International Journal of Database Theory and Application
Vol.8, No.4 (2015), pp.91-106
http://dx.doi.org/10.14257/ijdta.2015.8.4.10

Product of Likelihood Ratio Scores Fusion of Dynamic Face and On-line Signature Based Biometrics Verification Application Systems

Soltane Mohamed

Electrical Engineering & Computing Department, Faculty of Sciences & Technology
Doctor yahia fares University OF MEDEA, 26000 MEDEA, ALGERIA
&
Laboratoire des Systèmes Électroniques Avancées (LSEA)
soltane.mohamed@univ-medea.dz;soltane.mohamed.3099@gmail.com&xor99@hotmail.com

Abstract

In this paper, the use of finite Gaussian mixture modal (GMM) based Expectation Maximization (EM) estimated algorithm for score level data fusion is proposed. 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, 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 mono modal biometric systems. Simulation show that finite mixture modal (GMM) is quite effective in modelling the genuine and impostor score densities, fusion based the resulting density estimates achieves a significant performance on eNTERFACE 2005 multi-biometric database based on dynamic face and signature modalities.

Keywords: *Biometry, Multi-Modal, Authentication, Face Recognition, Signature Verification, data Fusion, Adaptive Bayesian decision, GMM &EM*

1. Introduction

BIOMETRIC 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 [1]. 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 [2]. 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 [3]. 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-

ISSN: 2005-4270 IJDTA

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 mono-biometric systems.

Some works based on multi-modal biometric identity verification systems has been reported in literature. Ben-Yacoub et al. [4] evaluated five binary classifiers on combinations of faces and voice modalities (XM2VTS database). They found that (i) a support vector machine and Bayesian classifier achieved almost the same performances; and (ii) both outperformed Fisher's linear discriminant, a C4.5 decision tree, and a multilayer perceptron. Korves et al. [5] compared various parametric techniques on the BSSR1 dataset. That study showed that the Best Linear technique performed consistently well, in sharp contrast to many alternative parametric techniques, including simple sum of z-scores, Fisher's linear discriminant analysis, and an implementation of sum of probabilities based on a normal (Gaussian) assumption. Jain et al. [6] propose a framework for optimal combination of match scores that is based on the likelihood ratio test. The distributions of genuine and impostor match scores are modeled as finite Gaussian mixture model. The proposed fusion approach is general in its ability to handle (i) discrete values in biometric match score distributions, (ii) arbitrary scales and distributions of match scores, (iii) correlation between the scores of multiple matchers and (iv) sample quality of multiple biometric sources. The performance of complete likelihood ratio based fusion rule was evaluated on the three partitions of the NIST-BSSR1 database and the XM2VTS-Benchmark database. As expected, likelihood ratio based fusion leads to significant improvement in the performance compared to the best single modality on all the four databases. At a false accept rate (FAR) of 0:01%. Jain et al.[7] applied the sum of scores, max-score, and min-score fusion methods to normalized scores of face, fingerprint and hand geometry biometrics (database of 100 users, based on a fixed TAR). The normalized scores were obtained by using one of the following techniques: simple distance-to-similarity transformation with no change in scale (STrans), min–max, z-score, median-MAD, double sigmoid, tanh, and Parzen. They found that (a) the min–max, z-score, and tanh normalization schemes followed by a simple sum of scores outperformed other methods; (b) tanh is better than min-max and z-score when densities are unknown; and (c) optimizing the weighting of each biometric on a user-by-user basis outperforms generic weightings of biometrics. Kittler et al. [8] proposed a multi-modal person verification system, using three experts: frontal face, face profile, and voice. The best combination results are obtained for a simple sum rule. Snelick et al.[9] compared combinations of z-score, min-max, tanh and adaptive (two-quadrics, logistic and quadric-line-quadric) normalization methods and simple sum, min score, max score, matcher weighting, and user weighting fusion methods (database of about 1000 users, at a fixed FAR). They found that (a) fusing COTS fingerprint and face biometrics does outperform mono-modal COTS systems, but the high performance of mono-modal COTS systems limits the magnitude of the performance gain; (b) for open-population applications (e.g., airports) with unknown posterior densities, min-max normalization and simple-sum fusion are effective; (c) for closed-population applications (e.g. an office), where repeated user samples and their statistics can be accumulated, QLQ adaptive normalization and user weighting fusion methods are effective. Teoh et al. [10] Applied a modified *k*-NN and evidence theoretic *k*-NN classifier based on Dampster-safer theory, and it found that the best result is obtained using the evidence theoretic *k*-NN classifier as it introduces low FAR and FRR compared to both the ordinary and modified *k*-NN.

A multi-modal biometric verification system based on dynamic facial and on-line signature modalities is described in this paper. Both still face images and on-line signature biometrics are chosen due to their complementary characteristics, physiology, and behavior. 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 [5].

2. Authentication Traits

2.1 Face Extraction and Recognition

Face recognition, authentication and identification are often confused. Face recognition is a general topic that includes both face identification and face authentication (also called verification). On one hand, face authentication is concerned with validating a claimed identity based on the image of a face, and either accepting or rejecting the identity claim (one-to-one matching). On the other hand, the goal of face identification is to identify a person based on the image of a face. This face image has to be compared with all the registered persons (one-to-many matching). Thus, the key issue in face recognition is to extract the meaningful features that characterize a human face. Hence there are two major tasks for that: Face detection and face verification.

2.1.1 Face Detection: Face detection is concerned with finding whether or not there are any faces in a given image (usually in gray scale) and, if present, returns the image location and content of each face. This is the first step of any fully automatic system that analyzes the information contained in faces (e.g., identity, gender, expression, age, race and pose). While earlier work dealt mainly with upright frontal faces, several systems have been developed that are able to detect faces fairly accurately with in-plane or out-of-plane rotations in real time. For biometric systems that use faces as non-intrusive input modules, it is imperative to locate faces in a scene before any recognition algorithm can be applied. An intelligent vision based user interface should be able to tell the attention focus of the user (i.e., where the user is looking at) in order to respond accordingly. To detect facial features accurately for applications such as digital cosmetics, faces need to be located and registered first to facilitate further processing. It is evident that face detection plays an important and critical role for the success of any face processing systems.

On the results presented on this paper only size normalization of the extracted faces was used. All face images were resized to 130x150 pixels, applying a bi-cubic interpolation. After this stage, it is also developed a position correction algorithm based on detecting the eyes into the face and applying a rotation and resize to align the eyes of all pictures in the same coordinates. The face detection and segmentation tasks presented in this paper was performed based on 'Face analysis in Polar Frequency Domain' proposed by Yossi Z. et al. [11]. First it extracts the Fourier-Bessel (FB) coefficients from the images. Next, it computes the Cartesian distance between all the Fourier-Bessel transformation (FBT) representations and re-defines each object by its distance to all other objects. Images were transformed by a FBT up to the 30^{th}Bessel order and 6^{th}root with angular resolution of 3°, thus obtaining to 372 coefficients. These coefficients correspond to a frequency range of up to 30 and 3 cycles/image of angular and radial frequency, respectively. Figure 1 Shows the face and eyes detections for different users from the database, and Figure 2 Shows the face normalization for the same users.

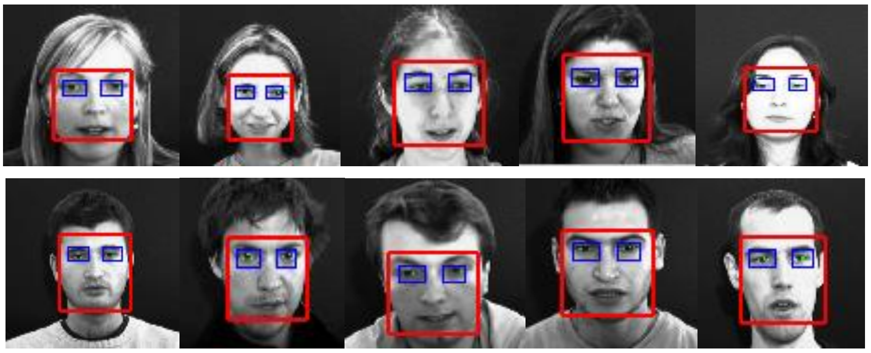

Figure 1. Face and Eyes Detections for Different Users

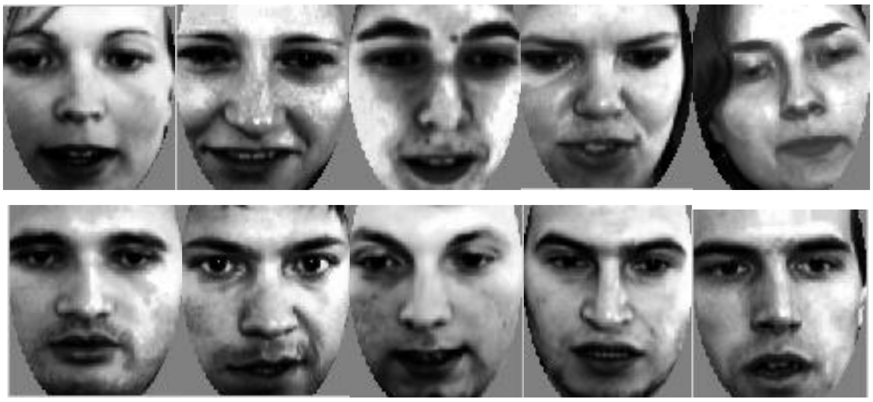

Figure 2. Face Normalization for the Above Users

Polar Frequency Analysis: The FB series is useful to describe the radial and angular components in images [11]. FBT analysis starts by converting the coordinates of a region of interest from Cartesian (*x*, *y*) to polar (*r*, *θ*). The *f* (*r*, *θ*) function is represented by the two-dimensional FB series, defined as:

$$f(r,\theta)=\sum_{i=1}^{\infty}\sum_{n=1}^{\infty}A_{n,i}J_n\left(\alpha_{n,i}r\right)\cos(n\theta)+\sum_{i=1}^{\infty}\sum_{n=1}^{\infty}B_{n,i}J_n\left(\alpha_{n,i}r\right)\sin(n\theta) \qquad (1)$$

where $J_n$is the Bessel function of order *n*, $f(R, \theta) = 0$ and $0 \le r \le R$. $\alpha_{n,i}$ is the i^{th}root of the $J_n$function, i.e. the zero crossing value satisfying $J_n(\alpha_{n,i}) = 0$ is the radial distance to the edge of the image. The orthogonal coefficients $A_{n,i}$ and $B_{n,i}$are given by:

$$A_{0,i}=\frac{1}{\pi R^2 J_1^2(\alpha_{n,i})}\int_{\theta=0}^{\theta=2\pi}\int_{r=0}^{r=R} f(r,\theta)rJ_n\left(\frac{\alpha_{n,i}}{R}r\right)dr\,d\theta \qquad (2)$$

$if B_{0,i} = 0\ and n = 0;$

$$\begin{bmatrix}A_{n,i}\\B_{n,i}\end{bmatrix}=\frac{2}{\pi R^2 J_{n+1}^2(\alpha_{n,i})}\int_{\theta=0}^{\theta=2\pi}\int_{r=0}^{r=R} f(r,\theta)rJ_n\left(\frac{\alpha_{n,i}}{R}r\right)\begin{bmatrix}\cos(n\theta)\\\sin(n\theta)\end{bmatrix}dr\,d\theta \qquad (3)$$

*if*n> 0.

An alternative method to polar frequency analysis is to represent images by polar Fourier transform descriptors. The polar Fourier transform is a well-known mathematical operation where, after converting the image coordinates from Cartesian to polar, as described above; a conventional Fourier transformation is applied. These descriptors are directly related to radial and angular components, but are not identical to the coefficients extracted by the FBT.

2.1.2 Face Verification

Feature Extraction: The so-called “Eigen faces” method [12] is one of the most popular methods for face recognition. It is based on the Principal Components Analysis (PCA) of the face images in a training set. The main idea is that since all human faces share certain common characteristics, pixels in a set of face images will be highly correlated. The K-L (Karhunen-Loeve) transform can be used to project face images to a different vector space that is of reduced dimensionality where features will be uncorrelated. In the new space nearest neighbor classifiers can be used for classification. Euclidean distances *d* in the projection space are mapped into the [0,1] interval of the real line using the mapping function: *f = d / (1+d).* It is easily seen that *f* is also a metric with distance values in [0,1]. Thus, the decomposition of a face image into an Eigen face space provides a set of features. The maximum number of features is restricted to the number of images used to compute the KL transform, although usually only the more relevant features are selected, removing the ones associated with the smallest eigenvalues. Two different approaches, database training stage and the operational stage [12]. The concept verification system is illustrated in figure 4.

The Training Stage: Face spaces are eigenvectors of the covariance matrix corresponding to the original face images, and since they are face-like in appearance, they are so is called Eigen faces.

Consider the training set of face images be $i_1, i_2, \dots, i_m$; the average face of the set is defined as:

$$\bar{i} = \frac{1}{M}\sum_{j=1}^{M} i_j \tag{4}$$

Where M is the total number of images.

Each face differs from the average by the vector $\emptyset_n = i_n - \bar{i}$. A covariance matrix is constructed where:

$$C = \sum_{j=1}^{M} \emptyset_j \emptyset_j^T = AA^T \tag{5}$$

Where$A = [\emptyset_1\ \emptyset_2 \dots \emptyset_M]$.

Then, the eigenvectors v_k and the eigenvalues λ_k with a symmetric matrix $\boldsymbol{C}$are calculated. v_kDetermines the linear combination of $\boldsymbol{M}$ difference images with $\emptyset$ to form the Eigen faces:

$$u_l = \sum_{k=1}^{M} v_{lk} \emptyset_k l = 1, \dots, M \tag{6}$$

From these Eigen faces, $K(< M)$ Eigenfaces are selected corresponding to the Khighest eigenvalues.

At the training stage, a set of normalized face images, $\{i\}$, that best describe the distribution of the raining facial images in a lower dimensional subspace (Eigen face) is computed by the following operation:

$$\omega_k = u_k(i_n - \bar{i}) \tag{7}$$

Where $n = 1, \dots, M$and $k = 1, \dots, K$.

After that, the training facial images are projected onto the Eigen space, $\boldsymbol{\Omega_i}$, to generate representations of the facial images in Eigenface

$$\boldsymbol{\Omega}_i = (\omega_{n1}, \omega_{n2}, \dots, \omega_{nk}) \tag{8}$$

Where$n = 1, 2, \dots, M$.

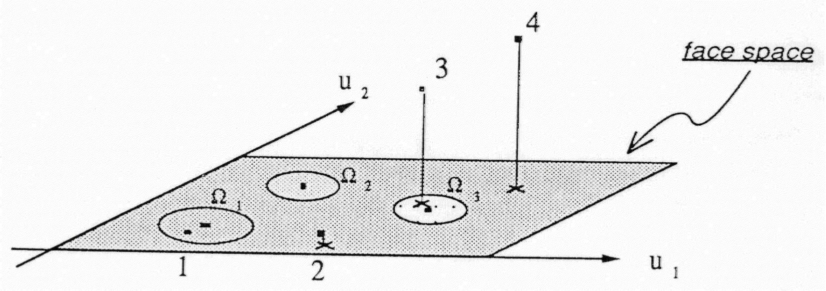

Figure 3. Simplified Version of the Face Space Illustrating the Four Results of the Projection of an Image onto the Face Space

In this case there are two Eigen faces, u1 and u2 [12]

The Operational Stage: This approach is based on the same principles as standard PCA, explained in the training stage. The difference is that an Eigen face space is extracted for each user. Thus, when a claimant wants to verify its identity, its vectored face image is projected exclusively into the claimed user Eigen face space and the corresponding likelihood is computed. The advantage of this approach is that it allows a more accurate model of the user's most relevant information, where the first Eigen faces are directly the most representative user's face information. Another interesting point of this method is its scalability in terms of the number of users. Adding a new user or new pictures of an already registered user only requires computing or re-computing the specific Eigen face space, but not the whole dataset base as in the standard approach. For verification systems, the computation of the claimant's likelihood to be an specific user is independent on the number of users in the dataset. On the contrary, for identification systems, the number of operations increases in a proportional way with the number of users, because as many projections as different users are required. In the verification system described in this article, the independent user Eigen face approach has been chosen. Each user's Eigen face space was computed which 16 frames extracted from the database still faces.

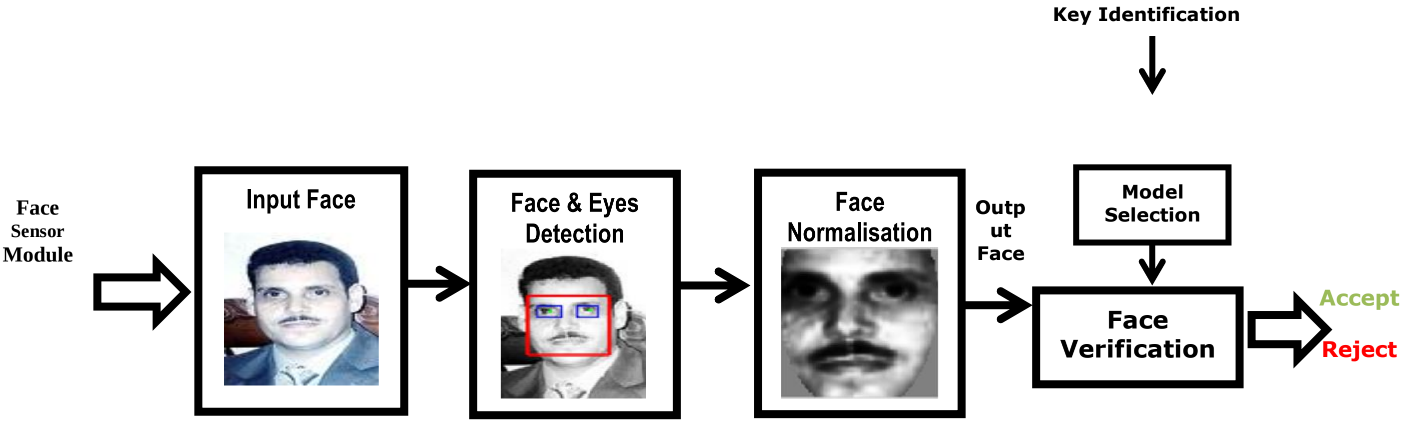

Figure 4. Face Verification Concept System

2.2 Signature Verification Systems

Handwritten signature is one of the first accepted civilian and forensic biometric identification technique in our society [13-15]. Human verification is normally very accurate in identifying genuine signatures. A signature verification system must be able to detect forgeries and at the same time reduce rejection of genuine signatures. The signature verification problem can be classified into categories: offline and online. Offline signature verification does not use dynamic information that is used extensively in online signature verification systems. This paper investigates the problem of offline signature verification. The problem of offline signature verification has been faced by taking into account three different types of forgeries: random forgeries, produced without knowing either the name of the signer or the shape of his signature; simple forgeries, produced knowing the name of the signer but without having an example of his signature; and skilled forgeries, produced by people who, looking at an original instance of the signature, attempt to imitate it as closely as possible.

Figure 5. Wacom Graphire3 Digitizing TabletPC

Feature Extraction: The coordinate trajectories(x_n, y_n) and pressure signal p_n are the components of the unprocessed feature vectors $u_n = [x_n, y_n, p_n]^T$ extracted from the signature signal [13-14, 16], where $n =1,...,N_s$and N_s is the duration of the signature in time samples. Signature trajectories are then pre-processed by subtracting the centre of mass followed by rotation alignment based on the average path tangent angle. An extended set of discrete-time functions are derived from the pre-processed trajectories consisting of sample estimations of various dynamic properties. As s result, the parameterised signature O consists in the sequence of feature vectors $o_n = [x_n, y_n, p_n, \theta_n, v_n, \dot{x}_n, \dot{y}_n]^T$, $n =1,...,N_s$, where the upper dot notation represents an approximation to the first order time derivative and $\boldsymbol{\theta}$ ***and*** $\boldsymbol{v}$ stand respectively for path tangent angle, path velocity magnitude.

$$v_i = \sqrt{\dot{x}_i^2 + \dot{y}_i^2} \quad and \quad \theta_i = arctan(\dot{y}_i, \dot{x}_i) \qquad (9)$$

$$and \dot{x}_i = x_i - x_{i-1} \ \ and \ \dot{y}_i = y_i - y_{i-1} \qquad (10)$$

A whitening linear transformation is finally applied to each discrete-time function so as to obtain zero mean and unit standard deviation function values. Seven dimensional feature vectors are used for GMM processing described in the following section. Figure 7 shows x-, y-, p- and velocity signals of an example signature.

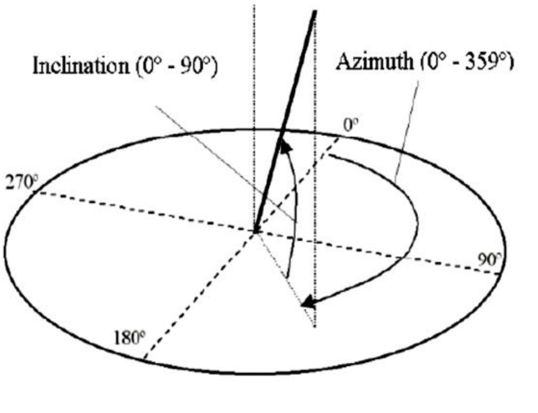

**Figure 6. Azimuth and Inclination Angles of the Pen Respect to the Plane of the Graphic Card
GD-0405U from Wacom Graphire3 Digitizing TabletPC**

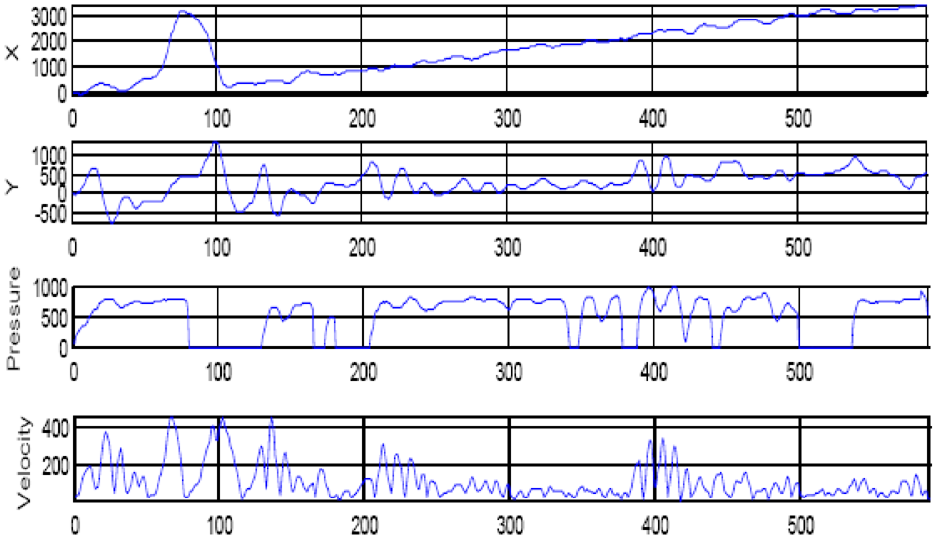

Figure 7. Signals (x-, y- position, pen pressure and velocity) of one Signature Fragment

3. Multimodal Biometric Fusion Decision

The process of biometric user authentication can be outlined by the following steps [17]: a) acquisition of raw data, b) extraction of features from these raw data, c) computing a score for the similarity or dissimilarity between these features and a previously given set of reference features and d) classification with respect to the score, using a threshold. The results of the decision processing steps are *true* or *false* (or *accept/reject*) for verification purposes or the user identity for identification scenarios.

The fusion of different signals can be performed 1) at the raw data or the feature level, 2) at the score level or 3) at the decision level. These different approaches have advantages and disadvantages. For *raw data* or *feature level* fusion, the basis data have to be compatible for all modalities and a common matching algorithm (processing step c) must be used. If these conditions are met, the separate feature vectors of the modalities easily could be concatenated into a single new vector. This level of fusion has the advantage that only one algorithm for further processing steps is necessary instead of one for each modality. Another advantage of fusing at this early stage of processing is that no information is lost by previous processing steps. The main disadvantage is the demand of compatibility of the different raw data of features. The fusion at *score level* is performed by computing a similarity or dissimilarity (distance) score for each single modality. For joining of these different scores, normalization should be done. The straightforward and most rigid approach for fusion is the decision level. Here, each biometric modality results in its own decision; in case of a verification scenario this is a set of *trues* and *falses*. From this set a kind of voting (majority decision) or a logical *AND* or *OR* decision can be computed. This level of fusion is the least powerful, due to the absence of much information. On the other hand, the advantage of this fusion strategy is the easiness and the guaranteed availability of all single modality decision results. In practice, score level fusion is the best-researched approach, which appears to result in better improvements of recognition accuracy as compared to the other strategies.

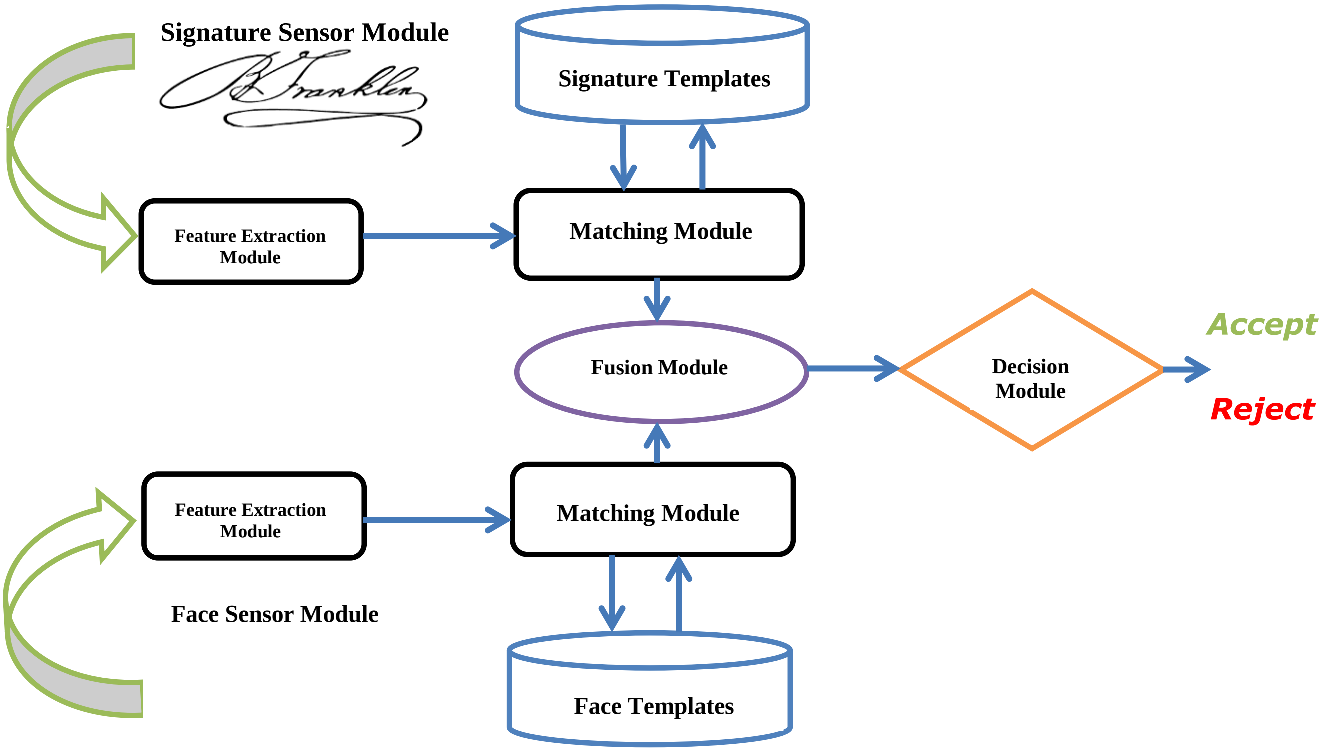

Figure 8. Score Fusion Based Multimodal Biometric Verification System

3.1 Adaptive Bayesian Method Based Score Fusion

Let $X = [X_1, X_2, \dots, X_K]$ denote the match scores of K different biometric matchers, where X_k is the random variable representing the match score of the k^{th}matcher, $k = 1, 2, \dots, K$. Let $f_{\text{gen}}(x)$ and $f_{\text{imp}}(x)$ be the conditional joint densities of the K match scores given the genuine and impostor classes, respectively, where $x = [x_1, x_2, \dots, x_K]$. Suppose we need to assign the observed match score vector X to genuine or impostor class. Let Ψ be a statistical test for testing H_0: *X corresponds to an impostor* against H_1: *X corresponds to a genuine user*. Let $\Psi(x) = i$ imply that we decide in favor of H_i, $i = 0, 1$. The probability of rejecting H_0 when H_0 is true is known as the *false accept rate* (*size* or *level* of the test). The probability of correctly rejecting H_0 when H_1 is true is known as the *genuine accept rate*. The Neyman-Pearson theorem [18][19] states that:

1) For testing H_0 against H_1, there exists a test Ψ and a constant η such that:

$$P(\Psi(X) = 1 | H_0) = \propto \quad (11)$$

and

$$\Psi(x) = \begin{cases} 1, & when \frac{f_{gen}(x)}{f_{imp}(x)} \geq \eta \\ 0, & when \frac{f_{gen}(x)}{f_{imp}(x)} \geq \eta \end{cases} \qquad (12)$$

2) If a test satisfies equations (11) and (12) for some η, then it is the *most powerful test* for testing H_0 against H_1 at level$\propto$.

According to the Neyman-Pearson theorem, given the false accept rate (FAR) $\propto$, the *optimal* test for deciding whether a score vector X corresponds to a genuine user or an impostor is the likelihood ratio test given by equation (12). *For a fixed FAR, it can select a threshold η such that the likelihood ratio test maximizes the genuine accept rate (GAR)*. Based on the Neyman-Pearson theorem, we are guaranteed that *there does not exist any other decision rule with a higher GAR*. However, this optimality of the likelihood ratio test is guaranteed only when the underlying densities are known. In practice, it estimate the densities $f_{gen}(x)$ and $f_{imp}(x)$ from the training set of genuine and impostor match scores, respectively and the performance of likelihood ratio test will depend on the accuracy of these estimates [17,20].

3.1.1 Estimation of Match Score Densities: Gaussian mixture model (GMM) has been successfully used to estimate arbitrary densities and it is used for estimating the genuine and impostor score densities [21,22].

Let $\Phi^K(x;\ \mu, \sum)$ be the K-variate Gaussian density with mean vector µand covariance matrix $\sum$, i.e.,

$\Phi^K(x;\ \mu, \sum) = (2\pi)^{-K/2}|\sum|^{-1/2}\, exp\left(-\frac{1}{2}(x-\mu)^T {\sum}^{-1}(x-\mu)\right)$. The estimates of $f_{gen}(x)$ and $f_{imp}(x)$ are obtained as a mixture of Gaussians as follows.

$$\hat{f}_{gen}(x) = \sum_{j=1}^{M_{gen}} P_{gen,j}\Phi^K\left(x;\ \mu_{gen,j}\ ,{\sum}_{gen,j}\right) \tag{13}$$

$$\hat{f}_{imp}(x) = \sum_{j=1}^{M_{imp}} P_{imp,j}\Phi^K\left(x;\ \mu_{imp,j}\ ,{\sum}_{imp,j}\right) \tag{14}$$

Where $M_{gen}(M_{imp})$ is the number of mixture components used to model the density of the genuine (impostor) scores, $p_{gen,j}$ $(p_{imp,j})$ is the weight assigned to the j^{th} mixture component in $\hat{f}_{imp}(x)\left(\hat{f}_{imp}(x)\right), \sum_{j=1}^{M_{gen}} P_{gen,j} = \sum_{j=1}^{M_{imp}} P_{imp,j} = 1$. Selecting the appropriate number of components is one of the most challenging issues in mixture density estimation; while a mixture with too many components may result in over-fitting, a mixture with too few components may not approximate the true density well. The GMM fitting algorithm automatically estimates the number of components and the component parameters using an EMalgorithm and the minimum message length criterion [21-22].

Maximum Likelihood Parameter Estimation: Given a set of observation data in a matrix X and a set of observation parameters θ the ML parameter estimation aims at maximizing the likelihood $L(\theta)$or log likelihood of the observation data $X = \{X_1, \dots, X_n\}$

$$\hat{\theta} = arg \max_{\theta} L(\theta).(15)$$

Assuming that it has independent, identically distributed data, it can write the above equations as:

$$L(\theta) = p(X|\theta) = p(X_1, \dots, X_n|\theta) = \prod_{i=1}^{n} p(X_i|\theta). \tag{16}$$

The maximum for this function can be find by taking the derivative and set it equal to zero, assuming an analytical function.

$$\frac{\partial}{\partial\theta} L(\theta) = 0. \tag{17}$$

The incomplete-data log-likelihood of the data for the mixture model is given by:

$$L(\theta) = log(X|\theta) = \sum_{i=1}^{N} log(x_i|\theta) \tag{18}$$

Which is difficult to optimize because it contains the log of the sum. If it considers X as incomplete, however, and posits the existence of unobserved data items $Y = \{y_i\}_{i=1}^{N}$whose values inform us which component densitygenerated each data item, the likelihood expression is significantly simplified. That is, it assume that $y_i \in \{1\,..\,K\}$ for each i, and $y_i = k$if the i-th sample was generated by the k-th mixture component. If it knows the values of Y, it obtains the complete-data log-likelihood, given by:

$$L(\theta, Y) = \log p(X, Y|\theta) \tag{19}$$

$$= \sum_{i=1}^{N} \log p(x_i, y_i|\theta) \tag{20}$$

$$= \sum_{i=1}^{N} \log\left(p(y_i|\theta)p(x_i|y_i, \theta)\right) \tag{21}$$

$$= \sum_{i=1}^{N}\left(\log\, p_{y_i} + \log g\left(x_i\middle|\mu_{y_i}, {\sum}_{y_i}\right)\right) \tag{22}$$

Which, given a particular form of the component densities, can be optimized using a variety of techniques [8].

EM Algorithm: The expectation-maximization (EM) algorithm [17-19, 22] is a procedure for maximum-likelihood (ML) estimation in the cases where a closed form expression for the optimal parameters is hard to obtain. This iterative algorithm guarantees the monotonic increase in the likelihood L when the algorithm is run on the same training database.

The probability density of the Gaussian mixture of k components in $\mathbb{R}^d$ can be described as follows:

$$\Phi(x) = \sum_{i=1}^{N} \pi_i \, \emptyset(x|\theta_i) \quad \forall x \in \mathbb{R}^d \, , \tag{23}$$

Where$\emptyset(x|\theta_i)$ is a Gaussian probability density with the parameters $\theta_i = (m_i, \sum_i)$, m_i is the mean vector and $\sum_i$ is the covariance matrix which is assumed positive definite given by:

$$\emptyset(x|\theta_i) = \emptyset(x|m_i, \sum_i) = \frac{1}{(2\pi)^{\frac{n}{2}}|\sum_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-m_i)^T \sum_i^{-1}(x-m_i)} \, , \tag{24}$$

And $\pi_i \in [0, \; 1](i = 1,2,\ldots,k)$ are the mixing proportions under the constraint $\sum_{i=1}^{k} \pi_i = 1$. If it encapsulate all the parameters into one vector: $\Theta_k = (\pi_1, \pi_2, \ldots, \pi_k, \theta_1, \theta_2, \ldots, \theta_k)$,then, according to Eq. (24), the density of Gaussian mixture can be rewritten as:

$$\Phi(x|\Theta_k) = \sum_{i=1}^{k} \pi_i \emptyset(x|\theta_i) = \sum_{i=1}^{k} \pi_i \emptyset(x|m_i, \sum_i). \tag{25}$$

For the Gaussian mixture modeling, there are many learning algorithms. But the EM algorithm may be the most well-known one. By alternatively implementing the E-step to estimate the probability distribution of the unobservable random variable and the M-step to increase the log-likelihood function, the EM algorithm can finally lead to a local maximum of the log-likelihood function of the model. For the Gaussian mixture model, given a sample data set $S = \{x_1, x_2, \cdot\cdot\cdot, x_N\}$as a special incomplete data set, the log-likelihood function can be expressed as follows:

$$\log p(S|\Theta_k) = \log \prod_{t=1}^{N} \emptyset(x_t|\Theta_k) = \sum_{t=1}^{N} \log \sum_{i=1}^{k} \pi_i \emptyset(x_t|\theta_i), \tag{26}$$

Which can be optimized iteratively via the EM algorithm as follows:

$$P(j|x_t) = \frac{\pi_j \emptyset(x_t|\theta_j)}{\sum_{i=1}^{k} \pi_i \emptyset(x_t|\theta_i)}, \tag{27}$$

$$\pi_j^{+} = \frac{1}{N} \sum_{t=1}^{N} P(j|x_t), \tag{28}$$

$$\mu_j^{+} = \frac{1}{\sum_{t=1}^{N} P(j|x_t)} \sum_{t=1}^{N} P(j|x_t) x_t, \tag{29}$$

$$\sum\nolimits_j^{+} = \frac{1}{\sum_{t=1}^{N} P(j|x_t)} \sum_{t=1}^{N} P(j|x_t)\left(x_t - \mu_j^{+}\right)\left(x_t - \mu_j^{+}\right)^T. \tag{30}$$

Although the EM algorithm can have some good convergence properties in certain situations, it certainly has no ability to determine the proper number of the components for a sample data set because it is based on the maximization of the likelihood.

4. Experiments and Results

The experiments were performed using a signature and still face database extracted from video, which is encoded in raw UYVY. AVI 640 x 480, 15.00 fps with uncompressed 16bit PCM audio; mono, 32000 Hz little endian. Uncompressed PNG files are extracted from the video files for feeding the face detection algorithms. Thirty subjects were used for the experiments in which twenty-six are males and four are females. For each subject, 30 signatures (with dat header) are used. Each line of a (.dat files) consists of four comma separated integer values for the sampled x- and y-position of the pen tip, the pen pressure and the timestamp (in ms); the lines with values of -1 for x, y and pressure represent a pen-up/pen-down event; The device used for recording the handwriting data was a Wacom Graphire3 digitizing tablet. Size of sensing surface is

127.6mm x 92.8mm. With spatial resolution of 2032 lpi (lines per inch), able to measure 512 degrees of pressure. The signature data is acquired with a non-fixed sampling rate of about 100Hz. The database obtained from eNTERFACE 2005 [23]. Thirty subjects were used for the experiments in which twenty-five are males and five are females. For face experts, ninety-six face images from a subject were randomly selected to be trained and projected into Eigen space, and the other twenty-four samples were used for the subsequent validation and testing. Similarly, twenty four signatures from a subject were randomly selected for training, and the other six samples were used for the subsequent validation and testing. Three sessions of the face database and signature database were used separately. Session one was used for training the signature and face experts. Eachexpert used ten mixture client models. To find the performance, Sessionstwo and three were used for obtaining expert opinions of known impostor and true claims.

Performance Criteria: The basic error measure of a verification system is false rejection rate (FRR) and false acceptance rate (FAR) as defined in the following equations:

False Rejection Rate (**FRR***i*): is an average of number of falsely rejected transactions. If ***n*** is a transaction and x(n) is the verification result where 1 is falsely rejected and 0 is accepted and ***N*** is the total number of transactions then the personal False Rejection Rate for user ***i*** is

$$FRR_i = \frac{1}{N}\sum_{n=1}^{N} x(n) \qquad (31)$$

False Acceptance rate (**FAR***i*) is an average of number of falsely accepted transactions. If ***n*** is a transaction and x(n) is the verification result where 1 is a falsely accepted transaction and 0 is genuinely accepted transaction and ***N*** is the total number of transactions then the personal False Acceptance Rate for user ***i*** is

$$FAR_i = \frac{1}{N}\sum_{n=1}^{N} x(n) \qquad (32)$$

Both FRR*i* and FAR*i* are usually calculated as averages over an entire population in a test. If P is the size of populations then these averages are

$$FRR = \frac{1}{P}\sum_{i}^{P} FRR_i \qquad (33)$$

$$FAR = \frac{1}{P}\sum_{i}^{P} FAR_i \qquad (34)$$

Equal Error Rate (**EER**), is an intersection where FAR and FRR are equal at an optimal threshold value. This threshold value shows where the system performs at its best.

As a common starting point, classifier parameters were selected to obtain performance as close as possibleto EER on clean test data (following the standard practice in theface and speaker verification area of using EER as a measure of expectedperformance). A good decision is to choose the decision threshold such as the false accept equal to the false reject rate. In this paper it uses the Detection Error Tradeoff (DET) curve to visualize and compare the performance of the system (see Figure 9).

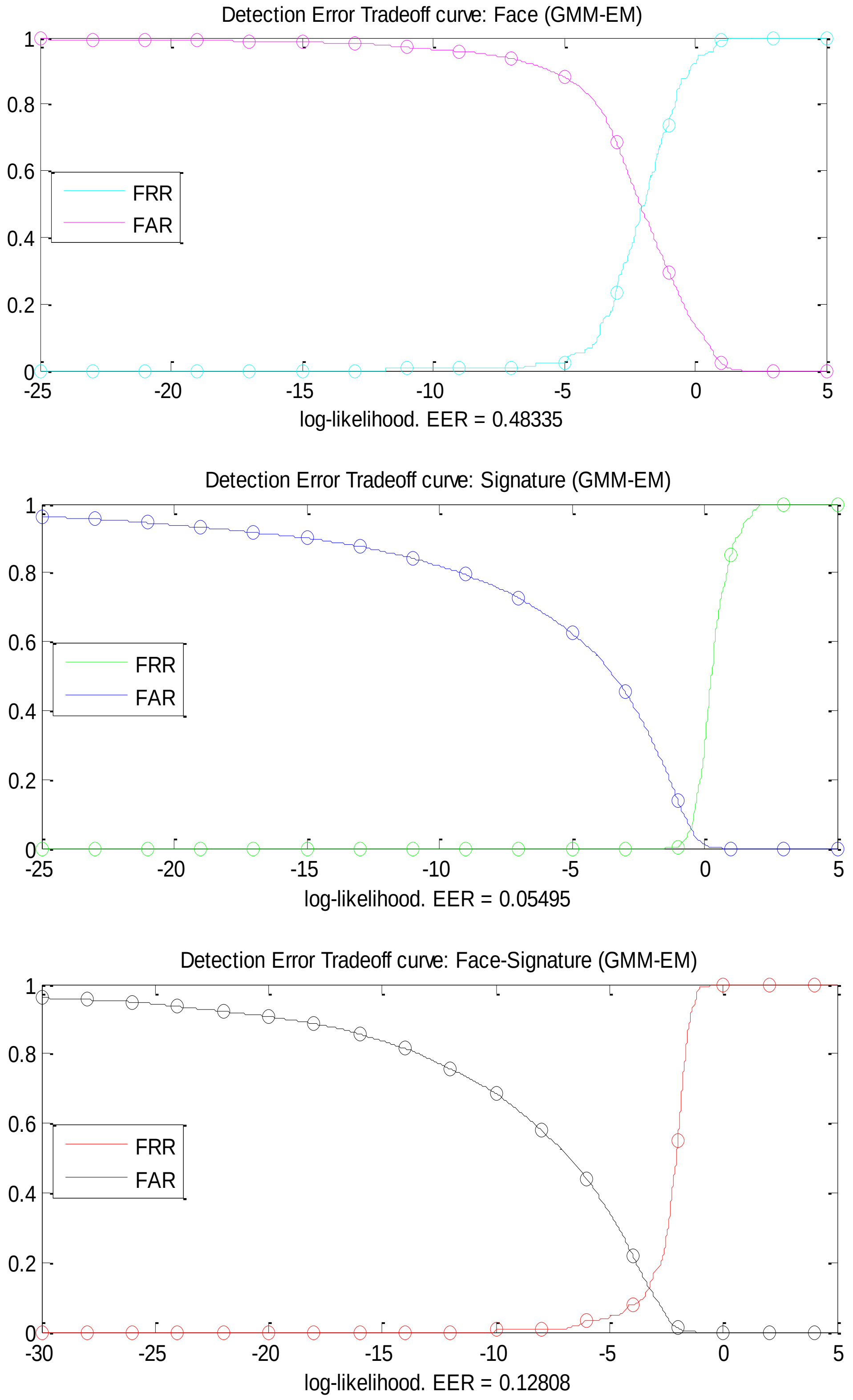

Figure 9. Detection Error Tradeoff Curves

5. Discussion and Conclusions

The question (?) that always arises is how it can obtain the “best” (in terms of accuracy); the solution is fusion. It is well known that the “best” fusion solution is one that satisfies the Neyman-Pearson (NP) criterion [25]. The NP criterion characterizes the fusion solution that maximizes the True Accept Rate (TAR) for a fixed value of False Accept Rate (FAR) [18,24,25].

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) [25]. 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 [25,26]. 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, EER denotes Equal Error Rate and it’s the point where (FAR=FRR). This threshold must be calculated afresh for each application, to adapt it to the specific population concerned [25,26]. This is done in general using a small database recorded for this purpose.

Performance capabilities have been traditionally shown in the form of ROC (receiver- or relative-operating characteristic) plots [25], in which the probability of a false-acceptance is plotted versus the probability of a false-rejection for varying decision thresholds. 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 [25].

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 [25]. 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 [25]. 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 [25,26].

This paper has presented a human authentication method combined dynamic face and on-line signature information in order to improve the problem of single biometric authentication, since single biometric authentication has the fundamental problems of high False Accept Rate (FAR) and False Reject Rate (FRR). It has presented a framework for fusion of match scores in multi-modal biometric system based on adaptive Bayesian method. The product of likelihood ratio based fusion rule with (GMM-EM) based density estimation achieves a significant recognition rates. As a result presented a combined authentication method can provide a stable authentication rate and it overcomes the limitation of a single mode system. Based on the experimental results, it has shown that EER can be reduced down significantly between the dynamic face and on-line signature mode and a combined face-signature mode.

References

[1] S. Gleni and P. Petratos, "DNA Smart Card for Financial Transactions", The ACM Student Magazine **(2004)**.

[2] G. Chetty and M. Wagner, "Audio-Visual Multimodal Fusion for Biometric Person Authentication and Liveness Verification", Australian Computer Society, Inc. This paper appeared at the NICTA-HCSNet Multimodal UserInteraction Workshop, **(2006)**; Sydney, Australia.

[3] N. Poh and S. Bengio, "Database, Protocol and Tools for Evaluating Score-Level Fusion Algorithms in Biometric Authentication", IDIAP RR 04-44, **(2004)**.

[4] S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, "Fusion of Face and Speech Data for Person Identity Verification", IEEE Transactions on Neutral Networks, vol.10, no.5, **(1999)**.

[5] H. J. Korves, L. D. Nadel, B. T. Ulery and D. M. B. Masi, "Multi-biometric Fusion: From Research to Operations", MTS MitreTek Systems, sigma summar, **(2005)**, pp.39-48.

[6] K. Nandakumar, Y. Chen, S. C. Dass and A. K. Jain, "Biometric Score Fusion: Likelihood Ratio, Matcher Correlation and Image Quality", **(2007)**.

[7] A. Jain, K. Nandakumar and A. Ross, "Score normalization in multimodal biometric systems", THE Journal of Pattern Recognition Society, (**2005)**.

[8] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas, "On combining classifiers", IEEETransactions on Pattern Analysis and Machine Intelligence, vol.20, no.3, **(1998)**, pp.226–239.

[9] 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)**, pp.450-455.

[10] A. Teoh, S. A. Samad and A. Hussain, "Nearest Neighbourhood Classifiers in Biometric Fusion", International Journal of The Computer, the internet and management, vol.12, no.1, **(2004)**, pp.23-36.

[11] Y. Zana, R. M. Cesar-Jr, R. S. Feris, M. Turk and G. Bebis et al., "Face Verification in Polar Frequency Domain: A Biologically Motivated Approach", ISVC 2005, LNCS 3804, Springer-Verlag Berlin Heidelberg, **(2005)**, pp.183–190.

[12] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol.3, no.1, **(1991)**, pp.71-86.

[13] M. Soltane, N. Doghmane and N. Guersi, "State of the Art: Signature Biometrics Verification", BRAIN. Broad Research in Artificial Intelligence and Neuroscience, vol.1, no.2, **(2010)**.

[14] C. Vielhauer, S. Schimke, V. Thanassis, Y. Stylianou, R. Creutzburg and J. H. Takala, "Fusion Strategies for Speech and Handwriting Modalities in HCI", Multimedia on Mobile Devices, Proc. of SPIE-IS&T Electronic Imaging, **(2005)**.

[15] I. S. I. Abuhaiba, "Offline Signature Verification Using Graph Matching", Turk J Elec Engin, Department of Electrical and Computer Engineering, Islamic University of Gaza, PALESTINE, vol.15, no.1, **(2007)**.

[16] J. Richiardi, J. Fierrez-Aguilar, J. Ortiga-Garcia and A. Drygajlo, "On-line signature verification resilience to packet loss in IP networks", second COST 275 WORKSHOP Biometrics on the Internet: Fundamentals, Advances and Applications, **(2004)**; University of Vigo, Vigo-Spain.

[17] K. Veeramachaneni, L. A. Osadciw and P. K. Varshney, "An Adaptive Multimodal Biometric Management Algorithm", IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews, vol.35, no.3, **(2005)**.

[18] V. Trees and L. Harry, "Detection, Estimation, and Modulation Theory", Part I, John Wiley and Sons, **(1968)**.

[19] Q. Yan and R. S. Blum, "Distributed Signal Detection under the Neyman-Pearson Criterion", IEEE Transactions on Informations Theory, vol.47, no.4, **(2001)**.

[20] K. Nandakumar, Y. Chen, S. C. Dassand and A. K. Jain, "Likelihood Ratio Based Biometric Score Fusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, **(2007)**.

[21] P. Paalanen, "Bayesian classification using gaussian mixcute model and EM estimation: implementation and comparisons", Information Technology Project, Lappeenranta, **(2004)**.

[22] P. Paalanen, J.-K. Kamarainen, J. Ilonen and H. Kälviäinen, "Feature Representation and Discrimination Based on Gaussian Mixture Model Probability Densities: Practices and Algorithms", Department of Information Technology, Lappeenranta University of Technology, P.O.Box 20, FI-53851 Lappeenranta, Finland, **(2005)**.

[23] Y. Stylianou, Y. Pantazis, F. Calderero, P. Larroy, F. Severin, S. Schimke, R.o Bonal, F. Matta and A. Valsamakis, "GMM-Based Multimodal Biometric Verification", eNTERFACE 2005 The summer Workshop on Multimodal Interfaces, **(2005)**; Facultè Polytechnique de Mons, Belgium.

[24] H. J. Korves, L. D. Nadel, B. T. Ulery and D. M. B. Masi, "Multi-biometric Fusion: From Research to Operations", MitreTek Systems, sigma summar, **(2005)**, pp.39-48.

[25] M. Soltane and B. Mimen, "Multi-modal Biometric Authentications: Concept Issues and Applications Strategies", International Journal of Advanced Science and Technology, vol.49, **(2012)**.

[26] 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)**.

Author

Mohamed Soltane (Assoc. Prof. Dr.), he 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 microcomputer based system design.