

Practical On-Line Signature Verification^{*}

J.M. Pascual-Gaspar, V. Cardenoso-Payo, and C.E. Vivaracho-Pascual

ECA-SIMM, Dpto. Informática, Universidad de Valladolid,
Campus Miguel Delibes s/n, 47011 Valladolid, Spain
pascual@gaspar.com, valen@infor.uva.es, cevp@infor.uva.es

Abstract. A new DTW-based on-line signature verification system is presented and evaluated. The system is specially designed to operate under realistic conditions, it needs only a small number of genuine signatures to operate and it can be deployed in almost any signature capable capture device. Optimal features sets have been obtained experimentally, in order to adapt the system to environments with different levels of security. The system has been evaluated using four on-line signature databases (MCYT, SVC2004, BIOMET and MyIDEA) and its performance is among the best systems reported in the state of the art. Average EERs over these databases lay between 0.41% and 2.16% for random and skilled forgeries respectively.

1 Introduction

Handwriting signature has been a mean of personal identification used for centuries and its usage in experimental scenarios as a dynamic biometric modality has proved to be as efficient as many physiological traits [1]. However, this biometric modality faces big challenges when applied to real practical scenarios, far away from ideal laboratory conditions. Working with a small number of genuine signatures for user enrolment is one of those, which motivates the Dynamic Signature Verification (DSV) system which is described and evaluated in this work. The following design guidelines have been applied in order to ensure good practical usability characteristics:

- The system shall use a reduced number of genuine signatures without compromising accuracy and no extra development data sets will be required.
- The signature feature sets should be easily computable and cross-device compatible. Small storage requirements would be desirable.
- The verification algorithm should allow easy and efficient implementation using freely available programming environments.
- The system must be flexible enough as to be adapted to different security restrictions.

^{*} This work has been partially supported by Consejería de Educación de la Junta de Castilla y León under research contract VA077A08.

2 System Design

This section describes the four building blocks that make up our system, according to Jain's classification of a biometric system in [2]: sensor module, feature extraction module, matcher module and system database.

2.1 Sensor Module

On-line signatures can be obtained through a variety of devices (digital pens, personal digital assistants (PDAs), Tablet-PCs, ...). Digital pen tablets usually provide higher spatial and temporal resolution at affordable prices. That's why many public available signature databases have been acquired using pen tablets, and in particular the four databases used in this work. As will be shown later, our experiments show that the results we report here could be similarly obtained using other devices with almost no modifications.

Pen tablets record temporal and writing gesture information into a sequence of vector samples, at a typical fixed rate of around 100Hz. Raw features f_k can be classified into positional and ergonomical: a) (*p*)ositional features $f_k^p \in \{x_k, y_k\}$, a 2D point in the path followed by the pen; b) (*e*)rgonomical features $f_k^e = \{p_k, a_k, i_k\}$, resulting from the hand-pen interaction. They include the *pressure* p exerted by the pen on the tablet and two pen orientation angles (*azimuth* a and *inclination* i).

2.2 Feature Extraction Module

A raw on-line signature S is represented by a discrete time sequence of N vectors of features, which number is calculated in terms of signature duration and sampling rate:

$$S = \{(x_t, y_t, p_t, a_t, i_t)\}_{t=1..N} = \{f_{k,t}\}_{t=1..N}^{k=1..5} \quad (1)$$

One of three alternative strategies is usually used by DSV systems to select the final feature set from the raw signature parameters:

1. Direct use of the raw features provided by the sensor [3,4]. Although this approach does not exclude some simple preprocessing tasks, emphasis usually focus on the matcher module.
2. Generate larger sets of derived features and empirically select the most effective using commonly accepted previous results [5,6]. This approach can lead to high-accuracy systems but it is more computationally intensive and the results could be database-dependent [7].
3. Select a combined set of raw and simple derived features which gives optimal performances [8]. This could be the best alternative for real-time, cross-database and cross-device practical applications when statistical classification techniques are applied to feature selection, as we have done in this work.

In our system the basic feature set $F = (x, y, p, a, i)$ was expanded to include first and second time derivatives, to make a final raw feature vector with 15 components: $\widehat{F} = \{F, \Delta F, \Delta\Delta F\} = \{f_k, df_k, ddf_k\}$ with $df_k = (f_{k,t+1} - f_{k,t})/\Delta t$ and $ddf_k = (df_{k,t+1} - df_{k,t})/\Delta t$.

Two normalization process were applied to get the final set of features from this raw set: a *geometrical translation* $N_1 : f_k^{N_1} = f_k^p - \mu_k^p$ to locate the geometric center of the signature at the origin of coordinates, and a *statistical normalization* based on z-norm scaling $N_2 : f_k^{N_2} = (f_k - \mu_k)/\sigma_k$ was applied to all features so that zero mean and unit variance was ensured for each feature in the vector, where $\mu_k = (\sum_{t=1}^N f_{k,t})/N$ and $\sigma_k = \sqrt{(\sum_{t=1}^N (f_{k,t} - \mu_k)^2)/(N - 1)}$ are the mean and standard deviation of k^{th} -feature respectively.

2.3 Matcher Module

The two most common alternatives used to determine the similarity between time series associated to on-line signatures are the *reference-based* and *model-based* approaches. Reference-based systems require storing several instances of genuine signatures in order to evaluate the intra-class variability (Fig.1c). Model-based systems do not need to store signature specimens, but just a compact representation of the parameters of the model. Both alternatives have been successfully used with similar performance results in the state-of-art systems [1]. In our system, we have chosen a reference-based approach using Dynamic Time Warping (DTW) for time series alignment. This approach combines high accuracy results [9,8] with efficient implementation under a wide spectrum of practical scenarios, which is a goal of our final DSV system.

DTW provides the optimal nonlinear alignment of two sequences of vectors, through a minimization of the overall accumulated distance along the aligning sequence. The distance between a reference signature $S_R = \{r_i\}_{i=1..N}$ and a test signature $S_T = \{t_j\}_{j=1..M}$ is calculated by filling a $DTW_{N+1 \times M+1}$ matrix following equation (2), after initialization of $DTW[0, 0] = 0$ and $DTW[i, 0] = DTW[0, j] = \infty \forall i, j \in [1, N]$:

$$DTW[i, j] = \underbrace{dist(i, j)}_{\text{current cost}} + \overbrace{\begin{cases} dist(i - 1, j) \\ dist(i, j - 1) \\ dist(i - 1, j - 1) \end{cases}}^{\text{accumulated cost}} \tag{2}$$

The distance between reference and test signatures will be stored at the upper right corner of the DTW matrix: $Dist(S_R, S_T) = DTW[N, M]$. The local distance $dist$ in equation 2 was the usual Euclidean vector distance $dist(f_i^R, f_j^T) = \sqrt{\sum_{k=1}^5 (f_{i,k}^R - f_{j,k}^T)^2}$.

To deal with intra-class variability, inherent to the signing process, a number of genuine signature samples should be stored for each user. Previous results show that five signatures is a reasonably low number and could still provide

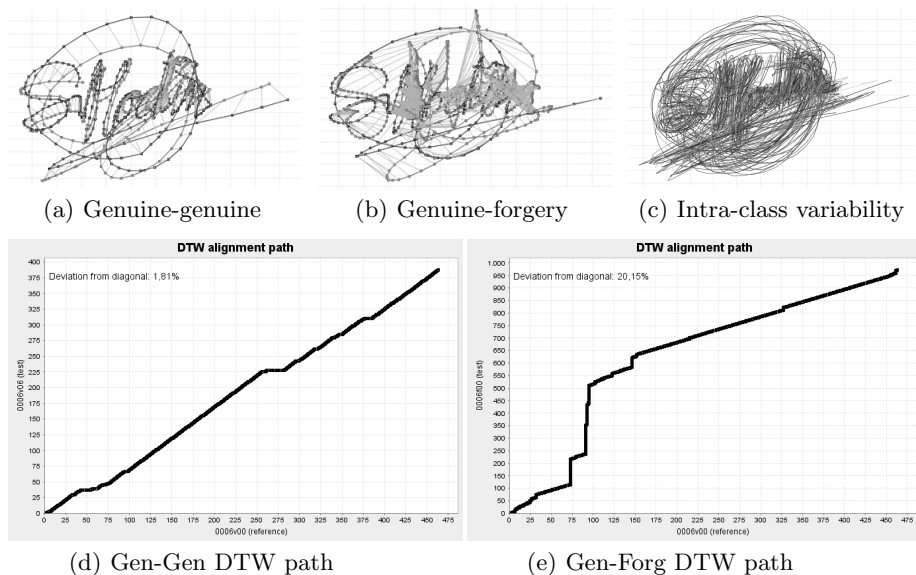


Fig. 1. Figures a) and b) illustrates the DTW alignment between pairs of genuine-genuine and genuine-forgery signatures and shows their respective alignment paths in d) and e). Visual appearance of the alignment path looks closer to a diagonal straight line when genuine samples are compared. Figure c) shows an example of intra-class variability.

good results in practical scenarios [10]. DTW distances between a test signature and the five reference signatures were combined using arithmetic mean and ten-fold cross-validation was carried out to compute a reliable average error value. Five different reference signatures were randomly chosen in each fold and all the remaining genuine signatures of the same user were used for testing.

The performance of our system will be presented using the typical EER measure, both for random and skilled forgery scenarios. This allows easy comparison with results obtained in referenced works for the same databases.

2.4 System Database Module

Four on-line signature databases compiled by different research groups have been used in this work (table 1). These databases were acquired using similar models of pen tablets, providing a common set of signature features. All these databases include skilled forgeries, which is mandatory to test the system in practical scenarios subjected to ‘professional forgers’ attacks. Another criterion to select these databases had to do with the number of experiments reported on them in the literature.

The signature databases set was split in two disjoint data sets: a) a *Development Data (DD)* set, including signatures from the 50 first users of MCYT-100 [11] (hereby MCYT-A), was used to get the optimal features sets, and b) a *Test*

Table 1. Main figures of the two data sets used in this work. Users for whom there were no forgeries available were not included in our experiments. Further details of each database can be found in their reference papers.

Dataset	Database	Users	Signatures		
			Genuines	Forgeries	Total
DD	MCYT-A	50	25	25	2500
TD	MCYT-B	50	25	25	2500
	SVC2004	40	20	20	1600
	BIOMET	84	15	17	2688
	MYIDEA	69	18	36	3726
Total		293	5802	7212	13014

Data (TD) set, with the remaining signatures of all other databases, was used to test the system. *TD* set contains the signatures of the last 50 users of MCYT-100 (hereby MCYT-B) and all signatures of SCV2004¹ [1], BIOMET [12] and MyIDEA [13].

3 Selection of Optimal Feature Sets

The tablets used to acquire the signatures we used in our experiments register both ink (visible) and pen-up strokes. This behaviour is hardware-dependent and many signature capturing devices could not provide hidden strokes. Thus, we removed them from each signature for the experiments. First challenge we faced to increase system performance was to properly and efficiently select an optimal set of features for further experimentation. To this end, both individual and combined feature evaluations have been carried out.

In a first stage, the 15 available raw features were evaluated individually on *DD* set to reduce the dimension of the combined features space. Fig. 2 proves that pen angle features (azimuth a and inclination i) show significantly poorer performance than the rest, for the three signal domains. These results suggest removal of these features, reducing feature space dimension from 15 to 9, which implies a 98% reduction of the number of possible feature combinations (from $2^{15} - 1 = 32767$ to $2^9 - 1 = 511$) and brings a computationally cheaper feature selection process at no relevant performance loss. Additionally, the remaining features (x, y coordinates and pressure p) made up a more common set of features available in the vast majority of ink capturing devices.

After an initial set of isolated features was determined, next step was to select the way to combine them into an almost-optimal final set of features. Three classical feature selection techniques were analyzed [14]: a) *Sequential Forward Selection* (SFS), b) *Sequential Backward Selection* (SBS) and c) *Plus l-take away r Selection* (PTA(l,r)).

SFS progressively incorporates most promising features into larger subsets in such a way that once a feature is added to the set, it cannot be discarded later. In terms of computational cost, this is the most attractive solution, because the

¹ Only the public available development set part was used -40 users-.



Fig. 2. Evaluation results for individual features. By averaging the random and skilled EER's, the best individual feature is the y-coordinate (3.22%) followed by the x-coordinate (3.61%). Pressure provides intermediate performance (11.12%) but azimuth (24.03%) and inclination (24.54%) perform worse than others and were taken out from the feature selection process.

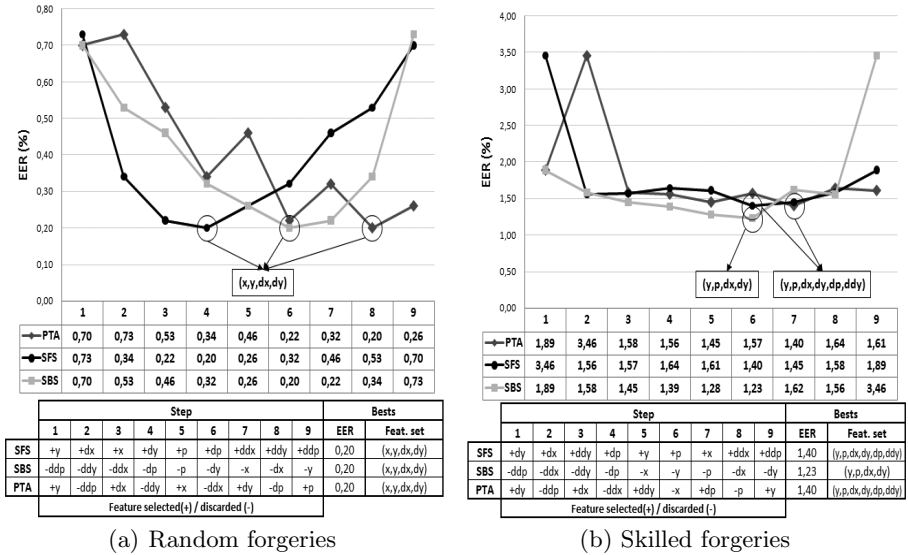


Fig. 3. Evaluation results for combined features. *DD* set for random and skilled scenarios. Curves plot EER evolution at each step of the feature selection process (middle tables below the curves show their numeric values). Tables at the bottom show the order in which features are incorporated or discarded. In both cases, minimum values are significantly lower than extremals. Error is reduced up to 73% (from 0.73% to 0.20%) for random forgeries. Error reduction is 64% (from 3.46% to 1.23%) for skilled forgeries.

Table 2. Standard vs optimal features sets for both scenarios. Minimum values for standard and optimal feature sets in each column are emphasized. Columns in the middle show average over all databases. Lighter shadowed cells show the best feature set for standard and optimal sets. Last row displays relative EER reduction between the best standard feature set and the optimal set; darker cells show average EER reduction of the average best standard set of features ($P + V$) versus the optimal set for each scenario.

		Random forgeries					Skilled forgeries				
Feat. set		mcyt-b	svc04	biomet	myidea	mean	mcyt-b	svc04	biomet	myidea	mean
Stand.	P	5.64	0.78	7.58	2.45	4.11	6.53	4.70	5.41	2.89	4.88
	$P + V$	3.25	0.40	4.47	3.52	2.91	4.21	4.15	3.69	3.25	3.83
	$P + V + A$	4.82	1.01	6.58	6.14	4.64	4.23	6.14	4.43	4.10	4.73
Opt.	F_{rd}^o	0.38	0.00	0.33	0.92	0.41	1.16	3.70	1.25	2.94	2.26
	F_{sk}^o	0.46	0.32	0.96	2.39	1.03	1.06	3.38	1.48	2.72	2.16
EER red.		88.3%	100.0%	92.6%	62.4%	86.0%	74.8%	18.6%	59.9%	5.9%	40.8%

Table 3. Comparison with selected reference systems. In bold are the bests EER(%) results for each database/scenario. Overall, our system provides better performance in all cases, using similar number of training signatures. System 8 provides better results for the random forgery scenario on MCYT.

Author	Year	MCYT		SVC2004		Biomet		MyIDea		Comments
		rd	sk	rd	sk	rd	sk	rd	sk	
1 Hennebert [16]	2007							2.7	7.3	- 6 training signatures - GMM-based algorithm - Results of the signature expert with time variability.
2 Humm [17]	2007							2.6	7.3	- 6 training signatures - HMM-based algorithm - Results with time variability schema.
3 Garcia-Salicetti [18]	2007	1.22	3.40							- 5 training signatures - HMM + Distance based algorithm - Best results from individual systems combination - Random test also included skilled forgeries
4 Van-Bao [6]	2007		3.37		4.83		2.33			- 5 training signatures - HMM-based algorithm - Viterbi path and likelihood fusion intra-algorithm
5 Pascual-Gaspar [19]	2007	2.09	6.14							- 3 training signatures - HMM-based algorithm - HMM with user-dependent structure
6 SVC2004 official results [1]	2004			3.02	6.90					- 5 training signatures - Best skilled system: DTW; Best random system: HMM - Results on development set (40 users) for Task2
7 Fierrez-Aguilar [20]	2005			0.15	6.91					- 5 training signatures - Local (DTW) and Regional (HMM) fusion - Bests results on development set (40 users) for Task2
8 Fierrez-Aguilar [5]	2005	0.24	2.12							- 5 training signatures - Global (Parzen WC) and local (HMM) experts fusion
9 Fierrez-Aguilar [21]	2007	0.05	0.74							- 10 training signatures* - HMM-based algorithm
Our system	2008	0.29	1.23	0.00	3.38	0.33	1.48	0.92	2.72	- 5 training signatures - DTW-based algorithm - Results with scenario-dependent optimal features - Results on MCYT-100

* Although system 9 uses a more training signatures, it is included for future comparisons due to its excellent results.

size of the evaluated sets of features are kept low up to the final steps of the procedure. SBS deletes one feature at a time, so that it cannot be brought back to the locally optimal subset once it has been discarded. Although this method is computationally expensive, defenders argue that it better takes into account inter-features dependencies [15].

A combined and more sophisticated technique, named *Plus l-take away r Selection* was also evaluated (using $l = r = 1$). This approach aims a balance between computational cost and more adequate treatment of inter-features dependencies. Additionally, this method avoids the nesting problems which arise in SFS and SBS feature selection solutions [14].

Figure 3-a) shows that the three described feature selection methods resulted in the same optimal feature set for the random forgeries scenario (x, y, dx, dy) . This indicates the robustness of this combination of features for low FRR (false rejection rate) scenarios. It also suggests that pressure feature could be simply ignored in low to medium security environments, which leads to a widening in the range of capturing devices which could be used for signature acquisition in those situations. For skilled forgeries, best results are obtained using the *SBS* method, which provides the optimal combination (y, dx, dy, p) , made up of geometric and pressure values. In this case, both *SFS* and *PTA* find a slightly worse solution than *SBS*, as would be initially expected, but with less computational cost. This result indicates that for high-security environments pressure information should be present, even when it has the side effect of a small increase of false rejections.

4 Benchmark Results

Two benchmark tests were carried out on the development data (DD) set to evaluate our system. Table 2 shows error results when evaluating with different features sets. Optimal sets for random (F_{rd}^o) and skilled (F_{sk}^o) scenarios are compared with other standard features sets which combine the three signal domains (*[P]osition*, *[V]elocity* and *[A]cceleration*). Using optimal sets for each scenario drastically improves the accuracy of the system for the verification task, specially in low and medium security scenarios, where no genuine user rejections are desired. Except for the BIOMET database, the optimal set of features for a given scenario (e.g. random or skilled) outperforms the results obtained with the set of features of the other one (e.g. skilled or random) in all cases.

Finally, table 3 compares performance results of our system with the ones in other recently published systems which used the same signature databases we used.

5 Conclusions

We described a new DTW-based on-line signature verification system specially designed to be used in practical scenarios. It does not need special hardware features to get good performance results, just geometric coordinates and, optionally, pen pressure. The system needs only a reduced number of signatures

from the user to bring excellent verification results. Depending on the security requirements, different feature sets could be selectively chosen. Benchmark experiments carried out over four popular on-line signature databases (MCYT, SVC2004, BIOMET and MYIDEA) prove that our system provides excellent results in terms of EER, specially for the skilled forgery scenario, where the system clearly outperforms other up-to-date systems in the literature under similar testing conditions.

References

1. Yeung, D., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First international signature verification competition. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 16–22. Springer, Heidelberg (2004)
2. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. *IEEE transactions on circuits and systems for video technology* 14(1) (2004)
3. Faundez-Zanuy, M.: On-line signature recognition based on VQ-DTW. *Pattern Recognition* 40(3), 981–992 (2007)
4. Vivaracho-Pascual, C., Faundez-Zanuy, M., Pascual, J.M.: An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances. *Pattern Recognition* 42(1), 183–193 (2009)
5. Fierrez-Aguilar, J., Nanni, L., Lopez-Peñalba, J., Ortega-Garcia, J., Maltoni, D.: An on-line signature verification system based on fusion of local and global information. In: Kanade, T., Jain, A., Ratha, N.K. (eds.) AVBPA 2005. LNCS, vol. 3546, pp. 523–532. Springer, Heidelberg (2005)
6. Ly, V.B., Garcia-Salicetti, S., Dorizzi, B.: On using the viterbi path along with hmm likelihood information for online signature verification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 37(5), 1237–1247 (2007)
7. Jain, A.K., Bolle, R.M., Pankanti, S.: *Biometrics: Personal Identification in Networked Society*. Springer, Heidelberg (2005)
8. Kholmatov, A., Yanikoglu, B.: Identity authentication using improved online signature verification method. *Pattern Recognition Letters* 26(15), 2400–2408 (2005)
9. Jain, A.K., Griess, F.D., Connell, S.D.: On-line signature verification. *Pattern Recognition* 35(12), 2963–2972 (2002)
10. Fierrez, J., Ortega-Garcia, J.: 12: On-line signature verification. In: *Handbook of Biometrics*, pp. 189–209. Springer, Heidelberg (2008)
11. Ortega-Garcia, J., Fierrez-Aguilar, J., Simon-Zorita, D., Gonzalez-Rodriguez, J., Hernaez, I., Igarza, J.J., Vivaracho, C., Escudero, D., Moro, Q.I.: MCYT baseline corpus: a bimodal biometric database. *IEE Proc. Visual Image Signal Processing* 150(6), 395–401 (2003)
12. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., les Jardins, J.L., Lunter, J., Ni, Y., Petrovska-Delacretaz, D.: BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In: Kittler, J., Nixon, M.S. (eds.) AVBPA 2003. LNCS, vol. 2688, pp. 845–853. Springer, Heidelberg (2003)
13. Dumas, B., Pugin, C., Hennebert, J., Petrovska-Delacretaz, D., Humm, A., Evequoz, F., Ingold, R., von Rotz, D.: MyIDEa - Multimodal biometrics database, description of acquisition protocols. In: *Proceedings of Third COST 275 Workshop (COST 275)*, Hatfield (UK), October 27–28, 2005, pp. 59–62 (2005)

14. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1), 4–37 (2000)
15. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* 3, 1157–1182 (2003)
16. Hennebert, J., Humm, A., Ingold, R.: Modelling spoken signatures with gaussian mixture model adaptation. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)* (2007)
17. Humm, A., Hennebert, J., Ingold, R.: Hidden markov models for spoken signature verification. In: *First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007*, pp. 1–6 (September 2007)
18. Garcia-Salicetti, S., Fierrez-Aguilar, J., Alonso-Fernandez, F., Vielhauer, C., Guest, R., Allano, L., Trung, T.D., Scheidat, T., Van, B.L., Dittmann, J., Dorizzi, B., Ortega-Garcia, J., Gonzalez-Rodriguez, J., di Castiglione, M.B., Fairhurst, M.: Biosecure reference systems for on-line signature verification: A study of complementarity. *Annals of Telecommunications, Special Issue on Multimodal Biometrics* 62(1-2), 36–61 (2007)
19. Pascual-Gaspar, J., Cardenoso-Payo, V.: On-line signature verification using hidden markov models with number of states estimation from the signature duration (September 2007)
20. Fierrez-Aguilar, J., Krawczyk, S., Ortega-Garcia, J., Jain, A.K.: Fusion of local and regional approaches for on-line signature verification. In: Li, S.Z., Sun, Z., Tan, T., Pankanti, S., Chollet, G., Zhang, D. (eds.) *IWBRS 2005. LNCS*, vol. 3781, pp. 188–196. Springer, Heidelberg (2005)
21. Fierrez, J., Ramos-Castro, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: HMM-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters* 28(16), 2325–2334 (2007)