IMPACT OF SIGNATURE LEGIBILITY AND SIGNATURE TYPE IN OFF-LINE SIGNATURE VERIFICATION

F. Alonso-Fernandez[a], M.C. Fairhurst[b], J. Fierrez[a] and J. Ortega-Garcia[a].

[a]Biometric Recognition Group - ATVS, Escuela Politecnica Superior - Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{fernando.alonso, julian.fierrez, javier.ortega}@uam.es

[b]Department of Electronics, University of Kent, Canterbury, Kent CT2 7NT, UK
{M.C.Fairhurst}@kent.ac.uk

ABSTRACT

The performance of two popular approaches for off-line signature verification in terms of signature legibility and signature type is studied. We investigate experimentally if the knowledge of letters, syllables or name instances can help in the process of imitating a signature. Experimental results are given on a sub-corpus of the MCYT signature database for random and skilled forgeries. We use for our experiments two machine experts, one based on global image analysis and statistical distance measures, and the second based on local image analysis and Hidden Markov Models. Verification results are reported in terms of Equal Error Rate (EER), False Acceptance Rate (FAR) and False Rejection Rate (FRR). [1]

1. INTRODUCTION

The handwritten signature is one of the most widely used individual authentication methods due to its acceptance in government, legal and commercial transactions as a method of identity verification [1, 2]. As a result, a number of algorithms have been proposed for automatic signature verification [3]. This work is focused on off-line verification, a pattern classification problem with a long history, involving the discrimination of signatures written on a piece of paper [4]. It is worth noting that even professional forensic document examiners perform a correct classification rate of only about 70%, confirming that this a challenging research area.

In this paper, we focus on *occidental* signatures, which typically consist of connected text (i.e. name) and/or some form of flourish. Sometimes, signatures only consist of a readable written name (e.g. American signatures). In other cases, as frequently happens in European countries, signatures may consist of only an elaborated flourish. In contrast to occidental signatures, oriental signatures consist of independent symbols. Examples of both oriental and occidental signatures can be found in the First International Signature Verification Competition [5].

Signature verification systems have been shown to be sensitive to some extent to signature complexity [6]. Easy to forge signatures result in increased False Acceptance Rate. Signature variability also has an impact in the verification rates attainable [7]. It can be hypothesized that these two factors, complexity and variability, are related in some way with signature legibility and signature type. Moreover, some studies have been concerned with the ability of humans in recognizing handwritten script [8, 9]. Knowledge about letters, syllables or name instances may help in the process of imitating a signature, which is not the case for an incomprehensible set of strokes that, in principle, are not related to any linguistic knowledge.

The main goal of this work is to evaluate the impact of signature legibility and signature type on the recognition rates of two popular approaches to off-line signature verification. In this paper, signature legibility and type are assessed by a human expert. Some examples are shown in Figs. 1 and 2. This process is not unreasonable in relation to off-line signature verification environments, where signature acquisition is typically performed by a human operator using a scanner or a camera [4].

Two machine experts with different approaches for feature extraction are used in the work reported here, as described in Section 2. The first is based on global image analysis and a minimum distance classifier as proposed in [10], and further developed in [11]. The second is based on local image analysis and left-to-right Hidden Markov Models as used in [12] but with a local parameterization derived from [10], and also detailed in [11]. The rest of this paper is orga-

[1]This work has been carried out while F. A.-F. was guest scientist at the University of Kent. This work has been supported by Spanish MCYT TEC2006-13141-C03-03 and by European Commission IST-2002-507634 Biosecure NoE projects. Author F. A.-F. thanks Consejeria de Educacion de la Comunidad de Madrid and Fondo Social Europeo for supporting his PhD studies. Author J. F. is supported by a Marie Curie Fellowship from the European Commission.

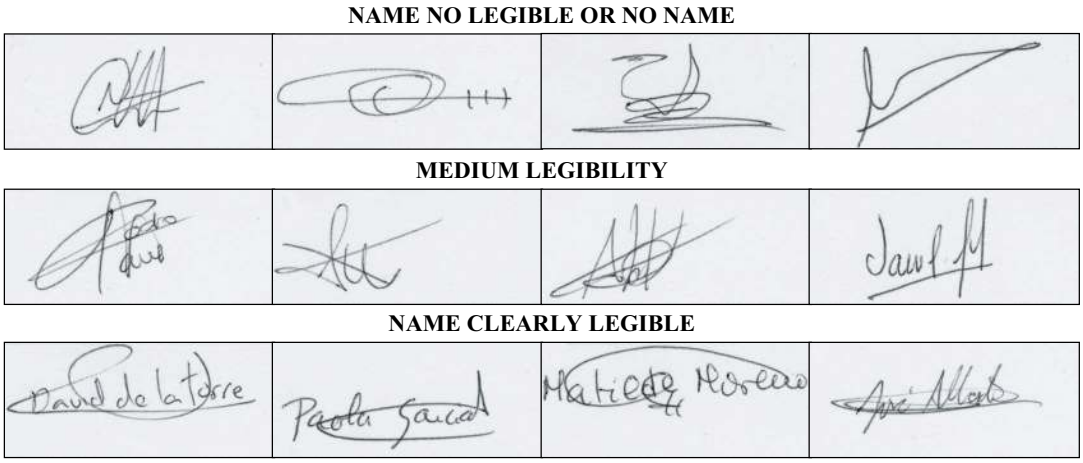

Fig. 1. Signature examples with different degrees of name legibility (from top to bottom).

Fig. 2. Signature examples of the four types encountered in the MCYT corpus (from left to right).

nized as follows. The experimental framework used, including the database, protocol and results, is described in Section 3. Some conclusions are finally drawn in Section 4.

2. MACHINE EXPERTS

In this section, the two machine experts used in this paper are described. They exploit information at two different levels: the first approach analyze the image in a holistic manner, wheres the second approach is based on features extracted locally. Additional details can be found in [11].

2.1. Based on global information

Input signature images are first *preprocessed* according to the following consecutive steps: binarization by global thresholding of the histogram [13], morphological closing operation on the binarized image [14], segmentation of the signature outer traces, and normalization of the image size to a fixed width of 512 pixels while maintaining the aspect ratio (see Fig. 3 for an example). Normalization of the image size is used to make the proportions of different realizations of an individual sample to be the same, whereas segmentation of the outer traces is carried out because a signature boundary typically corresponds to a flourish, which has high intra-user variability. For this purpose, left and right height-wide blocks having all columns with signature pixel count lower than threshold T_p and top and bottom width-wide blocks having all rows with signature pixel count lower than T_p are discarded.

A *feature extraction stage* is then performed, in which slant directions of the signature strokes and those of the envelopes of the dilated signature images are extracted using mathematical morphology operators [14], see Fig. 4. These descriptors are used as features for recognition as proposed in [10]. For slant direction extraction, the preprocessed signature image is eroded with 32 structuring elements, thus generating 32 eroded images. A slant direction feature sub-vector of 32 components is then generated, where each component is computed as the signature pixel count in each eroded image. For envelope direction extraction, the preprocessed signature image is successively dilated 5 times with each one of 6 linear structuring elements, thus generating 5×6 dilated images. An envelope direction feature sub-vector of 5×6 components is then generated, where each component is computed as the

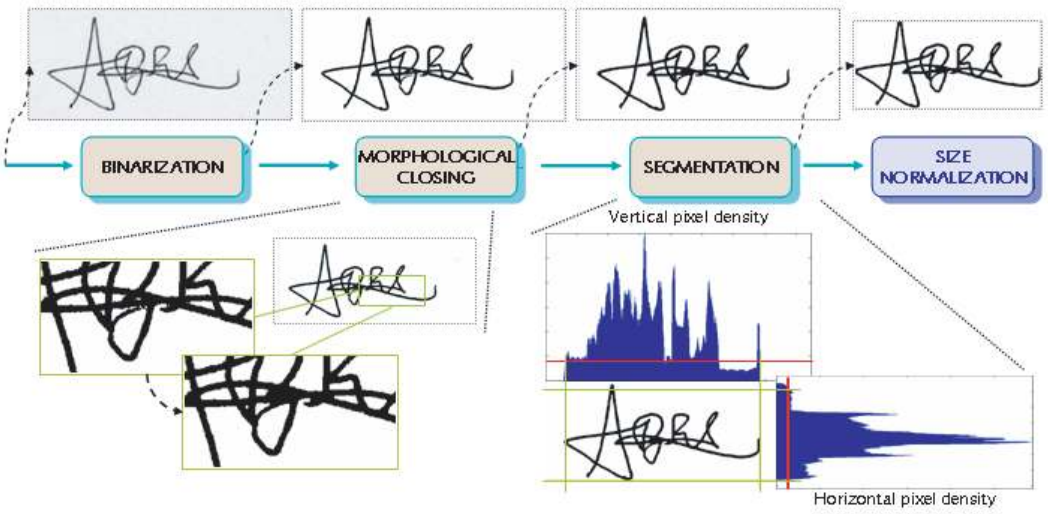

Fig. 3. Preprocessing stage performed in the global expert.

signature pixel count in the difference image between successive dilations. The preprocessed signature is finally parameterized as a vector $\mathbf{o}$ with 62 components by concatenating the slant and envelope feature sub-vectors. Each client (enrolee) of the system is represented by a statistical model $\lambda = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ which is estimated by using an enrolment set of K parameterized signatures $\{\mathbf{o}_1, ..., \mathbf{o}_K\}$. The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ denote mean and standard deviation vectors of the K vectors $\{\mathbf{o}_1, ..., \mathbf{o}_K\}$. In the *similarity computation stage*, the similarity score between a claimed model $\lambda = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ and a parameterized test signature $\mathbf{o}$ is computed as the inverse of the Mahalanobis distance [15].

2.2. Based on local information

In the *preprocessing stage*, images are first binarized and segmented as described in Section 2.1. Next, a *feature extraction step* is performed, in which slant directions and envelopes are locally analyzed using the approach described in Section 2.1, but applied to blocks. Preprocessed images are divided into height-wide blocks of 64 pixels width with an overlapping between adjacent blocks of 75%. The rightmost block is discarded. A signature is then parameterized as a matrix $\mathbf{O}$ whose columns are 62-tuples, each one corresponding to a block. Each client of the system is represented by a Hidden Markov Model λ (HMM) [16, 17], which is estimated by using an enrolment set of K parameterized signatures $\{\mathbf{O}_1, ..., \mathbf{O}_K\}$. A left-to-right topology of four hidden states with no transition skips between states is used in this work. Estimation of the model is made by using the iterative Baum-Welch procedure [16]. The *similarity computation* between a claimed model λ and a parameterized test signature $\mathbf{O}$ is computed by using the Viterbi algorithm [16, 17].

Legibility level	Number of users
Non-legible	18 users (24%)
Medium	19 users (25,33%)
Legible	38 users (50,67%)

Type	Number of users
Simple flourish	5 users (6,67%)
Complex flourish	13 users (17,33%)
Name + simple flourish	35 users (46,67%)
Name + complex flourish	22 users (29,33%)

Table 1. Distribution of users on the MCYT database based on name legibility and signature type.

3. EXPERIMENTAL FRAMEWORK

3.1. Database and protocol

We have used for the experiments a subcorpus of the MCYT bimodal database [18], which includes fingerprint and on-line signature data of 330 contributors. In the case of the signature data, skilled forgeries are also available. Imitators are provided the signature images of the client to be forged and, after an initial training period, they are asked to imitate the shape with natural dynamics. Signature data were acquired 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 the signature images were available on paper. Paper templates of 75 signers (and their associated skilled forgeries) have been digitized with a scanner at 600 dpi (dots per inch). The resulting subcorpus comprises 2250 signature images, with 15

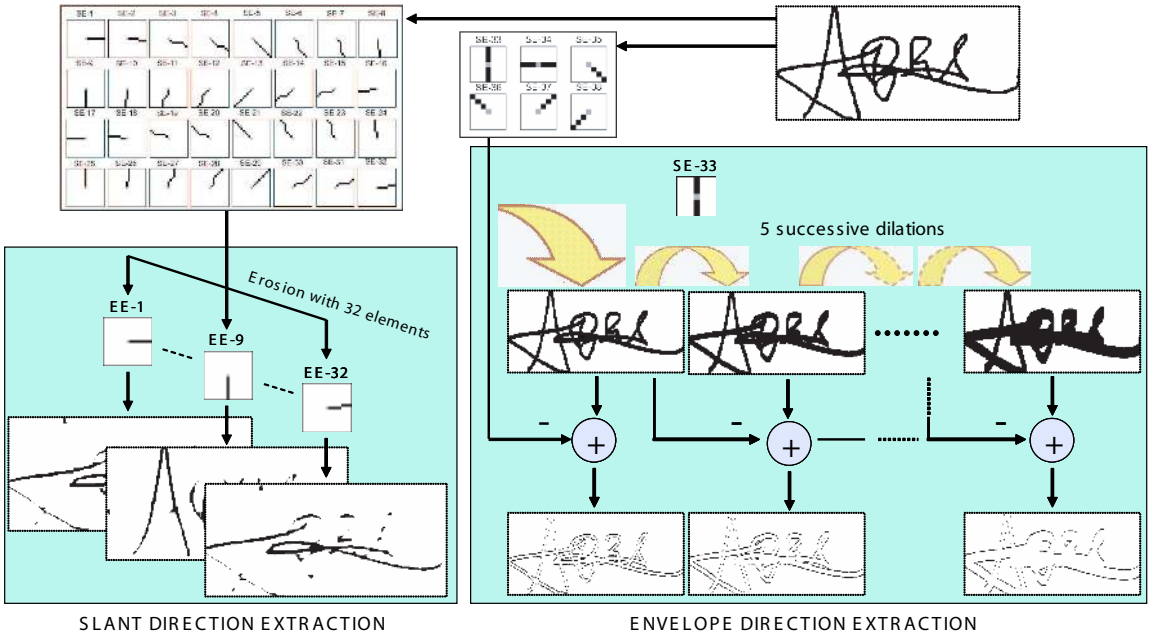

Fig. 4. Feature extraction stage performed in the global expert. Structuring elements used for slant direction extraction (SE-1 to SE-32) and envelope direction extraction (SE-33 to SE-38) are also shown. Origin of the element is indicated in gray. The area of SE-1 to SE-32 is 10 pixels and the angle between successive elements is approximately 11 degrees. The areas of SE-33/34 and SE-35/36/37/38 are 7 and 4 pixels respectively.

genuine signatures and 15 forgeries per user (contributed by 3 different user-specific forgers). Examples can be seen in Figs. 1 and 2.

The experimental protocol is as follows. The training set comprises either 5 or 10 genuine signatures (depending on the experiment under consideration). The remaining genuine signatures are used for testing. For a specific target user, casual impostor test scores are computed by using the genuine samples available from all the remaining targets. Real impostor test scores are computed by using the skilled forgeries of each target. As a result, we have $75 \times 10 = 750$ or $75 \times 5 = 375$ client similarity scores, $75 \times 15 = 1,125$ impostor scores from skilled forgeries, and $75 \times 74 \times 10 = 55,500$ or $75 \times 74 \times 5 = 27,750$ impostor scores from random forgeries.

In order to have an indication of the level of performance with an ideal score alignment between users, results here are based on using *a posteriori* user-dependent score normalization [6]. The score normalization function is as follows $s' = s - s_\lambda(client, impostor)$, where s is the raw score computed by the signature matcher, s' is the normalized matching score and $s_\lambda(client, impostor)$ is the user-dependent decision threshold at a selected point obtained from the genuine and impostor histograms of user λ. In the work reported here, we record verification results at three points: EER, FAR=10% and FRR=10%.

3.2. Results

All signers in the database used for our experiments are manually assigned a *legibility* label and a *type* label. One of three different *legibility* labels is assigned: i) name not legible or no name; ii) uncertain; and iii) name clearly legible. Examples are shown in Fig. 1. Condition ii) is used in the case that some characters of the name can be recognized but it is not possible to extract the name completely. In addition, four different *type* labels are assigned based on the following criterion: a) simple flourish; b) complex flourish; c) name + simple flourish; and d) name + complex flourish. Examples are shown in Fig. 2. It should be noted that signatures of class a) and b) are those assigned to the non-legible class. Similarly, signatures of class c) and d) are those assigned to the medium and legible classes. The distributions of signers in the database based on name legibility and signature type are shown in Table 1.

Table 2 shows the system performance based on name legibility for the two machine experts. Regarding skilled forgeries, we find that the best results are always obtained for the legible case. The non legible case results in no significant improvement in most cases or even worse performance with both machine experts. It could be expected that legible signatures result in worse performance, since they are easier to imitate, because imitators have some background knowledge of what they have to imitate. However, it is observed that legible signatures provide better performance than non legible ones. This may be due to the simplicity of most non-legible

EXPERT BASED ON GLOBAL INFORMATION

		Skilled forgeries				Random forgeries			
TR sign	point	**Non legible**	**Medium**	**Legible**	**Overall**	**Non legible**	**Medium**	**Legible**	**Overall**
5	EER	24.91	26.49	**21.58**	23.78	**8.41**	10.58	9.94	9.79
	FA=10	FR=45.56	FR=44.74	**FR=37.63**	41.47	**FR=11.11**	FR=13.16	FR=15.53	13.73
	FR=10	FA=39.81	FA=53.68	**FA=36.49**	40.44	**FA=13.09**	FA=19.06	FA=15.62	15.41
10	EER	21.11	25.17	**20.55**	22.13	6.57	9.47	**5.97**	7.26
	FA=10	FR=38.89	FR=42.11	**FR=36.32**	38.13	FR=6.67	FR=7.89	**FR=5.26**	6.27
	FR=10	FA=41.29	FA=47.72	**FA=32.28**	38.4	FA=11.46	FA=13.11	**FA=8.50**	10.32

EXPERT BASED ON LOCAL INFORMATION (HMM)

		Skilled forgeries				Random forgeries			
TR sign	point	**Non legible**	**Medium**	**Legible**	**Overall**	**Non legible**	**Medium**	**Legible**	**Overall**
5	EER	16.67	21.23	**16.54**	17.76	**4.45**	5.26	5.59	5.21
	FA=10	FR=35.00	FR=39.47	**FR=27.37**	32.4	**FR=1.67**	FR=4.21	FR=6.58	4.8
	FR=10	FA=24.82	FA=37.19	**FA=22.11**	26.84	**FA=4.14**	FA=4.58	FA=5.62	5.03
10	EER	16.67	20.00	**10.61**	14.44	**1.51**	2.28	3.27	2.74
	FA=10	FR=23.33	FR=31.58	**FR=18.42**	22.93	**FR=0.00**	FR=1.05	FR=4.74	2.67
	FR=10	FA=22.22	FA=32.63	**FA=16.84**	22.04	**FA=1.81**	FA=4.69	FA=4.35	3.82

Table 2. System performance based on name legibility. Results are given in %.

signatures.

Regarding random forgeries, we observe from Table 2 that for the expert based on global information, improvement achieved depends on the number of signatures used for enrolment. When using 5 signatures, the best results are obtained for the non legible case, whereas when using 10 signatures, the best results are for the legible signature case. On the other hand, for the machine expert based on local information, the best performance is always obtained for the non legible case.

System performance in relation to signature type is shown in Table 3. Regarding skilled forgeries, Table 2 shows that non legible signatures resulted in no significant improvement with either expert. If we divide non legible signatures into "simple flourish" and "complex flourish", we observe that complex flourish signatures result in improved performance. This could be because simple flourish signatures are easier to imitate than complex flourish ones. It is also worth noting that signatures classified as "name + simple flourish" result in better performance with the global expert, but a worse performance is obtained with the local expert. The opposite happens with the "name + complex flourish" samples. This could be because, since the local machine expert processes signature images by blocks, it better deals with most complex signatures such as the "name + complex flourish" ones. In complex signatures, there are regions of the signature image having various strokes crossing in several directions. The global machine expert is not able to deal satisfactorily with this case, since it processes the signature image as a whole.

Regarding random forgeries, we observe from Table 3 that signatures classified as "name + complex flourish" always result in worse performance with both machine experts. Signatures classified as "name + simple flourish" result in improved performance with the global expert, but worse performance is obtained with the local expert in most cases. The opposite happens with the "complex flourish" signatures. Also interestingly, simple flourish signatures always work well with the local expert, but this is not the case with the global expert, in which the performance becomes poorer as we increase the number of signatures for enrolment.

4. CONCLUSIONS

In this paper, we evaluate the impact of signature legibility and signature type on the recognition rates of off-line signature verification systems. For our experiments, we have used two machine experts that exploit information at two different levels. The first is based on global image analysis and a statistical distance measure, whereas the second is based on local image analysis and left-to-right Hidden Markov Models.

Regarding name legibility criteria, similar behaviour is found for both machine experts for the skilled forgeries experiments. The best results are always obtained for the legible case, whereas the non legible case results in no significant improvement, or even worse performance.

It could be expected that legible signatures result in worse performance for skilled forgeries, since they are easier to imitate, however this is not the case in our experiments. Characteristics such as signature complexity or stability could have clearer impact in the performance [7, 19] and this will be the target of future work. In our experiments, we observe that the most complex signatures ("name + complex flourish") are quite robust to skilled forgeries using the HMM system, although they are not suitable to discriminate between different signers (i.e. random forgeries). The opposite happens with the most simple signatures ("simple flourish").

Exploiting differences in performance of several matchers with respect to a measurable criteria can be used to improve verification rates, as shown in other biometric traits (e.g. see [20]). For instance, the steps of the recognition system can be adjusted or different matchers can be invoked based on the measured criteria.

EXPERT BASED ON GLOBAL INFORMATION

TR sign	point	Skilled forgeries: Simple flourish	Skilled forgeries: Complex flourish	Skilled forgeries: Name + simple fl.	Skilled forgeries: Name + complex fl.	Skilled forgeries: Overall	Random forgeries: Simple flourish	Random forgeries: Complex flourish	Random forgeries: Name + simple fl.	Random forgeries: Name + complex fl.	Random forgeries: Overall
5	EER	26.33	23.72	**20.33**	28.18	23.78	**4.14**	10.06	7.24	14.74	9.79
	FA=10	FR=68	FR=36.92	**FR=35.14**	FR=47.73	FR=41.47	**FR=0.00**	FR=15.38	FR=9.71	FR=22.73	FR=13.73
	FR=10	FA=37.33	FA=40.77	**FA=36**	FA=49.70	FA=40.44	**FA=2.89**	FA=17.06	FA=8.05	FA=29.21	FA=15.41
10	EER	**20**	21.12	22.32	22.41	22.13	7.97	6.94	**5.70**	9.53	7.26
	FA=10	FR=48	**FR=35.38**	FR=36.57	FR=40.91	FR=38.13	**FR=4.00**	FR=7.69	FR=4.57	FR=8.64	FR=6.27
	FR=10	FA=57.33	**FA=34.87**	FA=35.05	FA=42.12	FA=38.4	FA=19.43	**FA=8.41**	FA=8.68	FA=12.24	FA=10.32

EXPERT BASED ON LOCAL INFORMATION (HMM)

TR sign	point	Skilled forgeries: Simple flourish	Skilled forgeries: Complex flourish	Skilled forgeries: Name + simple fl.	Skilled forgeries: Name + complex fl.	Skilled forgeries: Overall	Random forgeries: Simple flourish	Random forgeries: Complex flourish	Random forgeries: Name + simple fl.	Random forgeries: Name + complex fl.	Random forgeries: Overall
5	EER	25.67	13.85	21.57	**12.58**	17.76	**4.00**	4.67	4.86	6.41	5.21
	FA=10	FR=52.00	FR=28.46	FR=36.29	**FR=24.10**	32.4	FR=2.00	**FR=1.54**	FR=5.14	FR=6.82	4.8
	FR=10	FA=42.67	FA=18.72	FA=33.52	**FA=17.58**	26.84	**FA=3.84**	FA=4.36	FA=4.90	FA=6.10	5.03
10	EER	25.33	12.82	15.33	**11.82**	14.44	**0.03**	2.08	1.71	4.84	2.74
	FA=10	FR=36.00	FR=18.46	FR=25.71	**FR=18.18**	22.93	**FR=0.00**	**FR=0.00**	FR=3.43	FR=3.64	2.67
	FR=10	FA=29.33	**FA=20.00**	FA=22.48	FA=21.21	22.04	**FA=0.22**	FA=2.39	FA=2.72	FA=7.26	3.82

Table 3. System performance based on signature type. Results are given in %.

5. REFERENCES

[1] M.C. Fairhurst, "Signature verification revisited: promoting practical exploitation of biometric technology," *Electronics and Communication Engineering Journal*, vol. 9, pp. 273–280, December 1997.

[2] A.K. Jain, A. Ross, S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 14, no. 1, pp. 4–20, 2004.

[3] G. Dimauro et al., "Recent advancements in automatic signature verification," *Proc. IWFHR*, pp. 179–184, 2004.

[4] R. Plamondon and S.N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. on PAMI*, vol. 22, no. 1, pp. 63–84, 2000.

[5] D.Y. Yeung et al., "SVC2004: First international signature verification competition," *Proc. ICBA, Springer LNCS-3072*, pp. 15–17, July 2004.

[6] J. Fierrez-Aguilar, J. Ortega-Garcia, and J. Gonzalez-Rodriguez, "Target dependent score normalization techniques and their application to signature verification," *IEEE Trans. SMC-C*, vol. 35, no. 3, 2005.

[7] C. Allgrove and M.C. Fairhurst, "Enrolment model stability in static signature verification," in *in Proc. IWFHR*, pp. 565–570, 2000.

[8] J.J. Brault, R. Plamondon, "A complexity measure of handwritten curves: Modeling of dynamic signature forgery," *IEEE Trans. SMC*, vol. 23, pp. 400–413, 1993.

[9] M.C. Fairhurst and E. Kaplani, "Perceptual analysis of handwritten signatures for biometric authentication," *IEE Proc. VISP*, vol. 150, pp. 389–394, 2003.

[10] L.L. Lee and M.G. Lizarraga, "An off-line method for human signature verification," in *Proc. ICPR*, 1996, p. 195198.

[11] J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, and J. Ortega-Garcia, "An off-line signature verification system based on fusion of local and global information," in *Proc. BIOAW, Springer LNCS-3087*, 2004, pp. 295–306.

[12] E. Justino, F. Bortolozzi, R. Sabourin, "Off-line signature verification using HMM for random, simple and skilled forgeries," *Proc. ICDAR*, pp. 1031–1034, 2001.

[13] N. Otsu, "A threshold selection method for gray-level histograms," *IEEE Trans. on SMC*, vol. 9, pp. 62–66, December 1979.

[14] R.C. Gonzalez and R.E Woods, *Digital Image Processing*, Addison-Wesley, 2002.

[15] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, 2003.

[16] L.R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, 1989.

[17] J. Ortega-Garcia, J. Fierrez-Aguilar, J. Martin-Rello, and J. Gonzalez-Rodriguez, "Complete signal modelling and score normalization for function-based dynamic signature verification," *Proc. AVBPA, Springer LNCS-2688*, pp. 658–667, 2003.

[18] J. Ortega-Garcia et al., "MCYT baseline corpus: a bimodal biometric database," *IEE Proc. VISP*, vol. 150, no. 6, pp. 395–401, December 2003.

[19] M.C. Fairhurst, E. Kaplani, and R.M. Guest, "Complexity measures in handwritten signature verification," *Proc. UAHCI*, pp. 305–309, 2001.

[20] J. Fierrez-Aguilar and Y. Chen and J. Ortega-Garcia and A.K. Jain, "Incorporating image quality in multi-algorithm fingerprint verification," *Proc. ICB, Springer LNCS-3832*, pp. 213-220, 2006.