

Adapted User-Dependent Multimodal Biometric Authentication Exploiting General Information

Julian Fierrez-Aguilar ^{a,*}, Daniel Garcia-Romero ^a,
Javier Ortega-Garcia ^a, Joaquin Gonzalez-Rodriguez ^a

^a*Escuela Politecnica Superior, Universidad Autonoma de Madrid,
Ctra. Colmenar km. 15, E-28049 Madrid, Spain*

Abstract

A novel adapted strategy for combining general and user-dependent knowledge at the decision-level in multimodal biometric authentication is presented. User-independent, user-dependent, and adapted fusion and decision schemes are compared by using a bimodal system based on fingerprint and written signature. The adapted approach is shown to outperform the other strategies considered in this paper. Exploiting available information for training the fusion function is also shown to be better than using existing information for post-fusion trained decisions.

Key words: Biometrics, multimodal, authentication, verification, user, local, global, support vector machine, fingerprint, signature

* Corresponding author. Tel.: +34-91-4973363; fax: +34-91-4972235

Email addresses: Julian.Fierrez@uam.es (Julian Fierrez-Aguilar),
Javier.Ortega@uam.es (Javier Ortega-Garcia), Joaquin.Gonzalez@uam.es

1 Introduction

The basic aim of biometrics (Bolle et al., 2004a) is to discriminate among subjects –in a reliable way and according to some target application– based on one or more signals derived from physical or behavioral traits, such as fingerprint, face, iris, voice, hand, or written signature. Authentication systems built upon only one of the above modalities may not fulfill the requirements of demanding applications in terms of universality, uniqueness, permanence, collectability, performance, acceptability, and circumvention. This has motivated the current interest in multimodal biometrics, in which several biometric traits are simultaneously used in order to make an identification decision (Maltoni et al., 2003; Jain et al., 2004).

A common practice in most of the reported works on multimodal biometrics is to combine the matching scores obtained from the unimodal systems by using simple rules (e.g., sum, product), statistical methods, or machine learning procedures (Brunelli and Falavigna, 1995; Bigun et al., 1997; Kittler et al., 1998; Hong and Jain, 1998; Ben-Yacoub et al., 1999; Chatzis et al., 1999; Verlinde et al., 2000). A remarkable characteristic of this approach, as compared to the feature-level combination techniques, is the possibility of designing structured multimodal systems by using existing unimodal recognition strategies (Maltoni et al., 2003). This multiple matcher approach is interesting not only for biometrics, but also for other pattern recognition areas (Jain et al., 2000; Roli et al., 2004).

In all the works referenced above, the fusion algorithms worked independently

(Joaquin Gonzalez-Rodriguez).

of the claimed identity (also referred to as general or global approaches hereafter). Recently, new research efforts have focused on user-dependent (also referred to as specific or local hereafter) score fusion schemes (Jain and Ross, 2002; Fierrez-Aguilar et al., 2003; Kumar and Zhang, 2003; Indovina et al., 2003; Fierrez-Aguilar et al., 2004; Wang et al., 2004; Toh et al., 2004). The basic aim of this approach is to cope with the fact that some traits do not work properly with some subjects for recognition purposes even though these traits can be highly discriminant among other subjects. This asseveration has been corroborated experimentally in a number of works. As an example, about 4% of the population have poor quality fingerprints that cannot be easily imaged by some of the existing sensors (Jain and Ross, 2004). Also, a number of speakers, the so-called *lambs* (Doddington et al., 1998), tend to have high individual speaker recognition error rate. This fact has also been pointed out regarding signature verification (Fierrez-Aguilar et al., 2005a).

In the present work, operational procedures exploiting user dependencies for multimodal biometrics are presented and evaluated on data from the MCYT bimodal corpus (Ortega-Garcia et al., 2003) using a non-biased experimental setup based on bootstrap sampling (Bolle et al., 2004b). Moreover, a novel adapted user-dependent strategy is introduced. The proposed technique is shown to overcome the severe training data scarcity problem commonly encountered in user-specific learning scenarios.

This paper is organized as follows. A detailed look at related work and the motivation for the proposed adapted user-specific fusion scheme is described in Section 2. The proposed approach is presented in Section 3. The baseline biometric systems based on fingerprint and on-line signature traits used in the bimodal experiments are introduced in Section 4. Experimental protocol and

results demonstrating the benefits of the proposed approach are reported in Section 5. Conclusions are finally drawn in Section 6.

2 Related work and motivation

The idea of exploiting user-specific parameters at the decision-level in multimodal biometrics has been studied by Jain and Ross (2002). In this preceding work, user-independent weighted linear combination of similarity scores was demonstrated to be improved by using either user-dependent weights or user-dependent decision thresholds, both of them computed by exhaustive search on testing data. Subsequently, a trained user-dependent scheme using Support Vector Machines (SVM) was presented by Fierrez-Aguilar et al. (2003) and evaluated using leave-one-out error estimates. The idea of Jain and Ross (2002) was also explored by Wang et al. (2004) using non-biased error estimation procedures. Other attempts to localized multimodal biometrics include the use of the claimed identity index as a feature for a global trained fusion scheme based on Neural Networks (Kumar and Zhang, 2003), computing user-dependent weights using *lambness* metrics (Indovina et al., 2003), and using personalized Fisher ratios (Poh and Bengio, 2005).

Toh et al. (2004) have recently proposed a taxonomy of decision-level approaches for multibiometrics. Existing multimodal fusion approaches are classified as global or local depending firstly on the fusion function (i.e., user-independent or user-dependent fusion strategies) and secondly on the decision making process (i.e., user-independent or user-dependent decision thresholds). Examples are global-learning-global-decision (GG) (Brunelli and Falavigna, 1995; Bigun et al., 1997; Kittler et al., 1998; Hong and Jain, 1998; Ben-Yacoub

et al., 1999; Chatzis et al., 1999; Verlinde et al., 2000), local-learning-global-decision (LG) (Jain and Ross, 2002; Fierrez-Aguilar et al., 2003; Kumar and Zhang, 2003; Indovina et al., 2003; Fierrez-Aguilar et al., 2004; Wang et al., 2004; Toh et al., 2004; Poh and Bengio, 2005), and similarly global-learning-local-decision (GL) (Jain and Ross, 2002; Toh et al., 2004), and local-learning-local-decision (LL) (Toh et al., 2004). In the present work we adhere to this taxonomy and extend it by incorporating new items: adapted-learning and adapted-decisions.

The use of general information in user-dependent fusion schemes has recently been introduced by Fierrez-Aguilar et al. (2004). In this case a computationally demanding batch SVM learning procedure was used. The focus of the present paper is to extend this preceding work by simplifying the batch training procedure and to compare the proposed method with existing approaches.

The idea of adapted learning is based on the fact that the amount of available training data in localized learning is usually not sufficient and representative enough to guarantee good parameter estimation/learning and generalization capabilities. To cope with this lack of robustness derived from partial knowledge of the problem structure, the use of robust adaptive learning/decision strategies based on “all” the available information has been proposed in related research areas (Lee and Huo, 2000). As an example of the underlying philosophy, we exploit the fact that general information of the problem (such as user-independent data) can constitute a rich source of information for user-specific recognition problems. In general, the relative balance between the prior knowledge (global) and the empirical data (local) is performed as a trade-off between both kinds of information.

Based on the related work and the above mentioned ideas, the aim of this paper is to develop an adapted-learning-global-decision (AG) fusion method incorporating the general knowledge available from pooling user-independent data. A counterpart global-learning-adapted-decision (GA) method is also introduced, using the same learning paradigm and amount of training data. The proposed methods are compared with existing procedures using a non-biased experimental setup on real multimodal biometric data.

3 Exploiting user specificities at the decision-level in multimodal biometrics

The proposed adapted local fusion scheme is derived from user-independent and user-dependent fusion strategies (Fierrez-Aguilar et al., 2003) based on SVM classifiers (Theodoridis and Koutroumbas, 2003). Firstly, the notation is established and a summary of SVM-based score fusion is provided. Global, local, and adapted fusion schemes are also described. Finally, global, local, and adapted decision making approaches are introduced for their use with the combined scores. The system model of multimodal biometric verification including global/local/adapted learning/decisions is depicted in Fig. 1.

3.1 Score-level multimodal fusion based on SVMs

Given a multimodal biometric verification system consisting of R different unimodal systems $r = 1, \dots, R$, each one computes a similarity score $x_r \in \mathbb{R}$ between an input biometric pattern and the enrolled pattern of the given claimant. Let the similarity scores, provided by the different unimodal systems,

be combined into a multimodal score $\mathbf{x} = [x_1, \dots, x_R]^T$. The design of a trained fusion scheme consists in the estimation of a function $f : \mathbb{R}^R \rightarrow \mathbb{R}$, based on empirical data, so as to maximize the separability of client $\{f(\mathbf{x})|\text{client attempt}\}$ and impostor $\{f(\mathbf{x})|\text{impostor attempt}\}$ fused score distributions.

Formally, let the training set be $X = (\mathbf{x}_i, y_i)_{i=1}^N$ where N is the number of multimodal scores in the training set, and $y_i \in \{-1, 1\} = \{\text{Impostor}, \text{Client}\}$. The principle of SVM relies on a linear separation in a high dimension feature space \mathbb{H} where the data have previously been mapped via $\Phi : \mathbb{R}^R \rightarrow \mathbb{H}; X \rightarrow \Phi(X)$, so as to take into account the eventual non-linearities of the problem (Vapnik, 2000). In order to achieve a good level of generalization capability, the margin between the separator hyperplane

$$\{\mathbf{h} \in \mathbb{H} | \langle \mathbf{w}, \mathbf{h} \rangle_{\mathbb{H}} + w_0 = 0\} \quad (1)$$

and the mapped data $\Phi(X)$ is maximized (where $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ denotes inner product in space \mathbb{H} , and $(\mathbf{w} \in \mathbb{H}, w_0 \in \mathbb{R})$ are the parameters of the hyperplane). The optimal hyperplane can be obtained as the solution of the following quadratic programming problem (Vapnik, 2000):

$$\min_{\mathbf{w}, w_0, \xi_1, \dots, \xi_N} \left(\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \right) \quad (2)$$

subject to

$$\begin{aligned} y_i (\langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle_{\mathbb{H}} + w_0) &\geq 1 - \xi_i, & i = 1, \dots, N \\ \xi_i &\geq 0, & i = 1, \dots, N \end{aligned} \quad (3)$$

where slack variables ξ_i are introduced to take into account the eventual non-

separability of $\Phi(X)$ and parameter C is a positive constant that controls the relative influence of the two competing terms.

The optimization problem in (2), (3) is typically solved using the Wolfe dual representation using the kernel trick (Theodoridis and Koutroumbas, 2003), i.e., the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle_{\mathbb{H}}$ is introduced avoiding direct manipulation of the elements of \mathbb{H} . In particular, a Radial Basis Function (RBF) kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2) \quad (4)$$

is used in this work. Other kernel choices used for multimodal biometrics include polynomial (Ben-Yacoub et al., 1999) and linear (Fierrez-Aguilar et al., 2005b) kernels.

The fused score s_T of a multimodal test pattern \mathbf{x}_T is defined as follows (Fierrez-Aguilar et al., 2003)

$$s_T = f(\mathbf{x}_T) = \langle \mathbf{w}, \Phi(\mathbf{x}_T) \rangle_{\mathbb{H}} + w_0 \quad (5)$$

which is a signed distance measure from \mathbf{x}_T to the separating surface given by the solution of the SVM problem.

As a result, the training procedure in (2), (3) and the fusion strategy in (5) are obtained for the problem of multimodal fusion.

3.2 Global, local and adapted fusion schemes

Global learning The training set $X_G = (\mathbf{x}_i, y_i)_{i=1}^{N_G}$ includes multimodal scores from a number of different clients and the resulting fusion rule $f_G(\mathbf{x})$ is applied globally at the operational stage regardless of the claimed identity.

Local learning A different fusion rule $f_{j,L}(\mathbf{x})$ is obtained for each client enrolled in the system $j = 1, \dots, M$ by using development scores X_j of the specific client j . At the operational stage, the fusion rule $f_{j,L}(\mathbf{x})$ of the claimed identity j is applied.

Adapted learning An adapted user-dependent fusion scheme is proposed trading off the general knowledge provided by the user-independent training set X_G , and the user specificities provided by the user-dependent training set X_j . To obtain the adapted fusion rule, $f_{j,A}(\mathbf{x})$, for user j , we propose to train both the global fusion rule, $f_G(\mathbf{x})$, and the local fusion rule, $f_{j,L}(\mathbf{x})$, as described above, and finally combine them as follows:

$$f_{j,A}(\mathbf{x}) = \alpha f_{j,L}(\mathbf{x}) + (1 - \alpha) f_G(\mathbf{x}) \quad (6)$$

where α is a trade-off parameter. This can be seen as a user-dependent fusion scheme adapted from user-independent information. The idea can also be extended easily to trained fusion schemes based on other classifiers. Worth noting, sequential algorithms to solve the SVM optimization problem in (2), (3) have already been proposed (Navia-Vazquez et al., 2001), and can be used to extend the proposed idea, first constructing the user-independent solution and then refining it by incorporating the local data.

3.3 Global, local and adapted decisions

Once a combined similarity score has been obtained using either a local or a global fusion function, it is compared to a decision threshold in order to accept/reject the identity claim being made. This decision making process can also be made locally or globally.

Global decision. Let the training set be $S_G = (s_i, y_i)_{i=1}^{N_G}$ be a set of labelled fused scores from a pool of known users. The decision rule

$$d_G(s) \begin{cases} > 0 \rightarrow \text{accepted} \\ \leq 0 \rightarrow \text{rejected} \end{cases} \quad (7)$$

is trained by using a 1 dimensional SVM as described in Section 3.1.

Local decision. A different decision function is used for each client enrolled in the system $j = 1, \dots, M$. Each function is trained by using a development set of fused scores of the specific client. At the operational stage, the decision function $d_{j,L}(s)$ of the client j being claimed is applied.

Adapted decision. An adapted decision criterion $d_{j,A}(s)$ is built similarly to Eq. 6 as follows

$$d_{j,A}(s) = \alpha d_{j,L}(s) + (1 - \alpha) d_G(s) \quad (8)$$

4 Baseline monomodal systems

Individual verification systems with standard performance have intentionally been used to make the comparison of subsequent fusion strategies easier. In

particular, the experiments have been carried out on our bimodal biometric verification system including the minutiae-based fingerprint verification subsystem described by Simon-Zorita et al. (2003) and the on-line signature verification subsystem based on temporal functions and Hidden Markov Models reported by Fierrez-Aguilar et al. (2005a). A brief description of both systems is given below.

4.1 Fingerprint recognition system

Image enhancement. The fingerprint ridge structure is reconstructed by using: *i*) grayscale level normalization, *ii*) orientation field calculation *iii*) interest region extraction, *iv*) spatial-variant filtering according to the estimated orientation field, *v*) binarization, and *vi*) ridge profiling.

Feature extraction. The minutiae pattern is obtained from the binarized profiled image as follows: *i*) thinning, *ii*) removal of structure imperfections from the thinned image, and *iii*) minutiae extraction. For each detected minutia, the following parameters are stored: *a*) the x and y coordinates of the minutia, *b*) the orientation angle of the ridge containing the minutia, and *c*) the x and y coordinates of 10 samples of the ridge segment containing the minutia. An example fingerprint image is shown in Fig. 2 together with the feature extraction steps.

Pattern comparison. Given a test and a reference minutiae pattern, a matching score x'_{finger} is computed. First, both patterns are aligned based on the minutia whose associated sampled ridge is most similar. The matching score is computed then by using a variant of the edit distance on polar coordinates and based on a size-adaptive tolerance box. When more than one reference

minutiae pattern per client model are considered, the maximum matching score obtained by comparing the test and each reference pattern is used.

Score normalization. In order to generate a similarity score x_{finger} between 0 and 1, the matching score x'_{finger} (greater than or equal to zero) is further normalized according to

$$x_{\text{finger}} = \tanh\left(c_{\text{finger}} \cdot x'_{\text{finger}}\right) \quad (9)$$

The parameter c_{finger} has been chosen heuristically on fingerprint data not used for the experiments reported here.

4.2 Signature recognition system

Feature extraction. Coordinate trajectories $(x[n], y[n])$, $n = 1, \dots, N_s$ and pressure signal $p[n]$, $n = 1, \dots, N_s$, are considered in the feature extraction process, where N_s is the duration of the signature in time samples (sampling frequency = 100 Hz.). Signature trajectories are first preprocessed by subtracting the center of mass followed by a rotation alignment based on the average path tangent angle. An extended set of discrete-time functions are derived from the preprocessed trajectories. As a result, the signature is parameterized as the following set of 7 discrete-time functions $\{x[n], y[n], p[n], \theta[n], v[n], \rho[n], a[n]\}$, $n = 1, \dots, N_s$, and first order time derivatives of all of them (θ , v , ρ and a stand respectively for path tangent angle, path velocity magnitude, log curvature radius and total acceleration magnitude). A linear transformation is finally applied to each discrete-time function so as to obtain zero mean and unit standard deviation function values.

Similarity computation. Given the parameterized enrollment set of signa-

tures of a client j , a left-to-right Hidden Markov Model λ_j is estimated. No transition skips between states are allowed and multivariate Gaussian Mixture density observations are used. On the other hand, given a test signature P (with a duration of N_s time samples) and a claimed identity j modelled as λ_j , the similarity matching score

$$x'_{\text{sign}} = \frac{1}{N_s} \log p(P|\lambda_j) \quad (10)$$

is obtained through Viterbi alignment of the test signature with the HMM (Theodoridis and Koutroumbas, 2003).

Score normalization. In order to generate a similarity score x_{sign} between 0 and 1, the matching score x'_{sign} (less than or equal to zero) is further normalized according to

$$x_{\text{sign}} = \exp\left(c_{\text{sign}} \cdot x'_{\text{sign}}\right) \quad (11)$$

The parameter c_{sign} has been chosen heuristically on signature data not used for the experiments reported here.

The processing stages are shown graphically for an example signature in Fig. 3.

5 Experiments

The problem in (2), (3) is solved in its dual representation by using the decomposition algorithm proposed by Osuna et al. (1997), and the interior point optimization solver proposed by Vandervei (1999). Main SVM parameters are as follows: $C = 100$ for client scores, $C = 50$ for impostor scores, and $\sigma = 0.05$.

5.1 Database description

In our experiments we use 10 samples of one finger and 17 signatures of each of the first 75 subjects from the MCYT biometric database (Ortega-Garcia et al., 2003).

In order to highlight the benefits of the proposed approaches in an scenario showing user-dependencies, lowest quality finger was used for 10% of the users and highest quality finger was used for the remaining users. The quality labeling was done manually by a human expert (Simon-Zorita et al., 2003).

For each user, 3 fingerprints are used for fingerprint enrollment and the other 7 are used for testing. A near worst-case scenario has been considered by using as impostor data, for each user, the best 10 impostor fingerprints from a pool of 750 different fingers. For each user, 10 user signatures are used for signature enrollment, the other 7 user signatures are used for testing, and 10 skilled forgeries from 5 different impostors are used as impostor testing data.

As a result, data for evaluating the proposed fusion strategies consist of 75×7 user and 75×10 impostor bimodal attempts in a near worst-case scenario.

5.2 Multimodal experimental procedure

Several methods have been described in the literature in order to maximize the use of the information embedded in the training samples during a test (Jain et al., 2000; Theodoridis and Koutroumbas, 2003). Regarding localized multimodal fusion, some of the methods used include resubstitution (Jain and Ross, 2002), holdout (Kumar and Zhang, 2003; Wang et al., 2004; Toh et al.,

2004), and variants of jackknife sampling using the leave-one-out principle (Fierrez-Aguilar et al., 2003).

In particular, when dealing with localized learning we are confronted with severe data scarcity. This has been overcome by Toh et al. (2004) by augmenting the training set with noisy samples and by Fierrez-Aguilar et al. (2004) by using a robust error estimation method based on bootstrap sampling (Duda et al., 2001; Bolle et al., 2004b). In this work we follow either one of these two experimental approaches:

Global learning/decision: Bootstrap data sets have been created by randomly selecting M users from the training set with replacement. This selection process has been independently repeated 300 times to yield 300 different bootstrap data sets. Each data set is used then to generate either a user-independent fusion rule or a user-independent decision function. In the latter case, a non-trained sum rule fusion function is assumed and the selected training data is used for training the decision function on combined scores. Testing is finally performed on the remaining users not included in each bootstrap data set.

Local learning/decision: For each user, 75 bootstrap data sets have been created randomly selecting N samples with replacement forcing each class client/impostor to have at least one sample. For each user and bootstrap data set, a different fusion rule (or a decision function on summed scores) is constructed. Testing is performed on the remaining samples not included in the bootstrap data set.

Adapted learning/decision: Bootstrap sampling of users is performed as in the global case yielding 300 global bootstrap data sets (GBD). Multimodal scores of the remaining users not included in each GBD are then

sampled as in the local case. This yields 75 local bootstrap data sets (LBD) per GBD and per client not included in the GBD. Training of the fusion function (or the decision function on summed scores) is performed using the LBD and associated GBD from which the user was left out. Testing is performed on the remaining samples not included in each LBD.

5.3 Results

Comparative results of global, local, and adapted fusion/decision functions are given in Fig. 4.

In Fig. 4 (a) we plot the verification performance of the bimodal authentication system using the proposed trained SVM-based global fusion approach (GG) for an increasing number of clients in the fusion function training set. Individual performances of the signature and fingerprint subsystems, and the non-trained sum rule fusion approach are also shown for reference. In this case, baseline equal error rate of the simple fusion approach based on sum rule, 2.28% EER, is improved to 1.39% by using the global SVM-based trained fusion scheme ($M = 74$ users for training the fusion function).

In Fig. 4 (c) we compare local approaches for training either the fusion function or the decision function. It is shown that using training data for learning local fusion functions (1.23% EER for $N = 16$ training samples per user) is significantly better than using a simple common fusion rule and exploiting existing development data for training localized decisions (2.17% EER). Worth noting, the local fusion approach (1.23% EER) also outperforms the global fusion strategy in Fig. 4 (a) (1.39% EER) when enough training samples for

building the user-specific fusion functions are available (approximately more than 10 in this experiment).

In Fig. 4 (b) we show the verification results of the proposed adapted approaches. In this case, $M = 74$ clients (global) and $N = 16$ samples per client (local) are used for training and α is varied, hence trading off the influence of the global and local information for training the fusion/decision functions. As a result, a minimum of 1.85% EER is found for $\alpha = 0.75$ in the case of sum rule fusion and adapted decisions, outperforming the local decision scheme in Fig. 4 (c) (2.17%). Adapted fusion outperforms all other strategies lowering the error rate down to 0.80% EER also for $\alpha = 0.75$.

Trade-off verification performances for the above mentioned experiments are depicted in Fig. 5 as DET curves (Martin et al., 1997). In particular, a highly remarkable relative improvement of 42% in the EER with respect to the user-independent fusion approach is achieved by using the proposed adapted fusion method. The severe and very common problem of training data scarcity in the user-dependent fusion strategy is also relaxed by the proposed scheme, resulting in a relative improvement of 35% in the EER compared to the raw user-dependent fusion strategy.

In order to visualize the discriminative capability of SVM classifiers in the above described fusion approaches, client and impostor scatter plots of signature and fingerprint scores before fusing are plotted in Fig. 6 (a). A data set of the bootstrap error estimation process is considered and global, local and adapted fusion function boundaries (i.e., $f(\mathbf{x}) = 0$) are depicted. For the same data set of the bootstrap sampling process, global, local, an adapted decision boundaries on summed scores (i.e., $f(s) = 0$) are shown in Fig. 6 (b).

It can be seen in both cases how the proposed adapted scheme helps in classifying correctly a client test sample in which the fingerprint score is significantly lower than the local client training scores. In this case, training data scarcity in the local approach leads to a wrong decision, i.e., it is not likely that this attempt comes from a client based on the training data with the local approach. Considering the general knowledge with the adapted scheme leads to a correct decision, i.e., based on the general knowledge provided by other users, we can expect client attempts with low fingerprint score and very high signature score.

6 Conclusions and future work

User-dependent approaches to multimodal biometric verification have been reviewed, and the taxonomy proposed by Toh et al. (2004) based on global/local learning/decision has been extended by incorporating adapted strategies. Operational methods for learning the fusion/decision functions based on Support Vector Machines have been described. Most remarkably, a novel adapted scheme for learning/decision has been introduced based on both the general knowledge provided by pooling user-independent data, and the local characteristics of the user at hand. The proposed approach has been experimentally shown to overcome the training data scarcity problem encountered very often in user-dependent learning scenarios.

A set of comparative experiments have been conducted using: *i*) a bimodal biometric verification system based on fingerprint (Simon-Zorita et al., 2003) and on-line signature (Fierrez-Aguilar et al., 2005a) traits, *ii*) real bimodal biometric data from the MCYT database (Ortega-Garcia et al., 2003), and *iii*)

a novel experimental protocol based on a worst-case scenario and bootstrap error estimates (Bolle et al., 2004b).

For the scenario described in this work, and when enough training data is available for the trained approaches, the following set of experimental findings have been obtained: *i*) trained fusion/decision outperforms non-trained simple approaches such as sum rule, *ii*) for the same amount of training data, local learning of the fusion functions outperforms localized trained decisions on summed scores, *iii*) local learning outperforms global learning, *iv*) adapted learning by using both global information from a pool of users and user-specific training data outperforms all other approaches. Most remarkably, we report some indications of the critical “enough training data” issue when comparing the trained to the not trained, and the global to the local approaches.

Future work will involve exploring other sources of errors and dependencies in multimodal biometrics, for example biometric signal quality (Fierrez-Aguilar et al., 2005b), and developing adapted schemes to compensate for them. Finally, even though we have focused on multimodal biometrics, the proposed techniques can be applied to other pattern recognition problems using multiple matcher approaches.

Acknowledgements

This work has been supported by the Spanish Ministry for Science and Technology under projects TIC2003-09068-C02-01 and TIC2003-08382-C05-01. J. F.-A. thanks Consejeria de Educacion de la Comunidad de Madrid and Fondo Social Europeo for supporting his doctoral research.

References

- Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E., 1999. Fusion of face and speech data for person identity verification. *IEEE Trans. on Neural Networks* 10 (5), 1065–1074.
- Bigun, E. S., Bigun, J., Duc, B., Fischer, S., 1997. Expert conciliation for multi modal person authentication systems by bayesian statistics. In: Bigun, J., Chollet, G., Borgefors, G. (Eds.), *Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA*. Springer LNCS-1206, pp. 291–300.
- Bolle, R. M., Connell, J. H., Pankanti, S., Ratha, N. K., Senior, A. W., 2004a. *Guide to Biometrics*. Springer Verlag.
- Bolle, R. M., Ratha, N. K., Pankanti, S., 2004b. Error analysis of pattern recognition systems—the subsets bootstrap. *Computer Vision and Image Understanding* 93, 1–33.
- Brunelli, R., Falavigna, D., 1995. Person identification using multiple cues. *IEEE Trans. on Pattern Anal. and Machine Intell.* 17 (10), 955–966.
- Chatzis, V., Bors, A. G., Pitas, I., 1999. Multimodal decision-level fusion for person authentication. *IEEE Trans. on System, Man, and Cybernetics, part A* 29 (6), 674–680.
- Doddington, G., Liggett, W., Martin, A., Przybocki, M., Reynolds, D., 1998. Sheeps, goats, lambs and wolves: A statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation. In: *Proc. of Intl. Conf. on Speech and Language Processing, ICSLP*.
- Duda, R. O., Hart, P. E., Stork, D. G., 2001. *Pattern Classification*. Wiley.
- Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J., 2004. Exploiting general knowledge in user-dependent fusion

- strategies for multimodal biometric verification. In: Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, ICASSP. Vol. 5. pp. 617–620.
- Fierrez-Aguilar, J., Ortega-Garcia, J., Garcia-Romero, D., Gonzalez-Rodriguez, J., 2003. A comparative evaluation of fusion strategies for multimodal biometric verification. In: Kittler, J., Nixon, M. S. (Eds.), Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA. Springer LNCS-2688, pp. 830–837.
- Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., 2005a. Target dependent score normalization techniques and their application to signature verification. *IEEE Trans. on Systems, Man and Cybernetics, part C* 35 (3), (to appear).
- Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Bigun, J., 2005b. Discriminative multimodal biometric authentication based on quality measures. *Pattern Recognition* 38 (5), 777–779.
- Hong, L., Jain, A. K., 1998. Integrating faces and fingerprints for personal identification. *IEEE Trans. on Pattern Anal. and Machine Intell.* 20 (12), 1295–1307.
- Indovina, M., Uludag, U., Snelick, R., Mink, A., Jain, A., 2003. Multimodal biometric authentication methods: A COTS approach. In: Proc. of Workshop on Multimodal User Authentication, MMUA. (available at <http://mmua.cs.ucsb.edu/>).
- Jain, A. K., Duin, R. P. W., Mao, J., 2000. Statistical pattern recognition: A review. *IEEE Trans. on Pattern Anal. and Machine Intell.* 22 (1), 4–37.
- Jain, A. K., Ross, A., 2002. Learning user-specific parameters in a multibiometric system. In: Proc. of the IEEE Intl. Conf. on Image Processing, ICIP. Vol. 1. pp. 57–60.

- Jain, A. K., Ross, A., 2004. Multibiometric systems. *Communications of the ACM* 47 (1), 34–40.
- Jain, A. K., Ross, A., Prabhakar, S., 2004. An introduction to biometric recognition. *IEEE Trans. on Circuits and Systems for Video Technology* 14 (1), 4–20.
- Kittler, J., Hatef, M., Duin, R., Matas, J., 1998. On combining classifiers. *IEEE Trans. on Pattern Anal. and Machine Intell.* 20 (3), 226–239.
- Kumar, A., Zhang, D., 2003. Integrating palmprint with face for user authentication. In: *Proc. of Workshop on Multimodal User Authentication, MMUA*. (available at <http://mmua.cs.ucsb.edu/>).
- Lee, C.-H., Huo, Q., 2000. On adaptive decision rules and decision parameter adaptation for automatic speech recognition. *Proceedings of the IEEE* 88 (8), 1241–1269.
- Maltoni, D., Maio, D., Jain, A. K., Prabhakar, S., 2003. *Handbook of Fingerprint Recognition*. Springer Verlag, Ch. Multimodal Biometric Systems, pp. 233–255.
- Martin, A., Doddington, G., Kamm, T., Ordowski, M., Przybocki, M., 1997. The DET curve in assessment of decision task performance. In: *Proc. of ESCA Eur. Conf. on Speech Comm. and Tech., EuroSpeech*. pp. 1895–1898.
- Navia-Vazquez, A., Perez-Cruz, F., Artes-Rodriguez, A., Figueiras-Vidal, A.-R., 2001. Weighted least squares training of support vector classifiers leading to compact and adaptive schemes. *IEEE Trans. on Neural Networks* 12 (5), 1047–1059.
- Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J.-J., Vivaracho, C., Escudero, D., Moro, Q.-I., 2003. MCYT baseline corpus: A bimodal biometric database. *IEE Proceedings Vision, Image and Signal Processing*

- 150 (6), 395–401.
- Osuna, E., Freund, R., Girosi, F., 1997. An improved training algorithm for support vector machines. In: Proc. of IEEE Workshop on NNSP. pp. 276–285.
- Poh, N., Bengio, S., 2005. An investigation of f-ratio client-dependent normalisation on biometric authentication tasks. In: Proc. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, ICASSP. Vol. 1. pp. 721–724.
- Roli, F., Kittler, J., Windeatt, T. (Eds.), 2004. Multiple Classifier Systems, Fifth International Workshop. Vol. 3077 of Lecture Notes on Computer Science. Springer.
- Simon-Zorita, D., Ortega-Garcia, J., Fierrez-Aguilar, J., Gonzalez-Rodriguez, J., 2003. Image quality and position variability assessment in minutiae-based fingerprint verification. IEE Proceedings Vision, Image and Signal Processing 150 (6), 402–408.
- Theodoridis, S., Koutroumbas, K., 2003. Pattern Recognition. Academic Press.
- Toh, K.-A., Jiang, X., Yau, W.-Y., 2004. Exploiting local and global decisions for multimodal biometrics verification. IEEE Trans. on Signal Processing 52, 3059–3072.
- Vandervei, R. J., 1999. LOQO: An interior point code for quadratic programming. Optimization Methods and Software 12, 451–484.
- Vapnik, V. N., 2000. The Nature of Statistical Learning Theory. Springer.
- Verlinde, P., Chollet, G., Acheroy, M., 2000. Multi-modal identity verification using expert fusion. Information Fusion 1 (1), 17–33.
- Wang, Y., Wang, Y., Tan, T., 2004. Combining fingerprint and voice biometrics for identity verification: An experimental comparison. In: Zhang, D., Jain, A. K. (Eds.), Proc. of Intl. Conf. on Biometric Authentication, ICBA.

Springer LNCS-3072, pp. 663–670.

Yeung, D.-Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G., 2004. SVC2004: First International Signature Verification Competition. In: Zhang, D., Jain, A. K. (Eds.), Proc. of Intl. Conf. on Biometric Authentication, ICBA. Springer LNCS-3072, pp. 16–22.

Figure captions:

Fig. 1. System model of multimodal biometric verification. Global, local, and adapted approaches for score fusion and decision making are also depicted.

Fig. 2. Fingerprint feature extraction process.

Fig. 3. Graphical sketch of the processing stages of the on-line signature verification system.

Fig. 4. Equal error rates of global (a), adapted (b), and local (c) approaches for multimodal fusion based on SVMs.

Fig. 5. Verification performance of global, local, and adapted approaches for multimodal fusion based on SVMs.

Fig. 6. Training/testing scatter plot and decision boundaries of global, local, and adapted approaches for multimodal fusion based on SVMs (one iteration of the bootstrap-based error estimation process).

Figure 1:

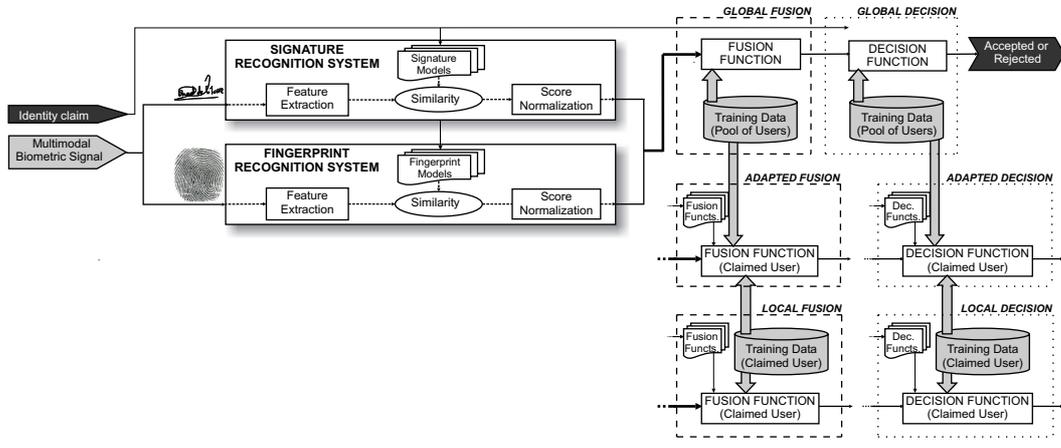


Figure 2:



Figure 3:

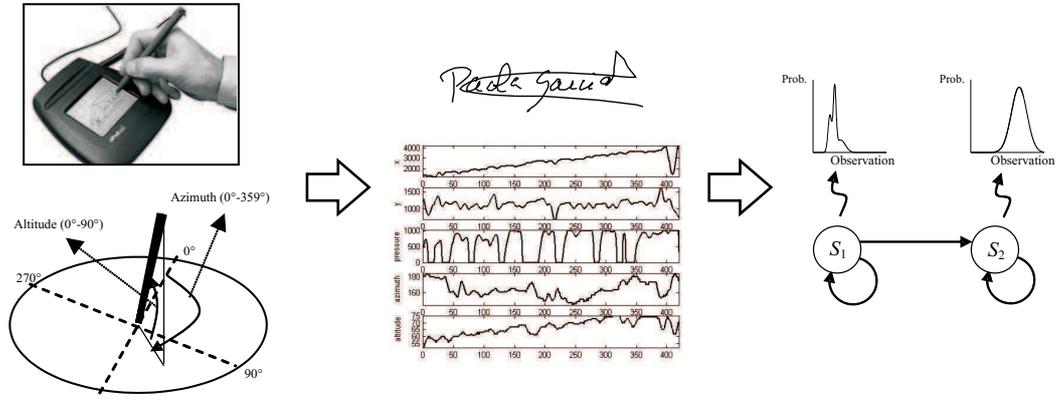
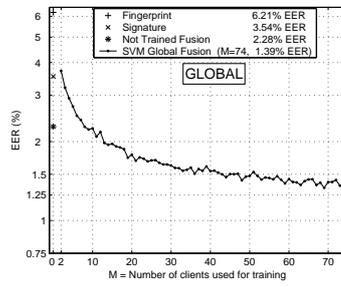
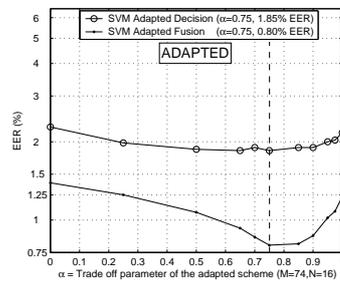


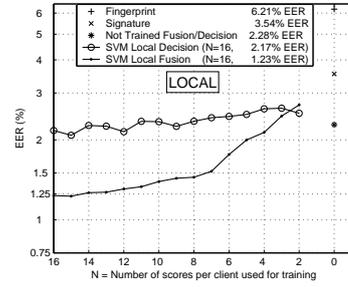
Figure 4:



(a)



(b)



(c)

Figure 5:

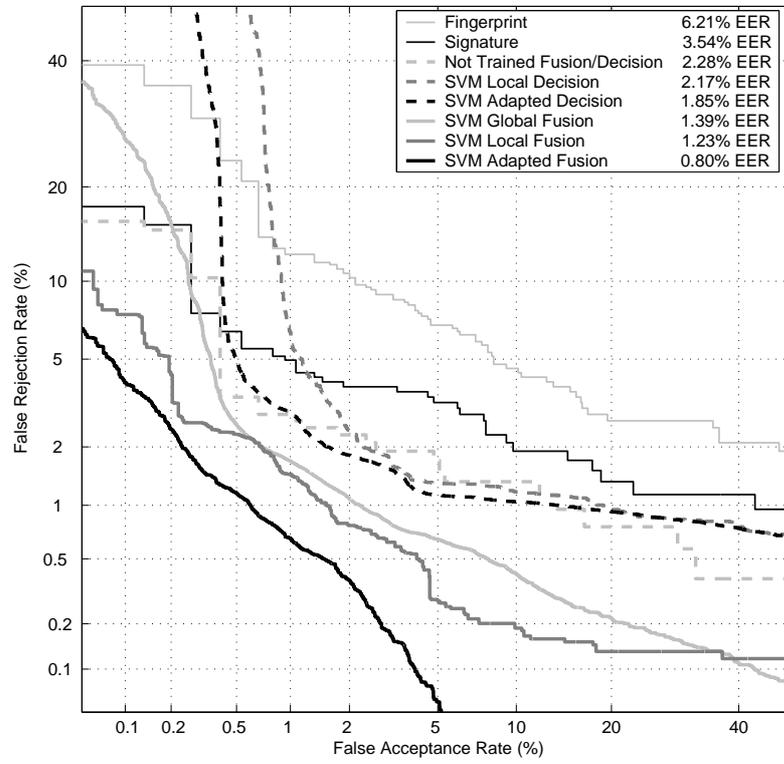
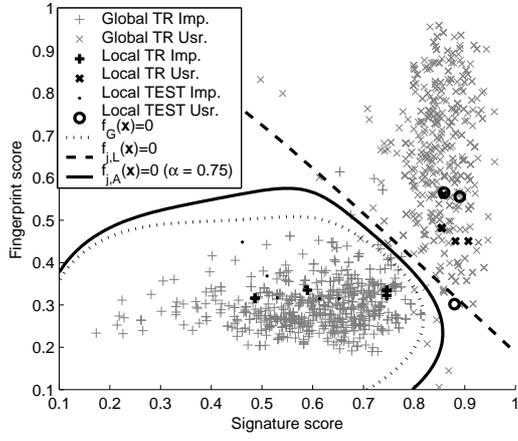
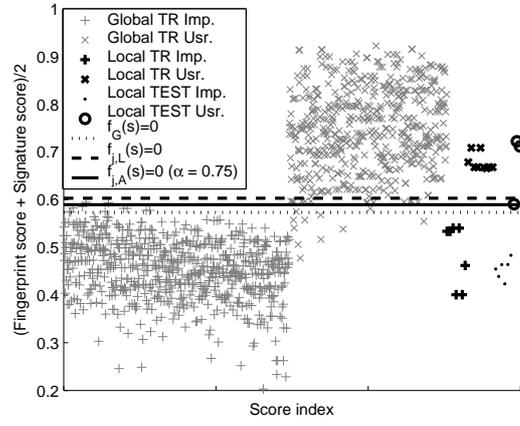


Figure 6:



(a)



(b)

JULIAN FIERREZ-AGUILAR received the M.S. degree in Electrical Engineering in 2001, from Universidad Politecnica de Madrid. Since 2004 he is with Universidad Autonoma de Madrid, where he is currently working towards the Ph.D. degree on multimodal biometrics. His research interests include signal and image processing, pattern recognition and biometrics. He was the recipient of the Best Poster Award at AVBPA 2003 and led the development of the UPM signature verification system ranked 2nd in SVC 2004.

DANIEL GARCIA-ROMERO received his B.S. in Electrical Engineering (with highest honors) in 2000 and his M.S. in Electrical Engineering in 2004, both from Universidad Politecnica de Madrid, Spain. His research interest are focused on signal and image processing, pattern recognition and biometrics. He led the development of the ATVS-UPM speaker recognition system for the NIST 2002 and NIST 2004 evaluations.

JAVIER ORTEGA-GARCIA received the Ph.D. degree in electrical engineering in 1996 from Universidad Politecnica de Madrid. He is currently an Associate Professor at Universidad Autonoma de Madrid. His research interests are focused on forensic acoustics and biometrics signal processing. He has participated in several scientific and technical committees, and has chaired “Odyssey-04, The ISCA Speaker Recognition Workshop”.

JOAQUIN GONZALEZ-RODRIGUEZ received the Ph.D. degree in electrical engineering in 1999 from Universidad Politecnica de Madrid. He is currently and Associate Professor at Universidad Autonoma de Madrid. His research interests are focused on signal processing, biometrics and forensics. He is an invited member of ENFSI and has been vice-chairman for “Odyssey-04, The ISCA Speaker Recognition Workshop”.