

Universal Background Models for Dynamic Signature Verification

M. Martinez-Diaz, J. Fierrez, J. Ortega-Garcia

Abstract—The applicability of Universal Background Models as a score normalization technique is studied for the case of dynamic signature verification. This technique is commonly used in speaker verification systems. Background Models are tested in two different systems based on global features: one based on Parzen Windows and another based on adapted Gaussian Mixture Models. Experiments are carried out in the large MCYT database (16,500 signatures from 330 users) revealing a significant improvement in the overall system performance, specially in the casual impostor scenario.

I. INTRODUCTION

The increasing need of automatic personal identification in our society has motivated the raise of biometrics as a convenient approach since no tokens or passwords must be kept by the user [1]. Signature is among the most socially accepted biometric traits, and it is commonly used for document validation, and legal and financial transactions (e.g. bank checks). Thus, many approaches for automatic signature verification have been studied [2]–[4]. Nevertheless, signature verification is affected by some factors which still make it a challenging task. These are the large intra-class variability, as signers tend to vary their signature over periods of time, and the existence of skilled forgers, which can perform signature imitations leading to a very low inter-class variability.

Signature verification systems can be classified in two types: *off-line* systems use only static information extracted from the signature still images while on the other hand *on-line* (or dynamic) systems take advantage of dynamic information from the signature (e.g. velocity, pressure, etc.) which must be captured by a digitizing tablet or equivalent device. Traditionally, the on-line or dynamic approach has provided better performance than the off-line approach [2].

Two main approaches exist in dynamic signature verification systems [2]. *Feature-based* systems extract holistic feature vectors from the signatures and compare them in order to obtain similarity measures. *Function-based* systems consider the signing process as a discrete time function and perform the matching process based on elastic or statistical matching of sequences such as Dynamic Time Warping (DTW) or Hidden Markov Models (HMM) [3]. Feature-based signature

verification systems are considered in this work. In such systems, similarity between signature samples is generally computed as a distance measure or using statistical pattern classification techniques [5]. These techniques consider the verification task as a two-class problem, where the user input signature must be classified as genuine or as non-genuine (i.e. belonging to an impostor). Decisions are made by comparing the similarity score between the input and the enrolled signatures with a given threshold [6]. Similarity scores can be normalized to allow fusion from multiple scores (e.g. multi-matchers) or to increase the system performance [6], [7].

In this work, a score normalization approach for dynamic signature verification based on Universal Background Models (UBM) [8], which consist in statistical models of an “average” user, is proposed. In the verification process, the user signature is compared to his claimed template and the resulting score is normalized by its similarity to the UBM. In Fig. 1 the signature verification system model, including the UBM score normalization step, is depicted. This technique has led to very good results in speaker verification systems [9], [10]. However, it has not been applied to signature verification systems to the extent of our knowledge. Our proposed UBM is tested on an existing system based on Parzen Windows and on a new system based on adapted Gaussian Mixture Models (GMM) adapted from the UBM. GMMs have been applied in signature verification both for function-based [11], [12] and feature-based approaches [13]. Nevertheless, none of the existing works, to the best of our knowledge, have applied model adaptation.

This paper is structured as follows. A theoretical background for score normalization is given in Sect. II, UBMs and model adaptation are presented in Sect. III, verification systems are described in Sect. IV. The experimental protocol and results are given in Sect. V and conclusions are finally drawn in Sect. VI.

II. THEORETICAL BACKGROUND

Several normalization techniques have been proposed with the aim of re-scaling scores produced by different matchers to a common domain in order to allow their fusion [7]. These approaches don't generally affect the individual matcher performance, but the combined performance of a multi-biometric system. This is because the score normalization function is the same for all enrolled subjects and inputs. User-dependent score normalization approaches for signature verification are studied in [6], and are demonstrated to improve the performance of a single matcher. The present work is focused on *test-dependent score normalization*.

M. Martinez-Diaz, J. Fierrez, and J. Ortega-Garcia are with Biometric Recognition Group - ATVS, Esc. Politecnica Superior - Univ. Autonoma de Madrid C/ Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain {marcos.martinez, julian.fierrez, javier.ortega}@uam.es

This work has been supported by the Spanish Ministry of Education under project TEC2006-13141-C03-03 and by the European Network of Excellence Biosecure (IST-2002-507634).

J. Fierrez is supported by a Marie Curie Fellowship from the European Commission.

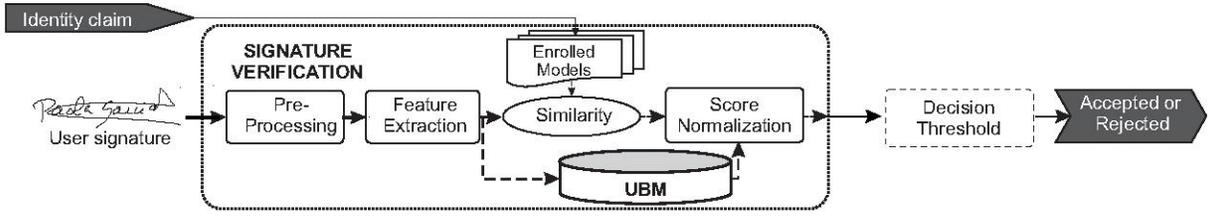


Fig. 1. Signature Verification System architecture showing the score normalization step based on UBM.

Let's state the verification task as a two-class $H_C, H_{\bar{C}}$ problem, each representing an hypothesis:

- H_C : the signature is from the claimed user C .
- $H_{\bar{C}}$: the signature is not from the claimed user C .

Then, the optimum decision is taken using the likelihood ratio test or, equivalently, the log-likelihood ratio test [5]:

$$\log p(\mathbf{x} | H_C) - \log p(\mathbf{x} | H_{\bar{C}}) \begin{cases} > \log \theta, & \text{accept } H_C \\ < \log \theta, & \text{reject } H_C \end{cases} \quad (1)$$

where \mathbf{x} is the input feature vector and θ is a fixed decision threshold. It is a common practice to compute the match score between the test vector \mathbf{x} and the target user feature statistical model λ_C as the likelihood function

$$s = \log p(\mathbf{x} | \lambda_C). \quad (2)$$

So, following the previous notation, a normalized score s_n may be obtained as follows:

$$s_n = s - \log p(\mathbf{x} | \lambda_{\bar{C}}). \quad (3)$$

where $\lambda_{\bar{C}}$ represents a model of the rest of potential users or attackers to the system.

This type of normalization is known as *test-dependent normalization* [6] and requires a pool of users for training the normalization function. Other approaches that depend only on the claimed user and not on the test input are called *user- or target-dependent score normalization* techniques [6]. Universal Background Models provide a convenient approach to approximate the distribution $p(\mathbf{x} | \lambda_{\bar{C}})$.

III. METHODS

A. Universal Background Models

The UBM proposed in this work is a Gaussian Mixture Model (GMM) trained with feature vectors from a pool of users [8]. These vectors are obtained from a database, and are no longer used for the experiments. The number of vectors must be large enough to cover a representative user space. The UBM must be trained once and can be then used for all the claiming users. GMMs model a statistic distribution as a linear combination of d -dimensional Gaussian probability density functions:

$$p(\mathbf{x} | \lambda_{\bar{C}}) = \sum_{i=1}^N \omega_i p_i(\mathbf{x}) \quad (4)$$

where

$$p_i(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \right\}.$$

In order to be a valid pdf, the weights must satisfy $\sum_{i=1}^N \omega_i = 1$. The parameters to be estimated are then $\{\omega_i, \boldsymbol{\mu}_i, \Sigma_i\}$, $i = 1, \dots, N$, where N is the number of Gaussian components specified by the system designer. The covariance matrices are chosen to be diagonal, as full matrices don't usually provide an advantage in the model approximation [9]. The model parameters can be estimated from a pool of user features using the Expectation Maximization (EM) algorithm [5].

B. Adapted Gaussian Mixture Models

User models can be derived from the UBM by employing the user feature vectors and a form of Bayesian learning or Bayesian adaptation [9]. This approach allows to derive a user-specific GMM by combining the UBM and the information provided by the specific user feature vectors. The UBM parameters are updated using a modified MAP (*Maximum A Posteriori*) estimation. The following steps are performed.

First, the probabilistic alignment between the user training vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_T\}$ and the UBM is computed:

$$P(i | \mathbf{x}_t) = \frac{\omega_i p_i(\mathbf{x}_t)}{\sum_{j=1}^N \omega_j p_j(\mathbf{x}_t)} \quad (5)$$

The sufficient statistics for the weight, mean and variance are then computed:

$$n_i = \sum_{t=1}^T \omega_i P(i | \mathbf{x}_t) \quad (6)$$

$$E_i(\mathbf{x}) = \frac{1}{n_i} \sum_{t=1}^T P(i | \mathbf{x}_t) \mathbf{x}_t \quad (7)$$

$$E_i(\mathbf{x}^2) = \frac{1}{n_i} \sum_{t=1}^T P(i | \mathbf{x}_t) \mathbf{x}_t^2 \quad (8)$$

where $\mathbf{x}^2 = \mathbf{x}\mathbf{x}^T$ and \mathbf{x}^T denotes transpose. Finally, the parameters are updated in the following manner:

$$\hat{\omega}_i = [\alpha n_i / T + (1 - \alpha) \omega_i] \gamma \quad (9)$$

$$\hat{\boldsymbol{\mu}}_i = \alpha E_i(\mathbf{x}) + (1 - \alpha) \boldsymbol{\mu}_i \quad (10)$$

$$\hat{\boldsymbol{\sigma}}_i^2 = \alpha E_i(\mathbf{x}^2) + (1 - \alpha) (\boldsymbol{\sigma}_i^2 - \boldsymbol{\mu}_i^2) - \hat{\boldsymbol{\mu}}_i^2 \quad (11)$$

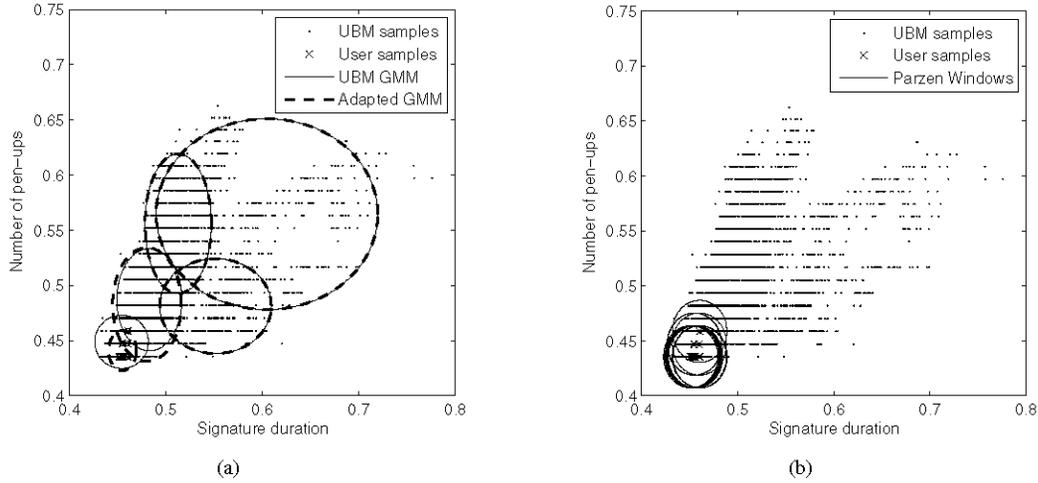


Fig. 2. (a) Bi-dimensional example of adapted GMM vs. UBM with 5 centers for two example features and 20 training signatures. It can be observed that only two components have been significantly adapted. (b) For the same signer and number of training signatures, model based on Parzen Windows. The 2σ contour is depicted for each Gaussian component and features have been normalized. Some signatures of the user considered are shown in Fig. 3.



Fig. 3. Two genuine signatures (left and center) and one skilled forgery (right) and examples of associated captured signals.

where $\mu^2 = \mu\mu^T$, $\sigma^2 = \text{diag}(\Sigma)$, and γ is a constant computed to ensure that $\sum_{i=1}^N \hat{\omega}_i = 1$. The parameter α is an adaptation coefficient that controls the contribution of the old and new parameters to the adapted model and is defined as

$$\alpha = \frac{n_i}{n_i + r} \quad (12)$$

where r is a fixed relevance factor [9] that must be specified. Note that mixture components with low n_i won't be almost adapted to the user parameters, while mixtures with high n_i will have their parameters considerably modified by the user data. The main advantage of this method is that the user model and the UBM are “coupled”, in the sense that the comparison of a given feature vector with the user's GMM and the UBM provides a measure of the user's specificities against the average signer.

IV. VERIFICATION SYSTEMS

The verification systems studied in this work use global features extracted from the captured signature signals. Each signature is represented by 40 features, of which a detailed description may be found in [14]. Some examples

of features are: signature duration, number of pen-ups, direction histograms and acceleration and velocity statistics. The features are normalized using tanh-estimators [7] to a range of (0, 1). Two approaches are taken in order to create the user's signature model: Parzen Windows [5] and adapted GMMs.

A. Parzen Window models

The approach based on Parzen Windows is the same presented in [14]. Gaussian Windows are selected in order to estimate the feature vector probability density function λ_C^{PW} of a given user C . The number of Gaussian Windows is equal to the amount of signatures selected for the training phase (5 or 20, in our experiments). The similarity score between an input vector \mathbf{x}_T and the user's model λ_C^{PW} is then computed as

$$s_{PW} = \log p(\mathbf{x}_T | \lambda_C^{PW}). \quad (13)$$

Note that this model based on Gaussian Parzen Windows can be considered as a particular case of GMM, with equal weights $\omega_i = 1/N$, $i = 1, \dots, N$, and a covariance matrix that is a scalar multiple of the identity matrix $\Sigma_i = \sigma^2 \mathbf{I}$.

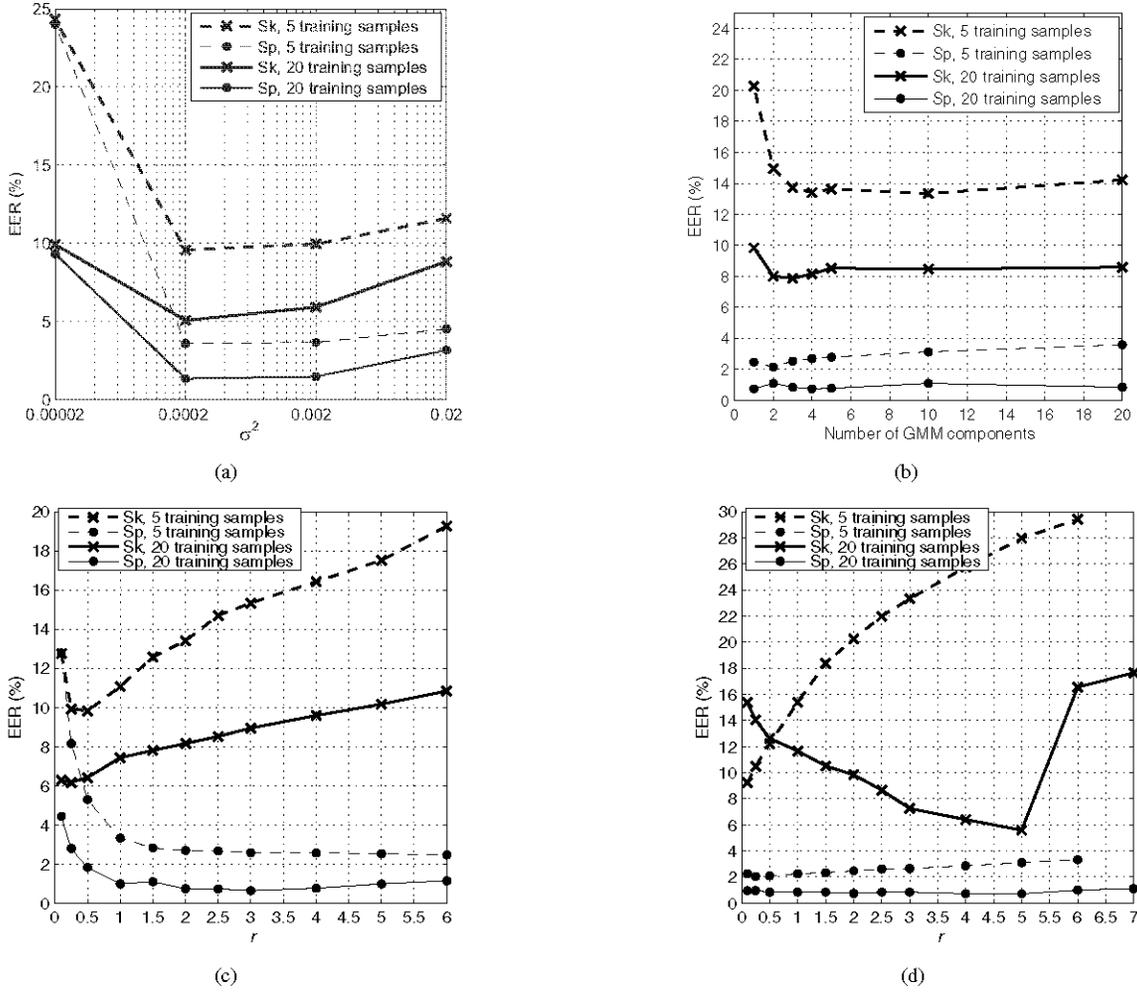


Fig. 4. Effect of the parameters on the system performance. (a) Effect of σ^2 in Parzen Windows, (b) Effect of the number of GMM components with $r = 2$, (c) Effect of r using 4 GMM components; and (d) Effect of r using 1 GMM component. Note that figures have different scales.

B. Adapted GMMs

The proposed adapted GMM system is designed as follows: once the UBM has been calculated, it is adapted to each user C with (5)-(12) leading to an estimation of the user's model λ_C^{AG} . The optimal parameter r is studied in the experiments section. The similarity score s_{AG} is computed as in the case of Parzen Windows:

$$s_{AG} = \log p(\mathbf{x}_T | \lambda_C^{AG}) \quad (14)$$

An example of a bi-dimensional Parzen Window model and an adapted GMM for two example global features and 20 training signatures of a user is shown in Fig. 2. Example signatures of this user and the related dynamic functions are shown in Fig. 3

V. EXPERIMENTS

A. Database and Experimental Protocol

The signature corpus of the bimodal MCYT database [15] is used for the experiments. This database consists of 330 users, with 25 genuine signatures and 25 skilled forgeries

per user, which have been performed by other users that have been let to train for each forgery. One half of the signature corpus is used to train the UBM (25×165 signatures). From the remaining half, either 5 or 20 genuine signatures per user are left to train the user models, while the remaining genuine signatures are used for testing purposes (leading to 20×165 or 5×165 genuine test scores respectively). This is done to assess the impact of the training set size in the system performance. The 25 skilled forgeries per user are used for testing (i.e. 25×165 skilled impostor scores). Simple forgery scores (the case where an impostor claims to be another user but traces his own signature) are computed by comparing one signature from each user with another from the rest of users in the evaluation test (i.e. 164×165 simple forgery scores).

In the experiments, we investigate the best UBM parameters for different training set sizes in the two systems under study. In the case of Parzen Windows, the Gaussian window covariance matrix constant value σ^2 is analyzed, since in [14] no information about σ^2 is provided. For the adapted GMMs, the effect of the number of components N and of the parameter r is studied.

TABLE I
SYSTEM PERFORMANCE USING ADAPTED GMMs FOR 20 AND 5 TRAINING SAMPLES. SP AND SK DENOTE SIMPLE AND SKILLED FORGERIES
RESPECTIVELY.

Approach	N	r	20 Training Samples		5 Training Samples	
			$EER_{sp}(\%)$	$EER_{sk}(\%)$	$EER_{sp}(\%)$	$EER_{sk}(\%)$
Adapted GMM / UBM	4	2	0.74	8.16	2.69	13.41
Adapted GMM / UBM	4	3	0.64	8.95	2.58	15.34
Adapted GMM / UBM	4	1	0.99	7.43	3.33	11.07
Adapted GMM / UBM	1	5	0.72	5.59	3.11	27.96
Adapted GMM / UBM	1	0.25	0.97	14.05	2.06	10.49
Adapted GMM	4	3	4.19	6.79	11.25	13.21
Adapted GMM	1	0.25	0.85	4.73	2.61	8.11

TABLE II
SYSTEM PERFORMANCE USING PARZEN-WINDOW MODELS AND UBM NORMALIZATION FOR 20 AND 5 TRAINING SAMPLES AND DIFFERENT
NUMBER OF GAUSSIAN COMPONENTS IN THE UBM GMM. SP AND SK DENOTE SIMPLE AND SKILLED FORGERIES RESPECTIVELY.

Approach	N	20 Training Samples		5 Training Samples	
		$EER_{sp}(\%)$	$EER_{sk}(\%)$	$EER_{sp}(\%)$	$EER_{sk}(\%)$
Parzen / UBM	4	0.85	6.10	2.45	9.59
Parzen / UBM	2	0.97	4.85	2.48	8.70
Parzen / UBM	1	0.76	5.50	2.30	9.33

TABLE III
SYSTEM PERFORMANCE USING PARZEN WINDOWS FOR A DIFFERENT NUMBER OF TRAINING SAMPLES. SP AND SK DENOTE SIMPLE AND SKILLED
FORGERIES RESPECTIVELY.

Approach	σ^2	N	20 Training Samples		5 Training Samples	
			$EER_{sp}(\%)$	$EER_{sk}(\%)$	$EER_{sp}(\%)$	$EER_{sk}(\%)$
Parzen Windows	0.0002	N/A	1,36	5,08	3,60	9,58
Parzen / UBM	0.0002	1	0.76	5.50	2.30	9.33
Adapted GMM / UBM	N/A	1	($r = 5$) 0.72	($r = 5$) 5.59	($r = 0.25$) 2.06	($r = 0.25$) 10.49

B. Experimental Results

a) *Reference system – Parzen Windows without score normalization:* As a starting point, the performance of the system based on Parzen Windows with no score normalization is studied. The effect of the value of σ^2 in the system performance is depicted in Fig. 4.a. As it can be seen, the best configuration for all EERs is $\sigma^2 = 0.0002$. Lower values of σ^2 lead to very poor results due to the over-fitting phenomenon, caused by too narrow Parzen Windows with very little overlap, each one centered on a training sample. This effect is much more pronounced in the case of 5 training signatures. The EER values for the best configuration are shown in the first row of Table III.

b) *Adapted GMMs and UBM normalization:* The effect of the number N of GMM components with a fixed relevance factor $r = 2$ is first studied. It is shown in Fig. 4.b that, depending on the scenario being considered (5 or 20 training signatures, skilled or random forgeries), the optimum number of Gaussian components varies from 1 to 4. The EER for the 4-component, which gives a good performance for all scenarios is shown in the first row of Table I.

The effect of the relevance factor r is studied for the case of 4 and 1 components, as they provide two reasonable working points. The system EER as a function of r is presented in Figs. 4.c and 4.d for $N = 4$ (best suited for 20 training samples) and $N = 1$ (best suited for 5 training samples) respectively. For the case of 4 components, it can be observed that as r increases, the EER for skilled forgeries increases after a local minimum while in the case of simple forgeries, it remains stable after an initial steep descent. This may be caused by the fact that the model becomes less adapted to the user's specificities as r increases, causing it to be more generalist and thus easier to imitate. On the other hand, even with large r , the adapted model provides enough user-specific information to be discriminative among simple forgeries. In the case of $N = 1$ (Fig. 4.d), the most remarkable difference with $N = 4$ (Fig. 4.c) can be seen in the EER for skilled forgeries, which is higher in general but experiments a local minimum around $r = 5$, lower than the best possible configuration with 4 centers. A selected set of results using adapted GMMs is presented in Table I, both with and without UBM score normalization. As can be seen, the performance improvement over the reference

system based on Parzen Windows is achieved mainly due to the UBM normalization and not due to the adaptation.

When comparing the results of the adapted GMMs with UBM normalization to the reference system based on Parzen Windows, a notable reduction of the EER is detected for the case of simple forgeries, while the EER for skilled forgeries increases slightly.

c) Parzen Window models and UBM normalization:

We now study the effect of UBM normalization on the system based on Parzen Windows. In Table II a notorious improvement in the performance of the Parzen-Window based system can be observed. UBM normalization allows to reduce the EER for simple forgeries while keeping the EER for skilled forgeries at similar values than without normalizing. This effect may be produced by the fact that UBM normalization allows to highlight the user's signature most characteristic features vs. the global pool of background signatures, improving the performance for simple forgeries. The relative invariability of the performance with skilled forgeries may be explained by the fact that the highlighted user specificities are indeed the ones that the forgers are not able to reproduce. Consequently, the overall performance of Parzen Windows is significantly improved by including UBM normalization and is higher than the one for adapted GMMs. The experimental results suggest that the improvement in the system's performance using adapted GMM and Parzen Windows with UBM is primarily caused by the UBM normalization, as occurred with the adapted GMMs.

Finally, we compare in Table III the baseline reference system based on Parzen Windows, and the best configurations achieved for adapted GMMs and Parzen Windows with UBM score normalization. The performance improvement of both approaches over the baseline system is similar, with slightly better exploitation of the user specificities among different signers (i.e., better EER_{sp}) by the adapted GMM for small training set size (5 signatures), at the cost of slightly worse EER_{sk} and increased computational complexity.

VI. CONCLUSIONS AND FUTURE WORK

A *test-dependent* score normalization technique with Universal Background Models has been applied to two on-line signature verification systems using global features. This approach allows to normalize the matcher's output score by a factor dependent on the similarity between the input vector and the UBM, which is a Gaussian Mixture Model of an "average" user built from a pre-selected pool of users.

Normalization has been tested on an existing system based on Parzen Windows and on a novel system based on adapted GMMs. Adapted GMMs allow to compute a user statistical model based on a weighted combination of the UBM parameters and the user's feature vectors and are widely used in speaker verification systems.

Experiments have been carried out on the MCYT database comprising 16500 signatures from 330 contributors and the influence of the different types of parameters has been

studied. It has been found that UBM normalization allows to a remarkable improvement on the system's performance in the scenario of simple forgeries, while not substantially affecting the performance for skilled forgeries. Parzen Windows with UBM normalization have been found to have a similar performance than adapted GMMs, while being less computationally demanding. On the other hand, the adapted GMM approach outperformed the Parzen Windows in exploiting the user-specificities among different signers with small training set size. Finally, we have observed that the performance improvement in the adapted GMM approach comes mainly from the UBM score normalization step.

Future work includes the use of UBM score normalization and model adaptation in signature verification systems based on Hidden Markov Models (HMM).

REFERENCES

- [1] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 2, pp. 125–143, 2006.
- [2] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification: The state of the art," *Pattern Recognition*, vol. 22, no. 2, pp. 107–131, 1989.
- [3] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 63–84, 2000.
- [4] J. Fierrez and J. Ortega-Garcia, *Handbook of Biometrics*, chapter On-line signature verification, Eds. A. K. Jain and A. Ross and P. Flynn (in press), 2007.
- [5] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, Wiley-Interscience, 2001.
- [6] J. Fierrez-Aguilar, J. Ortega-Garcia, and J. Gonzalez-Rodriguez, "Target dependent score normalization techniques and their application to signature verification," *IEEE Trans. on Systems, Man and Cybernetics, part C*, vol. 35, no. 3, pp. 418–425, 2005.
- [7] A. K. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270–2285, 2005.
- [8] D. A. Reynolds, "Comparison of background normalization methods for text-independent speaker verification," in *Proc. of 5th European Conf. on Speech Communication and Technology (Eurospeech)*, 1997, vol. 2, pp. 963–966.
- [9] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian Mixture Models," *Digital Signal Processing*, vol. 10, pp. 19–41, 2000.
- [10] F. Bimbot, J. F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meignier, T. Merlin, J. Ortega-Garcia, D. Petrovska-Delacretaz, and D. A. Reynolds, "A tutorial on text-independent speaker verification," *EURASIP Journal on Applied Signal Processing*, vol. 2004:4, pp. 430–451, 2004.
- [11] J. Richiardi and A. Drygajlo, "Gaussian mixture models for on-line signature verification," in *Intl Multimedia Conf., Proc. ACM SIGMM workshop on Biometrics methods and applications*, 2003, pp. 115–122.
- [12] S. Garcia-Salicetti, J. Fierrez-Aguilar, F. Alonso-Fernandez, C. Vielhauer, R. Guest, L. Allano, T. Doan Trung, T. Scheidat, B. Ly Van, J. Dittmann, B. Dorizzi, J. Ortega-Garcia, J. Gonzalez-Rodriguez, M. Bacile di Castiglione, and M. Fairhurst, "Biosecure reference systems for on-line signature verification: A study of complementarity," *Annals of Telecommunications, Special Issue on Multimodal Biometrics*, vol. 62, no. 1-2, pp. 36–61, January-February 2007.
- [13] J. Richiardi, H. Ketabdar, and A. Drygajlo, "Local and global feature selection for on-line signature verification," in *Proc. IAPR 8th International Conference on Document Analysis and Recognition, ICAR*, Seoul, Korea, August-September 2005.
- [14] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia, and D. Maltoni, "An on-line signature verification system based on fusion of local and global information," in *Proc. of IAPR Intl. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA*. 2005, pp. 523–532, Springer LNCS-3546.
- [15] J. Ortega-Garcia, Fierrez-Aguilar, et al., "MCYT baseline corpus: a bimodal biometric database," *IEE Proc. Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 391–401, 2003.