An Off-line Signature Verification System Based on Fusion of Local and Global Information

J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, and J. Ortega-Garcia

Biometrics Research Lab., ATVS
Universidad Politecnica de Madrid, Spain
{jfierrez,jortega}@diac.upm.es

Abstract. An off-line signature verification system based on fusion of two machine experts is presented. One of the experts is based on global image analysis and a statistical distance measure while the second one is based on local image analysis and Hidden Markov Models. Experimental results are given on a subcorpus of the large MCYT signature database for random and skilled forgeries. It is shown experimentally that the machine expert based on local information outperforms the system based on global analysis in all reported cases. The two proposed systems are also shown to give complementary recognition information which is exploited with a simple fusion strategy based on the sum rule.

1 Introduction

Automatic extraction of identity cues from personal traits (e.g., signature, fingerprint, speech, and face image) has given raise to a particular branch of pattern recognition, biometrics, where the goal is to infer identity of people from biometric data [1, 2]. The increasing interest on biometrics is related to the number of important applications where a correct assessment of identity is a crucial point.

This work is focused on off-line signature recognition (i.e., only signature image is given as input information for recognition) [3]. The importance of this biometric characteristic (because of its widespread use and legal and social acceptability) and the inherent variability of the signing process has motivated great research efforts in the last decade [4]. As a result, various pattern recognition strategies have been applied to this problem such as minimum distance classifiers [5], dynamic programming [6], neural networks [7], and Hidden Markov Models [8, 9].

One major research trend in biometric verification is the successful exploitation of the different information levels embodied in the biometric signal at hand. Examples can be found regarding fingerprint verification, where the combined utilization of global texture features and local minutiae-based descriptors has been shown to improve verification performance [10], or speaker verification, where the combination of prosodic descriptors and local acoustic features outperforms the state-of-the-art acoustic approach [11]. Regarding off-line signature

D. Maltoni and A.K. Jain (Eds.): BioAW 2004, LNCS 3087, pp. 295–306, 2004.

recognition, some works on multi-expert approaches are [12] and [13]. For a theoretical framework on classifier combination and a comparative evaluation of common fusion strategies, we refer the reader to [14] and [15] respectively.

Two different machine experts have been built in this work. The first machine expert is described in Sect. 2. It is based on global image analysis and a minimum distance classifier as proposed in [5]. The second machine expert is described in Sect. 3. It is based on local image analysis and left-to-right Hidden Markov Models as used in [8] but with a novel local parameterization derived from the work in [5]. A simple fixed fusion strategy based on the sum rule [14] is used as the fusion algorithm as described in Sect. 4. Experimental procedure and results are given in Sect. 5. Some conclusions are finally extracted in Sect. 6.

2 Machine Expert Based on Global Information

2.1 Preprocessing

Input signature images are preprocessed in three consecutive stages as follows:

Binarization: Input images are first binarized by using the histogram-based global thresholding algorithm proposed in [16]. A morphological closing operation with a 3×3 squared structuring element [17] is applied then to the binarized image.

Segmentation: Signature is then segmented by using a bounding box which removes outer traces (this is because signature boundary normally corresponds to flourish with high intra-user variability). Left and right height-wide blocks having all columns with signature pixel count lower than threshold T_p, respectively top and bottom width-wide blocks having all rows with signature pixel count lower than T_p are discarded ($T_p = 15$ in the reported experiments).

Normalization: Database used for experiments have been acquired on a restricted size grid, so intra-user rotation variability is expected to be low and no rotation normalization is applied. Segmented signatures are resized in order to have a width of 512 pixels while maintaining the aspect ratio.

2.2 Feature Extraction

Slant directions of the signature strokes and those of the envelopes of the dilated signature images are used as features for recognition as proposed in [5]. These descriptors are extracted by using mathematical morphology operators [17].

Slant directions are analyzed as follows. The preprocessed signature image is eroded with the structuring elements shown in Fig. 1 thus generating 32 eroded binary images (the area of the structuring elements is 10 pixels and the angle between successive elements is approximately 11°). A slant direction feature subvector with 32 components is then generated, where each component is computed as the signature pixel count in each eroded image.

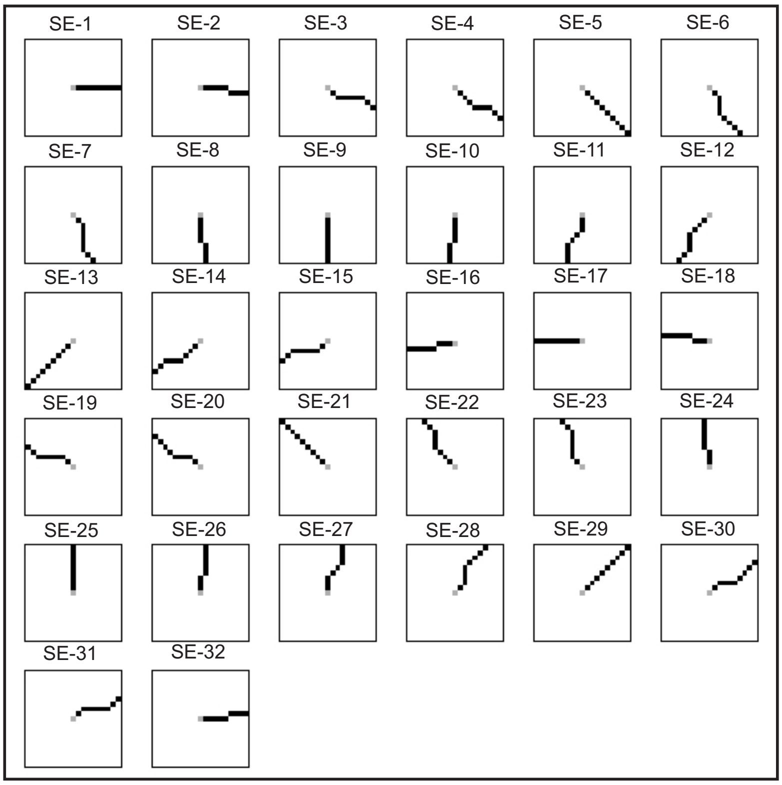


Fig. 1. Structuring elements used for slant direction feature extraction. Origin of the structuring element is indicated in light gray.

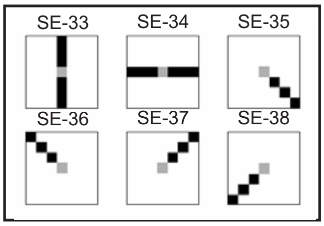


Fig. 2. Structuring elements used for envelope feature extraction. Origin of the structuring element is indicated in light gray.

Envelope is analyzed as follows. Preprocessed signature image is successively dilated 5 times with each one of the 6 structuring elements in Fig. 2 (areas of SE-33/34 and SE-35/36/37/38 are 7 and 4 pixels respectively). An envelop feature sub-vector with 5×6 components is finally generated, where each component is computed as the signature pixel count in the difference image between successive dilations.

Preprocessed signature is finally parameterized as a column vector $\mathbf{o}$ with 62 coordinates by concatenating slant and envelop feature sub-vectors.

2.3 Similarity Computation

Each client of the system or target $\mathcal{T}$ is represented by a statistical model $\lambda_{\mathcal{T}} = \{\boldsymbol{\mu}_{\mathcal{T}}, \boldsymbol{\sigma}_{\mathcal{T}}\}$, where $\boldsymbol{\mu}_{\mathcal{T}}$ and $\boldsymbol{\sigma}_{\mathcal{T}}$ denote mean and deviation vectors respectively. These are estimated by using an enrollment set of K parameterized signatures $\{\mathbf{o}_{\mathcal{T}}^{(1)}, \ldots, \mathbf{o}_{\mathcal{T}}^{(K)}\}$ as follows:

$$\boldsymbol{\mu}_{\mathcal{T}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{o}_{\mathcal{T}}^{(k)} \tag{1}$$

$$\sigma_{\mathcal{T}j} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(o_{\mathcal{T}j}^{(k)} - \mu_{\mathcal{T}j} \right)^2} \quad j = 1, \ldots, 62 \tag{2}$$

where $o_{\mathcal{T}j}$, $\mu_{\mathcal{T}j}$ and $\sigma_{\mathcal{T}j}$ are the jth coordinates of $\mathbf{o}_{\mathcal{T}}$, $\boldsymbol{\mu}_{\mathcal{T}}$ and $\boldsymbol{\sigma}_{\mathcal{T}}$ respectively.

Similarity score between a parameterized test signature $\mathbf{o}$ and a claimed model $\lambda_{\mathcal{T}} = \{\boldsymbol{\mu}_{\mathcal{T}}, \boldsymbol{\sigma}_{\mathcal{T}}\}$ is computed as the inverse of the Mahalanobis distance [18]:

$$s_{global}\left(\mathbf{o}, \lambda_{\mathcal{T}}\right) = \frac{1}{\sqrt{\sum_{j=1}^{62} \frac{1}{\sigma_{\mathcal{T}j}^2} \left(o_j - \mu_{\mathcal{T}j}\right)^2}} \tag{3}$$

3 Machine Expert Based on Local Information

3.1 Preprocessing

Input signature images are preprocessed in two consecutive steps as follows:

Binarization: Input images are binarized as in Sect. 2.1.
Segmentation: Signature is segmented by using the bounding box procedure as described in Sect. 2.1 but with a lower signature pixel count threshold ($T_p = 5$ in the experiments).

3.2 Feature Extraction

Slant directions and envelopes are locally analyzed by using an approach similar to the one summarized in Sect. 2.2 but applied to blocks.

Preprocessed signature images are divided into height-wide blocks of 64 pixels width with an overlapping between adjacent blocks of 75%. Rightmost block is discarded. Feature extraction is carried out as in Sect. 2.2 within each block. In this way, each signature is represented by a matrix $\mathbf{O}$ whose T columns are 62-tuples each one corresponding to one block.

3.3 Similarity Computation

In this case a Hidden Markov Model (HMM) is used to represent each client or target $\mathcal{T}$ of the system. Basically, a HMM models a doubly stochastic process governed by a finite state Markov Chain and a set of random functions. At discrete time instant t an observation symbol $\mathbf{o}_t$ is generated according to the random function associated to the current state [19]. The model is hidden in the sense that the underlying state which generated each symbol cannot be deduced from simple symbol observation.

Formally, a HMM is described as follows:

N**:** which is the number of hidden states $\{S_1, \ldots, S_N\}$. The state at discrete time t is denoted as q_t. For the experiments reported in this work, $N = 4$ is used.

A**:** which is the state transition matrix $A = \{a_{ij}\}$ where

$$a_{ij} = \Pr\left(q_{t+1} = S_j \middle| q_t = S_i\right) \qquad 1 \leq i, j \leq N \tag{4}$$

B**:** which are the parameters of the observation symbol density functions. In this work, the observation density function in state j is modelled as a mixture of M multi-variate Gaussian densities

$$b_j(\mathbf{o}) = \sum\nolimits_{m=1}^{M} c_{jm} N\left(\mathbf{o}, \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}\right) \qquad 1 \leq j \leq N \tag{5}$$

Thus $B = \{c_{jm}, \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}\}$, $1 \leq j \leq N$, $1 \leq m \leq M$. For the experiments reported in this work, $M = 8$ is used.

π**:** which is the initial state distribution $\boldsymbol{\pi} = \{\pi_j\}$ where

$$\pi_j = \Pr(q_1 = S_j) \qquad 1 \leq j \leq N \tag{6}$$

A left-to-right topology with no transition skips between states is used in this work. In order to estimate the model $\lambda_{\mathcal{T}} = \{\boldsymbol{\pi}_{\mathcal{T}}, A_{\mathcal{T}}, B_{\mathcal{T}}\}$ for client $\mathcal{T}$, an enrollment set of K parameterized signatures $\{\mathbf{O}_{\mathcal{T}}^{(1)}, \ldots, \mathbf{O}_{\mathcal{T}}^{(K)}\}$ is used. Estimation is carried out maximizing

$$\prod\nolimits_{k=1}^{K} \Pr\left(\mathbf{O}_{\mathcal{T}}^{(k)} \middle| \lambda_{\mathcal{T}}\right) \tag{7}$$

by using the iterative Baum-Welch procedure [19]. Initial estimations of the model parameters needed for the iterative training algorithm are computed using homogenous length duration state splitting and vector quantization as described in [20].

The similarity score of an input signature parameterized as $\mathbf{O} = \{\mathbf{o}_1, \ldots, \mathbf{o}_T\}$ claiming the model $\lambda_{\mathcal{T}}$ is computed as

$$s_{local}(\mathbf{O}, \lambda_{\mathcal{T}}) = \frac{1}{T} \log \Pr(\mathbf{O}|\lambda_{\mathcal{T}}) \tag{8}$$

by using the Viterbi algorithm [19].

4 Fusion of Global and Local Information

Two sound theoretical frameworks for combining classifiers with application to biometric verification are described in [21] and [14], the former from a risk analysis point of view [22] and the later according to the statistical pattern recognition theory [23]. Both of them concluded that the weighted average is a good way of combining the similarity scores provided by the different experts or information levels. Machine learning approaches have also been applied for combining biometric classifiers [24]. The weights, or the parameters of the learning machines, can be estimated off-line using development data, thus leading to *fixed* fusion schemes, or can be estimated by using training data, thus leading to *trained* fusion schemes. Both approaches have their advantages and disadvantages as discussed in [25].

In this work, a fixed fusion strategy based on the sum rule is used. Similarity scores of the global and local experts are averaged after being linearly mapped into the range $[0, 1]$. For estimating the mapping function, a development corpus comprising 25 signers and their associated forgers is used. Development corpus is not used in subsequent experiments.

5 Experiments

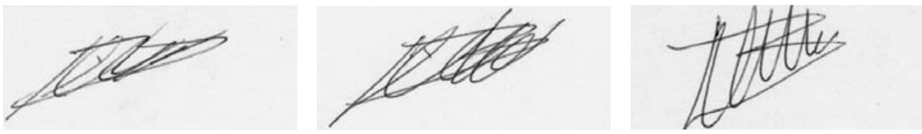
5.1 Database and Experimental Protocol

A subcorpus of the larger MCYT bimodal database [26] is used for the experiments. MCYT database encompasses fingerprint (optical and capacitive acquisition) and on-line signature data (x, y, pressure, azimuth and altitude trajectories of the pen) of 330 contributors from 4 different Spanish sites. In case of signature, high skilled forgeries are also available (forgers are provided the signature images of the clients to be forged and, after training with them several times, they are asked to imitate the shape with natural dynamics, i.e., without breaks or slowdowns)

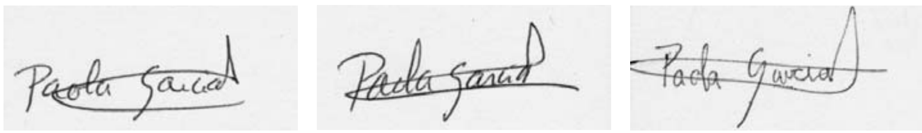
Signature information were acquired in MCYT project by using an inking pen and paper templates over a pen tablet (each signature is written within a 1.75×3.75 cm^2 frame), so signature images were available on paper. Paper templates of 75 signers (and their associated skilled forgeries) have been randomly selected and digitized with an scanner at 600 dpi (dots per inch). Resulting subcorpus comprises 2250 signature images, with 15 genuine signatures and 15 forgeries per user (contributed by 3 different user-specific forgers). Some examples of genuine signatures (left and central columns) and forgeries (right column) are given in Fig. 3 for the four types of signatures encountered in MCYT corpus, namely: simple flourish, complex flourish, name with simple flourish and name with complex flourish.

Signature subcorpus is further divided into training and test sets in order to give holdout system performance error figures [27]. In case of considering skilled forgeries, training set comprises either 5 or 10 genuine signatures (it depends

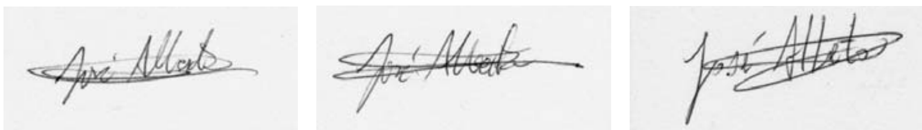
(a) Signature consisting of simple flourish.



(b) Signature consisting of complex flourish.



(c) Signature consisting of written name and simple flourish.



(d) Signature consisting of written name and complex flourish.

Fig. 3. Signature examples from MCYT corpus. Genuine signatures (left and central columns) and skilled forgeries (right column) are depicted for the four types of signatures encountered in MCYT corpus (from (a) to (d)).

on the experiment) and test set consist of the remaining samples (i.e., 10 or 5 genuine signatures and 15 forgeries per user). As a result, either $75 \times 5 = 375$ or $75 \times 10 = 750$ client similarity scores and $75 \times 15 = 1125$ impostor similarity scores are used for reporting results. In case of considering random forgeries (i.e., impostors are claiming others' identities but are using their own signatures), client similarity scores are as above and the number of impostor similarity scores is $75 \times 74 \times 5 = 27750$.

In verification or authentication (the problem addressed here) a claim is made concerning the identity of a person and the biometric system has to take the binary decision of accepting or rejecting it based on the information extracted from the considered biometric trait regarding a predetermined threshold. In a verification context, two situations of error are possible: an impostor is accepted (false acceptance, FA) or the correct user is rejected (false rejection, FR). Performance measures of verification systems are related to the frequency with which these situations of error happen. One common performance measure is the so-called EER (equal error rate) which is the point attained when FA and FR rates coincide [2].

In this work, the overall system performance when *a posteriori* user-independent decision thresholds are used is reported by means of DET plots [28], which are graphical representations of FA vs. FR rates with a particular axis scaling. Average EER tables when using *a posteriori* user-dependent thresholds are also given following the operational procedure proposed in [29] for computing the individual EER for each user. For more details on *a priori* and *a posteriori* decision thresholding techniques and their application to signature verification, we refer the reader to [30].

5.2 Results

Verification performance of the proposed global, local and combined off-line signature machine experts are given for *a posteriori* user-dependent decision thresholding in Table 1 (skilled forgeries) and Table 2 (random forgeries).

Table 1. System performance on **skilled forgeries** for *a posteriori* user-dependent decision thresholding. Average EERs are given in %

	5 training samples	10 training samples
Global Expert	21.84	18.93
Local Expert	14.51	12.22
Combined Expert	11.00	9.28

Table 2. Verification performance on **random forgeries** for *a posteriori* user-dependent decision thresholding. Average EERs are given in %

	5 training samples	10 training samples
Global Expert	8.64	5.12
Local Expert	4.35	1.59
Combined Expert	2.69	1.14

Worth noting, the system based on local analysis and HMMs outperforms the system based on global analysis and Mahalanobis distance in all cases, even when the enrollment set size is small (5 signatures). The two systems are also shown to provide complementary information for the verification task, as the combined

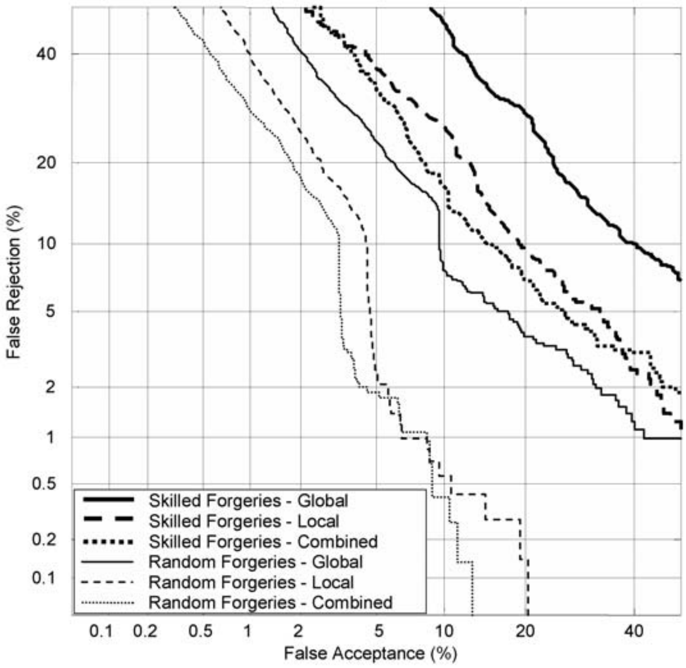


Fig. 4. Verification performance for *a posteriori* user-independent decision thresholding (5 training samples)

system always outperforms the best performing individual expert. In particular, when considering skilled forgeries and 10 training samples, the system based on global analysis leads to 18.93% EER while the proposed combined system leads to 9.28% EER, so a relative improvement of 51% is obtained as compared to published works in this area (the global expert follows closely the system proposed in [5]). In case of random forgeries and 10 training samples, 5.12% EER and 1.14% EER are obtained for global and combined experts respectively. In this case, a relative improvement of 78% is obtained.

Verification performances of individual and combined systems for *a posteriori* user-independent decision thresholds are plotted in Fig. 4 for 5 training signatures, respectively in Fig. 5 for 10 training signatures. The above pointed out comments also hold in this case, i.e., the expert based on local analysis outperforms the other one based on global description and both of them provide complementary information for the verification task.

Another effect which can be observed is the fact that error figures with user-dependent decision thresholding (see Tables 1 and 2) are lower than error figures when user-independent decision thresholds are used (see Figs. 4 and 5). This effect has also been noticed in other signature verification works [31], [30].

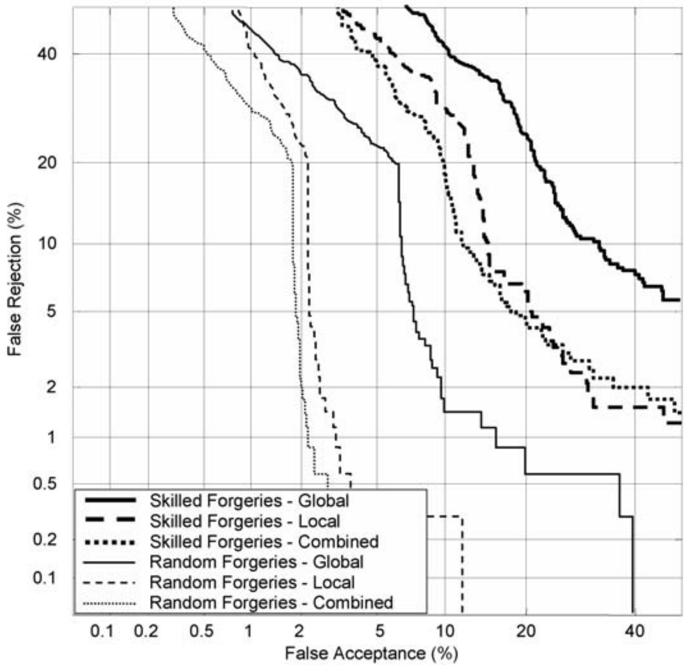


Fig. 5. Verification performance for *a posteriori* user-independent decision thresholding (10 training samples)

6 Conclusions

An off-line signature recognition system based on fusion of local and global analysis of input images has been described.

Slant directions of the signature strokes and those of the envelopes of the dilated signature images are used as features for recognition. Global analysis is based on extracting features for the whole preprocessed signature image and a statistical distance measure. Local analysis is based on extracting features per height-wide blocks and left-to-right Hidden Markov Models.

Experimental results are given on a subcorpus of the large MCYT signature database comprising 2250 different signature images from 75 contributors. Verification performance on random and skilled forgeries is given for user-dependent and user-independent decision thresholds. The machine expert based on local information is shown to outperform the system based on global analysis in all reported cases. The two proposed systems are also shown to give complementary recognition information which is exploited with a simple fixed fusion strategy based on the sum rule. Relative improvements in the verification performance as high as 51% (skilled forgeries) and 78% (random forgeries) are obtained as compared to published works.

Future work includes the evaluation of the proposed system on the full MCYT database [26] and the exploitation of the revealed user-dependencies either by using a trained fusion strategy [32], [33] or by using target-dependent score normalization techniques [30].

7 Acknowledgements

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

References

[1] Jain, A.K., Bolle, R., Pankanti, S.: Biometrics - Personal Identification in a Networked Society. Kluwer (1999)
[2] Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. on Circuits and Systems for Video Technology **14** (2004) 4–20
[3] Plamondon, R., Lorette, G.: Automatic signature verification and writer identification: The state of the art. Pattern Recognition **22** (1989) 107–131
[4] Plamondon, R., Srihari, S.N.: On-line and off-line handwriting recognition: A comprehensive survey. IEEE Trans. on PAMI **22** (2000) 63–84
[5] Lee, L.L., Lizarraga, M.G.: An off-line method for human signature verification. In: Proc. of the Intl. Conf. on Pattern Recognition, ICPR. (1996) 195–198
[6] Guo, J.K., Doermann, D., Rosenfeld, A.: Local correspondence for detecting random forgeries. In: Proc. of ICDAR. (1997) 319–323
[7] Huang, K., Yan, H.: Off-line signature verification based on geometric feature extraction and neural network classification. Pattern Recognition **30** (1997) 9–19
[8] Justino, E., Bortolozzi, F., Sabourin, R.: Off-line signature verification using HMM for random, simple and skilled forgeries. In: Proc. of ICDAR. (2001) 1031–1034
[9] Muramatsu, D., Matsumoto, T.: An HMM signature verifier incorporating signature trajectories. In: Proc. of ICDAR. (2003) 438–442
[10] Ross, A., Jain, A.K., Reisman, J.: A hybrid fingerprint matcher. Pattern Recognition **36** (2003) 1661–1673
[11] Campbell, J.P., Reynolds, D.A., Dunn, R.B.: Fusing high- and low-level features for speaker recognition. In: Proc. of EuroSpeech. (2003) 2665–2668
[12] Bajaj, R., Chaudhury, S.: Signature verification using multiple neural classifiers. Pattern Recognition **30** (1997) 1–7
[13] Cordella, C.P., Foggia, P., Sansone, C., Vento, M.: Document validation by signatures: A serial multi-expert approach. In: Proc. of ICDAR. (1999) 601–604
[14] Kittler, J., Hatef, M., Duin, R., Matas, J.: On combining classifiers. IEEE Trans. on Pattern Anal. and Machine Intell. **20** (1998) 226–239
[15] Fierrez-Aguilar, J., Ortega-Garcia, J., Garcia-Romero, D., Gonzalez-Rodriguez, J.: A comparative evaluation of fusion strategies for multimodal biometric verification. In: Proc. of AVBPA, LNCS-2688, Springer (2003) 830–837
[16] Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. on Systems, Man and Cybernetics **9** (1979) 62–66

[17] Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Addison-Wesley (2002)
[18] Theodoridis, S., Koutroumbas, K.: Pattern Recognition. Academic Press (2003)
[19] Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE **77** (1989) 257–286
[20] Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: A function-based on-line signature verification system exploiting statistical signal modeling. Intl. Journal of Image and Graphics (2004) (accepted).
[21] Bigun, E.S., Bigun, J., Duc, B., Fischer, S.: Expert conciliation for multimodal person authentication systems by Bayesian statistics. In: Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA. (1997) 291–300
[22] Bigun, E.S.: Risk analysis of catastrophes using experts' judgments: An empirical study on risk analysis of major civil aircraft accidents in europe. European J. Operational Research **87** (1995) 599–612
[23] Duda, R.O., Hart, P.E., Stork, D.G.: Pattern classification. Wiley (2001)
[24] Verlinde, P., Chollet, G., Acheroy, M.: Multi-modal identity verification using expert fusion. Information Fusion **1** (2000) 17–33
[25] Duin, R.P.W.: The combining classifier: to train or not to train? In: Proc. of the IAPR Intl. Conf. on Pattern Recognition, ICPR, IEEE CS Press (2002) 765–770
[26] Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., et al.: MCYT baseline corpus: A bimodal biometric database. IEE Proceedings Vision, Image and Signal Processing **150** (2003) 395–401
[27] Jain, A.K., Duin, R., Mao, J.: Statistical pattern recognition: A review. IEEE Trans. on Pattern Anal. and Machine Intell. **22** (2000) 4–37
[28] Martin, A., Doddington, G., Kamm, T., Ordowski, M., Przybocki, M.: The DET curve in assessment of decision task performance. In: Proc. of ESCA Eur. Conf. on Speech Comm. and Tech., EuroSpeech. (1997) 1895–1898
[29] Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., , Jain, A.K.: FVC2000: Fingerprint Verification Competition. IEEE Trans. on PAMI **24** (2002) 402–412
[30] Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Target dependent score normalization techniques and their application to signature verification. In: Proc. of ICBA, LNCS, Springer (2004) (to appear).
[31] Jain, A.K., Griess, F., Connell, S.: On-line signature verification. Pattern Recognition **35** (2002) 2963–2972
[32] Jain, A.K., Ross, A.: Learning user-specific parameters in a multibiometric system. In: Proc. of ICIP. Volume 1. (2002) 57–60
[33] Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Exploiting general knowledge in user-dependent fusion strategies for multimodal biometric verification. In: Proc. of ICASSP, IEEE (2004) (to appear).