Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 2325–2334
Pattern Recognition Letters
www.elsevier.com/locate/patrec

HMM-based on-line signature verification: Feature extraction and signature modeling

Julian Fierrez *, Javier Ortega-Garcia, Daniel Ramos, Joaquin Gonzalez-Rodriguez

Biometric Recognition Group – ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, C/Francisco Tomas y Valiente 11, 28049 Madrid, Spain

Received 22 March 2007; received in revised form 28 June 2007
Available online 1 August 2007

Communicated by A.M. Alimi

Abstract

A function-based approach to on-line signature verification is presented. The system uses a set of time sequences and Hidden Markov Models (HMMs). Development and evaluation experiments are reported on a subcorpus of the MCYT bimodal biometric database comprising more than 7000 signatures from 145 subjects. The system is compared to other state-of-the-art systems based on the results of the First International Signature Verification Competition (SVC 2004). A number of practical findings related to feature extraction and modeling are obtained.

Keywords: Biometrics; On-line signature; Feature extraction; Hidden Markov Model; Recognition; Verification

1. Introduction

Automatic signature verification has been an intense research area because of the social and legal acceptance and widespread use of the written signature as a personal authentication method (Plamondon and Lorette, 1989; Leclerc and Plamondon, 1994; Plamondon and Srihari, 2000), and still is a challenging problem. This is mainly due to the large intra-class variations and, when considering forgeries, small inter-class variations.

This paper deals with on-line signature verification. On-line refers here to the use of the time functions of the dynamic signing process (e.g., position trajectories, or pressure versus time), which are obtained using acquisition devices like touch screens or digitizing tablets.

Many different approaches have been considered in the literature in order to extract discriminative information from on-line signature data (Plamondon and Lorette, 1989). The existing methods can broadly be divided into two classes: (i) feature-based approaches, in which a holistic vector representation consisting of a set of global features is derived from the signature trajectories (Lee et al., 1996; Ketabdar et al., 2005), and (ii) function-based approaches, in which time sequences describing local properties of the signature are used for recognition (Nalwa, 1997; Fairhurst, 1997; Jain et al., 2002; Li et al., 2006), e.g., position trajectory, velocity, acceleration, force, or pressure (Lei and Govindaraju, 2005). Although recent works show that feature-based approaches are competitive with respect to function-based methods in some conditions (Fierrez-Aguilar et al., 2005b), the latter methods have traditionally yielded better results.

Function-based approaches can be classified into local and regional methods. In local approaches, the time functions of different signatures are directly matched by using elastic distance measures such as dynamic time warping (Munich and Perona, 2003; Kholmatov and Yanikoglu, 2005; Faundez-Zanuy, 2007). In regional methods, the time functions are converted to a sequence of vectors describing regional properties. One of the most popular regional

* Corresponding author. Tel.: +34 914973363; fax: +34 914972235.
E-mail addresses: julian.fierrez@uam.es (J. Fierrez), javier.ortega@uam.es (J. Ortega-Garcia), daniel.ramos@uam.es (D. Ramos), joaquin.gonzalez@uam.es (J. Gonzalez-Rodriguez).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.07.012

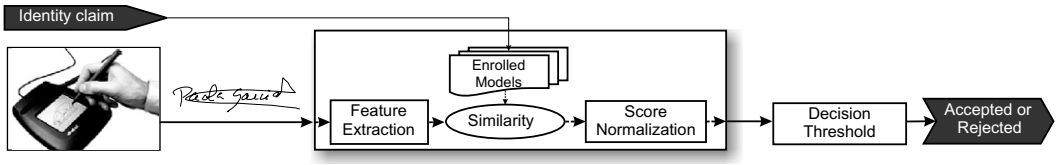

Fig. 1. Architecture of the proposed on-line signature verification system.

approaches is the method based on Hidden Markov Models (Yang et al., 1995; Kashi et al., 1997; Dolfing et al., 1998). In most of these cases, the HMMs modeled stroke-based sequences. In this paper, we study the application of HMMs to time sequences directly based on the dynamic functions, extending the work reported in (Ortega-Garcia et al., 2003a).

The system architecture of the proposed on-line signature verification system is depicted in Fig. 1.

The system described in this paper was submitted by the *Biometric Recognition Group – ATVS* to the First International Signature Verification Competition 2004 with very good results (Yeung et al., 2004), which will be summarized in Section 4.6.

The paper is organized as follows. In Section 2, the feature extraction process is presented. The statistical modeling based on HMMs is described in Section 3. The experimental setup and the database used are presented in Section 4, which are followed by the experimental results regarding feature extraction, HMM configuration, training strategy, and a comparison with other state-of-the-art systems based on SVC 2004 results. Some conclusions are finally drawn in Section 5.

2. Feature extraction

2.1. Basic functions

The signature representation considered in this work is based on the following five time sequences: horizontal x_n and vertical y_n position trajectories, azimuth γ_n and altitude ϕ_n of the pen with respect to the tablet, and pressure signal p_n. The value $n = 1,\ldots,N$ is the discrete time index given by the acquisition device and N is the time duration of the signature in sampling units. Although pen inclination trajectories have shown some discriminative capabilities in other works (Hangai et al., 2000; Sakamoto et al., 2001; Pacut and Czajka, 2001), the use of these two functions in our system degrades the verification performance (as it is shown in Section 4.4.1). As a result, the basic function set consists of x_n, y_n and p_n.

2.2. Geometric normalization

A signature acquisition process on a restricted size frame is assumed (Fierrez-Aguilar et al., 2004). As a result, users are supposed to be consistent in size and writing dynamics. Moreover, a geometric normalization consisting of position normalization followed by rotation alignment is applied.

Position normalization consists in aligning the center of mass

$$[\bar{x},\bar{y}]^{\mathrm{T}} = (1/N)\sum_{n=1}^{N}[x_n,y_n]^{\mathrm{T}} \tag{1}$$

of the different signatures, where $[\cdot]^{\mathrm{T}}$ denotes transpose. Rotation normalization consists in aligning the average path tangent angle

$$\alpha = (1/N)\sum_{n=1}^{N}\arctan(\dot{y}_n/\dot{x}_n) \tag{2}$$

of the different signatures, where the upper dot notation denotes first order time derivative.

2.3. Extended functions

After geometric normalization, four extended sequences are derived from the basic function set. Previous results with other dynamic sequences have shown good levels of performance (Ortega-Garcia et al., 2003a). In the present work, four dynamic sequences have been used as extended functions, namely (Nelson and Kishon, 1991):

- Path-tangent angle: $\theta_n = \arctan(\dot{y}_n/\dot{x}_n)$
- Path velocity magnitude: $v_n = \sqrt{\dot{x}_n^2 + \dot{y}_n^2}$
- Log curvature radius: $\rho_n = \log(1/\kappa_n) = \log(v_n/\dot{\theta}_n)$, where κ_n is the curvature of the position trajectory and $\log(\cdot)$ is applied in order to reduce the range of function values.
- Total acceleration magnitude: $a_n = \sqrt{t_n^2 + c_n^2}$, where $t_n = \dot{v}_n$ and $c_n = v_n \cdot \dot{\theta}_n$ are respectively the tangential and centripetal acceleration components of the pen motion.

In all cases, (discrete) time derivatives have been computed by using the second-order regression described in Section 2.4. The complete instantaneous function-based feature set, including three basic and four extended time sequences is as follows:

$$\boldsymbol{u}_n = [x_n, y_n, p_n, \theta_n, v_n, \rho_n, a_n]^{\mathrm{T}}, \quad n = 1,\ldots,N, \tag{3}$$

where N is the time duration of the considered signature in sampling units.

2.4. Time derivatives

First order time derivatives of complete instantaneous function-based feature sets are highly effective as discrimi-

native parameters regarding verification with other behavioral traits (Soong and Rosenberg, 1988), so we have decided to incorporate time derivatives in our system. Because of the discrete nature of the above-mentioned functions, first order time derivatives are calculated by using a second-order regression (Young et al., 2002), expressed through operator Δ:

$$\dot{f}_n \approx \Delta f_n = \frac{\sum_{\tau=1}^{2} \tau (f_{n+\tau} - f_{n-\tau})}{2 \cdot \sum_{\tau=1}^{2} \tau^2}. \quad (4)$$

In this way, each parameterized signature is formally described as a matrix $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N]$, where $\boldsymbol{v}_n = [\boldsymbol{u}_n^{\mathrm{T}}, (\Delta \boldsymbol{u}_n)^{\mathrm{T}}]^{\mathrm{T}}$, $n = 1, \ldots, N$.

2.5. Signal normalization

A final signal normalization, oriented to obtain zero mean and unit standard deviation function values, is applied:

$$\boldsymbol{o}_n = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{v}_n - \boldsymbol{\mu}), \quad n = 1, \ldots, N, \quad (5)$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are respectively the sample mean and sample diagonal covariance matrix of vectors $\boldsymbol{v}_n$, $n = 1, \ldots, N$.

As a result, each signature is represented by a matrix $\boldsymbol{O} = [\boldsymbol{o}_1, \ldots, \boldsymbol{o}_N]$ comprising 14 normalized discrete time sequences.

3. Signature modeling

3.1. Background on Hidden Markov Models

Hidden Markov Models were introduced in the pattern recognition field as a robust method to model the variability of discrete time random signals where time or context information is available (Rabiner, 1989). Some previous works using HMMs for signature verification include (Yang et al., 1995; Kashi et al., 1997; Dolfing et al., 1998). Basically, the HMM represents a doubly stochastic process governed by an underlying Markov chain with finite number of states and a set of random functions each of which is associated with the output observation of one state (Yang et al., 1995). At discrete instants of time n, the process is in one of the states and generates an observation symbol according to the random function corresponding to that current state. The model is hidden in the sense that the underlying state which generates each symbol cannot be deduced from simple symbol observation.

Formally, a HMM is described as follows:

- H, which is the number of hidden states $\{S_1, S_2, \ldots, S_H\}$. The state at discrete time n will be denoted as q_n.
- The state transition matrix $\boldsymbol{A} = \{a_{ij}\}$ where

$$a_{ij} = P(q_{n+1} = S_j | q_n = S_i), \quad 1 \leqslant i, j \leqslant H. \quad (6)$$

- The observation symbol probability density function in state j, $b_j(\boldsymbol{o})$, $1 \leqslant j \leqslant H$.
- The initial state distribution $\boldsymbol{\pi} = \{\pi_i\}$ where

$$\pi_i = P(q_1 = S_i), \quad 1 \leqslant i \leqslant H. \quad (7)$$

3.2. HMM configuration

In our system, the observation symbol probabilities $b_j(\boldsymbol{o})$ have been modeled as mixtures of M multi-variate Gaussian densities:

$$b_j(\boldsymbol{o}) = \sum_{m=1}^{M} c_{jm} p(\boldsymbol{o} | \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}), \quad 1 \leqslant j \leqslant H, \quad (8)$$

where $p(\boldsymbol{o}|\boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm})$ is a multi-variate Gaussian distribution with mean $\boldsymbol{\mu}_{jm}$ and diagonal covariance matrix $\boldsymbol{\Sigma}_{jm}$, and the coefficients are restricted to $\sum_{m=1}^{M} c_{jm} = 1$, for $1 \leqslant j \leqslant H$. Thus, the observation symbol density functions can be parameterized as $B = \{c_{jm}, \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}\}$, $1 \leqslant j \leqslant H, 1 \leqslant m \leqslant M$.

The set $\lambda = \{\boldsymbol{\pi}, \boldsymbol{A}, B\}$ models the K training signatures of a given subject. The similarity score of an input signature $\boldsymbol{O} = [\boldsymbol{o}_1, \ldots, \boldsymbol{o}_N]$ claiming the identity λ is calculated as $(1/N)\log P(\boldsymbol{O}|\lambda)$ by using the Viterbi algorithm (Rabiner, 1989).

The client model λ is estimated with K training signatures $\{\boldsymbol{O}^{(1)}, \ldots, \boldsymbol{O}^{(K)}\}$, where $\boldsymbol{O}^{(k)} = [\boldsymbol{o}_1^{(k)}, \ldots, \boldsymbol{o}_{N_k}^{(k)}]$, $k = 1, \ldots, K$, by means of the following iterative strategy:

(A) Initialize λ. Each one of the training signatures $\boldsymbol{O}^{(k)}$, $1 \leqslant k \leqslant K$, is divided into H segments $\boldsymbol{S}_1^{(k)}, \ldots, \boldsymbol{S}_H^{(k)}$ where

$$\begin{aligned} \boldsymbol{S}_i^{(k)} &= [\boldsymbol{o}_{(i-1)\lceil N_k/H \rceil + 1}^{(k)}, \boldsymbol{o}_{(i-1)\lceil N_k/H \rceil + 2}^{(k)}, \ldots, \boldsymbol{o}_{(i)\lceil N_k/H \rceil}^{(k)}], \quad 1 \leqslant i \leqslant H-1, \\ \boldsymbol{S}_H^{(k)} &= [\boldsymbol{o}_{(H-1)\lceil N_k/H \rceil + 1}^{(k)}, \boldsymbol{o}_{(H-1)\lceil N_k/H \rceil + 2}^{(k)}, \ldots, \boldsymbol{o}_{N_k}^{(k)}], \end{aligned} \quad (9)$$

and $\lceil \cdot \rceil$ denotes equal or higher nearest integer. Observations from the segments $\boldsymbol{S}_i^{(1)}, \boldsymbol{S}_i^{(2)}, \ldots, \boldsymbol{S}_i^{(K)}$ are clustered into M groups by using the k-means algorithm (Theodoridis and Koutroumbas, 2003a) and the samples from cluster m are used to calculate, according to the maximum likelihood criterion (Theodoridis and Koutroumbas, 2003b), the initial parameters $B = \{c_{im}, \boldsymbol{\mu}_{im}, \boldsymbol{\Sigma}_{im}\}$, $1 \leqslant i \leqslant H$, $1 \leqslant m \leqslant M$. Initial $\boldsymbol{A}$ takes into account a left-to-right topology without skipping state transitions as represented in Fig. 2. Thus, $a_{ij} = 0$ for $i > j$ and $j > i + 1$, $a_{ii} = (O_i - 1)/O_i$ and $a_{i,i+1} = 1/O_i$, where O_i is the number of observations in the segments $\boldsymbol{S}_i^{(1)}, \boldsymbol{S}_i^{(2)}, \ldots, \boldsymbol{S}_i^{(K)}$. The initial state distribution $\boldsymbol{\pi} = \{\pi_1, \pi_2, \ldots, \pi_H\}$ is set up as $\{1, 0, \ldots, 0\}$.

(B) Re-estimate a new model $\bar{\lambda}$ from λ by using the Baum-Welch re-estimation equations (Rabiner, 1989), which guarantee that:

$$\prod_{k=1}^{K} P(\boldsymbol{O}^{(k)} | \bar{\lambda}) \geqslant \prod_{k=1}^{K} P(\boldsymbol{O}^{(k)} | \lambda). \quad (10)$$

(C) Replace λ by $\bar{\lambda}$ and go to step (B) until:

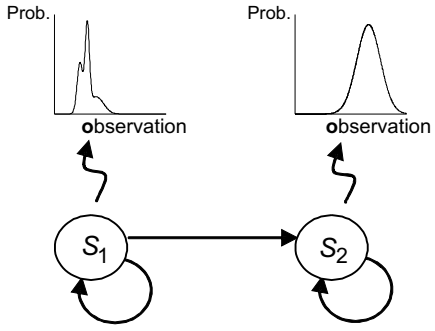

Fig. 2. Topology of the Hidden Markov Model: left-to-right without skipping state transitions.

$$\prod_{k=1}^{K} P(\boldsymbol{O}^{(k)}|\bar{\lambda}) - \prod_{k=1}^{K} P(\boldsymbol{O}^{(k)}|\lambda) \leqslant \Theta. \quad (11)$$

where the threshold Θ is chosen heuristically and the maximum number of iterations is limited to ten.

In the experiments that follows, the training algorithm typically converges after five iterations.

4. Experiments

In the present section, we describe the signature database used and the variability factors considered in its design. This is followed by a discussion of various aspects of performance assessment in biometric verification systems. Finally, we present the experimental results.

4.1. MCYT database

For the experiments reported in this paper we used data from the MCYT bimodal biometric database. In this section we provide a brief description of this corpus. For more details we refer the reader to (Ortega-Garcia et al., 2003b).

The MCYT bimodal biometric database consists of fingerprint and on-line signature modalities. In order to acquire the dynamic signature sequences, a WACOM pen tablet was employed, model INTUOS A6 USB. The pen tablet resolution is 2540 lines per inch (100 lines/mm), and the precision is ±0.25 mm. The maximum detection height is 10 mm (so pen-up movements are also considered), and the capture area is 127 mm (width) × 97 mm (height). This tablet provides the following discrete-time dynamic sequences: (i) position x_n in x-axis, (ii) position y_n in y-axis, (iii) pressure p_n applied by the pen, (iv) azimuth angle γ_n of the pen with respect to the tablet, and (v) altitude angle ϕ_n of the pen with respect to the tablet. The sampling frequency is set to 100 Hz. Taking into account the Nyquist sampling criterion and the fact that the maximum frequencies of the related biomechanical sequences are always under 20–30 Hz (Baron and Plamondon, 1989), this sampling frequency leads to a precise discrete-time signature representation. The capture area is further divided into 37.5 mm (width) × 17.5 mm (height) blocks which are used as frames for acquisition (Fierrez-Aguilar et al., 2004).

Signature corpus comprises genuine and shape-based skilled forgeries with natural dynamics. The forgeries were generated by contributors to the database imitating other contributors. For this task they were given the printed signature to imitate and were asked not only to imitate the shape but also to generate the imitation without artifacts such as time breaks or slowdowns. Fig. 3 shows some example signatures.

The acquisition procedure was as follows. User n wrote a set of five genuine signatures, and then five skilled forgeries of client $n-1$. This procedure was repeated four more times imitating previous users $n-2$, $n-3$, $n-4$ and

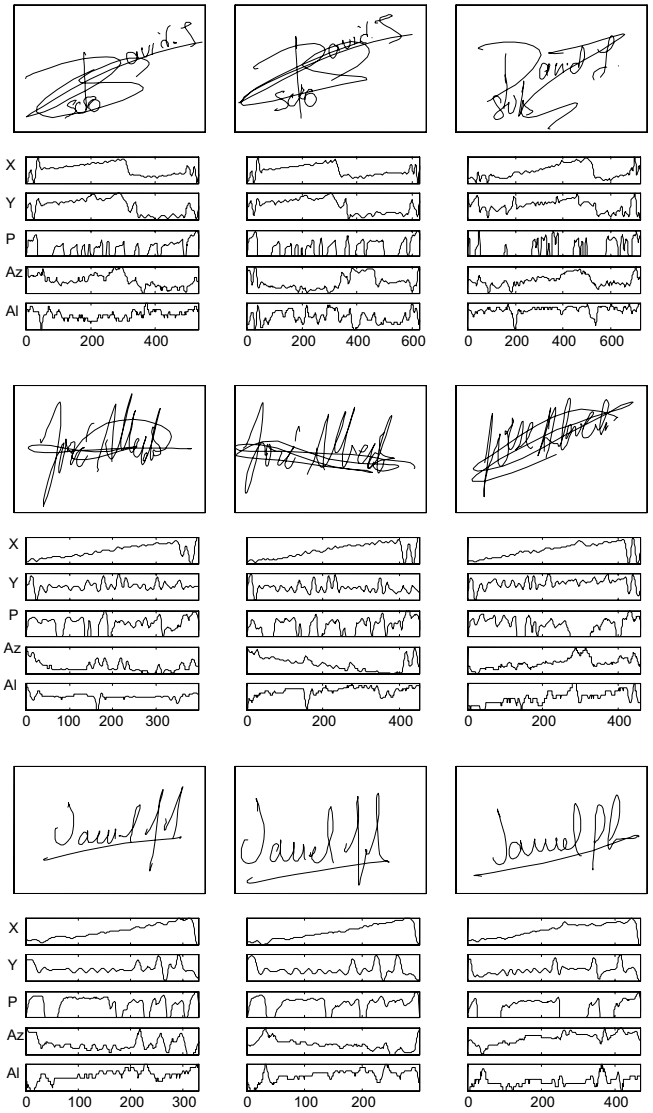

Fig. 3. Signatures from MCYT database corresponding to three different subjects. For each subject, the two left signatures are genuine and the one on the right is a skilled forgery. Plots below each signature correspond to the available information, namely: position trajectories (horizontal x, and vertical y), pressure (p), and pen inclination (*azimuth* and *altitude* angles).

$n-5$. Summarizing, signature data for each client user n include 25 samples of his/her own signature and 25 skilled forgeries produced by users $n+1$, $n+2$, $n+3$, $n+4$, and $n+5$. Taking into account that the signer was concentrated in a different writing task between genuine signature sets, the variability between client signatures from different *acquisition sets* (also called *sessions* throughout the paper) is higher than the variability of signatures within the same acquisition set.

4.2. Score normalization

Typical representations of the verification performance of a given biometric system are the ROC and DET plots (Martin et al., 1997). Those representations consider all possible users of the system with the same scoring scale. This common scale representation is equivalent to the use of a global decision threshold for all clients of the system. User-specific decision thresholds produce results that in average usually outperform the global user-independent decision approach (Jain et al., 2002; Ortega-Garcia et al., 2003a). Some score normalization schemes have been proposed in order to exploit this fact (Auckenthaler et al., 2000), as the following target-dependent score normalization technique based on EERs (TDSN-EER), which consists in normalizing the score $s(\boldsymbol{O},\lambda^{\mathscr{T}})$ obtained from a test signature $\boldsymbol{O}$ claiming the model $\lambda^{\mathscr{T}}$ by means of (Fierrez-Aguilar et al., 2005c):

$$s_n(\boldsymbol{O},\lambda^{\mathscr{T}}) = s(\boldsymbol{O},\lambda^{\mathscr{T}}) - s_{\mathrm{EER}}(\mathscr{G},\mathscr{I}), \tag{12}$$

where s_{EER} is the decision threshold at the empirical EER point for a given set of genuine $\mathscr{G}$ and impostor $\mathscr{I}$ scores corresponding to the subject at hand. In this work, the above described TDSN-EER has been applied *a posteriori* (i.e., $\mathscr{G}$ and $\mathscr{I}$ score sets are the same used for system performance evaluation). This is done to minimize the effects of the score misalignments across users when studying the different modules of the system.

In an operating system, the statistics necessary to estimate the parameters of the normalization functions have to be estimated from the available training data, using *a priori* score normalization schemes (Fierrez-Aguilar et al., 2005c). Interestingly, the application of these *a priori* schemes in SVC 2004 proved to be quite useful in spite of the training data scarcity (only five training signatures) and the challenging scenario: invented signatures written without visual feedback, time span between sessions of at least one week and only between-sessions comparisons, etc. (Yeung et al., 2004). In particular, the application of a target-centric score normalization approach provided a 18% performance improvement to our HMM system for skilled forgeries while maintaining the performance for random forgeries. This can be observed in the competition results, summarized in Section 4.6. The systems 19a and 19b submitted by the authors were identical except for an *a priori* score normalization step (a = without score normalization, b = with score normalization). The score normalization step in this case was the *a priori* score normalization technique rotTC-3 described in (Fierrez-Aguilar et al., 2005c).

4.3. Experimental protocol

For the experiments presented here, we have used two disjoint subsets of the MCYT database, the first one for development (50 subjects) and the second one for evaluation (145 subjects). All the signatures of each subject are used in both cases (25 genuine signatures and 25 skilled forgeries).

The client signatures not used for training are used for testing. The skilled forgeries are used only for testing. In case of testing against random forgeries and for a particular subject, all client signatures from the remaining subjects are used during development (i.e., $50\times49\times25$ impostor trials), but only two signatures from each of the remaining subjects are used during evaluation (i.e., $145\times144\times2$ impostor trials).

Experiments have been carried out according to the following procedure. We first study the effects of function set selection, modeling and training strategy on the verification performance using the development set. Results are given as DET plots applying the *a posteriori* target-dependent score normalization technique described in Section 4.2. From these development experiments, and considering the EER as the cost function to minimize, an enhanced system configuration is obtained. Finally, verification performance of the enhanced system is provided on the evaluation set. Separate results are reported for random and skilled forgeries.

4.4. Development experiments

4.4.1. Feature extraction

The initial configuration for the experiments is as follows (Ortega-Garcia et al., 2003a): four HMM states, eight mixtures per state and five training signatures from different sessions. Results for different functions are shown in Fig. 4.

In Fig. 4a we compare some basic function sets. The EER decreases from 10.37% to 4.54% when pressure signal is included to the basic position trajectory information but increases to 4.84% and 6.28% when altitude and azimuth signals are respectively considered.

In Fig. 4b we show the progressive improvement in verification performance when extended functions are included one by one to the basic set $\{x,y,p\}$ (4.54% EER). When path tangent angle θ, path velocity magnitude v, log curvature radius ρ, and total acceleration magnitude a are progressively included, we obtain the following EERs: 2.57%, 1.99%, 1.44%, and 0.68%. The set composed by these 7 functions $\{x,y,p,\theta,v,\rho,a\}$ will be referred to as w.

4.4.2. Training strategy

The initial configuration for the experiments is as follows: $w+\Delta w$ functions, four HMM states and eight

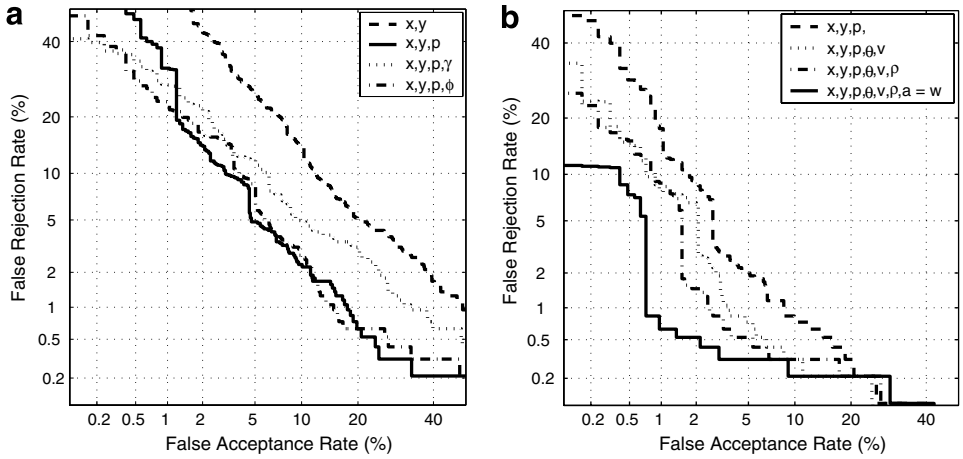

Fig. 4. Feature extraction experiments. Verification performance results for skilled forgeries are given for various function sets including: position trajectories x and y, pressure p, azimuth γ, altitude ϕ, path tangent angle θ, path velocity magnitude v, log curvature radius ρ, and total acceleration magnitude a.

mixtures per state. Results for different training strategies, considering only skilled forgeries, are shown in Fig. 5.

In Fig. 5a we evaluate the performance improvement of our system for increasing number of training signatures when training and testing signatures come from different sessions. In order to do so, training signatures are selected from the minimum number of sessions possible (i.e., 1 set for 1–5 training signatures, 2 sessions for 6–10 signatures, and so on). As a result, the EER does not improve significantly for more than five training samples.

In Fig. 5b we evaluate the performance improvement in our system when the number of training signatures is increased and the training and testing signatures come from the same sessions. For that purpose, we have selected training signatures from the maximum possible number of sessions (i.e., 1 session for 1 training signature, 2 sessions for 2 training signatures, 3 sessions for 3 training signatures, 4 sessions for 4 training signatures and 5 sessions for 5–20 training signatures). As a result, the EER improves significantly with the first 5 training signatures (0.85% EER) and keeps improving for a higher number of signatures (0.05% EER for 20 training samples).

Finally, we test the verification performance for a fixed number of training signatures from an increasing number of sessions. In this case we provide a DET plot (see Fig. 6) where, testing with two sessions, the fixed five training samples have been selected from the other one (0.84% EER), two (0.82% EER) and three (0.48% EER) sessions, respectively. As can be observed, verification performance improves when training data come from different sessions.

4.4.3. Signature modeling

The initial configuration for the experiment is $w + \Delta w$ functions and five training signatures from the first acquisition set. In Table 1 we show the results for different HMM parameters including the degenerated case of the single state HMM, which is equivalent to modeling based on Gaussian Mixture Models (Richiardi and Drygajlo, 2003). From this table we observe that for a given complexity of the model $\approx H \times M$ (Richiardi and Drygajlo, 2003),

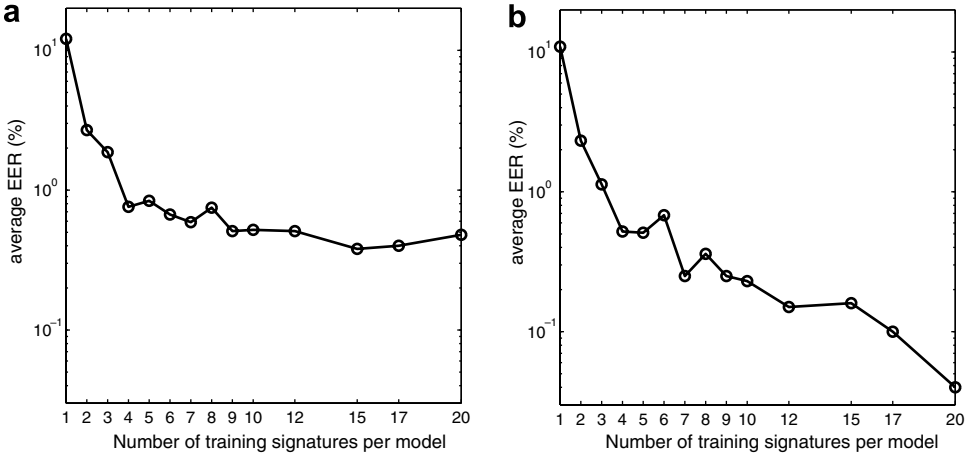

Fig. 5. Training strategy experiments. Verification performance results are given for skilled forgeries with increasing number of training signatures: (a) training and testing signatures from different sessions, (b) training and testing signatures from the same sessions.

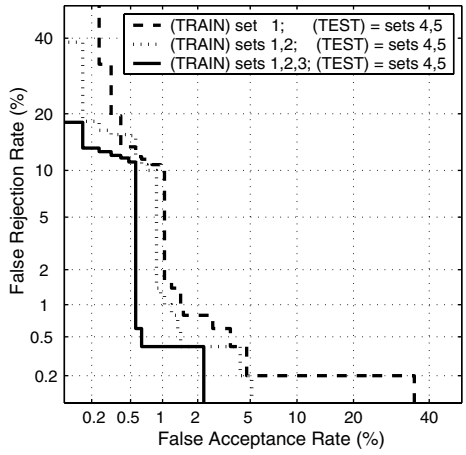

Fig. 6. Training strategy experiments. Verification performance results for skilled forgeries for a fixed number of training signatures from an increasing number of sessions.

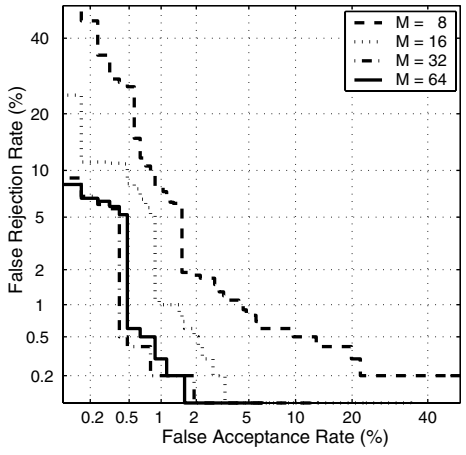

Fig. 7. Signal modeling experiments. Verification performance results are given for an increasing number of Gaussian mixtures per state M, being the number of states fixed $H = 2$ (skilled forgeries).

Table 1
EER (in %) for different HMM configurations (skilled forgeries)

H	$M=1$	$M=2$	$M=4$	$M=8$	$M=16$	$M=32$	$M=64$	$M=128$
1					1.74	1.05	0.70	0.71
2				1.51	0.74	**0.30**	0.44	
4			1.64	0.87	0.52	0.48		
8		1.81	0.79	0.76	0.35			
16		1.20	0.96	0.74				
32	1.28	0.97						

H = number of states; M = number of Gaussian mixtures per state.

i.e., a diagonal in Table 1, the configuration $H = 2$ is usually the best.

The best HMM parameters for our system are $H = 2$ and $M = 32$. This result is in accordance with the recent trend of reducing the number of states in HMM-based on-line signature verification systems (Richiardi and Drygajlo, 2003). This can be explained because the majority of signatures in the MCYT database consist of written name and flourish (Fierrez-Aguilar et al., 2004). Both parts usually have different dynamic statistics that the 2-state modeling approach may be exploiting. Another explanation of the degraded performance when using a large number of states is the data scarcity in on-line signatures as compared to speech recognition, where HMM is the dominant approach.

Finally, we represent in Fig. 7 DET plots for $H = 2$ with different values of M, where the verification performance improvement for increasing values of M can be seen, until the optimum $M = 32$ is reached.

4.5. Evaluation experiments

Summarizing the results of the experiments reported above, we conclude that the best system configuration is the following: (i) feature extraction based on seven discrete-time functions (x trajectory, y trajectory, pressure, path tangent angle, path velocity magnitude, log curvature radius and total acceleration magnitude) and their first order time derivatives, and (ii) HMM recognition with 2 states and 32 Gaussian mixtures per state.

We finally show error rates on the evaluation set using as training data 10 signatures per model from different sessions both for skilled, Fig. 8a, and random forgeries, Fig. 8b. Verification performance is shown for three *a posteriori* target-dependent score normalization techniques, namely: no score normalization, z-norm (Auckenthaler et al., 2000), and target-dependent score normalization based on individual EERs (EER-TDSN) as described in Section 4.2. The EERs are reported in Table 2. As a result, we observe (EER for skilled forgeries is specified):

- Score normalization using the z-norm technique degrades the verification performance (4.44%) with respect to no score normalization (3.36%).
- Verification performance is highly improved when using the target-dependent score normalization technique (0.78%), which is equivalent to the use of user-dependent decision thresholds.

4.6. Comparison with the state-of-the-art

The First International Signature Verification Competition (SVC) was organized in 2004, and provided a public benchmark for signature verification (Yeung et al., 2004). The development corpus of the extended task (pen coordinate, orientation and pressure information are available) is available through the competition website.[1] This corpus consists of 40 sets of signatures. The evaluation set is comprised of 60 additional sets of signatures. Each set contains 20 genuine signatures from one contributor (acquired in two separate sessions) and 20 skilled forgeries from five other contributors. The SVC database is especially

[1] http://www.cs.ust.hk/svc2004/.

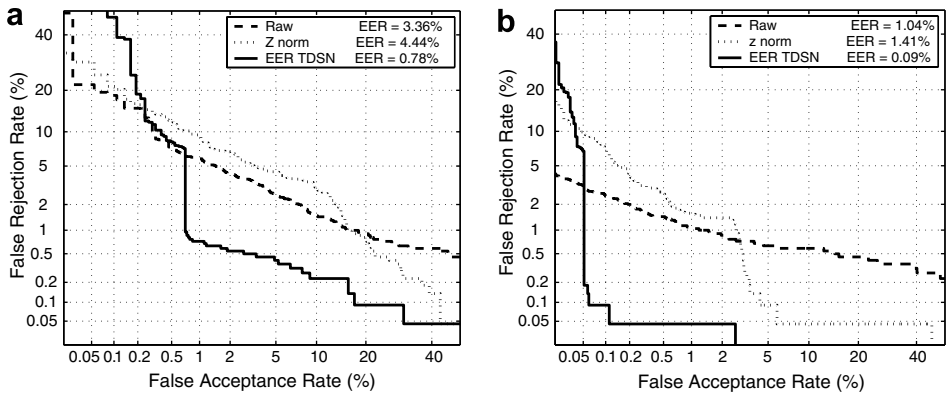

Fig. 8. Verification performance results on the evaluation set for skilled (a) and random forgeries (b).

Table 2
System performance on the evaluation set (EER in %)

Type of forgery	Raw scores	*z*-norm	EER-TDSN
Skilled	3.36	4.44	0.78
Random	1.04	1.41	0.09

challenging due to several factors, including: (i) no visual feedback when writing (acquisition was conducted by using a WACOM tablet with a Grip Pen), (ii) subjects used invented signatures different to the ones used in daily life in order to protect their personal data, (iii) skilled forgers imitated not only the shape but also the dynamics, and (iv) time span between training and testing signatures was at least one week. The signatures are in either English or Chinese (see Fig. 9).

The authors participated to SVC 2004 with the system described in the present paper (team ID 19). A summary of the evaluation results for the extended task is given in Tables 3 and 4, which correspond to the development and the evaluation sets, respectively (Yeung et al., 2004). Training data consisted of five signatures randomly selected from the first acquisition session, and testing data consisted of the 10 genuine signatures of the second session, together with 20 skilled or random forgeries. As indicated in Section 4.2, the different systems submitted by the authors (a, b, and c) were identical except for an *a priori* score normalization step: a = without score normalization, b and c = different types of user-dependent score normalization (Fierrez-Aguilar et al., 2005c).

In the evaluation set, the proposed HMM system was ranked second for skilled forgeries and first for random forgeries. It was only outperformed by the winner of the competition, which was based on a dynamic time warping approach (Kholmatov and Yanikoglu, 2005). Interestingly,

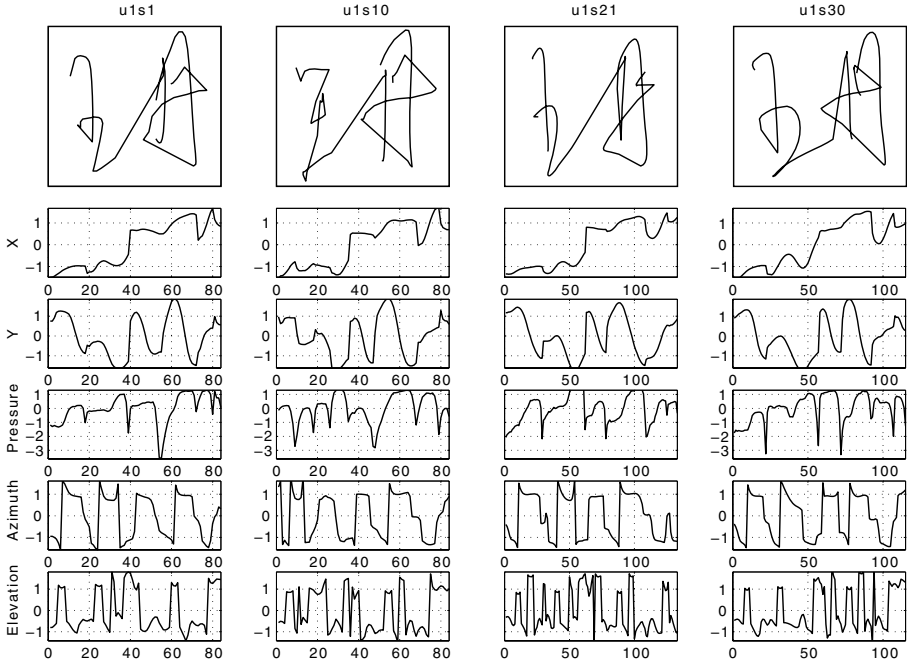

Fig. 9. Signature examples from SVC 2004 corpus. For a particular subject, two genuine signatures (left columns) and two skilled forgeries (right columns) are given. Plots of the coordinate trajectories, pressure signal and pen orientation functions are also given.

Table 3
EER statistics in % (Yeung et al., 2004) for the extended task of SVC 2004 (development set, 40 subjects)

Team ID	Skilled forgeries			Random forgeries		
	Average	SD	Maximum	Average	SD	Maximum
19b	6.90	9.45	50.00	3.02	3.65	15.00
19c	6.91	9.42	50.00	3.02	3.65	15.00
6	6.96	11.76	65.00	3.47	4.23	20.00
29	7.64	12.62	60.00	4.45	6.84	50.00
19a	8.90	11.72	71.00	3.08	3.74	15.00
14	11.29	13.47	70.00	4.41	5.35	28.57
18	15.36	13.88	60.00	6.39	7.03	45.00
17	19.00	14.43	70.00	4.29	4.69	30.00
3	20.01	18.44	76.19	5.07	8.13	44.44
4	21.89	17.05	73.33	8.75	9.71	48.72

SD denotes standard deviation.

Table 4
EER statistics in % (Yeung et al., 2004) for the extended task of SVC 2004 (evaluation set, 60 subjects)

Team ID	Skilled forgeries			Random forgeries		
	Average	SD	Maximum	Average	SD	Maximum
6	2.89	5.69	30.00	2.55	5.66	50.00
19b	5.01	9.06	50.00	1.77	2.92	10.00
19c	5.13	8.98	51.00	1.79	2.93	10.00
19a	5.91	9.42	50.00	1.70	2.86	10.00
14	8.02	10.87	54.05	5.19	8.57	52.63
18	11.54	12.12	50.00	4.89	6.65	45.00
17	12.51	13.01	70.00	3.47	5.53	30.00
4	16.34	14.00	61.90	6.17	9.24	50.00

SD denotes standard deviation.

it has been recently shown that the HMM approach outperforms an implementation of the DTW approach used by the winner when enough training signatures are available (Fierrez-Aguilar et al., 2005a). More comparative experiments with other state-of-the-art systems can be found in (Garcia-Salicetti et al., 2006).

5. Conclusions

An on-line signature verification system based on dynamic time functions and Hidden Markov Models has been presented.

In the experiments, we have explored various aspects of feature extraction and modeling. With respect to HMM modeling we have shown that the best configuration is 2 HMM states and 32 Gaussian mixtures per state. Experiments on the training strategy have shown that training with multi-session data remarkably improves the verification performance and five training signatures are enough for obtaining robust models.

Our verification performance results are 0.74% and 0.05% EER for skilled and random forgeries, respectively (with *a posteriori* user-dependent decision thresholds). These results were achieved on a database of 145 subjects comprising 3625 client signatures, 3625 skilled forgeries and 41,760 random impostor attempts. The system described in this work was submitted by the *Biometric Recognition Group – ATVS* to SVC 2004 (Yeung et al., 2004), from which selected results in comparison with other systems have been also reported.

Acknowledgements

This work has been supported by the Spanish projects TIC2003-08382-C05-01 and TEC2006-13141-C03-03, and by the European NoE Biosecure. The postdoctoral research of J.F. is supported by a Marie Curie Outgoing International Fellowship. Authors wish to thank Jorge Martin-Rello for his valuable development work.

References

Auckenthaler, R., Carey, M., Lloyd-Tomas, H., 2000. Score normalization for text-independent speaker verification systems. Digit. Signal Process. 10, 42–54.

Baron, R., Plamondon, R., 1989. Acceleration measurement with an instrumented pen for signature verification and handwriting analysis. IEEE Trans. Instrum. Meas. 38 (6), 1132–1138.

Dolfing, J.G.A., Aarts, E.H.L., van Oosterhout, J.J.G.M., 1998. On-line signature verification with Hidden Markov Models. Proc. Internat. Conf. on Pattern Recognition, ICPR. IEEE CS Press, pp. 1309–1312.

Fairhurst, M.C., 1997. Signature verification revisited: Promoting practical exploitation of biometric technology. IEE Electron. Comm. Eng. J. 9 (6), 273–280.

Faundez-Zanuy, M., 2007. On-line signature recognition based on VQ-DTW. Pattern Recognition 40, 981–992.

Fierrez-Aguilar, J., Alonso-Hermira, N., Moreno-Marquez, G., Ortega-Garcia, J., 2004. An off-line signature verification system based on fusion of local and global information. Post-ECCV Workshop on Biometric Authentication BIOAW. In: LNCS, vol. 3087. Springer, pp. 295–306.

Fierrez-Aguilar, J., Krawczyk, S., Ortega-Garcia, J., Jain, A.K., 2005a. Fusion of local and regional approaches for on-line signature verification. Proc. Internat. Workshop on Biometric Recognition Systems, IWBRS. In: LNCS, vol. 3781. Springer, pp. 188–196.

Fierrez-Aguilar, J., Nanni, L., Lopez-Penalba, J., Ortega-Garcia, J., Maltoni, D., 2005b. An on-line signature verification system based on fusion of local and global information. Proc. IAPR Internat. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA. In: LNCS, vol. 3546. Springer, pp. 523–532.

Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., 2005c. Target dependent score normalization techniques and their application to signature verification. IEEE Trans. Systems Man Cybernet. C 35 (3), 418–425.

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., Castiglione, M.B., Fairhurst, M., 2006. Biosecure reference systems for on-line signature verification: A study of complementarity. Ann. Telecomm. 62 (1–2) (Special Issue Multimodal Biometrics).

Hangai, S., Yamanaka, S., Hanamoto, T., 2000. On-line signature verification based on altitude and direction of pen movement. Proc. IEEE Internat. Conf. on Multimedia and Expo, ICME 1, 489–492.

Jain, A., Griess, F., Connell, S., 2002. On-line signature verification. Pattern Recognition 35 (12), 2963–2972.

Kashi, R.S., Hu, J., Nelson, W.L., Turin, W., 1997. On-line handwritten signature verification using Hidden Markov Model features. In: Proc. 4th Internat. Conf. on Document Analysis and Recognition, ICDAR, vol. 1. IEEE CS Press, pp. 253–257.

Ketabdar, H., Richiardi, J., Drygajlo, A., 2005. Global feature selection for on-line signature verification. In: Proc. 12th Internat. Graphonomics Society Conf.

Kholmatov, A., Yanikoglu, B., 2005. Identity authentication using improved online signature verification method. Pattern Recognition Lett. 26 (15), 2400–2408.

Leclerc, F., Plamondon, R., 1994. Automatic signature verification: The state of the art, 1989–1993. Internat. J. Pattern Recognition Artificial Intell. 8 (3), 643–660.

Lee, L.L., Berger, T., Aviczer, E., 1996. Reliable on-line human signature verification systems. IEEE Trans. Pattern Anal. Machine Intell. 18 (6), 643–647.

Lei, H., Govindaraju, V., 2005. A comparative study on the consistency of features in on-line signature verification. Pattern Recognition Lett. 26 (15), 2483–2489.

Li, B., Zhang, D., Wang, K., 2006. On-line signature verification based on NCA (null component analysis) and PCA (principal component analysis). Pattern Anal. Appl. 8, 345–356.

Martin, A., Doddington, G., Kamm, T., Ordowski, M., Przybocki, M., 1997. The DET curve in assessment of decision task performance. In: Proc. ESCA Eur. Conf. on Speech Comm. and Tech., EuroSpeech, pp. 1895–1898.

Munich, M.E., Perona, P., 2003. Visual identification by signature tracking. IEEE Trans. Pattern Anal. Machine Intell. 25 (2), 200–217.

Nalwa, V.S., 1997. Automatic on-line signature verification. Proc. IEEE 85 (2), 215–239.

Nelson, W., Kishon, E., 1991. Use of dynamic features for signature verification. Proc. IEEE Internat. Conf. Systems Man Cybernet. 1, 201–205.

Ortega-Garcia, J., Fierrez-Aguilar, J., Martin-Rello, J., Gonzalez-Rodriguez, J., 2003a. Complete signal modeling and score normalization for function-based dynamic signature verification. Proc. IAPR Internat. Conf. on Audio- and Video-based Person Authentication, AVBPA. In: LNCS, vol. 2688. Springer, pp. 658–667.

Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., et al., 2003b. MCYT baseline corpus: A bimodal biometric database. IEE Proc. Vision Image Signal Process. 150 (6), 395–401.

Pacut, A., Czajka, A., 2001. Recognition of human signatures. Proc. IEEE Joint Internat. Conf. Neural Networks, IJCNN 2, 1560–1564.

Plamondon, R., Lorette, G., 1989. Automatic signature verification and writer identification – the state of the art. Pattern Recognition 22 (2), 107–131.

Plamondon, R., Srihari, S.N., 2000. On-line and off-line handwriting recognition: A comprehensive survey. IEEE Trans. Pattern Anal. Machine Intell. 22 (1), 63–84.

Rabiner, L.R., 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proc. IEEE 77 (2), 257–286.

Richiardi, J., Drygajlo, A., 2003. Gaussian Mixture Models for on-line signature verification. In: Proc. of ACM SIGMM Workshop on Biometric Methods and Applications, WBMA, pp. 115–122.

Sakamoto, D., Morita, H., Ohishi, T., Komiya, Y., Matsumoto, T., 2001. On-line signature verification algorithm incorporating pen position, pen pressure and pen inclination trajectories. Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing, ICASSP 2, 993–996.

Soong, F.K., Rosenberg, A.E., 1988. On the use of instantaneous and transitional spectral information in speaker recognition. IEEE Trans. Acoust. Speech Signal Process. 36 (6), 871–879.

Theodoridis, S., Koutroumbas, K., 2003a. Pattern Recognition. Academic Press, pp. 531–533.

Theodoridis, S., Koutroumbas, K., 2003b. Pattern Recognition. Academic Press, pp. 28–31.

Yang, L., Widjaja, B.K., Prasad, R., 1995. Application of Hidden Markov Models for signature verification. Pattern Recognition 28 (2), 161–170.

Yeung, D.-Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G., 2004. SVC 2004: First International Signature Verification Competition. Proc. Internat. Conf. on Biometric Authentication, ICBA. In: LNCS, vol. 3072. Springer, pp. 16–22.

Young, S., et al., 2002. The HTK Book Version 3.2.1. Cambridge University Engineering Department, p. 65. <http://htk.eng.cam.ac.uk/>.