Legendre polynomials based feature extraction for online signature verification. Consistency analysis of feature combinations

Marianela Parodi *, Juan C. Gómez

Laboratory for System Dynamics and Signal Processing, FCEIA, Universidad Nacional de Rosario, CIFASIS, CONICET, Argentina

ARTICLE INFO

ABSTRACT

In this paper, feature combinations associated with the most commonly used time functions related to the signing process are analyzed, in order to provide some insight on their actual discriminative power for online signature verification. A consistency factor is defined to quantify the discriminative power of these different feature combinations. A fixed-length representation of the time functions associated with the signatures, based on Legendre polynomials series expansions, is proposed. The expansion coefficients in these series are used as features to model the signatures. Two different signature styles, namely, Western and Chinese, from a publicly available Signature Database are considered to evaluate the performance of the verification system. Two state-of-the-art classifiers, namely, Support Vector Machines and Random Forests are used in the verification experiments. Error rates comparable to the ones reported over the same signature datasets in a recent Signature Verification Competition, show the potential of the proposed approach. The experimental results, also show that there is a good correlation between the consistency factor and the verification errors, suggesting that consistency values could be used to select the optimal feature combination.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Automatic signature verification has long been considered an important research area in the field of biometrics [1–4]. Signature verification is the most popular method for identity verification. Signatures are recognized as a legal means of verifying an individual's identity by financial and administrative institutions. In addition, it is a non-invasive biometric technique, and people are familiar with the use of signatures for identity verification in their everyday life.

Two categories of signature verification systems can be distinguished taking into account the acquisition device, namely, offline and online systems. For offline verification systems, only the image of the signature is available, while for online systems, dynamic information acquired during the signing process, such as x and y pen coordinates and pen pressure, is available. The interest in the online approach for signature verification has increased in recent years due to the widespread use of electronic pen-input devices, such as digitizer tablets and PDAs. In addition, it would be reasonable to expect that the incorporation of dynamic information acquired during the signing process would make signatures more difficult to forge and, in this way, the online verification systems more reliable than the offline ones.

In online systems, the signature is parameterized by several discrete time functions, *e.g.*, pen coordinates, pen pressure and, when available, pen inclination angles. Researchers have long argued about the effectiveness of these different time functions for verification purposes. During the First International Signature Verification Competition (SVC2004), the results using only pen coordinates outperformed those adding pen pressure and pen inclination angles [5]. Since then, several works have been presented concerning the best set of features to model the signatures. In [6], the authors state that using only pen coordinates leads to better results than incorporating the pen pressure. The time variability between training and testing data acquisition sessions is considered in [7], where it is concluded that pen pressure is the most unreliable feature, pen inclination angles are too unstable, and pen coordinates are the most robust time functions in the presence of a long term time variability. On the other hand, some works show improvements when combining pen coordinates with pen pressure and inclination angles [8]. The conflicting results observed in the literature make the discussion still open. In a preliminary work by the present authors [9], some feature combinations based on the pen coordinates and the pen pressure, are studied. The conclusions in [9] are in line with the idea that combining pen coordinates with the pen pressure leads to a verification performance improvement.

* Correspondence to: CIFASIS, 27 de Febrero 210 bis, (S2000EZP) Rosario, Argentina. Tel.: +54 341 4237248x335; fax: +54 341 4821772x3.
E-mail addresses: parodi@cifasis-conicet.gov.ar (M. Parodi), gomez@cifasis-conicet.gov.ar (J.C. Gómez).

A desirable property for any feature is to have high consistency in the sense that the feature values of the genuine signatures should be close to each other while the ones of genuine and forged signatures should be not. A well defined consistency model would allow to quantify the discriminative power of the features and to predict their effectiveness for verification purposes. A consistency model was first introduced in [10,11]. In [10], the consistency model is used to select an optimal subset of global features from a larger global feature set. In [11], several local and global features are compared on the basis of their consistency, resulting pen coordinates and some derived features the most reliable ones. The lack of a widely used consistency model in the literature, makes its study an interesting issue. In [9], a new consistency factor is introduced. The proposed feature combinations are compared based on their consistency factor values, being the feature combinations containing the pen pressure the most reliable ones.

An important factor that deserves more investigation is the influence of the cultural origin of the signatures in the performance of the verification systems. To the best of the authors' knowledge, there are not many works in the literature that consider non-Western signatures such as Chinese, Japanese, Arabic, etc. In [12], an updated survey of non-English and non-Latin signature verification systems can be found. Non-Western signatures do have different shapes and the writing style is different to the Western one. For instance, the Chinese handwriting style consists of one or more multi-trace characters, most of them being phono-semantic compounds, composed by two parts: the radical, which is often a simplified pictograph and suggests the character's general meaning and a phonetic indicator. Originally, Chinese pictographs conveyed their meaning through pictorial resemblance to a physical object. Although in modern Chinese this resemblance is no longer clear, Chinese characters are still pictorial symbols. Among the literature of non-Western signature verification, more attention has been given to Chinese signatures than to Japanese, Arabic, Persian or Indian signatures. Offline [13,14], as well as online [15] verification systems have been presented in the literature for Chinese signature verification. Further, the Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011) held within ICDAR 2011 [16], introduced a new Database containing Chinese signatures, encouraging the researches to work on this type of data. On the other hand, Japanese and Arabic signatures, among others, have not been investigated so much. Japanese signatures consist of different component characters spaced from each other. There is not much work done on this type of data [17,18], and it is mostly focused in offline data. Arabic script is written from right to left in a cursive style. Although a lot of research has been carried out on Arabic handwriting recognition, not much work has been carried out on Arabic signature verification. In [19], an offline verification system for Arabic signatures is presented. For a verification system to have a widespread acceptance it should take into account these different writing styles. As pointed out in [12], there are still many challenges in this research area.

In this paper, the coefficients in the Legendre series approximations of the time functions associated with the signatures are used as features to model them. The time functions considered in this paper are the pen coordinates, pressure, velocities and acceleration, as well as the log curvature radius, which are the most commonly used functions in the literature for online signature verification [20,21]. A consistency factor is proposed to quantify the discriminative power of different combinations of the time functions related to the signing process. Two different signature styles are considered, namely, Western and Chinese, of a publicly available Signature Database. Two state-of-the-art classifiers, namely, Support Vector Machines (SVMs) and Random Forests (RFs), are used to perform the verification experiments.

This approach of representing the time functions using Legendre polynomials was first introduced by the present authors in the conference paper [9]. Only few feature combinations were studied there, and a qualitative study of the consistency of the feature combinations was performed. In the present paper, more time functions are considered and a thorough analysis of all the possible feature combinations is carried out. In addition, a quantitative study correlating the consistency factor with the verification errors is performed.

The main contributions of this paper are the following:

- A feature extraction approach based on Legendre series representation of the time functions associated with the signatures is proposed. To the best of the authors' knowledge this is the first time that this approach is used in the context of signature verification.
- A consistency factor is proposed to quantify the discriminative power of different combinations of the time functions associated with the signing process. A thorough study of all the possible feature combinations is carried out, and the pros and cons of these different combinations are analyzed.
- A quantitative study of the relationship between the proposed consistency factor and the verification performance of a feature combination is performed based on correlation analysis.
- The experiments are performed on one of the most recent signature datasets, containing Western and Chinese signatures, which have been used in one of the latest signature verification competitions. To quantify the verification performance, the EER (Equal Error Rate) and the cost of the log-likelihood ratios $\hat{C}_{llr}$ are reported.

The paper is organized as follows. The feature extraction approach is described in Section 2. In Section 3 the proposed consistency model is introduced. In Section 4 the Database is described. Section 5 is devoted to the description of the experiments, in particular, Sections 5.1 and 5.2 focus on the consistency computation and the verification experiments, respectively. In Section 6 the results are presented and discussed. Finally, some concluding remarks are given in Section 7.

2. Feature extraction

Several methods have been proposed in the literature for online signature verification. They differ basically in the way they perform the feature extraction and in the classification approach they employ. The different features can be classified into local features, calculated for each point in the time sequence, and global features, calculated from the whole signature. Many researchers accept that approaches based on local features achieve better performance than the ones based on global features, but still there are others who favor the use of global features [21,22]. When using global features, feature vectors have a fixed amount of components regardless the signature length. This represents an advantage since it makes the comparison between two signatures easier with respect to the case of having different feature vector lengths. Moreover, a fixed-length model of the signatures can be required for certain biometric applications [23,24]. In [25], a fixed-length representation of the signatures is proposed based on the Fast Fourier Transform (FFT). In this paper, the approximation of the different time functions by orthogonal polynomials, introduced by the authors in [9], is employed to obtain a fixed-length representation of the signature.

2.1. Basic functions

Typically, the measured data consists of three discrete time functions: pen coordinates x and y, and pen pressure p. Depending

on the acquisition device, the pen altitude and azimuth angles could also be available. In addition to the raw data, some other dynamic functions, such as, x and y velocities and accelerations and log curvature radius can also be computed from them.

2.2. Normalization

Depending on the given space to sign, signatures can be written in different sizes, signers can place the signatures anywhere they want in the sheet of paper and many times they would sign in a rotated angle with respect to the one they usually sign. This makes size, translation and rotation normalization fundamental preprocessing tasks.

- Size normalization:
  A width normalization is performed on the x and y pen coordinates of the signature. The width of the signature is previously fixed while the height is left to take the corresponding value in order to keep the original height-to-width ratio. Then, the x and y pen coordinates are modified as

$$x_{sn}(n) = \frac{x^o(n) - x_{min}}{x_{max} - x_{min}} \cdot W_{new}, \qquad (1)$$

$$y_{sn}(n) = \frac{y^o(n) - y_{min}}{y_{max} - y_{min}} \cdot H_{new}, \qquad (2)$$

  where $(x^o(n), y^o(n))$ are the original point coordinates and $(x_{sn}(n), y_{sn}(n))$ are the corresponding ones after size normalization, W_{new} is the new fixed width and H_{new} is the resulting new height computed as

$$H_{new} = H^o \cdot \frac{W_{new}}{W^o}, \qquad (3)$$

  being W^o and H^o the original width and height, respectively.
- Translation normalization:
  In order to normalize the position of the signatures, they are centered by subtracting the corresponding mean values from the original x and y pen coordinates, that is

$$x(n) = x_{sn}(n) - x_{mean}, \qquad (4)$$

$$y(n) = y_{sn}(n) - y_{mean}, \qquad (5)$$

  where $(x_{sn}(n), y_{sn}(n))$ are the size normalized point coordinates and $(x(n), y(n))$ are the corresponding ones after translation normalization.
- Rotation normalization:
  There are conflicting views in the literature regarding rotation normalization. Some authors perform a correction in the main direction of the signature, rotating it until it has the direction of a predetermined baseline [20]. However, it has been argued that the main direction of the signature is a distinctive feature and so compensating it would result in loss of useful discriminative information. Since there exist a significant variability in the main direction of the signature for a given author, rotation compensation would make the system less robust. In line with these ideas, no rotation compensation is performed in this paper.

Another widely used preprocessing technique is resampling. Due to the acquisition process, the measured data may contain noise or gaps introduced during the recordings. Resampling is used to correct these acquisition artifacts and, in addition, to get a fixed-length resampled time function. Several works in the literature use resampling to remove redundant points from the measured signals [26,27]. In [27], the effect of different resampling techniques on the verification performance is studied. The authors state that resampling has several advantages such as reducing the storage requirements and increasing the simplicity without compromising and even improving the system performance. On the other hand, many other works in the literature do not use resampling as a preprocessing step [6,11,20,25]. Moreover, in [6,25] the authors observed that using resampling leads to worse verification performances, since it implies a significant loss of information. They conclude that it is convenient not to use resampling and that the disadvantage of not having a fixed-length signal, is not that important. In the present paper, the proposed feature extraction technique delivers a fixed-length feature vector, so that no resampling of the original time functions is required.

2.3. Extended functions

Several extended functions that can be computed from the acquired functions have been used in the literature. In [6], the incremental variations of the x and y pen coordinates are proposed. In [21], several time functions, such as, the x and y velocities and accelerations and the log curvature radius, among others, are used as well as their first and second derivatives. In this paper, the x and y velocities (v_x and v_y, respectively), the total acceleration a_T and the log curvature radius ρ [20] are computed from the basic function set of x and y pen coordinates and pen pressure p. Let $n = 1, \ldots, L_{sign}$ be the discrete time index of the measured functions and L_{sign} the time duration of the signature in sampling units, then the above mentioned extended functions are computed as

- x velocity: $v_x(n) = \dot{x}(n)$
- y velocity: $v_y(n) = \dot{y}(n)$
- Total acceleration: $a_T(n) = \sqrt{t^2(n) + c^2(n)}$, where $t(n) = \dot{v}_T(n)$, being $v_T(n) = \sqrt{v_x^2(n) + v_y^2(n)}$ and $c(n) = v_T(n) \cdot \dot{\theta}(n)$, being $\theta(n) = arctan(\dot{y}(n)/\dot{x}(n))$.
- Log curvature radius: $\rho(n) = \log(v_T(n)/\dot{\theta}(n))$

In all cases, the time derivatives are computed as [20]

$$\dot{f}(n) \approx \Delta f(n) = \frac{\sum_{\tau=1}^{2} \tau(f(n+\tau) - f(n-\tau))}{2 \cdot \sum_{\tau=1}^{2} \tau^2}. \qquad (6)$$

2.4. Time function approximation via Legendre orthogonal polynomials

Many techniques have been proposed in the literature to approximate the time functions associated with the signing process. In [25], the Fourier Transform was used while in [28], the Wavelet Transform is proposed. In [29], an approach based on Legendre polynomials is introduced to represent handwritten mathematical symbols. The authors state that Legendre polynomials have the benefit that the coefficients can be computed in a small fixed number of arithmetic operations. In addition, the coefficients could be computed resorting to function moments at the end of each stroke so that the feature extraction could be performed in real time. In that work, the coefficients in the Legendre polynomials series expansions of the time functions were computed resorting to the function moments. In the present paper, the coefficients in the Legendre series approximations of the time functions associated with the signatures are used as features to model them. In this case, the coefficients are computed resorting to least squares techniques.

2.4.1. Orthogonal polynomials series expansions

A family of functions $\{g_i\}$ in (in general) an infinite dimensional functional space $H([a, b])$, defined in the domain $[a, b]$, is said to be orthonormal with respect to an inner product $\langle \cdot, \cdot \rangle$ in $H([a, b])$ if $\langle g_i, g_j \rangle = \delta_{ij}$, where δ_{ij} is the Kronecker delta.

Provided the inner product space $H([a, b])$ is complete with respect to the metric induced by the inner product, a set of orthonormal basis functions $\{h_i\}_{i=0}^{\infty}$ can be defined. In this case, any function $f \in H([a, b])$ can be uniquely represented by a series expansion in the orthonormal basis, that is

$$f = \sum_{i=0}^{\infty} \alpha_i h_i, \qquad (7)$$

where

$$\alpha_i = \langle f, h_i \rangle. \qquad (8)$$

It is not difficult to prove that the best (in the sense of the metric induced by the inner product) approximation of $f \in H([a, b])$ in an N-dimensional subspace is given by

$$f \approx \sum_{i=0}^{N} \alpha_i h_i. \qquad (9)$$

2.4.2. Time function approximation by Legendre orthogonal polynomials

The idea here is to approximate the time functions measured during the signature acquisition stage by a finite series expansion in orthonormal polynomials in the interval $[0, 1]$, and to use the series expansion coefficients as features. Particularly, Legendre polynomials are considered in this paper. In this case, the approximation equation (9) becomes

$$f(t) \approx \sum_{i=0}^{N} \alpha_i L_i(t), \qquad (10)$$

where $L_i(t)$ are the orthonormal Legendre polynomials[1] normalized to the interval $[0, 1]$.[2]

Since the time functions $f(t)$ are unknown, the coefficients in the truncated series expansions (10) cannot be computed as in (8) but rather they have to be estimated from a set of M (usually larger than $N + 1$) samples of the function at the time instants $\{t_1, t_2, \ldots, t_M\}$.

In matrix form, Eq. (10) at the time instants $\{t_1, t_2, \ldots, t_M\}$ can be written as

$$\underbrace{\begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_M) \end{bmatrix}}_{\mathbf{f}} = \underbrace{\begin{bmatrix} L_0(t_1) & L_1(t_1) & \cdots & L_N(t_1) \\ L_0(t_2) & L_1(t_2) & \cdots & L_N(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ L_0(t_M) & L_1(t_M) & \cdots & L_N(t_M) \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}}_{\boldsymbol{\alpha}} \qquad (11)$$

It is well known that the solution $\hat{\boldsymbol{\alpha}}$, in the least squares sense, of the overdetermined system of Eq. (11) is given by $\hat{\boldsymbol{\alpha}} = \mathbf{L}^{\dagger}\mathbf{f}$, where $\mathbf{L}^{\dagger} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T$, stands for the left pseudo-inverse of $\mathbf{L}$.

To illustrate the above estimation procedure, the x and y pen coordinates associated with a signature, and the corresponding approximations using Legendre polynomials with orders $N=21$, $N=15$ and $N=10$, are shown in Fig. 1.

The Best FIT[3] between the measured and the approximated time functions, for the above mentioned Legendre polynomial orders, are given in Table 1. Experimental results showed that further increasing the polynomial orders does not substantially
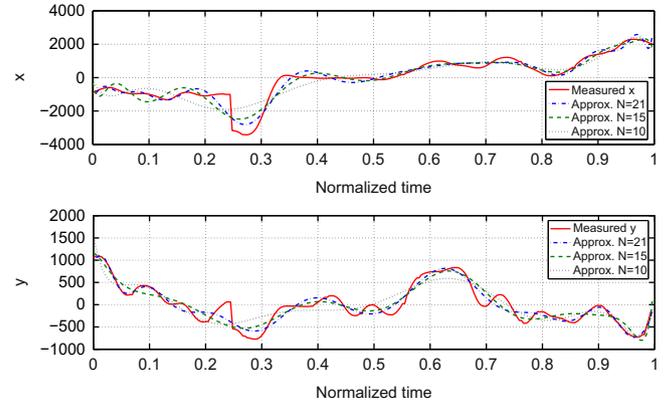
---

[1] The polynomials are orthonormal with respect to the standard inner product

$$\langle h_i(t), h_j(t) \rangle = \int_0^1 h_i(\tau) h_j(\tau) d\tau.$$

[2] Typically, Legendre polynomials are defined in the interval $[-1, 1]$.

[3] The best FIT is defined as

$$\text{Best FIT} = 100 \left(1 - \frac{\|x - x_{approx}\|}{\|x - x_{mean}\|} \right). \qquad (12)$$

.

Fig. 1. Time functions: x and y pen coordinates (red solid line) and their corresponding approximations by Legendre polynomials with orders $N=21$ (blue dashed line), $N=15$ (green dash-dotted line) and $N=10$ (black dotted line). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 1
Best FIT between the measured and the approximated time functions.

N	21	15	10
FIT_x (%)	77.7955	68.9708	57.6664
FIT_y (%)	70.7341	62.9579	53.3995

improve the approximation accuracy. This is an expected result, taking into account the bias-variance tradeoff inherent to least squares estimation from noisy data. As can be observed from Table 1, for all the considered Legendre polynomials orders the best FITs corresponding to the x coordinate time functions outperform the ones associated with the y coordinate time functions. This is probably due to the fact that the y coordinate time functions are, in general, more discontinuous than the x coordinate ones, and therefore they are more difficult to represent with the Legendre polynomials.

3. Consistency measure

An important property of a feature is its discriminative capability. Features associated with genuine signatures should be close to each other while distances between features associated with genuine and forged signatures should be relatively large. This property is usually called *consistency* of the feature.

A measure of consistency based on the features would be difficult to compute since they may have different lengths. It is then more reasonable to define a consistency measure based on the distances among features and not on the features themselves.

In this paper, the consistency of a given feature will be computed based on the statistics of the intraclass (for the genuine signature class) and interclass (between the genuine and forged signature classes) distances. A consistency factor d, for each signer, could then be defined as follows:

$$d = \frac{\mu_D(C_g, C_f) - \mu_D(C_g, C_g)}{\sqrt{\sigma_D^2(C_g, C_g)} + \sqrt{\sigma_D^2(C_g, C_f)}}, \qquad (13)$$

where C_g and C_f stand for the genuine and the forged classes, respectively, and where $\mu_D(C_g, C_g)$ and $\sigma_D^2(C_g, C_g)$, and $\mu_D(C_g, C_f)$ and $\sigma_D^2(C_g, C_f)$ are the sample means and sample variances of the genuine intraclass distances and the genuine-forged interclass distances, respectively.
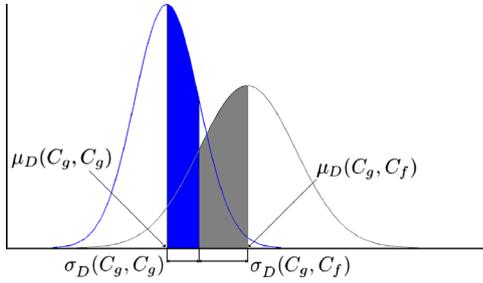
Fig. 2. Genuine intraclass distance distribution (blue) and genuine-forged interclass distance distribution (grey). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

The consistency factor in (13) is normalized in such a way that, under the assumption of Gaussian distributions for the involved distances, it equals 1 when the means $\mu_D(C_g, C_f)$ and $\mu_D(C_g, C_g)$ are separated by the sum of the respective standard deviations. This is illustrated in Fig. 2. The larger the consistency factor, the more consistent the features are.

An alternative definition of consistency, based on the Fisher ratio [30]

$$J = \frac{(\mu_D(C_g, C_f) - \mu_D(C_g, C_g))^2}{\sigma_D^2(C_g, C_g) + \sigma_D^2(C_g, C_f)}, \tag{14}$$

would be the following:

$$\tilde{d} = \sqrt{J} = \frac{\mu_D(C_g, C_f) - \mu_D(C_g, C_g)}{\sqrt{\sigma_D^2(C_g, C_g) + \sigma_D^2(C_g, C_f)}}. \tag{15}$$

Here, the interpretation of the normalization in Fig. 2 is not possible any more. The definition in (15) has been employed by several authors (see for instance [11,21]), while the definition in (13) has been introduced by the present authors in [9].

4. Signature database

The publicly available SigComp2011 Dataset [31] presented within ICDAR 2011 is used. It has two separate datasets, one containing genuine and forged Western signatures (Dutch ones) and the other one containing genuine and forged Chinese signatures. The available forgeries are skilled forgeries, which are simulated signatures in which forgers (different signers than the reference one) are allowed to practice the reference signature for as long as they deem it necessary. Fig. 3(a) and (b) shows the offline (left) and the online (right) versions of typical examples of genuine (top) and forged (bottom) signatures, from the Dutch and Chinese datasets, respectively. The data was collected from realistic and forensically relevant scenarios. The signatures were acquired using a ballpoint pen on paper (WACOM Intuos3 A3 Wide USB Pen Tablet), which is the natural writing process. This is in contrast to the approach of other researchers who tested signatures produced on a PDA or with a Wacom-stylus on a glass or plastic surface.

Each of the datasets in the SigComp2011 Database is divided into two sets, namely, the Training Set and the Testing Set. The Dutch (left) and the Chinese (right) datasets are described in Table 2.[4]

The measured data consists of three discrete time functions: pen coordinates x and y, and pen pressure p. In addition to this

raw data, the extended functions described in Section 2.3 are computed.

5. Evaluation protocol

In [11], the most common time functions were individually compared based on a consistency model. In this paper, the consistency factor introduced in (13) is used to evaluate the discriminative capability of the feature vectors composed by the Legendre polynomial coefficients associated with all possible combinations of the considered time functions (x, y, p, v_x, v_y, a_T, ρ).

The correlation between the consistency factor and the corresponding verification error for each feature combination is analyzed. It is important to study this relationship since, in case of a high correlation, the consistency factor would be used to single out the best feature combination, in terms of the verification error. This, without the need of explicitly computing the error.

5.1. Consistency computation

The consistency factor quantifies the discriminative power of a particular combination of time functions. Based on this value, it is possible to select the most suitable combination of time functions to be used in a verification system. This selection has to be done in the training stage. The consistency factor should then be computed with the signatures available during this stage. It is the common case that, when training a verification system, skilled forgeries are not available. For this reason, the consistency factor for a particular signer will be computed using the genuine signatures corresponding to all the remaining signers as forgeries. This will result in larger consistency factors compared with the ones that would be obtained using skilled forgeries for each signer. In any case, since the Database does contain skilled forgeries, the consistency factor will also be computed using them, for comparison purposes.

The datasets in the SigComp2011 Database are divided into two sets, namely, the Training Set and the Testing Set (see Table 2). The consistency factor will then be calculated as mentioned in the previous paragraph over the Dutch and Chinese Training Sets. The correlation between the consistency factor and the verification errors will be computed over the Dutch and Chinese Training Sets as well.

5.2. Verification performance evaluation

Two well known state-of-the-art classifiers, SVMs [33] and RFs [34], are used to assess the verification performance of the different time function combinations.

For each of the datasets, namely, Dutch and Chinese, the optimization of the meta-parameters of the system is performed over the corresponding Training Set while the corresponding Testing Set is used for independent testing purposes.

The tuning parameters to adjust are the order of the Legendre polynomials[5] and the internal parameters of the classifiers. For SVMs, the parameters[6] are the scale σ^2 in the Radial Basis Functions (RBF) kernel,[7] and the regularization parameter C. For RFs, the parameters are the number of trees to grow and the number of randomly selected splitting variables to be considered at each node.[8]

---

[4] The amount of genuine and forged signature samples may differ from those in [31] since when making signatures available for the research community some of them were missing [32].

[5] To select the optimal order, this parameter was varied from 1 to 25.

[6] Optimized, within the range 10^{-10} to 10^{10}, using tune.svm of the e1071 Package [35].

[7] The RBF kernel is defined as $K(\mathbf{x}(n), \mathbf{x}(k)) = e^{\|\mathbf{x}(n) - \mathbf{x}(k)\|^2 / \sigma^2}$.

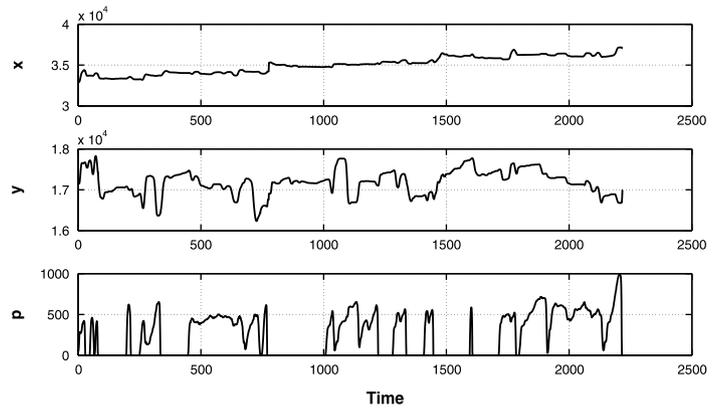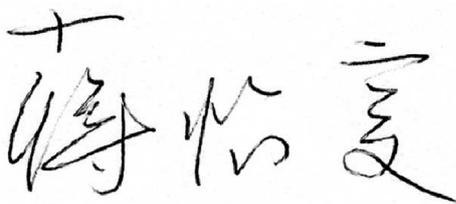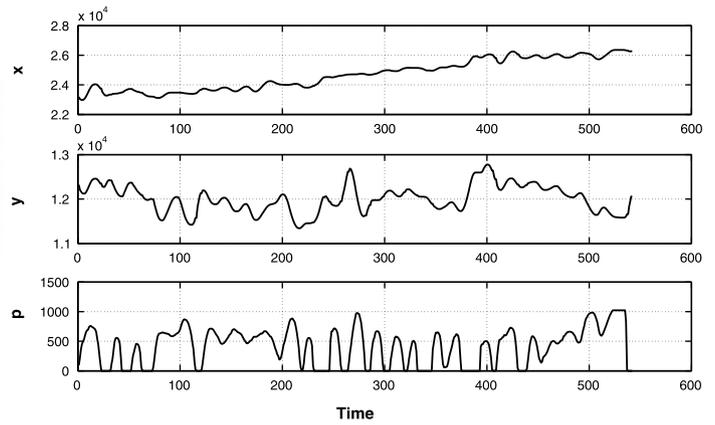[8] In general, the default values are a good choice for these parameters.
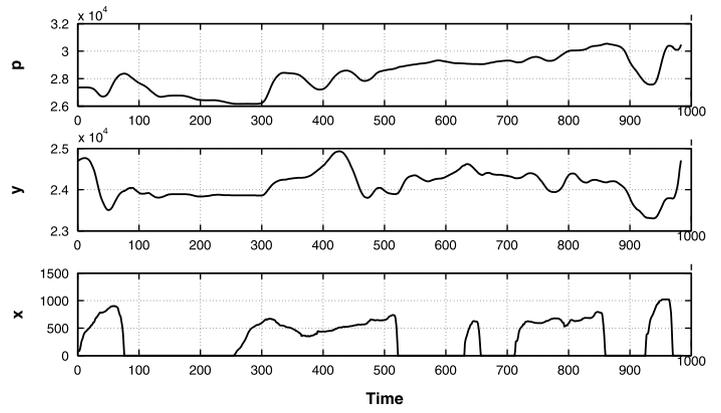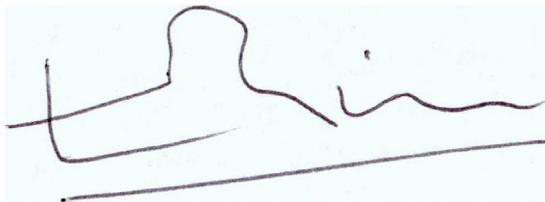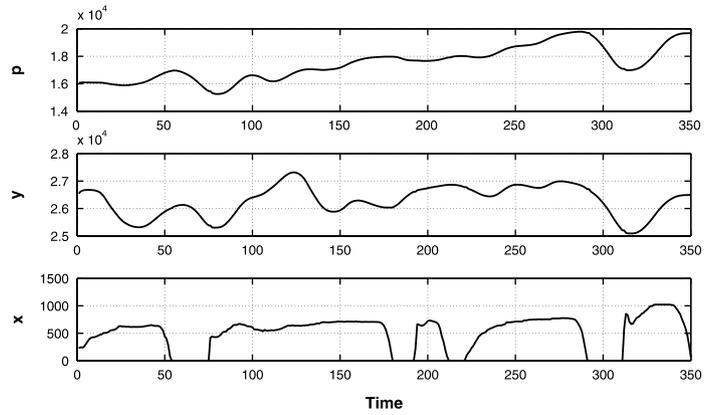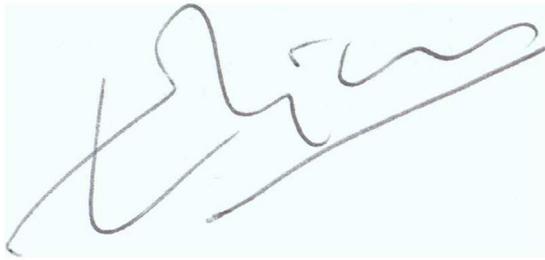
Fig. 3. Offline (left) and online (right) versions of a typical genuine (top) and forged (bottom). (a) Dutch signature and (b) Chinese signature.

Table 2
Online Dutch (left) and Chinese (right) datasets.

Dutch dataset			Chinese dataset		
Authors	Genuines	Forgeries	Authors	Genuines	Forgeries
Training set			Training set		
10	240	119	10	230	429
Testing set			Testing set		
54	1296	611	10	219	461

Table 3
Tested feature combinations.

Code	1	2	3	4	5
Feature Comb.	xy	p	$v_x v_y$	a_T	ρ
Code	6	7	8	9	10
Feature Comb.	xyp	$xyv_x v_y$	xya_T	$xy\,\rho$	$pv_x v_y$
Code	11	12	13	14	15
Feature Comb.	pa_T	$p\,\rho$	$v_x v_y a_T$	$v_x v_y \rho$	$a_T \rho$
Code	16	17	18	19	20
Feature Comb.	$xypv_x v_y$	$xypa_T$	$xyp\,\rho$	$xyv_x v_y a_T$	$xyv_x v_y \rho$
Code	21	22	23	24	25
Feature Comb.	$xya_T \rho$	$pv_x v_y a_T$	$pv_x v_y \rho$	$pa_T \rho$	$v_x v_y a_T \rho$
Code	26	27	28	29	30
Feature Comb.	$xypv_x v_y a_T$	$xypv_x v_y \rho$	$xypa_T \rho$	$xyv_x v_y a_T \rho$	$pv_x v_y a_T \rho$
Code	31				
Feature Comb.	$xypv_x v_y a_T \rho$				

To obtain statistically significant results, a 5-fold cross-validation (5-fold CV) is performed over the Testing Set to estimate the testing errors. For each instance of the 5-fold CV, a signature model is trained for each signer, using only genuine signatures. To train the signature model for a particular signer, the genuine class consists in the genuine signatures of the signer available in the corresponding training set of the 5-fold CV, while the forged class consists in the genuine signatures of all the remaining signers in the dataset available in the same training set. The genuine and forged signatures of the signer under consideration available in the corresponding testing set of the 5-fold CV are used for testing. Only skilled forgeries are considered to calculate the testing errors. Random signatures, that is, signatures that belong to another signer, are also commonly used to test the verification systems. Nevertheless, in the present paper, they are not considered for testing since they seldom appear in real situations.

To evaluate the performance, the EER is calculated, using the Bosaris toolkit [36], from the Detection Error TradeOff (DET) Curve as the point in the curve where the FRR (False Rejection Rate) equals the FAR (False Acceptance Rate). The cost of the log-likelihood ratios $\hat{C}_{llr}$ and its minimal possible value $\hat{C}_{llr}^{min}$ [36] are computed using the toolkit as well. A smaller value of $\hat{C}_{llr}^{min}$ indicates a better performance of the system. Using these measurements to evaluate the performance of a signature verification system is proposed in [31], where the importance of computing the likelihood ratios was highlighted since they make Forensic Handwriting Experts (FHEs) able to combine the results obtained from an automatic verification system with other evidence presented in a court of law [22].

6. Results and discussion

For the sake of compactness of the notation, the different combinations of the considered time functions (x, y, p, v_x, v_y, a_T, ρ), are coded as indicated in Table 3. Note the reader that whenever the x coordinate is considered, also the y coordinate is included in the combination, since no preferential direction would be expected to exist a priori. The same holds for the x and y velocities.

6.1. Consistency analysis

The proposed consistency factor d (in (13)) was computed for each of the combinations listed in Table 3, using only the genuine signatures, over the 10 authors in the Dutch Training Set and over the 10 authors in the Chinese Training Set. Figs. 4 and 5 show the boxplots associated with the consistency factors for each feature combination over the 10 signers in the Training Set (left), and a detail of the boxplots associated with the two best (larger consistency factors) and two worst combinations (right), for the Dutch and Chinese datasets, respectively. The different feature combinations have been included in a nonincreasing order of the consistency factor, from left (most consistent) to right (least consistent).

Fig. 4 shows that the most consistent combinations for the Dutch data are the one containing the pressure, the total acceleration and the log curvature radius ($pa_T\rho$), and the one containing the pen coordinates, the pressure and the total acceleration ($xypa_T$). The overlapping notches in the boxplots would indicate that the difference between the medians is not statistically significant. This is not the case if one compares the best (left most boxplot) and the worst (right most boxplot) combinations, where no overlapping is present, indicating that the difference between the medians is statistically significant. Similar comments, *mutatis mutandi*, hold for the boxplots in Fig. 5 corresponding to the Chinese data. For this dataset, the most consistent combinations are the one containing the pen coordinates and the pressure (xyp), and the one containing the pen coordinates, the pressure and the velocities ($xypv_x v_y$).

The discriminative capability of the pen pressure has long been questioned in the literature of online signature verification. In particular, since the results presented during SVC 2004 suggested that using the pen pressure did not improve the verification performance. The results obtained in this paper, show that whenever the pen pressure is incorporated to a feature combination, the consistency factor is improved in almost all the cases, for both datasets. This indicates that the pen pressure, when combined with the other time functions, is a reliable feature. This is not the case if the pen pressure is employed as the unique feature. This result agrees with the one presented in [11] where the authors used only one feature at a time to compute the consistency. However, using only one time function is not a very realistic case. These observations agree with many other reported results in the literature (see for example [8,21]), where it is stated that the pen pressure is a useful feature to distinguish between signers when used in combination with other time functions. In addition, the fact that the most consistent combinations (shown in Figs. 4 and 5), for both datasets, contain the pen pressure, could suggest that the reliability of the pen pressure is not highly influenced by the considered cultural origin of the signatures, depending mainly on the signer.

For the Dutch data, in addition to the pen pressure, the acceleration is also present in the most consistent combinations. Moreover, as it is the case of the pen pressure, when the acceleration is incorporated to a particular feature combination it improves the consistency factor in most of the cases. Again, the acceleration by itself is not a highly consistent feature, in agreement with [11]. For the Chinese data, the x and y pen coordinates are present in the most consistent combinations, and they improve the consistency in most of the cases when incorporated to a particular feature combination. The agreement with [11] in the case of using the pen coordinates alone, is also found in this case.
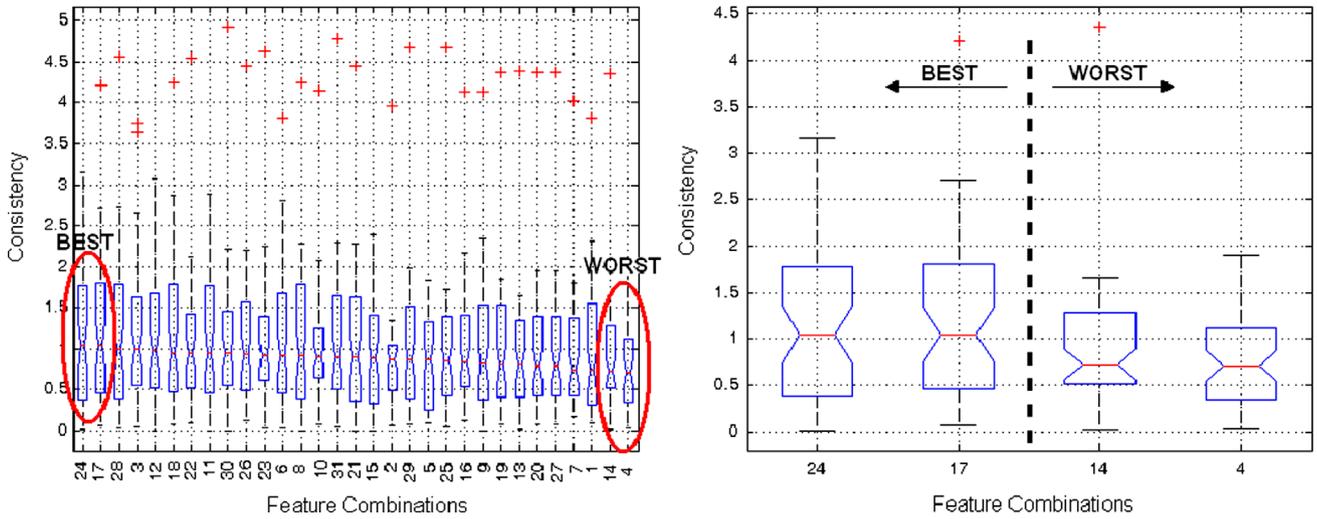
Fig. 4. Boxplots for the consistency factor over the 10 authors in the Training Set for the Dutch data (left), and detail of the two most and the two least consistent feature combinations (right).
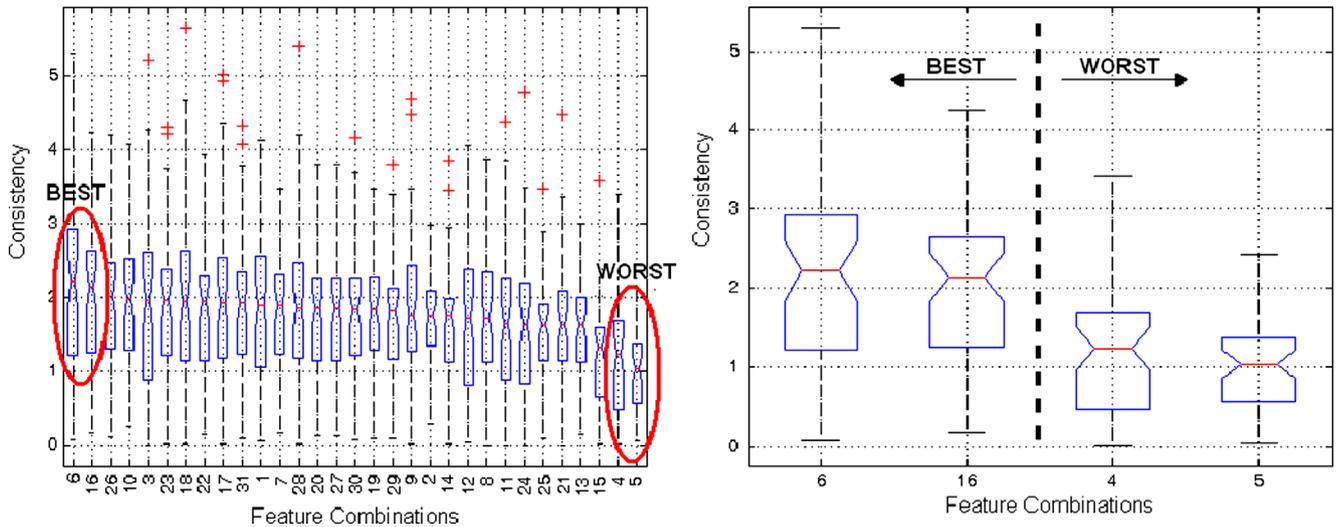

Fig. 5. Boxplots for the consistency factor over the 10 authors in the Training Set for the Chinese data (left), and detail of the two most and the two least consistent feature combinations (right).

The position information is likely to be more consistent for the Chinese data than for the Dutch data. Chinese signature style is, in most of the cases, close to the Chinese handwriting style, consisting of one or more multi-trace characters, while Western signatures can adopt several different styles. Due to the nature of Chinese characters, it is likely that the position information (x and y pen coordinates) has more discriminative power than in the case of Dutch data. On the other hand, for the Dutch data, the information about the changes in the velocity of writing (the acceleration) is likely to be more consistent. Dutch signatures are, in general, irregularly shaped, then it is likely that the position information is less important than the information regarding the rate of change of the position. Even further, the changes in the velocity seem to be more important than the velocity itself. The acceleration points out these changes, revealing typical hesitations of the forgers. This may not be the case for the Chinese signatures because as they are signatures containing separated characters, the velocity presents lot of changes and zero velocity moments in the genuine as well as in the forged signatures. Then, a hesitation of the forger is difficult to distinguish from a stop in the natural writing process of a genuine signer.

In order to assess the ability of the proposed consistency factor d to predict the verification performance of the system (in the sense of the classifiers error rates), the correlation between the consistency factor d and the verification error $\hat{C}_{llr}^{min}$ is calculated for the feature combinations listed in Table 3. Spearman's correlation coefficient [37] is used to quantify this correlation. The results are shown in Figs. 6 and 7, for the Dutch and Chinese data, respectively, where the feature combinations have been included in a nondecreasing order of the correlation coefficient, from left (highest absolute correlation) to right (lowest absolute correlation). An alternative measure of correlation would be the Pearson correlation coefficient [37] which is design to measure the strength of the linear dependence between two variables. While the Pearson correlation coefficient is limited to linear dependence, Spearman's correlation coefficient is more general, in the sense that it measures the monotonic dependence between two variables (not restricted to a linear function), being the best option for the current analysis.

Fig. 6 shows that, for the Dutch data, the correlation between the proposed consistency factor d and the $\hat{C}_{llr}^{min}$, and the correlation between the consistency factor $\tilde{d}$ (in (15)) and the $\hat{C}_{llr}^{min}$, are very

Fig. 6. Spearman's correlation coefficient between the Consistency factors (d and $\tilde{d}$) and the $\hat{C}_{llr}^{min}$ calculated for the signatures in the Dutch Training Set using the RF classifier (top) and the SVM classifier (bottom).
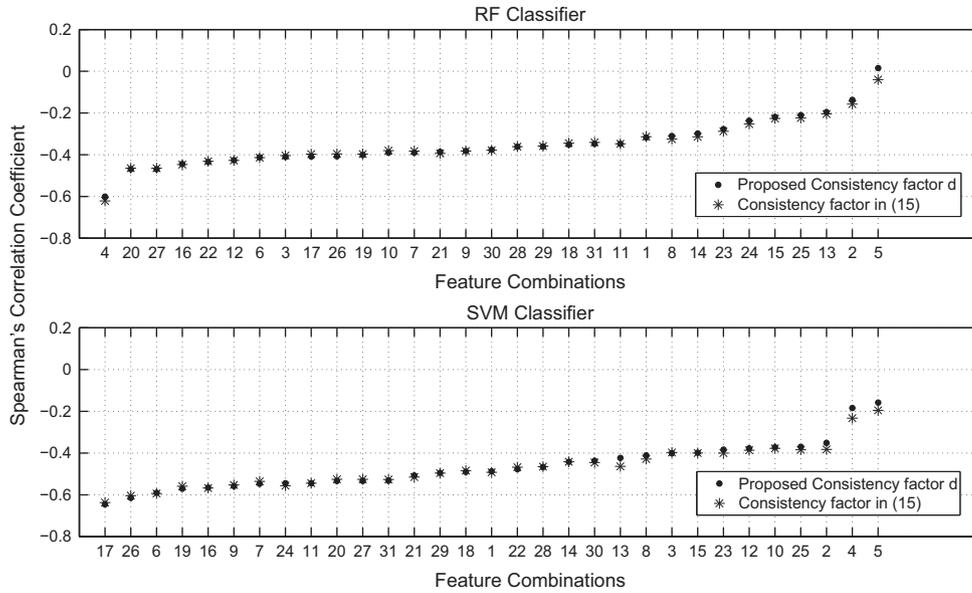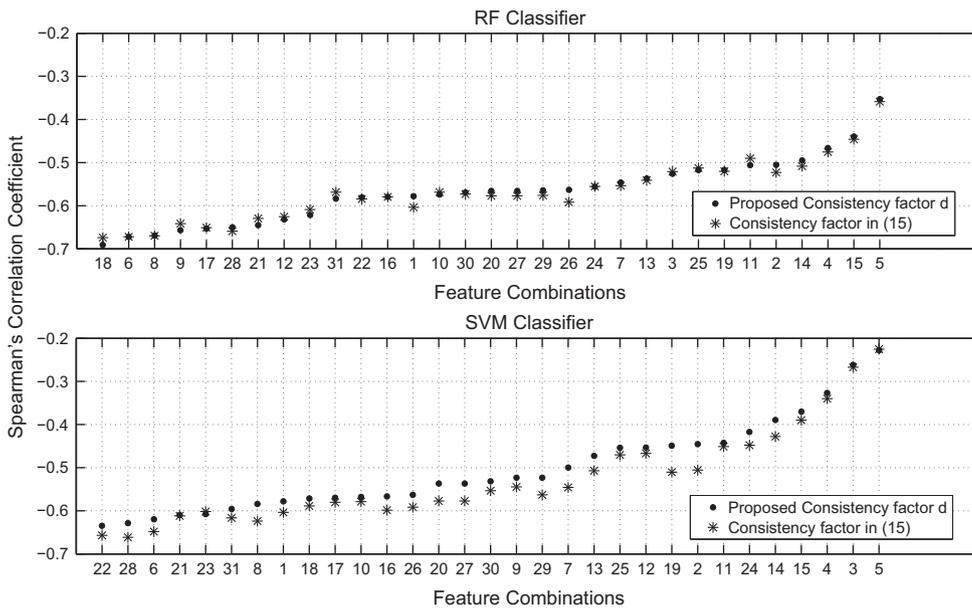

Fig. 7. Spearman's correlation coefficient between the Consistency factors (d and $\tilde{d}$) and the $\hat{C}_{llr}^{min}$ calculated for the signatures in the Chinese Training Set using the RF classifier (top) and the SVM classifier (bottom).

similar. On the other hand, Fig. 7 shows that, for the Chinese data, the correlation between the consistency factor and the $\hat{C}_{llr}^{min}$ is slightly better in the case of using the consistency factor $\tilde{d}$ (in (15)), specially when using the SVM classifier. Nevertheless, the best correlations obtained by d (-0.6003 and -0.6457 for the RF and SVM classifiers, respectively) in the case of the Dutch data and the ones in the case of the Chinese data (-0.6911 and -0.6347 for the RF and SVM classifiers, respectively) are indicating an acceptable correlation between the consistency factor and the verification error. That is, d could be used as an indicative of the verification performance to be expected for a particular feature combination (and a particular classifier), for both datasets. From Figs. 6 and 7 it can be seen that the correlation between the consistency factor and the verification error is highly dependent on the classifier being used. This correlation is an effective and

maybe more accurate indicative of the verification performance than the consistency factor itself. Nevertheless, it is important to highlight that if the best feature combination is to be chosen based either on the consistency or the correlation, the consistency factor has the advantage of not being dependant on the classifier being used.

While the decision about which feature combination is to be used must be necessarily made based on consistency factors computed with genuine signatures, in real situations the verification system is likely to be subjected to skilled forgeries. For the consistency factor, computed during the training stage, to be reliable, it must have a high correlation with the consistency factor computed resorting to skilled forgeries. Since the Dutch and Chinese Training Sets do contain skilled forgeries, this correlation can be computed. Spearman's correlation coefficient is employed to quantify this. Figs. 8 and 9 show the boxplots for Spearman's

correlation coefficient over all the feature combinations listed in Table 3, for d and $\tilde{d}$, for the Dutch and Chinese Training Sets, respectively. From Figs. 8 and 9 it can be observed that the correlation values using the consistency factor d are better than the ones using $\tilde{d}$. This is an important advantage of the proposed consistency factor d. Despite the fact that the correlation for the Dutch data (Fig. 8) is higher than for the Chinese data (Fig. 9), both

correlation coefficients are high enough to allow the use of the consistency factor computed using only genuine signatures in a real situation in which skilled forgeries are not available to train the system but they are present to test the system. Of course, this is strongly dependant on the forgeries quality.

6.2. Verification performance analysis

The verification performance for each combination in Table 3 is quantified by the *EER*, $\hat{C}_{llr}$ and $\hat{C}_{llr}^{min}$, over the Dutch and Chinese Testing Sets. The experiments were performed using the state-of-the-art classification techniques RF and SVM. For the RF classifier, the number of trees was set to 500 and the number of randomly selected splitting variables was equal to $\sqrt{P}$, where P is the dimension of the feature vector, for both datasets. For the SVM classifier the internal parameters were set to the optimal values $\sigma^2 = 10^7$ and $C = 1$, for the Dutch dataset and $\sigma^2 = 10^7$ and $C = 10$ for the Chinese dataset. Finally, the order of the Legendre polynomials was set to $N = 21$, for both datasets. Figs. 10 and 11 show the verification errors corresponding to the Dutch and the Chinese Testing Sets, respectively, when using RF (left) and SVM (right) as the classifiers, for all the feature combinations in Table 3. In Figs. 10 and 11, the feature combinations have been included in a nondecreasing order of the errors, from left (best performance) to right (worst performance).

Fig. 10 shows that, for the Dutch data, whenever the pen pressure or the acceleration is included in a feature combination, the error rate is improved, independently of the classifier being used. In addition, these two features are also the ones that improve the consistency of a feature combination whenever they are incorporated (see Section 6.1). This fact is important because it means that whenever an improvement in the consistency factor is made by incorporating one of these features to a feature combination it can be expected an improvement in the verification performance of that combination. Then, in these cases, there would be no need to explicitly compute the error rate in order to evaluate its improvement, because it would be predicted by the improvement in the consistency factor. Further, note that the inclusion of the x and y velocities is also helpful since they are contained in the feature combinations that perform better, for both classifiers. For the Chinese data (Fig. 11) it is also the case that whenever the pen pressure or the acceleration is included in a feature combination, the error rate is improved, independently of the classifier being used. In the case of incorporating the pen pressure, there exist a correspondence between an improvement in the consistency factor and an improvement of the verification error. Like for the Dutch data, the x and y velocities are also useful, for both classifiers. Finally, it is important to remark that, for this data, the x and y pen coordinates are included in most of the best combinations

Fig. 8. Boxplots for Spearman's correlation coefficient between d (left) and $\tilde{d}$ (right) computed using only genuine signatures and using skilled forgeries from the Dutch Training Set, over all the combinations tested.

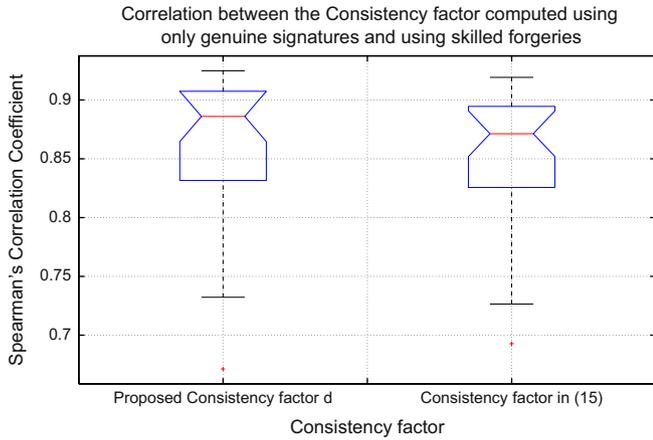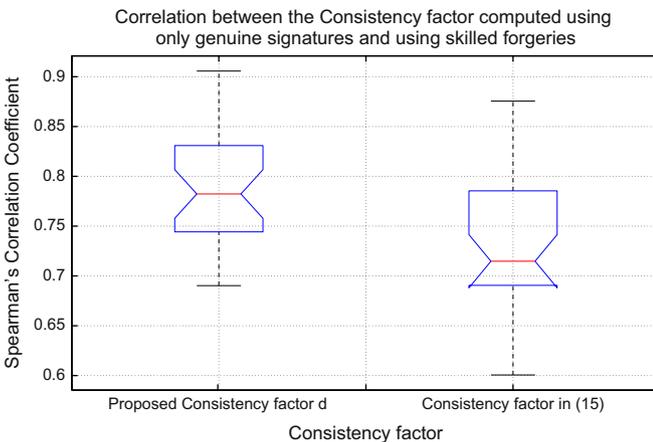**Fig. 9.** Boxplots for Spearman's correlation coefficient between d (left) and $\tilde{d}$ (right) computed using only genuine signatures and using skilled forgeries from the Chinese Training Set, over all the combinations tested.
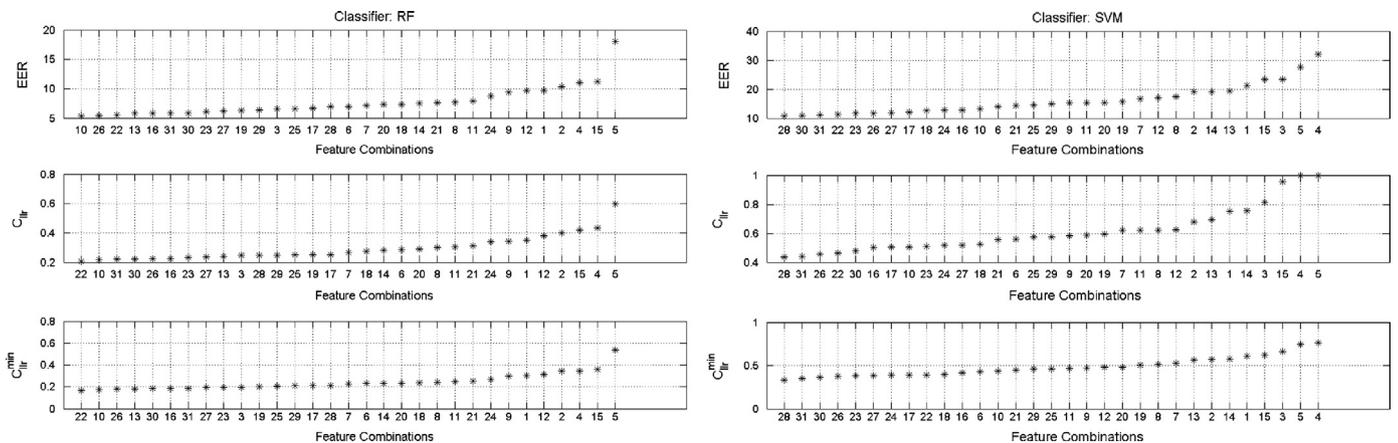

Fig. 10. *EER* (top), $\hat{C}_{llr}$ (middle) and $\hat{C}_{llr}^{min}$ (bottom), for the Dutch Testing Set when using RF (left) and SVM (right) as classifiers.
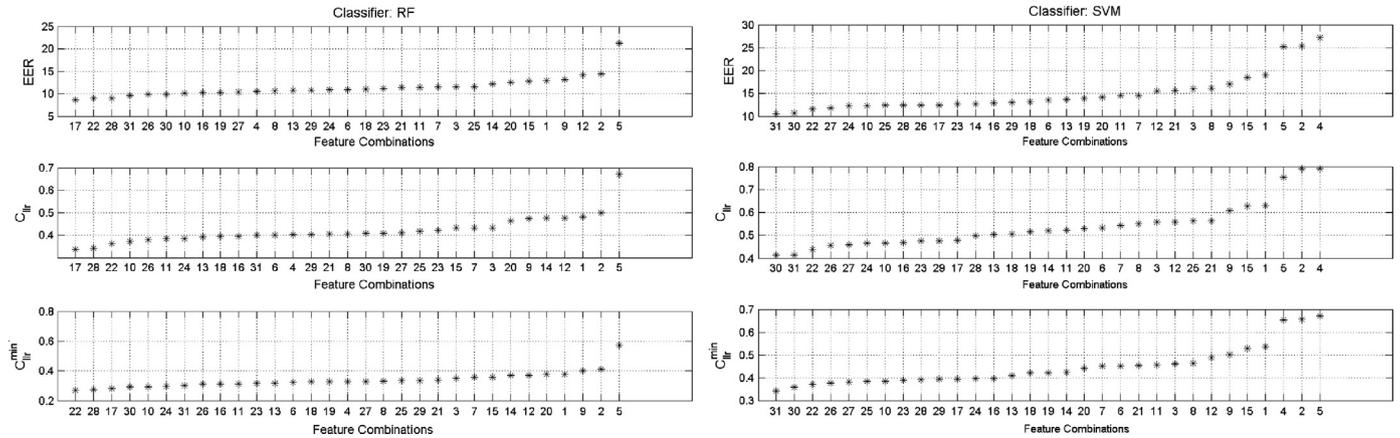
Fig. 11. *EER* (top), $\hat{C}_{llr}$ (middle) and $\hat{C}_{llr}^{min}$ (bottom), for the Chinese Testing Set when using RF (left) and SVM (right) as classifiers.

Table 4
Best verification results for the Dutch (left) and Chinese (right) datasets, for both classifiers.

Dutch dataset					Chinese dataset				
Features	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$	Features	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$
$pv_xv_ya_T$	RF	5.5	0.2039	0.1652	$pv_xv_ya_T$	RF	8.93	0.3620	0.2722
$xypa_T\rho$	SVM	10.68	0.4368	0.3323	$xypv_xv_ya_T\rho$	SVM	10.54	0.4139	0.3419

(in the sense of the verification errors), for both classifiers. This is in line with the analysis in Section 6.1, where it was argued that the position information (x and y pen coordinates) is likely to have more discriminative power for this type of data.

Incorporating the pen pressure has been shown to improve the verification results, independently of the classifier being used and of the type of data being considered. The independence of the classifier reveals a high discriminative power of the feature by itself. This agrees with the observations in Section 6.1 where it was shown that incorporating the pen pressure to a feature combination improves its consistency. Although in the present paper only Dutch (as an example of Western signatures) and Chinese signatures are considered, the independence of the data is very important since it means that the pen pressure could be a useful feature for any signature style. Of course, more data from different cultures have to be analyzed in order to make further conclusions, but this observation can be a very promising starting point.

Table 4 shows the best results for the Dutch (left) and Chinese (right) data obtained using RF and SVM as classifiers. Regarding the Dutch data, the best results are obtained by the feature combinations $pv_xv_ya_T$ (for RF) and $xypa_T\rho$ (for SVM), while for the Chinese data, the best results are obtained by the feature combinations $pv_xv_ya_T$ (for RF) and $xypv_xv_ya_T\rho$ (for SVM). This makes sense since the pen pressure and the acceleration are reliable features, for both datasets, from both points of view, viz., consistency and verification performance. In the case of using the RF classifier, including the x and y velocities helps to get better results, while in the case of using the SVM classifier it is necessary to incorporate more features to get better results. In addition, the results obtained when using RF are better than the ones obtained with SVM. To the best of the authors' knowledge, there are no conclusive results regarding which one, between RF and SVM, is the best classifier, independently of the chosen features, in applications of handwriting recognition. For instance, the results in [38] show that SVM outperforms RF as a classifier, for the particular features (different from the ones chosen here) considered in that paper, while in [9] the results using RF outperform the ones using SVM.

Table 5
Best verification results for the Dutch (left) and Chinese (right) datasets.

Features	Class.	Dutch dataset			Chinese dataset		
		EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$
$pv_xv_ya_T$	RF	5.5	0.2039	0.1652	8.93	0.362	0.2722
System		Acc.	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$	Acc.	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$
Commercial		96.27	0.2589	0.1226	93.17	0.4134	0.2179
1st. non-commercial		93.49	0.4928	0.2375	84.81	0.5651	0.3511

From Figs. 10 and 11 it can be noticed that for the Chinese data more features are needed than for the Dutch data in order to get better results. In addition, the results for the Dutch signatures are better than those for the Chinese ones. Generally speaking, Chinese signatures appear to be more complex than Dutch signatures, in the sense that they have multiple separated characters composed by multiple traces leading to discontinuities in the time functions associated with the signing process, then it is not surprising that more features are needed to model them in order to reach better verification results and that these results are not as good as in the case of the Dutch data. This is in line with the observations in [31], indicating that Chinese signatures are more challenging and that a lot of research has to be done on this type of data.

For the purposes of comparison, the best results obtained in this paper are shown in Table 5 together with the best commercial and non-commercial systems in the SigComp2011 competition [31] for the Dutch (left) and Chinese (right) data. Since the EER was not reported in the competitions results in [31], the accuracy (Acc) has been included instead. The comparison is then performed taking into account $\hat{C}_{llr}^{min}$. Even though the results are not as good as the corresponding to the best commercial system (*xyzmo*,[9] see [31]), they

---

[9] http://www.xyzmo.com.

Table 6
Verification results over the Dutch Testing Set for the best, regarding consistency and correlation, feature combinations.

Comb.	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$	Comb.	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$
Most consistent					**Highest corr.: cons. gen. and skilled for.**				
$pa_T\rho$	RF	8.75	0.3379	0.2695	$xya_T\rho$	RF	7.65	0.3109	0.2525
	SVM	12.78	0.5154	0.3889		SVM	14.25	0.5563	0.4460
$xypa_T$	**RF**	**6.65**	**0.2531**	**0.2122**	$xypa_T$	**RF**	**6.65**	**0.2531**	**0.2122**
	SVM	12.04	0.5056	0.3901		**SVM**	**12.04**	**0.5056**	**0.3901**
$xypa_T\rho$	RF	6.91	0.2478	0.2131	$pv_xv_ya_T\rho$	RF	5.84	0.2208	0.1862
	SVM	**10.68**	**0.4368**	**0.3323**		SVM	10.94	0.4782	0.3644
Highest corr.: cons. and $\hat{C}_{llr}^{min}$ (RF)					**Highest corr.: cons. and $\hat{C}_{llr}^{min}$ (SVM)**				
a_T	RF	11.06	0.4336	0.3450	$xypa_T$	RF	6.65	0.2531	0.2122
	SVM	31.95	1.000	0.7665		SVM	12.04	0.5056	0.3901
$xyv_xv_y\rho$	RF	7.29	0.2885	0.2307	$xypv_xv_ya_T$	**RF**	**5.44**	**0.2247**	**0.1791**
	SVM	15.31	0.5857	0.4823		**SVM**	**11.63**	**0.4567**	**0.3758**
$xypv_xv_y\rho$	**RF**	**6.20**	**0.2347**	**0.1957**	xyp	RF	6.96	0.2869	0.2303
	SVM	**11.88**	**0.5176**	**0.3854**		SVM	13.92	0.5577	0.4294

Table 7
Verification results over the Chinese Testing Set for the best, regarding consistency and correlation, feature combinations.

Comb.	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$	Comb.	Class.	EER	$\hat{C}_{llr}$	$\hat{C}_{llr}^{min}$
Most consistent					**Highest corr.: cons. gen. and skilled for.**				
xyp	RF	10.93	0.3998	0.3239	v_xv_y	RF	11.50	0.4336	0.3505
	SVM	13.54	0.5320	0.4523		SVM	16.06	0.5583	0.4612
$xypv_xv_y$	RF	10.27	0.3948	0.3116	$pv_xv_y\rho$	RF	11.17	0.4238	0.3161
	SVM	12.91	0.4673	0.3965		SVM	12.63	0.4765	0.3884
$xypv_xv_ya_T$	**RF**	**9.85**	**0.3786**	**0.3112**	pv_xv_y	**RF**	**10.07**	**0.3727**	**0.2955**
	SVM	**12.41**	**0.4547**	**0.3752**		**SVM**	**12.36**	**0.4654**	**0.3842**
Highest corr.: cons. and $\hat{C}_{llr}^{min}$ (RF)					**Highest corr.: cons. and $\hat{C}_{llr}^{min}$ (SVM)**				
$xyp\ \rho$	RF	11.01	0.3940	0.3262	$pv_xv_ya_T$	**RF**	**8.93**	**0.3620**	**0.2722**
	SVM	**13.11**	**0.5052**	**0.4202**		**SVM**	**11.57**	**0.4377**	**0.3703**
xyp	**RF**	**10.93**	**0.3998**	**0.3239**	$xypa_T\rho$	RF	9.02	0.3408	0.2739
	SVM	13.54	0.5320	0.4523		SVM	12.39	0.4969	0.3906
xya_T	RF	10.59	0.4058	0.3308	xyp	RF	10.93	0.3998	0.3239
	SVM	16.16	0.5505	0.4626		SVM	13.54	0.5320	0.4523

would have ranked first among the non-commercial systems and second among all the participants.

For the purposes of summarizing, Tables 6 and 7 show the verification errors, quantified by the *EER*, $\hat{C}_{llr}$ and $\hat{C}_{llr}^{min}$, for the three most consistent feature combinations (top left section of the tables), the three ones that obtained the highest correlation between the consistency factor and the verification error $\hat{C}_{llr}^{min}$ using RF (bottom left section) and using SVM (bottom right section), and the three ones that obtained the highest correlation between the consistency factor computed using only genuine signatures and using skilled forgeries (top right section), over the Dutch and Chinese Testing Sets, respectively. The best results in Tables 6 and 7 using RF and SVM classifiers, are highlighted in boldface.

From Tables 4 and 6, it can be observed that using consistency as the criterion for feature selection leads to the feature combination with the best error rate which corresponds, in this case, to the SVM classifier (shaded cells in Table 6). Similarly, from Tables 4 and 7, it can be observed that using the correlation between the consistency and the $\hat{C}_{llr}^{min}$ for SVM as the feature

selection criterion leads to the feature combination with the best verification performance which corresponds, in this case, to the RF classifier (shaded cells in Table 7). Although the consistency and correlation measures in Tables 6 and 7 not always lead to the best combination, that is the one with the lowest error rate, they lead to combinations with low errors, and so they can be used as feature selection criteria.

7. Conclusions

All the possible feature combinations associated with the most commonly used time functions related to the signing process were analyzed, in order to provide some insight on their actual discriminative power for online signature verification. A consistency factor was defined to quantify the discriminative power of these different feature combinations. A fixed-length representation, based on Legendre polynomials series expansions, was used to represent the time functions associated with the signatures. Two different signature styles, namely, Western and Chinese, from

a publicly available Signature Database were considered to evaluate the performance of the verification system.

The experimental results show a good correlation between the consistency factor computed using only genuine signatures and using skilled forgeries. This robustness property is important since, in real applications, skilled forgeries are not available in the training phase but they do appear when testing the system. In addition, the experimental results also show a good correlation between the consistency values and the verification errors, suggesting that the former could be used to select the optimal (i.e., leading to the smallest verification error) feature combination.

Based on the defined consistency factor, the most consistent feature combinations were determined, for the two different signature styles being considered. These optimal combinations induced verification errors close to the smallest ones.

The experimental results show that the pen pressure improves the consistency factor and the verification errors whenever it is incorporated to a feature combination, for both signature styles. In this sense, the pen pressure proved to be a reliable feature that should be incorporated to any feature set.

The use of Legendre polynomials to model the time functions associated with the signatures proved to be a good choice, resulting in verification performances comparable to those of other state-of-the-art verification systems, tested on the same datasets. In addition, the use of Legendre coefficients as features results in a fixed-length feature vector avoiding the need for length normalization.

Conflict of interest

None declared.

References

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

[2] F. Leclerc, R. Plamondon, Automatic signature verification: the state of the art 1989–1993, International Journal on Pattern Recognition and Artificial Intelligence 8 (3) (1994) 643–660.

[3] R. Plamondon, S. Srihari, On-line and off-line handwriting recognition: a comprehensive survey, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (1) (2000) 63–84.

[4] D. Impedovo, G. Pirlo, Automatic signature verification: the state of the art, IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews 38 (5) (2008) 609–635.

[5] D. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, G. Rigoll, SVC2004: First international signature verification competition, in: International Conference on Biometric Authentication, Hong Kong, 2004, pp. 16–22.

[6] A. Kholmatov, B. Yanikoglu, Identity authentication using improved online signature verification method, Pattern Recognition Letters 26 (2005) 2400–2408.

[7] N. Houmani, S. Garcia-Salicetti, B. Dorizzi, On assessing the robustness of pen coordinates, pen pressure and pen inclination to time variability with personal entropy, in: IEEE Third International Conference on Biometrics: Theory, Applications, and Systems, Washington DC, USA, 2009.

[8] D. Maramatsu, T. Matsumoto, Effectiveness of pen pressure, azimuth, and altitude features for online signature verification, in: International Conference on Biometrics, 2007, pp. 503–512.

[9] M. Parodi, J. Gómez, Online signature verification based on Legendre series representation. consistency analysis of different feature combinations, in: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2012), Vol. 7441 of Lecture Notes in Computer Science, Springer, Berlin, 2012, pp. 715–723.

[10] L. Lee, T. Berger, E. Aviczer, Reliable on-line human signature verification systems, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (6) (1996) 643–647.

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

[12] S. Pal, M. Blumenstein, U. Pal, Non-english and non-latin signature verification systems: a survey, in: First International Workshop on Automated Forensic Handwriting Analysis, Beijing, China, 2011.

[13] J. Ji, Z. Lu, X. Chen, Similarity computation based on feature extraction for off-line Chinese signature verification, in: Sixth International Conference on Fuzzy Systems and Knowledge Discovery, Tianjin, China, 2009, pp. 291–295.

[14] J. wen Ji, X. su Chen Off-line Chinese signature verification segmentation and feature extraction, in: International Conference on Computational Intelligence and Software Engineering, 2009, pp. 1–4.

[15] T.H. Rhee, S.J. Cho, J.H. Kim, On-line signature verification using model-guided segmentation and discriminative feature selection for skilled forgeries, in: Sixth International Conference of Document Analysis and Recognition, Seattle, USA, 2001.

[16] ICDAR 2011, International Conference on Document Analysis and Recognition, Beijing, China, 19–21 September, 2011.

[17] K. Ueda, Investigation of off-line Japanese signature verification using a pattern matching, in: Seventh International Conference of Document Analysis and Recognition, Edinburgh, Scotland, 2003.

[18] M. Yoshimura, I. Yoshimura, Investigation of a verification system for Japanese countersignatures on traveler's cheques, Transactions of the IEICE J80-D-II (7) (1998) 1764–1773.

[19] M. Ismail, S. Gad, Off-line arabic signature recognition and verification, Pattern Recognition 33 (10) (2000) 1727–1740.

[20] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Ramos-Castro, J. Gonzalez-Rodriguez, HMM-based on-line signature verification: feature extraction and signature modelling, Pattern Recognition Letters 28 (2007) 2325–2334.

[21] J. Richiardi, H. Ketabdar, A. Drygajlo, Local and global feature selection for on-line signature verification, in: Eighth International Conference on Document Analysis and Recognition, Seoul, Korea, 2005.

[22] D.R.J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, J. Ortega-Garcia, Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems, Forensic Science International 155 (2005) 126–140.

[23] P. Tuyls, A. Akkermans, T. Kevenaar, G. Schrijen, A. Bazen, R. Veldhuis, Practical biometric authentication with template protection, in: Fifteenth International Conference on Pattern Recognition, Barcelona, Spain, 2000, pp. 45–47.

[24] H. Xu, R. Veldhuis, T. Kevenaar, A. Akkermans, A. Bazen, Spectral minutiae: a fixed-length representation of a minutiae set, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008.

[25] B. Yanikoglu, A. Kholmatov, Online signature verification using fourier descriptors, EURASIP Journal on Advances in Signal Processing (2009) 230–275.

[26] A.K. Jain, F.D. Griess, S.D. Connell, On-line signature verification, Pattern Recognition 35 (2002) 2963–2972.

[27] M.R.F.M. Martinez-Diaz, J. Fierrez, J. Ortega-Garcia, On the effects of sampling rate and interpolation in HMM-based dynamic signature verification, in: Ninth International Conference on Document Analysis and Recognition, Curitiba, Brazil, 2007, pp. 1113–1117.

[28] H. Chang, D. Dai, P. Wang, Y. Xu, F. Si, S. Huang, Online signature verification using wavelet transform of feature function, Journal of Information & Computational Science 9 (11) (2012) 3135–3142.

[29] O. Golubitsky, S. Watt, Distance-based classification of handwritten symbols, International Journal of Document Analysis and Recognition 13 (2) (2010) 133–146.

[30] C. Bishop, Pattern Recognition and Machine Learning, Springer, Berlin, NY, 2006.

[31] M. Liwicki, M. Malik, C. den Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, B. Found, Signature verification competition for online and offline skilled forgeries (SigComp2011), in: Eleventh International Conference on Document Analysis and Recognition, Beijing, China, 2011.

[32] SigComp2011, ICDAR 2011 Signature verification competition, ⟨http://www.iapr-tc11.org/mediawiki/index.php/Datasets_List⟩ (2011).

[33] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, NY, 1995.

[34] L. Breiman, Random Forests, Technical Report, Statistics Department, University of California, Berkeley (2001).

[35] E. Dimitriadou, K. Hornik, F. Leisch, D. Meyer, A. Weingessel, Misc. functions of the department of statistics (e1071), TU Wien (2010).

[36] N. Brümmer, J. du Preez, Application-independent evaluation of speaker detection, Computer Speech & Language 20 (2006) 230–275.

[37] J.D. Gibbons, S. Chakraborti, Nonparametric Statistical Inference, Marcel Dekker, Inc., New York, Basel, 2003.

[38] I. Bärbăntan, R. Potolea, Enhancements on a signature recognition problem, in: IEEE Sixth International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, Romania, 2010, pp. 141–147.

Marianela Parodi graduated as Electronic Engineer at the National University of Rosario (UNR), Argentina, in 2009, ranking at the top of her class. She is currently a Engineering Ph.D. Student at the same University, working on Automatic Signature Verification, supported by a Scholarship from CONICET (National Research Council). She is also part time teaching assistant at the UNR.

Juan C. Gómez has an Electronic Engineering degree from the National University of Rosario, Argentina (1983), and a Ph.D. in E&CE from the University of Newcastle, Australia (1998). He is Professor at the UNR and Director of the Laboratory for System Dynamics and Signal Processing. His research interests are in the areas of Multimedia Signal Processing, Pattern Recognition and System Identification.