base DET curves of scenario 1.

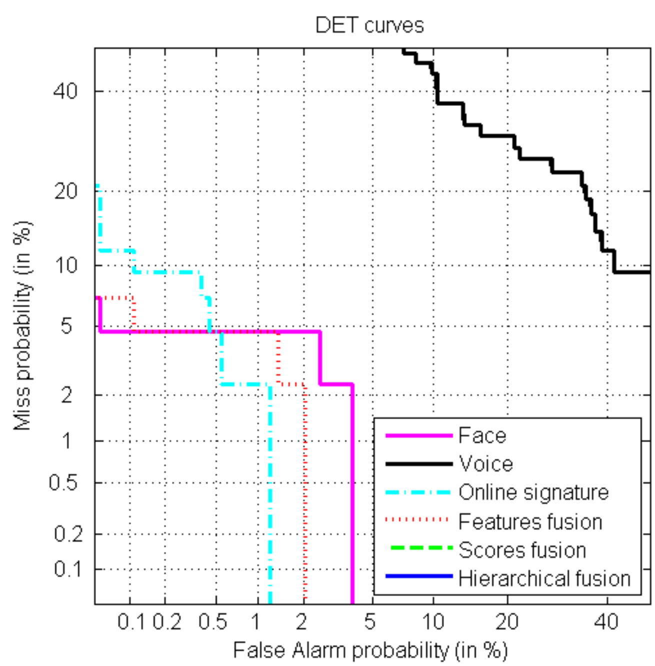


Figure 50. PCA base DET curves of scenario 2.

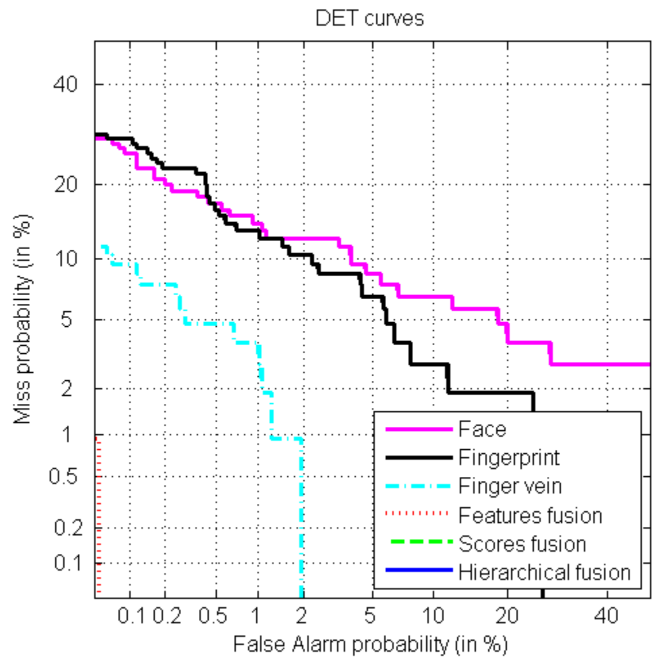


Figure 51. PCA base DET curves of scenario 3.

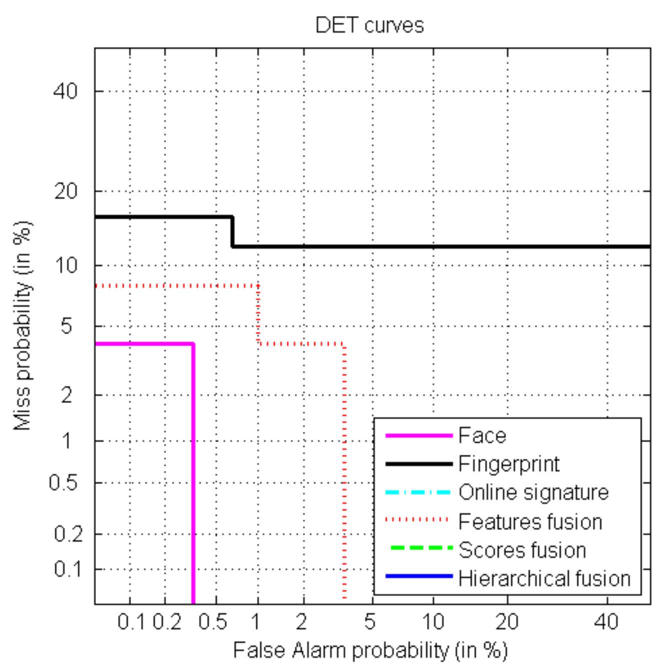


Figure 52. LDA based DET curves of scenario 1.

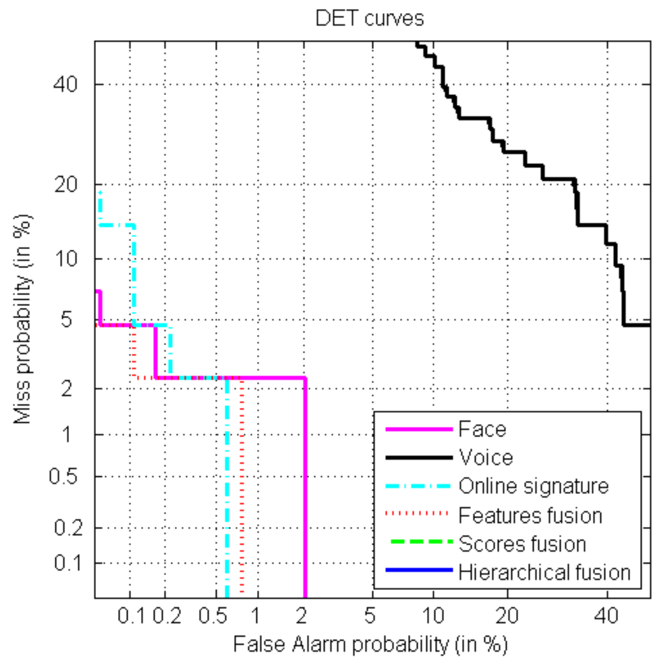


Figure 53. LDA based DET curves of scenario 2.

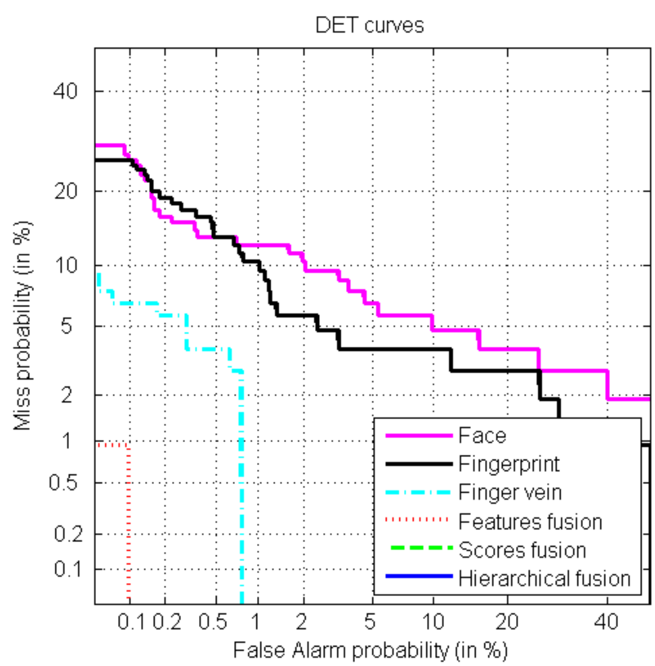


Figure 54. LDA based DET curves of scenario 3.

Appropriate selection of parameters for the feature fusion based approach has provided better recognition performance than the monomodal based approach. It has been shown that the feature based system increased the recognition rate. Encouraging initial results of feature base approach motivate further research in order to exploit user specificities in the fusion stage of multimodal biometric recognition systems.

In addition, different fusion techniques are explored, such as score fusion. The experiments highlight the benefits of using the max-of-scores rule for score fusion. In fact, the combination modalities based on a max-of-scores analysis outperformed the combination based on the min and sum-of-scores based strategies as well as the single modality-based authentication systems such as face and online signature biometric systems. Furthermore, the experimental results reveal an improvement of the monomodal biometric system compared with the best classifiers that use a single modality (online signature/finger vein).

It is noted that the ROC curves of hierarchical and score fusion strategies are identical, because that they generate the highest recognition rate using both of PCA and LDA.

On the contrary of ROC and DET curves, where hierarchical and scores fusion strategies obtained the same verification performance, the hierarchical fusion strategy produces less errors than scores fusion strategy in the task of identification, it is clear that the hierarchical fusion based multimodal systems could benefit from the advantages of the other fusion strategies to improve the method's overall efficiency.

From these figures, the hierarchical fusion improves the performance, which indicates the high effectiveness of multimodal biometrics. This strategy significantly improves performances of both features and scores fusion strategies.

Table 5. Comparison with existing methods

<table>
<tr><th colspan="3">Method</th><th>EER (%)</th><th>Database</th><th>Population</th></tr>
<tr><td colspan="3">Face + sign (sum rule based) [14]</td><td>> 5</td><td>XM2VTS</td><td>50</td></tr>
<tr><td colspan="3">SVM (face,sign) user-indep [14]</td><td>> 10</td><td>XM2VTS</td><td>50</td></tr>
<tr><td colspan="3">SVM classifier (face and sign) [15]</td><td>5.54</td><td>Biosecure</td><td>500</td></tr>
<tr><td colspan="3">Hierarchical Fusion of Face and Iris [16]</td><td>/</td><td>Their own database</td><td>82</td></tr>
<tr><td colspan="3">A two level hierarchical fusion of face images + SVM classifier [17]</td><td>2.86</td><td>Equinox face database</td><td>91</td></tr>
<tr><td colspan="3">MinMax + GMM (face, voice & sign) [107]</td><td>2.39,1.54, 2.3</td><td>PDA</td><td>60</td></tr>
<tr><td colspan="3">Fused feature with PCA (face, voice & sign) [108]</td><td>1</td><td>Their own</td><td>20</td></tr>
<tr><td rowspan="6">The proposed</td><td rowspan="2">Feature fusion</td><td>PCA</td><td>0.03</td><td rowspan="6">SDUMLA-HMT</td><td rowspan="6">106</td></tr>
<tr><td>LDA</td><td>0.05</td></tr>
<tr><td rowspan="2">Score fusion</td><td>PCA</td><td>0.02</td></tr>
<tr><td>LDA</td><td>0.01</td></tr>
<tr><td rowspan="2">Hierarchical fusion</td><td>PCA</td><td>0.01</td></tr>
<tr><td>LDA</td><td>0.01</td></tr>
</table>

Table. 5 gives a comparison of existing methods and the proposed method. It is evident from this table that the results obtained by the proposed method are comparatively much better than the existing methods. The lowest EER obtained is 0.01%. For verification purposes, the scores fusion based method gives same results as the hierarchical fusion but in terms of identification purposes, as it is shown in Fig. 19, the best performance is the one obtained by the hierarchical strategy of fusion. Thus, it is evident that the results are comparatively outperforming other methods.

## *8 Summary*

Any multimodal biometric system is expected to achieve reduced equal error rate (EER). To release this requirement, the hierarchical strategy is used for biometric traits fusion. The main contributions of this thesis are summarized as follows. The proposed method, firstly, reduces the fusion template size by the fusion of features of two biometric traits instead of three at the same time, and then by using principal component analysis (PCA) or Linear discriminant analysis (LDA) for data reduction to improve system speed. The hierarchical fusion strategy being the combination of features fusion and scores fusion. The reduced feature vector size thus improves the system speed without paying any significant cost in terms of accuracy. Secondly, the fusion strategy used in the proposed method has reduced the equal error rate significantly. Finally, this work compares the obtained results with few of the existing methods.

# Conclusion and future work

This chapter is a summary of the proposed framework for multimodal biometric based on hierarchical fusion of face, fingerprints, voice and/or online signature, finger vein modalities. Nevertheless, there still many possibilities to be investigated in future work.

## 1 Summary and contribution

In the literature, there are very few teams that are working in the field of biometric hierarchical fusion strategy.

In this thesis a new vision is introduced for a highly accurate biometric system which combines face, fingerprint, voice, finger vein and/or online signature authentication systems in order to optimize the accuracy and performance. The proposed approach is based on hierarchical multilevel biometric fusion integration: feature-level fusion and matching at score-level. The hierarchical biometric fusion provides most of the overall performance improvement for the whole multimodal biometric system.

In addition, different fusion techniques are explored, such as the feature fusion and score fusion for face, fingerprint, voice, finger vein and/or online signature data. A multimodal biometric system that merges evidence from these modalities is released. The experiments highlight the benefits of using the hierarchical strategy for multimodal based authentication. In fact, the combination of modalities based on hierarchical strategy outperformed the combination based on feature or score fusion, as well as the single modality based authentication systems such as face, fingerprint, finger vein, voice and/or online signature biometric systems. Furthermore, the experimental results reveal an improvement of the monomodal biometric system compared with the best classifiers that use a single modality (online signature, finger vein).

Moreover, a statistical-motivated experimental process has been introduced and applied in order to compare best referenced fusion based and monomodal based biometric recognition rate by means of DET plots. Cosine Mahalanobis distance has been proposed to classify fused vectors that have been derived.

## *2 Future work*

The work presented in this thesis leads to a several possible future works that can be investigated. The fusion at feature level in multimodal biometric can be extended by several ideas by improving the performance of feature extraction, selection and combination and especially by localizing the region of interest, fact that was not considered in this thesis.

Features of face, signature, voice, fingerprint and finger vein are computed for the fusion process. However, some of them may consist low frequency information thus will have high discrimination power, while the others may contain high frequency information thus not effective for discrimination. In the future work, pre-processing of the information from each biometric trait can be done to select the area that has high discrimination power and eliminate the zone with redundant features. This technique may increase the useful information in the fused feature vector when the less informative zone is removed from the feature vector.

Regarding face, finger vein and fingerprint images, the proposed feature extraction and fusion method is designed to deal with grayscale images. The input image is treated at grayscale level. This method could be extended in the future to treat images at color level that might contain more useful information. The Fusion of information extracted from the red, green and blue components of an image might generate a consistent fused feature vectors.

This work can be extended to other biometric traits such as iris, and gait. The hierarchical fusion framework discussed in this thesis can be generalized so as to be employed with different biometric traits. Some of the biometrics traits require similar processing and feature extraction technique as discussed in this thesis. However, other biometric traits such as online signature and voice require a different method to transform the features to a compatible form for fusion process. The score generated from these types of biometric traits could then be used in the hierarchical fusion process.

# Bibliography


[1] A. K. Jain and A. Ross, "Multibiometric systems," *Communications of the ACM,* vol. 47, no. 1, pp. 34-40, 2004.

[2] Y. Chen, S. C. Dass and A. K. Jain, "Fingerprint quality indices for predicting authentication performance," *Audio-and Video-Based Biometric Person Authentication,* pp. 160-170, 2005.

[3] The National Institute of Standards and Technology (NIST), "NIST report to the United States Congress. Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability," November 2002. [Online]. Available: ftp://sequoyah.nist.gov/pub/nist_internal_reports/NISTAPP_Nov02.pdf.

[4] T. Matsumoto, H. Matsumoto, K. Yamada and S. Hoshino, "Impact of artificial gummy fingers on fingerprint systems," *Electronic Imaging,* pp. 275-289, April 2002.

[5] T. Van der Putte and J. Keuning, "Biometrical fingerprint recognition: don't get your fingers burned," *Smart Card Research and Advanced Applications,* pp. 289-303, 2000.

[6] Hitachi, *Finger* Vein *Authentication- White Paper,",* 2006.

[7] M. Naoto, A. Nagasaka and M. Takafumi, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Machine Vision and Applications,* pp. 194 -203, 2004.

[8] S. Khellat-kihel, R. abrishambaf, N. Cardoso, J. Monteiro and M. Benyettou, "Finger Vein Recognition Using Gabor Filter and Support Vector Machine," in *IEEE* INTERNATIONAL *IMAGE PROCESSING APPLICATIONS AND SYSTEMS*, 2014.

[9] N. Morizet*, Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris,* Télécom ParisTech, 2009.

[10] A. K. Jain, P. J. Flynn and A. A. Ross, Handbook of biometrics, Springer, 2008.

[11] F. Perronnin and J.-L. Duglay, "Introduction a la biométrie. Authentification des individus par traitement Audio Vidéo," *Traitement du signal,* vol. 19, no. 4, 2002.

[12] R. Beveridge and M. Kirby, "Biometrics and Face Recognition," *IS&T Colloquium,* p. 25, 2005.

[13] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha and A. W. Senior, "The relation between the ROC curve and the CMC," in *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, 2005.

[14] J. Fiérrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero and J. Gonzalez-Rodriguez, "A comparative evaluation of fusion strategies for multimodal biometric verification.," *Audio-and Video-based Biometric Person Authentication,* pp. 830-837, 2003.

[15] B. Dorizzi, Multi-*biometrics: Score Fusion Strategies,* Université d'EVRY, 2009.

[16] Z. Xiaobo, Z. Sun and T. Tan, "Hierarchical fusion of face and iris for

personal identification," in *IEEE 20th International Conference on Pattern Recognition*, 2010.

[17] R. Singh, M. Vatsa and A. Noore, "Hierarchical fusion of multi-spectral face images for improved recognition performance," *Information Fusion,* vol. 9, no. 2, pp. 200-210, 2008.

[18] W. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, "Face Recognition: A Literature Survey," *ACM Computing Surveys,* p. 399–458, 2003.

[19] R. Singh, M. Vatsa, A. Ross and A. Noore, "Performance Enhancement of 2D Face Recognition via Mosaicing," in *4th IEEE Workshop on Automatic Identification Advanced Technologies*, 2005.

[20] X. Liu and T. Chen, "Pose Robust Face Recognition Based on Mosaicing – An Example Usage of Face In Action (FIA) Database," in *the IEEE International Conference on Computer Vision and Pattern Recognition*, 2004.

[21] L. Wiskott, J. M. Fellous, N. Kruger and C. Von Der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 19, no. 7, p. 775–779, 1997.

[22] A. K. Jain, D. Maltoni, D. Maio and S. Prabhakar, Handbook of Fingerprint Recognition, New York: Springer, 2003.

[23] A. K. Jain, S. Prabhakar, L. Hong and S. Pankanti, "Filterbank-Based Fingerprint Matching," *IEEE Transactions on Image Processing,* vol. 9, p. 846–859, 2000.

[24] Y. He, J. Tian, L. Li, H. Chen and X. Yang, "Fingerprint matching based on global comprehensive similarity," *IEEE Transactions on Pattern Analysis*

*Machine Intelligence,* vol. 28, no. 6, p. 850–862, 2006.

[25] M. Tico and P. Kuosmanen, "Fingerprint matching using an orientation-based minutia descriptor," *IEEE Transactions on Pattern Analysis Machine Intelligence,* vol. 25, no. 8, p. 1009–1014, 2003.

[26] J. Gu, J. Zhou and C. Yang, "Fingerprint recognition by combining global structure and local cues," IEEE *Transactions on Image Processing,* vol. 15, no. 7, pp. 1942-1964, 2006.

[27] D. Wan and J. Zhou, "Fingerprint recognition using model-based density map," *IEEE Transactions on Image Processing,* vol. 15, no. 6, p. 1690–1696, 2006.

[28] Y. Elmir, Z. Elberrichi and R. Adjoudj, "Liquid State Machine based Fingerprint Identification," *Australian Journal Of Basic and Applied Sciences,* vol. 5, no. 5, pp. 857-865, 2011.

[29] J. F. Yang, Y. S. Shi and J. L. Yang, "Person identification Based on finger-vein features," *Computers in Human Behavior,* vol. 28, pp. 1565-1570, 2011.

[30] J. C. Hashimoto, "Finger vein authentication technology and its future," in *Symposium on VLSI Circuits Digest of Technical Papers*, Honolulu, USA, 2006.

[31] T. Yanagawa, S. Aoki and T. Ohyama, "Human finger vein images are diverse and its patterns are useful for personal identification," *Kyushu University MHF Preprint Series,* pp. 1-7, 2007.

[32] N. Miura, A. Nagasaka and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Machine Vision and Applications,* vol. 15, p. Machine Vision

and Applications, 2004.

[33] Z. Zhang, S. Ma and X. Han, "Multiscale feature extraction of finger-vein patterns based on curvelet and local interconnection structure neural network," in *18th international conference on pattern recognition*, Hong Kong, China, 2006.

[34] W. Song, T. Kim, H. C. Kim, J. H. Choi, H. Kong and S. Lee, "A finger-vein verification system using mean curvature," *Pattren Recognition Letters,* vol. 32, pp. 1541-1547, 2011.

[35] E. C. Lee, H. C. Lee and K. R. Park, "Finger vein recognition using minutia-based alignment and local binary pattern-based feature extraction," *International Journal of Imaging Systmes and Technology,* vol. 19, pp. 179-186, 2009.

[36] E. C. Lee, H. Jung and D. Kim, "New finger biometric method using near infrared imaging," *Sensors,* vol. 11, p. Sensors, 2011.

[37] A. Petpon and S. Srisuk, "Face recognition with local line binary pattern," in *the Fifth* International *Conference on Image and Graphics*, Xi‟an, China, 2009.

[38] A. R. Bakhtiar, W. S. Chai and A. S. Shahrel, "Finger vein recognition using local line binary pattern," *Sensors,* vol. 11, pp. 11357-11371, 2011.

[39] L. Yu, Y. Sook, J. X. Shan and S. P. Dong, "Finger Vein Identification Using Polydirectional Local Line Binary Pattern," in *IEEE ICTC*, 2014.

[40] H. Hermansky, B. Hanson and H. Wakita, "Perceptually based linear predictive analysis of speech," in *the IEEE International Conference on*

*Acoustics, Speech and Signal Processing*, 1985.

[41] S. Kajarekar, L. Ferrer, K. Sonmez, J. Zheng, E. Shriberg and A. Stolcke, "Modeling NERFs for speaker recognition," in *IEEE Odyssey*, Toledo, Spain, 2004.

[42] H. Hermansky and N. Morgan, "Rasta processing of speech," *IEEE Transactions on* Speech *and Audio Processing,* vol. 2, no. 4, p. 578–589, 1984.

[43] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification: The state of the art," *Pattern Recognition,* vol. 22, no. 2, p. 107–131, 1989.

[44] F. Leclerc and R. Plamondon, "Automatic signature verification: The state of the art, 1989–1993," *Internatiional Journal of Pattern Recognition and Artificial Intelligence,* vol. 8, no. 3, p. 643–660, 1994.

[45] R. Plamondon and S. N. Srihari, "Survey, On-line and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. PAMI,* vol. 22, no. 1, p. 63–84, 2000.

[46] L. L. Lee, T. Berger and E. Aviczer, "Reliable on-line human signature verification systems," *IEEE Trans. on PAMI,* vol. 18, no. 6, p. 643–647, 1996.

[47] H. Ketabdar, J. Richiardi and A. Drygajlo, "Global feature selection for online signature verification," in *12th International Graphonomics Society*, 2005.

[48] V. S. Nalwa, "Automatic on-line signature verification," Proceedings of IEEE, vol. 85, no. 2, pp. 215-239, 1997.

[49] M. C. Fairhurst, "Signature verification revisited: Promoting practical exploitation of biometric technology," *IEE Electronics and Communication Engineering Journal,* vol. 9, no. 6, p. 273–280, 1997.

[50] A. K. Jain, F. D. Griess and S. D. Connell, "On-line signature verification," *Pattern Recognition,* vol. 35, no. 12, p. 2963–2972, 2002.

[51] B. Li, D. Zhang and K. Wang, "On-line signature verification based on NCA (Null Component Analysis) and PCA (Principal Component Analysis)," *Pattern Analysis and Application,* vol. 8, p. 345–356, 2006.

[52] H. Lei and V. Govindaraju, "A comparative study on the consistency of features in on-line signature verification," *Pattern Recognition Letters,* vol. 26, no. 15, p. 2483–2489, 2005.

[53] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia and D. Maltoni, "An on-line signature verification system based on fusion of local and global Information," in *AVBPA*, 2005.

[54] A. K. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. on PAMI,* vol. 19, no. 2, p. 153–158, 1997.

[55] L. Lee, On-Line Systems for Human Signature Verification, Cornell University, 1992.

[56] M. Fairhurst and P. Brittan, "An evaluation of parallel strategies for feature vector construction in automatic signature verification systems," *International Journal of Pattern Recognition and Artificial. Intelligence,* vol. 8, no. 3, p. 661–678, 1994.

[57] J. Brault and R. Plamondon, "A complexity measure of handwritten curves: Modeling of dynamic signature forgery," *IEEE Trans. on SMC,* vol. 23, no. 2, p. 400–413, 1993.

[58] M. Fairhurst, E. Kaplani and R. Guest, "Complexity measures in handwritten signature verification," in *1st Int. Conf. on Universal Access in HumanComputer Interaction*, 2001.

[59] V. Struc and V. Pavesic, "The Complete Gabor-Fisher Classifier for Robust Face Recognition," *EURASIP Advances in Signal Processing,* vol. 2010, p. 26, 2010.

[60] V. Struc and V. Pavesic, "Gabor-Based Kernel Partial-Least-Squares Discrimination Features for Face Recognition," *Informatica (Vilnius),* vol. 20, no. 1, pp. 115-138, 2009.

[61] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, no. 5, pp. 725-737, 2006.

[62] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing,* vol. 11, no. 4, p. 467–476, 2002.

[63] W. Kienzle, M. O. Franz, B. Scholkopf and G. H. Bakir, "Face Detection---Efficient and Rank Deficient," *Advances in Neural Information Processing Systems,* pp. 673-680, 2004.

[64] Y. Mami, Reconnaissance de locuteurs par localisation dans un espace de locuteurs de référence, Paris: Télécom ParisTech, 2003.

[65] N. D. Minh, "An automatic speaker recognition system," Audio Visual Communications Laboratory, Swiss Federal Institute of Technology,, Lausanne, Switzerland, 2001.

[66] C. Cornaz, U. Hunkeler and V. Velisavljevic, "An automatic speaker recognition system," Digital Signal Processing Laboratory, Federal Institute of Technology, Lausanne, Switzerland, 2003.

[67] A. Shukla, J. Dhar, C. Prakash, D. Sharma, R. K. Anand and S. Sharma, "Intelligent Biometric System using PCA and R-LDA," in *Proceedings of WRI Global Congress on Intelligent Systems*, 2009.

[68] A. H. Boualleg, C. Bencheriet and H. Tebbikh, "Automatic Face recognition using neural network-PCA," in *Proceedings of Information and Communication Technologies*, April 2006..

[69] H. Kong, X. Li, L. Wang, E. K. Teoh, J. G. Wang and R. Venkateswarlu, "Generalized 2D principal component analysis," in *Proceedings of International Joint Conference on Neural Networks*, 2005.

[70] W. S. Yambor, "Analysis of PCA-Based and Fisher Discriminant-Based Image Recognition Algorithms," Colorado, July 2000.

[71] A. M. Martinez and K. A. C., "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, no. 2, pp. 228-233, 2001.

[72] J. Yang, Y. Yu and W. Kunz, "An Efficient LDA Algorithm for Face Recognition," in *Proceeding of the Sixth International Conference on Control, Automation, Robotics and Vision*, 2000.

[73] D. S. Jeong, H.-A. Park, K. R. Park and J. Kim, "Iris recognition in mobile

phone based on adaptive gabor filter," in *Advances in Biometrics. Springer*, Berlin Heidelberg, 2005.

[74] A. Ross, K. Nandakumar and A. K. Jain, Handbook of Multibiometrics., Springer, 2006.

[75] K. Nandakumar, *Multibiometric systems: Fusion strategies and template security,* Michigan State University, 2008.

[76] L. Hong, A. K. Jain and S. Pankanti, "Can Multibiometrics Improve Performance?," in *IEEE Workshop on Automatic Identification Advanced Technologies*, New Jersey, USA, 1999.

[77] K. I. Chang, K. W. Bowyer and P. J. Flynn, "An Evaluation of Multimodal 2D+3D Face Biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, no. 4, p. 619–624, 2005.

[78] R. Brunelli and D. Falavigna, "Person Identification Using Multiple Cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 17, no. 10, p. 955–966, 1995.

[79] D. Maltoni, D. Maio, A. K. Jain and S. Prabhakar, Handbook of fingerprint recognition, Springer, 2009.

[80] M. I. Ahmad, Feature *extraction and information fusion in face and palmprint multimodal biometrics.,* 2013.

[81] A. K. Jain and A. Ross, "Fingerprint mosaicking," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002.

[82] A. A. Ross and R. Govindarajan, "Feature level fusion of hand and face biometrics," *Defense and Security. International Society for Optics and*

*Photonics,* 2005.

[83] R. Raghavendra, B. Dorizzi, A. Rao and G. H. Kumar, "Particle swarm optimization based fusion of near infrared and visible images for improved face verification," *Pattern Recognition,* vol. 44, no. 2, pp. 401-411, 2011.

[84] R. Raghavendra, B. Dorizzi, A. Rao and G. H. Kumar, "Designing efficient fusion schemes for multimodal biometric systems using face and palmprint," *Pattern Recognition,* vol. 44, no. 5, pp. 1076-1088, 2011.

[85] G. L. Marcialis and F. Roli, "Fingerprint verification by fusion of optical and capacitive sensors," *Pattern Recognition Letters,* vol. 25, no. 11, pp. 1315-1322, 2004.

[86] A. Jain, K. Nandakumar and A. Ross, "Score normalization in multimodal biometric systems," *Pattern recognition,* vol. 38, no. 12, pp. 2270-2285, 2005.

[87] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern recognition letters,* vol. 24, no. 13, pp. 2115-2125, 2003.

[88] L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,* vol. 27, no. 5, pp. 553-568, 1997.

[89] L. Lam and C. Y. Suen, "Optimal combinations of pattern classifiers," *Pattern Recognition Letters,* vol. 16, no. 9, pp. 945-954, 1995.

[90] J. Kittler, M. Hatef, R. P. Duin and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, no. 3, pp. 226-239, 1998.

[91] Y. Elmir, Z. Elberrichi and R. Adjoudj, "Score Level Fusion Based Multimodal Biometric Identification (Fingerprint & Voice)," in *IEEE International Conference of Sciences of Electronics, Technologies of Information and Telecommunications*, Sousse, Tunisia, 2012.

[92] S. Al-Maadeed, W. Ayoubi, A. Hassaine, A. Almejali, A. Al-yazeedi and R. Al-atiya, "Arabic Signature Verification Dataset," in *The 13th International Arab Conference on Information Technology*, Amman, Jordan, 2012.

[93] A. Oueld Bahamou and B. Bouchareb, *Un systeme biométrique Multimodal,* Adrar, Algeria: University of Adrar, 2013.

[94] C. Sanderson., Biometric Person Recognition: Face, Speech and Fusion., VDM-Verlag, 2008.

[95] Y. Yilong, L. Lili and S. Xiwei, "SDUMLA-HMT: A Multimodal Biometric Database," in *The 6th Chinese Conference on Biometric Recognition*, Beijing, China, 2011.

[96] R. Beveridge, D. Bolme, M. Teixeira and B. Draper, "The CSU face identification evaluation system user's guide: version 5," May 2003.

[97] Y. Elmir, Z. Elberrichi and R. Adjoudj, "Multimodal Biometric Using a Hierarchical Fusion of a Person's Face, Voice, and Online Signature," *Journal of Information Processing Systems,* vol. 10, no. 4, p. 555~567, December 2014.

[98] Y. Elmir, S. Al-Maadeed, A. Amira and A. Hassaine, "A Multi-modal Face and Signature Biometric Authentication System Using a Max-of-Scores Based Fusion," in *19th International Conference on Neural Information Processing*, Doha, Qatar, 2012.

[99] Y. Elmir, S. Al-Maadeed, A. Amira and A. Hassaine, "Multi-modal biometric authentication system using face and online signature fusion," in *Qatar Foundation Annual Research Forum*, Doha, QATAR, 2012.

[100] Y. Elmir, Z. Elberrichi and R. Adjoudj, "A Hierarchical Fusion Strategy based Multimodal Biometric System," in *The International Arab Conference on Information Technology*, Khartoum, Soudan, 2013.

[101] R. Adjoudj, Authentification *Automatique par Identification & Reconnaissance dans un Système de Haute Sécurité,* Sidi Bel Abbes: University of Djilalli Liabès, 2006.

[102] S. S. Iyengar, L. Prasad and H. Min, Advances in distributed sensor technology, Prentice-Hall, 1995.

[103] X. Liu and T. Chen, "Geometry-assisted statistical modeling for face mosaicing.," in *International Conference on Image Processing*, 2003.

[104] A. Ross, S. Shah and J. Shah, "Image versus feature mosaicing: A case study in fingerprints.," in *Defense and Security Symposium.*, 2006.

[105] A. A. Ross and R. Govindarajan, "Feature Level Fusion Using Hand and Face Biometrics," in Proceedings *of SPIE Conference on Biometric Technology for Human Identification II*, Orlando, USA, March 2005.

[106] C. Sanderson and B. C. Lovell, "Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference," *Lecture Notes in Compuer Science,* vol. 5558, pp. 199-208, 2009.

[107] L. Allano, A. C. Morris, H. Sellahewa, S. Garcia-Salicetti, J. Koreman, S. Jassim and B. ... & Dorizzi, "Nonintrusive multibiometrics on a mobile device:

a comparison of fusion techniques.," in *Defense and Security Symposium.*, 2006.

[108] L. Fang-Jun and L. Lan, "Fusing multi-biometrics authorization with PCA.," in *4th IEEE International Conference on Biomedical Engineering and Informatics*, 2011.