

Published in IET Biometrics
 Received on 15th February 2013
 Revised on 2nd July 2013
 Accepted on 17th July 2013
 doi: 10.1049/iet-bmt.2013.0025

Special Issue: Selected papers from International Conference
 on Frontiers in Handwriting Recognition (ICFHR 2012)



ISSN 2047-4938

Orthogonal function representation for online signature verification: which features should be looked at?

Marianela Parodi¹, Juan C. Gómez¹, Marcus Liwicki², Linda Alewijnse³

¹Laboratory for System Dynamics and Signal Processing, FCEIA Universidad Nacional de Rosario, and CIFASIS, Argentina

²German Research Center for Artificial Intelligence (DFKI), 67663 Kaiserslautern, Germany

³Netherlands Forensic Institute, The Hague, The Netherlands

E-mail: parodi@cifasis-conicet.gov.ar; parodi.marianela@gmail.com

Abstract: In this study, several feature combinations are studied to analyse their relevance for online signature verification. Different time functions associated with the signing process are analysed in order to provide some insight on their actual discriminative power. This analysis could also help forensic handwriting experts (FHEs) to further understand the signatures and the writer's behaviour. Among the different feature combinations analysed, a set of features which seems to be relevant for signature analysis by FHEs is particularly considered. The feasibility of developing a system which could complement the FHEs work is evaluated. Two different approximations of the analysed time functions are proposed, one based on the Legendre polynomials and another based on the wavelet decomposition. The coefficients in these orthogonal series expansions of the time functions are used as features to model them. Two different signature styles are considered, namely, Western and Chinese, of one of the most recent publicly available signature databases. The experimental results are promising, in particular for the features that seem to be relevant for the FHEs, since the obtained verification error rates are comparable with the ones reported in the state-of-the-art over the same datasets.

1 Introduction

Automatic signature verification has long been considered an important research area in the field of biometrics [1–5]. Signature verification is the most popular method for identity verification. Signatures are recognised 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, whereas for online systems, dynamic information acquired during the signing process, such as x and y pen coordinates, pen pressure and pen inclination angles: azimuth and altitude, is available. The interest in the online approach for signature verification has increased in recent years because of the widespread use of electronic pen-input devices, such as digitiser tablets and personal digital assistants (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. Nevertheless, there are certain applications that demand the use of the offline approach.

forensic handwriting experts (FHEs) often only have the offline data available in their daily casework. To perform a forensic signature comparison, it is necessary to work with offline data, while online data can be used to perform biometric person verification/identification. Furthermore in the future, it might occur that FHEs will also have to deal with online signatures.

There are several aspects to evaluate when designing a signature verification system. The First International Signature Verification Competition (SVC2004) [6] addressed the researchers' need for benchmark databases and benchmarking rules for comparing different signature verification systems. Two online datasets were presented, one containing only the signatures' pen coordinates and the other one containing also the pen pressure and inclination angles. Researchers have long argued about the effectiveness of these different time functions for verification purposes. During SVC2004, the results were better when using only pen coordinates than when adding the pen pressure and inclination angles. Since then, several works have been presented concerning the best set of features to model the signatures. In [7], the authors state that using only pen coordinates lead to better results than incorporating the pen pressure. In [8], the authors compute the consistency of different features, resulting the pen coordinates more consistent than the pen pressure. On the other hand, some works showed improvements when

combining pen coordinates with pen pressure and inclination angles [9]. The results in [10, 11] and a preliminary work by the present authors [12], among others, are also in line with the idea that combining pen coordinates with the pen pressure lead to a verification performance improvement. The BioSecure Signature Evaluation Campaign (BSEC 2009) [13] presented two large online databases, which allowed the participants to evaluate the impact of mobility acquisition conditions, time variability and quality of the signatures in the verification performance. The time variability impact was also evaluated in [14], where the authors 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. An analysis of the quality of online signatures based on the sigma-lognormal model is carried out in [15], where the experiments show a high potential of certain kinematic features for signature quality assessment. The ICDAR2009 Signature Verification Competition [16] and the Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011) [17] introduced new forensic-like online and offline datasets, in order to make a step towards bridging the gap between automated performances and FHEs' comparisons. Further, in SigComp2011, methods used by FHEs to assess the value of the evidence were applied. In addition, the influence of the cultural origin of the signatures in the systems' performance, was also taken into account in SigComp2011, by providing two different datasets, containing Dutch and Chinese signatures, respectively. This is an important aspect that deserves more investigation. 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 [18], 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 Hindi ones. Offline [19, 20] and online [21] verification systems have been presented in the literature for Chinese signature verification. Further, the Chinese dataset presented in [22] encouraged researches to work on this type of data. In [23, 24], Japanese offline signature verification systems are presented. A new competition is going to be held within ICDAR 2013, providing a new Japanese signature database. In [25], an offline verification system for Arabic signatures is presented, whereas in [26] such a system is presented for Hindi signatures.

As mentioned above, several aspects can be studied for an online signature verification system. The conflicting results observed in the literature regarding which is the best set of time functions to model the signatures, make this discussion an open and interesting issue. In addition, to bridge the gap between the pattern recognition (PR) and the FHEs

communities is also an important task, being crucial for the automatic signature verification approach to be useful to FHEs. It would then be interesting to address these two aspects, by investigating different online feature combinations to determine their actual discriminative power, focusing the study, specially, in a feature set containing features which are relevant to FHEs. This analysis could help FHEs to further understand the signatures and the writer's behaviour. FHEs work with the offline specimens of the signature, so it is not possible for them to look at online features. Nevertheless, they have plenty of experience in the interpretation of some dynamic information that can be inferred from the offline signature [27–29]. FHEs try to understand the forgery process from the forgers' point of view. To make a good simulation of a signature, there will be a tradeoff between accuracy and velocity [30]. If the forger wants to produce an accurate copy of the specimen signature, it will likely be written slowly. This will result in bad line fluency and hesitations, artefacts that will be visible for the FHE. On the other hand, if the forger focusses on the writing velocity to make the forgery more fluent, he will aim at a variation that fits within the writer's variability. The curvature is hard to copy, and the forger will likely use slightly different curves at specific spots. Pen pressure is strongly dependent on the writing surface, then it is very difficult to be inferred from the offline specimen of the signature. A monotonous pressure over all the signature or the presence of more pressure in unusual places can also be signs of forged signatures. Earlier literature has been presented regarding the use of features motivated by FH examination [31, 32]. Nevertheless, much work needs still to be done in order to bridge the gap between the FHEs needs and the automatic signature verification developments.

The main aim of this paper is to study different online feature sets selected based on different criteria. Among them, a set of features that are relevant for signature analysis by FHEs is considered with particular attention. This analysis could help FHEs to further understand the signatures and the writer's behaviour, and could make the developed automatic signature verification system useful to them. Two different approximation schemes of the time functions are proposed, one based on the Legendre polynomials expansions and the other on wavelet decomposition. The coefficients in these orthogonal expansions of the time functions are used as features to model them. For evaluation purposes, in line with the intention of bridging the gap between PR and FHEs communities, a forensic-like publicly available signature database is used. Furthermore, to quantify the verification performance, in addition of computing the equal error rate (EER), the cost of the log-likelihood ratios \hat{C}_{llr} are computed, which allow FHEs to give an opinion on the strength of the evidence. Finally, the used signature database contains two different signature styles, namely, Western and Chinese, making it possible to analyse the influence of the cultural origin of the signatures in the performance of the verification system.

The main contributions of this paper are as follows:

- Several online feature sets selected based on different criteria are studied to give an insight on their actual discriminative power. This study is intended to also be interesting for FHEs.
- A set of automatically selected features is compared with a set of features which are meaningful for FHEs.

- A new feature extraction approach based on orthogonal series expansions of the time functions associated with the signatures is proposed. To the best of the authors' knowledge, this is the first time that the approach based on the Legendre polynomials is used in the context of signature verification.
- The experiments are performed on one of the most recent forensic-like publicly available signature datasets, containing Western and Chinese signatures, which have been used in one of the latest signature verification competitions. The FHEs needs regarding the performance evaluation of the systems are taken into account by computing the cost of the log-likelihood ratios \hat{C}_{llr} , which allow FHEs to give an opinion on the strength of the evidence.

The paper is organised as follows. The time functions and preprocessing tasks are described in Section 2. The feature extraction approach is described in Section 3, in particular, Subsections 3.1 and 3.2 focus on the Legendre polynomials based approximations and on wavelet-based approximations of the time functions, respectively. In Section 4, the database is described. Section 5 describes the different feature combinations proposed, whereas Section 6 is devoted to the description of the experiments. In Section 7, the experimental results are presented and discussed. Finally, some concluding remarks are given in Section 8.

2 Time functions and preprocessing

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 Normalisation

Depending on the given space to sign, signatures can be written in different sizes, writers can place them 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 normalisation fundamental preprocessing tasks. A width normalisation 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. The signatures are centred with respect to the origin of coordinates, by subtracting the corresponding mean values from the original x and y pen coordinates. Regarding rotation normalisation, conflicting views can be found in the literature. Some authors perform a correction in the main direction of the signature, rotating it until it has the direction of a predetermined baseline [33]. 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 exists 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. Owing to the acquisition process, the measured data may contain noise or gaps introduced during the signing process. Resampling is used to correct these acquisition artefacts and, in addition, to obtain a fixed-length resampled time function. Several works in the literature use resampling to remove redundant points from the measured signals [34, 35]. In [35], 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 simplifying the computation, 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 [7, 8, 33, 36]. Moreover in [7, 36], the authors observed that using resampling leads to worst 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 techniques are based on approximations of the time functions associated with the signing process by the Legendre polynomials and wavelets. The former delivers a fixed-length feature vector, so that no resampling of the original time functions is required, while for the latter the obtained feature vector length would depend on the length of the original time function. Then in the case of the approximation using wavelets, 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 [7], the incremental variations of the x and y pen coordinates are proposed. In [11], 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-order time derivatives. In this paper, the path velocity magnitude v_T , the path-tangent angle θ , the total acceleration a_T and the log curvature radius ρ [33] are computed from the basic function set composed of x and y pen coordinates and pen pressure p . Let $n = 1, \dots, L_{\text{sign}}$ be the discrete-time index of the measured functions and L_{sign} be the time duration of the signature in sampling units, then the above-mentioned extended functions are computed as:

- Path velocity magnitude: $v_T(n) = \sqrt{\dot{x}^2(n) + \dot{y}^2(n)}$.
- Path-tangent angle: $\theta(n) = \arctan(\dot{y}(n)/\dot{x}(n))$.
- Total acceleration: $a_T(n) = \sqrt{\dot{v}_T^2(n) + c^2(n)}$, where $c(n) = v_T(n)\dot{\theta}(n)$.
- Log curvature radius: $\rho(n) = \log[v_T(n)/\dot{\theta}(n)]$.

In all cases, the first-order time derivatives are computed as [33]

$$\dot{f}(n) \simeq \Delta f(n) = \frac{\sum_{\tau=1}^2 \tau(f(n+\tau) - f(n-\tau))}{2 \sum_{\tau=1}^2 \tau^2} \quad (1)$$

In this paper, the initial set of features will be composed by the x and y pen coordinates, the pressure p , the above-mentioned extended functions, namely, v_T , θ , a_T and ρ , their first-order time derivatives (these are the features proposed in [33]), and their second-order time derivatives

computed as in (1) from the corresponding first-order time derivatives.

3 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 favour the use of global features [11, 37]. 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. Several works in the literature have proposed a fixed-length representation of the signatures, among them, [36] where the authors employ the fast Fourier transform, can be mentioned. Further, a fixed-length model of the signatures can be required for certain biometric applications [38, 39].

In this paper, two different fixed-length representations of the signatures are proposed. One is based on the approximation of the different time functions by the Legendre orthogonal polynomials, introduced by the present authors in [12], and the other on their representation using wavelets.

3.1 Time function approximation via the Legendre polynomials

In this subsection, models of the time functions associated with the signing process, based on the Legendre series approximations, are presented. The coefficients of the Legendre series are computed resorting to least squares techniques. A similar approach for the representation of handwritten mathematical symbols has been proposed in [40]. 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.

3.1.1 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 $\{h_i\}_{i=1}^{\infty}$ 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

$$f = \sum_{i=1}^{\infty} \alpha_i h_i \quad (2)$$

where

$$\alpha_i = \langle f, h_i \rangle \quad (3)$$

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 the N -dimensional subspace spanned by the first N basis functions h_i is given by

$$f \simeq \sum_{i=1}^N \alpha_i h_i \quad (4)$$

3.1.2 Coefficient estimation: 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 (4) becomes

$$f(t) \simeq \sum_{i=1}^N \alpha_i L_i(t) \quad (5)$$

where $L_i(t)$ are the orthonormal (with respect to the standard inner product $\langle h_i(t), h_j(t) \rangle = \int_0^1 h_i(\tau) h_j(\tau) d\tau$) Legendre polynomials normalised to the interval $[0, 1]$. Typically, the Legendre polynomials are defined in the interval $[-1, 1]$.

Since the time functions $f(t)$ are unknown, the coefficients in the truncated series expansions (5) cannot be computed as in (3), but rather they have to be estimated from a set of M (usually larger than N) samples of the function at the time instants $\{t_1, t_2, \dots, t_M\}$.

In matrix form, (5) at the time instants $\{t_1, t_2, \dots, t_M\}$ can be written as

$$\underbrace{\begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_M) \end{bmatrix}}_f = \underbrace{\begin{bmatrix} L_1(t_1) & L_2(t_1) & \cdots & L_N(t_1) \\ L_1(t_2) & L_2(t_2) & \cdots & L_N(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ L_1(t_M) & L_2(t_M) & \cdots & L_N(t_M) \end{bmatrix}}_L \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}}_{\alpha} \quad (6)$$

It is well known that the solution $\hat{\alpha}$, in the least squares sense, of the overdetermined system of (6) is given by $\hat{\alpha} = L^\dagger f$, where $L^\dagger = (L^T L)^{-1} L^T$ stands for the left pseudo-inverse of L .

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

The Best FIT, defined as

$$\text{Best FIT} = 100 \left(1 - \frac{\|x - x_{\text{approx}}\|}{\|x - x_{\text{mean}}\|} \right) \quad (7)$$

between the measured and the approximated time functions, for the above-mentioned Legendre polynomial orders, are given in Table 1. It can be observed that a reasonable FIT is obtained for $N=21$. Experimental results show that further increasing the polynomial orders does not substantially improve the approximation accuracy. This is

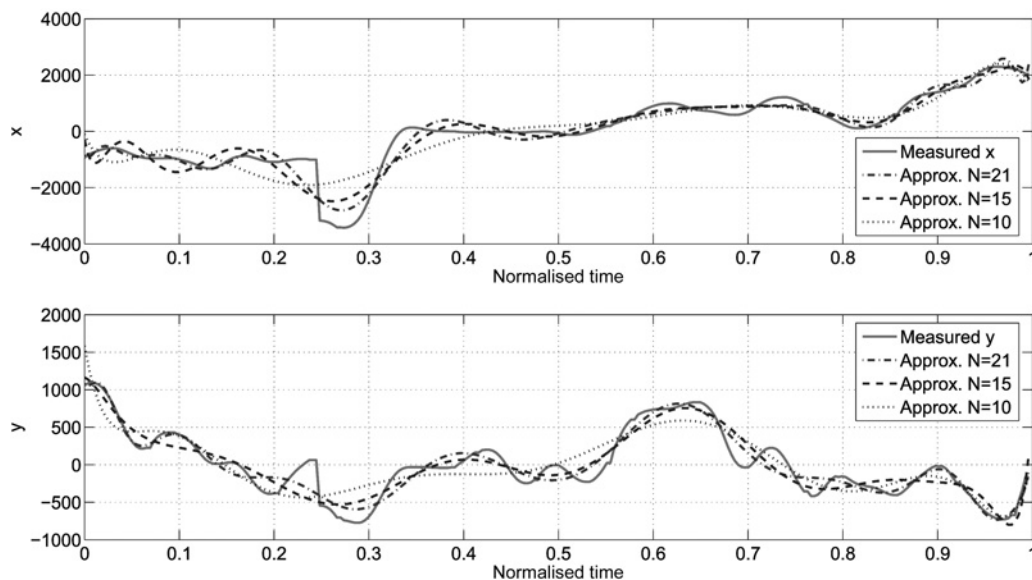


Fig. 1 Time functions: x and y pen coordinates (solid line) and their corresponding approximations by the Legendre polynomials with orders $N = 21$ (dashed line), $N = 15$ (dash-dotted line) and $N = 10$ (dotted line)

an expected result, taking into account the bias-variance tradeoff inherent to least squares estimation from noisy data.

3.2 Time function approximation via wavelets

An orthonormal wavelet bases in an inner product space can be generated by dilation and translation of a mother wavelet [41]. An approximation of a function f similar to the one in (4) can also be performed using these wavelet bases. This linear approximation can be improved if one chooses 'a posteriori' the N bases h_i , depending on the function f , in such a way to minimise the approximation error. This is done by choosing the set of N bases that have the largest inner product amplitudes $|\langle f, h_i \rangle|$. The approximation would then be as follows

$$f \simeq \sum_{i \in I_N} \alpha_i h_i \quad (8)$$

where I_N is an index set containing the indices corresponding to the largest inner products amplitudes.

This results in a non-linear approximation scheme, since the approximation vectors change with the function f . Since the amplitude of the inner products in a wavelet bases is related to the regularity of the signal, the approximation scheme is equivalent to constructing and adaptive approximation grid, whose resolution is locally adapted to the signal regularity. For signals with isolated singularities, the wavelet-based approximation is more precise than a linear scheme, which maintains the same resolution over the whole signal support.

Table 1 Best FIT between the measured and the approximated time functions using the Legendre polynomials

N	21	15	10
FIT _{x} , %	77.7955	68.9708	57.6664
FIT _{y} , %	70.7341	62.9579	53.3995

For the case of discrete-time functions, the inner products are computed resorting to the discrete wavelet transform (DWT) [41], which decomposes the signal at different levels of resolutions, splitting it in low-frequency ('approximation') and high-frequency ('details') components.

The idea is to perform a multilevel decomposition of the time functions using the DWT and to use the approximation coefficients to represent them. As mentioned in Section 2, resampling of the time functions, previous to the DWT decomposition, is needed in order to have a fixed-length feature vector. An approach where the detail coefficients, instead of the approximation ones, are used to represent the time functions within the framework of online signature verification was presented in [42].

Fig. 2 schematically depicts a filter bank representation of the multilevel (level of resolution $\ell = 2$) decomposition (left) and the approximate reconstruction (right) of the discrete-time function $f(n)$ using the DWT. Note that only the approximation coefficients are used for the reconstruction of the signal in the right side of Fig. 2. In the figure, LP represents a low-pass filter having the scaling function as its impulse response, whereas HP represents a high-pass filter having the mother wavelet as its impulse response. As mentioned above, the DWT ℓ -level approximation coefficients a_ℓ will be used to model the corresponding time function.

To illustrate the above approximation procedure, the x - and y pen coordinates associated with the same signature analysed in the case of the Legendre approximations in Fig. 1, and the corresponding approximations using the DWT (with the db4 wavelet [41]) with levels of resolutions $\ell = 1$, $\ell = 2$ and $\ell = 3$ are shown in Fig. 3. The time functions were resampled so that the resulting length is 256.

The Best FIT between the measured and the approximated time functions, for the above-mentioned levels of resolution, are given in Table 2. Also shown are the lengths of the resulting feature vectors for the different levels of resolution. It can be observed that reasonable FITs are obtained for the three levels of resolution considered in the table. The design parameter will then be the length of the resulting feature vector, which will determine the level of resolution to be used.

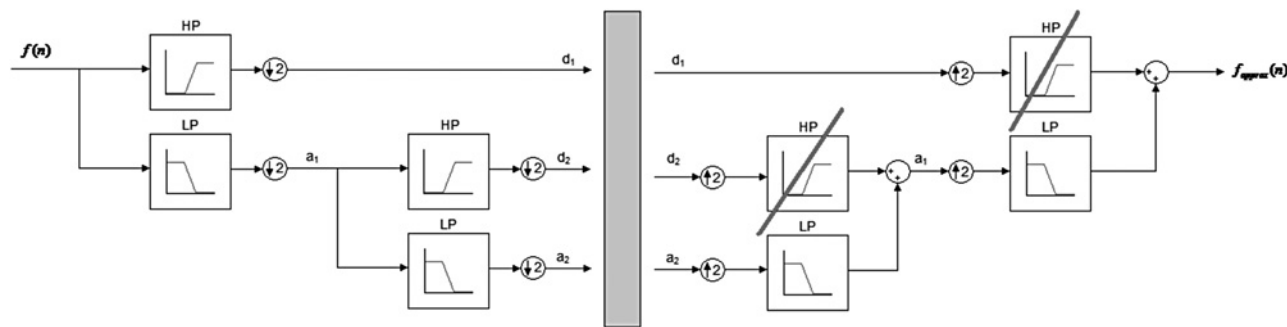


Fig. 2 Filter bank scheme of the multilevel ($\ell = 2$) decomposition (left) and the approximate reconstruction (right)

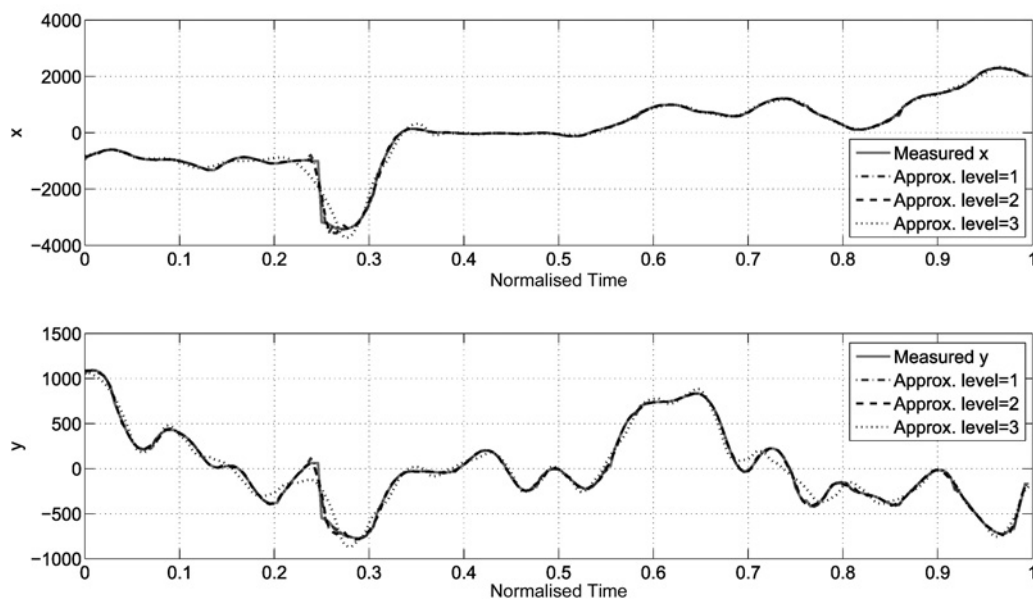


Fig. 3 Time functions: x and y pen coordinates (solid line) and their corresponding approximations by the DWT (db4) with levels of resolutions $\ell = 1$ (dashed line), $\ell = 2$ (dash-dotted line) and $\ell = 3$ (dotted line)

Table 2 Best FIT between the measured and the approximated time functions using the DWT (db4). The lengths of the resulting feature vectors are shown in the last row

ℓ	1	2	3
FIT _{x} %	94.1014	93.2690	86.9960
FIT _{y} %	94.9623	93.8578	82.9524
length a_ℓ	131	69	38

The computational load corresponding to the computation of the wavelet coefficients, is of the order of 0.1 s on a standard PC, which is a reasonable time in comparison with the remaining processing stages. A similar comment holds for the computation of the Legendre coefficients considered in Section 3.1.

4 Signature database

The publicly available SigComp2011 dataset [17] presented within ICDAR 2011 is used. It has two separate datasets, one containing genuine and forged Western signatures (Dutch ones) and the other containing genuine and forged Chinese signatures. The available forgeries are skilled forgeries, which are simulated

signatures, in which forgers (different writers than the reference one) are allowed to practice the reference signature for as long as they deem it necessary. The signatures were acquired using a ballpoint pen on paper (WACOM Intuos3 A3 Wide universal serial bus (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 dataset is divided into two sets, namely, the training and testing sets. The Dutch (left) and the Chinese (right) datasets are described in Table 3. Note that the amount of genuine and forged signature samples may differ from those in [17] since when making signatures available for the research community some of them were missing [43].

The measured data consist 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 Proposed feature combinations

The idea here is to address the problem of feature selection, based on their actual discriminative power, for online

Table 3 Online Dutch (left) and Chinese (right) datasets

Dutch dataset			Chinese dataset		
authors 10	Training set genuines 240	forgeries 119	authors 10	Training set genuines 230	forgeries 429
authors 54	testing set genuines 1296	forgeries 611	authors 10	testing set genuines 219	forgeries 461

signature verification. The main purpose of the proposed experiments in Section 6 is to compare different feature combinations chosen according to different criteria. The initial set of features, as mentioned in Section 2.3, consists of: x , y , p , v_T , θ , a_T and ρ , their first-order time derivatives, dx , dy , dp , dv_T , $d\theta$, da_T and $d\rho$, and their second-order time derivatives d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T and $d^2\rho$. In this paper, different feature combinations selected based on different criteria are analysed in an attempt to give some insight on the actual discriminative power of the features. The following feature combinations will be considered in this paper:

- ‘FHE-based features’: as mentioned in Section 1, FHEs work with the static image of the signature, so it is not possible for them to look at online features; however, they can infer some dynamic properties from the signature image, to some extent. The velocity is a distinctive feature for FHEs. So is the curvature, since it is hard to copy. On the other hand, the acceleration, the pen position (it can be established by striae and inkless starts) and the pressure are less useful for them. The local pressure by itself is not valuable for them, because external factors such as surface and writing material can easily influence it. However, pressure fluctuations could be interesting for them.

In this paper, the online features that will be considered as the ones an FHE would look at are: the velocity [v_T (magnitude) and θ (direction)], the curvature (ρ) and the first-order time derivative of the pressure (dp). These features were selected based on the above-described criteria corresponding to FHEs who examine Latin scripts. It is likely that for FHEs who examine Chinese scripts, the criteria would be different, but this will not be analysed in the present paper.

- ‘Automatically selected features’: it would be interesting to compare the performance obtained when the selection of the features is based on the FHEs criterion and when it is based on an automatic feature selection technique. In order to make this comparison, an automatic feature selection based on the variable importance provided by the random forest (RF) algorithm is also used in this paper. As already mentioned in Section 4, the datasets in the SigComp2011 database are divided into the training and testing sets [see Table 3 for the Dutch (left) and Chinese (right) data]. The feature selection is performed over the training sets for both datasets.

- ‘Features in [33]’: the initial set of features proposed in this paper is based on a set of widely used online features in the literature. It is important to analyse the performance of the proposed verification system on the same features that have already been used by other authors. For this reason, the feature combination used in [33] is also considered here.

- ‘Original set of features’: the measured (x , y and p) and the extended functions (v_T , θ , a_T and ρ).
- ‘All features’: the whole set of initial features, that is, x , y , p , v_T , θ , a_T and ρ , their first-order time derivatives, dx , dy , dp , dv_T , $d\theta$, da_T and $d\rho$ and their second-order time derivatives d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T and $d^2\rho$.

6 Evaluation protocol

To assess the verification performance of the different time function combinations proposed in this paper, a well-known state-of-the-art classifier, namely, RF [44], is used. In a previous work by the present authors [12], this classifier and one based on support vector machines (SVMs) were analysed on the same datasets considered here. The preliminary results presented in [12] show that the RF classifier outperforms the one based on SVMs regarding the verification error rates, and it is for this reason that only RF classifiers are considered in the present paper. For each of the datasets, namely, Dutch and Chinese, the optimisation 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.

For the case of the Legendre polynomials representations, the tuning parameter to adjust is the order of the Legendre polynomials. To select the optimal order, this parameter was varied from 1 to 25. For the representation based on DWT approximations, the user has to choose the mother wavelet, the length of the resampled functions and the level of resolution for the approximation. The length of the resulting feature vector is determined by the length of the resampled functions and the level of resolution. Regarding the RF classifier, the parameters to adjust are the number of trees to grow and the number of randomly selected splitting variables to be considered at each node. In general, the default values are a good choice for these parameters.

To obtain statistically significant results, a 5-fold cross-validation (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 writer, using only genuine signatures. To train the signature model for a particular writer, the genuine class consists in the genuine signatures of the writer available in the corresponding training set of the 5-fold CV, whereas the forged class consists in the genuine signatures of all the remaining writers in the dataset available in the same training set. The genuine and forged signatures of the writer 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 writer, 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 [45], from the detection error tradeoff curve as the point in the curve where the false rejection rate equals the false acceptance rate. The cost of the log-likelihood ratios \hat{C}_{llr} and its minimal possible value \hat{C}_{llr}^{\min} [45] 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 [17], where the importance of computing the likelihood ratios was highlighted since they make FHEs give an opinion on the strength of the evidence [37], although they are not in the position to make a leap of faith and judge about guilt or no guilt.

7 Results and discussion

The verification performance for each combination described in Section 5 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 well known state-of-the-art classification technique mentioned in Section 6: RF. For this 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. The order of the Legendre polynomials was set to $N=21$ since, as already mentioned, further increase of the order does not improve the approximation accuracy. For the case of the approximation based on wavelets, the time functions were resampled resulting in a normalised length of 256. The resolution level was set to 3, in order to

obtain a feature vector of a reasonable length. Two different wavelets, namely, db4 and bio6.8, were employed for the approximations.

The verification error rates for the different feature combinations listed in Section 5, are shown in Tables 4 and 5, for the Dutch and Chinese data, respectively. For the purposes of comparison, the verification results for the best commercial and non-commercial systems in the SigComp2011 competition (see [17] for a description of these systems) are also included in the last two rows of Tables 4 and 5.

From Tables 4 and 5, it can be seen that the combination using all the features and the combination using the automatically selected features obtain the best results, for both datasets. This shows that the feature selection done by the RF algorithm is a meaningful one. In addition, the length of the resulting feature vector is smaller than the corresponding to the set including all the features. Table 6 summarises the sets of automatically selected features for the three different approximation techniques and for the Dutch (top section of the table) and Chinese (bottom section) datasets.

Regarding the verification performance of the features selected based on the FHEs criterion, it is important to observe that despite the fact that the combination is not among the ones that obtain the best performances, it achieves an acceptable performance for both datasets. This result is promising since these features have a meaningful interpretation by the FHEs (see Section 5). This would suggest that in case the verification system has to be limited to take into account only FHE-based features, its performance would not be substantially deteriorated. In fact,

Table 4 Verification results for the dutch dataset

Features	Dutch dataset								
	Legendre polynomials			db4 wavelets			Bio6.8 wavelets		
	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}
FHEs' feat.	7.23	0.2624	0.2187	9.59	0.3408	0.2966	7.43	0.3221	0.2681
selec. feat.	5.01	0.1903	0.1594	6.58	0.2426	0.2049	5.68	0.2295	0.1971
selec. feat. (4)	5.53	0.2277	0.1908	9.67	0.3365	0.2948	8.86	0.3433	0.2894
Feat. in [33]	5.18	0.2023	0.1746	6.87	0.2466	0.2144	6.06	0.2553	0.2067
$x y p v_7 \theta a_{TP}$	6.4	0.2402	0.2077	8.04	0.2984	0.2485	7.24	0.2714	0.2309
all feat.	5	0.1879	0.1537	6.78	0.2491	0.2055	6.16	0.2320	0.1947
system		Acc.			\hat{C}_{llr}			\hat{C}_{llr}^{\min}	
commercial		96.27			0.2589			0.1226	
first non-commercial		93.49			0.4928			0.2375	

Table 5 Verification results for the Chinese dataset

Features	Chinese dataset								
	Legendre polynomials			db4 wavelets			Bio6.8 wavelets		
	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}	EER	\hat{C}_{llr}	\hat{C}_{llr}^{\min}
FHEs' feat.	11.61	0.4158	0.3171	10.27	0.3454	0.2760	9.4	0.3772	0.2742
selec. feat.	7.95	0.3313	0.2651	7.455	0.2962	0.2483	6.9	0.2874	0.2443
selec. feat. (4)	9.28	0.3784	0.2914	9.91	0.3948	0.3265	12.33	0.4655	0.3710
feat. in [33]	9.38	0.3825	0.2764	8.41	0.3338	0.2621	8.12	0.3271	0.2619
$x, y, p, v_T, \theta, a_T, p$	9.23	0.3539	0.3030	9.08	0.3548	0.2945	9.91	0.3727	0.3086
all feat.	8.66	0.3175	0.2647	7.9	0.3126	0.2476	7.26	0.2974	0.2421
system		Acc.			\hat{C}_{llr}			\hat{C}_{llr}^{\min}	
commercial		93.17			0.4134			0.2179	
first non-commercial		84.81			0.5651			0.3511	

Table 6 Optimal feature set selected by the automatic feature selection for each of the feature extraction approaches for the Dutch (top) and Chinese (bottom) datasets

Dutch	
feature extraction approach	optimal feature set
Legendre polynomials	$p, dx, dp, x, d^2p, a_T, y, dp, d^2y, d^2x, v_T, d^2\theta, dy, p$
wavelets db4	$x, a_T, y, v_T, p, dp, p, dx, \theta, dy, d^2x, d^2y, dv_T$
wavelets Bio6.8	$a_T, x, v_T, y, p, dy, dx, dp, \theta, d^2x, p, d^2y$
Chinese	
feature extraction approach	optimal feature set
Legendre polynomials	$dx, d^2x, dp, p, y, d^2p, dy, dv_T, d\theta$
wavelets db4	$yxpv_Ta_Tdydx^2y\theta dpd^2xd\theta^2pdv_Tdpd^2\theta$
wavelets Bio6.8	$x, y, v_T, p, a_T, dy, p, dx, dp, \theta, d^2x, d^2y, d\theta, d^2p, dv_T$

if all the features that FHEs look at could be implemented (which is hard to do since some features used by FHEs are not appropriately defined to be computed automatically), the performance might even be better. Moreover, taking into account the other results in the state-of-the-art reported over the same datasets presented in the last two rows of Tables 4 and 5, it can be concluded that the performance of the proposed systems using only the FHE-based features will still be comparable with the ones reported in the state-of-the-art.

The results obtained when using the features automatically selected outperform the ones obtained when using the FHE-based features, for both datasets. This is probably due to the fact that the automatic feature selection is not limited with respect to the size of the selected set. Nevertheless, the automatic feature selection is always (with an exception in the case of the Legendre approximation for the Chinese data) keeping the FHE-based features among their selected ones. This means that the FHE-based features are considered important by the automatic feature selection algorithm. In addition, it is shown that the FHE-based features are not good enough to achieve the best results, but need to be combined with other important features to achieve that. The fact that the automatic feature selection technique can keep the FHEs features can be useful since the selected feature set will contain features that are fully understandable and have been thoroughly investigated by FHEs. If a system is thought to interact with the FHEs, it is important to use a set of features that are generally accepted by them.

The fact that the results obtained when using the automatically selected feature sets outperform the ones obtained when using the FHE-based features is probably because of, as mentioned above, the larger size of the selected feature sets. In order to compare feature sets of the same size, the results using only the four most important features in the automatically selected feature sets were also computed. These results were included in Tables 4 and 5 [Feat. select. (4)]. It can be observed that whenever using the Legendre approach to extract the features, the results obtained by the first four features in the automatic selection are better than the ones obtained by the FHE-based features, while whenever using the DWT approach, the situation is the opposite. The first four automatically

selected features do not coincide with the FHE-based features. Only dp is kept among the first four features for the Legendre approach and v_T is kept for the DWT approach. Owing to the feature selection technique, it is likely that the first features in the ranking are not the best features by themselves, but they are good features when combined with other features.

The discriminative power of the pen pressure has long been questioned in the literature of online signature verification. This discussion, as already mentioned in Section 1, has particularly been intensified since SVC2004 was held [6]. The results obtained in this paper show that the pen pressure and its first-order time derivative are always among the best features since the automatic feature selection always include them into the optimal sets of features. This holds for the Dutch and Chinese datasets. Moreover, for the latter, the second-order time derivative of the pen pressure is also included in the selected feature sets. These results would suggest that the pen pressure is a useful feature when combined with other features. This observation agrees with many other reported results in the literature, such as [9–12], where it is stated that the pen pressure is a useful feature to distinguish between writers when used in combination with other time functions. The best set of features for any of the feature extraction approaches being used and any of the datasets being considered contains the pen pressure and its first-order time derivative. These results show that the pen pressure is useful regardless of the feature extraction being used here and, maybe more interesting, regardless of the dataset being considered in this paper. This could be suggesting that the reliability of the pen pressure is not highly influenced by the considered cultural origin of the signatures, depending mainly on the writer. Since in the present paper only Dutch (as an example of Western signatures) and Chinese signatures are considered, it is mandatory to analyse more data from different cultures in order to make further conclusions, but this observation could be a promising starting point.

In [33], the first-order time derivatives are considered into the initial feature set since it has been shown that they are highly effective as discriminative parameters regarding verification with other behavioural traits, such as speech. In [11], not only the first- but the second-order time derivatives of an original set of features are analysed. The results obtained when only the original set of features is used are clearly outperformed by the ones obtained when the features proposed in [33] are used, that is, incorporating the first-order time derivatives to the original set. This observation, agrees with the ones done in [33], and shows that the set of first-order time derivatives is useful to reach better verification results. Nevertheless, to incorporate the second-order time derivative seems not to be that significantly helpful, since incorporating them to the features used in [33] does not improve the verification results as much as incorporating the first-order time derivatives to the original feature set. Then, to use the first-order time derivatives contributes to reach a better model of the signatures, but to use the second-order time derivatives does not contribute so much. This contribution of the first-order time derivatives could be related to the power of the time derivatives to model the details. Nevertheless, it is important to note that, in addition to highlighting the details, the derivatives can also increase spurious information such as noise and other signal artefacts. Then, taking into account the first-order time derivative of a signal may contribute, but the second-order

time derivative may introduce non-useful information. This is particularly notable in the cases of the time derivatives of the extended time functions, namely, v_t , θ , a_t and p that are more complex signals than the measured ones (x , y and p).

From Tables 4 and 5, it can be seen that the results obtained when using the automatically selected feature sets outperform the ones obtained when using the set of features used in [33]. In line with the observations done in the previous paragraph, it can be observed that the main difference between these feature combinations is the way in which they incorporate the time derivatives to the original set of features. The first-order time derivatives of the measured functions, that is, dx , dy and dp are included in the optimal feature sets. On the other hand, dv_t and $d\theta$ are not always included, dp is only included once and da_t is not included in any optimal feature set. Instead of including these first-order time derivative features, the second-order time derivatives of the most simple time functions, that is, d^2x , d^2y and d^2p are included. This shows that it is likely more useful to use the second-order time derivative of simple features than the first-order time derivatives of more complex ones. As mentioned above, the time derivatives are capable to highlight details, but also useless information. If the features are not so representative or they are unstable, the first-order time derivatives of these features will be nothing but noisy data. On the other hand, the second-order time derivatives of the simple features will still keep useful information and will be capable to highlight the details without increasing drastically the artefacts of the original signal.

The best results for the Dutch data (Table 4) are reached when using the Legendre polynomials approximation of the time functions. Dutch signatures are likely to be written in a

continuous way, so that the time functions associated with these signatures should be smooth and so polynomials, the Legendre ones in this paper, can accurately approximate them. In the case of using the approximation based on DWT, the results are not as good as in the case of using the Legendre approximation. The best results for the Chinese data (Table 5) are obtained when using the DWT approximations of the time functions. Chinese signature style is, in most of the cases, close to the Chinese handwriting style, consisting of one or more separated multi-trace characters. This causes the time functions associated with the Chinese signatures to have several discontinuities. Then, a polynomial approximation, such as the one based on the Legendre polynomials used here, is not good enough to model this type of signals.

The best results for the Dutch and Chinese data are comparable with the recent results reported in the literature over the same datasets. Taking into account the results for the best commercial and non-commercial systems in SigComp2011 (last two rows of Tables 4 and 5), it is worth to note that even though the best results for the Dutch and Chinese data are not as good as the corresponding to the best commercial system [xyzmo (<http://www.xyzmo.com>)], they would have ranked first among the non-commercial systems and second among all the participants. Finally, the verification results obtained for the Dutch signatures are better than those for the Chinese ones, confirming the observations in [17, 12], and indicating that Chinese signatures are more challenging and that a lot of research has to be done on this type of data.

The potential, regarding the discriminative capability, of the FHE-based features has been discussed above.

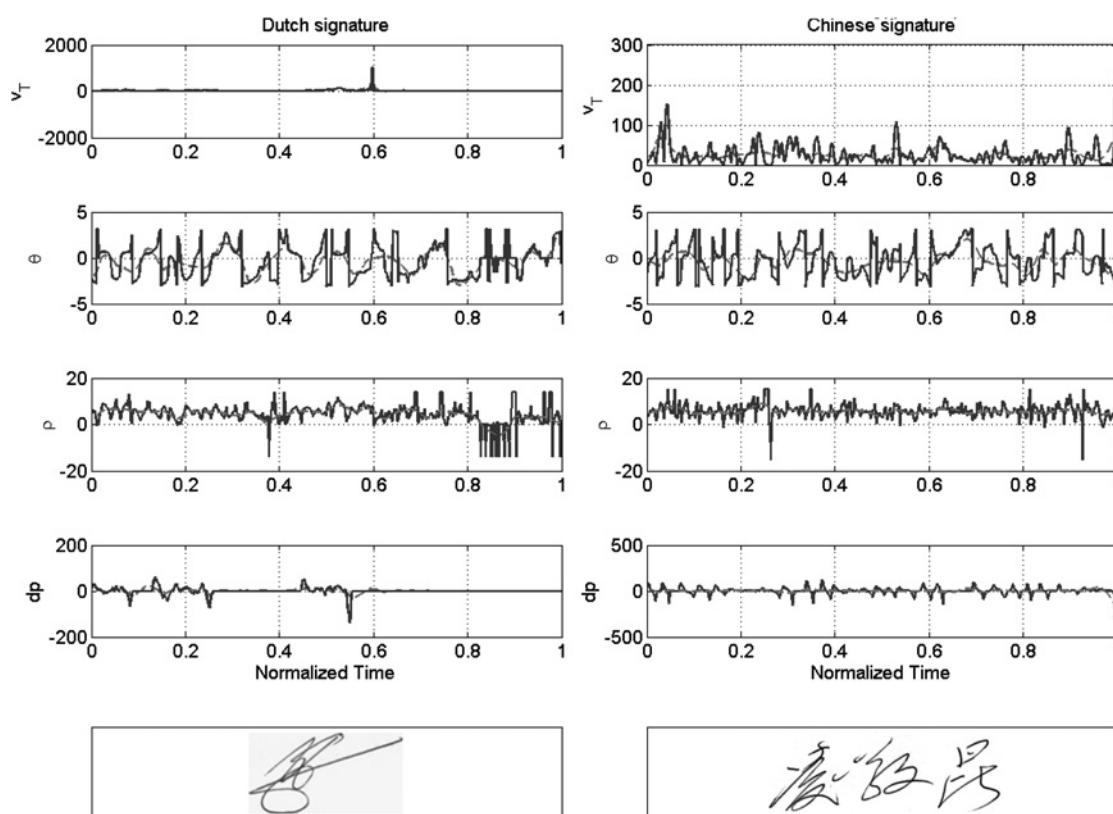


Fig. 4 Wrongly classified signatures with FHEs features: original v_t (first row), θ (second row), ρ (third row) and dp (fourth row) (solid line) and their corresponding approximations by DWT (db4) (dashed line), for the Dutch (left) and Chinese (right) datasets. Bottom row: associated signature images

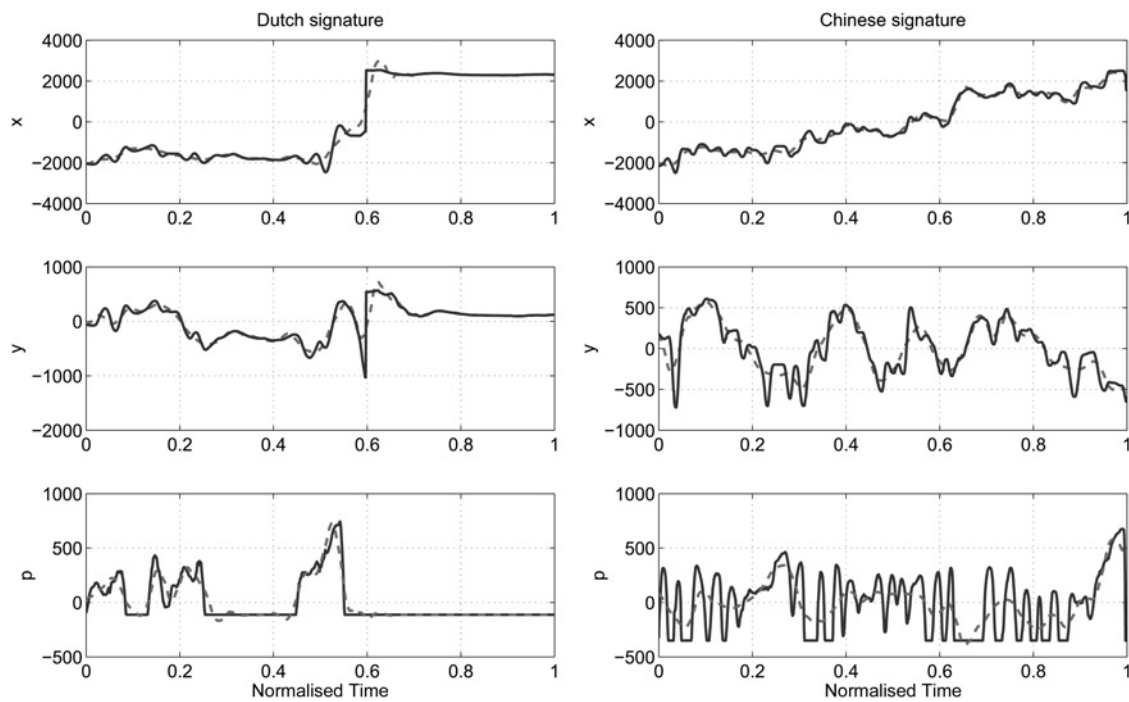


Fig. 5 Wrongly classified signatures with FHEs features: original x (top), y (middle) and p (bottom) (solid line) and their corresponding approximations by DWT (db4) (dashed line), for the Dutch (left) and Chinese (right) datasets

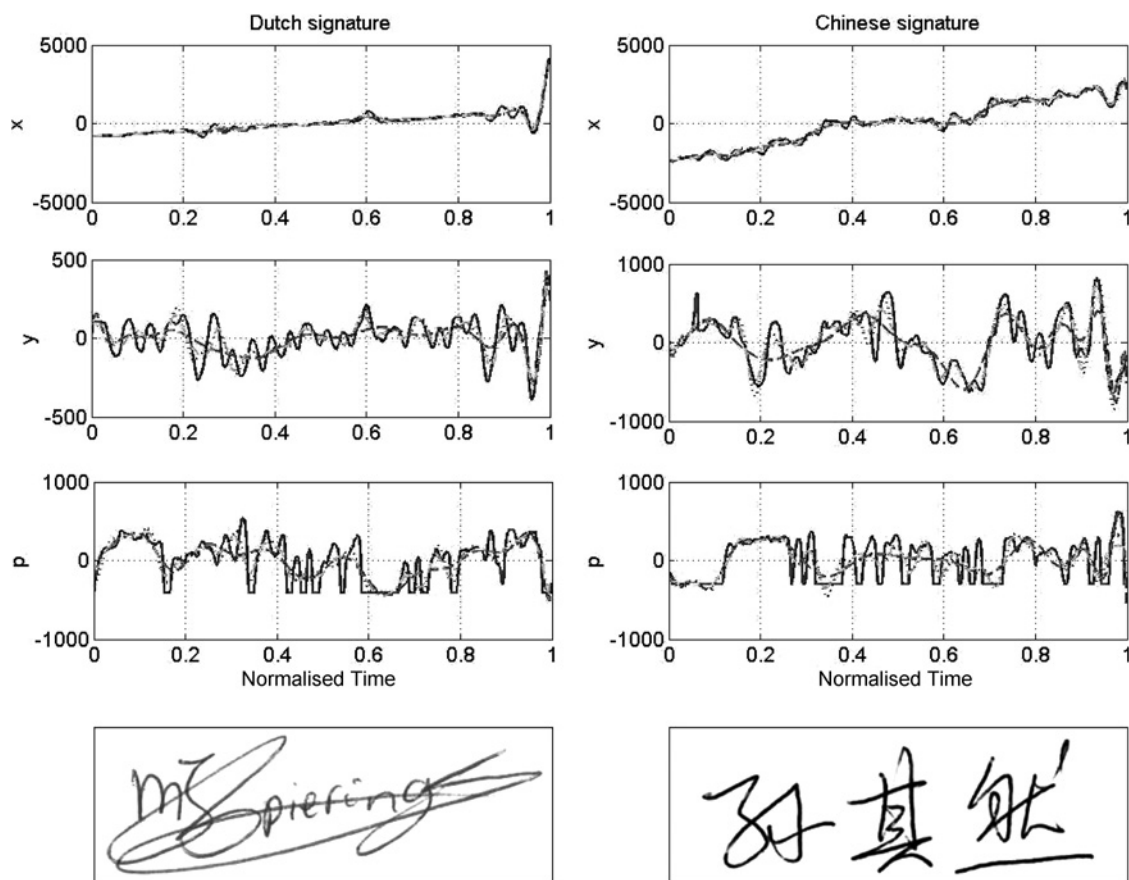


Fig. 6 Wrongly classified signatures: original x (top), y (middle) and p (bottom) (solid line), their corresponding approximations by the Legendre polynomials (dashed line), DWT (db4) (dash-dotted line) and DWT (bio6.8) (dotted line), for the Dutch (left) and Chinese (right) datasets. Bottom row: associated signature images

Nevertheless, there are signatures corresponding to certain authors in the Dutch and Chinese datasets that are wrongly classified using these features, whereas they are correctly classified with other feature combinations. In order to analyse these particular cases, the FHE-based features, that is, v_T , θ , ρ and dp for a sample of the signatures that are wrongly classified, together with their corresponding approximations computed with DTW ('db4') are shown in Fig. 4, for a Dutch (left) and a Chinese (right) genuine signature, respectively. Also shown in the figure are the associated signature images (bottom row). It can be seen that these time functions are not smooth, causing that the approximations proposed in this paper cannot model them accurately. In the case of using the Legendre approximations, a possible solution would be to increase the order of the polynomials, but it has already been mentioned that further increasing it does not improve the approximation accuracy. In the case of using the DWT approximation, an alternative would be to incorporate the detailed coefficients to the model. Experiments showed that this actually improves the approximation accuracy. Despite the fact that this would increase the length of the feature vector, this would not be a limitation here since the FHE-based feature sets include only four features.

As mentioned above, these particular signatures are correctly classified when using another feature combination. Features such as x , y and p , that are simpler than the FHE-based ones, are shown for these signatures in Fig. 5, for the Dutch (left) and Chinese (right) cases, respectively.

These signals are smooth and they can be more accurately approximated by the techniques proposed here. Then, incorporating these type of features (those associated with smooth time functions), would result in better verification results. This is the case for the signatures analysed here, since they were only wrongly classified by the FHE-based features, being correctly classified when using the other feature combinations proposed in this paper.

Finally, there are signatures corresponding to certain authors in the Dutch and Chinese datasets that are wrongly classified by any of the feature combinations proposed here. In Fig. 6, the x , y and p time functions corresponding to a sample of the wrongly classified cases are shown together with their corresponding approximations based on the Legendre polynomials and DWT, for a Dutch (left) and a Chinese (right) genuine signature, respectively. In addition, in Fig. 7, the same time functions corresponding to a sample of the correctly classified cases are shown together with their corresponding approximations, for a Dutch (left) and a Chinese (right) genuine signature, respectively. Also shown in Figs. 6 and 7 are the associated signature images (bottom rows).

From Figs. 6 and 7, it can be observed that the approximations for the time functions in Fig. 6 are significantly worse than the ones in Fig. 7. This is suggesting that the bad performance in these cases is because of the poor quality approximation obtained with the proposed approximation techniques. If the analysis is focused on the pressure signal in Figs. 6 and 7, it can be noted that the

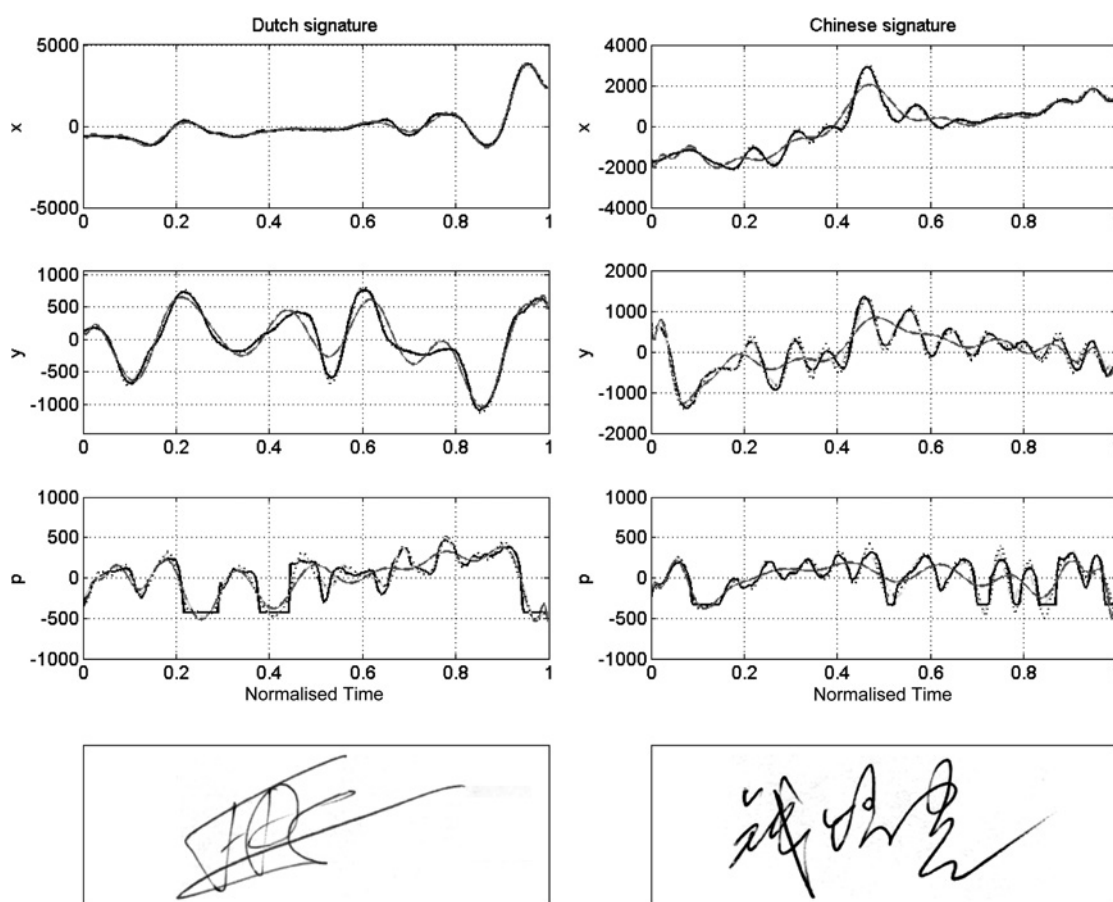


Fig. 7 Correctly classified signatures: original x (top), y (middle) and p (bottom) (solid line), their corresponding approximations by the Legendre polynomials (dashed line), DWT (db4) (dash-dotted line) and DWT (bio6.8) (dotted line), for the Dutch (left) and Chinese (right) datasets. Bottom row: associated signature images

approximation is not good in the case of a signal having pulses with a large amplitude and a short duration. Regarding the Legendre polynomials approximations, it is well known that polynomial approximations are not capable to model this type of signals, while regarding the wavelet approximations, a good approximation is not achieved since the detail coefficients (associated with the high-frequency components of the signal) are being neglected. It is important to note that such type of signals usually are associated with the p function and, eventually, to the y function, and the functions that are derived from them. Note the reader that, despite the fact that the proposed approximations cannot handle pulses that have large amplitude and a short duration, they can handle wider pulses. Fortunately, this is, most of the times, the case for the pressure signals.

The analysed time functions in Fig. 6 are associated with the cases, in which the proposed verification system has not a good performance. It has been observed that the shortcomings are because of the poor quality approximation obtained in these cases with the approximation techniques proposed in this paper. The problem could be solved by increasing the order of the Legendre polynomials, or incorporating the detail coefficients in the wavelet approach. Nevertheless, as already mentioned, further increasing the order of the polynomials does not improve the results and incorporating the detail coefficients would increase the length of the time function model and then would limit the number of features to be selected to model the signatures.

8 Conclusions

Different feature combinations were analysed, for two different signature styles, namely, Western and Chinese, in order to provide some insight on the actual discriminative power of the features for online signature verification. This could also help FHEs for further understanding the signatures and the writer's behaviour.

The experimental results showed that the proposed signature verification system could be integrated into toolkits that could be used by FHEs for helping them to analyse and understand the signatures. In particular, the following observations are worth to be done:

- The features that an FHE would look at when doing his daily work proved to reach good results, comparable with the ones in the state-of-the-art reported on the same datasets. This result is promising since these features have a meaningful interpretation and are generally accepted by the FHEs.
- The comparison between the verification performance obtained when using the FHE-based features and the automatically selected features showed that a better verification performance is obtained in the latter case. Despite this fact, the set of automatically selected features does contain the set of FHE-based features; this is important since it makes the system capable to interact with FHEs.

It was shown that the automatic feature selection proposed in this paper was a meaningful one, since the results obtained when using the automatically selected features were the best ones (together with the ones obtained when using all the features), for both datasets. That is, a valuable set of features could be selected, from the proposed original feature set, for the two different signature styles.

The discriminative power of the pen pressure has particularly been analysed, and it was shown to be a useful feature for both datasets. Although in this paper only Dutch and Chinese signatures are considered, the independence of the data found here is very important since this could be suggesting that the reliability of the pen pressure is not highly influenced by the considered cultural origin of the signatures, depending mainly on the writer. Of course, it is mandatory to analyse more data from different cultures in order to make further conclusions, but this observation could be a promising starting point. Finally, it is important to note that the discriminative power of the pen pressure can easily be influenced by the signature acquisition process. If this process is done in carefully controlled conditions, the pressure is likely to be a very distinctive feature. As a future work, it would be interesting to study its discriminative power in an FHEs casework context.

The use of the proposed time function approximations, namely, the one based on the Legendre polynomials and the one based on wavelet decomposition, proved to be a good choice since the obtained error rates using both of them are comparable with those of other state-of-the-art verification systems, tested on the same datasets.

The verification results obtained for the Dutch signatures are better than the ones obtained for the Chinese ones, indicating that Chinese signatures are more challenging and that a lot of research has to be done on this type of data.

9 References

- 1 Plamondon, R., Lorette, G.: 'Automatic signature verification and writer identification – the state of the art', *Pattern Recognit.*, 1989, **22**, (2), pp. 107–131
- 2 Leclerc, F., Plamondon, R.: 'Automatic signature verification: the state of the art–1989–1993', *Int. J. Pattern Recognit. Artif. Intell.*, 1994, **8**, (3), pp. 643–660
- 3 Plamondon, R., Srihari, S.N.: 'On-line and off-line handwriting recognition: a comprehensive survey', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (1), pp. 63–84
- 4 Impedovo, D., Pirlo, G.: 'Automatic signature verification: the state of the art', *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, 2008, **38**, (5), pp. 609–635
- 5 Impedovo, D., Pirlo, G., Plamondon, R.: 'Handwritten signature verification: new advancements and open issues'. Proc. 13th Int. Conf. Frontiers in Handwriting Recognition, Bari, Italy, September 2012, pp. 367–372
- 6 Yeung, D.-Y., Chang, H., Xiong, Y., *et al.*: 'SVC2004: first international signature verification competition'. Proc. Int. Conf. Biometric Authentication, Hong Kong, 2004, pp. 16–22
- 7 Kholmatov, A., Yanikoglu, B.: 'Identity authentication using improved online signature verification method', *Pattern Recognit. Lett.*, 2005, **26**, pp. 2400–2408
- 8 Lei, H., Govindaraju, V.: 'A comparative study on the consistency of features in on-line signature verification', *Pattern Recognit. Lett.*, 2005, **26**, pp. 2483–2489
- 9 Maramatsu, D., Matsumoto, T.: 'Effectiveness of pen pressure, azimuth, and altitude features for online signature verification'. Proc. Int. Conf. Biometrics, 2007, pp. 503–512
- 10 Pascual-Gaspar, J.M., Faundez-Zanuy, M., Viveracho, C.: 'Fast on-line signature recognition based on VQ with time modeling', *Eng. Appl. Artif. Intell.*, 2011, **24**, (2), pp. 368–377
- 11 Richiardi, J., Ketabdar, H., Drygajlo, A.: 'Local and global feature selection for on-line signature verification'. Proc. Eighth Int. Conf. Document Analysis and Recognition, Seoul, Korea, 2005
- 12 Parodi, M., Gómez, J.C., Liwicki, M.: 'Online signature verification based on legendre series representation. Robustness assessment of different feature combinations'. Proc. 13th Int. Conf. Frontiers in Handwriting Recognition, Bari, Italy, September 2012, pp. 377–382
- 13 Houmani, N., Mayoue, A., Garcia-Salicetti, S., *et al.*: 'Biosecure signature evaluation campaign (BSEC'2009): evaluating online signature algorithms depending on the quality of signatures', *Pattern Recognit.*, 2012, **45**, (3), pp. 993–1003

- 14 Houmani, N., Garcia-Salicetti, S., Dorizzi, B.: 'On assessing the robustness of pen coordinates, pen pressure and pen inclination to time variability with personal entropy'. Proc. IEEE Third Int. Conf. Biometrics: Theory, Applications, and Systems, Washington DC, USA, 2009
- 15 Galbally, J., Fierrez, J., Martinez-Diaz, M., Plamondon, R.: 'Quality analysis of dynamic signature based on the sigma-lognormal model'. Proc. 11th Int. Conf. Document Analysis and Recognition, Beijing, China, September 2011, pp. 633–637
- 16 Blankers, V.L., van den Heuvel, C.E., Franke, K.Y., Vuurpijl, L.G.: 'The ICDAR 2009 signature verification competition'. Proc. Tenth Int. Conf. Document Analysis and Recognition, Barcelona, Spain, July 2009, pp. 1403–1407
- 17 Liwicki, M., Malik, M.I., den Heuvel, C.E., *et al.*: 'Signature verification competition for online and offline skilled forgeries (SigComp2011)'. Proc. 11th Int. Conf. Document Analysis and Recognition, Beijing, China, September 2011
- 18 Pal, S., Blumenstein, M., Pal, U.: 'Non-English and Non-Latin signature verification systems: A survey'. Proc. First Int. Workshop on Automated Forensic Handwriting Analysis, Beijing, China, 2011
- 19 Lv, H., Wang, W., Wang, C., Zhuo, Q.: 'Off-line chinese signature verification based on support vector machines', *Pattern Recognit. Lett.*, 2005, **26**, pp. 2390–2399
- 20 Ji, J.-W., Chen, X.-S.: 'Off-line Chinese signature verification segmentation and feature extraction'. Proc. Int. Conf. Computational Intelligence and Software Engineering, 2009, pp. 1–4
- 21 Zou, M., Tong, J., Liu, C., Lou, Z.: 'On-line signature verification using local shape analysis'. Proc. 7th Int. Conf. Document Analysis and Recognition, Edinburgh, Scotland, UK, 2003
- 22 ICDAR 2011: Int. Conf. Document Analysis and Recognition, Beijing, China, 19–21 September, 2011
- 23 Ueda, K.: 'Investigation of off-line Japanese signature verification using a pattern matching'. Proc. Seventh Int. Conf. Document Analysis and Recognition, Edinburgh, Scotland, 2003
- 24 Yoshimura, M., Yoshimura, I.: 'Investigation of a verification system for Japanese countersignatures on traveler's cheques'. *Trans. IEICE*, 1998, **J80-D-II**, (7), pp. 1764–1773
- 25 Ismail, M.A., Gad, S.: 'Off-line Arabic signature recognition and verification', *Pattern Recognit.*, 2000, **33**, (10), pp. 1727–1740
- 26 Pal, S., Blumenstein, M., Pal, U.: 'Hindi off-line signature verification'. Proc. 13th Int. Conf. Frontiers in Handwriting Recognit., Bari, Italy, September 2012, pp. 373–378
- 27 Ballantyne, K.N., Bird, C.L., Found, B., Rogers, D.: 'Dynamic features of naturally written, disguised and forged handwritten text'. Proc. 15th Conf. Int. Graphonomics Society., Cancun, Mexico, June 2011
- 28 Will, E.J.: 'Inferring relative speed of handwriting from the static trace', *J. Forensic Doc. Examination*, 2012, **22**, pp. 55–62
- 29 Caligiuri, M., Mohammed, L.: 'The neuroscience of handwriting' (CRC Press, Florida, 2012)
- 30 O'Reilly, C., Plamondon, R.: 'Agonistic and antagonistic interaction in speed/accuracy tradeoff: a delta-lognormal perspective', *Hum. Mov. Sci.*, 2013, doi:10.1016/j.humov.2012.07.005
- 31 Oliveira, L.S., Santos, C.R., Bortolozzi, F., Justino, E.: 'Off-line signature verification based on forensic questioned document examination approach'. Proc. Symp. Applied Computing, Seoul, Korea, March 2007
- 32 Pervouchine, V., Leedham, G.: 'Extraction and analysis of forensic document examiner features used for writer identification', *Pattern Recognit.*, 2007, **40**, (3), pp. 1004–1013
- 33 Fierrez-Aguilar, J., Ortega-Garcia, J., Ramos-Castro, D., Gonzalez-Rodriguez, J.: 'HMM-based on-line signature verification: Feature extraction and signature modelling', *Pattern Recognit. Lett.*, 2007, **28**, pp. 2325–2334
- 34 Jain, A.K., Griess, F.D., Connell, S.D.: 'On-line signature verification', *Pattern Recognit.*, 2002, **35**, pp. 2963–2972
- 35 Freire, M.R., Martinez-Diaz, M., Fierrez, J., Ortega-Garcia, J.: 'On the effects of sampling rate and interpolation in HMM-based dynamic signature verification'. Proc. Ninth Int. Conf. Document Analysis and Recognition, Curitiba, Brazil, 2007, pp. 1113–1117
- 36 Yanikoglu, B., Kholmatov, A.: 'Online signature verification using fourier descriptors', *EURASIP J. Adv. Signal Process.*, 2009, pp. 230–275
- 37 Ramos-Castro, D., Gonzalez-Rodriguez, J., Fierrez-Aguilar, J., Ortega-Garcia, J.: 'Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems', *Forensic Sci. Int.*, 2005, **155**, pp. 126–140
- 38 Tuyls, P., Akkermans, A.H.M., Kevenaar, T.A.M., Schrijen, G.-J., Bazen, A.M., Veldhuis, R.N.J.: 'Practical biometric authentication with template protection'. Fifth Int. Conf. Audio- and Video-Based Personal Authentication (AVBPA), volume 3546 of Lect. Notes in Comput. Sci., Springer-Verlag, Berlin, 2005, pp. 436–446
- 39 Xu, H., Veldhuis, R.N.J., Kevenaar, T.A.M., Akkermans, A.H.M., Bazen, A.M.: 'Spectral minutiae: a fixed-length representation of a minutiae set'. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops, 2008
- 40 Golubitsky, O., Watt, S.M.: 'Distance-based classification of handwritten symbols', *Int. J. Doc. Anal. Recognit.*, 2010, **13**, (2), pp. 133–146
- 41 Daubechies, I.: 'Ten lectures on wavelets' (SIAM: Society for Industrial and Applied Mathematics, Pennsylvania, 1992)
- 42 Chang, H., Dai, D., Wang, P., Xu, Y., Si, F., Huang, S.: 'Online signature verification using wavelet transform of feature function', *J. Inf. Comput. Sci.*, 2012, **9**, (11), pp. 3135–3142
- 43 SigComp2011. ICDAR 2011 Signature verification competition. Available at http://www.iapr-tc11.org/mediawiki/index.php/Datasets_List, 2011
- 44 Breiman, L.: 'Random forests'. Technical report, Department of Statistics, University of California, Berkeley, 2001
- 45 Brümmer, N., du Preez, J.: 'Application-independent evaluation of speaker detection', *Comput. Speech Lang.*, 2006, **20**, pp. 230–275