International Journal of Advanced Science and Technology
Vol. 48, November, 2012

Multi-Modal Biometric Authentications: Concept Issues and Applications Strategies

Mohamed Soltane and Mimen Bakhti

Electrical Engineering & Computing Department, Faculty of Sciences & Technology Doctor Yahia Fares University of MEDEA, 26000 MEDEA, Algeria

soltane.mohamed@univ-medea.dz, *mimen.bakhti@univ-medea.dz*

Abstract

As the information age matures, biometric identification technology will be at the heart of computer interaction with humans and the biosphere in which they reside. Automated biometric systems for human identification measure a "signature" of the human body, compare the resulting characteristic to a database, and render an application dependent decision. These biometric systems for personal authentication and identification are based upon physiological or behavioral features which are typically distinctive, although time varying, such as fingerprints, hand geometry, face, voice, lip movement, gait, and iris patterns. Multi-biometric systems, which consolidate information from multiple biometric sources, are gaining popularity because they are able to overcome limitations such as non-universality, noisy sensor data, large intra-user variations and susceptibility to spoof attacks that are commonly encountered in uni-biometric systems. In this paper, it addresses the concept issues and the applications strategies of multi-biometric systems.

Keywords: *Biometrics, Multi-Modal, Authentication, Recognition, Identification, Verification, E-Learning, human-computer Interaction, Biometrics Fusion*

1. Introduction

A biometric is a physical or biological feature or attribute that can be measured [68]. Biometrics authentication (BA) (*Am I whom I claim I am?*) involves confirming or denying a person's *claimed identity* based on his/her physiological or behavioral characteristics [1]. BA is becoming an important alternative to traditional authentication methods such as keys ("something one has", i.e., by possession) or PIN numbers ("something one knows", i.e., by knowledge) because it is essentially "who one is", i.e., by biometric information. Therefore, it is not susceptible to misplacement or forgetfulness [22].

Biometrics is a Greek composite word stemming from the synthesis of bio and metric, meaning life measurement. In this context, the science of biometrics is concerned with the accurate measurement of unique biological characteristics of an individual in order to securely identify them to a computer or other electronic system. Biological characteristics measured usually include fingerprints, voice patterns, retinal and iris scans, face patterns, and even the chemical composition of an individual's DNA [4].

Authentication is the process of an individual claiming to have a certain identity, and then biometrically validating the users' identity is what they claim it to be. This is a one-to-one search process, a direct comparison between the claimants' features and the data regarding their features, with a certain error margin to allow for minor temporal factors or sensor discrepancies [3].

2. Biometrics Historical Background

The use of Biometrics is often regarded as a groundbreaking concept, coming straight out of modern science fiction literature. Nevertheless, there are numerous historical events that prove that the idea of using physical or behavioral characteristics for identification existed in ancient civilizations as well.

One of the oldest and most basic examples of a characteristic that is used for recognition by humans is the face. Since the beginning of civilisation, humans have used faces to identify known (familiar) and unknown (unfamiliar) individuals. This simple task became increasingly more challenging as populations increased and as more convenient methods of travel introduced many new individuals into once small communities. The concept of human-to-human recognition is also seen in behavioral predominant biometrics such as speaker and gait recognition. Individuals use these characteristics, somewhat unconsciously, to recognise known individuals on a day-to-day basis [13].

The first recorded historic incident is reported to have taken place in ancient Egypt, during the construction of the great pyramid of Khufu. Faced with a huge logistical challenge, the administrators in charge of providing food supplies to the workforce devised a system, by which every worker in a unit was assigned to go to the food warehouse once a month to receive his food allowance for that month. The administrators kept records of every worker’s name, age, place of origin, work unit, occupation and the last date on which the worker received his allowance. The collected data was used for verification of identity, when a worker appeared in the food warehouse to claim his allowance. As violations were discovered (some workers claimed multiple/false identities), the administrators decided to include physical and behavioral characteristics on the record as well [14, 13].

Another interesting technique was first used in the Babylonian era, where hand imprints were used to “prove the authenticity of certain engravings and works” [14, 13], a concept that was revisited in 1823 by the Czech Jan Evangelista Purkinje, who noticed that unique patterns were formed by sweat excreted on a person’s hand [14, 13]. This concept was further refined in 1888, by Juan Vucetich, an Argentinean police officer, who was the first to take fingerprints on ink as an identification method. 1893, Sir Francis Galton finally demonstrated that no two fingerprints are alike, even in cases of identical twins [15, 13].

True biometric systems began to emerge in the latter half of the twentieth century, coinciding with the emergence of computer systems. The nascent field experienced an explosion of activity in the 1999s and began to surface in everyday applications in the early 2000s [13].

3. Biometric Systems

A biometric system is essentially a pattern-recognition system that recognizes a person based on a feature vector derived from a specific physiological or behavioral characteristic that the person possesses [62, 66]. Biometric characteristics can be roughly broken up into two categories; Invasive and Non-Invasive. The following characteristics are just a subset of those in use and development.

3.1 Invasive Biometric Characteristics

Invasive biometric identification requires the subject to perform an action in order to be identified. Usually used for authentication.

Iris Recognition: The visual texture of the human iris is determined by the chaotic morphogenetic processes during embryonic development and is supposed to be unique for each person and each eye. An iris image is typically acquired using a non-contact imaging process:

capturing an iris image involves cooperation from the user, both to register the image of iris in the central imaging area and to ensure that the iris is at a predetermined distance from the focal plane of the camera. The iris identification technology is believed to be extremely accurate and fast [6]. Iris recognition was originally thought up on 28th August 1986 by Leonard Flom and Aran Safir who patented the process in a broad sense, hampering any further development by the wider field, until 1993 when they allow John Daug-man to develop an algorithm for use, which was patented in 1994. Primarily used for authentication, Iris recognition involves a high resolution picture being taken of a sub-ject's eye and comparing it to a data set. The blood vessels in an Iris have complete uniqueness across the population, as they are determined randomly during gestation. However, as with many biometric techniques, there is no way for the system to verify the vitality of the component [3].

Fingerprint Recognition: Fingerprints are probably the more extensively studied biometric. Uniqueness, permanence, easy acquisition and the small size of the acquisition devices (at least the electronic ones) make fingerprints one of the most popular person identification methods [59]. Fingerprint recognition has been around since 1880, however, digital fingerprint identification systems have only emerged much more recently. The epidermis (skin) of fingers and palms are unique to an individual. The most common use of this technology is for authentication, the hand is placed upon a sensory pad, which reads the ridges of the epidermis for use as the features. Fingerprinting can also be used non-invasively, removing a fingerprint from an object and scanning into a computer. In a recent development, the vitality of the appendage pressed onto the sensor can be determined by detecting the perspiration between the ridges, this helps combat faked latex fingerprints and the use of high resolution printed images of fingers, which still affects some cheaper systems [3].

Hand Geometry Recognition: Hand geometry biometric systems are becoming very popular for verification purposes. Although hand geometry is not as unique as other biometrics (e.g., fingerprints), it is permanent and has not been related for criminal prosecution; therefore it is an acceptable method for verification for the great public. In person identification systems hand geometry has been used mostly as a complement to fingerprints. Hand geometry bio-metrics fall into two main categories: geometric measurements and contour description. The automatic extraction of geometric measurements from a hand geometry image is a rather dif-ficult error pruned task. The method is more appropriate in a semi-automatic environment where a human user indicates the prominent points in the hand contour. Contour description methods have in general lower accuracy but they are more robust in automatic authentication processes [59].

Palm Print Recognition: The palms of the human hands contain pattern of ridges and valleys much like the fingerprints. The area of the palm is much larger than the area of a finger and, as a result, palm prints are expected to be even more distinctive than the fingerprints. Since palm print scanners need to capture a large area, they are bulkier and more expensive than the fingerprint sensors. Human palms also contain additional distinctive features such as principal lines and wrinkles that can be captured even with a lower resolution scanner, which would be cheaper [65]. Finally, when using a high-resolution palm print scanner, all the features of the palm such as hand geometry, ridge and valley features (e.g., minutiae and singular points such as deltas), principal lines, and wrinkles may be combined to build a highly accurate biometric system [66].

DNA “fingerprinting”: DNA (deoxyribonucleic acid) is the well-known double helix structure present in every human cell. A DNA sample is used to produce either a DNA fingerprint or a DNA profile [63]. The molecular structure of DNA can be imagined as a zipper with each tooth represented by one of the letters: A (Adeline), C (Cytosine), G (Guanine), T (Thymine) and with opposite teeth forming one of two pairs, either A-T or G-C. The information in DNA is determined by the sequence of letters along the zipper [60]. This method takes advantage of the different biological pattern of the DNA molecule between individuals. Unique differences in the banding pattern of the DNA fragments occur. DNA prints were first used in 1983 in United Kingdom [61]. DNA fingerprinting is unpopular for authentication; it is only commonly used to compare two samples to check if they are from the same person. The reason it is not more widespread is many see it as a violation of their privacy. It is also computationally complex and thus time intensive to perform. DNA is unique among the majority of the population; however, DNA is not always unique between monozygotic twins [3].

Signature: The way a person signs her/his name is known to be a characteristic of that individual, Signatures are a *behavioral biometric*, evolve over a period of time and are influenced by physical and emotional conditions of the signatories: this makes signature recognition a very challenging biometric recognition problem [6]. Signature is not new as it has long been the means by which we validate all our legal documents. However, absolute validation of signatures is a different matter, and is much more difficult. Some systems use pens with motion-sensing and pressure-sensing devices inside. In this case, a special pen is used that contains a bi-axial accelerometer to measure changes in force in the x and y direction. A force sensor measures the variations in downward (z-axis) force. A person enrolls into the system by signing his or her name a number of times. The computer reads and analyzes the dynamic motions produced by the signer during each signature. Software senses the pen’s movements and extracts significant templates. These may include signing speed, sharpness of loops, and changes in pressure. These templates form a profile that is compared to a profile stored on the user’s card or in a central database. A good match validates the user [64].

3.2 Non-Invasive Biometric Characteristics

Non-invasive biometric identification does not require that the subject be aware they are being identified; some of these techniques can even be applied at a later date. However, any of these techniques can be applied in a more direct form.

Gait Recognition: Gait recognition owes its origins to Shakespeare “Highest Queen of state, Great Juno comes; I know her by her gait” [3]. Gait is the peculiar way one walks and is a complex spatio-temporal biometric. Gait is not supposed to be very distinctive, but is sufficiently discriminatory to allow verification in some low-security applications. Gait is a behavioral biometric and may not remain invariant, especially over a long period of time, due to fluctuations in body weight, major injuries involving joints or brain, or due to inebriety. Acquisition of gait is similar to acquiring a facial picture and, hence, may be an acceptable biometric. Since gait-based systems use the video-sequence footage of a walking person to measure several different movements of each articulate joint, it is input intensive and computationally expensive [66].

Ear Recognition: Ears have been shown to be one of the most unique physiological features on the human body [3]. What makes them so useful for biometrics is that they retain their shape throughout life and are fairly static. However, ear biometrics can suffer heavily from occlusion of the ears by hair, jewelers etc. There are currently no ear recognition systems in use. However, it is a topic of extensive current research.

Face Recognition: Face is one of the most acceptable biometrics, because it is one of the most common methods of identification that humans use in their visual interactions and acquisition of faces is non-intrusive. Face recognition is of concern in un-attended authentication applications [6]. Face recognition systems detect patterns, shapes, and shadows in the face, perform feature extraction and recognition of facial identity. In the broader view, it include all types of facial processing such as tracking, detection, analysis and synthesis. The most popular approach is based on Eigen faces, that represent the differences between the face under recognition and the enrolled ones in the database. The principle component analysis using higher-order statistics is the underlying mathematics for this facial pattern recognition. Many biometric systems are confused when identifying the same person smiling, aged, with various accessories (moustache, glasses), and/or in badly lit conditions. For robustness of recognition, advanced techniques such as morphable models and expression-invariant face representation methods. On the other hand, facial recognition tools can be improved by training on a set of synthetic facial expressions and appearance/environment variations generated from real facial images [17].

Voice Recognition: Voice is a combination of physiological and behavioral biometrics [66]. Voice is the natural means of communication for human beings thus making it the most convenient to use biometric. In addition, voice needs inexpensive equipment for capturing and can be deployed in a variety of telephone-based or internet-based applications where other biometrics are impossible to be deployed [59]. Voice recognition has already been used to replace number entry on certain Sprint systems. This kind of voice recognition is related to (yet different from) speech recognition. While speech recognition technology interprets what the speaker says, speaker recognition technology verifies the speaker's identity [7]. Voice authentication or speaker recognition uses a microphone to record the voice of a person. The recorded voice is digitised and then used for authentication. The speech can be acquired from the user enunciating a known text (text dependent) or speaking (text independent) [35]. A text-dependent voice recognition system is based on the utterance of a fixed predetermined phrase. A text-independent voice recognition system recognizes the speaker independent of what she speaks. A text-independent system is more difficult to design than a text-dependent system but offers more protection against fraud [66]. The captured speech is then enhanced and unique features extracted to form a voice template. There are two types of templates: stochastic templates and model templates. Stochastic templates require probabilistic matching techniques such as the popular Hidden Markov Model and results in a measure of likelihood of the observation given the template. For model templates, the matching techniques used are deterministic. The observation is assumed to be similar to the model, albeit some distortion. Matching result is obtained by measuring the minimum error distance when the observation is aligned to the model. The matching techniques popularly used for model templates include Dynamic Time Warping algorithm, Vector Quantisation and Nearest Neighbors algorithm [35].

Odor Recognition: It is known that each object exudes an odor that is characteristic of its chemical composition and this could be used for distinguishing various objects. A whiff of air surrounding an object is blown over an array of chemical sensors, each sensitive to a certain group of (aromatic) compounds. A component of the odor emitted by a human (or any animal) body is distinctive to a particular individual. It is not clear if the invariance in the body odor could be detected despite deodorant smells, and varying chemical composition of the surrounding environment [66]. Odor recognition is a recent development in the bio-metric field. Human odors are apparently unique, and the nasal sensory equipment is very sensitive to minor changes in scent. However, the technology required to develop a system which can mimic the human nose and identify by smell is still a long way off [3].

Figure 1 shows most of the physiological/anatomical and behavioral characteristics that are being used for biometric recognition. Examples of anatomical traits include face, fingerprint, iris, palmprint, hand geometry and ear shape. Gait, signature and keystroke dynamics are some of the behavioral characteristics that can be used for person authentication. Voice can be considered either as an anatomical or as a behavioral trait because certain characteristics of a person's voice such as pitch, bass/tenor and nasality are due to physical factors like vocal tract shape, and other characteristics such as word or phoneme pronunciation (e.g., dialect), use of characteristic words or phrases and conversational styles are mostly learned.

Figure 1. Characteristics (Modalities) that are being Used for Biometric Recognition: Anatomical include iris, retina, face, fingerprint, palm print, hand geometry and ear Shape, while keystroke dynamics, signature, gait and lip motion are some of behavioral characteristics. Voice can be considered either as an anatomical or as behavioral characteristic and DNA as a Biological Chemical Characteristic.

4. Comparison of Commonly Used Biometrics

There are at least ten biometric techniques commercially available and new techniques are in the stage of research and development. What conditions must be fulfilled for a biological measurement to become a biometric? Any human physiological or behavioral characteristics can become a biometric provided the following properties are fulfilled (extended version of [9]). In Yun's paper "The '123' of Biometric Technology" [35] a relative comparison between performance from different biometric technologies is evaluated (See table 1.). This table shows each technique's aptitude for the following criteria:

Universality: This means that every person should have the characteristics. It is really difficult to get 100% coverage. There are mute people, people without fingers or with injured eyes. All these cases must be handled.

Uniqueness: This means that no two persons should be the same in terms of the biometric characteristics. Fingerprints have a high discrimination rate and the probability of two persons with the same iris is estimated as low as 1 : 1052. Identical twins, on the other side, cannot be easily distinguished by face recognition and DNA-analysis systems.

Permanence: This means that the characteristics should be invariant with time. While the iris usually remains stable over decades, a person's face changes significantly with time. The signature and its dynamics may change as well and the finger is a frequent subject to injuries.

Collectability: This means that the characteristics must be measured quantitatively and obtaining the characteristics should be easy. Face recognition systems are not intrusive and obtaining of a face image is easy. In the contrast the DNA analysis requires a blood or other bodily sample. The retina scan is rather intrusive as well.

Performance: This refers to the achievable identification/verification accuracy and the resources and working or environmental conditions needed to achieve an acceptable accuracy. The crossover accuracy of iris-based systems is under 1% and the system is able to compare over 4 10^6 iris codes in one second. The crossover accuracy of some signature dynamics systems is as high as 25% and the verification decision takes over one second.

Acceptability: This indicates to what extend people are willing to accept the biometric system. Face recognition systems are personally not intrusive, but there are countries where taking pictures of persons is not viable. The retina scanner requires an infrared laser beam directed through the cornea of the eye. This is rather invasive and only few users accept this technology.

Circumvention: This refers to how difficult it is to fool the system by fraudulent techniques. An automated access control system that can be easily fooled with a fingerprint model or a picture of a user's face does not provide much security.

Table 1. Comparison of Various Biometric Technologies [8, 9]

Biometrics	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand Geometry	M	M	M	H	M	M	M
Keystroke Dynamics	L	L	L	M	L	M	M
Hand vein	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retina	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Facial Thermogram	H	H	L	H	M	H	H
DNA	H	H	H	L	H	L	L

H–High, M–Medium, L–Low

5. Biometric Functionalities

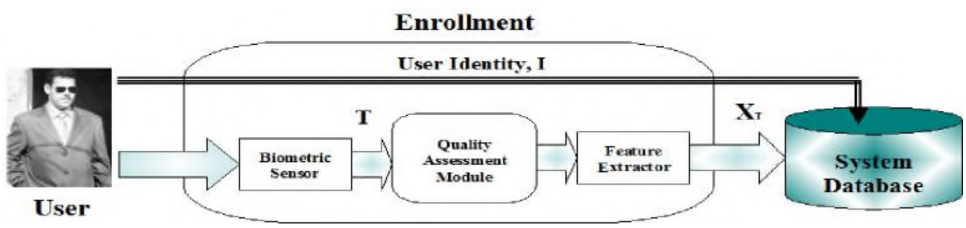
The functionalities provided by a biometric system can be categorised as *verification* and *identification.* Figure 2 shows the enrollment and authentication stages of a bio-metric system operating in the verification and identification modes.

Enrollment: A user is added to the biometric system. A certain number of biometric presentation of a particular user are acquired, preprocessed, transformed into features, and post processed, then used to train a user model and adapt (retrain) the world model if necessary. The user model along with impostor presentations may be used to obtain a threshold for that user. The new model is then stored, along with the threshold for that user if needed [10].

Verification (one to one matching): The claim to a user's identity causes the presented biometric data to be compared against the claimed user's model. Thus, the biometric data is acquired, preprocessed, transformed into features, and post processed, before being matched with the claimed user's model and the resulting score being compared with the stored threshold computed for the claimed user or a generic threshold value [10].

Identification (one to many matching): A database of user models is searched for the most likely source of the biometric presentation. Thus, the biometric data is acquired, preprocessed, transformed into features, and post processed, before being matched with all the user models of interest. The user model that obtains the highest score with respect to the presentation is suggested to be the source of the presentation [10].

Watch List: In the watch list task, the biometric system determines if the individual's biometric identifier matches a biometric identifier of someone on the watch list [58]. The biometric system's output is the list of the individuals that matches the unknown individual's biometric identifier sorted by the similarity score in ascending order. If the individual to be found is present on the watch list, and the individual with the highest similarity score returned by the biometric system is the correct individual, then we say that the biometric system has correctly detected and identified the person. The probability of making correct detect and identify decision is called correct detect and identify rate. If the individual to be found is present on the watch list and the resulting list does not include the correct person, or the individual with the highest similarity score returned by the biometric system is not the correct individual, then we say that the biometric system has made a false alarm error. The probability of making false alarm is called false alarm rate [58]. Figure 3 below consider watch list case, when an image (a probe) of a woman is an input to a face recognition system [58].



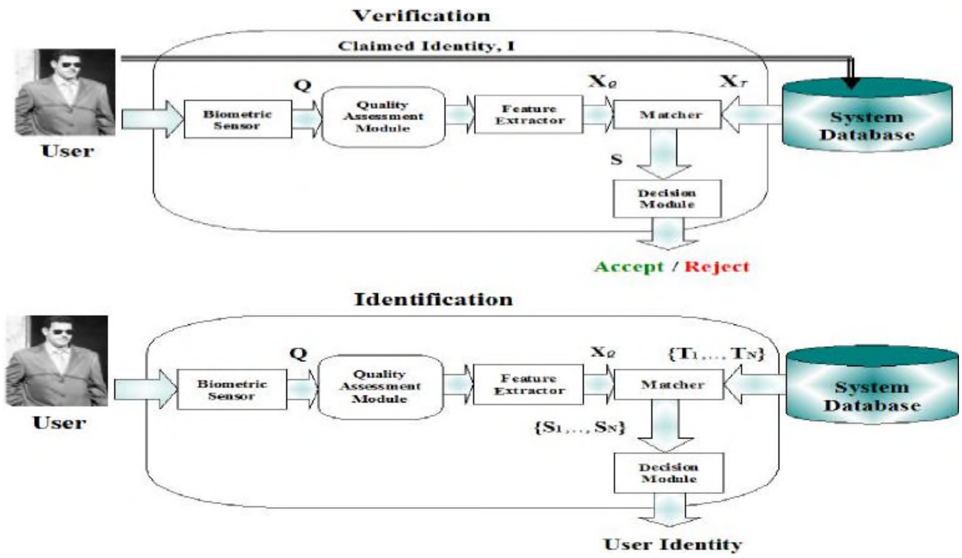


Figure 2. Information Flow in Biometric Systems. Here, T represents the biometric sample obtained during enrollment, Q is the query biometric sample obtained during recognition, X_T and X_Q are the template and query feature sets, respectively and S represents the match score and N is the number of users enrolled in the database.

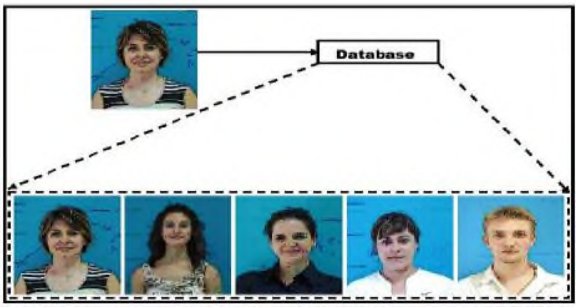


Figure 3. Watch list Example

6. Performance of Biometric Systems

Biometric systems are often evaluated solely on the basis of recognition system performance. But it is important to note that other factors are involved in the deployment of a biometric system. One factor is the quality and ruggedness of the sensors used. Clearly the quality of the sensors used will affect the performances of the associated recognition algorithms. What should be evaluated is therefore the sensor/algorithm combination, but this is difficult because often the same sensors are not used in both the enrolment and test phases. In practice therefore the evaluation is made on the basis of the recognition algorithm's resistance to the use of various types of sensor (interoperability problem). Another key factor in determining the acceptability of a biometric solution is the quality of the associated communication interface. In addition to ease of use, acquisition speed and processing speed are key factors, which are in many cases not evaluated in practice.

In the case of a verification system, two error rates are evaluated which vary in opposite directions: the **false rejection rate FRR** (rejection of a legitimate user called “the client”) and the **false acceptance rate FAR** (acceptance of an impostor). In Figure 4. are drawn the distributions of clients and impostors according to the response of the system which in general is a real number (likelihood) [11, 12].

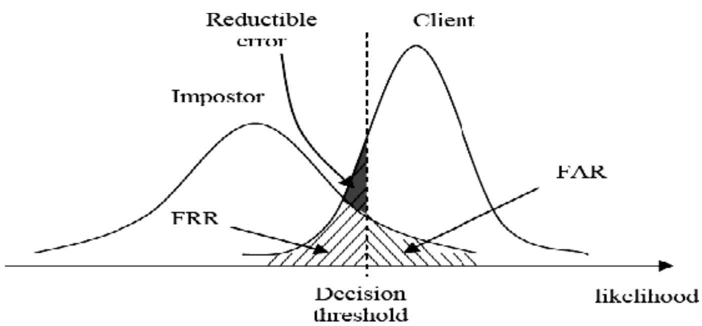


Figure 4. False Rejection Rate and False Acceptance Rate of a Biometric Verification System

The decision of acceptance or rejection of a person is thus taken by comparing the answer of the system to a threshold (called the decision threshold). The values of FAR and FRR are thus dependent on this threshold which can be chosen so as to reduce the global error of the system [12].

The decision threshold must be adjusted according to the desired characteristics for the application considered. High security applications require a low FAR which has the effect of increasing the FRR, while Low security applications are less demanding in terms of FAR, Figure 4. EER denotes Equal Error Rate (FAR=FRR). This threshold must be calculated afresh for each application, to adapt it to the specific population concerned [12]. This is done in general using a small database recorded for this purpose.

An operational biometric system makes a trade-off between false match rate (FMR) and false non-match rate (FNMR). In fact, both FMR and FNMR are functions of the system threshold *t:* If the system's designers decrease *t* to make the system more tolerant to input variations and noise, FMR increases. On the other hand, if they raise *t* to make the system more secure, then FNMR increases accordingly. We can depict system performance at all operating points (thresholds *t*) in the form of a *receiver operating characteristic* (ROC) curve. An ROC curve plots FMR against (1 – FNMR) or FNMR for various values of threshold *t* (see Figure 5) [62].

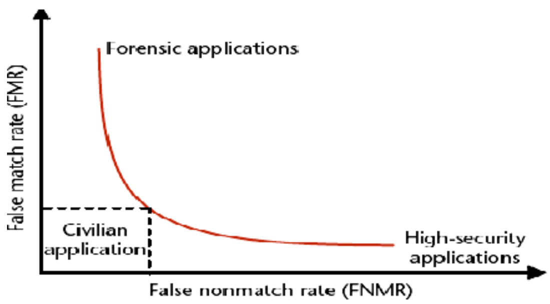


Figure 5. Receiver Operating Characteristic (ROC) Curve: Different biometric application types make different trade-offs between the false match rate and false non-match

rate (FMR and FNMR). Lack of understanding of the error rates is a primary source of confusion in assessing system accuracy in vendor and user communities alike [62].

Performance capabilities have been traditionally shown in the form of ROC (receiver- or relative-operating characteristic) plots, in which the probability of a false-acceptance is plotted versus the probability of a false-rejection for varying decision thresholds. An example of an ROC plot is given in Figure 6(a), where the desired area is at the lower left of the plot, in which both types of errors are minimized. Unfortunately, with ROC plots, curves corresponding to well-performing systems tend to bunch together near the lower left corner, impeding a clear visualization of competitive systems [16].

More recently, a variant of an ROC plot, the detection error tradeoff (DET) plot has been used, which plots the same tradeoff using a normal deviate scale. This has the effect of moving the curves away from the lower left corner when performance is high and producing linear curves, making system comparisons easier. In Figure 6(b), the DET plot corresponding to the same data in the ROC plot in Figure 6(a) is shown [16].

Although the complete DET curve is needed to fully describe system error tradeoffs, it is desirable to report performance using a single number. Often the equal-error-rate (EER), the point on the DET curve where the FA rate and FR rate are equal, is used as this single summary number. However, the suitability of any system or techniques for an application must be determined by taking into account the various costs and impacts of the errors and other factors such as implementations and lifetime support costs and end-user acceptance issues [16].

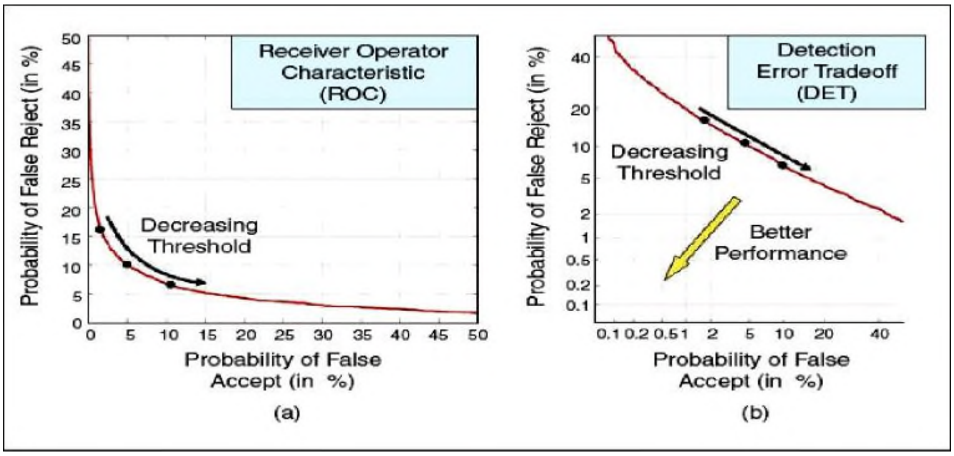


Figure 6. Example of Verification Performance Comparison for Same Hypothetical Systems, A and B, for both (a) ROC and (b) DET plots [16]

There is a tradeoff between the probability of correct detect and identify rate and the false alarm rate. If we increase the probability of correct detect and identify rate, the false alarm rate will increase [58]. A Watch list Receiver Operating Characteristic curve is used to show the relationship between the probability of correct detect and identify rate and the false alarm rate. Figure 7. [58] below is an example of the Watch list ROC curve. Selection of a watch list threshold will depend on what is trying to be accomplished. In practice, most applications that operate in the watch list task can be grouped into five operational areas:

a) **Extremely low false alarm:** In this application, any alarm requires immediate action. This could lead to public disturbance and confusion. An alarm and subsequent action may give away the fact that surveillance is being performed and how, and may minimize the possibility of catching a future suspect [58].

b) **Extremely high probability of detect and identify**: In this application, we are mostly concerned with detecting someone on the watch list; false alarms are a secondary concern and will be dealt with according to pre-defined procedures [58]

c) **Low false alarm and detect/identify:** In this application we are more concerned with lower false alarms and can deal with low detect/identify [58].

d) **High false alarm and detect/identify:** In this application we are more concerned with higher detect/identify performance and can deal with a high false alarm rate as well [58].

e) **No threshold:** User wants all results with confidence measures on each for investigation case building [58].

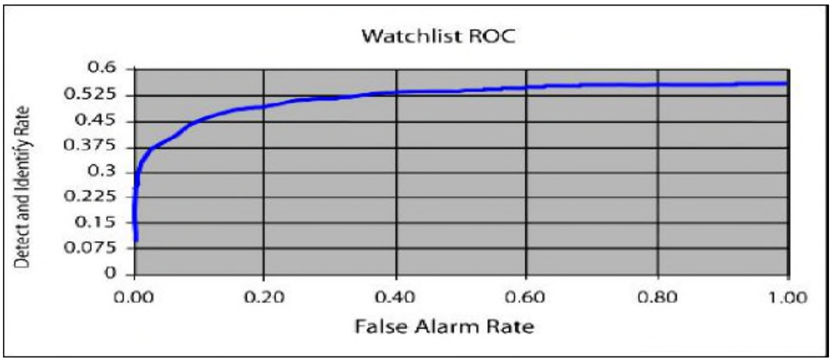


Figure 7. Example Watch List ROC [58]

7. Uni-modal Biometric Systems

Although there have been some attempts to directly compare performance of different biometric modalities (for example, see Figure 8), it is still difficult to characterize the performance of different biometric recognition systems in a consistent way. One reason for this is that there are many factors that produce degradation in recognition performance, and these factors are not homogeneous throughout biometric modalities. Factors like the type of application, enrollment and/or testing scenario, size of the population under study, controlled situations versus uncontrolled situations, etc., introduce a heterogeneous assessment framework. There are also intra-modality factors (e.g., type of acquisition device, fingertip position, or finger humidity in fingerprint matching; illumination, pose, face artifacts, or background in face recognition; transmission channel, noise, or type of handset in speaker verification; online versus offline acquisition, or degree of skill of forgeries in signature biometrics, etc.) that are almost impossible to equalize between modalities to make absolute comparisons. One observed theme in these cases is that performance tends to improve with increasing constraints on the application (more biometric samples, less distortion/noise/artifacts, well-performing acquisition devices, cooperative users, etc.). Determining acceptable performance for a particular application will depend on the benefit of replacing any current verification procedure, the threat model (claimant to impostor attempts), and the relative costs of errors [16].

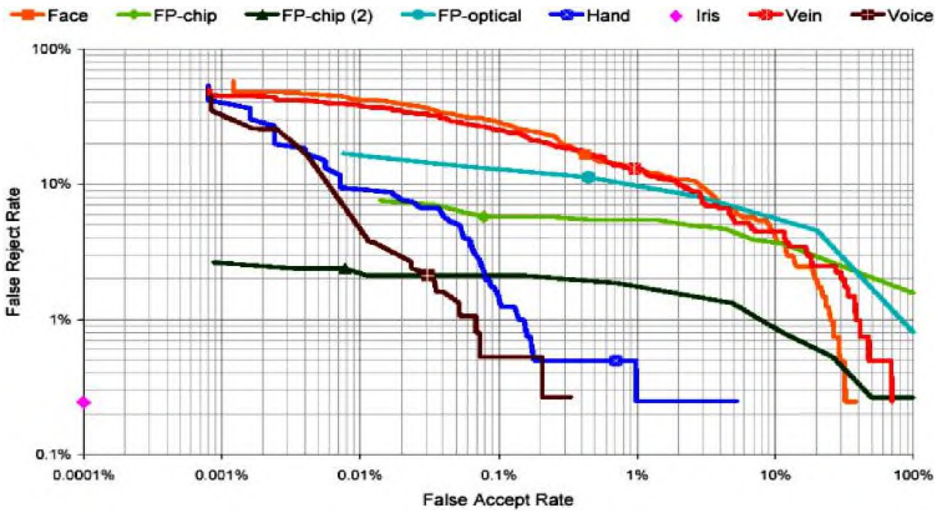


Figure 8. DET plot showing biometric performance for six different modalities, namely, face, fingerprint, hand geometry, iris, veins, and voice. Regarding fingerprints, results include performance with three different sensors, two of them being capacitive chips and the third being optical [16, 34].

8. Bi-Modal Biometric Systems

From the hypothesis that every biometric indicator has some limitations; For example, it is estimated that approximately 5% of the population does not have "legible" fingerprints. A multimodal system, which combines the decisions made by a number of independent biometrics indicators, can overcome some of these limitations. A multimodal system is generally more robust to fraudulent technologies, because it is more difficult to forge multiple biometric characteristics than to forge a single biometric characteristic [19]. In multimodal systems, complementary input modalities provide the system with non-redundant information whereas redundant input modalities allow increasing both the accuracy of the fused information by reducing overall uncertainty and the reliability of the system in case of noisy information from a single modality. Information in one modality may be used to disambiguate information in the other ones. The enhancement of precision and reliability is the potential result of integrating modalities and/or measurements sensed by multiple sensors [2].

In this section it present an overview result of recent research in the fields of speaker verification and face verification and bi-modal verification. A Fusion Systems of Bi-Modal Biometric Based Face & Voice Verification at different Levels is illustrated below (see Figure 9).

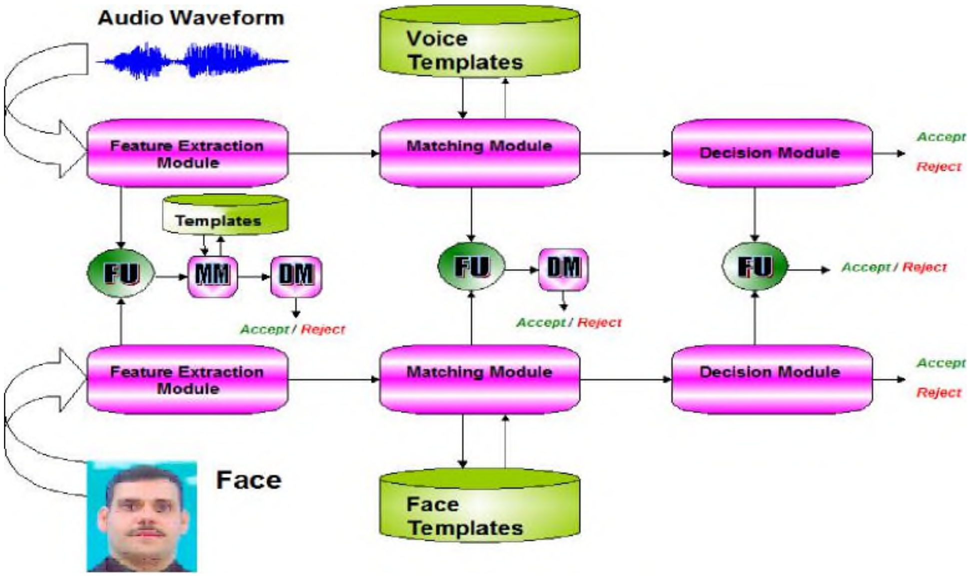


Figure 9. A Bi-Modal Biometric System Showing the Three Levels of Fusion: FU: Fusion Module, MM: Matching Module and DM: Decision Module

8.1 Databases and Protocols

8.1.1 The XM2VTS database and the Lausanne Protocols [22]

The XM2VTS database [36, 24] contains synchronised video and speech data from 295 subjects, recorded during four sessions taken at one month intervals. On each session, two recordings were made, each consisting of a speech shot and a head shot. The speech shot consisted of frontal face and speech recordings of each subject during the recital of a sentence.

The database is divided into three sets: a training set, an evaluation set and a test set. The training set (LP Train) was used to build client models, while the evaluation set (LP Eval) was used to compute the decision thresholds (as well as other hyper-parameters) used by classifiers. Finally, the test set (LP Test) was used to estimate the performance.

The 295 subjects were divided into a set of 200 clients, 25 evaluation impostors and 70 test impostors. There exists two configurations or two different partitioning approaches of the training and evaluation sets. They are called Lausanne Protocol I and II, denoted as LP1 and LP2 in this paper. In both configurations, the test set remains the same. Their difference is that there are three training shots per client for LP1 and four training shots per client for LP2. Table 2 is the summary of the data.

Note that LP Eval's of LP1 and LP2 are used to calculate the optimal thresholds that will be used in LP Test. Results are reported only for the test sets, in order to be as unbiased as possible (using an a priori selected threshold). More details can be found in [22].

Table 2. The Lausanne Protocols of XM2VTS Database. The last column shows the terms used in the fusion protocols. LP Eval corresponds to the Fusion protocols' development set while LP Test corresponds to the Fusion Protocols' evaluation set [22].

Data Sets	Lausanne Protocols		Fusion Protocols
	LP1	LP2	
LP Train client accesses	3	4	NIL
LP Eval client accesses	600 (3 x 200)	400 (2 x 200)	Fusion dev
LP Eval impostor accesses	40,000 (25 x 8 x 200)		Fusion dev
LP Test client accesses	400 (2 x 200)		Fusion eva
LP Test impostor accesses	112,000 (70 x 8 x 200)		Fusion eva

8.1.2 The BANCA Database [27]

The BANCA (biometric access control for networked and e-commerce applications) database [25, 37] was designed in order to test multi-modal identity authentication with various acquisition devices (2 cameras and 2 microphones) and under several scenarios (controlled, degraded and adverse). For 5 different languages (English, French, German, Italian and Spanish), video and speech data were collected for 52 subjects (26 males and 26 females), i.e. a total of 260 subjects. Each language - and gender - specific population was itself subdivided into 2 groups of 13 subjects (denoted g1 and g2). Each subject participated to 12 recording sessions, each of these sessions containing 2 records: 1 true *client access* (T) and 1 informed 3 *impostor attack* (I). For the image part of the database, there is 5 shots per record. The 12 sessions were separated into 3 different scenarios.

In the BANCA protocol, we consider that the true client records for the first session of each condition is reserved as training material. In all our experiments, the client model training is done on at most these 3 records. We consider the following protocols, namely Matched Controlled (Mc) and Pooled test (P) protocol, where one controlled session is used for client training and, the same controlled conditions sessions for Mc, and all conditions sessions for P, are used for client and impostor testing [27].

8.1.3 The VidTIMIT Database [18]

The VidTIMIT database consists of audio recordings and video sequences of 43 subjects (19 female and 24 male), reciting short sentences from the test section of the NTIMIT corpus [39] in three sessions with an average delay of a week between sessions, allowing for appearance and mood changes. Each person utters ten sentences. The first two sentences are the same for all subjects, while the remaining eight are generally different for each person. All sessions contain phonetically balanced sentences. In addition to the sentences, the subjects were asked to move their heads left, right, up, then down, in order to obtain head rotation sequence. The AV biometric systems that utilize the VidTIMIT corpora are described in [38].

8.2 Baseline System Description

8.2.1 Face and Speech Features [22]

The face baseline experts are based on the following features:

1. FH: normalised face image concatenated with its RGB Histogram (thus the abbreviation FH) [40].

2. DCTs: DCTmod2 features [41] extracted from face images with a size of 40 32 (rows columns) pixels. The Discrete Cosine Transform (DCT) coefficients are calculated from an 8 8 window with horizontal and vertical overlaps of 50%, i.e., 4 pixels in each direction. Neighbouring windows are used to calculate the \delta" features. The result is a set of 35 feature vectors, each having a dimensionality of 18. (s indicates the use of this small image compared to the bigger size image with the abbreviation b.)[22].

3. DCTb: Similar to DCTs except that the input face image has 80 64 pixels. The result is a set of 221 feature vectors, each having a dimensionality of 18 [22].

The speech baseline experts are based on the following features:

1. LFCC: The Linear Filter-bank Cepstral Coefficient (LFCC) [42] speech features were computed with 24 linearly-spaced filters on each frame of Fourier coefficients sampled with a window length of 20 milliseconds and each window moved at a rate of 10 milliseconds. 16 DCT coefficients are computed to decorrelate the 24 coefficients (log of power spectrum) obtained from the linear filter-bank. The first temporal derivatives are added to the feature set [22].

2. PAC: The Phase Auto-Correlation Mel Filter-bank Cepstral Coefficient (PAC-MFCC) features [43] are derived with a window length of 20 milliseconds and each window moves at a rate of 10 milliseconds. 20 DCT coefficients are computed to decorrelate the 30 coefficients obtained from the Mel-scale filter-bank. The first temporal derivatives are added to the feature set [22].

3. SSC: Spectral Subband Centroid (SSC) features, originally proposed for speech recognition [44], were used for speaker authentication in [45]. It was found that mean-subtraction could improve these features significantly. The mean-subtracted SSCs are obtained from 16 coefficients. The parameter, which is a parameter that raises the power spectrum and controls how much influence the centroid, is set to 0.7 [46, 22].

8.2.2 Baseline System Classifiers [22]

1. (FH, MLP) Features are normalised Face concatenated with Histogram features. The client dependent classifier used is an MLP with 20 hidden units. The MLP is trained with geometrically transformed images [47, 22].

2. (DCTs, GMM) The face features are the DCTmod2 features calculated from an input face image of 40 32 pixels, hence, resulting in a sequence of 35 feature vectors each having 18 dimensions. There are 64 Gaussian components in the GMM. The world model is trained using all the clients in the training set [48, 22].

3. (DCTb, GMM) Similar to (DCTs, GMM), except that the features used are DCTmod2 features calculated from an input face image of 8064 pixels. This produces in a sequence of 221 feature vectors each having 18 dimensions. The corresponding GMM has 512 Gaussian components [48, 22].

4. (DCTs, MLP) Features are the same as those in (DCTs, GMM) except that an MLP is used in place of a GMM. The MLP has 32 hidden units [48]. Note that in this case a training example consists of a big single feature vector with a dimensionality of 35 18. This is done by simply concatenating 35 feature vectors each having 18 dimensions [22].

5. (DCTb, MLP) The features are the same as those in (DCTb, GMM) except that an MLP with 128 hidden units is used. Note that in this case the MLP in trained on a single feature vector with a dimensionality of 221 18 [48, 22].

and for the speech experts:

1. (LFCC, GMM) This is the Linear Filter-bank Cepstral Coefficients (LFCC) obtained from the speech data of the XM2VTS database. The GMM has 200 Gaussian components, with the minimum relative variance of each Gaussian fixed to 0.5, and the MAP adaptation weight equals 0.1. This is the best known model currently available [49] under clean conditions [22].

2. (PAC, GMM) The same GMM configuration as in LFCC is used. Note that in general, 200-300 Gaussian components would give about 1% of difference of HTER [49]. This system is particularly robust to very noisy conditions (less than 6 dBs, as tested on the NIST2001 one-speaker detection task) [22].

3. (SSC, GMM) The same GMM configuration as in LFCC is used [46]. This system is known to provide an optimal performance under moderately noisy conditions (18-12 dBs, as tested on NIST2001 one-speaker detection task) [22].

Table 3. HTER Results of Combining Two Baseline Experts [22]

(a) Fusion with different modalities for LP1.

No.	Fusion candidates	HTER
1	((FH,MLP)(LFCC,GMM))	0.782
2	((FH,MLP)(PAC,GMM))	1.120
3	((FH,MLP)(SSC,GMM))	0.871
4	((DCTs,GMM)(LFCC,GMM))	0.543
5	((DCTs,GMM)(PAC,GMM))	1.436
6	((DCTs,GMM)(SSC,GMM))	1.149
7	((DCTb,GMM)(LFCC,GMM))	0.511
8	((DCTb,GMM)(PAC,GMM))	1.021
9	((DCTb,GMM)(SSC,GMM))	0.752
10	((DCTs,MLP)(LFCC,GMM))	0.840
11	((DCTs,MLP)(PAC,GMM))	1.138
12	((DCTs,MLP)(SSC,GMM))	1.333
13	((DCTb,MLP)(LFCC,GMM))	1.523
14	((DCTb,MLP)(PAC,GMM))	3.664
15	((DCTb,MLP)(SSC,GMM))	3.108

(b) Fusion with different modalities for LP1.

No.	Fusion candidates	HTER
1	((FH,MLP)(DCTs,GMM))	1.280
2	((FH,MLP)(DCTb,GMM))	1.122
3	((FH,MLP)(DCTs,MLP))	1.513
4	((FH,MLP)(DCTb,MLP))	1.960
5	((LFCC,GMM)(SSC,GMM))	1.595
6	((PAC,GMM)(SSC,GMM))	4.225

(c) Fusion with different classifiers for LP1

No.	Fusion candidates	HTER
1	((DCTs,GMM)(DCTs,MLP))	2.388
2	((DCTb,GMM)(DCTb,MLP))	3.063

(d) Fusion with different modalities for LP2.

No.	Fusion candidates	HTER
1	((FH,MLP)(LFCC,GMM))	1.122
2	((FH,MLP)(PAC,GMM))	1.513
3	((FH,MLP)(SSC,GMM))	1.960
4	((DCTb,GMM)(LFCC,GMM))	1.836
5	((DCTb,GMM)(PAC,GMM))	2.388
6	((DCTb,GMM)(SSC,GMM))	3.672

(e) Fusion with different feature sets for LP2

No.	Fusion candidates	HTER
1	((FH,MLP)(DCTb,GMM))	1.280
2	((LFCC,GMM)(SSC,GMM))	3.063
3	((PAC,GMM)(SSC,GMM))	2.934

Table 4. Baseline Results (in terms of HTER) on XM2VTS [24], BANCA [25] and IDIAP Databases [26, 27]

	XM2VTS (LP 1)	BANCA (Mc)	BANCA (P)	IDIAP
Face	**1.67**	**5.77**	**18.96**	**7.61**
Speech	**1.14**	**4.32**	**12.29**	**3.15**
Fusion	**0.48**	**4.32**	**9.99**	**1.49**

Table 5. Sample Audio Visual Person Recognition Systems [18]

System	Features		Database	Non-ideal Conditions	Expert	AV Fusion Method	Recognition Mode*
	Acoustic	Visual					
Chibelushi *et al.* [50]	MFCCs	shape-based (PCA,LDA, concatenation)	10 speakers [50]	white noise at different SNRs	ANNs	opinion fusion (weighted summation)	TD/ID
Brunelli and Flavigna [51]	MFCCs+ A^{**}+AA	appearance-based	89 speakers 3 sessions	none	VQ	opinion fusion (weighted product)	TI/ID
Ben-Yacoub *et al.* [52]	LPCs	appearance-based	XM2VTS	none	HMMs sphericity measure [52]	post classifier using binary classifier (SVM, Baysian classifier, FLD, decision tree and MLP)	TD+ TI/VER

Sanderson and Paliwal [38]	MFCCs+A	appearence-based (PCA)	VidTIMIT	white and operations room noise at different SNRs	GMMs	weighted summation, concatenation, adaptive weighted summation, SVM, Bayesian classifier	TI/VER
Hazen *et al.* [53]	MFCCs	appearance-based	35 speakers [53]	data recorded on a handheld device	SVMs	opinion fusion (weighted summation)	TD/ID
Fox *et al.* [54,55]	MFCCs+A	appearance-based (DCT)	XM2VTS	white noise at different SNRs	HMMs	feature-level concatenation, opinion fusion (weighted summation)	TD/ID
Nefian *et al.* [56]	MFCCs+A+AA	appearance-based (PCA+LDA)	XM2VTS	white noise at different SNRs	Coupled HMMs embedded HMMs	midst-mapping fusion, opinion fusion (weighted summation)	TD/ID
Kanak *et al.* [57]	MFCCs+A+AA	appearance-based (PCA)	38 speakers [57]	white noise at different SNRs	HMMs	concatenation, opinion fusion (Bayesian fusion)	TD/ID

* TD: text-dependent; TI: text-independent ** A- first derivative
VER: verification; ID: identification AA-second derivative

Chibelushi, et. al., [50] developed an AV biometrics system that utilizes acoustic information and static visual information contained in face profiles. They utilized an AV database that consists of audio recordings and face images of ten speakers [50]. The images are taken at different head orientations, image scales, and subject positions. They combined acoustic and visual information utilizing weighted summation fusion. Their system achieved an EER of 3.4%, 3.0%, and 1.5% when only speech information, only visual information, or both acoustic and visual information were used, respectively [18]. Brunelli and Falavigna [51] developed a text-independent speaker identification system that combines audio-only speaker identification and face recognition. They utilized an AV database that consists of audio recordings and face images of 89 speakers collected in three sessions [51]. The system provides five classifiers, two acoustic and three visual. The two acoustic classifiers correspond to two sets of acoustic features (static and dynamic) derived from the short time spectral analysis of the speech signal. Their audio-only speaker identification system is based on vector quantization (VQ). The three visual classifiers correspond to the visual classifying features extracted from three regions of the face, i.e., eyes, nose, and mouth. The individually obtained classification scores are combined using the weighted product approach. The identification rate of the integrated system is 98%, compared to the 88% and 91% rates obtained by the audioonly speaker recognition and face recognition systems, respectively [18]. Ben-Yacoub, et. al., [52] developed both text-dependent and text-independent AV speaker verification systems by utilizing acoustic information and frontal face visual information from the XM2VTS database. They utilized elastic graph matching in order to obtain face matching scores. They investigated several binary classifiers for postclassifier opinion fusion, namely, SVM, Bayesian classifier, Fisher's linear discriminant, decision tree, and multilayer perceptron (MLP). They obtained the best results utilizing SVM and Bayesian classifiers, which also outperformed single modalities [18]. Sanderson and Paliwal [38] utilized speech and face information to perform text-independent identity verification. They extracted appearance-based visual features by performing PCA on the face image window containing the eyes and the nose. The acoustic features consisted of MFCCs and their corresponding deltas and maximum autocorrelation values, which capture pitch and voicing information. A voice activity detector (VAD) was used to remove the feature vectors which represent silence or background noise, while a GMM classifier was used as a modality (speech or face) expert, to obtain opinions from the speech features. They performed elaborate analysis and evaluation of several nonadaptive and adaptive approaches for information fusion and compared them in noisy and clean audio conditions with respect to overall verification performance on the VidTIMIT database. The fusion methods they analyzed include weighted summation, Bayesian classifier, SVM, concatenation, adaptive weighted summation, and proposed piecewise linear postclassifier, and modified Bayesian postclassifier. The utilized fusion methods take into account how the distributions of opinions are likely to change due to noisy conditions, without making a direct assumption about the type of noise present in the testing features. The results are reported in terms of total error (TE), defined as TE = FAR + FRR. They concluded that the performance of most of the nonadaptive fusion systems was similar and that it degraded in noisy conditions [18]. Hazen, et. al., [53] developed a text-dependent speaker authentication system that utilizes lower quality audio and visual signals obtained by a handheld device. They detected 14 face components and used ten of them, after normalization, as visual features for a face recognition algorithm which utilizes SVMs. They achieved 90% reduction in speaker verification EER when fusing face and speaker identification information [18].

9. Multi-biometric Systems

Biometric systems that integrate information at an early stage of processing are believed to be more effective than those systems which perform integration at a later stage. Since the feature set contains richer information about the input biometric data than the matching score or the output decision of a matcher, fusion at the feature level is expected to provide better recognition results. However, fusion at this level is difficult to achieve in practice because (i) the feature sets of the various modalities may not be compatible (e.g., Eigen-coefficients of face and minutiae set of finger), and (ii) most commercial biometric systems do not provide access to the feature sets (nor the raw data) which they use in their products. Fusion at the decision level is considered to be rigid due to the availability of limited information. Thus, fusion at the match score level is usually preferred, as it is relatively easy to access and combine the scores presented by the different modalities [20].

9.1 Fusion Levels

A generic biometric system has four important modules (see Figure 10): (a) Fusion at the sensor level: the sensor module which captures the *trait* in the form of raw biometric *data*; (b) Fusion at the data or feature level: Either the data itself or the feature sets originating from multiple sensors/sources are fused. (c) Fusion at the match score level: The scores generated by multiple classifiers pertaining to different modalities are combined. (d) Fusion at the decision level: The final output of multiple classifiers are consolidated via techniques such as majority voting.

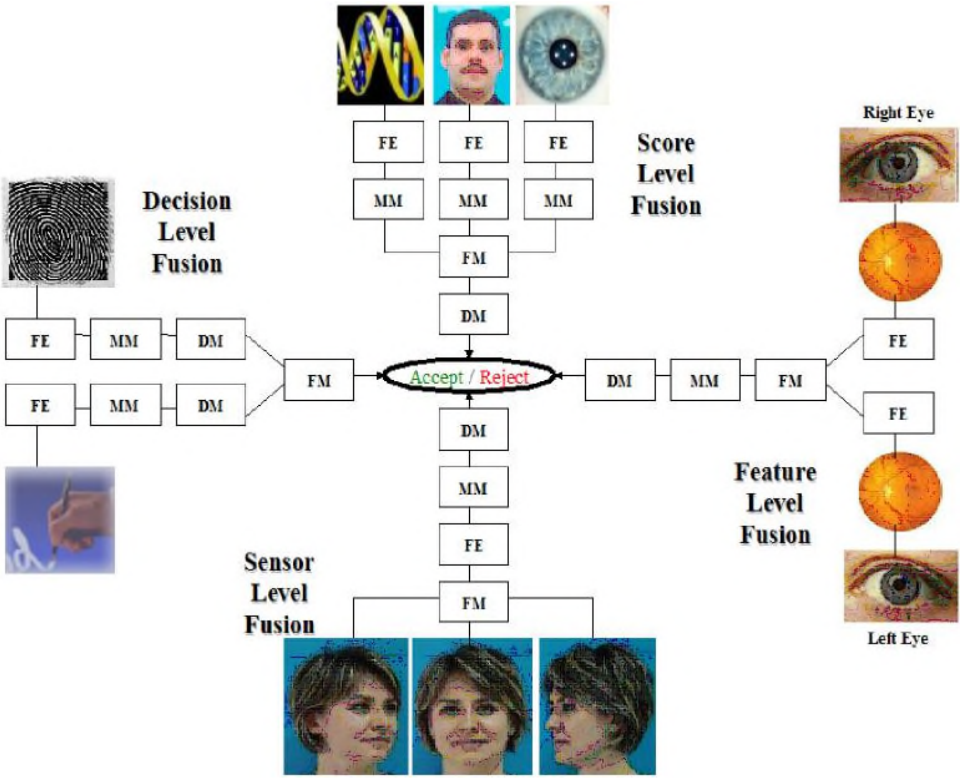


Figure 10. Various Levels Fusion in Multi-biometric Systems. Most multimodal biometric systems fuse information at the match score level or the decision level. FE: feature extraction module; MM: matching module; DM: decision-making module and FU: fusion module.

9.2 Fusion Scenarios

Depending on the number of traits, sensors, and feature sets used, a variety of scenarios are possible in a multimodal biometric system (Figure 11.): (a) Multiple sensors: A single biometric modality is acquired by using a number of sensors. One example is multiple face cameras for creating a 3D input face or for combining the output scores of the different baseline face images. (b) Multiple algorithms: A single biometric input is processed with different feature extraction algorithms in order to create templates with different information content. One example is processing fingerprint images according to minutiae and texture-based representations. (c) Multiple instances: A single biometric modality but multiple parts of the human body are used. One example is the use of multiple fingers in fingerprint verification. (d) Repeated instances: The same biometric modality and instance is acquired with the same sensor multiple times. One example is the sequential use of multiple impressions of the same finger in fingerprint verification. (e) Multiple modalities: Multiple biometric modalities are combined. This also known as multimodal biometrics.

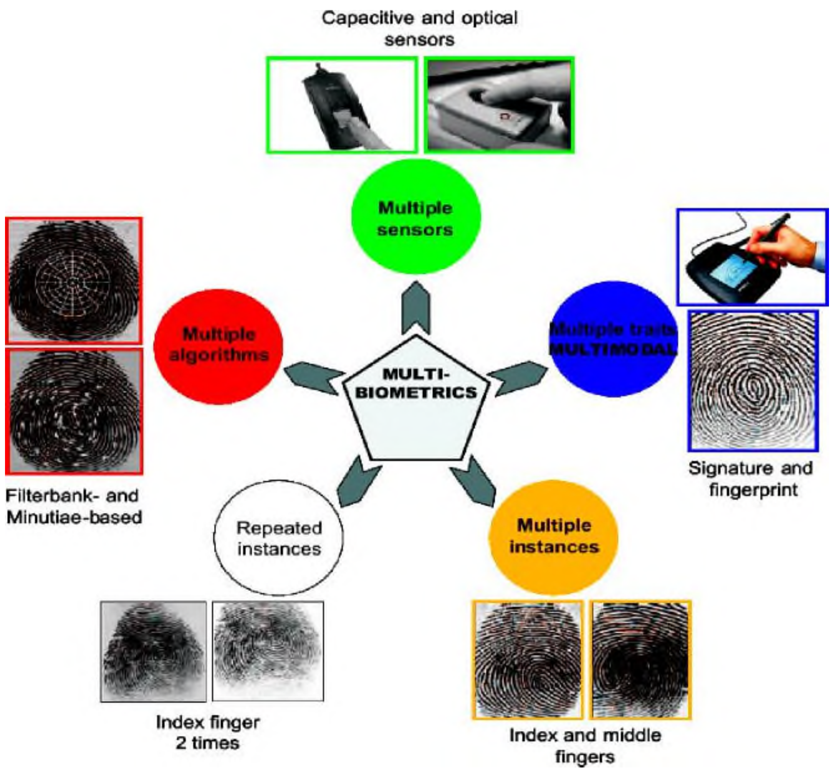


Figure 11. Scenarios in a Multimodal Biometric System [66, 70].

9.3 Score Normalisation Techniques [74]

This section, present six normalization methods, it denote a raw matching score as from the set of all scores for that matcher, and the corresponding normalized score as .

Min-Max (MM): This method maps the raw scores to the [0, 1] range. The quantities *max(S)* and *min(S)* specify the end points of the score range:

(1)

Z-score (ZS): This method transforms the scores to a distribution with mean of 0 and standard deviation of 1. The operators *mean()* and *std()*denote the arithmetic mean and standard deviation operators, respectively:

(2)

The median and median absolute deviation (MAD) [77]: are insensitive to outliers and the points in the extreme tails of the distribution. Hence, a normalization scheme using median and MAD would be robust and is given by

$$s'_k = \frac{s_k - median}{MAD},$$

(3)

where MAD=median*(|sk−median|)*. However, the median and the MAD estimators have a low efficiency compared to the mean and the standard deviation estimators, i.e., when the score distribution is not Gaussian, median and MAD are poor estimates of the location and scale parameters. Therefore, this normalization technique does not retain the input distribution and does not transform the scores into a common numerical range.

Tanh (TH): This method is among the so-called *robust* statistical techniques [75, 74]. It maps the raw scores to the (0, 1) range:

(4)

Adaptive (AD): The errors of individual biometric matchers stem from the overlap of the genuine and impostor score distributions. It characterise this overlap region by its center and its width . To decrease the effect of this overlap on the fusion algorithm, it propose to use an adaptive normalization procedure that aims to increase the separation of the genuine and impostor distributions, while still mapping the scores to [0,1] range. Previously, test normalization (T-norm) [76, 74] that can be thought of as adaptive normalization considering impostor scores is proposed. The adaptive normalization is formulated as ,

where denotes the mapping function that is applied to the MM normalized scores, . It has considered the following three choices for the function . These functions use two parameters of the overlapping region, and , which can be either provided by the vendors or estimated by the system integrator. In this work, it estimates these parameters.

Two-Quadrics (QQ) [74]: This function is composed of two quadratic segments that change the concavity at (Figure 12(a)):

$$n_{AD} = \begin{cases} \frac{1}{c} n_{MM}^2, & n_{MM} \leq c \\ c + \sqrt{(1-c)(n_{MM}-c)}, & \text{otherwise} \end{cases} \quad (5)$$

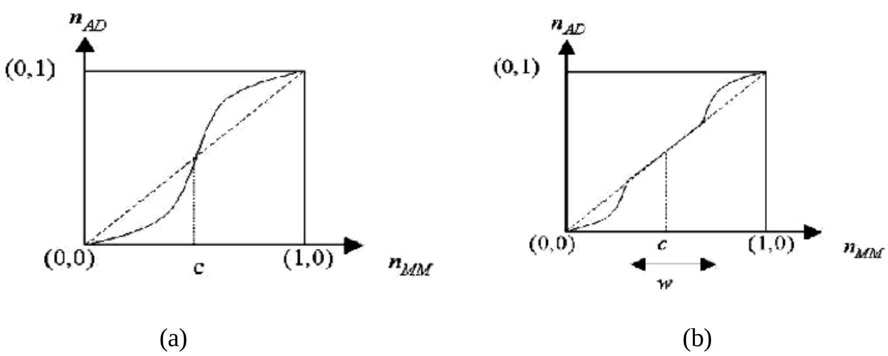


Figure 12. Mapping Functions for QQ and QLQ Adaptive Normalizations

For comparison, the identity function, , is also shown by the dashed lines in Figure 12.

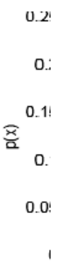
Logistic (LG) [74]: Here, takes the form of a logistic function. The general shape of the curve is similar to that shown for function QQ in Figure 12 (a). It is formulated as:

(6)

Where the constants and are calculated as and . Here, is equal to the constant , which is selected to be a small value (0.01 in this study).
Note that, due to this specification, the inflection point of the logistic function occurs at , the center of the overlap region.

Quadric-Line-Quadric (QLQ) [74]: The overlap zone, with center and width , is left unchanged while the other regions are mapped with two quadratic function segments (Figure 12(b)):

(7)



(a)

Figure 13. Conditional Distribution of Genuine and Impostor Scores: (a) Face (distance score); (b) Fingerprint (similarity score) [77]

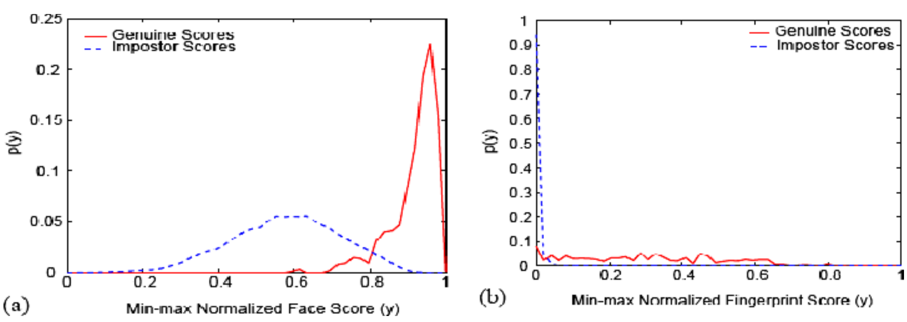


Figure 14. Distribution of Genuine and Impostor Scores after min–max Normalization: (a) Face; (b) Fingerprint [77]

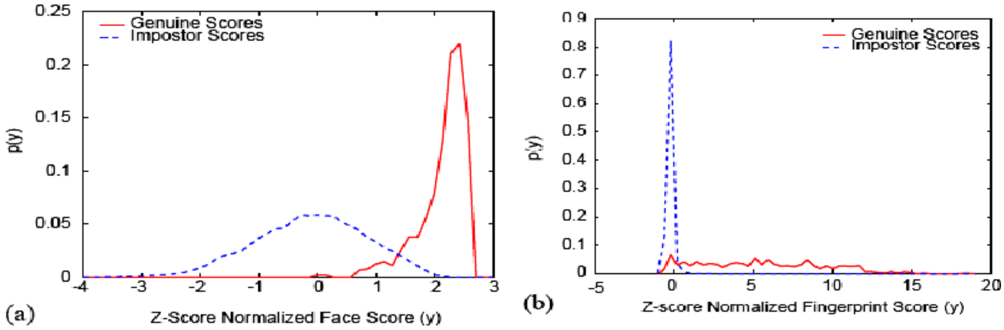


Figure 15. Distribution of Genuine and Impostor Scores after z-score Normalization: (a) Face; (b) Fingerprint [77]

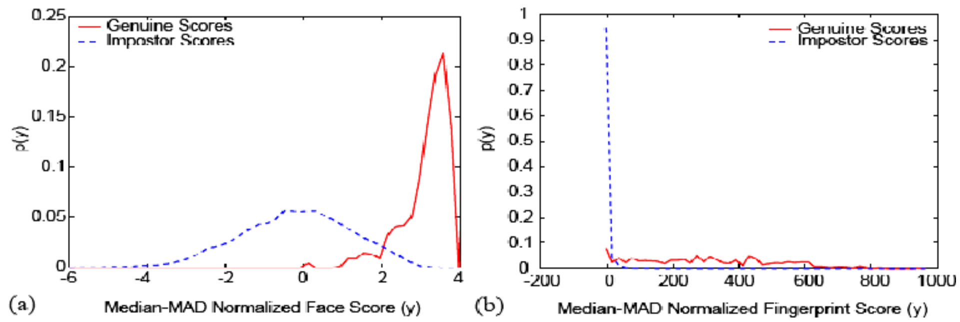


Figure 16. Distribution of Genuine and Impostor Scores after median-MAD Normalization: (a) Face; (b) Fingerprint [77]

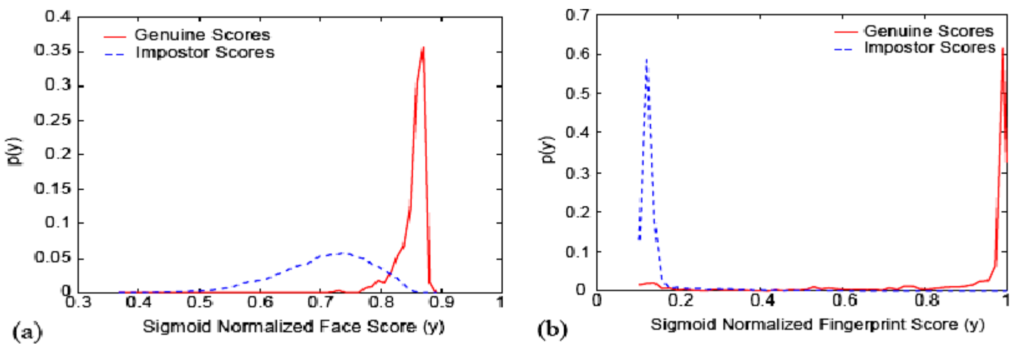


Figure 17. Distribution of Genuine and Impostor Scores after Double Sigmoid Normalization: (a) Face; (b) Fingerprint [77]

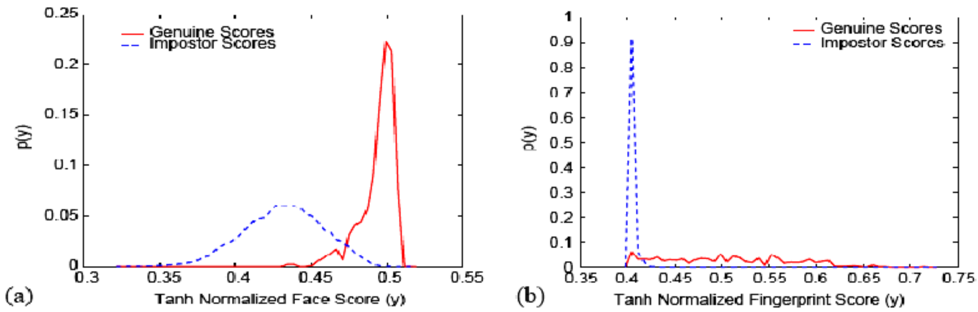


Figure 18. Distribution of Genuine and Impostor Scores after Tanh Normalization: (a) Face; (b) Fingerprint [77]

9.4 Score Fusion Methods [74]

Five different fusion methods it presented here, namely simple-sum, min-score, max-score, matcher weighting and user weighting. The quantity represents the normalized score for matcher ($m= 1, 2, \ldots, M$, where is the number of matchers) applied to user ($i=1, 2, \ldots, I$, where is the number of individuals in the database). The fused score for user is denoted as .

Simple-Sum (SS): $f_i = \sum_{m=1}^{M} n_i^m, \forall i$

Min-Score (MIS): $f_i = min(n_i^1, n_i^2, \ldots, n_i^M), \forall i$

Max-Score (MAS): $f_i = max(n_i^1, n_i^2, \ldots, n_i^M), \forall i$

Matcher Weighting (MW) [74]: Weights are assigned to the individual matchers based on their Equal Error Rates (EER's). Denote the EER of matcher *m* as , $m= 1, 2, \ldots, M$. Then, the weight associated with matcher *m* is calculated as:

$$w^m = \frac{\left(1/\sum_{m=1}^{M}\frac{1}{e^m}\right)}{e^m}. \quad (8)$$

Note: tl $0 \le w^m \le 1, \forall m, \ \sum_{m=1}^{M} w^m = 1$ and the weights are inversely proportional to the corresponding errors; the weights for *more accurate* matchers are higher than those of *less accurate* matchers. The MW fused score for user is calculated as:

(9)

User Weighting (UW) [74]: The User Weighting fusion method assigns weights to individual matchers that may be different for different users. Jain and Ross [78] proposed a similar scheme, but they exhaustively searched a coarse sampling of the weight space, where weights are multiples of 0.1 in the range [0, 1]. Their method can be prohibitively expensive if the number of fused matchers, , is high, since the weight space is ; further, coarse sampling as used in [78] may not find the optimal weight set. In our method, the UW fused score for user is calculated as:

(10)

Where represents the weight of matcher for user . The calculation of these user-dependent weights is based on the *wolf-lamb* concept introduced by Doddington et al. [79] for unimodal speech biometrics. They label the users who can be imitated easily as *lambs* (namely, impostors can provide biometric data similar to that of lambs); *wolves* on the other hand are those who can successfully imitate some other users. Lambs and wolves decrease the performance of biometric systems since they lead to false accepts. It extend these notions to multimodal biometrics by developing a metric of *lambness* for every pair of user and matcher, (,). This lambness metric is then used to calculate the weights for biometric fusion. Thus, if user is a lamb (can be imitated easily by some wolves) in the space of matcher , the weight associated with this matcher is decreased for user . The main aim is to decrease the lambness of user in the space of combined matchers. It assume that for every (,) pair, the mean and standard deviation of the associated genuine and impostor distributions are known (or can be estimated, as is done in this study). It denote the means of these distributions as and , respectively, and denote the standard deviations as and , respectively. It use the d-prime metric [80] as a measure of the separation of these two distributions in formulating the lambness metric for user and matcher as:

$$d_i^m = \frac{\mu_i^m(gen)-\mu_i^m(imp)}{\sqrt{(\sigma_i^m(gen))^2+(\sigma_i^m(imp))^2}} \tag{11}$$

If is small, user is a lamb for some wolves and if is large, is not a lamb. It structures the user weights to be proportional to this lambness metric as follows:

(12)

Note that

10. Score Level Fusion in Multi-biometric Systems

One of the challenges in combining match scores is that scores from different matchers may not be homogeneous. Consider the scores provided by the two face matchers in the NIST-Face database (described below). The scores from the first face matcher are in the range [-1; 1], whereas scores from the second face matcher are in the range [0; 100] (see Figure 19) [71, 33]. The match scores of different matchers can be either distance or similarity measures, they may follow different probability distributions and the accuracy of the matchers may be quite different.

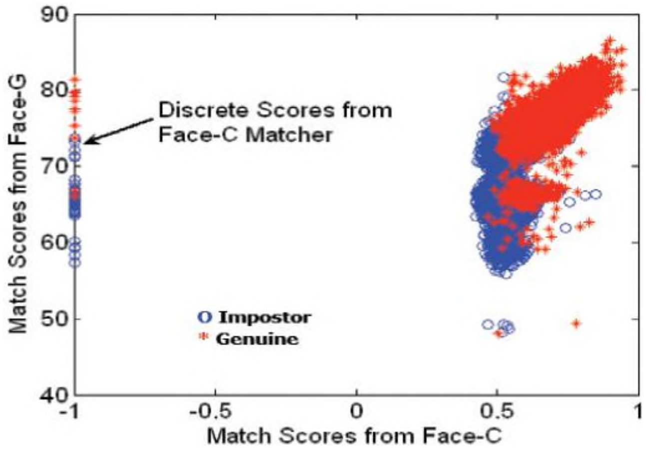


Figure 19. Match Scores from the Two Face Matchers in NIST-BSSR1 Database. Score ranges are different; C: [-1,1], G: [0,100] ; Statistical distributions are different. In addition, they have continuous and discrete components; Scores from the matchers are correlated [23].

NIST BSSR1 Set 1 data includes four 517x517 matrices of matcher scores based on biometric samples collected from 517 individuals [21]. Each matrix of scores is called a similarity matrix. Each score represents the one-to-one comparison of biometric samples collected from either one individual or two individuals. There is one BSSR1 Set 1 similarity matrix associated with each of the following: right index fingerprints, left index fingerprints, frontal face images scored by vendor matcher "C," and frontal face images scored by vendor matcher "G." The fingerprint matcher scores come from the NIST Verification Test Bed (VTB). It has produced fusion results for the six different combinations of multi-biometric data. The examples in the Figure 20 focus primarily on fusion results obtained using the face matcher C and the right index fingerprint scores [28, 21]. A two public domain databases are summarised in Table 6. XM2VTS-Benchmark and NIST-BSSR1.

Table 6. Summary of Multi-biometric Databases: NIST-BSSR1 [21], XM2VTS-Benchmark [22]

Database	Biometric Traits	No. of Matchers (K)	No. of Users
NIST-Multimodal	Fingerprint (2 Fingers) Face (2 Matchers)	4	517
NIST-Fingerprint	Fingerprint (2 Fingers)	2	6000
NIST-Face	Face (2 Matchers)	2	3000
XM2VTS-Benchmark	Face (5 Matchers) Speech (3 Matchers)	8	295

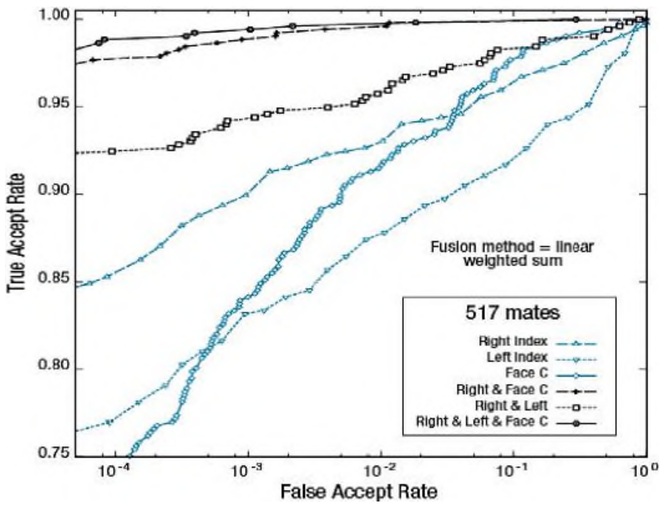
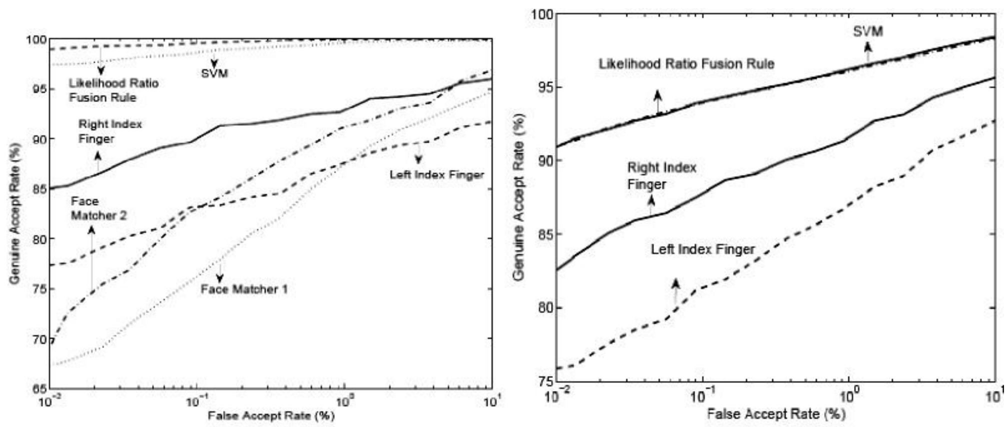
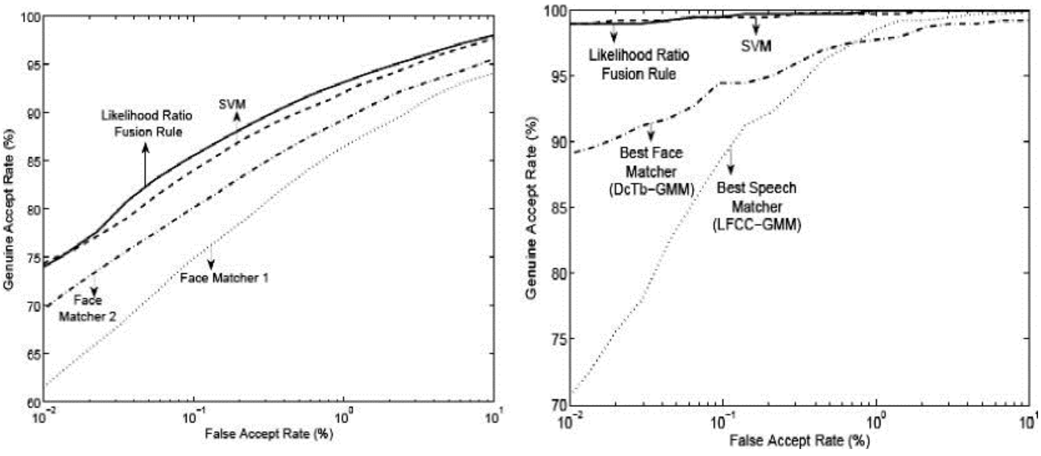


Figure 20. ROC for Multi-modal Fusion (NIST BSSR1 Database) [28]

Figure 20 depicts the ROCs for the original BSSR1 data for the face C and right and left index fingerprint matcher scores. In addition, the figure depicts the results of fusing right and left index fingerprint scores, face C and right index fingerprint scores, and face C with both index fingerprint scores. In all three cases, the results are based on using the best linear combination method optimized for FAR = 10-4. For the data analyzed, fusion produced a significant degree of improvement over using one modality. Moreover, when face C and right index fingerprint data are fused, the resulting TAR is greater than when fusing the left index fingerprint scores with the right index fingerprint scores. As depicted in Figure 20, when face C and right index fingerprint data are fused, the resulting TAR is about 97.5% at FAR = 10^{-4}, compared to 92.5% when fusing the more accurate left index fingerprint scores with the right index fingerprint scores. As mentioned previously, a significant improvement was expected since the multi-biometric (face and fingerprint scores) data for the mates is uncorrelated (r^2 = 0.0008). The ROC resulting from fusing both right and left index fingerprint with the face C data shows further increased improvement. For example, at FAR = 10^{-4} the false reject rate (FRR), which is 1 minus the TAR, is reduced by about 50% compared to fusing only the right index fingerprint score with the face C score [28].



(a) (b)



(c) (d)

Figure 21. Performance of Likelihood Ratio based Fusion Rule and SVM-based Fusion on (a) NIST-Multimodal, (b) NIST-Fingerprint, (c) NISTFace and (d) XM2VTS-Benchmark Databases. Although there are 8 matchers in the XM2VTS-Benchmark database, only the ROC curves of the best face matcher (DCTb-GMM) and the best speech matcher (LFCC-GMM) are shown in (d) for clarity [71, 82].

10. Applications Strategies of Biometric Systems

Literature and public perception of biometrics often focus on security applications [67]. For example, governments seek to prevent terrorist attacks, and public is concerned about the danger of surveillance. However, the application scope of biometric technologies is much wider and can be classified in vertical and horizontal segments. Whereas the vertical sovereign segment (government and authorities) is driven by the security benefit, other applications are driven by security and convenience, e.g. the customization of services. These vertical segments are the non-sovereign public sector, the private sector (investment goods) and consumer sector (consumer goods). This vertical distinction will be used to structure the following sections [67].

In practical terms, biometrics will be used mainly for four purposes [68]: law enforcement, physical access control (including border control), logical access control and convenience. Traditionally, the most widespread use of biometrics has been in law enforcement. Fingerprints have been used since the 19th century, and more recently DNA analysis has become routine in assisting criminal investigations. It is due to this history that many citizens associate enrolment in biometric systems with criminals and hence tend to resent it. Therefore, it is important to underline that law enforcement is only one among many possible application areas.

Law enforcement is however until now the only area where large-scale applications have been in use for some time. Physical access control based on biometrics has so far been mostly limited to private companies' premises, i.e. small-scale applications. However, there are a number of trials underway or recently completed, many of which are at airports, which have tested biometrics access with large numbers of customers, rather than employees. Most importantly, on the government side the integration of biometrics into passports and visas will for the first time create truly large-scale physical access control applications [68].

Logical access control (in particular online identity) is forecast to be a fast growing use of biometrics. With more and more transactions such as e-banking, e-commerce and e-government taking place online, biometrics offer a promising way of establishing secure identities especially when face-to-face contact between the participants in the transaction is not possible. This is particularly important for high-value financial transactions and for the transmission of confidential data (for example tax returns). Logical access control will also include access to entitlements offline, such as social security pay-outs [68].

Finally, convenience applications include all uses of biometrics where individuals voluntarily participate because they find it advantageous to do so. This would include ambient intelligence applications such as personally-adjusted home lighting or e-toys, but also participation in biometric applications offered by private actors, such as shops, sports clubs or other, where participation is not mandatory [68].

Figure 22. Biometrics Application Examples. (a) Digital Persona's fingerprint verification system provides personal recognition for computer and network login. (b) Indivos manufactures a fingerprint based point-of-sale (POS) terminal that verifies customers before charging their credit cards. (c) BioThentica's fingerprint-based door lock restricts access to premises. (d) The Inspass immigration system, developed by Recognition Systems and installed at major airports in the US, uses hand geometry verification technology [62].

E-learning: Human-Computer Interaction (HCI) is a research area aiming at making the interaction with computer systems more effective, easier, safer and more seamless for the users [2]. A challenge in today's Internet is providing easy collaboration across administrative boundaries. Using and sharing resources between individuals in different administrative domains should be just as easy and secure as sharing them within a single domain. E-learning continues to grow at a tremendous rate. Aiming to achieve state-of-the-art and continuously up-to-dated set of e-learning tools we characterize a task that goes beyond the capability of a single research and technology group. In my opinion, it is mandatory that a joint development effort must exist and be followed by a process for adoption or definition of e-learning standards regarding architecture, language, data format and content objects. E-learning has emerged in the recent times as an alternate solution to the traditional education. One of the main drawbacks of the e-learning is the difficulty to do online user verification. A number of reasons, resource owners may want to restrict the access to their resources to certain user groups, or provide user-specific contents. In both cases **authentication** and **authorization** is required. Figures 23 and 24 are Systems Without and With Authentication and Authorization Infrastructure (AAI) successively [5].

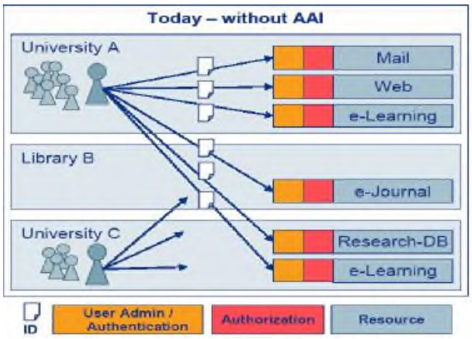


Figure 23. Systems without Authentication and Authorization Infrastructure (AAI) L5]

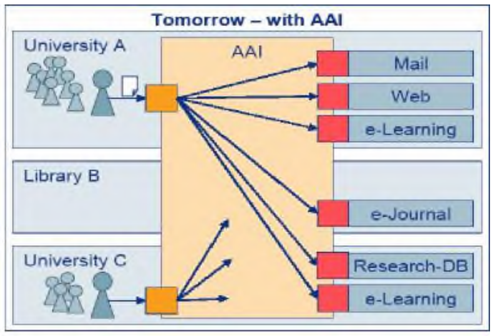


Figure 24. Systems with Authentication and Authorization Infrastructure (AAI) L5]

Future Applications of Biometric Systems: The objective of the biometric multidimensional targets presented in this paper, is to broaden the scope of thinking on the future of biometrics and to raise key issues that might at present be overlooked. Four dimensions are depicted: biometrics at the borders, in the health sector, in business and in everyday life. They can be placed on a continuum ranging from public-sector applications, to private applications with little or no government involvement. Privacy, security, usability and user acceptance concerns differ according to the environment [68].

Dimension 1: The everyday life dimension depicts a day in the life of a traditional family, in the form of a diary entry by the teenage son. The everyday life dimension draws attention to one basic fact about biometric technologies: that they can never be 100% secure. There is a trade-off between allowing impostors through the system (false accept) and denying access or services to legitimate users (false reject); the choice of threshold will depend on the nature of the application.

Dimension 2: The use of biometrics in business environments can be for various purposes: internal (e.g. for employees) and external (e.g. with clients, other companies). The business dimension is presented as a memo to the senior management of a large multinational supermarket chain which has embraced the use of biometrics but is concerned that it is not reaping the expected benefits (access control, auditing working hours, and customer loyalty). It shows that back-up/alternative procedures are important and that biometric access systems are only as secure as their weakest link, which is, in this case as in most cases, human. The business dimension describes how users concerned about their privacy may reject biometrics when there is little perceived added value for them.

Dimension 3: The health dimension presents an exchange of e-mails between two doctors in different countries. Strong identification is essential in the health sectors retrieving medical histories, administering medicine, handing out prescriptions, and carrying out medical procedures, all rely on the correct identification of the individual. In addition there is a strong need for privacy given the sensitive nature of medical data. These two requirements make the health sector a very likely field for the application of biometrics.

Dimension 4: Biometrics at the borders is likely to occur within the shortest timeframe as concrete plans for this application already exist and it needs challenging for the concept issues sophisticated. For examples are biometric visa and biometric ticket.

11. Conclusion

Security is not enforced by focusing on a single parameter. Instead of solving a one-dimensional problem, a secure environment requires multiple dimensions of critical check points. Secure authentication is provided by multiple parameters. One parameter is a security token an individual uniquely possesses, such as a physical key or a smart card. Another parameter is an item an individual uniquely knows, such as a PIN. An additional parameter is an individual's unique biological characteristic, such as DNA or an iris code.

Biometric technology could make a huge positive impact into society, if it is correctly utilised to increase the robustness of security systems across the world. This would help to cope with the rising levels of fraud, crime and terrorism. However, Biometrics has the potential to hinder society by allowing governments and corporations a massive amount of personal information regarding individuals.

To integrate fully biometric identification systems will be a lengthy process, but the technology has the potential to change the way the world works, no more passwords and smart cards, just using your body as your key. However, biometrics has been usefully applied for matters of lower importance, time monitoring systems and industry authentication systems. As the progress of technology increases, it is assured that biometrics can be effectively applied to important systems.

There is no doubt that biometrics is the next stage of ubiquitous security technology in our increasingly paranoid, authoritarian society. However, there is still much to be done: customers are scared off by high failure-to-enroll and false non-match rates as well as incompatibilities. Furthermore, system security as a whole needs more care to be taken of.

Future improvements in acquisition technology and algorithms as well as the availability of industry standards will certainly assure a bright future for biometrics. Will this be the end of traditional password or token-based systems. Certainly not biometrics is not the perfect solution either, it is just a good trade-off between security and ease of use.

Multi-biometric systems alleviate a few of the problems observed in uni-modal biometric systems. Besides improving matching performance, they also address the problems of non-universality and spoofing. Multi-biometric systems can integrate information at various levels, the most popular one being fusion at the matching score level where the scores output by the individual matchers are integrated. Performance gain is pronounced when uncorrelated traits are used in a multimodal system. Incorporating user-specific parameters can further improve performance of these systems. With the widespread deployment of biometric systems in several civilian and government applications, it is only a matter of time before multimodal biometric systems begin to impact the way in which identity is established in the 21st century.

12. Further Directions

The question arises how we can obtain the “best” (in terms of accuracy) fusion solution. It is well known that the “best” fusion solution is one that satisfies the Neyman-Pearson (NP) criterion. The NP criterion characterizes the fusion solution that maximizes the TAR for a fixed value of FAR [28, 29].

It’s usefull to apply the concept of Quantum Computation theory, proposed in the 80’s. Uses the strange properties of quantum systems and has been demonstrated on a small scale. As it’s known from quantum logic, the concept of quantum bits (qubits) is based on the superposition of quantum states [30]. A Quantum Computing (QC) is a Motivation method of computation that uses a dynamic process governed by the Schrodinger Equation (SE) [31].

The history of quantum information processing goes as far as to well-known pioneer John von Neumann. Quantum computing offers enormous parallelism. The size of computational state space grows exponentially with its physical size, i.e. the number of used quantum bits (qubits). Qubits can be realized by any two-dimensional quantum system. The polarization of photons (vertical/horizontal) or 1/2-spin momentum of particles (up/down) are examples of the ones that are best explored. An isolated quantum system corresponds to a Hilbert space, which is a suitable mathematical framework for describing processes of quantum mechanics. Single qubit with two orthogonal states (denoted by |0> and |1>) is told to live in Hilbert space H_2. While the classical bit can be only in one of the two states 0 or 1, a qubit can take any linear superposition of |0> and |1>. A tensor product of n qubits is called an n-qubit quantum register with corresponding Hilbert space H_{2^n}. If an n-qubit register cannot be written as a tensor product of smaller Hilbert spaces the qubits are told to be entangled. Entanglement is a powerful quantum primitive that allows for even very distant parts of the system to be strongly tied. Its applications vary from speed-up of classical computations (e.g. polynomial time integer-factoring) or quantum key generation to teleportation. Entanglement is the main reason why quantum computers cannot be efficiently simulated by classical ones [32].

References

[1] G. Chetty and M. Wagner, "Audio-Visual Multimodal Fusion for Biometric Person Authentication and Liveness Veriftion", Copyright © 2006, Australian Computer Society, Inc., Proceedings of the 2005 NICTA-HCSNet Multimodal User Interaction Workshop, vol. 57, **(2005)**, pp. 17-24.

[2] A. Corradini, M. Mehta, N. O. Bernsen, J. C. Martin, S. Abrilian, Multimodal Input Fusion in Human-Computer Interaction, On the Example of the NICE Project 2003, http://perso.limsi.fr/Individu/martin/research/articles/NATO-ASI_1.pdf, **(2003)**.

[3] I. Simpson, "Biometrics: Issues and Applications", 6th Annual Multimedia Systems, Electronics and Computer Science, University of Southampton, **(2006)**.

[4] S. Gleni and P. Petratos, "DNA Smart Card for Financial Transactions", The ACM Student Magazine, **(2004)**, http://www.acm.org.

[5] Authentication and Authorization Infrastructure (AAI) in a nutshell, The Swiss Education & Research Network, http://www.switch.ch/aai.

[6] http://biolab.csr.unibo.it/Research.asp.

[7] P. J. Phillips, A. Martin and C. L. W. M. Przybocki, "An Introduction to Evaluating Biometric Systems", National Institute of Standards and Technology, **(2000)**, IEEE.

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

[9] A. Jain, et. al., "BIOMETRICS: Personal Identification in Networked Society", Kluwer Academic Publishers, **(1999)**, ISBN 0-7923-8345-1.

[10] D. Dessimoz, J. Richiardi, C. Champod and A. Drygajlo, "Multimodal Biometrics for Identity Documents", State-of-the-Art, Research Report PFS 341-08.05 (Version 2.0),UNIL Univesite de Lausanne (EPFL), **(2006)** June.

[11] B. Dorizzi, P. Lamadeleine, C. Guerrier and J. L. Les Jardins, "Biométrie: Techniques et usages", Revue des sciences et techniques de l'ingenieur, **(2004)** April.

[12] B. Dorizzi, "Biometrics at the frontiers, assessing the impact on Society Technical impact of Biometrics", Background paper for the Institute of Prospective Technological Studies, DG JRC – Sevilla, European Commission, **(2005)** January.

[13] R. Bailey, D. Blackburn, J. Dunn, V. Lively, J. Mayer-Splain and S. Sexton, "Biometrics History", National Science and Technology Council (NSTC), http://www.biometrics.gov & http://www.ostp.gov/nstc, **(2006)** August.

[14] J. Ashbourn, "Biometrics: Advanced Identity Verification", Springer, London, **(2000)**.

[15] S. Baird, "Biometrics", The Technology Teacher, **(2002)** February.

[16] J. Ortega-Garcia, J. Bigun, D. Reynolds and J. Gonzalez-Rodriguez, "Authentication Gets Personal with Biometrics", IEEE Signal Processing Magazine, vol. 21, no. 2, **(2004)** March, pp. 50-62.

[17] P. S. P. Wang and S. N. Yanushkevich, "Biometric Technologies and Applications", Proc. IASTED2007, Austria, **(2007)**, pp. 226-231.

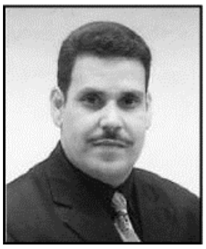
[18] P. S. Aleksic and A. K. Katsaggelos, "Audio-Visual Biometrics", Proceedings of the IEEE, vol. 94, no. 11, **(2006)** November.

[19] L. A. Hornak, "The Center for Identification Technology Research, Summary", An IUCRC Planning Grant Proposal, West Virginia University, Michigan State University and San Jose State University & Marshall University.
[20] A. Ross and A. K. Jain, "Multimodal Biometrics: An Overview", Appeared in Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), **(2004)** September, pp. 1221-1224.
[21] National Institute of Standards and Technology, "NIST Biometric Scores Set - release 1," http://www.itl.nist.gov/iad/894.03/biometricscores, **(2004)**.
[22] N. Poh and S. Bengio, "Database, Protocol and Tools for Evaluating Score-Level Fusion Algorithms in Biometric Authentication", IDIAP RR 04-44, **(2004)** August, a IDIAP, CP 592, 1920 Martigny, Switzerland.
[23] A. K. Jain, et. al., "Quality-Based Fusion in Multi-Biometric Systems", NIST Quality Workshop, Michigan State University, http://biometrics.cse.msu.edu, **(2006)** March.
[24] http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb/.
[25] http://www.ee.surrey.ac.uk/banca/.
[26] S. Marcel, "Face Detection and Bi-Modal Authentication", IDIAP Research Institute, Martigny, Switzerland, http://www.idiap.ch, **(2005)**, February.
[27] S. Marcel, J. Mariethoz, Y. Rodriguez and F. Cardinaux, "Bi-Modal Face and Speech Authentication: A BioLogin Demonstration System", IDIAP–RR 06-18 TO APPEAR IN Proceedings of the Second Workshop on Multimodal User Authentication, **(2006)** March.
[28] H. J. Korves, L. D. Nadel, B. T. Ulery and D. M. B. Masi, "Multi-biometric Fusion: From Research to Operations", MTS MitreTek Systems, http://www.mitretek.org/home.nsf/Publications/SigmaSummer2005, **(2005)**, pp. 39-48.
[29] V. Trees and L. Harry, "Detection, Estimation, and Modulation Theory", Part I, John Wiley and Sons, **(1968)**.
[30] P. K. Amiri and S. A. Shahdi, "Proposing a new approach to implementing Fuzzy systems using Quantum Computational Methods", http://arxiv.org/vc/quant-ph/papers/0210/0210037v1.pdf, **(2002)**.
[31] A. N. Al-Rabadi and G. G. Lendaris, "Artificial Neural Networks Implementation Using Many-Valued Quantum Computing", DOI:10.1109/IJCNN.2003.1224069 ISBN: 0-7803-7898-9In proceeding of: Neural Networks, 2003. Proceedings of the International Joint Conference on, vol. 4, **(2003)** August.
[32] M. Dobsicek, J. Kolar and R. Lorencz, "Quantum Technologies for Trust and Security", 130.203.133.150/viewdoc/summary?doi=10.1.1.130.2221.
[33] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, "Likelihood Ratio Based Biometric Score Fusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, **(2007)** February, pp. 342-347.
[34] J. P. Campbell, "Phonetic, Idiolectic, and Acoustic Speaker Recognition", IEEE Distinguished Lecture, IEEE Signal Processing Society, INRS Telecommunications, U. Québec, **(2002)** April 19.
[35] Y. W. Yun, "The '123' of Biometric Technology", Synthesis Journal, **(2002)**, pp. 83-95.
[36] J. Matas, M. Hamouz, K. Jonsson, J. Kittler, Y. Li, C. Kotropoulos, A. Tefas, I. Pitas, T. Tan, H. Yan, F. Smeraldi, J. Begun, N. Capdevielle, W. Gerstner, S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, "Comparison of Face Verication Results on the XM2VTS Database", In Proc. 15th Int'l Conf. Pattern Recognition, vol. 4, **(2000)**, pp. 858-863, Barcelona.
[37] E. Bailly-Bailliere, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariethoz, J. Matas, K. Messer, V. Popovici, F. Poree, B. Ruiz and J. -P. Thiran, "The BANCA database and evaluation protocol", in Proc. Audio-and Video-Based Biometric Person Authentication, Guilford, **(2003)**, pp. 625-638.
[38] C. Sanderson and K. K. Paliwal, "Identity verification using speech and face information", Digital Signal Processing, vol. 14, no. 5, **(2004)**, pp. 449-480.
[39] C. Sanderson and K. K. Paliwal, "Noise compensation in a person verification system using face and multiple speech features", Pattern Recognition, vol. 36, no. 2, **(2003)** February, pp. 293–302.
[40] S. Marcel and S. Bengio, "Improving Face Verication Using Skin Color Information", In Proc. 16th Int. Conf. on Pattern Recognition, page unknown, Quebec, **(2002)**.
[41] C. Sanderson and K. K. Paliwal, "Fast Features for Face Authentication Under Illumination Direction Changes", Pattern Recognition Letters, vol. 24, no. 14, **(2003)**, pp. 2409-2419.
[42] L. Rabiner and B. -H. Juang, "Fundamentals of Speech Recognition", Oxford University Press, **(1993)**.
[43] S. Ikbal, H. Misra and H. Bourlard, "Phase Auto-Correlation (PAC) derived Robust Speech Features", In Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP-03), **(2003)**, pp. 133-136, Hong Kong.
[44] K. K. Paliwal, "Spectral Subband Centroids Features for Speech Recognition", In Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), vol. 2, **(1998)**, pp. 617-620, Seattle.
[45] N. Poh, C. Sanderson and S. Bengio, "An Investigation of Spectral Subband Centroids For Speaker Authentication", In Springer LNCS-3072, Int'l Conf. on Biometric Authentication (ICBA), **(2004)**, pp. 631-639, Hong Kong.
[46] N. Poh, C. Sanderson and S. Bengio, "An Investigation of Spectral Subband Centroids For Speaker Authentication. Research Report 03-62, IDIAP, Martigny, Switzerland, 2003.

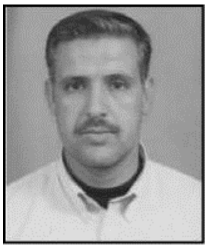
[47] S. Marcel and S. Bengio, “Improving Face Verication Using Skin Color Information”, In Proc. 16th Int. Conf. on Pattern Recognition, page unknown, Quebec, **(2002)**.
[48] F. Cardinaux, C. Sanderson and S. Marcel, “Comparison of MLP and GMM Classiers for Face Verication on XM2VTS”, In Springer LNCS-2688, 4th Int'l Conf. Audio- and Video-Based Biometric Person Authentication (AVBPA'03), **(2003)**, pp. 911-920, Guildford.
[49] N. Poh and S. Bengio, “Noise-Robust Multi-Stream Fusion for Text-Independent Speaker Authentication”, In The Speaker and Language Recognition Workshop (Odyssey), **(2004)**, pp. 199-206, Toledo.
[50] C. C. Chibelushi, F. Deravi and J. S. Mason, “Voice and facial image integration for speaker recognition”, in Proc. IEEE Int. Symp. Multimedia Technologies Future Appl., Southampton, U.K., **(1993)**.
[51] R. Brunelli and D. Falavigna, “Person identification using multiple cues”, IEEE Trans. Pattern Anal. Machine Intell., vol. 10, **(1995)** October, pp. 955–965.
[52] S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, “Fusion of face and speech data for person identity verification”, IEEE Trans. Neural Networks, vol. 10, **(1999)**, pp. 1065-1074.
[53] T. J. Hazen, E. Weinstein, R. Kabir, A. Park and B. Heisele, “Multi-modal face and speaker identification on a handheld device”, in Proc. Works Multimodal User Authentication, Santa Barbara, CA, **(2003)**, pp. 113-120.
[54] N. A. Fox, R. Gross, P. de Chazal, J. F. Cohn and R. B. Reilly, “Person identification using automatic integration of speech, lip, and face experts”, in Proc. ACM SIGMM 2003 Multimedia Biometrics Methods and Applications Workshop (WBMA’03), Berkeley, CA, **(2003)**, pp. 25-32.
[55] N. A. Fox and R. B. Reilly, “Audio-visual speaker identification based on the use of dynamic audio and visual features”, in Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication, Guildford, U.K., **(2003)**, pp. 743-751.
[56] A. V. Nefian, L. H. Liang, T. Fu and X. X. Liu, “A Bayesian approach to audio-visual speaker identification”, in Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication, Guildford, U.K., **(2003)**, pp. 761-769.
[57] A. Kanak, E. Erzin, Y. Yemez and A. M. Tekalp, “Joint audio-video processing for biometric speaker identification”, in Proc. Int. Conf. Acoustic, Speech Signal Processing, Hong Kong, China, **(2003)**, pp. 561-564.
[58] D. M. Blackburn, “BIOMETRICS 101”, Version 3.1, Federal Bureau of Investigation, **(2004)** March, http://www.biometricscatalog.org.
[59] A. Kounoudes, N. Tsapatsoulis, Z. Theodosiou and M. Milis, “POLYBIO: Multimodal Biometric Data Acquisition Platform and Security System”, Lecture Notes In Computer Science, vol. 5372, **(2008)**, pp. 216-227.
[60] D. Betch, “DNA Fingerprint in Human Health and Society”, Biotechnology Information Series (Bio-6).
[61] J. Wambaugh, “The Blooding”, William Morrow, N.Y. **(1989)**.
[62] S. Prabhakar, S. Pankanti, A. K. Jain, “Biometric Recognition: Security and Privacy Concerns”, IEEE Security & Privacy Magazine, vol. 1, no. 2, **(2003)** March-April, pp. 33-42.
[63] http://www.statewatch.org/news/2005/mar/Report-IPTS-Biometrics-for-LIBE.pdf.
[64] K. Daimi and K. Snyder, “Security Requirements for Multimodal Biometric Systems”, in Proc. Security and Management, **(2006)**, pp.72-77.
[65] J. Clark and A. Yuille, “Data Fusion for Sensory Information Processing Systems”, Boston, MA: Kluwer, **(1990)**.
[66] A. K. Jain, A. Ross and S. Prabhakar, “An Introduction to Biometric Recognition”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, **(2004)** January.
[67] S. Schulz, “The German Biometric Strategy Platform: Biometrics State of the Art, Industry Strategy Development, and Platform Conception”, Gregor Költzsch MBA Leila Ambrosio, **(2005)** May.
[68] I. Maghiros, Y. Punie, S. Delaitre, E. Lignos, C. Rodríguez, M. Ulbrich, M. Cabrera, B. Clements, L. Beslay, and R. van Bavel, “Biometrics at the Frontiers: Assessing the impact on Society”, Institute for Prospective Technological Studies, http://www.jrc.es , EC-DG JRC-IPTS © European Communities, **(2005)**.
[69] T. Shirai, “Biometrics in Japan”, Science and Technology Section, British Embassy, Tokyo, **(2003)** March.
[70] S. Prabhakar and A. K. Jain, “Decision-level fusion in fingerprint verification”, Pattern Recognition, vol. 35, no. 4, **(2002)**, pp. 861-874.
[71] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, “Biometric Score Fusion: Likelihood Ratio, Matcher Correlation and Image Quality,” DRAFT, **(2007)** March.
[72] http://www2.citer.wvu.edu.
[73] M. E. Schuckers, C. J. Knickerbocker, “Documentation for Program for Rate Estimation and Statistical Summaries (PRESS)”, PRESS v1.0, **(2004)** September 8.
[74] R. Snelick, U. Uludag, A. Mink, M. Indova and A. Jain, “Large Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, **(2005)** March, pp. 450-455.
[75] P. J. Huber, “Robust Statistics”, Wiley, **(1981)**.
[76] R. Auckenthaler, M. Carey and H. Lloyd-Thomas, “Score Normalization for Text-Independent 58Speaker Verification Systems”, Digital Signal Processing, vol. 10, **(2000)**, pp. 42-54.

[77] A. Jain, K. Nandakumar and A. Ross, "Score normalization in multimodal biometric systems", The Journal of Pattern Recognition Society, Elsevier, vol. 38, **(2005)**, pp. 2270-2285.
[78] A. K. Jain and A. Ross, "Learning User-Specific Parameters in a Multibiometric System", Proc. IEEE International Conference on Image Processing (ICIP), Rochester, NY, **(2002)** September, pp. 57-60.
[79] G. Doddington, W. Liggett, A. Martin, M. Przybocki and D. Reynolds, "Sheeps, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation", Proc. ICSLD 98, Sydney, Australia, **(1998)** November.
[80] R. M. Bolle, S. Pankanti and N. K. Ratha, "Evaluation Techniques for Biometrics-based Authentication Systems (FRR)", Proc. 15th International Conference on Pattern Recognition (ICPR), vol. 2, **(2000)** September, pp. 831-837.
[81] S. C. Dass, K. Nandakumar and A. K. Jain, "A Principled Approach to Score Level Fusion in Multimodal Biometric Systems", in Proc. Of AVBPA, Rye Brook, USA, **(2005)** July, pp. 1049-1058.
[82] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, "Likelihood Ratio Based Biometric Score Fusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, **(2007)**.

Authors



Mohamed SOLTANE (Assoc. Prof. Dr.) received the M.Eng. degree in Electronics from Badji-Mokhtar University of Annaba-Algeria, in 1995 and the M.Sc. degree in Electrical and Electronics Engineering from UKM Malaysia in 2005, and the Ph.D. degrees in Electronics from Badji-Mokhtar University of Annaba-Algeria, in 2010. He is currently an associate professor at Electrical Engineering & Computing Department, Faculty of Sciences & Technology, DOCTOR YAHIA FARES UNIVERSITY OF MEDEA, ALGERIA. His research interests include statistical pattern recognition, biometric authentication, cryptography and quantum computing, computer vision and machine learning and micro-computer based system design.



Mimen BAKHTI (Assist. Prof.) received the M.Sc. degree in Electronics major Micro-Electronics from University of Science & Technology of ORAN (U.S.T.O) Oran-Algeria, in 1999. He has been with the Electrical Engineering & Computing Department, Faculty of Sciences & Technology, DOCTOR YAHIA FARES UNIVERSITY OF MEDEA, ALGERIA, as an Assistant Professor since 2001. His research interests include photonics; signal processing, computer vision and machine learning. And currently he is working toward his Ph.D. thesis.